Banco de Portugal

EUROSISTEMA

Estudos e Documentos de Trabalho

Working Papers

7 | 2007

CREDIT RISK DRIVERS: EVALUATING THE CONTRIBUTION OF FIRM LEVEL INFORMATION AND OF MACROECONOMIC DYNAMICS

Diana Bonfim

March 2007

Please address correspondence to
Economic Research Department
Email: dbonfim@bportugal.pt

BANCO DE PORTUGAL

Economic Research Department
Av. Almirante Reis, 71-6th floor
1150-012 Lisboa

www.bportugal.pt

Printed and distributed by
Administrative Services Department
Av. Almirante Reis, 71-2nd floor
1150-012 Lisboa

Number of copies printed
300 issues

Legal Deposit no. 3664/83
ISSN 0870-0117
ISBN 972-9479-72-0

Credit risk drivers: evaluating the contribution of firm level information and of macroeconomic dynamics *

Diana Bonfim†
Banco de Portugal

Abstract

Understanding why some firms default, while others do not, is an important issue for the assessment of financial stability. In this domain, it may be interesting to understand if credit risk is driven mostly by idiosyncratic firm characteristics or by systematic factors, which simultaneously affect all firms. In order to empirically examine the determinants of loan default, we begin by exploring the links between credit risk and macroeconomic developments at an aggregate level. The results obtained seem to confirm the hypothesis that in periods of economic growth, which are sometimes accompanied by strong credit growth, there may be some tendency towards excessive risk-taking, even though the imbalances created in such periods only become apparent when economic growth slows down. After examining the determinants of credit risk at an aggregate level, we focus our attention on an extensive dataset with detailed financial information for more than 30.000 firms. The results obtained suggest that default probabilities are influenced by several firm-specific characteristics, such as their financial structure, profitability and liquidity, as well as by their recent sales performance or their investment policy. When time-effect controls or macroeconomic variables are taken into account together with the firms' characteristics, the results seem to improve substantially. Hence, though the firms' financial and operational situation has a central role in explaining default probabilities at the micro level, overall macroeconomic conditions are also very important when assessing default probabilities over time.

JEL Codes: G21, G33, E32, C25, C41.

Keywords: credit risk, default probability, corporate loans, probit models, duration analysis

*I am grateful to Paula Antão, Paulo Soares Pinho, António Antunes, Nuno Ribeiro, Pedro Portugal and Miguel Ferreira, for their helpful comments and suggestions. All remaining errors are my responsability.
†Email: dbonfim@bportugal.pt. Address: Av. Almirante Reis, 71, 6^o, 1150-012 Lisboa, Portugal.

1 Introduction

Understanding the determinants of credit risk is a major issue for financial stability. Banks and other financial intermediaries try to maximise their profits, which requires an accurate pricing of the risks contained in their assets portfolios. Thereby, given the weight loans to firms have on banks' assets, understanding why do some firms default, while others do not, may be a very important question to address. A clearer understanding of credit risk drivers may be an important contribution to help predict if and when will a firm default on its credit liabilities. Against this background, it is interesting to understand if credit default risk is mostly driven by idiosyncratic or by systematic factors (or both). On one hand, firm-specific characteristics should clearly be determinant on their decision to default on bank loans. On the other hand, it has become clearer that macroeconomic developments may also have an important role in the evolution of credit risk over time. Under this setup, the main purpose of this paper is to empirically examine the determinants of corporate credit default, taking simultaneously into account firm-specific data as well as macroeconomic information.

The results obtained suggest that there are some important links between credit risk and macroeconomic developments. In fact, periods of strong economic growth, which are sometimes accompanied by robust credit growth, are sometimes followed by an increase in default rates, possibly as a consequence of imbalances generated in those periods. Nevertheless, when micro information is used to assess the determinants of loan default, it becomes clearer that it is ultimately the firms' own financial situation that will determine whether they will default on their loan commitments. However, it is important to consider that firms' financial ratios also vary over the business cycle. Hence, some firms may be more vulnerable to systematic shocks, explaining the fluctuation of default rates over the business cycle.

The remainder of the paper proceeds as follows. We begin by reviewing some of the literature in credit risk modelling in Section 2. We also briefly present the modelling setup underlying the empirical work which will be developed further ahead. In Section 3 we try to understand some of the links between credit risk and macroeconomic developments at an aggregate level. For that purpose, we look at correlations between the cyclical components of credit overdue and of a large set of macroeconomic and financial variables. In Section 4 we finally look at firm-specific evidence. In this section, we begin by describing the panel dataset used in our work. In this extensive dataset we have information for more than 30.000 Portuguese firms for the period comprised between 1996 and 2002. This dataset contains information on firms' credit liabilities as well as detailed accounting information for each firm. In order to explore the determinants of loan default at a micro level, we use two different econometric techniques. We begin by using discrete choice models to understand *why* do some firms default, but later we complement our analysis using duration models. The introduction of the time dimension encompassed in duration models may provide some additional evidence on the timing of loan default, thereby addressing the question of *when* do firms default. Finally, Section 5 presents some concluding remarks.

2 A brief insight into credit risk modelling

2.1 Review of the literature

One of the first major contributions in credit risk modelling literature was provided by Altman (1968). In this paper, the author introduced a credit scoring model, usually known as Z-score. This score is a linear combination of explanatory variables and their respective coefficients, obtained through discriminant analysis. The set of independent variables comprises several financial ratios, namely 1) working capital to total assets, 2) retained earnings to total assets, 3) earnings before interest and taxes over total assets, 4) the ratio between the market value of equity and the book value of debt and, finally, 5) sales over total assets. This score may be computed at firm-level, being used to try to predict whether a firm will become bankrupt or not, within a certain period of time. There are, however, some caveats in this approach. For instance, the conclusions obtained with these models may be, to some extent, sample-specific. Hence, their application to different samples may be somewhat limited.

Another early and important contribution to credit risk modelling was given by Merton (1974). In this seminal paper, Merton introduced the idea of applying option pricing theory to the valuation of risky bonds and loans (by modelling loans as zero-coupon bonds with fixed maturities). In this model, a borrower will have an incentive to default whenever the market value of the firm becomes lower than the amount borrowed. The value of the default option (or the value of the risky loan) will depend on five variables: the market value of firm's assets, the amount borrowed, the short-term interest rate, the volatility of the firm's assets and the loan maturity. There are several more recent models that draw on this simple modelling framework introduced by Merton, such as Tudela and Young (2003), for instance.

However, it is possible to identify several drawbacks in Merton's structural model. First of all, it is usually pointed out that we cannot observe the firm's market value of assets or the volatility of the firm's equity value. These must be approximated using available information, usually from equity prices, implying that Merton-type models usually can only be applied to quoted firms. It can also be argued that in these models, most of the information is subsumed in share prices and, as a consequence, it is not possible to understand which factors may be more relevant in determining default probabilities. Furthermore, as mentioned by Bunn and Redwood (2003), Merton-type models usually provide good ordinal rankings of companies, but fail to provide accurate probabilities of default. Another drawback of the Merton model is that, in practice, firms do not always default when the market value of the firm becomes lower than the amount borrowed. As a consequence, it would be more accurate to determine default probabilities for each value of the ratio between the amount borrowed and the market value of assets. This is, in part, what underlies the concept of empirical expected default probabilities (EDFs) in Moody's KMV model (for further details on this model see Saunders and Allen (2002) and Moody's (2004)).

In the late 90's, discussions concerning the design of the new international bank capital accord, usually known as Basel II, generated a renewed interest in credit risk modelling. The new capital accord (Basel Committee on Banking Supervision (2004)) proposes the use of credit risk models to determine banks' capital requirements. Banks can use internal (or external) rating models to classify borrowers according to their risk. Capital requirements can then be determined based on such credit exposure, instead of being constant per credit type, as under the previous accord. Under this new regulatory setup, it becomes crucial to accurately measure credit risk. On the one hand, banks must hold enough capital to limit risks for depositors and to reduce insolvency risks. However, on the other hand, holding excessive capital is costly and limits efficiency. This recent surge in credit risk modelling, to some extent associated with Basel II, is leading to several new contributions. A brief overview of some of the most important contributions in this field may be found in Crouhy *et al* (2000), in Gordy (2000) or, more recently, in Saunders and Allen (2002) or Duffie and Singleton (2003). In order to simplify the description of these recent models, we can try to group them according with their required inputs. We can identify three different groups of models, using this criteria: i) models which rely mostly on accounting variables, ii) models which use mostly market information, and iii) models which use macroeconomic variables or which consider default correlation issues.

The first group of models borrows from Altman's (1968) work, even though such variables can be used under different modelling techniques. Some recent work in this domain includes Bernhardsen (2001), Eklund *et al* (2001), Bunn and Redwood (2003) or Benito *et al* (2004). It should be borne in mind, however, that most of the models here mentioned do not rely solely on accounting information.

In the second group of models (those which rely mostly on market information), we can include Merton-type approaches to credit risk modelling (see, for instance, Tudela and Young (2003), Gersbach and Lipponer (2003) or even Moody's KMV model (2004)), as well as other modelling setups, such as Jarrow and Turnbull (1995), Shumway (2001) or Couderc and Renault (2005). The major drawback of such models is that, as they rely on market information, usually they can only be applied to quoted companies. From a prudential regulation viewpoint, one might be more interested in assessing the risk in a complete loan portfolio, which may contain a substantial part of non-quoted firms (and which are sometimes the riskier companies). This is an important issue when modelling credit risk in Portugal, for instance, as most corporate borrowers are non-quoted SMEs, which rely mostly on banks as providers of external funds. Nevertheless, the use of aggregate market information (such as overall stock market performance or interest rates) may provide useful insights, even when evaluating credit risk at such firms.

Finally, we can identify a third set of credit risk models as those which use macroeconomic variables or consider default correlation issues. Discussions resulting from the implementation of Basel II made clear that credit risk varies over time and, most notably, it varies with overall macroeconomic conditions. The main idea is that most risk is built up during upturns, when banks apply more loose credit standards. However, most of the risk materialises only when the economy hits a downturn. Some authors, such as Pederzoli and Torricelli (2005), Jiménez and Saurina (2006), Kent and D'Arcy (2001) or Borio *et al* (2001) argue that high default rates during recessions are just a materialisation of the risk that is built up during expansions, most notably when strong economic growth is accompanied by the creation of unsustainable financial imbalances. Moreover, Korajczyk and Levy (2003) observe that leverage tends to increase during periods of favourable economic conditions only for financially constrained firms, whereas for firms without relevant financial constraints leverage tends to be counter-cyclical. Hence, banks should have an over-the-cycle perspective when approving (and pricing) loans in phases of economic growth (most notably when they are accompanied by strong credit growth), in order to minimise the impact of possible future losses. It should therefore be emphasised that there is a large difference between potential and observed risk. Wilson (1998), who developed CreditPortfolioView (McKinsey's credit risk model), was one of the first authors to emphasise the role macroeconomic variables could have in explaining credit defaults, using a multi-factor model of systematic default risk. Bangia *et al* (2002) also had a crucial role in demonstrating the importance of macroeconomic developments in credit risk. These authors were the first to build transitions matrices conditional on business cycle conditions. Allen and Saunders (2003) provide a survey of cyclical effects in existing credit risk models. Other authors who tried to consider business cycle conditions in credit risk models include Lis *et al* (2000), Nickell *et al* (2000), Kent and D'Arcy (2001), Lowe (2002), Berger and Udell (2003), Carling *et al* (2002, 2004), Jiménez and Saurina (2006) or Figlewski *et al* (2006) [1]. Furthermore, taking into account macroeconomic variables helps us to deal with default correlation issues. In fact, empirical evidence presented by Rosch (2003) seems to suggest that default correlations between borrowers increase in economic downturns, as several borrowers are simultaneously hit by negative systematic shocks.

Our objective is to evaluate simultaneously the effects of some of these dimensions of corporate credit risk. In order to achieve such objective, we will consider firm-specific accounting information, as well as macroeconomic and financial data, trying to understand the relative importance of idiosyncratic and systematic risk factors in the default process.

2.2 Modelling default probabilities

The theoretical modelling setup underlying the empirical analysis which will be developed in Section 4 draws to some extent on previous work done by Rosch (2003) and Hamerle *et al* (2004). Under this modelling framework, we model the default event of firm i in period t as a random variable Y_{it} such that:

$$Y_{it} = \begin{cases} 1 & \text{if firm } i \text{ defaults in } t \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

The time-discrete hazard rate can be defined as:

$$\lambda_{it} = \text{Prob } (Y_{it} = 1) \tag{2}$$

We will consider two vectors of explanatory variables. The first one is a set of firm-specific variables, which shall account for idiosyncratic risk (Z_{it}). This vector will include contemporaneous and lagged variables regarding several dimensions of the firm's financial situation, such as age, size, asset growth, profitability, leverage and liquidity. The second vector comprises a set of aggregate time-varying regressors, which intend to account for systematic risk (X_t).

[1] There is a recent contribution by Koopman *et al* (2005) which also considers systematic and firm-specific risk components. However, in this paper the systematic component is not captured by observable macroeconomic variables, but rather by the estimation of an unobserved dynamic component.

These may include variables such as GDP growth, industrial production, confidence levels, credit growth, interest rates, equity prices (and their volatility) and bond spreads.

We can use a two-state one factor return generating model, which can be used under the framework of Basel II to calibrate risk weights (Rosch (2003) develops an application using a similar modelling setup). The discrete-time process for the return on a firm's assets (R_{it}), in a given period, follows a one-factor model defined as:

$$R_{it} = \sqrt{\rho} X_t + \sqrt{(1-\rho)} Z_{it} \tag{3}$$

where $X_t \sim N(0,1)$, $Z_{it} \sim N(0,1)$ (normalised returns assumption).

The exposure to the common factor is given by $\sqrt{\rho}$. If we consider that the idiosyncratic component is independent from the systematic factor, as well as independent across borrowers, then ρ measures the correlation between the normalised asset returns of any two borrowers.

However, these assumptions may be too strict. In fact, empirical evidence suggests that the idiosyncratic component is not independent from the systematic factor: when overall economic conditions deteriorate, it is likely that, for instance, profitability will decline simultaneously for several firms, most notably in sectors more sensitive to business cycle conditions. Furthermore, the assumption of a unique common factor may be too poor if we want to fully understand which factors are more important in driving default probabilities. So, taking into account these considerations, the return model may be slightly adapted, yielding:

$$R_{it} = \Gamma X_t + \Delta Z_{it} \tag{4}$$

where Γ and Δ are parameter vectors, which can be estimated through a linear panel model such as:

$$R_{it} = \alpha + \gamma X_t + \delta Z_{it} + u_{it} \tag{5}$$

where α is the model constant term, γ and δ are the estimates of Γ and Δ, and u_{it} is the error-term.

The borrower will default if his returns fall below a given threshold c_{it}:

$$R_{it} \leq c_{it} \Leftrightarrow Y_{it} = 1 \tag{6}$$

The realisation of the risk drivers X_t and Z_{it} and of the default indicator Y_{it} is observable, but the returns R_{it} are not. The link between the risk factors and the default probability can be accomplished with a threshold model. So, we can redefine the probability of default at time t for borrower i (time-discrete hazard) as:

$$\begin{aligned} \lambda_{it} &= \text{Prob}\,(Y_{it} = 1) = \text{Prob}\,(R_{it} \leq c_{it}) = \\ &= \text{Prob}\,(\Gamma X_t + \Delta Z_{it} \leq c_{it}) = \phi(c_{it}) \end{aligned} \tag{7}$$

where $\phi(.)$ denotes the cumulative standard normal distribution function.

Taking into account the estimated linear panel model we can write:

$$
\begin{aligned}
\lambda_{it}(X_t, Z_{it}) &= \text{Prob}\,(Y_{it} = 1 \mid X_t, Z_{it}) = \\
&= \text{Prob}\,(R_{it} \leq c_{it} \mid X_t, Z_{it}) = \\
&= \text{Prob}\,(\alpha + \gamma X_t + \delta Z_{it} + u_{it} \leq c_{it} \mid X_t, Z_{it}) = \\
&= \text{Prob}\,(u_{it} \leq c_{it} - \alpha - \gamma X_t - \delta Z_{it} \mid X_t, Z_{it}) = \\
&= F(\tilde{\alpha} + \tilde{\gamma} X_t + \tilde{\delta} Z_{it})
\end{aligned}
\tag{8}
$$

where $F(.)$ denotes the cumulative distribution function of the error term, $\tilde{\alpha} = c_{it} - \alpha$ (assuming $c_{it} = c, \forall_{it}$), $\tilde{\gamma} = -\gamma$ and $\tilde{\delta} = -\delta$.

Before exploring the information available at the firm-level, we will begin by trying to draw some conclusions on the relationship between credit risk and macroeconomic information at an aggregate level. Hence, this modelling setup will be applied further ahead, in Section 4, where we will thoroughly use micro data to assess the determinants of default probabilities at the firm-level.

3 Credit risk and macroeconomic dynamics: an aggregate approach

Before looking at evidence provided by firm-level data, we will try to understand some of the links between credit risk and macroeconomic developments at an aggregate level. In order to achieve such objective, we built up correlation matrices between the cyclical components of credit overdue and of a large set of macroeconomic and financial variables. These matrices may provide a clearer understanding of the cyclical comovement between credit overdue and other variables, which can later be used as explanatory variables under a regression analysis framework, together with firm-specific variables.

