

# Firm default probabilities revisited

**António Antunes**

Banco de Portugal and NOVA SBE

**Homero Gonçalves**

Banco de Portugal

**Pedro Prego**

Banco de Portugal

April 2016

## Abstract

This article describes a tool to assess the creditworthiness of the Portuguese non-financial firms. In its design, the main goal is to find factors explaining the probability that any given firm will have a significant default episode vis-à-vis the banking system during the following year. Using information from the central credit register for period 2002–2015 and a comprehensive balance sheet data set for period 2005–2014, we develop a method to select explanatory variables and then estimate binary response models for ten strata of firms, defined in terms of size and sector of activity. We use this methodology for the classification of firms in terms of one-year probability of default consistent with typical values of existing credit rating systems, in particular the one used within the Eurosystem. We provide a brief characterisation of the Portuguese non-financial sector in terms of probabilities of default and transition between credit rating classes. (JEL: C25, G24, G32)

---

## Introduction

This article describes a tool to assess the creditworthiness of the Portuguese non-financial firms. The main goal is to find factors explaining the probability that any given firm will have a significant default episode vis-à-vis the banking system during the following year. The output of this tool is a probability of default in banking debt with a one-year horizon. This value is then mapped into a masterscale where companies are grouped into homogeneous risk classes. The fact that credit quality is assessed only in terms of banking debt is essentially not limiting our analysis for two reasons. First, most credit in Portugal is granted by banks. Only a few large firms typically issue market debt. Second, defaults in issued debt should be highly correlated with defaults in bank loans.

---

Acknowledgements: We thank Lucena Vieira for skilfully supplying the data, Manuel Lingo and Florian Resch (Oesterreichische Nationalbank) for sharing with us their expertise in the design of credit rating systems, and our colleagues at the Statistics Department and Economics and Research Department who helped us in this project.

E-mail: [aantunes@bportugal.pt](mailto:aantunes@bportugal.pt); [hgoncalves@bportugal.pt](mailto:hgoncalves@bportugal.pt); [pmprego@bportugal.pt](mailto:pmprego@bportugal.pt)

Credit Quality Step	Upper default probability limit
1 & 2	0.1
3	0.4
4	1.0
5	1.5
6	3.0
7	5.0
8	100

TABLE 1. Credit Quality Steps within the Eurosystem. All values in percentage.

Source: ECB.

Each risk class will be labeled by a “credit rating” and in the rest of this article we will refer to a risk class using its label. A credit rating is then a synthetic indicator reflecting several features (e.g. solvency, liquidity, profitability) that measure the firm’s ability to fulfill its financial commitments.

In the current exercise the Eurosystem’s taxonomy will be used, where a credit rating is designated by “Credit Quality Step”. Table 1 presents the different risk classes and the associated upper limits of the probability of default. See ECB (2015) for additional details.

This article is partly based on previous efforts made in Martinho and Antunes (2012), but there is a vast policy and scholarly literature on the topic (see, for example, Coppens *et al.* 2007; Lingo and Winkler 2008; Figlewski *et al.* 2012), as well as a variety of documents produced by public and private institutions, including the European Central Bank (ECB), the European Banking Authority (EBA), Fitch Ratings, Moody’s and Standard & Poors.

Credit ratings are used in a variety of situations. The most obvious one relates to the banks’ credit allocation process. Ratings are indeed an important tool for lenders to select the borrowers according to their predefined risk appetite and to determine the terms of a loan. A higher credit ranking usually means better financing terms, including lower costs and access to more diversified instruments such as, for instance, securities markets.

Periods of broader materialisation of credit risk, like the one recently experienced in Portugal, put even more emphasis on the relevance of the firms’ credit assessment process. Data for 2015 show that the total debt of non-financial corporations in Portugal represents 115% of GDP, one of the highest values in the euro area. A considerable share of this debt is in banks’ balance sheets, where non-financial corporations were responsible for close to 28% of the total bank credit (bank loans and debt securities). The quality of these credits has been deteriorating substantially over the last years, putting pressure on the banks’ results and capital requirements. Between December 2008 and December 2015 the non-performing loans ratio of non-financial corporations increased from 2.2% to 15.9%. In the same period the share of

companies with overdue loans rose 10 percentage points to 29% in December 2015.

