

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

Diana Bonfim**

1. INTRODUCTION

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. Moreover, from a financial stability viewpoint, it is interesting to understand if credit default risk is mostly driven by idiosyncratic factors (which may allow for the identification of a set of characteristics common to riskier firms) or by systematic factors, which simultaneously affect all firms (thus having more widespread impacts on financial stability, given that several banks can suffer sizeable losses in their credit portfolios at the same time). 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. The empirical results obtained in the literature suggest that there are some important links between credit risk and macroeconomic developments (see, for instance, Pederzoli and Torricelli (2005), Jiménez and Saurina (2006) or *Bonfim* (2007)). In fact, periods of strong economic growth, which are sometimes accompanied by robust credit growth, are sometimes followed, with some lag, by an increase in default rates at the aggregate level, possibly as a consequence of imbalances generated in those periods.

Under this setup, the main purpose of this article is to empirically examine the determinants of corporate credit default in loans granted by the Portuguese banking system, taking simultaneously into account firm-specific data as well as macroeconomic information. Using micro information, which includes very detailed firm-level data, it becomes clear that the firms' financial situation will influence the probability of default on their loan commitments. Profitability, liquidity, solvency and recent investment and sales performance seem to offer a valuable contribution in explaining default probabilities. 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, the results obtained with this work suggest that 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 probabilities over time.

The remainder of the paper proceeds as follows. *Section 2* describes the main datasets used in this work, which include information for more than 30.000 Portuguese firms. In *Section 3* we present some summary statistics which are relevant for the characterisation of the sample. Afterwards, in *Section 4* we briefly present the modelling setup underlying the empirical work which will be developed in

* The analysis, opinions and findings in this article represent the views of the author and are not necessarily those of the Banco de Portugal. The author is grateful to Paula Antão, Paulo Soares Pinho, António Antunes, Nuno Ribeiro, Pedro Portugal, Mário Centeno, Ana Cristina Leal and Nuno Alves for their helpful comments and suggestions. Any remaining errors and omissions are the author's responsibility.

** Economics and Research Department, Banco de Portugal.

Sections 5 and 6, which present, respectively, the results obtained using discrete choice models and duration models. Finally, in Section 7 we present some concluding remarks.

2. DATA

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 renegotiated, or if it is an off-balance sheet risk, such as the unused parts of credit lines or bank guarantees).¹ Using end-of-year data for the period between 1996 and 2002, there are 203.655 observations regarding loans granted to non-financial corporations.² In turn, the Central Balance Sheet Database contains detailed accounting data on a large sample of Portuguese firms, being used mostly for statistical and economic analysis purposes.³ Reporting by firms in this database is not compulsory. The sample of firms covers to an acceptable degree the Portuguese universe, although some bias may exist towards larger firms, which are almost totally covered.⁴ In this dataset there are 153.581 annual observations for the period between 1996 and 2002. By merging the two datasets there are 113.119 observations, which regard 33.084 different firms.

We constructed several ratios and indicators to evaluate firms' financial situation, namely in what concerns their profitability, financial structure, leverage, productivity, liquidity and investment. In order to guarantee the quality of results, several filters were applied to the data. First, ratios for which the denominator equalled (or was close to) zero, as well as ratios for which a negative numerator was combined with a negative denominator (leading to ratios with a positive sign), were eliminated from the analysis, in order to avoid misleading results. Moreover, in order to prevent outliers from distorting the analysis, observations below the 1st percentile and above the 99th percentile were replaced with the value of the corresponding percentile.

3. CHARACTERISATION OF FIRMS IN DEFAULT

Only a small percentage of firms in the sample has credit overdue (around 3 per cent of firms). The higher default rates are recorded in fishing, mining, tourism and restaurants and manufacturing (Table 1). With regard to the size of firms, the highest default frequencies are recorded by medium-sized firms, closely followed by larger 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

(1) The main purposes and features of the Central Credit Register of Banco de Portugal are described in http://www.bportugal.pt/internet/prod/publish/cadernos/responsabilidades_credito_e.pdf.

(2) Reporting credit institutions aggregate information on loans with similar status for each firm (i.e. information is not reported on a loan-by-loan basis). In order to merge the two datasets used in this article, the CRC records were aggregated within firms. As a consequence, each observation is defined as a pair firm-year, summing up all credit liabilities for a given firm in each year.

(3) The main purposes and features of the Central Balance Sheet Database of Banco de Portugal are described in http://www.bportugal.pt/internet/prod/publish/cadernos/central_balancos_e.pdf.

(4) In 2005 firms included in the Central Balance Sheet Database accounted for around 61 per cent of the gross value added of the Portuguese economy and for around 35 per cent of the total size of the workforce in Portugal. However, while for large enterprises coverage stood at around 82 per cent of the size of the workforce, for small and medium-sized enterprises this coverage stood at around 20 per cent. For more details on the coverage of this database, see *Table G.1.2* of the Statistical Bulletin of Banco de Portugal.

Table 1

DEFAULT FREQUENCIES BY FIRM SECTOR AND SIZE		
By sector of activity		
	No of obs.	Default rate (%)
Fishing	277	11.19
Mining	1 084	5.17
Agriculture	3 487	2.81
Manufacturing	41 427	3.76
Utilities	355	1.97
Construction	14 020	3.25
Commerce	31 793	1.83
Tourism and restaurants	1 405	4.70
Transportation and communications	6 004	2.60
Real estate activities	2 319	2.29
Education	249	3.21
Healthcare	331	1.21
By size		
	No of obs.	Default rate (%)
Micro	39 725	2.67
Small	42 608	2.72
Medium	16 548	4.21
Large	4 320	3.89

Note: The size of enterprises was defined according to Commission Recommendation of 6 May 2003 (2003/631/EC), taking into account the number of employees and sales and services turnover.