3.1 Data and methodology

In order to evaluate the relationship between credit risk and macroeconomic developments, we gathered a large set of macroeconomic and financial time series[2]. In this analysis, credit default is measured as credit and interest which have become overdue within the last 3 to 6 months[3]. There is, however, one caveat in using this measure of credit overdue: it is not possible to separately assess the evolution of non-financial corporations' and households' credit overdue. To partly overcome this issue, estimations were also performed using the stock of non-performing loans of non-financial corporations, though this stock variable should not perform so well in capturing the dynamics of new credit overdue.

Macroeconomic and financial series include information on national accounts, inflation, labour market data, loans, loan loss provisions, interest rates and stock market prices. All time

[2] Data sources are listed in Appendix A.

[3] An alternative would be to use instead credit overdue for a period shorter than 3 months, but this variable may display some spurious volatility. Nevertheless, all the estimations described below were also performed considering this shorter horizon of credit overdue. The results proved to be generally robust.

series, considered at a quarterly frequency, were detrended using the Hodrick-Prescott filter[4]. The cyclical components obtained through filtering were then used to compute correlations with our aggregate credit risk measure, considering several time lags.

3.2 Some results

As mentioned above, the analysis of the cyclical components of several macroeconomic and financial variables (and of their correlation with non-performing loans) may shed some light on the links between credit risk and overall macroeconomic developments. A large set of time series was taken into account. Table 1 reports some of the most significant correlation coefficients obtained for the period comprised between 1990Q1 and 2004Q4[5].

First of all, the correlation between loans to non-financial corporations at t and credit overdue at $t+5$ is quite high and positive (0.64), as illustrated in the first panel of Table 1. This evidence helps to support the hypothesis that most credit risk is built up during periods of strong credit growth, materialising only when the economy hits a downturn, as discussed above (see Pederzoli and Torricelli (2005) or Jiménez and Saurina (2006), for instance). In turn, the correlation between loans to non-financial corporations at t and credit overdue at $t-8$ is also relatively high, but it is now negative (-0.41), implying that a strong growth in credit overdue at t is correlated with a contraction in total credit at $t+8$. This may suggest that banks apply tighter standards on loan approval after a period in which non-performing loans increase significantly. Moreover, in a period of economic slowdown, loan demand is expected to remain subdued.

The cyclical component of provisions has, as expected, a positive correlation with new credit overdue. The correlation is stronger when provisions for general credit risks are added up to specific provisions (total provisions)[6].

The cyclical component of GDP displays a positive leading correlation with the cycle of credit overdue (the strongest correlation is seen between GDP at t and new credit overdue at $t+8$). This result implies that a period of robust economic growth is usually followed by an increase in new credit overdue, with a lag of at least two years. This result is also important to confirm the hypothesis that in periods of economic growth there may be some tendency towards excessive risk-taking, which materialises in an increase of credit overdue only when the economy hits a downturn. The negative contemporaneous correlation is particularly strong if we consider the stock of non-financial corporations' non-performing loans instead of the flow of new credit overdue. In sum, in periods of strong economic growth imbalances may be building up. According to our results, these imbalances start to be gradually reflected in new credit overdue with a lag of at least two years. Then, the growth of new credit overdue is progressively reflected in an increase of the stock of non-performing loans. When the cyclical component of non-performing loans reaches its peak (that is, when credit risk fully materialises), the cycle

[4] The smoothing parameter was set to be 1600.

[5] The correlation coefficients displayed in Table 1 refer to the correlation of each variable with the cyclical component of the (log) amount of credit overdue for a period comprised between 3 and 6 months, taking into account quarterly data.

[6] Provisions for general credit risks are a function of total credit granted, whereas specific provisions are built up when credit overdue is recorded in banks' balance sheets.

Table 1

Correlation coefficient of x_t with credit overdue$_{t+i}$

x_t:	i = -8	i = -7	i = -6	i = -5	i = -4	i = -3	i = -2	i = -1	i = 0	i = 1	i = 2	i = 3	i = 4	i = 5	i = 6	i = 7	i = 8
Loans																	
Loans to non-financial corp.	-0.41	-0.32	-0.28	-0.18	-0.03	0.15	0.27	0.31	0.41	0.52	0.56	0.62	0.62	0.64	0.58	0.50	0.43
Agriculture	-0.22	-0.07	0.08	0.19	0.37	0.46	0.41	0.35	0.36	0.21	0.11	0.18	0.19	0.26	0.27	0.31	0.22
Mining	-0.27	-0.21	-0.12	-0.03	0.11	0.27	0.33	0.36	0.35	0.25	0.20	0.27	0.30	0.43	0.49	0.53	0.48
Manufacturing	-0.28	-0.17	-0.14	-0.07	0.05	0.19	0.30	0.33	0.38	0.47	0.52	0.64	0.63	0.55	0.46	0.46	0.45
Utilities	-0.38	-0.29	-0.17	-0.05	-0.01	0.04	0.09	0.08	0.15	0.23	0.31	0.41	0.37	0.48	0.54	0.60	0.51
Construction	-0.46	-0.38	-0.29	-0.14	0.02	0.15	0.26	0.33	0.41	0.45	0.50	0.54	0.53	0.54	0.51	0.40	0.26
Services	-0.38	-0.32	-0.31	-0.23	-0.08	0.14	0.25	0.30	0.41	0.55	0.58	0.60	0.61	0.65	0.60	0.47	0.40
Provisions																	
Specific loan loss provisions	0.48	0.30	0.29	0.46	0.50	0.25	0.16	0.14	0.09	-0.29	-0.44	-0.48	-0.45	-0.39	-0.43	-0.23	-0.02
Total provisions	0.12	-0.01	-0.13	0.06	0.24	0.17	0.17	0.33	0.52	0.42	0.40	0.43	0.54	0.57	0.45	0.47	0.47
National accounts																	
Private consumption	-0.40	-0.38	-0.40	-0.30	-0.17	-0.12	-0.04	0.12	0.27	0.21	0.23	0.40	0.52	0.56	0.59	0.69	0.74
Durables	-0.48	-0.39	-0.43	-0.36	-0.27	-0.26	-0.26	-0.19	-0.06	-0.05	0.03	0.17	0.33	0.45	0.45	0.57	0.66
Non-durables	-0.26	-0.29	-0.30	-0.20	-0.09	-0.02	0.09	0.27	0.40	0.32	0.31	0.47	0.53	0.52	0.55	0.62	0.63
Public consumption	-0.46	-0.47	-0.43	-0.36	-0.24	-0.11	-0.01	0.11	0.23	0.27	0.33	0.42	0.52	0.63	0.69	0.71	0.67
GFCF	-0.61	-0.58	-0.50	-0.46	-0.40	-0.42	-0.38	-0.33	-0.27	-0.15	0.04	0.24	0.38	0.53	0.59	0.58	0.53
Exports	-0.10	-0.03	0.03	-0.07	-0.21	-0.37	-0.42	-0.49	-0.54	-0.38	-0.29	-0.19	-0.09	-0.01	0.04	0.04	0.10
Goods	0.03	0.13	0.20	0.11	-0.05	-0.24	-0.37	-0.49	-0.54	-0.40	-0.34	-0.29	-0.19	-0.10	-0.06	-0.10	-0.05
Services	-0.37	-0.41	-0.34	-0.40	-0.42	-0.43	-0.32	-0.27	-0.31	-0.20	-0.04	0.10	0.18	0.20	0.23	0.32	0.36
Imports	-0.38	-0.31	-0.25	-0.31	-0.36	-0.47	-0.41	-0.40	-0.46	-0.34	-0.15	-0.04	0.14	0.29	0.38	0.37	0.42
Goods	-0.37	-0.28	-0.22	-0.27	-0.34	-0.44	-0.39	-0.40	-0.47	-0.36	-0.20	-0.09	0.09	0.26	0.36	0.34	0.41
Services	-0.16	-0.28	-0.28	-0.30	-0.23	-0.33	-0.31	-0.13	-0.05	-0.01	0.21	0.32	0.35	0.26	0.30	0.35	0.21
GDP	-0.52	-0.45	-0.39	-0.26	-0.14	-0.16	-0.20	-0.16	-0.04	0.08	0.14	0.35	0.42	0.45	0.47	0.56	0.57
Other economic indicators																	
Coincident indic. for econ. act.	-0.57	-0.60	-0.60	-0.59	-0.56	-0.54	-0.52	-0.45	-0.36	-0.26	-0.13	0.01	0.16	0.26	0.33	0.40	0.44
Inflation																	
CPI growth	0.08	0.24	0.17	0.15	0.21	0.28	0.29	0.45	0.58	0.54	0.37	0.37	0.42	0.35	0.16	0.11	0.15
Bank interest rates																	
Interest rate on firms	-0.15	-0.15	-0.17	-0.21	-0.24	-0.24	-0.22	-0.18	-0.13	0.10	0.24	0.27	0.35	0.29	0.25	0.25	0.30
Interest rate housing	-0.20	-0.22	-0.23	-0.27	-0.29	-0.26	-0.22	-0.17	-0.13	0.12	0.29	0.32	0.39	0.33	0.29	0.26	0.28
Interest rate households other	-0.03	-0.05	-0.12	-0.19	-0.22	-0.24	-0.24	-0.23	-0.20	0.02	0.15	0.19	0.28	0.22	0.20	0.22	0.29
Stock market data																	
PSI Geral	-0.35	-0.31	-0.36	-0.46	-0.38	-0.31	-0.39	-0.41	-0.37	-0.17	-0.07	0.00	0.16	0.11	0.03	0.06	0.27
PSI 20	-0.32	-0.26	-0.29	-0.36	-0.25	-0.14	-0.20	-0.18	-0.10	0.04	0.06	0.10	0.22	0.33	0.23	0.11	0.26
Bond yields																	
Gov bond DE 5	-0.22	-0.11	0.00	0.12	0.16	0.02	-0.10	-0.09	-0.10	-0.15	-0.15	0.04	0.22	0.33	0.39	0.42	0.34
Gov bond DE 10	-0.04	0.07	0.14	0.23	0.28	0.13	-0.02	-0.02	-0.03	-0.12	-0.16	-0.01	0.16	0.27	0.32	0.35	0.30
Gov bond EMU 10	-0.05	0.00	0.00	0.02	-0.02	-0.16	-0.26	-0.26	-0.26	-0.09	0.06	0.13	0.27	0.24	0.25	0.27	0.34

Note: This table reports correlation coefficients between the cyclical component of each listed variable at t (x_t) and the cyclical component of credit overdue at different time periods (credit overdue$_{t+i}$), using quarterly information for the period comprised between 1990Q1 and 2004Q4 (except for Government bond yields, for which only slightly shorter time series are available). This definition of credit overdue comprises credit and interest overdue for more than 3 and less than 6 months. The highest correlation for each variable is highlighted in grey. The coincident indicator for economic activity refers to the cyclical component underlying the construction of this business cycle indicator.

of GDP is at its trough, resulting in a strong negative contemporaneous correlation between these two variables[7].

In what concerns GDP components, the results are rather mixed. On the one hand, the cyclical component of private (and public) consumption displays a positive leading correlation with the cycle of new credit overdue. On the other hand, the cyclical components of investment (measured by gross fixed capital formation), imports and exports display a negative lagged correlation. This may suggest that credit imbalances are usually more associated with consumption-driven expansions, though this conclusion may be very sensitive to the time interval considered.

Bank interest rates display a positive correlation with the cyclical component of new credit overdue and are leading variables. In fact, their strongest correlation is seen at $t+4$, suggesting that an increase in credit overdue is often preceded by an interest rate increase. This result may be associated, on one hand, with the increase of interest rates during periods of stronger and prolonged economic growth, which are sometimes followed by an increase in credit overdue, as discussed above. Additionally, a sizeable increase in interest rates implies a higher debt service, which may put some strain on highly leveraged firms. On the other hand, when interest rates increase significantly, adverse selection problems may become more frequent, implying higher default rates some periods afterwards[8].

Government bond yields display a pattern similar to that of bank interest rates. Stock market indices exhibit a negative correlation, implying that positive developments in stock market prices, which usually reflect a broad-based improvement in firms' financial condition, are usually associated with lower default ratios, as should be expected.

The links between credit default and macroeconomic developments will be further explored in the next section, by taking simultaneously into account firm-specific and macroeconomic variables under a regression analysis framework. The insight provided by the analysis of cyclical components will then be helpful in choosing the set of explanatory variables to be considered.

4 The contribution of firm level information to understand loan default

In the previous section we discussed some of the determinants of loan default at an aggregate level, using macroeconomic and financial time series. However, firm level data may provide a much richer insight of credit risk drivers. Even though macroeconomic and financial conditions may offer a valuable contribution to explain credit risk at an aggregate level, it is the firm's specific financial situation that will ultimately determine whether it will default on its liabilities.

[7]Most of the empirical literature on credit risk modelling focuses on the cross-section rather than on the time-series dimension of credit risk. An exception is a recent work by Koopman and Lucas (2005), which uses a multivariate unobserved components approach to evaluate the dynamic behaviour of business failures, credit spreads and GDP in the United States, using considerably long time series (1933-1997). The authors find evidence of negative co-cyclicality between GDP and business failures for long business cycles (with an average duration of 11 years) and a positive relationship between the cyclical component of business failures and credit spreads.

[8]Borrowers with projects which entail relatively low risks may consider that interest rates are higher than what is deemed adequate to ensure minimum profitability levels, thus introducing a potential bias in banks' loan portfolios towards riskier borrowers.

As a consequence, taking into account firm-specific characteristics, such as leverage, profitability or solvency, may provide a clearer understanding of which factors drive credit risk. Moreover, as argued by Benito *et al* (2004) or Bunn and Redwood (2003), the use of firm-level data may improve the precision and accuracy of empirical estimates, by exploring how the combination of factors at the micro-level influences credit risk, taking advantage from the greater variability in such datasets. Furthermore, firm-level data allow us to take into account distributional issues, instead of focusing only on mean behaviour. For example, it is possible to separately consider credit risk for different economic sectors, for different firm size cohorts or even to separately look at firms which have previously defaulted.

In this section we will explore an extensive and detailed dataset which comprises information on more than 30.000 Portuguese firms. We will begin by describing the dataset, presenting some revealing summary statistics. Then we will briefly describe the econometric methodology used. First we use discrete choice models to better understand what drives firms' loan defaults. Afterwards, we complement our analysis using duration models. The time dimension encompassed in duration models allows us to focus on the time it takes for a loan to default, rather than simply considering whether or not firms default.

4.1 Data and summary statistics

The microeconomic dataset used in this work comprises two distinct datasets held by Banco de Portugal, namely, the Central Credit Register and the Central Balance Sheet Database. The Central Credit Register provides information on all credit exposures above 50 euro in Portugal. The information contained in this database is reported by credit institutions (reporting is mandatory) and its main objective is sharing information between participant institutions, in order to improve their credit risk assessment and management. This database contains monthly information on loans granted to firms and households, including their current status (it is possible to know whether credit has become overdue, if it was written-off banks' balance sheets, if it was renegotiated or if it is an off-balance sheet risk, such as the unused parts of credit lines or bank guarantees)[9]. Using end-of-year data for the period comprised between 1996 and 2002, we have 203.655 observations[10]. The Central Balance Sheet Database provides detailed accounting information for a large sample of Portuguese firms, being used mostly for economic and statistical purposes. We use annual data, though quarterly data is also available for a smaller set of firms. Reporting is not compulsory, but the sample is considered to be representative. Nevertheless, there may exist some bias towards larger firms. Even though this bias represents a shortcoming of this database, it still is an extremely rich and unique dataset on non-financial corporations. Furthermore, even though smaller companies may, in some cases, present higher default probabilities, they do not usually hold large amounts of debt. The systemic implications of a firm's failure result not only from its default probability, but mostly from its debt at risk (default probability multiplied by the firm's total debt). For instance,

[9] Reporting banks aggregate information on loans with similar status for each firm (information is not reported on a loan by loan basis). There is no information on loan maturity, collateral or interest rates. In what concerns loan maturity, nearly half of the loans granted to non-financial corporations in the period under analysis had maturities above one year, taking into account aggregate statistics.

[10] In order to merge the two datasets, loans were aggregated within firms. Hence, one observation is defined as a pair firm-year, summing up all credit liabilities for a given firm in each year.

Benito *et al* (2004), using a similar database for Spanish firms, find that overall financial risk in the non-financial corporate sector is highly concentrated in a few large firms, not because of their high default probability, but rather because of their size. As a consequence, even though micro and small firms may be under-represented in this sample, they account only for a small part of banks' total credit exposures. In this dataset we have 153.581 observations for the period comprised between 1996 and 2002. Merging the two databases we obtain a dataset containing 113.119 observations, comprising 33.084 firms.

We constructed several ratios and indicators to evaluate each firms' financial situation, namely in what concerns their profitability, financial structure, leverage, productivity, liquidity and investment. A detailed description of the ratios used to assess firms' financial conditions is presented in Appendix B. In order to avoid spurious results, all the ratios for which the denominator equalled (or was close to) zero were considered as missing values. Furthermore, some ratios in which a negative denominator (such as equity) was combined with a negative numerator (such as profits) were also considered as missing values, in order to avoid misleading results. In both cases, in order to avoid loosing information, these variables were excluded from the regression analysis whenever possible and equivalent ratios without such problems were taken into account. Finally, in variables where there were significant outliers, we replaced observations above the 99th percentile with the value of that percentile (the same procedure was applied to observations below the 1st percentile, whenever necessary).