Early warning systems that can help predict future defaults are therefore of utmost relevance to support, at the banks' individual level, the credit allocation process and, at the aggregated level, the analysis of the financial stability of the overall banking system. Credit ratings are useful because they allow regulators and other agents in the market to identify potential problems that may be forthcoming in particular strata of firms—for example, defined in terms of activity sector or size. This is particularly important in an environment where banks' incentives in terms of reporting accurately and consistently probabilities of defaults of firms have been challenged. For example, Plosser and Santos (2014) show that banks with less regulatory capital systematically assign lower probabilities of default to firms than banks with more regulatory capital. This underreporting then implies that, for a loan with the same firm, different banks will constitute different levels of capital.

Credit ratings can also be useful as input for stress tests in order to evaluate the impact that changes in the economic environment may have on the financial sector performance. These measures can be used to estimate expected losses within a given time frame and are therefore key instruments for the risk management of financial institutions as well as for supervisory purposes. For this last purpose, it is important as well to have a benchmark tool to validate the capital requirements of each financial institution.

The existence of independent credit assessment systems also supports investment. As investment opportunities become more global and diverse, it is increasingly difficult to decide not only on which countries but also on which companies resources should be allocated. Measuring the ability and willingness of an entity to fulfil its financial commitments is key for helping make important investment decisions. Oftentimes, investors base part of their decisions on the credit rating of the company. For lenders it is difficult to have access and to analyse detailed data about each individual company presenting an investment opportunity. These grades are used as well to design structured financial products and as requirements for inclusion of securities portfolios eligible for collateral in various operations of the financial institutions.

The existence of this kind of indicator is also important for the borrower as it can provide better access to funding. Moreover, management and company owners can also use credit ratings to get a quick idea of the overall health of a company and for a direct benchmark with competitors.

Under the Eurosystem's decentralised monetary policy framework, national central banks grant credit to resident credit institutions. In order to protect the Eurosystem from financial risk, eligible assets<sup>1</sup> must be posted

---

1. Eligible collateral for refinancing operations includes not only securities but also credit claims against non-financial corporations.

as collateral for all lending operations. The Eurosystem Credit Assessment Framework (ECAAF) defines the procedures, rules and techniques which ensure that the Eurosystem requirement of high credit standards for all eligible assets is met. Credit assessment systems can be used to estimate non-financial corporations' default risk. On the one hand, this credit assessment dictates whether credit institutions can pledge a certain asset against these enterprises as collateral for monetary policy operations with the national central bank. On the other hand, in the case of eligible assets, the size of the haircut is also based on the credit rating.<sup>2</sup>

For economic analysis, credit ratings are particularly relevant to evaluate the monetary policy transmission mechanism and to gauge the health of quality of credit flowing to the economy through the financial system. For instance, this tool can be used to evaluate if companies with the same level of intrinsic risk are charged the same cost by the banks or if there are additional variables determining the pricing of loans. There are a number of theories explaining these differences, typically in terms of asymmetries of information or the level of bank capital (see, for example, Santos and Winton 2015, and also Plosser and Santos 2014). It is also particularly interesting to compare firms from different countries of the euro area and quantify the component of the interest rate that can be attributed to the company risk, and the part stemming from other reasons, namely problems in the monetary policy transmission mechanism or country-specific risk. The data used by credit assessment systems is also valuable to identify sustainable companies that are facing problems because of lack of finance. This information can be used to help design policy measures to support companies that have viable businesses but whose activity is constrained by a weak financial system.

For statistical purposes the use of credit ratings is straightforward. Indeed, any statistic based on individual company data can be broken down into risk classes. For example, it can be valuable to compile interest rate statistics by risk class of the companies or to simply split the total bank credit by risk classes.

In order to describe a rating system suitable for the uses described above, this article is structured as follows. First, the data are presented and the default event is defined based on the available data and appropriate conventions. Second, the methodology underpinning the rating system is described. Then a calibration exercise is performed to fine-tune the model to the credit assessment system used within the Eurosystem. Fourth, some results are presented in terms of model-estimated and observed default rates and transitions among credit risk classes. Finally, a conclusion is provided.