(2004) find that smaller firms are more likely to default.⁵ In turn, Pain and Vesala (2004) and Bernhardsen (2001) conclude that any systemic effect of firm size on default is relatively small, particularly after controlling for firm's characteristics.⁶ Such results imply that two firms with very similar financial indicators are not likely to present significantly different default probabilities, even if they have considerably different sizes. Finally, Benito *et al.* (2004), using a sample of Spanish firms, 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, given that reporting to the Central Balance Sheet Database is not mandatory (and therefore the response rate by firms under strong financial pressure is expected to be relatively lower).⁷

Taking into account that one of the main purposes of this article is to understand what drives credit risk at the firm-level, it may be relevant to compare some features of firms with and without default. Table 2 presents some of the main results obtained, showing that, in fact, there are several aspects that distinguish firms in default. For instance, average profitability levels in these firms are much lower, sales growth is far lower, and the dependence on external funding sources (solvency ratio, defined as the ratio between equity and total assets) seems to be significantly higher. Moreover, firms with problems in

(5) Bhattacharjee *et al.* (2002) and Bunn and Redwood (2003) use samples of firms from the United Kingdom, Eklund *et al.* (2001) use a database of Norwegian firms and, finally, Jiménez and Saurina (2004) use Banco de España databases similar to those used in this article.

(6) The above-mentioned studies use a database comprising European firms and Norwegian firms respectively.

(7) Antunes *et al.* (2005) conclude that default rates tend to be lower for firms with more credit, on the basis of available information in Central Credit Register, which contributes to strengthen the hypothesis that small enterprises in financial distress should be sub-represented in the Central Balance Sheet Database.

Table 2

SUMMARY STATISTICS – COMPARISON OF FIRMS WITH AND WITHOUT DEFAULTS				
	Mean values for non-defaulting firms in t	Mean values for defaulting firms in t	Welch test - Ho: no difference	
			Ha: difference not zero Pr(T > t)	Significantly different mean (Y/N)
ROA	0.5	-4.9	0.00	Y
Sales growth	12.9	5.7	0.00	Y
Solvency ratio	23.2	1.1	0.00	Y
Total credit as a % of assets	12.5	16.9	0.00	Y
Leverage	76.8	98.9	0.00	Y
Investment rate	2.6	-2.5	0.00	Y
Liquidity ratio	119.0	86.5	0.00	Y
Number of observations	100 117	3 084		

Note: ROA (return on assets) defined as net income as a percentage of assets; Sales growth defined as the year-on-year rate of change in sales and services; solvency ratio assessed by the ratio of equity to assets; total credit refers to total debt recorded at Central Credit Register for each firm; leverage is defined as the ratio of the firm's liabilities to assets; the investment rate refers to the annual change in net fixed assets as a percentage of sales and services provided; finally, the liquidity ratio assesses the amount recorded under cash and deposits, third-party debt, inventories and marketable securities as a percentage of third-party debt.

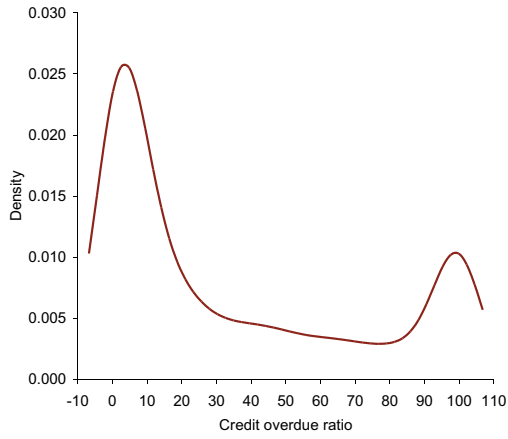
paying their debt also have, on average, lower investment rates, as well as less favourable liquidity indicators. The Welch test makes it possible to determine whether mean values in these two groups of firms are significantly different. The results of this statistical test confirm that all mean values for firms in default are statistically different from mean values for the remaining firms.

The empirical distribution of the default ratio of firms included in the sample is clearly two-peaked: either firms record only small amounts of credit overdue, which should reflect transitory episodes of delinquency, or they default on nearly all their debt, which should be a situation closer to bankruptcy (Chart 1). It may be interesting to notice that the former is particularly frequent for larger firms, whereas the latter mostly respects to smaller firms. In this sense, although larger firms have relatively higher default frequencies, default episodes recorded by these firms usually have a very small magnitude.

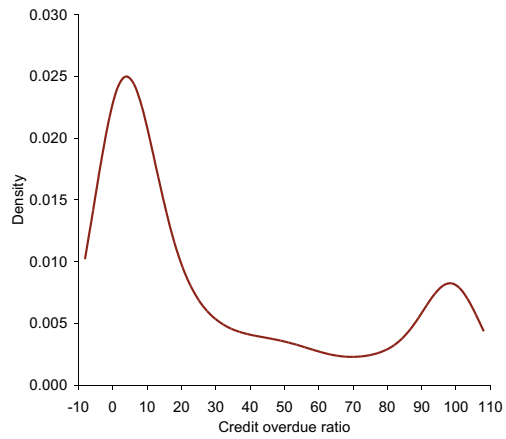
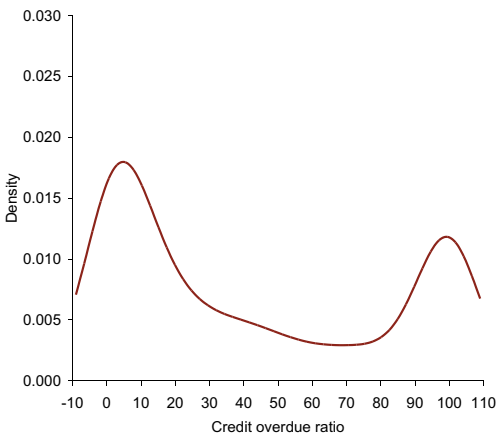
By estimating conditional transition matrices, it is possible to assess default probabilities for different time horizons ($t+1$, $t+2$, ...). These matrices suggest that for firms with no previous default episodes, default probabilities increase steadily over time. In turn, for firms that have recorded some default episode, default probabilities are always far higher than those of firms that have never had problems in paying their debt commitments. This implies that firms with a past record of credit overdue are more likely to default again in the future than firms that never defaulted before.