In Table 2 we present some descriptive statistics, in order to shortly characterize the sample of firms under analysis. On average, each firm in the sample has 54 employees, implying that we are dealing mostly with small and medium enterprises. Net profits show a remarkable dispersion, ranging from extremely negative values to considerably high positive values. Profitability ratios are, on average, relatively low. Average return on equity stands at 5.2 per cent and return on assets is negative for nearly a quarter of the firms included in the sample. Sales growth rates also show a high dispersion within the sample. The same is true for the solvency ratio, defined as equity as a percentage of total assets. The lower the value of this ratio, the higher is the firm's dependence on external funding sources. Credit accounts for 11.5 per cent of firms' assets, on average, though there are very different situations in the sample (this ratio varies between 0 and 80 per cent). Capital and labour productivity, which assess the productive contribution of the inputs used by firms, also vary considerably within the sample. The capital-labour coefficient, measured as the ratio between tangible assets and the number of employees, allows to control for the firm's capital intensity, which should vary considerably across specific economic activities. The investment rate, measured as new investment (deducted from depreciations) over total sales, is relatively low (2.3 per cent, on average). This ratio should reflect, to some extent, the investment and growth policy followed by each firm, though it should be conditioned by the sector in which the firm operates or by its age, for instance. On average, firms included in the sample are 16 years old. The liquidity ratio, which evaluates short-term assets as a percentage of the firm's debt, is significantly above 100 per cent, on average, implying that a large part of the firms included in the sample displays sustainable liquidity levels during the period considered. Only a small percentage of firms in the sample has credit overdue. In fact, on average, credit overdue represents only 1 per cent

Table 2

Panel a - Summary statistics - all sample

	N	mean	sd	min	p5	p25	p50	p75	p95	max	skewness	kurtosis	Baseline regression sample N	Baseline regression sample mean
Total assets	113,119	8,902,194	125,000,000	339	57,746	231,911	688,261	2,404,954	19,200,000	13,000,000,000	55.4	4,264	71,058	8,139,104
Employees	113,119	54	286	1	2	5	13	35	188	18688	33.1	1,570	71,058	55
Net profits	113,119	196,367	7,502,392	-359,000,000	-128,450	155	6,569	36,836	497,501	1,130,000,000	45.7	5,898	71,058	129,289
ROA	113,119	0.5	14	-200.0	-15.9	0.0	1.2	4.1	14.0	200.0	-4.9	69.7	71,058	0.8
ROE	100,662	5.2	46	-282.0	-44.3	0.8	6.0	16.8	55.8	145.4	-3.1	22.0	64,665	4.8
Oper. results as % of equity	100,662	25.8	63	-225.2	-34.9	4.3	17.1	38.6	116.8	351.5	1.3	13.7	64,665	24.5
Sales growth	80,035	12.8	49	-100.0	-38.0	-7.7	5.3	21.1	81.4	308.8	3.3	19.1	71,058	12.9
Solvency ratio	113,119	23.5	33	-139.1	-25.6	10.7	24.6	40.5	71.4	100.0	-1.7	10.0	71,058	24.6
Total credit as a % assets	113,119	11.5	16	0.0	0.0	1.2	4.7	14.7	47.1	80.2	2.1	7.6	71,058	14.9
Labour productivity	113,119	118,798	181,643	3	11,840	29,043	58,423	128,071	420,464	1,231,148	3.9	20.9	71,058	121,514
Capital productivity	112,332	2,492	6,709	0	71	262	642	1,750	9,574	51,393	5.6	37.1	70,735	2,318
K_L coefficient	113,119	22,962	44,404	0	613	3,675	9,476	23,175	81,972	328,962	4.8	29.6	71,058	23,371
Investment rate	80,035	2.3	16	-49.1	-12.9	-2.3	-0.1	3.4	25.0	98.3	2.7	18.6	71,058	2.5
Firm age (years)	113,119	16	16	0	2	6	12	21	48	100	2.5	10.6	71,058	18
Liquidity ratio	112,618	125	116	0	27	73	102	133	295	865	4.1	23.5	71,058	120
Liquid assets to total assets	112,655	68	24	0	21	52	73	89	98	100	-0.7	2.6	71,058	68
Credit overdue	113,119	5,327	211,443	0	0	0	0	0	0	44,600,000	118	20,021	71,058	995
Credit overdue as % total credit	103,201	1.03	9	0.0	0.0	0.0	0.0	0.0	0.0	100	9.8	102.5	71,058	0.30
Dummy credit overdue	103,201	0.03	0	0.0	0.0	0.0	0.0	0.0	0.0	1	5.5	31.5	71,058	0.02

Note: total assets, net profits and credit overdue are presented in euros. ROA, ROE, operational results as % of equity, sales growth, the solvency ratio, total credit as a % of assets, capital and labour productivity, the investment rate, the liquidity ratios and credit overdue as a % of total credit are displayed as percentages. K_L coefficient displayed as euro per person.

Panel b - Summary statistics - default frequencies by firm size and sector

Default frequencies by sector

	No. obs.	Default frequency (%)
Fishing	277	11.19
Mining	1084	5.17
Agriculture	3487	2.81
Manufacturing	41427	3.76
Utilities	355	1.97
Construction	14020	3.25
Commerce	31793	1.83
Tourism & restaurants	1405	4.70
Transports & commun.	6004	2.60
Real estate	2319	2.29
Education	249	3.21
Healthcare	331	1.21

Default frequencies by size

	No. obs.	Default frequency (%)
Micro	39725	2.67
Small	42608	2.72
Medium	16548	4.21
Large	4320	3.89

of total bank loans[11]. The mean value of the dummy variable credit overdue (which takes the value 1 when a firm records a loan default) can be interpreted as a historical default probability, standing at 3 per cent during the period under analysis (we observe 3084 defaults, using end of year data). The higher default frequencies are seen in fishing, mining, tourism and restaurants, and manufacturing, as illustrated in the second panel of Table 2.

The highest default frequencies are recorded by medium-sized firms, closely followed by large firms, as depicted in the last panel of Table 2[12]. Hence, firms in default seem to be slightly bigger firms, contrary to what is usually seen in the literature. For instance, Bhattacharjee *et al* (2002), Bunn and Redwood (2003), Eklund *et al* (2001) and Jiménez and Saurina (2004) find

[11] Taking into account only those firms which actually default, credit overdue represents, on average, 34 per cent of their total loans.

[12] However, the average amount of credit overdue as a percentage of total credit is higher for medium and micro firms. For robustness, different definitions of firm size were tested, but these results remain valid.

that smaller firms are more likely to default. In turn, Pain and Vesala (2004) and Bernhardsen (2001) conclude that any systemic effect of firm size on default is relatively small. Furthermore, there is also contrary evidence on the impact of firm size in the literature. According to Moody's (2004), larger firms default less often, but when financial statement ratios are taken into account, the impact of the size advantage declines. Hence, a small firm with healthy financial ratios should not be riskier than a large firm with comparable financial statements. Finally, Benito *et al* (2004) obtain a result similar to ours, observing a positive relationship between firm size and default rates (the authors argue that their database may be biased towards "good" companies, which may also be a problem in our database).

One of our main objectives is to understand what drives credit risk at the firm-level. This can be partly accomplished by looking separately at summary statistics for firms which record a loan default at t, comparing them with the remaining firms. In Table 3 we present the mean values for these two groups of firms for several potentially interesting variables. A brief analysis confirms that firms with loan defaults seem in fact to differ from other firms. The profitability of firms in default is, as expected, considerably lower (their return on assets is, on average, negative). Furthermore, sales growth is also lower for these firms, though positive. The solvency ratio suggests that firms in default are more dependent of external funding sources. In fact, the ratio of credit to total assets and the leverage ratio confirm that these are more indebted firms. Investment rates for troubled firms are also considerably lower. The liquidity ratio is also lower (and below 100 per cent, on average) for firms in default, revealing that these firms are usually subject to stronger liquidity pressures, as should be expected. On average, firms in default are slightly older, which is not quite what should be expected. Again, there is mixed evidence in the literature in what concerns the impact of firm age. Younger firms should be more sensitive to external shocks and should be expected to show higher bankruptcy probabilities, as argued by Eklund *et al* (2001), for instance. In turn, Shumway (2001) finds no evidence of duration dependence in bankruptcy probabilities (firm age is never statistically significant, after controlling for other firm characteristics). The positive correlation between firm age and default frequencies in our sample may reflect positive duration dependence: the longer the firm is at risk, the higher should be its default probability. This issue will be discussed in more detail with the results obtained using duration models. Finally, taking into account mean values for both groups, the results obtained confirm that firms in default are, on average, slightly larger than the remaining firms in the sample.

In order to more accurately test if these variables are in fact different for firms which default, we also present in Table 3 the results of a mean comparison Welch test[13]. For all variables considered, the mean values for firms in default are statistically different from the mean values observed for firms without default (with the exception of total assets, for which the mean value is not statistically different between the two groups of firms)[14]. Hence, the set of variables considered in this table may contribute to explain why do some firms default, under a regression analysis framework. Furthermore, we also computed pairwise correlations for all the

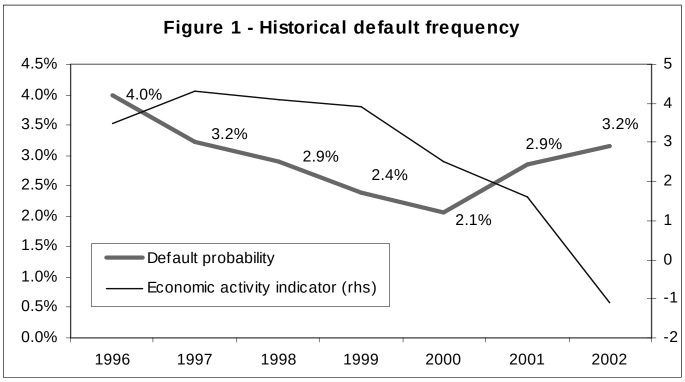
[13] This test should be more adequate than the more common mean comparison Satterthwaite test, given that variances should differ between the two groups of firms.

[14] Pairwise mean comparison tests performed between micro, small, medium and large firms confirmed that differences in average default frequencies between these groups are statistically significant.

Table 3

Summary statistics - comparing firms with and without defaults

	Mean values for firms with no default at t	Mean values for firms in default at t	Welch test - Ho: diff = 0 t-ratio	Degrees of freedom	diff = mean (no def.) - mean (def.)	Ha: diff not 0 Pr(|T| > |t|)	Mean is significantly different (Y/N)
ROA	0.5	-4.9	15.18	3178	5.4	0.00	Y
Sales growth	12.9	5.7	5.74	2252	7.2	0.00	Y
Solvency ratio	23.2	1.1	26.32	3171	22.1	0.00	Y
Total credit as a % assets	12.5	16.9	-12.02	3209	-4.4	0.00	Y
Leverage	76.8	98.9	-26.32	3171	-22.1	0.00	Y
Investment rate	2.6	-2.5	11.89	2248	5.1	0.00	Y
Liquidity ratio	119.0	86.5	20.75	3356	32.5	0.00	Y
Firm age	16.3	18.6	-7.52	3252	-2.3	0.00	Y
Total assets	9123577	9771957	**-0.17**	3149	-648380	0.87	**N**
Employees	53.6	59.3	-2.47	4389	-5.7	0.01	Y
Number of observations	100117	3084					

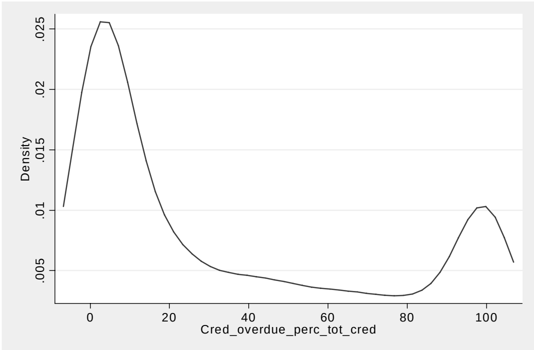


variables in the dataset, identifying which pairwise correlations are significant at a 5 per cent significance level. This correlation matrix was used as a guidance tool to choose relevant firm-specific and macroeconomic variables, as well as to identify possible multicollinearity problems between explanatory variables.

Figure 1 illustrates the evolution of historical default frequencies during the sample period, depicted against the economic activity coincident indicator. Until 2000, there was a steady decline in default frequencies, accompanied by positive economic developments. The deterioration of economic conditions was then mirrored (with some lag) by an increase in observed default frequencies, as well as in the amount of credit overdue as a percentage of total credit.

The empirical distribution of this latter ratio is depicted in Figure 2, using a gaussian kernel density. The distribution of this ratio is clearly two-peaked: either firms record only small amounts of credit overdue (as a percentage of their total credit liabilities), which may reflect transitory episodes of delinquency, or they default on nearly all their debt, which should be a situation closer to bankruptcy. In this domain, it may be interesting to notice the differences seen when firm size is taken into account. As mentioned above, large and medium-sized firms display higher default rates, in contrast to what is usually found in the literature. Nevertheless, the empirical distribution of the ratio between credit overdue and total credit is remarkably different for firms with different sizes, as illustrated in Figure 3. In fact, whereas the distribution for micro firms is clearly two-peaked (which is, to a lesser extent, also true for small firms), the distribution for medium and, most notably, for large firms is clearly single-peaked. This result

Figure 2
Empirical distribution function of the credit overdue ratio
(gaussian kernel density)



may suggest that even though larger firms display higher default frequencies in our sample, these usually reflect small and, most likely, transitory episodes of loan default. Larger firms may have fewer difficulties in overcoming credit problems in part because banks may be more willing to renegotiate impaired loans, in order to avoid sizeable losses.

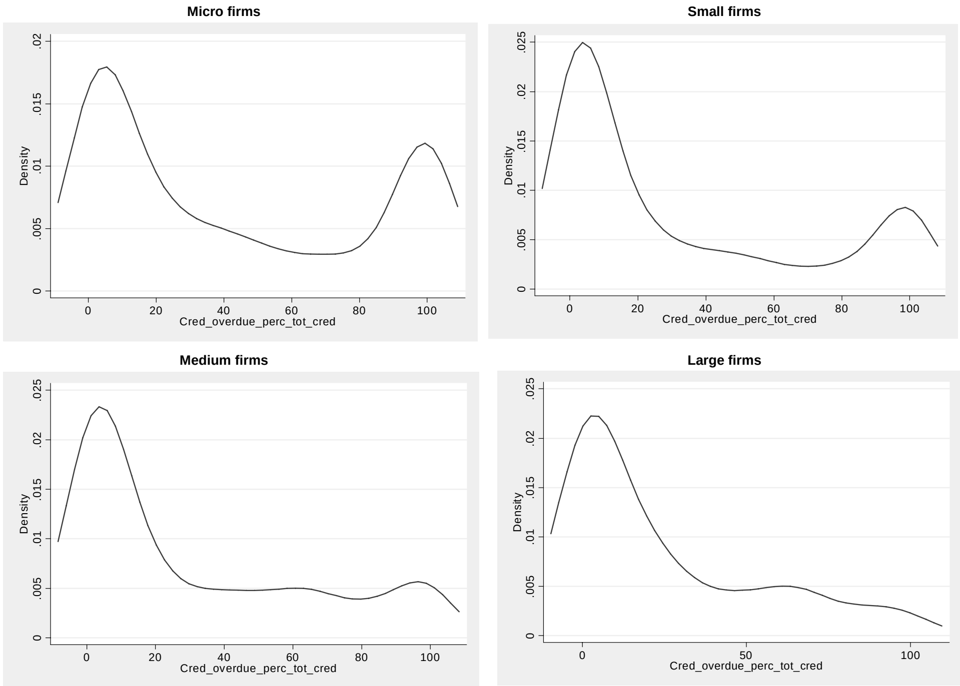
By setting up a transition matrix, it is possible to evaluate historical default probabilities at different time horizons. In Table 4 (panel A) are presented average default frequencies at different time horizons and for different years. Default probabilities are fairly stable during the first years (decreasing slightly until $t + 3$), but increase considerably afterwards. Such pattern may signal positive duration dependence: the longer the firm is at risk, the higher should be its default probability. By defining conditional transition probabilities, we can trace separately the evolution of the risk profile of firms which are in default at t (Table 4.B). We can see that this evolution is extremely different for firms with and without default at t. For firms which are not in default at t, default probabilities are clearly increasing over time, which was not clear when we considered all firms in the sample. In turn, for firms which have defaulted at t, recovery probabilities (defined as 100 per cent less the default probability) are markedly increasing and larger than 75 per cent after 6 years. Finally, we can also build conditional transition matrices for firms without any prior default (in the sample period) and also for firms which did not default at t, but which recorded at least one previous default episode in the sample period, as depicted in Table 4.C (this latter group has a very limited number of observations). Default probabilities seem to be slightly lower for firms without any previous default in the sample period, though their time evolution is similar to those with no default at t. In turn, for firms which are not in default at t but had some previous default in the sample period, default probabilities are considerably higher, implying that firms with a past record of credit overdue are more likely to default again in the future than firms that never defaulted before.

4.2 Econometric methodology

A common approach in empirical credit risk literature is to use standard discrete choice models, such as logit or probit models (see, for instance, Benito *et al* (2004), Bernhardsen (2001), Bunn and Redwood (2003), Hamerle (2004) or Campbell *et al* (2005))[15]. These models can be used

[15] For details on discrete choice modelling under a panel data framework see Wooldridge (2002), Hsiao (1986), Baltagi (1995) and Maddala and Rao (1996).