---

2. To assess the credit quality of collateral, the Eurosystem takes into account information from credit assessment systems belonging to one of four sources: (i) external credit assessment institutions (ECAI); (ii) national central banks' in-house credit assessment systems (ICAS); (iii) counterparties' internal ratings-based systems (IRB); and (iv) third-party providers' rating tools (RT).

## Data

The analysis in this article uses Banco de Portugal's annual *Central de Balanços* (CB) database—which is based on *Informação Empresarial Simplificada* (IES), an almost universal database with detailed balance sheet information of Portuguese firms—and the *Central de Responsabilidades de Crédito* (CRC), the Portuguese central credit register. CB contains yearly balance sheet and financial statements from virtually all Portuguese corporate firms, both private and state owned, since 2005 until 2014, which is the most recent year available. One of the main benefits of using CB is the ability to perform the analysis at the micro level. CRC records all credit institutions' exposures to Portuguese firms and households at monthly frequency, providing firm- and individual-level information on all types of credit and credit lines. For the purpose of this analysis, the time span ranges from 2002 until 2015.

In this article only private non-financial firms with at least one relationship vis-à-vis the financial sector were considered, which for the sake of simplicity will only be referred to as firms. The main reason for the exclusion of firms with no bank borrowing is that the aim is to estimate default probabilities. In addition, on the CB side observations regarding self-employed individuals and firms that reported incomplete or incoherent data, such as observations with negative total assets or negative business turnover, were excluded. As for the CRC, only information regarding performing and non-performing loans was considered, and credit lines, write-offs and renegotiated credit were disregarded. Moreover, all firm-bank relationships below €50 and firms that had an exposure to the financial system as a whole (aggregated over all the firm-bank relationships) below €10,000 were excluded.

### *Default definition*

A firm is considered to be “in default” towards the financial system if it has 2.5 per cent or more of its total outstanding loans overdue. The “default event” occurs when the firm completes its third consecutive month in default. A firm is said to have defaulted in a given year if a default event occurred during that year. It is possible for a single firm to record more than one default event during the period of analysis but, in order to make sure we are not biasing the sample towards firms with recurrent defaults, we exclude all observations of the firm after the first default event.

We only include firms that either are new to the financial system during the sample period (that is, firms which did not have banking relationships before 2005, possibly because they did not even exist) or have a history of three years with a clean credit record. We exclude firms that enter the CRC database immediately in default.

#	Size
1	Micro
2	Small, medium and large
#	Industry
1	Manufacturing, mining and quarrying
2	Construction and real estate activities
3	Wholesale and retail trade and the primary sector
4	Utilities, transports and storage
5	Services

TABLE 2. Size and industry groups of firms.

Source: Banco de Portugal.

### *Data treatment and definitions of variables*

In order to increase group homogeneity, we split the sample into micro firms and all other firms (i.e., small, medium and large firms). These two groups were further divided based on the firms' classification into thirteen industry NACE groups. Some industries were bundled according to their affinity, as was for instance the case of the real estate sector and the construction sector. We ended up with five groups of industries (manufacturing, mining and quarrying; construction and real estate activities; wholesale and retail trade and the primary sector; utilities, transports and storage; services) and two groups for size (micro firms; all other firms), in a total of ten groups of firms to be used in the econometric estimations. See Table 2.

The CB database contains detailed balance sheet data of Portuguese non-financial firms. For the purpose of this analysis, only a subset of CB's variables were used. The large pool of variables can be categorised into specific groups such as leverage, profitability, liquidity, capital structure, dimension, and a residual group which corresponds to variables related with the balance sheet ratios that do not fit in any of the groups previously defined. All the level variables are scaled by dividing them by either the firm's total assets, current liabilities or total liabilities, depending on the case. We never use denominators that can have negative values as that would create significant discontinuities when the denominator is close to zero. To account for the possible influence of the economy as a whole on a specific firm, we consider a small set of macro factors: nominal and real GDP growth, total credit growth and the aggregate corporate default rate. This choice was motivated by previous literature on the topic; for example, Figlewski *et al.* (2012) have found that real GDP growth and the corporate default rate help explain transitions across rating classes. Table 3 summarises the subset of CB variables and the macro factors used in this analysis.