Chart 1

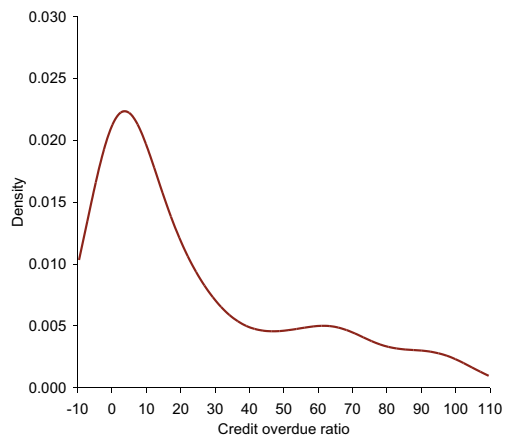
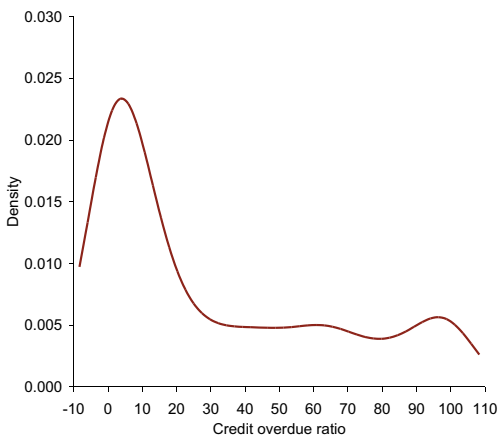
EMPIRICAL DISTRIBUTION OF THE CREDIT OVERDUE RATIO
(Gaussian kernel density)



Micro enterprises Small enterprises



Medium-sized enterprises Large enterprises



4. METHODOLOGY

A default episode of firm i in period t may be modelled as a random variable Y_{it} , such that:

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

The default probability would thus be defined as:

$$\Pr(Y_{it} = 1) = \Pr(R_{it} < c_{it})$$

where R_{it} represents the return on a firm's assets. Thus, a firm is likely to default if its returns fall below a given threshold c_{it} .

In this article two types of explanatory variables are considered. The first one is a set of firm-specific variables, which shall account for idiosyncratic risk (Z_{it}). The second vector comprises a set of aggregate time-varying regressors, common to all firms, which intend to account for the systematic risk component (X_t). If we consider that the firm's returns may be defined through a linear combination of these vectors, the following applies:

$$\Pr(Y_{it} = 1) = \Pr(R_{it} < c_{it}) = \Pr(X_t + Z_{it} + u_{it} < c_{it}) = F(-X_t - Z_{it} - u_{it} + c_{it})$$

where $F(\cdot)$ denotes the cumulative distribution function of the residual. This model's parameters may be estimated using a discrete choice model, namely a probit or a logit. The use of this type of econometric model is relatively common in the literature on credit risk, and therefore this will be the methodology initially used in this article for the econometric estimation of factors driving loan default.

However, it may be relevant to understand not only if a firm will default, but also when will that default eventually occur. The timing of loan default is important for perform a complete risk assessment, as well as for accurate loan pricing and provisioning. Duration models, which only recently started to be applied to credit risk models, directly model the survival time of a loan, taking as a dependent variable the time until default.⁸

Under the duration modelling framework, we define T as the time until a loan defaults⁹. The hazard function can be defined as the probability of a firm defaulting within the infinitely short interval $[t, t+dt)$, conditional on not having defaulted before:

$$h(t) = \lim_{dt \rightarrow 0} \frac{\Pr(t < T < t + dt | T > t)}{dt}$$

The duration distribution function can be defined as $F(t) = \Pr(T < t)$. The survival function is the probability of surviving up to t , and can be defined as:

$$S(t) = \Pr(T > t) = 1 - F(t) = \exp\left(-\int_0^t h(s) ds\right)$$

(8) For examples of studies applying this econometric estimation technique to credit risk models, see Banasik *et al.* (1999), Carling *et al.* (2007), Couderc and Renault (2005) or Shumway (2001).
 (9) Lancaster (1990) presents a detailed and rigorous description of several issues associated with duration models.

Whenever T has an exponential distribution, the hazard function $h(t)$ is constant. When that is not observed, the underlying process is said to exhibit duration dependence. If $\frac{h'(t)}{h(t)} > 0$, there is positive duration dependence, which implies that, in our framework, the probability of default increases over the life of the firm. Otherwise, duration dependence should be negative, implying that the longer the firm has remained without defaulting, the lower should be its default probability.

5. MAIN ECONOMETRIC RESULTS OBTAINED USING DISCRETE CHOICE MODELS

Table 3 presents some of the main results obtained using discrete choice models, based on a random-effects probit model.¹⁰ The choice of variables to be used was based on the estimation of correlation matrices between the set of available variables, as well as on the abovementioned mean comparison statistical tests. The first model includes a relatively diversified set of variables, which broadly characterises firms' financial situation. 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, given that higher profitability levels should reflect a solid financial situation of the firm and, as a consequence, imply 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 should have relatively low credit risk. Moreover, firms with high investment rates are also likely to have lower default probabilities. In fact, it seems reasonable to admit that firms under strong financial pressure are not expected to engage in investment projects. Finally, the liquidity indicator, defined as short-term assets as a percentage of the firm's total debt, also displays a negative coefficient, implying that firms facing stronger liquidity constraints may have higher difficulties in paying their debt commitments.