Figure 3
Empirical distribution function of the credit overdue ratio by firm size
(gaussian kernel density)



to empirically examine credit risk drivers, assessing their relative importance in determining whether firms default on their credit liabilities. Recalling equation 8, we want to estimate a linear panel model such as:

$$\lambda_{it}(X_t, Z_{it}) = \text{Prob}\,(Y_{it} = 1 \mid X_{t,} Z_{it}) = \\ = F(\tilde{\alpha} + \tilde{\gamma} X_t + \tilde{\delta} Z_{it}) \tag{9}$$

The model to be estimated will depend on the assumption made on the error distribution function $F(.)$. Assuming a standard normal distribution function $\Theta(.)$ yields a probit model such that:

$$\lambda_{it}(X_t, Z_{it}) = \Theta(\tilde{\alpha} + \tilde{\gamma} X_t + \tilde{\delta} Z_{it}) \tag{10}$$

If instead we assumed that the error term follows a logistic distribution, we would have:

$$\lambda_{it}(X_t, Z_{it}) = \frac{\exp(\tilde{\alpha} + \tilde{\gamma} X_t + \tilde{\delta} Z_{it})}{1 + \exp(\tilde{\alpha} + \tilde{\gamma} X_t + \tilde{\delta} Z_{it})} \tag{11}$$

The parameters $\tilde{\alpha}$, $\tilde{\gamma}$ and $\tilde{\delta}$ can be estimated by maximising the expected value of the likelihood function:

$$E\left[L(\tilde{\alpha}, \tilde{\gamma}, \tilde{\delta})\right] = \prod_{t=1}^{T} \prod_{i=1}^{N} \left[F(\tilde{\alpha} + \tilde{\gamma} X_t + \tilde{\delta} Z_{it})^{Y_{it}} (1 - F(\tilde{\alpha} + \tilde{\gamma} X_t + \tilde{\delta} Z_{it}))^{(1-Y_{it})}\right] \tag{12}$$

Table 4

A - Transition matrix

Default probabilities at different time horizons (%)	t	t+1	t+2	t+3	t+4	t+5	t+6
1996	**3.99**	3.51	3.29	2.61	2.29	3.41	3.73
1997	**3.23**	3.13	2.54	2.26	3.39	3.63	-
1998	**2.90**	2.51	2.22	3.30	3.52	-	-
1999	**2.38**	2.12	3.14	3.41	-	-	-
2000	**2.07**	3.00	3.26	-	-	-	-
2001	**2.86**	3.24	-	-	-	-	-
2002	**3.16**	-	-	-	-	-	-
Average	**2.99**	**2.95**	**2.89**	**2.86**	**3.04**	**3.52**	**3.73**

B - Conditional transition matrix

Default probabilities at different time horizons (%)	t	t+1	t+2	t+3	t+4	t+5	t+6
Firms with no default at t	0.00	1.63	2.06	2.23	2.57	3.07	3.34
Firms in default at t	100.00	54.62	42.41	36.72	31.45	30.04	23.97
Average	2.99	2.95	2.89	2.86	3.04	3.52	3.73

C - Transition matrix for firms with and without default

Default probabilities at different time horizons for firms without any prior default (%)	t	t+1	t+2	t+3	t+4	t+5	t+6
1996	0.00	1.52	2.13	1.81	1.65	2.84	3.34
1997	0.00	1.53	1.51	1.44	2.72	3.19	-
1998	0.00	1.19	1.23	2.61	2.98	-	-
1999	0.00	0.83	2.19	2.67	-	-	-
2000	0.00	1.95	2.42	-	-	-	-
2001	0.00	1.91	-	-	-	-	-
2002	0.00	-	-	-	-	-	-
Average	0.00	1.48	1.89	2.09	2.42	3.01	3.34

Default prob. at different time horizons for firms without default at t but with prior defaults (%)	t	t+1	t+2	t+3	t+4	t+5	t+6
1996	-	-	-	-	-	-	-
1997	0.0	8.5	8.3	6.9	16.0	13.1	-
1998	0.0	9.5	12.9	14.6	17.6	-	-
1999	0.0	10.1	13.7	15.9	-	-	-
2000	0.0	12.8	16.8	-	-	-	-
2001	0.0	12.9	-	-	-	-	-
2002	0.0	-	-	-	-	-	-
Average	0.0	10.9	13.4	13.4	16.9	13.1	-

Discrete choice models may provide an interesting assessment of the determinants of loan default, helping to determine whether or not a firm with given characteristics is likely to default. However, it would also be important to focus on the time dimension of default, understanding not only if a firm will default, but also when will that eventually occur. The timing of loan default is important for establishing a complete risk evaluation, as well as for accurate loan pricing and provisioning. Duration models directly model the survival time of a loan, taking as a dependent variable the time until default. Although not so common, there are some recent applications of survival analysis to credit risk modelling, such as Banasik *et al* (1999), Carling *et al* (2002, 2004) or Couderc and Renault (2005)[16].

Duration models may provide some advantages over discrete choice models, given that they can more easily incorporate the progressive deterioration of a firm's financial situation before default, as they control for each firm's time at risk, as argued by Shumway (2001). In addition, empirical evidence suggests that there may be duration dependence in default risk: firm age (or time at risk) may be an important explanatory variable, as found by Carling *et al* (2002, 2004) or Saretto (2004)[17]. Therefore, traditional logit and probit models, which imply constant hazard rates, may be less accurate than duration models[18]. Furthermore, the explicit introduction of the time dimension in duration models may provide better results when taking into account macroeconomic variables, as argued by Bhattacharjee *et al* (2002). Despite all the advantages provided by duration models, their application to our dataset is somewhat limited, given that there is a strong left censoring problem: most firms in the dataset were created before 1996, implying that firms' time at risk is, for most observations, much larger than the observation period. Though econometric software can handle this, it still limits the conclusions to be drawn from duration models. Hence, duration models will be used mostly to complement and verify the empirical findings obtained with discrete choice models.

Under the duration modelling framework, we define T as the time until a loan defaults[19]. The hazard function can be defined as the probability of a firm defaulting on a short interval $[t, t+dt)$, conditional on not having defaulted before:

$$h(t) = \lim_{dt \to 0} \frac{\text{Prob}(t \leq T < t + dt \mid T \geq t)}{dt} \tag{13}$$

The hazard function represents an instantaneous rate of default per unit of time. The duration distribution function can be defined as $F(t\ )$=Prob$(T < t)$. The survival function is the probability of surviving up to t, and can be defined as:

$$S(t) = \text{Prob}(T \geq t) = 1 - F(t) = \exp\left\{-\int_0^t h(s)d(s)\right\} \tag{14}$$

[16] Shumway (2001), Bhattacharjee *et al* (2002), Roszbach (2004), Saretto (2004) and Antunes (2005) also use duration models, but to address slightly different issues of credit risk.

[17] As discussed in the previous sub-section, in our dataset there seems to be evidence in favour of positive duration dependence, given that older firms display higher default frequencies.

[18] Nevertheless, the inclusion of a duration variable, such as firm age, in logit or probit models, should yield results similar to those obtained with duration models.

[19] Lancaster (1990) provides one of the most complete presentations of duration models. Wooldridge (2002) also provides a brief introduction to these models under a panel data framework.

Whenever T has an exponential distribution, the hazard function $h(t)$ is constant. When the hazard function is not constant, the underlying process is said to exhibit duration dependence. If $\frac{\delta h(t)}{\delta t} > 0$, $\forall t$, there is positive duration dependence, which implies that, in our framework, the probability of default increases with time, for firms which have never defaulted before. If, on the contrary, $\frac{\delta h(t)}{\delta t} < 0$, $\forall t$, there is evidence in favour of negative duration dependence (the longer the firm has remained without defaulting, the lower should be its default probability).

If firms were homogenous, the setup described above could be directly applied. However, we want to focus on the opposite assumption, understanding which firms' specific characteristics determine their default probabilities, as well as their timings. As a consequence, assuming that we have two vectors of time-varying covariates, X_t and Z_{it} (a systematic and a firm-specific component), we must slightly adapt the specifications presented above, such that:

$$h(t, X(t), Z(t)) = \lim_{dt \to 0} \frac{\text{Prob}(t \leq T < t + dt \mid T \geq t, X(t+dt), Z(t+dt))}{dt} \tag{15}$$

where $X(t)$ and $Z(t)$ are the covariates path up to t[20].

We begin our survival regression analysis by using a Cox proportional hazard model, such that:

$$h(t, X(t), Z(t)) = \kappa(X(t), Z(t))h_0(t) \tag{16}$$

where $\kappa(.)$ is a non-negative function of $X(t)$ and $Z(t)$, and $h_0(t)$ is defined as the baseline hazard, which is common to all firms (individual hazard functions differ from each other proportionally, as a function of $\kappa(X(t), Z(t))$. This is a partly non-parametric approach, given that we can estimate unknown parameters of $\kappa(.)$ without specifying the form of the baseline hazard. Under this setup, the regressors do not affect the shape of the overall hazard function, conditioning only the relative failure risk of each firm.

We then extend our analysis by estimating parametric duration models, using several different distribution functions (namely, exponential, Weibull, Gompertz, lognormal, and log-logistic).

4.3 Results

4.3.1 Results obtained using discrete choice models

In Table 5 we present some of the results obtained using a random-effects probit. The first results presented in this table focus on 71.058 observations, for 24.668 different firms, though the full sample comprises 113.119 observations for 33.084 firms (on average, we have 3 years of observations for each firm). This difference results from using variables constructed with information on the previous year (such as sales growth or the investment rate), which excludes from the regressions all observations for 1996, as well as those which do not have two consecutive years of information. Furthermore, several observations have missing values in some of the

[20]It is important to make a distinction between exogenous and endogenous regressors. According to Lancaster (1990), a covariate process $\{x(t)\}$ is exogenous for T if and only if $\text{Prob}(X(t, t+dt) \mid T \geq t + dt, X(t)) = \text{Prob}(X(t, t+dt \mid X(t))$. This means that any regressor whose path is determined independently of whether any particular agent has defaulted or not is exogenous. In our work, all variables used (both time-variant and time-invariant) will be considered exogenous.

variables used, being also naturally excluded from the regression analysis (which explains the slightly different number of observations in some of the models presented). Additionally, we are excluding from the regression analysis firms in default for at least two consecutive years, considering only their first default observation, in order to evaluate only new transitions into the default state (if, however, a firm defaults twice or more during the sample period, but in non-consecutive years, these defaults will be considered as new transitions).

Taking into account the mean comparisons between firms with and without default presented in Table 3, we started by performing some estimations using a limited set of variables. The pairwise correlations previously computed were also taken into account, not only to identify which firm variables are more correlated with default frequencies, but also to avoid possible multicollinearity problems. In the first model presented in Table 5, the set of explanatory variables comprises sales growth, return on assets (ROA), a solvency ratio, an investment rate and a liquidity indicator[21]. Sales growth displays a negative coefficient, suggesting that firms with stronger sales growth rates should have lower default probabilities. Profitability seems to offer an important contribution in explaining why do some firms default, exhibiting also a negative coefficient, as should be expected (more profitable firms should have a more solid financial situation and, consequently, display lower default probabilities). The solvency ratio, which is defined as the ratio between equity and total assets, also suggests that firms with healthier financial conditions are less likely to default on their loan commitments. Moreover, firms with stronger investment rates also show lower default probabilities. In fact, it seems reasonable to admit that firms under financial pressure are not expected to engage in large investment projects. Finally, the liquidity ratio, defined as short-term assets as a percentage of the firm's total debt, has a negative impact on default probabilities, implying that firms facing stronger liquidity constraints may have higher difficulties in paying their debt commitments, which is consistent with the results obtained by Bunn and Redwood (2003) or Benito *et al* (2004), for instance.

Even though the firm-specific variables taken into account seem to play an important role in predicting loan default, they should be seen as contingent on the firm's size, as well as on the sector in which it operates, given that some variables may be more or less important for different types of firms. Therefore, in model 2 we added sector dummies to our first specification (omitting the dummy variable for manufacturing firms). The results for these sector dummies suggest that there may be some differences in credit risk drivers across different sectors. Overall, the coefficients associated with firms' financial ratios remain robust. Though macroeconomic variables will be introduced further ahead, we will include for now year dummies (the omitted dummy is the year 2001), in order to control for any possible systematic effects (model 3). The fact that most of the coefficients for year dummies are significant gives support to the hypothesis that macroeconomic developments may also be important in explaining loan default, as thoroughly discussed in Section 3. Finally, we also included size dummies (model 4). Though micro and small firms seem to have lower default probabilities than larger firms, confirming the results obtained with descriptive statistics, these differences are not statistically significant. Therefore, even though larger firms display higher default frequencies in our sample, after

[21] As mentioned above, all firm-specific variables are described with more detail in Appendix B.

Table 5 - Probit regressions (dependent variable : dummy credit overdue)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
Sales growth	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	0.000	-0.001
	-2.28	*-2.68*	*-2.19*	*-2.20*	*-2.16*	*-1.79*	*-2.18*	*-2.52*	*-0.47*	*-1.97*
ROA	-0.005	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004
	-4.73	*-4.34*	*-3.96*	*-3.95*	*-3.92*	*-3.75*	*-3.93*	*-3.97*	*-4.05*	*-3.66*
Solvency ratio	-0.004	-0.004	-0.005	-0.005	-0.005	-0.006	-0.006	-0.006	-0.007	-0.006
	-6.52	*-7.15*	*-7.56*	*-7.35*	*-7.36*	*-11.16*	*-11.23*	*-11.24*	*-11.87*	*-11.09*
Investment rate	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005		-0.005	-0.005
	-5.38	*-5.35*	*-5.01*	*-4.99*	*-4.99*	*-4.44*	*-4.82*		*-5.18*	*-4.52*
Liquidity ratio	-0.002	-0.001	-0.001	-0.001	-0.001					
	-5.24	*-4.56*	*-4.47*	*-4.48*	*-4.51*					
Firm age					0.001					
					0.63					
Capital productivity						0.000				
						-2.64				
K_L coefficient							0.000			
							4.11			
Share of tangible assets								-0.002		
								-1.49		
Turnover ratio									-0.003	
									-12.01	
Available collateral (aprox.)										0.001
										1.51
Small				-0.044	-0.035	-0.048	-0.006	-0.035	-0.034	-0.044
				-0.52	*-0.41*	*-0.58*	*-0.07*	*-0.42*	*-0.41*	*-0.53*
Micro				-0.013	-0.001	-0.027	0.014	-0.011	-0.059	-0.025
				-0.15	*-0.01*	*-0.32*	*0.16*	*-0.13*	*-0.69*	*-0.29*
Medium				-0.026	-0.022	-0.025	0.007	-0.015	-0.005	-0.023
				-0.30	*-0.25*	*-0.28*	*0.07*	*-0.17*	*-0.06*	*-0.27*
Fishing		0.351	0.358	0.363	0.360	0.400	0.340	0.431	0.234	0.369
		1.48	*1.42*	*1.45*	*1.43*	*1.60*	*1.35*	*1.74*	*0.93*	*1.46*
Mining		0.215	0.222	0.223	0.224	0.238	0.204	0.240	0.148	0.228
		1.62	*1.57*	*1.57*	*1.58*	*1.68*	*1.43*	*1.71*	*1.05*	*1.60*
Agriculture		-0.197	-0.191	-0.195	-0.194	-0.181	-0.239	-0.182	-0.306	-0.194
		-2.14	*-1.95*	*-1.98*	*-1.96*	*-1.82*	*-2.38*	*-1.85*	*-3.07*	*-1.94*
Utilities		-0.423	-0.492	-0.500	-0.492	-0.434	-0.786	-0.446	-0.622	-0.473
		-1.22	*-1.34*	*-1.36*	*-1.34*	*-1.19*	*-2.06*	*-1.26*	*-1.70*	*-1.29*
Construction		0.030	0.040	0.039	0.041	0.028	0.030	0.019	-0.027	0.035
		0.64	*0.81*	*0.78*	*0.82*	*0.55*	*0.59*	*0.37*	*-0.54*	*0.67*
Commerce		-0.303	-0.329	-0.332	-0.332	-0.333	-0.353	-0.356	-0.199	-0.337
		-7.28	*-7.34*	*-7.26*	*-7.25*	*-7.17*	*-7.72*	*-7.78*	*-4.34*	*-7.09*
Tourism and restaurants		-0.120	-0.151	-0.152	-0.154	-0.119	-0.172	-0.107	-0.177	-0.152
		-0.87	*-1.03*	*-1.04*	*-1.05*	*-0.82*	*-1.17*	*-0.75*	*-1.21*	*-1.02*
Transports and communications		-0.011	-0.019	-0.023	-0.023	-0.017	-0.046	-0.027	0.052	-0.030
		-0.17	*-0.26*	*-0.32*	*-0.32*	*-0.24*	*-0.64*	*-0.38*	*0.73*	*-0.41*
Real estate		-0.334	-0.496	-0.502	-0.499	-0.498	-0.574	-0.535	-0.585	-0.505
		-2.36	*-3.28*	*-3.32*	*-3.29*	*-3.32*	*-3.80*	*-3.60*	*-3.91*	*-3.36*
Education		0.308	0.194	0.190	0.193	0.227	0.232	0.189	0.166	0.190
		1.10	*0.65*	*0.64*	*0.65*	*0.76*	*0.78*	*0.63*	*0.55*	*0.64*
Healthcare		-0.124	-0.286	-0.287	-0.284	-0.257	-0.268	-0.266	-0.253	-0.277
		-0.42	*-0.91*	*-0.92*	*-0.91*	*-0.82*	*-0.86*	*-0.86*	*-0.81*	*-0.88*
1997			-0.303	-0.303	-0.302	-0.313	-0.291	-0.312	-0.284	-0.313
			-5.61	*-5.59*	*-5.56*	*-5.76*	*-5.38*	*-5.76*	*-5.25*	*-5.76*
1998			-0.229	-0.230	-0.228	-0.235	-0.220	-0.236	-0.206	-0.235
			-4.55	*-4.55*	*-4.50*	*-4.65*	*-4.36*	*-4.68*	*-4.09*	*-4.65*
1999			-0.340	-0.341	-0.339	-0.342	-0.330	-0.343	-0.329	-0.342
			-6.38	*-6.37*	*-6.34*	*-6.39*	*-6.18*	*-6.44*	*-6.15*	*-6.39*
2000			-0.390	-0.390	-0.390	-0.393	-0.389	-0.391	-0.391	-0.393
			-6.51	*-6.51*	*-6.51*	*-6.57*	*-6.50*	*-6.56*	*-6.50*	*-6.56*
2001										
2002			0.006	0.006	0.005	0.002	-0.001	0.011	-0.013	0.002
			0.12	*0.12*	*0.11*	*0.04*	*-0.03*	*0.21*	*-0.26*	*0.05*
Constant	-2.377	-2.296	-2.184	-2.153	-2.175	-2.245	-2.336	-2.048	-1.907	-2.304
	-36.82	*-35.88*	*-29.42*	*-20.17*	*-19.27*	*-21.55*	*-21.80*	*-11.85*	*-18.61*	*-21.12*
Number of observations	71058	71058	71058	71058	71058	71078	71406	71078	71406	71078
Number of firms	24668	24668	24668	24668	24668	24589	24731	24589	24731	24589
Log-likelihood	-5574.7	-5531.8	-5484.1	-5483.7	-5483.5	-5468.2	-5503.7	-5481.5	-5404.0	-5471.1
Log-likelihood of the constant only model, for this sample	-5746.11	-5746.11	-5746.11	-5746.11	-5746.11	-5721.08	-5763.56	-5719.51	-5763.56	-5763.56
Pseudo-R2	0.030	0.037	0.046	0.046	0.046	0.044	0.045	0.042	0.062	0.051
Observations per group										
min	1	1	1	1	1	1	1	1	1	1
average	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9
max	6	6	6	6	6	6	6	6	6	6
Wald Chi2	286.9	333.2	346.3	347.0	346.7	348.7	356.6	338.1	412.5	345.8
Prob > Chi2	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
rho	0.341	0.336	0.397	0.396	0.396	0.397	0.398	0.392	0.389	0.399
Prob >= chibar2	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*