Measures of:	Variables
Leverage	Financial debt; Bank debt; Interest paid
Profitability	Value-added per worker; Profit / Loss; EBIT; Cash flow; EBITDA
Liquidity	Cash; Current liabilities
Capital structure	Equity; Current assets; Tangible assets
Dimension	Total assets; Age; Turnover; Employees
Other idiosyncratic	Wages; Trade debt
Macroeconomy	Aggregate default rate; Credit growth; Nominal GDP growth; Real GDP growth

TABLE 3. Summary of variables used in the regressions.

Source: Banco de Portugal. Precise definition of variables available upon request.

As previously mentioned, firms that had negative total assets, liabilities or turnover were removed from the analysis. Additionally, firms with total assets, turnover or the number of employees equal to zero were excluded. In order to cope with values for skewness and kurtosis far from what would be expected under the Normal distribution, strictly positive variables were transformed into their logarithms in order to reduce skewness. Because this transformation is not applicable to variables that can be negative, the set of variables was expanded with the ranks of all variables normalised between 0 and 1. The rank transformation was applied within each year-size-industry group to increase homogeneity. A final group of well-behaved variables was kept unchanged. This included variables expressed in shares and macro variables.

## Methodology

In this study, we develop an approach based on a multi-criteria system of variable selection out of a large pool of potential variables. We build upon the methodology used by Imbens and Rubin (2015) of explanatory variables selection through maximum likelihood estimation. This methodology selects variables in an iterative process based on the explanatory prediction power that each variable is able to provide. A variable under scrutiny will be included if the increase in explanatory power is above a certain threshold. We adapt this approach for our own purposes.

### *Selection of explanatory variables*

More specifically, we start by estimating a base model with fixed effects for size (only for non micro-sized firms) and for activity sector (at a disaggregation level of a few sectors per industry). For each variable of the

initial pool of  $N$  variables, we estimate a model with the fixed effects plus that variable. These regressions will then be compared to the base model by using a likelihood ratio (LR) test. The algorithm then picks the variable associated to the model with the highest likelihood statistic under the condition that it is above the initial likelihood at a 5% significance level; this corresponds to an LR ratio of at least 3.84.

The process is then repeated but the base model is now the model with the fixed effects plus the variable picked in the previous step. The next variable is to be chosen among the remaining pool of  $N - 1$  variables, but from this second step on we add criteria other than the requirement in terms of the LR. These criteria address potential problems stemming from a completely agnostic inclusion of variables. More specifically, the following conditions are added in order for the candidate variable to be included in the model:

1. It must have linear and non-linear correlation coefficients with any of the variables already present in the model lower than 0.5. This condition aims at avoiding potential problems of multicollinearity.
2. It has to be statistically significant at the 5% level in the new regression, while all of the previously included variables must remain statistically significant. This is to avoid that non significant variables survive in the final model specification.
3. It has to be such that the new model estimate improves the AUROC criterion<sup>3</sup> relative to its previous value. In addition, the new model estimate also has to improve the AIC information criterion. This condition addresses the potential problem of over-fitting the model, as this criterion penalises the inclusion of parameters.

The process ends when none of the remaining variables in the set of potential variables fulfills all the conditions 1–3 or, to avoid the proliferation of parameters, a maximum of ten variables has been reached. In order to maintain the approach as replicable and as simple as possible, a Logit specification was chosen.

All ten models (one for each combination between two size categories and five industries) were estimated by pooling the existing observations together, spanning the period from 2005 to 2014 in terms of the balance sheet information. All explanatory variables pertain to the end of the current year  $t$ . The dependent variable is defined as an indicator of the default event during year  $t + 1$ . Note that when the restriction on the maximum number of variables is removed none of the ten models includes more than 13 variables. Moreover, when analysing the evolution of the AUROC with each variable added it

---

3. AUROC stands for “area under the Receiver Operator Characteristic”. See Lingo and Winkler (2008) and Wu (2008) for the definition and the stochastic properties of this synthetic measure.



is possible to see that this benchmark tends to flatten out before the tenth variable; see Figure 1.