Even though the firm-specific variables taken into account seem to play an important role in predicting loan default for the firms included in this sample, 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. In order to control for specific effects associated with such firm characteristics, all the regressions presented in this article include economic sector and firm size dummies as control variables. Results suggest that there are indeed substantial differences in the determinants of default probabilities in various sectors (the estimation of separate regressions for each economic sector in the sample confirms this hypothesis). Estimated models confirm that fishing and mining firms will tend to have higher default probabilities, as shown in Table 1. With regard to the inclusion of control variables for firm size, the results are not so clear. Descriptive statistics for the entire sample suggested that smaller firms show slightly lower default rates than larger firms, which is not consistent with the results usually obtained in the literature on credit risk modelling.¹¹ Estimated regressions confirm this result, although differences between firms with different dimensions are not statistically significant. This result implies that, although larger firms display slightly higher default probabilities, after controlling for several firm characteristics, firm size does not have a statistically significant effect on default probabilities. As a consequence, two firms with similar financial characteristics should present identical default probabilities, even if they have very different sizes. Finally, given that macroeconomic variables will only be introduced further ahead, year dummies were also included as control variables, in order to control for any possible systematic effects. Most coefficients associated with these year

(10) Results presented in this section are a brief summary of those presented in *Bonfim* (2007).

(11) This result may be conditioned by the sample bias towards firms with better credit quality, which should be over-represented, as discussed in *Section 3*.

Table 3

PROBIT REGRESSIONS (DEPENDENT VARIABLE: DUMMY CREDIT OVERDUE)						
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Sales growth	-0.001	-0.001	-0.001	-0.001	-0.001	0.000
	<i>-2.19</i>	<i>-2.20</i>	<i>-2.16</i>	<i>-2.52</i>	<i>-1.97</i>	<i>-0.47</i>
ROA	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004
	<i>-3.96</i>	<i>-3.95</i>	<i>-3.92</i>	<i>-3.97</i>	<i>-3.66</i>	<i>-4.05</i>
Solvency ratio	-0.005	-0.005	-0.005	-0.006	-0.006	-0.007
	<i>-7.56</i>	<i>-7.35</i>	<i>-7.36</i>	<i>-11.24</i>	<i>-11.09</i>	<i>-11.87</i>
Investment rate	-0.005	-0.005	-0.005		-0.005	-0.005
	<i>-5.01</i>	<i>-4.99</i>	<i>-4.99</i>		<i>-4.52</i>	<i>-5.18</i>
Liquidity ratio	-0.001	-0.001	-0.001			
	<i>-4.47</i>	<i>-4.48</i>	<i>-4.51</i>			
Firm age			0.001			
			<i>0.63</i>			
Share of tangible fixed assets				-0.002		
				<i>-1.49</i>		
Available collateral (approx.)					0.001	
					<i>1.51</i>	
Turnover ratio						-0.003
						<i>-12.01</i>
Small		-0.044	-0.035	-0.035	-0.044	-0.034
		<i>-0.52</i>	<i>-0.41</i>	<i>-0.42</i>	<i>-0.53</i>	<i>-0.41</i>
Micro		-0.013	-0.001	-0.011	-0.025	-0.059
		<i>-0.15</i>	<i>-0.01</i>	<i>-0.13</i>	<i>-0.29</i>	<i>-0.69</i>
Medium-sized		-0.026	-0.022	-0.015	-0.023	-0.005
		<i>-0.30</i>	<i>-0.25</i>	<i>-0.17</i>	<i>-0.27</i>	<i>-0.06</i>
Fishing	0.358	0.363	0.360	0.431	0.369	0.234
	<i>1.42</i>	<i>1.45</i>	<i>1.43</i>	<i>1.74</i>	<i>1.46</i>	<i>0.93</i>
Mining	0.222	0.223	0.224	0.240	0.228	0.148
	<i>1.57</i>	<i>1.57</i>	<i>1.58</i>	<i>1.71</i>	<i>1.60</i>	<i>1.05</i>
Agriculture	-0.191	-0.195	-0.194	-0.182	-0.194	-0.306
	<i>-1.95</i>	<i>-1.98</i>	<i>-1.96</i>	<i>-1.85</i>	<i>-1.94</i>	<i>-3.07</i>
Utilities	-0.492	-0.500	-0.492	-0.446	-0.473	-0.622
	<i>-1.34</i>	<i>-1.36</i>	<i>-1.34</i>	<i>-1.26</i>	<i>-1.29</i>	<i>-1.70</i>
Construction	0.040	0.039	0.041	0.019	0.035	-0.027
	<i>0.81</i>	<i>0.78</i>	<i>0.82</i>	<i>0.37</i>	<i>0.67</i>	<i>-0.54</i>
Commerce	-0.329	-0.332	-0.332	-0.356	-0.337	-0.199
	<i>-7.34</i>	<i>-7.26</i>	<i>-7.25</i>	<i>-7.78</i>	<i>-7.09</i>	<i>-4.34</i>
Tourism and restaurants	-0.151	-0.152	-0.154	-0.107	-0.152	-0.177
	<i>-1.03</i>	<i>-1.04</i>	<i>-1.05</i>	<i>-0.75</i>	<i>-1.02</i>	<i>-1.21</i>
Transportation and communications	-0.019	-0.023	-0.023	-0.027	-0.030	0.052
	<i>-0.26</i>	<i>-0.32</i>	<i>-0.32</i>	<i>-0.38</i>	<i>-0.41</i>	<i>0.73</i>
Real estate activities	-0.496	-0.502	-0.499	-0.535	-0.505	-0.585
	<i>-3.28</i>	<i>-3.32</i>	<i>-3.29</i>	<i>-3.60</i>	<i>-3.36</i>	<i>-3.91</i>
Education	0.194	0.190	0.193	0.189	0.190	0.166
	<i>0.65</i>	<i>0.64</i>	<i>0.65</i>	<i>0.63</i>	<i>0.64</i>	<i>0.55</i>
Healthcare	-0.286	-0.287	-0.284	-0.266	-0.277	-0.253
	<i>-0.91</i>	<i>-0.92</i>	<i>-0.91</i>	<i>-0.86</i>	<i>-0.88</i>	<i>-0.81</i>
1997	-0.303	-0.303	-0.302	-0.312	-0.313	-0.284
	<i>-5.61</i>	<i>-5.59</i>	<i>-5.56</i>	<i>-5.76</i>	<i>-5.76</i>	<i>-5.25</i>
1998	-0.229	-0.230	-0.228	-0.236	-0.235	-0.206
	<i>-4.55</i>	<i>-4.55</i>	<i>-4.50</i>	<i>-4.68</i>	<i>-4.65</i>	<i>-4.09</i>
1999	-0.340	-0.341	-0.339	-0.343	-0.342	-0.329
	<i>-6.38</i>	<i>-6.37</i>	<i>-6.34</i>	<i>-6.44</i>	<i>-6.39</i>	<i>-6.15</i>
2000	-0.390	-0.390	-0.390	-0.391	-0.393	-0.391
	<i>-6.51</i>	<i>-6.51</i>	<i>-6.51</i>	<i>-6.56</i>	<i>-6.56</i>	<i>-6.50</i>
2002	0.006	0.006	0.005	0.011	0.002	-0.013
	<i>0.12</i>	<i>0.12</i>	<i>0.11</i>	<i>0.21</i>	<i>0.05</i>	<i>-0.26</i>
Constant	-2.184	-2.153	-2.175	-2.048	-2.304	-1.907
	<i>-29.42</i>	<i>-20.17</i>	<i>-19.27</i>	<i>-11.85</i>	<i>-21.12</i>	<i>-18.61</i>
Number of observations	71 058	71 058	71 058	71 078	71 078	71 406
Number of firms	24 668	24 668	24 668	24 589	24 589	24 731
Pseudo-R²	0.046	0.046	0.046	0.042	0.051	0.062
Wald Chi²	346.3	347.0	346.7	338.1	345.8	412.5
<i>Prob > Chi²</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>
rho	0.397	0.396	0.396	0.392	0.399	0.389
<i>Prob >= chibar²</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>