Note: z-scores in italics. All models estimated using a random-effects probit estimator, where the dependent variable is the dummy credit overdue. All ratios used are described in Appendix B. The pseudo-R2 is a measure of goodness of the fit, being computed as function of the model's log-likelihood and of the log-likelihood of the constant-only model, for the sub-sample used in each estimation. The Wald test evaluates the overall statistical significance of the estimated coefficients. Finally, rho measures the proportion of the total variance resulting from the panel-level variance component.

controlling for the firm's financial situation, the effect of firm size on default probabilities does not seem to remain significant.

Departing from this latter model, we tried several other possibly interesting variables. For instance, in Section 4.1, we had concluded that firms with default were, on average, older than the remaining firms in the sample. However, firm age does not seem to be statistically significant under a regression analysis framework (model 5). We also tried to take into account some productivity measures, such as capital productivity, measured as the ratio of sales to tangible assets (model 6). Though significant, its marginal contribution to explain loan default is rather small. Nevertheless, it helps to confirm that more productive firms should have, on average, lower default probabilities (though the productivity measure used can be highly sensitive to the sector in which the firm operates). Given the correlation between this indicator and sales growth, the latter ceases to be significant in this estimation. We also tried to consider whether capital intensity could help predict default (model 7). Even though the associated coefficient is very small, there seems to be evidence that firms more intensive on capital than on labour should display slightly higher default probabilities. Another variable considered was the share of tangible assets on firms' total non-financial fixed assets. This variable displays a negative coefficient, implying that the higher the share of tangible assets, the lower is the default probability, after controlling for the firm's economic sector. Nevertheless, the estimated coefficient for this variable is hardly statistically significant. Additionally, firms with higher turnover ratios (defined as sales to assets) are, as expected, firms with lower default risk (nevertheless, this variable is, to some extent, correlated with sales growth, which ceases to be significant when the turnover ratio is introduced in the regression). Given that the database does not provide information on the collateral used to guarantee loans, we tried to build an approximate measure of total available collateral (tangible assets as a percentage of total assets), but it did not prove to be significant in the estimated regression models.

Though most of the variables discussed above have some explanatory power in predicting loan default, we should focus our analysis on a limited set of variables, which comprehensively cover the more important dimensions of the firm's situation. Model 4 seems to provide a reasonable compromise between these two aspects, taking into account the firm's profitability, its sales evolution, its financial structure, its recent investment policy and its liquidity position, after controlling for size and economic sector, as well as for time-effects. Hence, this model will be considered as our baseline specification and all further extensions will be built upon it.

In the lower part of Table 5 some additional information on the estimations performed is displayed. Both the log-likelihood and the pseudo-R2 do not change significantly across the different specifications presented, suggesting that most of these variables have similar contributions in predicting loan default[22]. According to the Wald test reported, coefficients are overall significant for the models considered. We also report ρ, which provides a measure of the proportion of the total variance resulting from the panel-level variance component (when ρ is zero,

[22]The pseudo-R2 is a measure of goodness of the fit, being computed as $\frac{-\pi_0-(-\pi)}{-\pi}$, where π_0 is the log-likelihood of the constant-only model, for the sample used in the estimation, and π is the log-likelihood of the estimated regression. This ratio is a measure of the percentage of the variance on the dependent variable that is captured by the model.

the panel-level variance component is irrelevant and hence the panel estimator should be equal to the pooled estimator).

As mentioned above, the estimations presented in Table 5 were obtained using a random-effects probit. For robustness purposes, we also estimated some of these models using alternative estimation procedures. We first used a population-averaged estimator instead of the random-effects estimator, obtaining minor differences in the estimated coefficients, though without any qualitative changes[23]. We have also estimated the same population-averaged model using robust variance estimates, yielding minor changes in some z-scores (it is not possible to compute robust variances with a random-effects estimator). In addition, these models were also estimated using a logit instead of a probit model. The estimated coefficients are, as expected, slightly less than the double of those estimated with probits. Qualitatively, there are no significant changes when using random-effects and population-averaged estimators within a logit model. With logit models it also becomes possible to use fixed-effects estimators, though the results are relatively disappointing, given that only those firms that eventually default contribute to the likelihood function (as a result, the model focuses only on 771 firms and both the magnitude and the significance of the estimated coefficients change considerably). Finally, we also tried to estimate the baseline model without imposing a panel data structure. More specifically, we estimated three additional simple probit regressions: two using clustered standard errors (one of them clustering by firms and other clustering by years) and one without any clustering procedure (which would imply admitting that all observations are independent both across time and within each firm). The results are broadly similar, with one single exception: sales growth is not statistically significant when observations are clustered only by firm and when the clustering procedure is ignored.

The different model specifications outlined in Table 5 help to identify some of the firm-specific determinants of loan default. However, it should also be of interest to evaluate how the firm's past performance affects its default probability, which could help predicting future defaults. Moreover, given that firm-specific data is usually available with a considerable lag, it becomes crucial to try to assess whether a firm is likely to become stressed in the future by evaluating its current financial situation. Departing from the baseline specification presented above, in Table 6 we present some additional regressions, using all firm-specific variables lagged by one, two, three and four years, respectively. When all firm variables are lagged by one and two years, the results are mainly robust. Most of the coefficients on firm characteristics preserve the same signs. The most notable exception is the investment rate, which ceases to be significant when lagged. Moreover, the estimated coefficient for sales growth is not statistically significant when more than two lags are considered, suggesting that only the most recent sales performance truly conditions firms' default probabilities. There seems to be an increase in the marginal effect of profitability on credit risk, and, conversely, a decrease in the relative importance of the solvency ratio. Hence, sustained poor profitability ratios over time are a

[23] For a general model, the main difference between random-effects and population-averaged estimators is that the former fit the model $\mathrm{Prob}(Y_{it} = 1 \mid X_{it}, u_i) = F(X_{it}\beta + u_i)$, whereas population-averaged estimators fit the model $\mathrm{Prob}(Y_{it} = 1 \mid X_{it}) = G(X_{it}\beta^*)$. The subtle difference is that β and β^* are different population parameters: while the former takes into account the same firm for different values of the regressors, the latter focuses on average firm values (implying that $E(Y_{it} \mid X_{it}) = E(Y_{it} \mid X_i), \forall t$). For further details on population-averaged models (also known as generalised estimating equations (GEE) approach) please see Wooldridge (2002).

Table 6 - Probit regressions

		Baseline specification	All firm variables lagged: 1 year	2 years	3 years	4 years	Models with several simultaneous lags	
Sales growth	***t***	-0.001	-0.001	0.000	0.001	0.001	-0.003	
		-2.20	*-2.60*	*0.23*	*1.23*	*0.99*	*-5.54*	
	t-1						-0.001	-0.001
							-2.84	*-2.59*
ROA	***t***	-0.004	-0.005	-0.006	-0.006	-0.003		
		-3.95	*-3.59*	*-3.09*	*-2.22*	*-1.32*		
	t-1							-0.005
								-3.58
	t-2						-0.003	
							-2.01	
Solvency ratio	***t***	-0.005	-0.003	-0.003	-0.002	-0.002	-0.007	
		-7.35	*-3.60*	*-3.21*	*-1.83*	*-1.68*	*-7.39*	
	t-1							-0.003
								-3.61
	t-2						0.003	
							3.22	
Investment rate	***t***	-0.005	0.000	0.002	0.000	0.000	-0.005	
		-4.99	*0.17*	*1.40*	*0.22*	*-0.10*	*-3.62*	
Liquidity ratio	***t***	-0.001	-0.002	-0.001	-0.002	-0.001	-0.002	
		-4.48	*-4.68*	*-3.25*	*-3.23*	*-2.39*	*-3.99*	
	t-1							-0.002
								-4.81
Constant		-2.153	-2.085	-2.130	-1.951	-1.756	-2.092	-2.083
		-20.17	*-17.31*	*-14.28*	*-10.88*	*-14.92*	*-16.90*	*-17.32*
Number of observations		71058	46608	30924	19831	12139	45335	46608
Number of firms		24668	17169	12135	8623	7346	16662	17169
Log-likelihood		-5483.7	-3732.2	-2557.2	-1802.0	-1323.2	-3598.3	-3732.2
Log-likelihood of the constant only model, for this sample		-5746.11	-3879.97	-2659.01	-1870.56	-1354.59	-3797.12	-3879.97
Pseudo-R2		0.046	0.038	0.038	0.037	0.023	0.052	0.038
Observations per group								
min		1	1	1	1	1	1	1
average		2.9	2.7	2.5	2.3	1.7	2.7	2.7
max		6	5	4	3	2	5	5
Wald Chi2		347.0	196.4	119.0	65.4	55.7	250.2	196.2
Prob > Chi2		*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
rho		0.396	0.357	0.347	0.244	0.000	0.362	0.358
Prob >= chibar2		*0.00*	*0.00*	*0.00*	*0.02*	*1.00*	*0.00*	*0.00*

Note: z-scores in italics. All regressions include the control dummies for size, sector and year presented in Table 5. All models estimated using a random-effects probit estimator, where the dependent variable is the dummy credit overdue. All ratios used are described in Appendix B. The pseudo-R2 is a measure of goodness of the fit, being computed as function of the model's log-likelihood and of the log-likelihood of the constant-only model, for the sub-sample used in each estimation. The Wald test evaluates the overall statistical significance of the estimated coefficients. Finally, rho measures the proportion of the total variance resulting from the panel-level variance component.

strong sign of firm distress, yielding possibly high future default probabilities. When variables are lagged by three and, most notably, by four years, there is a clear decrease in the model's quality (most variables are no longer significant and the pseudo-R2 decreases considerably), suggesting that the firm's recent performance is, as expected, much more relevant to explain loan default than its "historical" background.

In addition, we also tried to estimate similar models using simultaneously several time lags. First we lagged all variables up to four years, considering also the contemporaneous information, and then we gradually dropped those lags which proved not be significant. Then we tried a more restricted approach, considering only up to three year lags (and no contemporaneous information). The results are consistent with those previously described. In both cases, only one and two year lags turn out to be statistically significant, confirming that using more than three year lags gives the model much less accuracy. Profitability seems to have the higher lagged explanatory power, though the liquidity and solvency ratios also provide interesting information when lagged by one year (however, the solvency ratio shows a rather counter-intuitive positive coefficient at $t-2$). Again, the investment rate fails to be significant when lagged.

As discussed above, there may be substantial differences in the determinants of loan default for firms of different ages, different sizes or different economic sectors. To better understand such differences, we estimated separate regressions for separate groups of firms. Separating the firms into three different age groups (young, average and mature firms), it is possible to conclude that the estimated model fits better older firms (most notably those which are more than 15 years old) than younger ones. In fact, sales growth does not seem to be an important determinant of loan default for young and average firms. Furthermore, profitability, solvency and liquidity are not significant for younger firms. In fact, for these firms, only the investment rate remains significant. These results suggest that start-up firms have relatively different determinants of loan default, reflecting in part that their capital structure (and, to some extent, their profitability) should necessarily be different than that of older firms. Taking into account firms with different dimensions, the estimates performed suggest that the baseline model has a better fit for small and medium firms. Sales growth and liquidity do not seem to be important in explaining default probabilities of micro firms. In contrast, the only clearly significant explanatory variable for large firms is the liquidity ratio. Estimating the model separately for different economic sectors also yields some interesting results. The performance of the model for manufacturing firms is remarkably good. For commerce firms, some of the variables considered do not seem to be significant, namely sales growth, liquidity and, to a lesser extent, the profitability ratio. For some of the other sectors considered, the performance of the model is weaker, in part because there are fewer observations. Nonetheless, these results confirm that default probabilities are driven by different factors in different economic sectors. From all the variables considered, the solvency ratio seems to be the most robust variable in predicting default probabilities across different sectors, suggesting that the firm's capital structure is an important determinant of loan default in most economic activity sectors, whereas sales growth or profitability, for instance, are more important determinants of loan default in some sectors than in others.

In Section 3 we discussed some of the links between loan default and macroeconomic and financial developments, at an aggregate level. In this section we have considered several firm characteristics that may contribute to understand why some firms default. Now, finally, we will try to simultaneously assess the role played by macroeconomic factors, together with firms' specific characteristics, by adding a set of macroeconomic variables to our panel data regressions. We considered a relatively large set of variables, taking into account some of the conclusions drawn in Section 3. Some of the variables tested in the regressions were GDP (level and growth), the coincident economic activity indicator, exports, private consumption, gross fixed capital formation, employment, loan growth, an exchange rate index, 10-year bond yields, the yield curve slope, banks interest rates applied on loans to firms, and stock market prices variation. Some of these variables did not prove to be significant or displayed unexpected signs. The most insightful results are presented in Table 7. From all the variables considered, the most important seem to be the GDP growth rate (with a negative contemporaneous impact on default probabilities, in agreement with what was discussed previously), the coincident economic activity indicator (which also evaluates economic conditions), loan growth (which also displays a negative coefficient) and, finally, stock market prices variation (implying, as