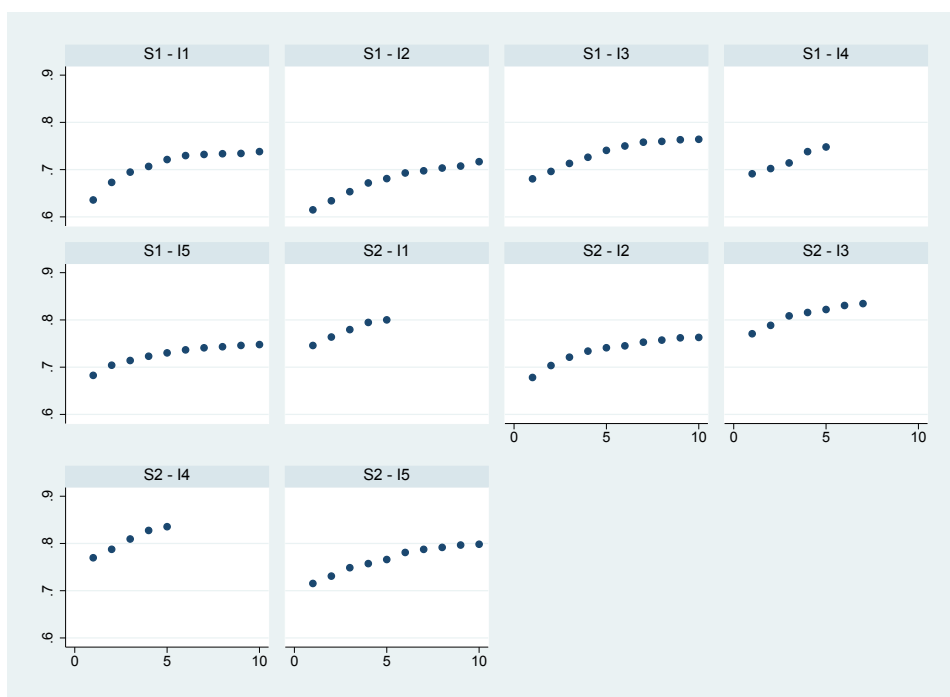


FIGURE 1: The AUROC as a function of the number of variables selected according to the methodology defined in the text. S# means size group # and I# means industry #; see Table 2 for details.

Source: Banco de Portugal and authors' calculations.

### *A summary of the results*

After applying the proposed methodology to our data set, we obtained ten estimated Logit models; Table 4 displays some information characterising them.<sup>4</sup> A first observation is the overall consistent goodness-of-fit, which can be gauged by the AUROC.<sup>5</sup> These values lie in the range 0.72–0.84 and reject comfortably the hypothesis that the models are not distinguishable from

4. In practice we did not use the original variables, except in cases where they represented shares or growth rates, because the algorithm always chose the transformed variables (logarithm or rank).

5. For a critique of the AUROC as a measure of discriminatory power in the context of model validation, see Lingo and Winkler (2008).

Group	Obs.	Defaults	Def. ratio	# variables	AUROC	Brier Score
S1 - I1	58063	3000	5.17%	<b>10</b>	0.738	0.047
S1 - I2	53543	2965	5.54%	10	0.717	0.050
S1 - I3	178178	7696	4.32%	<b>10</b>	0.764	0.039
S1 - I4	2681	121	4.51%	5	0.748	0.041
S1 - I5	123048	5336	4.34%	<b>10</b>	0.748	0.040
S2 - I1	98065	3887	3.96%	5	0.800	0.035
S2 - I2	58325	3861	6.62%	10	0.763	0.057
S2 - I3	96738	3062	3.17%	7	0.835	0.028
S2 - I4	3903	128	3.28%	5	0.836	0.030
S2 - I5	73782	2476	3.36%	10	0.798	0.031
Overall	746326	32532	4.36%	n.a.	0.777	0.0393

TABLE 4. A summary of the Logit estimations for ten strata of firms. Values in bold mean that the procedure was stopped due to the limit on explanatory variables. S# means size group # and I# means industry #; see Table 2 for details.

Source: Banco de Portugal and authors' calculations.

a random classifier. Also, in each model the Brier score, a measure of the goodness of fit, is considerably small. The Spiegelhalter (1986) test applied to each model (not reported) also indicates that the level predicted for the probability of default is consistent with the observed defaults.