Note: z-scores in italics. All models estimated using a random-effects probit estimator, where the dependent variable is the dummy credit overdue. Coefficients refer to marginal effects. In what concerns size, sector and year dummies, the omitted variables were large enterprises, manufacturing firms and 2001 respectively. The pseudo-R² is a measure of the goodness of the fit. The Wald test evaluates the overall statistical significance of the estimated coefficients. Finally, rho assesses the proportion of the total variance resulting from the panel-level variance component.

dummies are significant, which gives support to the hypothesis that macroeconomic developments should also be important in explaining loan default.

In addition to the variables used in the first model (sales growth, return on assets, solvency ratio, investment rate and liquidity indicator), other possible specifications were considered, which are also shown in Table 3. Firm age is not a statistically significant variable for determining 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. 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. Turnover ratios, defined as sales over assets, also seem to have a relatively significant explanatory power and show a negative sign. This variable confirms the evidence presented by sales growth (which is not significant in this specification, given the strong correlations between the two variables), suggesting that the current buoyancy of firm activity is an important signal of its financial soundness.

The different model specifications outlined in Table 3 help to identify some of the firm-specific determinants of loan default in a given moment of time. However, it should also be of interest to evaluate how the firm's past performance affects its current default probability. Such information may help to improve the ability to forecast default probabilities, given that it helps to better understand to what extent the deterioration of some financial indicators may imply an increase in the firm's credit risk in the near future. Moreover, understanding such dynamic relationships is also important due to the time lag usually associated with the release of accounting data, particularly for non-quoted firms, which may hamper the monitoring of credit risk developments over time. In this sense, Table 4 shows the baseline model defined in Table 3 with explanatory 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. Obtaining this result is very important, given that it implies that variables identified as particularly relevant for determining the current risk of the firm also make it possible to detect in advance possible financial difficulties in a horizon of up to two years. 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 strong sign of firm distress, yielding possibly high future default probabilities. When variables are lagged by three or four years there is a clear decrease in the model's quality (most variables are no longer significant and the model's goodness of fit, assessed by the pseudo-R², decreases considerably),¹² suggesting that the firm's recent performance is, as expected, much more relevant to explain loan default than its historical background over a longer horizon.

In addition, we also considered simultaneously several time lags, in order to capture in a more integrated manner the dynamic effect of the firm's financial situation on credit risk. Overall, the results are consistent with those previously described, as in both cases only one and two year lags turn out to be statistically significant. The results confirm that profitability seems to have the highest lagged explana-

(12) The pseudo-R² is a measure of the goodness of the fit, being computed as $\frac{l_0 - l}{l_0}$, where l_0 is the log-likelihood of the constant-only model $Y_t = \alpha_0 + \epsilon_t$, and l 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.