Table 7 - Probit regressions with macroeconomic variables

	Baseline specification without time dummies	Baseline specification with time dummies	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	All firm and macro variables lagged: 1 year Model 5	1 year Model 6	2 years Model 5	2 years Model 6
Sales growth	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	0.000	0.000
	-2.67	*-2.20*	*-2.12*	*-2.14*	*-2.21*	*-2.33*	*-2.18*	*-2.21*	*-0.72*	*-2.65*	*-2.59*	*0.23*	*0.23*
ROA	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.005	-0.005	-0.006	-0.006
	-4.30	*-3.95*	*-3.93*	*-3.96*	*-3.90*	*-4.16*	*-3.94*	*-3.94*	*-1.63*	*-3.60*	*-3.58*	*-3.09*	*-3.09*
Solvency ratio	-0.004	-0.005	-0.005	-0.005	-0.005	-0.004	-0.005	-0.005	-0.004	-0.003	-0.003	-0.003	-0.003
	-7.06	*-7.35*	*-7.37*	*-7.35*	*-7.34*	*-7.23*	*-7.37*	*-7.32*	*-3.08*	*-3.57*	*-3.61*	*-3.21*	*-3.21*
Investment rate	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.006	0.000	0.000	0.002	0.002
	-5.35	*-4.99*	*-4.99*	*-4.99*	*-4.91*	*-5.25*	*-4.97*	*-4.95*	*-2.45*	*0.19*	*0.17*	*1.40*	*1.40*
Liquidity ratio	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.003	-0.002	-0.002	-0.001	-0.001
	-4.52	*-4.48*	*-4.46*	*-4.47*	*-4.50*	*-4.44*	*-4.49*	*-4.49*	*-4.28*	*-4.71*	*-4.68*	*-3.25*	*-3.25*
Interest rate on loans to firms								0.026			0.111		0.117
								2.26			*4.10*		*1.66*
Yield curve slope (10 y - 3 m)							-0.159			0.043		-0.884	
							-3.43			*0.25*		*-2.85*	
Loan growth					-0.023		-0.019			-0.026	0.043	-0.129	
					-8.34		*-6.02*			*-1.45*	*3.39*	*-2.74*	
Stock market price variation						-0.002	-0.002			-0.001		-0.029	
						-4.86	*-3.48*			*-0.41*		*-4.27*	
GDP growth rate			-0.087						-0.141				
			-7.54						*-6.47*				
Coincident indicator BP				-0.061				-0.075			-0.325		-0.284
				-7.14				*-7.07*			*-5.96*		*-3.11*
Sales growth * GDP growth rate									0.000				
									-0.16				
ROA * GDP growth rate									0.000				
									-0.16				
Solvency ratio * GDP growth rate									0.000				
									-0.35				
Investment rate * GDP growth rate									0.000				
									0.26				
Liquidity ratio * GDP growth rate									0.001				
									2.81				
1997		-0.303											
		-5.59											
1998		-0.230											
		-4.55											
1999		-0.341											
		-6.37											
2000		-0.390											
		-(6.51)											
2001													
2002		0.006											
		0.12											
Constant	-2.241	-2.153	-2.093	-2.192	-1.872	-2.274	-1.755	-2.321	-1.935	-1.660	-2.983	2.664	-5.821
	-23.26	*-20.17*	*-20.38*	*-21.40*	*-17.64*	*-22.45*	*-14.57*	*-19.71*	*-16.78*	*-7.36*	*-9.15*	*1.90*	*-3.82*
Number of observations	71058	71058	71058	71058	71058	71058	71058	71058	71058	46608	46608	30924	30924
Number of firms	24668	24668	24668	24668	24668	24668	24668	24668	24668	17169	17169	12135	12135
Log-likelihood	-5531.2	-5483.7	-5500.3	-5503.9	-5494.1	-5518.6	-5487.0	-5501.4	-5495.5	-3754.0	-3732.2	-2557.2	-2557.2
Log-likelihood of the constant only model, for this sample	-5746.11	-5746.11	-5746.11	-5746.11	-5746.11	-5746.11	-5746.11	-5746.11	-5746.11	-3879.97	-3879.97	-2659.01	-2659.01
Pseudo-R2	0.037	0.046	0.043	0.042	0.044	0.040	0.045	0.043	0.044	0.032	0.038	0.038	0.038
Likelihood ratio test: model vs baseline without time dummies	-	95.0	61.8	54.7	74.3	25.2	88.4	59.7	-	-	-	-	-
Observations per group													
min	1	1	1	1	1	1	1	1	1	1	1	1	1
average	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.9	2.7	2.7	2.5	2.5
max	6	6	6	6	6	6	6	6	6	5	5	4	4
Wald Chi2	333.8	347.0	330.3	327.3	345.7	323.3	344.3	338.3	336.2	181.6	196.4	119.0	119.0
Prob > Chi2	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
rho	0.336	0.396	0.393	0.392	0.384	0.371	0.395	0.383	0.395	0.331	0.357	0.347	0.347
Prob >= chibar2	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*

Note: z-scores in italics. All regressions include the control dummies for size and sector presented in Table 5. All models estimated using a random-effects probit estimator, where the dependent variable is the dummy credit overdue. All ratios used are described in Appendix B. The pseudo-R2 is a measure of goodness of the fit, being computed as function of the model's log-likelihood and of the log-likelihood of the constant-only model, for the sub-sample used in each estimation. The Wald test evaluates the overall statistical significance of the estimated coefficients. Rho measures the proportion of the total variance resulting from the panel-

previously discussed, that positive developments in stock market prices, which usually reflect an improvement in firms' financial conditions, are associated with lower default probabilities). All these variables display relatively high marginal effects on default probabilities.

These first regressions were estimated by taking into account each macroeconomic variable separately, in order to minimise the losses in terms of information provided by firm heterogeneity. However, we also tried to take into account the joint effect of different macroeconomic and financial variables (models 5 and 6 in Table 7). In model 5 we considered loan growth, stock market prices variation and the slope of the yield curve. The slope of the yield curve, which may reflect, to some extent, expectations on future economic growth, has a strong negative marginal effect on default probabilities. When the variables considered in model 5 are lagged by one year (together with the firm-specific variables), the results are relatively disappointing, given that none of them remains statistically significant. Surprisingly, when we consider two year lags the results improve significantly (the marginal effect of the yield curve slope increases, confirming the forward-looking properties of this variable). Given the poor performance of this model when one year lags are taken into account, we estimated a different model (model 6), now considering interest rates on bank loans (which show, as expected, a positive contemporaneous coefficient), the coincident economic activity indicator and loan growth (which is automatically dropped when variables are used contemporaneously, given their high correlation). This model yields much better results when lagged by one year, but when two-year lags are used only the coincident indicator remains significant. It is interesting to notice that, contrary to what was suggested when we focused on aggregate time series in Section 3, the coefficient of economic growth (here proxied by the economic activity indicator) does not become positive when lagged by two years. Hence, even if at an aggregate level it seems to be clear that significant imbalances are created in periods of strong economic growth, after controlling for firm-specific characteristics this relationship is no longer apparent, possibly reflecting an asymmetric behaviour of firms with different characteristics during different phases of the credit cycle.

Given that firms' financial ratios are also subject to sizeable fluctuations over the business cycle, we tried to explicitly model these co-movements by adding to the model interactions between firm-specific variables and the GDP growth rate (model 7). The only significant interaction variable is the one associated with the liquidity indicator, suggesting that these interactions do not play a crucial role in explaining default probabilities. The GDP growth rate remains significant, but the coefficient associated with sales growth and ROA cease to be statistically significant in this model. However, when these two variables are excluded from the regression, their respective interaction with the GDP growth rate turns out significant.

As mentioned above, all the macroeconomic variables display relatively high marginal effects on default probabilities. To accurately assess the importance of macroeconomic conditions on default probabilities, we should begin by comparing the model without any time controls to the model with time dummies and to the models with macroeconomic variables. One important thing to notice is that the estimated coefficients and the z-scores for the firm specific variables almost do not change in all these specifications. This result suggests that the relevant information contained in macroeconomic variables is largely independent from that contained in firm specific variables. The incremental information provided by the inclusion of the time dimension can be confirmed by the significant increase in the pseudo-R2 of the model which includes time dummies, by comparison with a model without time controls. The substitution of these time dummies by specific macroeconomic variables (which can only capture part of the variation enclosed in time dummies) does not yield significant changes in the model's overall goodness of fit, suggesting that these macroeconomic variables can capture an important part of the time variation implicit in the year dummies. In order to more accurately test the role performed by the inclusion of time effects in the determination of default probabilities, we also performed likelihood ratio tests. The inclusion of year dummies allows for a significant increase in the likelihood of the model. When only one macroeconomic variable is considered (models 1 to 4), the change in the likelihood of the model is also very significant. Loan growth and GDP growth rate are the variables which have a higher impact on the model's likelihood. In fact, their explanatory power is not much lower than that of linear time controls. The additional explanatory power provided by the inclusion of three macroeconomic variables in model 5 is very similar to that of the time dummies. Hence, these results allow us to conclude that macroeconomic dynamics have an important additional (and independent) contribution in explaining why do firms default.

Finally, we performed some robustness checks, in order to test the validity of the results obtained. In the sample there are firms with multiple defaults and, as previously mentioned, only new transitions into the default state are being considered (if a firm is in default for more than two consecutive years, it is excluded from the regression as long as the default state persists). Taking into account the results obtained using conditional transition matrices, we have reasons to believe that firms with previous defaults may be riskier than other firms. To confirm this, we started by including in our sample all default observations (even if the firm is in default for more than two consecutive years). In this new sample, which includes more 391 firms, the results are generally robust, with the exception of profitability, which is no longer significant. Nevertheless, it is interesting to notice that the model's goodness of fit improves considerably. To better understand the differences in the behaviour of firms with past loan defaults, we estimated a separate regression only for firms which were in default at $t-1$. For the 1.236 firms considered in this regression, credit risk drivers seem to differ significantly from the ones considered in our base sample. In fact, the only firm-specific variables that remain statistically significant are the solvency ratio and the investment rate. We further extended this sample to include firms with at least one previous default during the sample period (even if not at $t-1$). This model performs slightly better (the pseudo-R2 increases considerably, as well as the proportion of the total variance captured by the panel level component). In addition to the solvency ratio and the investment rate, sales growth also becomes significant. Hence, firms with previous defaults which record relatively low solvency ratios, low investment rates and low sales growth should be much riskier than other firms.

Also for robustness purposes, we tested the impact of slightly changing the definition of the dependent variable, by considering that there was default only when credit overdue was above 100 euro, 1000 euro or 1 per cent of total debt, in order to focus only on more serious default problems. The estimation results remain broadly unchanged, though the model's explanatory power seems to improve slightly. The strikingly bimodal distribution of the credit overdue ratio which, as illustrated on Figure 2, displays either very high or very low values was also taken into account in the regressions, in order to test whether these different default events are driven by the same determinants. As discussed above, low credit overdue ratios should reflect mostly transitory episodes of deliquency, which may easily be reverted. When we only take into account default events in which this ratio is below 50 per cent of the firms' total bank debt, most of the variables considered retain their explanatory power. The only exception is sales growth, which is no longer significant. In turn, when only more serious default episodes are considered (credit overdue ratio above 50 per cent), both profitability and liquidity cease to be statistically significant.

Still for robustness purposes, we considered other modelling techniques, namely an ordered probit and a simple OLS with an alternative dependent variable. Concerning the ordered probit model, we defined different levels of default severity by constructing intervals for the ratio of credit overdue to total credit. The results are broadly consistent with those previously presented, showing only minor differences in the estimated coefficients. In addition to this, we considered an alternative model where the ratio of credit overdue to total credit was the dependent variable, instead of the binary dependent variable considered so far (this model was estimated within a simple panel data OLS framework). Again, the results are fairly robust, except in what concerns the liquidity ratio, which presents a counter-intuitive positive coefficient.

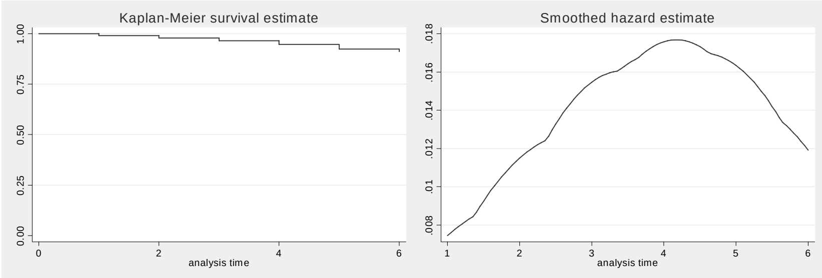
Recalling that we had controlled for outliers by setting observations above the 1st and 99th percentiles equal to the value of that percentile, we also tested the impact on the estimated regressions of running an alternative procedure for eliminating outliers, more specifically, by deleting the observations above or below those percentiles. The results are broadly consistent, but the change in the profitability coefficient, which becomes much stronger, should not be ignored. Finally, we also tested the introduction of some non-linearities in the model, by considering the squared value of some variables, as well as some interactions between variables. However, the marginal effect of the squared variables is almost negligible and does not seem to add much to the model. Moreover, the variable interactions tested were not statistically significant.

4.3.2 Results obtained using duration models

The way the data are organised and declared is very important in survival analysis models. Given that our database has strong left-censoring problems, this is a particularly important issue. In fact, most firms included in the sample were created before 1996, though in our database we do not have any information about their historical record, more specifically, we do not know whether those firms have defaulted before that year. This problem can be partly accounted for by declaring that firms are considered to be at risk since their creation date, though that failure risk can be observable only after the firm enters the sample (which may eventually be after 1996). In these models, our variable of interest will be the time until default, rather than a binary variable indicating whether the firm has defaulted or not. After organising the dataset according to these constraints, we are left with a sample of 32.966 firms, for which we have an average of 3.3 years of information. There are 1.921 observed defaults in this sample. The incidence rate, defined as the number of defaults divided by the total number of observations, is 1.8 per cent.

Given the left-censoring problems underlying our sample, we also tried to consider only those firms created from 1996 onwards, thus totally eliminating left-censoring. This implies focusing on a much smaller set of firms (3.284 firms, for which we observe only 94 defaults). The incidence rate for these firms is slightly lower, standing at 1.4 per cent. Figure 4 depicts several estimated functions for this subset of firms. The Kaplan-Meier survival estimate shows a steady decreasing trend, given that survival probabilities decrease over time. The most interesting results are those provided by the smoothed hazard estimate, suggesting that default probabilities are strongly increasing over time during the first 4 years of the firm's life. Afterwards, the hazard rate starts to decrease, resulting in a hump-shaped smoothed hazard estimate. These results shed some light on the previous discussions concerning the impact of

Figure 4
New firms



Note: analysis time in years.

firm age on default probabilities. In fact, it can be confirmed, to some extent, that default probabilities increase with firm age, though it is now clear that such increase is not linear through the firm's lifetime. Recalling from Section 4.2 that it can be said that there is positive duration dependence when $\frac{\delta h(t)}{\delta t} > 0$, $\forall t$ (as defined in equation 13), we cannot affirm that there is strictly positive duration dependence, given that for older firms we have $\frac{\delta h(t)}{\delta t} < 0$[24].

Within the framework of duration modelling, we estimated several regression models, in a spirit similar to that of discrete choice models. We started by fitting Cox proportional hazard models. The results obtained, presented in Table 8, are broadly similar to those obtained with probit models: firms with higher sales growth, higher profitability, higher solvency, higher investment rates, and better liquidity ratios display lower default probabilities (or, to be more precise, take a longer time to eventually default on their loan commitments). However, sales growth turns out to be clearly non-significant in the estimates performed when considering robust variance estimates. Hence, though sales growth may contribute to explain why some firms default, it does not seem to determine the time until default, at least under a Cox proportional hazard specification. In order to confirm the legitimate use of Cox models, we tested the proportional hazards assumption. Global and individual tests for the estimated regressions provide no evidence that the proportional hazards assumption is violated.

Given the strong left-censoring in the database, we also tested whether firms created from 1996 onwards were substantially different from others. In order to achieve that, we estimated a Cox model including a dummy variable for such firms (model 3 in Table 8). This dummy variable is far from being significant, suggesting that these firms do not substantially differ from the remaining firms in the sample. Nevertheless, to more deeply address this problem, we also estimated Cox regressions for this sub-sample, which are also displayed in Table 8. Both the solvency ratio and the investment rate cease to be significant. As argued above, these results suggest that start-up firms have relatively different determinants of loan default (in our sample, these firms show higher investment rates, as would be expected, as well as higher leverage ratios[25]). In order to complete our assessment, we tested the inclusion of other micro

[24] Estimating hazard rates for the full sample comprises significant problems, given the abovementioned left-censoring issue. Nevertheless, the estimates performed for the full sample also result in a hump-shaped hazard function. Default probabilities are clearly increasing during the first 25 years of the firm's life. Afterwards, default probabilities continue to increase, though at a less marked rate. Finally, for considerably older firms (more than 75 years), the hazard rate starts to decrease.

[25] Though higher leverage ratios are usually associated with higher default probabilities, as discussed above, in the first years of the firm's life a high level of indebtedness may be required to fund its initial investments,

Table 8 - Cox regressions (hazard ratios), robust

	Full sample			New firms								
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Sales growth	0.998			1.003								
	-1.72			*1.54*								
ROA	0.995	0.994	0.994	0.992	0.992	0.992	0.993	0.993	0.992	0.992	0.993	0.992
	-4.33	*-4.83*	*-4.84*	*-2.44*	*-2.31*	*-2.31*	*-2.02*	*-1.79*	*-2.36*	*-2.37*	*-2.17*	*-2.49*
Solvency ratio	0.995	0.995	0.995	1.003	1.003		1.003	1.000	1.005	1.003	1.002	1.003
	-4.59	*-4.56*	*-4.53*	*0.74*	*0.78*		*0.74*	*-0.06*	*1.29*	*0.70*	*0.59*	*0.77*
Investment rate	0.990	0.989	0.989	0.993	0.994	0.994	0.994	0.993	0.994	0.994	0.994	0.994
	-3.94	*-4.10*	*-4.12*	*-1.23*	*-1.02*	*-1.02*	*-1.00*	*-1.47*	*-1.04*	*-1.04*	*-1.08*	*-1.02*
Liquidity ratio	0.995	0.995	0.995	0.990	0.990	0.990	0.990	0.993	0.986	0.990	0.990	0.990
	-4.53	*-4.51*	*-4.54*	*-3.94*	*-4.04*	*-4.04*	*-3.89*	*-2.97*	*-5.01*	*-3.99*	*-3.97*	*-3.98*
Leverage						0.997						
						-0.78						
Share of tangible assets							0.994					
							-0.77					
Turnover ratio								0.996				
								-2.26				
Available collateral									0.994			
									-1.32			
Activity began after 1996 (Y/N)			0.962									
			-0.23									
GDP growth rate										1.030		
										0.24		
Loan growth											0.991	
											-0.39	
Stock market price variation												1.005
												0.98
Constant	-	-	-	-	-	-	-	-	-	-	-	-
Log pseudo likelihood	-7291.3	-7294.0	-7294.3	-434.1	-435.3	-435.3	-428.4	-429.9	-428.2	-437.2	-437.2	-436.8
No. of observations	76292	76292	76292	3847	3847	3847	3802	3847	3802	3847	3847	3847
No. of subjects	25690	25690	25690	2324	2324	2324	2297	2324	2297	2324	2324	2324
No. of failures	1000	1000	1000	68	68	68	67	68	67	68	68	68
Time at risk	76292	76292	76292	3847	3847	3847	3802	3847	3802	3847	3847	3847
Wald chi2	583.9	581.4	577.9	35.7	34.2	34.2	35.2	39.9	44.8	31.5	31.0	31.3
Prob > chi2	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*

Note: z-scores in italics. New firms are defined as those created from 1996 onwards. The regressions for the full sample include the control dummies for size, sector and year presented in Table 5. All models estimated using a Cox regression which evaluates the time until default, using robust variance estimates. All ratios used are described in Appendix B. An estimated coefficient lower than 1 should be interpreted as contributing a longer time until default eventually occurs. The Wald test evaluates the overall significance of the estimated coefficients.

and macro variables. Most of the variables tested do not seem to be statistically significant in the determination of the time until default of these start-up firms. The only relevant exception seems to be the turnover ratio. According to the results obtained with these regressions, firms with lower turnover ratios should default sooner than other firms. None of the macroeconomic variables tested is significant.