Although the methodology includes ten separate models there are several similarities among them. Table 5 presents a summary of the variables more often chosen using the procedure described above. Most importantly, the different models seem to have a core group of variables, even if they enter different models in slightly different variants: for instance, cash to total assets or cash to current assets as a measure of liquidity are always chosen, although they are never chosen together for the same model.

All ten models include a measure for profitability, alternating between cash-flow to total assets or earnings to total assets, and a measure for liquidity. Nine out of the ten models include the cost of credit as well as short-term liabilities, measured by current liabilities to total assets. Eight models include a measure for leverage and seven models include the weight of the employees' wage bill to total assets. Seven models select one macro factor among nominal GDP growth, total credit growth and the aggregate default rate. Finally, six models include the age of the firm and five models include a proxy for the firm's productivity as measured by value-added per worker.

Curiously, the weight of trade debt to total liabilities is also selected for five different models, all of them pertaining to micro-sized firms. This indicates that for this group of firms the behaviour of suppliers is particularly important.

Variable	# times selected	Avg. rank	Sign in regs.
r(Cash flow / Total assets)	6	1.0	-
r(Earnings / Total assets)	4	1.0	-
r(Interest paid / Financial debt)	9	3.1	+
r(Current liabilities / Total assets)	5	3.4	+
r(Age)	5	4.4	-
r(Wages / Total assets)	5	4.6	-
r(Cash / Current assets)	6	6.0	-
r(Financial debt / Total assets)	5	5.6	-
log(Current liabilities / Total assets)	4	4.5	+
r(Cash / Total assets)	4	5.8	-
r(Trade debt / Current liabilities)	5	7.2	+
log(Financial debt / Total assets)	2	3.0	+
r(Value-added per worker)	5	7.8	-
Nominal GDP growth	3	6.3	-
Credit growth	2	5.0	+
Agg. Default Rate	2	5.0	+
log(Equity / Total assets)	1	3.0	-
r(Bank debt / Total assets)	1	4.0	+
r(Employees)	2	8.5	+
log(Wages / Total assets)	2	9.0	-
log(Turnover)	2	9.5	-
log(Age)	1	7.0	-
r(Current assets / Total assets)	1	10.0	-

TABLE 5. Qualitative results of the variable selection procedure.  $r(\cdot)$  denotes the rank of the variable within the size-industry group in the current year;  $\log(\cdot)$  denotes the natural logarithm of the variable. The second column contains the number models for which the variable is chosen by the variable selection algorithm (out of a maximum of ten models). The third column contains the average rank (1 is first, 10 is tenth) with which the variable was chosen. The fourth column contains the sign of variable's coefficient in the Logit estimation of the default event. Variables are ordered by the product of the inverse of number of times chosen by the average rank, in ascending order. S# means size group # and I# means industry #; see Table 2 for details.

Source: Banco de Portugal and authors' calculations.

Another significant result is that the variables that are more often chosen by the algorithm are also among the first variables to be selected, which indicates that these variables have the largest contribution to the explanatory power of the model. In particular, the variables measuring profitability are the first to be picked by the algorithm in the ten different models.

Another important observation is that the coefficient of each variable always enters the model with the sign that would be expected, even though the algorithm does not impose any restriction to this effect. Moreover, when a variable is selected for more than one model the variable's coefficient sign is the same across those models.

## Rating class calibration

The next step in the setup of a rating tool system is to calibrate the model so that observed default rates of firms at any given credit category are consistent with the typical default rates used to define them (see Table 1). This step is usually needed because, while the average of the conditional model-estimated default probability should match the observed average default rate, this need not be so across different groups of firms, and in particular across rating classes. One basic requirement for the calibration that we want to perform is that overall the observed default rate is consistent with the conditional default rate stemming from the estimated models. While this requirement is generally fulfilled in-sample, one question remains: is the model conditional default probability consistent also across different categories of risk?