Table 4

PROBIT REGRESSIONS								
		Baseline specification	All firm variables lagged by:				Models with several simultaneous lags	
			1 year	2 years	3 years	4 years		
Sales growth	<i>t</i>	-0.001	-0.001	0.000	0.001	0.001	-0.003	
		<i>-2.20</i>	<i>-2.60</i>	<i>0.23</i>	<i>1.23</i>	<i>0.99</i>	<i>-5.54</i>	
	<i>t-1</i>						-0.001	-0.001
							<i>-2.84</i>	<i>-2.59</i>
ROA	<i>t</i>	-0.004	-0.005	-0.006	-0.006	-0.003		
		<i>-3.95</i>	<i>-3.59</i>	<i>-3.09</i>	<i>-2.22</i>	<i>-1.32</i>		
	<i>t-1</i>							-0.005
								<i>-3.58</i>
	<i>t-2</i>						-0.003	
							<i>-2.01</i>	
Solvency ratio	<i>t</i>	-0.005	-0.003	-0.003	-0.002	-0.002	-0.007	
		<i>-7.35</i>	<i>-3.60</i>	<i>-3.21</i>	<i>-1.83</i>	<i>-1.68</i>	<i>-7.39</i>	
	<i>t-1</i>							-0.003
								<i>-3.61</i>
	<i>t-2</i>						0.003	
							<i>3.22</i>	
Investment rate	<i>t</i>	-0.005	0.000	0.002	0.000	0.000	-0.005	
		<i>-4.99</i>	<i>0.17</i>	<i>1.40</i>	<i>0.22</i>	<i>-0.10</i>	<i>-3.62</i>	
Liquidity ratio	<i>t</i>	-0.001	-0.002	-0.001	-0.002	-0.001	-0.002	
		<i>-4.48</i>	<i>-4.68</i>	<i>-3.25</i>	<i>-3.23</i>	<i>-2.39</i>	<i>-3.99</i>	
	<i>t-1</i>							-0.002
								<i>-4.81</i>
Constant		-2.153	-2.085	-2.130	-1.951	-1.756	-2.092	-2.083
		<i>-20.17</i>	<i>-17.31</i>	<i>-14.28</i>	<i>-10.88</i>	<i>-14.92</i>	<i>-16.90</i>	<i>-17.32</i>
Number of observations		71 058	46 608	30 924	19 831	12 139	45 335	46 608
Number of firms		24 668	17 169	12 135	8 623	7 346	16 662	17 169
Pseudo-R ²		0.046	0.038	0.038	0.037	0.023	0.052	0.038
Wald Chi ²		347.0	196.4	119.0	65.4	55.7	250.2	196.2
Prob > Chi ²		<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>
rho		0.396	0.357	0.347	0.244	0.000	0.362	0.358
Prob >= chibar ²		<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.02</i>	<i>1.00</i>	<i>0.00</i>	<i>0.00</i>

Note: z-scores in italics. All regressions include the control dummies for size, sector and year presented in Table 3. All models estimated using a random-effects probit estimator, where the dependent variable is the dummy credit overdue. Coefficients refer to marginal effects. The pseudo-R² is a measure of the goodness of the fit. The Wald test evaluates the overall statistical significance of the estimated coefficients. Finally, rho assesses the proportion of the total variance resulting from the panel-level variance component.

tory power, though the liquidity and solvency ratios also provide interesting information when lagged by one year. Again, the investment rate fails to be significant when lagged.

As mentioned in the beginning of this article, empirical evidence suggests that developments in default probabilities over time should be largely associated with cyclical fluctuations of the economic activity. In this sense, it may be interesting to assess the contribution of firms' specific characteristics and of macroeconomic and financial conditions, by introducing macroeconomic variables in panel data regressions, as an alternative to simple time effects controls. The most insightful results are presented in Table 5. From all the variables considered, the most important seem to be the GDP growth rate or the

Table 5

PROBIT REGRESSIONS WITH MACROECONOMIC VARIABLES									
	Baseline specifica tion without time dummies	Baseline specifica tion with time dummies	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Sales growth	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001
	<i>-2.67</i>	<i>-2.20</i>	<i>-2.12</i>	<i>-2.14</i>	<i>-2.21</i>	<i>-2.33</i>	<i>-2.18</i>	<i>-2.21</i>	<i>-0.72</i>
ROA	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004	-0.004
	<i>-4.30</i>	<i>-3.95</i>	<i>-3.93</i>	<i>-3.96</i>	<i>-3.90</i>	<i>-4.16</i>	<i>-3.94</i>	<i>-3.94</i>	<i>-1.63</i>
Solvency ratio	-0.004	-0.005	-0.005	-0.005	-0.005	-0.004	-0.005	-0.005	-0.004
	<i>-7.06</i>	<i>-7.35</i>	<i>-7.37</i>	<i>-7.35</i>	<i>-7.34</i>	<i>-7.23</i>	<i>-7.37</i>	<i>-7.32</i>	<i>-3.08</i>
Investment rate	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.005	-0.006
	<i>-5.35</i>	<i>-4.99</i>	<i>-4.99</i>	<i>-4.99</i>	<i>-4.91</i>	<i>-5.25</i>	<i>-4.97</i>	<i>-4.95</i>	<i>-2.45</i>
Liquidity ratio	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.003
	<i>-4.52</i>	<i>-4.48</i>	<i>-4.46</i>	<i>-4.47</i>	<i>-4.50</i>	<i>-4.44</i>	<i>-4.49</i>	<i>-4.49</i>	<i>-4.28</i>
Interest rate on loans to firms								0.026	
								<i>2.26</i>	
Yield curve slope (10 y - 3 m)							-0.159		
							<i>-3.43</i>		
Loan growth					-0.023		-0.019		
					<i>-8.34</i>		<i>-6.02</i>		
PSI Geral variation						-0.002	-0.002		
						<i>-4.86</i>	<i>-3.48</i>		
GDP growth rate			-0.087						-0.141
			<i>-7.54</i>						<i>-6.47</i>
Coincident indicator of economic activity				-0.061				-0.075	
				<i>-7.14</i>				<i>-7.07</i>	
Sales growth * GDP growth rate									0.000
									<i>-0.16</i>
ROA* GDP growth rate									0.000
									<i>-0.16</i>
Solvency ratio* GDP growth rate									0.000
									<i>-0.35</i>
Investment rate * GDP growth rate									0.000
									<i>0.26</i>
Liquidity ratio* GDP growth rate									0.001
									<i>2.81</i>
1997		-0.303							
		<i>-5.59</i>							
1998		-0.230							
		<i>-4.55</i>							
1999		-0.341							
		<i>-6.37</i>							
2000		-0.390							
		<i>-6.51</i>							
2002		0.006							
		<i>0.12</i>							
Constant	-2.241	-2.153	-2.093	-2.192	-1.872	-2.274	-1.755	-2.321	-1.935
	<i>-23.26</i>	<i>-20.17</i>	<i>-20.38</i>	<i>-21.40</i>	<i>-17.64</i>	<i>-22.45</i>	<i>-14.57</i>	<i>-19.71</i>	<i>-16.78</i>
Number of observations	71 058	71 058	71 058	71 058	71 058	71 058	71 058	71 058	71 058
Number of firms	24 668	24 668	24 668	24 668	24 668	24 668	24 668	24 668	24 668
Pseudo-R²	0.037	0.046	0.043	0.042	0.044	0.040	0.045	0.043	0.044
Wald Chi²	333.8	347.0	330.3	327.3	345.7	323.3	344.3	338.3	336.2
<i>Prob > Chi²</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>
rho	0.336	0.396	0.393	0.392	0.384	0.371	0.395	0.383	0.395
<i>Prob >= chibar²</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>	<i>0.00</i>