Finally, in order to complete our analysis, we estimated parametric duration models, using several different distribution functions (namely, exponential, Weibull, Gompertz, lognormal, and log-logistic). The results for one of the estimated models for the sub-sample of start-up firms are displayed in Table 9. The estimated coefficients are broadly robust across the different distribution functions considered and do not differ substantially from those obtained using a Cox proportional hazard model. It should be noted that some of the estimated coefficients are displayed as proportional hazard ratios (PH), whereas others are presented as accelerated failure-time coefficients (AFT). The latter present signs opposite to those obtained with the Cox models because they have a different interpretation. Accelerated failure-time models change the time scale by a factor of $exp(-X_i\beta)$, in a general model. A positive coefficient implies an acceleration of time, which is the same as an increase in the expected waiting time until

without implying necessarily a higher default probability. Nevertheless, for an older firm, a highly leveraged financial structure, when combined with a deterioration in other financial ratios, may signal increased credit risk, as illustrated in the regressions presented for the full sample.

Table 9 - Parametric survival models for new firms, robust

	Exponential		Weibull		Gompertz	Lognormal	Log-logistic	Cox model
	PH	AFT	PH	AFT	PH	AFT	AFT	
Sales growth	-	-	-	-	-	-	-	-
ROA	0.993	0.007	0.989	0.003	0.989	0.004	0.003	0.992
	-2.25	*2.25*	*-2.94*	*2.36*	*-2.96*	*2.34*	*2.34*	*-2.31*
Solvency ratio	1.002	-0.002	1.005	-0.001	1.005	-0.001	-0.001	1.003
	0.71	*-0.71*	*1.22*	*-1.30*	*1.22*	*-0.58*	*-1.25*	*0.78*
Investment rate	0.994	0.006	0.996	0.001	0.996	0.000	0.001	0.994
	-1.04	*1.04*	*-0.72*	*0.78*	*-0.71*	*-0.15*	*0.72*	*-1.02*
Liquidity ratio	0.990	0.010	0.990	0.002	0.989	0.003	0.002	0.990
	-4.06	*4.06*	*-4.04*	*2.93*	*-4.07*	*3.23*	*2.94*	*-4.04*
Constant	-	3.151	-	2.093	-	2.443	2.079	-
		9.59		*13.42*		*11.36*	*13.63*	
Log-likelihood	-261.0	-261.0	-237.6	-237.6	-245.1	-241.3	-237.8	-435.3
No. of observations	3847	3847	3847	3847	3847	3847	3847	3847
No. of subjects	2324	2324	2324	2324	2324	2324	2324	2324
No. of failures	68	68	68	68	68	68	68	68
Time at risk	3847	3847	3847	3847	3847	3847	3847	3847
LR chi2	33.3	33.3	44.2	259.9	42.9	146.8	248.1	34.2
Prob > chi2	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*	*0.00*
AIC	542.06	542.06	497.20	497.20	512.15	504.68	497.58	

Note: z-scores in italics. New firms are defined as those created from 1996 onwards. All regressions include year control dummies. The models presented in this table were estimated parametrically, using the exponential, Weibull, Gompertz, log-normal and log-logistic distributions, using robust variance estimates. PH stands for proportional hazard ratios. In this case, an estimated coefficient lower than 1 should be interpreted as contributing to lower default probabilities or, more precisely, to a longer time until default eventually occurs. In turn, AFT stands for accelerated failure-time coefficients. A positive coefficient implies an acceleration of time, which is the same as an increase in the expected waiting time until default. All ratios used are described in Appendix B. The LR/Wald test evaluates the overall significance of the estimated coefficients. AIC stands for Akaike Information Criteria.

default. The Akaike information criteria (AIC) suggests that the Weibull and the log-logistic distributions are the ones which provide more accurate results[26].

Finally, an additional effort conducted to overcome the left-censoring problem was to gather information from the Central Credit Register on loan defaults observed between 1980 and 1995 for the firms included in the sample. As a result, 226 new defaults were taken into account. Using this new information, we still declare that firms are at risk since their creation date, though now we can observe their failure since 1980. Hence, if a firm defaulted between 1980 and 1995, it will now be excluded from the regressions, given that it failed before entering our observation window. Using this additional information allows to fully overcome the left-censoring problem, given that we can argue that a default that occurred before 1980 will hardly condition the firm's default probability from 1996 onwards. The results using this default history are broadly consistent with those obtained when the full default history was not taken into account. Hence, though we have concluded that firms with previous defaults are more likely to default again in the future, the inclusion of a longer default history does not seem to seriously affect regression results.

5 Concluding remarks

This work focused on the determinants of credit risk, both at an aggregate and at a firm-specific level. On one hand, we tried to understand how systematic factors, which simultaneously affect all firms, condition the evolution of aggregate default rates. On the other hand, we examined how firms' specific characteristics affect their default probabilities.

[26] The Akaike information criteria is computed as $AIC = -2\pi + 2 \times (c + p + 1)$, where π is the estimated log-likelihood of the model, c is the number of explanatory variables and p is the number of ancillary parameters (p is an output of the estimation).

We started by exploring the links between credit risk and macroeconomic developments at an aggregate level. The results obtained suggest that there are some important links between credit risk and macroeconomic developments. In fact, these results seem to confirm the hypothesis that in periods of economic growth, which are sometimes accompanied by strong credit growth, there may be some tendency towards excessive risk-taking. However, the imbalances created in such periods only become apparent when economic growth slows down.

After examining the determinants of credit risk at an aggregate level, we focused our attention on an extensive dataset with detailed financial information for more than 30.000 firms, which also includes their loan default record. The results obtained suggest that default probabilities are influenced by several firm-specific characteristics, such as their financial structure, profitability and liquidity, as well as by their recent sales performance or their investment policy. After controlling for the most relevant firm-characteristics, the firm's dimension does not seem to contribute to explain differences in default frequencies, though there are some important differences between economic sectors. Lagged information on the firm's financial situation over a short period also seems to be important in explaining why do some firms default on their loan commitments. Furthermore, the firm's default history should be taken into account in the assessment of its credit risk, given that firms which recorded loan defaults in the recent past seem to display much higher default probabilities than other firms.

Finally, when time-effect controls or macroeconomic variables are taken into account together with the firm-specific information, the results of the models seem to improve considerably. Hence, even though the determinants of loan default at the micro level are ultimately driven by the firms' specific financial situation, there are important relationships between overall macroeconomic conditions and default rates, which should be assessed from a financial stability perspective. Nevertheless, it is important to consider that ultimately the firms' financial situation is also exposed to systematic shocks in the economy, thereby affecting their individual default probabilities. Hence, some firms may be more vulnerable to those systematic shocks (as a consequence of the sector in which they operate or of their current financial situation), explaining the fluctuation of aggregate default rates over the business cycle.

Though this work envisaged a thorough empirical examination of credit risk drivers, both at the macro and micro level, many issues could still be addressed on further research. In what concerns econometric estimation procedures, a quantile regression framework could help to better explore some of the conclusions drawn when assessing differences between firms, though the application of this technique to panel datasets is still not straightforward. Another interesting research topic would be to explicitly model and estimate the interaction between macroeconomic developments and firm-specific characteristics. Finally, the extension of the sample period used for assessing credit risk drivers at the firm-level could provide interesting results, by taking into account at least one full business cycle.

References

Allen, L. and Saunders, A. (2003), "A survey of cyclical effects in credit risk measurement models", BIS Working Paper No.126.

Altman, E. (1968), "Financial ratios, discriminant analysis and the prediction of corporate bankruptcy", *The Journal of Finance*, Vol.23, No.4, 589-609.

Antunes, A. (2005), "Analysis of delinquent firms using multi-state transitions", Banco de Portugal Working Paper 5-05.

Baltagi, B. (1995), *Econometric analysis of panel data*, John Wiley & Sons.

Banasik, J., Crook, J.N. and Thomas, L.C (1999), "Not if but when will borrowers default", *The Journal of Operational Research Society*, Vol.50, No.12, 1185-1190.

Bangia, A., Diebold, F., Kronimus, A., Schagen, C. and Schuermann, T. (2002), "Ratings migration and the business cycle, with an application to credit portfolio stress testing", *Journal of Banking and Finance*, No.26, 445-474.

Basel Committee on Banking Supervision (2004), "International convergence of capital measurement and capital standards: a revised framework", BIS report, June 2004.

Benito, A., Javier Delgado, F. and Martínez Pagés, J. (2004), "A synthetic indicator of financial pressure for Spanish firms", Banco de España Working Paper No.411.

Berger, A. and Udell, G. (2003), "The institutional memory hypothesis and the procyclicality of bank lending behaviour", BIS Working Paper No.125.

Bernhardsen, E. (2001), "A model of bankruptcy prediction", Norges Bank Working Paper 2001/10.

Bhattacharjee, A., Higson, C., Holly, S. and Kattuman, P. (2002), "Macro economic instability and business exit: determinants of failures and acquisitions of large UK firms", Cambridge Working Papers in Economics 0206.

Borio, C., Furfin, C. and Lowe, P. (2001), "Procyclicality of the financial system and financial stability: issues and policy options", BIS Papers No.1, 1-57.

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

Campbell, J. Y., Hilscher, J. and Szilagyi, J. (2005), "In search of distress risk", Deutsche Bundesbank Discussion Paper No. 27/2005, Series 1: Economic Studies.

Carling, K., Jacobson, T., Lindé, J. and Roszbach, K. (2002), "Capital charges under Basel II: corporate credit risk modelling and the macroeconomy", Sveriges Riksbank Working Paper No.142.

Carling, K., Jacobson, T., Lindé, J. and Roszbach, K. (2004), "Corporate credit risk modelling and the macroeconomy", *mimeo* (forthcoming in *Journal of Banking and Finance*).

Couderc, F. and Renault, O. (2005), "Times-to-default: life cycle, global and industry cycle impacts", FAME Research Paper No.142.

Crouhy, M., Galai, D. and Mark, R. (2000), "A comparative analysis of current credit risk models", *Journal of Banking and Finance*, No.24, 59-117.

Duffie, D. and Singleton, K. (2003), *Credit risk: pricing, measurement and management*, Princeton University Press.

Eklund, T., Larsen, K. and Bernhardsen, E. (2001), "Model for analysing credit risk in the enterprise sector", Norges Bank Economic Bulletin Q3 01.

Figlewski, S., Frydman, H. and Liang, W. (2006), "Modeling the effect of macroeconomic factors on corporate default and credit rating transitions", *mimeo.*

Gersbach, H. and Lipponer, A. (2003), "Firm defaults and the correlation effect", *European Financial Management*, Vol.9, 361-378.

Gordy, M. (2000), "A comparative anatomy of credit risk models", *Journal of Banking and Finance*, No.24, 119-149.

Hamerle, A., Liebig, T. and Scheule, H. (2004), "Forecasting credit portfolio risk", Deutsche Bundesbank Discussion Paper No. 01/2004, Series 2: Banking and Financial Supervision.

Hsiao, C. (1986), *Analysis of panel data*, Econometric Society Monographs, Cambridge University Press.

Jarrow, R. and Turnbull, S. (1995), "Pricing derivatives on financial securities subject to credit risk", *The Journal of Finance*, Vol.50, No.1, 53-85.

Jiménez, G. and Saurina, J. (2004), "Collateral, type of lender and relationship banking as determinants of credit risk", *Journal of Banking and Finance*, No.28, 2191-2212.

Jiménez, G. and Saurina, J. (2006), "Credit cycles, credit risk and prudential regulation", *International Journal of Central Banking*, June, 65-98.

Kent, C. and D'Arcy, P. (2001), "Cyclical prudence - credit cycles in Australia", BIS Papers No.1, 58-90.

Koopman, S. J. and Lucas, A. (2005), "Business and default cycles for credit risk", *Journal of Applied Econometrics*, No.20, 311-323.

Koopman, S. J., Lucas, A. and Daniels, R. J. (2005), "A non-Gaussian panel time series model for estimating and decomposing default risk", De Nederlandsche Bank Working Paper No.55.

Korajczyk, R.A. and Levy, A. (2003), "Capital structure choice: macroeconomic conditions and financial constraints", *Journal of Financial Economics*, 68, 75-109.

Lancaster, T. (1990), *The econometric analysis of transition data*, Econometric Society Monographs, Cambridge University Press.

Lis, S., Pages, J. and Saurina, J. (2000), "Credit growth, problem loans and credit risk provisioning in Spain", Banco de España Working Paper No.18.

Lowe, P. (2002), "Credit risk measurement and procyclicality", BIS Working Paper No.116.

Maddala, G. and Rao, C. (1996), *Handbook of statistics 14: Statistical methods in finance*, North-Holland.

Merton, R. (1974), "On the pricing of corporate debt: the risk structure of interest rates", *The Journal of Finance*, Vol.29, No.2, 449-470.

Moody's (2004), *The Moody's KMV EDF Riskcalc v3.1 model*, Moody's KMV Company.

Nickell, P., Perraudin, W. and Varotto, S. (2000), "Stability of rating transitions", *Journal of Banking and Finance*, No.24, 203-227.

Pain, D. and Vesala, J. (2004), "Driving factors of credit risk in Europe", *mimeo*.

Pederzoli, C. and Torricelli, C. (2005), "Capital requirements and business cycle regimes: forward-looking modelling of default probabilities", *Journal of Banking and Finance*, No.29, 3121-3140.

Rosch, D. (2003), "Correlations and business cycles of credit risk: evidence from bankruptcies in Germany", *Financial Markets and Portfolio Management*, Vol.17, No.3, 309-331.

Roszbach, K. (2004), "Bank lending policy, credit scoring and the survival of loans", *The Review of Economics and Statistics*, Vol.86, Issue 4, 946-958.

Saretto, A. (2004), "Predicting and pricing the probability of default", *mimeo*.

Saunders, A. and Allen, L. (2002), *Credit risk measurement: new approaches to value at risk and other paradigms*, John Wiley & Sons.

Shumway, T. (2001), "Forecasting bankruptcy more accurately: a simple hazard model", *The Journal of Business*, Vol.74, No.1, 101-124.

Tudela, M. and Young, G. (2003), "Predicting default among UK companies: a Merton approach", Bank of England Financial Stability Review, June 2003.

Wilson, T. (1998), "Portfolio credit risk", Federal Reserve Board of New York Economic Policy Review, October 1998.

Wooldridge, J. M. (2002), *Econometric analysis of cross section and panel data*, MIT Press.

Appendix A: Macroeconomic and financial data

Macroeconomic and financial data

Series	Source	Available since
Loans and non-performing loans	Banco de Portugal (Monetary and Financial Statistics)	1979Q4
Credit overdue and provisions	Banco de Portugal	1990Q1
National accounts and labour market data (quarterly)	Banco de Portugal [1]	1977Q1
Consumer price index (CPI)	INE	1977Q1
Monthly coincident indicator for economic activity	Banco de Portugal [2]	1978Q1
Bank interest rates	Banco de Portugal [3]	1990Q1
Stock market data	Euronext	1990Q1
Government bond yields [4]	Reuters	1990Q1

Notes: 1) Castro, G. and Esteves, P. (2004), "Quarterly series for the Portuguese economy: 1977-2003", Banco de Portugal Economic Bulletin, June 2004.

2) Rua, A. (2004), "A new coincident indicator for the Portuguese economy", Banco de Portugal Economic Bulletin, June 2004.

3) Monetary and financial statistics and "New series on banks' interest rates: long series for the average rates on outstanding amounts", Banco de Portugal Economic Bulletin, December 2003.

4) Government bond yields for Germany available only since 1991Q1 (for 5 years bonds) and 1992Q1 (for 10 years bonds).