To answer this question, let us first define the concept of *z-score* in the context of our analysis. The Logit model used in the methodology described above is framed in terms of an unobserved latent variable which is then transformed into a number between 0 and 1, the probability of default. To keep the analysis simple, it suffices to say that the coefficients  $\beta$  of each one of the Logit models are estimated so that the probability of default is, to the extent possible, accurately given by

$$\Pr\{\text{default}_{t+1} = 1|x_t\} = \frac{1}{1 + e^{-x_t\beta}}$$

where  $\text{default}_{t+1}$  is an indicator of a default event occurring in year  $t + 1$ ,  $x_t$  is a (row) vector of regressors in year  $t$ —including a constant and variables characterising the firm and possibly the economy—and  $\beta$  is a (column) vector of coefficients. It is a property of these coefficients that the in-sample average of the predicted default rates (as computed by the equation above) is equal to the observed average default rate. The *z-score* of each observation is simply defined as the estimated value of the latent variable, that is,  $z_t = x_t\beta$ .

The answer to the question above is broadly positive. Figure 2 depicts the model-predicted default probabilities (the dash-dotted curve) along with average observed default rates (the dots in the graph). Each point represents the fraction of defaults for groups of firms with relatively similar *z-scores*. The lower (more negative) the *z-score*, the lower the estimated probability of default of the firm. We can see that using a Logit specification does a good job explaining the relationship between *z-scores* and observed default probabilities for groups of firms across the whole *z-score* distribution.

One way to try to improve the fit is to have a more flexible approach. While this procedure is not consistent with the estimation process, we view that as a fine-tuning exercise rather than something that invalidates the results obtained using regression analysis. The solid line is one such attempt: it is a semiparametric curve interpolating the dots. It is readily seen that the two curves (the Logit and the semiparametric) are really telling the same story, but

the semiparametric one lies above the Logit for very negative z-scores. This means that, for that range of z-scores, the semiparametric curve is going to be more conservative in assigning probabilities to firms.

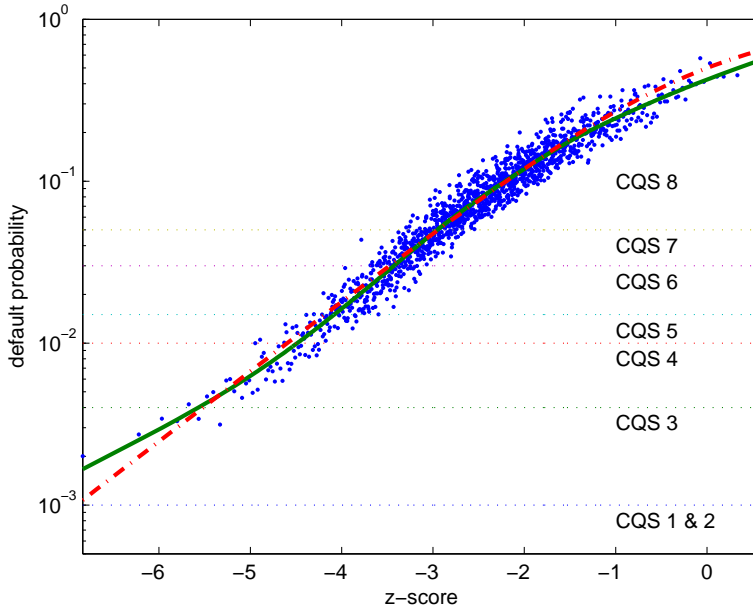


FIGURE 2: Probabilities of default of firms. Each dot represents the observed default rate for groups of firms with similar z-scores. Upper limits for default probabilities of each Credit Quality Step as defined by the Eurosystem also depicted.

Source: Banco de Portugal and authors' calculations.

We now provide additional details on the procedure of fitting the semiparametric curve to the dots, but the reader uninterested in mathematical details can safely skip the following section.