Note: z-scores in italics. All regressions include the control dummies for size and sector presented in Table 3. All models estimated using a random-effects probit estimator, where the dependent variable is the dummy credit overdue. Coefficients refer to marginal effects. The pseudo-R² is a measure of the goodness of the fit. The Wald test evaluates the overall statistical significance of the estimated coefficients. Finally, rho assesses the proportion of the total variance resulting from the panel-level variance component.

coincident economic activity indicator (with a negative contemporaneous impact on default probabilities, in agreement with what would be expected), loan growth (which also displays a negative coefficient), the interest rate on loans to firms (with a positive coefficient, as expected), and the change in stock market prices (implying that positive developments in stock market prices seem to be associated with lower default probabilities). 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. Results suggest that these interactions are not particularly significant in determining default probabilities (only the interaction between the liquidity indicator and the GDP growth rate is significant). In general, macroeconomic variables shown in Table 5 have a considerable explanatory power, with relatively strong marginal effects on default probabilities. Furthermore, the inclusion of control variables for time effects or macroeconomic variables significantly improves the model's goodness of fit.

6. MAIN ECONOMETRIC RESULTS OBTAINED USING DURATION MODELS

The application of duration models to our research helps to understand not only why do firms default, but also when is default more likely to occur. In order to estimate that, we take into account information on the firm's survival since its creation date. However, the database used in this research only includes data from 1996 onwards, originating significant left-censoring problems. 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). This information is taken into account in the estimation of the regressions. Alternatively, in order to fully eliminate left-censoring, some estimations only considered firms created after 1996. However, in this case we are analysing a very specific group of newly created firms, which may show credit risk determinants very different from those of the remaining firms. In general, these firms have, on average, higher investment rates and indebtedness levels, which is consistent with their life phase. Chart 2

Chart 2

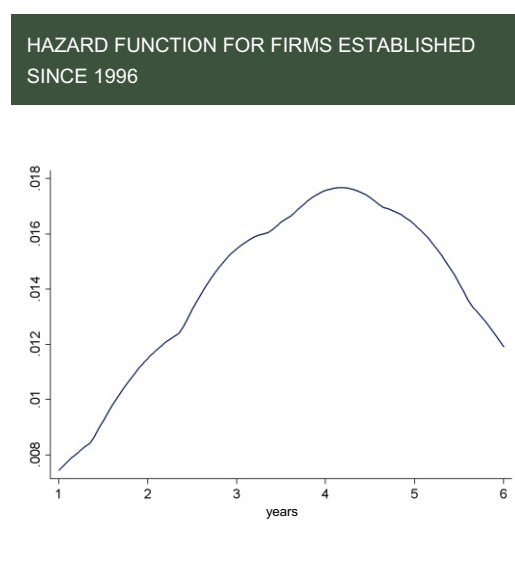


Table 6

COX REGRESSIONS (HAZARD RATIOS)