Appendix B: Microeconomic variables

Microeconomic variables

Name	Description
Profitability	
ROA	Net income / Total assets x 100
ROE	Net income / Equity x 100
Operational results as % of equity	Operational results / Equity x 100
Financial structure and leverage	
Solvency ratio	Equity / Total assets x 100
Total credit as a % assets	Total loans (credit register) / Total assets x 100
Total credit as a % equity	Total loans (credit register) / Equity x 100
Leverage	(Total assets - Equity) / Total assets x 100
Dummy "has long-term credit"	This variable assumes the value 1 whenever some of the firm's credit liabilities are recorded in the Credit Register as medium or long term
Productivity	
Labour productivity	Sales / Number of employees x 100
Capital productivity	Sales / Tangible assets x 100
K_L coefficient	Tangible assets / Number of employees (euro per person)
Investment	
Fixed assets	Tangible assets + Intangible assets + Financial assets
Investment rate	Annual variation in net fixed assets / Sales x 100
Share of tangible assets	Tangible assets as percentage of total tangible and intangible assets
Liquidity	
Liquidity ratio	(Bank deposits and cash + Debt receivables + Inventories + Short-term investments) / Debt payables x 100
Liquid assets to total assets	(Bank deposits and cash + Debt receivables + Inventories + Short-term investments) / Total assets x 100
Other	
Sales growth	Year-on-year growth rate of total sales
Credit overdue	Includes principal not paid for more than 30 days after its due date, as well as the respective interest and costs due. The dummy credit overdue takes the value 1 whenever there is a positive amount of credit overdue recorded in the Central Credit Register at the end of the year.
Available collateral (proxy)	Tangible assets / Total assets x 100
Turnover ratio	Sales / Total assets x 100
Firm size	Defined accordingly with the European Commision Recommendation of 6 May 2003 (2003/361/EC), by taking into account the number of employees and sales volume.

WORKING PAPERS

2000

1/00 UNEMPLOYMENT DURATION: COMPETING AND DEFECTIVE RISKS
— *John T. Addison, Pedro Portugal*

2/00 THE ESTIMATION OF RISK PREMIUM IMPLICIT IN OIL PRICES
— *Jorge Barros Luís*

3/00 EVALUATING CORE INFLATION INDICATORS
— *Carlos Robalo Marques, Pedro Duarte Neves, Luís Morais Sarmento*

4/00 LABOR MARKETS AND KALEIDOSCOPIC COMPARATIVE ADVANTAGE
— *Daniel A. Traça*

5/00 WHY SHOULD CENTRAL BANKS AVOID THE USE OF THE UNDERLYING INFLATION INDICATOR?
— *Carlos Robalo Marques, Pedro Duarte Neves, Afonso Gonçalves da Silva*

6/00 USING THE ASYMMETRIC TRIMMED MEAN AS A CORE INFLATION INDICATOR
— *Carlos Robalo Marques, João Machado Mota*

2001

1/01 THE SURVIVAL OF NEW DOMESTIC AND FOREIGN OWNED FIRMS
— *José Mata, Pedro Portugal*

2/01 GAPS AND TRIANGLES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

3/01 A NEW REPRESENTATION FOR THE FOREIGN CURRENCY RISK PREMIUM
— *Bernardino Adão, Fátima Silva*

4/01 ENTRY MISTAKES WITH STRATEGIC PRICING
— *Bernardino Adão*

5/01 FINANCING IN THE EUROSYSTEM: FIXED VERSUS VARIABLE RATE TENDERS
— *Margarida Catalão-Lopes*

6/01 AGGREGATION, PERSISTENCE AND VOLATILITY IN A MACROMODEL
— *Karim Abadir, Gabriel Talmain*

7/01 SOME FACTS ABOUT THE CYCLICAL CONVERGENCE IN THE EURO ZONE
— *Frederico Belo*

8/01 TENURE, BUSINESS CYCLE AND THE WAGE-SETTING PROCESS
— *Leandro Arozamena, Mário Centeno*

9/01 USING THE FIRST PRINCIPAL COMPONENT AS A CORE INFLATION INDICATOR
— *José Ferreira Machado, Carlos Robalo Marques, Pedro Duarte Neves, Afonso Gonçalves da Silva*

10/01 IDENTIFICATION WITH AVERAGED DATA AND IMPLICATIONS FOR HEDONIC REGRESSION STUDIES
— *José A.F. Machado, João M.C. Santos Silva*

2002

1/02 QUANTILE REGRESSION ANALYSIS OF TRANSITION DATA
— *José A.F. Machado, Pedro Portugal*

2/02 SHOULD WE DISTINGUISH BETWEEN STATIC AND DYNAMIC LONG RUN EQUILIBRIUM IN ERROR CORRECTION MODELS?
— *Susana Botas, Carlos Robalo Marques*

3/02 MODELLING TAYLOR RULE UNCERTAINTY
— *Fernando Martins, José A. F. Machado, Paulo Soares Esteves*

4/02 PATTERNS OF ENTRY, POST-ENTRY GROWTH AND SURVIVAL: A COMPARISON BETWEEN DOMESTIC AND FOREIGN OWNED FIRMS
— *José Mata, Pedro Portugal*

5/02 BUSINESS CYCLES: CYCLICAL COMOVEMENT WITHIN THE EUROPEAN UNION IN THE PERIOD 1960-1999. A FREQUENCY DOMAIN APPROACH
— *João Valle e Azevedo*

6/02 AN "ART", NOT A "SCIENCE"? CENTRAL BANK MANAGEMENT IN PORTUGAL UNDER THE GOLD STANDARD, 1854 -1891
— *Jaime Reis*

7/02 MERGE OR CONCENTRATE? SOME INSIGHTS FOR ANTITRUST POLICY
— *Margarida Catalão-Lopes*

8/02 DISENTANGLING THE MINIMUM WAGE PUZZLE: ANALYSIS OF WORKER ACCESSIONS AND SEPARATIONS FROM A LONGITUDINAL MATCHED EMPLOYER-EMPLOYEE DATA SET
— *Pedro Portugal, Ana Rute Cardoso*

9/02 THE MATCH QUALITY GAINS FROM UNEMPLOYMENT INSURANCE
— *Mário Centeno*

10/02 HEDONIC PRICES INDEXES FOR NEW PASSENGER CARS IN PORTUGAL (1997-2001)
— *Hugo J. Reis, J.M.C. Santos Silva*

11/02 THE ANALYSIS OF SEASONAL RETURN ANOMALIES IN THE PORTUGUESE STOCK MARKET
— *Miguel Balbina, Nuno C. Martins*

12/02 DOES MONEY GRANGER CAUSE INFLATION IN THE EURO AREA?
— *Carlos Robalo Marques, Joaquim Pina*

13/02 INSTITUTIONS AND ECONOMIC DEVELOPMENT: HOW STRONG IS THE RELATION?
— *Tiago V.de V. Cavalcanti, Álvaro A. Novo*

2003

1/03 FOUNDING CONDITIONS AND THE SURVIVAL OF NEW FIRMS
— *P.A. Geroski, José Mata, Pedro Portugal*

2/03 THE TIMING AND PROBABILITY OF FDI: AN APPLICATION TO THE UNITED STATES MULTINATIONAL ENTERPRISES
— *José Brandão de Brito, Felipa de Mello Sampayo*

3/03 OPTIMAL FISCAL AND MONETARY POLICY: EQUIVALENCE RESULTS
— *Isabel Correia, Juan Pablo Nicolini, Pedro Teles*

4/03 FORECASTING EURO AREA AGGREGATES WITH BAYESIAN VAR AND VECM MODELS
— *Ricardo Mourinho Félix, Luís C. Nunes*

5/03 CONTAGIOUS CURRENCY CRISES: A SPATIAL PROBIT APPROACH
— *Álvaro Novo*

6/03 THE DISTRIBUTION OF LIQUIDITY IN A MONETARY UNION WITH DIFFERENT PORTFOLIO RIGIDITIES
— *Nuno Alves*

7/03 COINCIDENT AND LEADING INDICATORS FOR THE EURO AREA: A FREQUENCY BAND APPROACH
— *António Rua, Luís C. Nunes*

8/03 WHY DO FIRMS USE FIXED-TERM CONTRACTS?
— *José Varejão, Pedro Portugal*

9/03 NONLINEARITIES OVER THE BUSINESS CYCLE: AN APPLICATION OF THE SMOOTH TRANSITION AUTOREGRESSIVE MODEL TO CHARACTERIZE GDP DYNAMICS FOR THE EURO-AREA AND PORTUGAL
— *Francisco Craveiro Dias*

10/03 WAGES AND THE RISK OF DISPLACEMENT
— *Anabela Carneiro, Pedro Portugal*

11/03 SIX WAYS TO LEAVE UNEMPLOYMENT
— *Pedro Portugal, John T. Addison*

12/03 EMPLOYMENT DYNAMICS AND THE STRUCTURE OF LABOR ADJUSTMENT COSTS
— *José Varejão, Pedro Portugal*

13/03 THE MONETARY TRANSMISSION MECHANISM: IS IT RELEVANT FOR POLICY?
— *Bernardino Adão, Isabel Correia, Pedro Teles*

14/03 THE IMPACT OF INTEREST-RATE SUBSIDIES ON LONG-TERM HOUSEHOLD DEBT: EVIDENCE FROM A LARGE PROGRAM
— *Nuno C. Martins, Ernesto Villanueva*

15/03 THE CAREERS OF TOP MANAGERS AND FIRM OPENNESS: INTERNAL VERSUS EXTERNAL LABOUR MARKETS
— *Francisco Lima, Mário Centeno*

16/03 TRACKING GROWTH AND THE BUSINESS CYCLE: A STOCHASTIC COMMON CYCLE MODEL FOR THE EURO AREA
— *João Valle e Azevedo, Siem Jan Koopman, António Rua*

17/03 CORRUPTION, CREDIT MARKET IMPERFECTIONS, AND ECONOMIC DEVELOPMENT
— *António R. Antunes, Tiago V. Cavalcanti*

18/03 BARGAINED WAGES, WAGE DRIFT AND THE DESIGN OF THE WAGE SETTING SYSTEM
— *Ana Rute Cardoso, Pedro Portugal*

19/03 UNCERTAINTY AND RISK ANALYSIS OF MACROECONOMIC FORECASTS: FAN CHARTS REVISITED
— *Álvaro Novo, Maximiano Pinheiro*

2004

1/04 HOW DOES THE UNEMPLOYMENT INSURANCE SYSTEM SHAPE THE TIME PROFILE OF JOBLESS DURATION?
— *John T. Addison, Pedro Portugal*

2005

1/05 SEAM: A SMALL-SCALE EURO AREA MODEL WITH FORWARD-LOOKING ELEMENTS
— *José Brandão de Brito, Rita Duarte*

2/05 FORECASTING INFLATION THROUGH A BOTTOM-UP APPROACH: THE PORTUGUESE CASE
— *Cláudia Duarte, António Rua*

3/05 USING MEAN REVERSION AS A MEASURE OF PERSISTENCE
— *Daniel Dias, Carlos Robalo Marques*

4/05 HOUSEHOLD WEALTH IN PORTUGAL: 1980-2004
— *Fátima Cardoso, Vanda Geraldes da Cunha*

5/05 ANALYSIS OF DELINQUENT FIRMS USING MULTI-STATE TRANSITIONS
— *António Antunes*

6/05 PRICE SETTING IN THE AREA: SOME STYLIZED FACTS FROM INDIVIDUAL CONSUMER PRICE DATA
— *Emmanuel Dhyne, Luis J. Álvarez, Hervé Le Bihan, Giovanni Veronese, Daniel Dias, Johannes Hoffmann, Nicole Jonker, Patrick Lünnemann, Fabio Rumler, Jouko Vilmunen*

7/05 INTERMEDIATION COSTS, INVESTOR PROTECTION AND ECONOMIC DEVELOPMENT
— *António Antunes, Tiago Cavalcanti, Anne Villamil*

8/05 TIME OR STATE DEPENDENT PRICE SETTING RULES? EVIDENCE FROM PORTUGUESE MICRO DATA
— *Daniel Dias, Carlos Robalo Marques, João Santos Silva*

9/05 BUSINESS CYCLE AT A SECTORAL LEVEL: THE PORTUGUESE CASE
— *Hugo Reis*

10/05 THE PRICING BEHAVIOUR OF FIRMS IN THE EURO AREA: NEW SURVEY EVIDENCE
— *S. Fabiani, M. Druant, I. Hernando, C. Kwapil, B. Landau, C. Loupias, F. Martins, T. Mathä, R. Sabbatini, H. Stahl, A. Stokman*

11/05 CONSUMPTION TAXES AND REDISTRIBUTION
— *Isabel Correia*

12/05 UNIQUE EQUILIBRIUM WITH SINGLE MONETARY INSTRUMENT RULES
— *Bernardino Adão, Isabel Correia, Pedro Teles*

13/05 A MACROECONOMIC STRUCTURAL MODEL FOR THE PORTUGUESE ECONOMY
— *Ricardo Mourinho Félix*

14/05 THE EFFECTS OF A GOVERNMENT EXPENDITURES SHOCK
— *Bernardino Adão, José Brandão de Brito*

15/05 MARKET INTEGRATION IN THE GOLDEN PERIPHERY – THE LISBON/LONDON EXCHANGE, 1854-1891
— *Rui Pedro Esteves, Jaime Reis, Fabiano Ferramosca*

2006

1/06 THE EFFECTS OF A TECHNOLOGY SHOCK IN THE EURO AREA
— *Nuno Alves , José Brandão de Brito , Sandra Gomes, João Sousa*

2/02 THE TRANSMISSION OF MONETARY AND TECHNOLOGY SHOCKS IN THE EURO AREA
— *Nuno Alves, José Brandão de Brito, Sandra Gomes, João Sousa*

3/06 MEASURING THE IMPORTANCE OF THE UNIFORM NONSYNCHRONIZATION HYPOTHESIS
— *Daniel Dias, Carlos Robalo Marques, João Santos Silva*

4/06 THE PRICE SETTING BEHAVIOUR OF PORTUGUESE FIRMS EVIDENCE FROM SURVEY DATA
— *Fernando Martins*

5/06 STICKY PRICES IN THE EURO AREA: A SUMMARY OF NEW MICRO EVIDENCE
— *L. J. Álvarez, E. Dhyne, M. Hoeberichts, C. Kwapil, H. Le Bihan, P. Lünnemann, F. Martins, R. Sabbatini, H. Stahl, P. Vermeulen and J. Vilmunen*

6/06 NOMINAL DEBT AS A BURDEN ON MONETARY POLICY
— *Javier Díaz-Giménez, Giorgia Giovannetti , Ramon Marimon, Pedro Teles*

7/06 A DISAGGREGATED FRAMEWORK FOR THE ANALYSIS OF STRUCTURAL DEVELOPMENTS IN PUBLIC FINANCES
— *Jana Kremer, Cláudia Rodrigues Braz, Teunis Brosens, Geert Langenus, Sandro Momigliano, Mikko Spolander*

8/06 IDENTIFYING ASSET PRICE BOOMS AND BUSTS WITH QUANTILE REGRESSIONS
— *José A. F. Machado, João Sousa*

9/06 EXCESS BURDEN AND THE COST OF INEFFICIENCY IN PUBLIC SERVICES PROVISION
— *António Afonso, Vítor Gaspar*

10/06 MARKET POWER, DISMISSAL THREAT AND RENT SHARING: THE ROLE OF INSIDER AND OUTSIDER FORCES IN WAGE BARGAINING
— *Anabela Carneiro, Pedro Portugal*

11/06 MEASURING EXPORT COMPETITIVENESS: REVISITING THE EFFECTIVE EXCHANGE RATE WEIGHTS FOR THE EURO AREA COUNTRIES
— *Paulo Soares Esteves, Carolina Reis*

12/06 THE IMPACT OF UNEMPLOYMENT INSURANCE GENEROSITY ON MATCH QUALITY DISTRIBUTION
— *Mário Centeno, Alvaro A. Novo*

13/06 U.S. UNEMPLOYMENT DURATION: HAS LONG BECOME LONGER OR SHORT BECOME SHORTER?
— *José A.F. Machado, Pedro Portugal e Juliana Guimarães*

14/06 EARNINGS LOSSES OF DISPLACED WORKERS: EVIDENCE FROM A MATCHED EMPLOYER-EMPLOYEE DATA SET
— *Anabela Carneiro, Pedro Portugal*

15/06 COMPUTING GENERAL EQUILIBRIUM MODELS WITH OCCUPATIONAL CHOICE AND FINANCIAL FRICTIONS
— *António Antunes, Tiago Cavalcanti, Anne Villamil*

16/06 ON THE RELEVANCE OF EXCHANGE RATE REGIMES FOR STABILIZATION POLICY
— *Bernardino Adao, Isabel Correia, Pedro Teles*

17/06 AN INPUT-OUTPUT ANALYSIS: LINKAGES VS LEAKAGES
— *Hugo Reis, António Rua*

2007

1/07 RELATIVE EXPORT STRUCTURES AND VERTICAL SPECIALIZATION: A SIMPLE CROSS-COUNTRY INDEX
— *João Amador, Sónia Cabral, José Ramos Maria*

2/07 THE FORWARD PREMIUM OF EURO INTEREST RATES
— *Sónia Costa, Ana Beatriz Galvão*

3/07 ADJUSTING TO THE EURO
— *Gabriel Fagan, Vítor Gaspar*

4/07 SPATIAL AND TEMPORAL AGGREGATION IN THE ESTIMATION OF LABOR DEMAND FUNCTIONS
— *José Varejão, Pedro Portugal*

5/07 PRICE SETTING IN THE EURO AREA: SOME STYLISED FACTS FROM INDIVIDUAL PRODUCER PRICE DATA
— *Philip Vermeulen, Daniel Dias, Maarten Dossche, Erwan Gautier, Ignacio Hernando, Roberto Sabbatini, Harald Stahl*

6/07 A STOCHASTIC FRONTIER ANALYSIS OF SECONDARY EDUCATION OUTPUT IN PORTUGAL
— *Manuel Coutinho Pereira, Sara Moreira*

7/07 CREDIT RISK DRIVERS: EVALUATING THE CONTRIBUTION OF FIRM LEVEL INFORMATION AND OF MACROECONOMIC DYNAMICS
— *Diana Bonfim*