	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 <i>-1.72</i>			1.003 <i>1.54</i>								
ROA	0.995 <i>-4.33</i>	0.994 <i>-4.83</i>	0.994 <i>-4.84</i>	0.992 <i>-2.44</i>	0.992 <i>-2.31</i>	0.992 <i>-2.31</i>	0.993 <i>-2.02</i>	0.993 <i>-1.79</i>	0.992 <i>-2.36</i>	0.992 <i>-2.37</i>	0.993 <i>-2.17</i>	0.992 <i>-2.49</i>
Solvency ratio	0.995 <i>-4.59</i>	0.995 <i>-4.56</i>	0.995 <i>-4.53</i>	1.003 <i>0.74</i>	1.003 <i>0.78</i>		1.003 <i>0.74</i>	1.000 <i>-0.06</i>	1.005 <i>1.29</i>	1.003 <i>0.70</i>	1.002 <i>0.59</i>	1.003 <i>0.77</i>
Investment rate	0.990 <i>-3.94</i>	0.989 <i>-4.10</i>	0.989 <i>-4.12</i>	0.993 <i>-1.23</i>	0.994 <i>-1.02</i>	0.994 <i>-1.02</i>	0.994 <i>-1.00</i>	0.993 <i>-1.47</i>	0.994 <i>-1.04</i>	0.994 <i>-1.04</i>	0.994 <i>-1.08</i>	0.994 <i>-1.02</i>
Liquidity ratio	0.995 <i>-4.53</i>	0.995 <i>-4.51</i>	0.995 <i>-4.54</i>	0.990 <i>-3.94</i>	0.990 <i>-4.04</i>	0.990 <i>-4.04</i>	0.990 <i>-3.89</i>	0.993 <i>-2.97</i>	0.986 <i>-5.01</i>	0.990 <i>-3.99</i>	0.990 <i>-3.97</i>	0.990 <i>-3.98</i>
Leverage						0.997 <i>-0.78</i>						
Share of tangible fixed assets							0.994 <i>-0.77</i>					
Turnover ratio								0.996 <i>-2.26</i>				
Available collateral (approx.)									0.994 <i>-1.32</i>			
Firm established since 1996 (Y/N)			0.962 <i>-0.23</i>									
GDP growth rate										1.030 <i>0.24</i>		
Loan growth											0.991 <i>-0.39</i>	
PSI Geral variation												1.005 <i>0.98</i>
Constant	-	-	-	-	-	-	-	-	-	-	-	-
Number of observations	76 292	76 292	76 292	3 847	3 847	3 847	3 802	3 847	3 802	3 847	3 847	3 847
Number of firms	25 690	25 690	25 690	2 324	2 324	2 324	2 297	2 324	2 297	2 324	2 324	2 324
Number of failures	1 000	1 000	1 000	68	68	68	67	68	67	68	68	68
Time at risk	76 292	76 292	76 292	3 847	3 847	3 847	3 802	3 847	3 802	3 847	3 847	3 847
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 include all firms established since 1996. Regressions for the full sample include the control dummies for size, sector and year presented in Table 3. All models estimated using a Cox regression which evaluates the time until default, using robust variance estimates. A coefficient lower than 1 should be interpreted as contributing to an increase in time until default eventually occurs. The Wald test evaluates the overall statistical significance of the estimated coefficients.

presents hazard functions for this group of firms, showing that their default probabilities increase up to the 4th year of the firm, gradually declining thereafter.¹³

Within the framework of duration modelling, we estimated several regression models, in a spirit similar to that of discrete choice models. The results obtained are broadly consistent with those obtained with probit models: firms with higher profitability, higher solvency, higher investment rates, and better liquidity ratios should take a longer time to eventually default on their loan commitments (Table 6). However, sales growth turns out to be clearly non-significant in these models. Hence, though sales growth may contribute to explain why some firms default, it does not seem to determine the time until default. 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 model including a dummy variable for such firms (model 3 in Table 6). This dummy variable is far from being significant, suggesting that these firms do not substantially differ from the remaining firms in the sample, after controlling for a set of firm's financial characteristics.

However, the only way to totally eliminate left-censoring in the sample is to exclude all firms that have not been observed since their creation date, estimating regressions only for the sub-group of firms created after 1996. The results of these estimations are also shown in Table 6. In general, the solvency ratio and the investment rate are no longer statistically significant, suggesting that default probabilities for start-up firms may have slightly different determinants than those of more mature firms. Several alternative specifications were considered, including the introduction of macroeconomic 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.

7. CONCLUDING REMARKS

The results obtained in the literature suggest that in periods of strong economic growth, which are sometimes accompanied by strong credit growth, there may be some tendency towards excessive risk-taking, amid some market optimism. Cumulative imbalances created in such periods will tend to become apparent only when economic activity slows down markedly. Against this background, this article examines the determinants of credit risk, taking into account firm-specific idiosyncratic factors, as well as systematic factors, which simultaneously affect all economic agents.

For that purpose, an extensive database including financial information for more than 30.000 Portuguese firms was thoroughly analysed. The results obtained suggest that default probabilities are affected 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, firm's size does not seem to significantly affect default probabilities. However, there are considerable differences across 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.

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

(13) For a discussion on the determinants of survival probabilities for young firms, see *Earinha* (2005). The results shown in Chart 2 are very similar to those obtained in that study.

situation, there are important relationships between overall macroeconomic conditions and default rates, which should be assessed from a financial stability perspective.

REFERENCES

- Antunes, A., Ribeiro, N. and Antão, P. (2005), "Estimating probabilities of default under macroeconomic scenarios", *Financial Stability Report*, Banco de Portugal.
- 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.
- 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.
- 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.
- Bonfim, D. (2007), "Credit risk drivers: evaluating the contribution of firm level information and macroeconomic dynamics", *Working Paper* No.7, Banco de Portugal.
- 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.
- Carling, K., Jacobson, T., Lindé, J. and Rozbach, K.(2007), "Corporate credit risk modeling and the macroeconomy", *Journal of Banking and Finance*, No. 31, 845-868.
- Couderc, F. and Renault, O. (2005), "Times-to-default: life cycle, global and industry cycle impacts", *FAME Research Paper* No. 142.
- Eklund, T., Larsen, K. and Bernhardsen, E. (2001), "Model for analysing credit risk in the enterprise sector", Norges Bank *Economic Bulletin* Q3 01.
- Farinha, L. (2005), "The survival of new firms: impact of idiosyncratic and environmental factors", *Financial Stability Report*, Banco de Portugal.
- 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.
- Lancaster, T. (1990), *The econometric analysis of transition data*, *Econometric Society Monographs*, Cambridge University Press.
- 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.
- Shumway, T. (2001), "Forecasting bankruptcy more accurately: a simple hazard model", *The Journal of Business*, Vol.74, No.1, 101-124.