### *Fitting the dots*

The dots in Figure 2 are empirical probabilities of default for groups of observations in the sample. Each dot in the graph represents a pair from the set of points  $S^n = \{(d_q^n, \hat{z}_q^n)\}_{q=1, \dots, Q^n}$ . These points were obtained as follows. First we sorted in ascending order all the z-scores (which are normalised and can be compared across the different groups of firms) of the sample. We then identified the first  $n$  defaults and set  $r_1^n$  as the order number of the observation with the  $n^{\text{th}}$  default. We grouped these observations in set  $A_1^n = \{z_1, \dots, z_{r_1^n}\}$ . We then computed the ratio  $\hat{d}_1^n = \frac{n}{\#A_1^n}$  and defined  $\hat{z}_1^n$  as the median of set  $A_1^n$ . We repeated the procedure for the next group of  $n$  defaults by finding

set  $A_2^n = \{z_{r_1^n+1}, \dots, z_{r_2^n}\}$ , default rate  $\hat{d}_2^n = \frac{n}{\#A_2^n}$  and median z-score  $\hat{z}_2^n$ . This process was carried out in a similar fashion until we exhausted all the observations, ending up with a total of  $Q^n$  pairs of empirical default rates and z-scores. Notice that, for all  $q$ ,  $\hat{z}_{q-1}^n \leq \hat{z}_q^n \leq \hat{z}_{q+1}^n$ , that is, these points are also sorted in ascending order in terms of the z-scores, although not necessarily in terms of default probabilities. Not all points were plotted in Figure 2; only a representative sample was.

One word about the choice of  $n$ . If this number is too small then the standard deviation of the estimated empirical probability will be relatively high. To see this, assume that the default event has a Binomial distribution within  $A_q^n$ , and take  $\hat{d}_q^n$  as an estimator for the default probability. Then, an estimate of the standard deviation of  $\hat{d}_q^n$  would be

$$\sqrt{\frac{\hat{d}_q^n(1 - \hat{d}_q^n)}{\#A_q^n - 1}}$$

which decreases with  $\#A_q^n$ . We picked  $n = 23$  in our simulations because, due to the relative scarcity of very negative z-scores (associated to relatively low probabilities of default), we wanted to have meaningful estimates for default rates even in high rating classes. With this choice we ended up with  $Q^{23}$  close to 1400. We later address the significance of the estimates obtained with this choice. The robustness of the general results of this analysis with respect to this choice is performed elsewhere. For commodity we will drop  $n$  from the notation described above.

In order to keep the analysis as standard and simple as possible, we fitted a smoothing spline to the points in the figure. The smoothing spline is a semiparametric curve that approximates a set of points in a graph while penalising the occurrence of inflexion points along the whole curve. More specifically, we chose the following specification:

$$s(\cdot) = \arg \min_p \sum_{q=1}^Q (\log(\hat{d}_q) - s(\hat{z}_q))^2 + (1 - p) \int_{\hat{z}_1}^{\hat{z}_Q} (s''(z))^2 dz.$$

In this formulation, function  $s : [\hat{z}_1, \hat{z}_Q] \rightarrow ] - \infty, 0]$  is a cubic spline defined over the set of points in  $S$ . A cubic spline is a set of cubic polynomials defined in intervals and “glued” together at the unique z-scores contained in  $S$ . By construction,  $s(\cdot)$  has continuous second derivative  $s''(\cdot)$  in all points. Parameter  $p$  governs the smoothness of the interpolating curve. If  $p$  is close to 1, one gets the so-called natural cubic interpolant, which passes through all the points in  $S$ .<sup>6</sup> If  $p$  is close to 0, the penalisation of the second derivative

---

6. Technically, if there are points in  $S$  with the same z-score, the natural interpolant passes through the average of the log default rates among all the points with the same z-score.

ensures that the solution will be the linear interpolant, which has zero second derivative.

The curve of the smoothing spline with  $p = 0.3$  is depicted in Figure 2 as the solid line.

One thing that is clear from Figure 2 is that the empirical default probability will still be a noisy measure: while each point represents the median z-score for the set of observations leading to a given number of observed defaults (23 defaults), it is possible to have groups of very similar firms—in the sense they have very similar z-scores—and still observe relatively different observed default rates among those groups of firms. That concern is addressed by the models' performance in terms of the AUROC, which has already been presented. In any case, the general shape of the cloud of points tells us that the analytical framework captures well the probability of default across firms: a random model would yield a cloud coalescing along an horizontal line in the graph at the unconditional observed default rate. The figure then underlines that even when large AUROC measures can be obtained, the default event is still a very uncertain event.

### *Defining credit quality classes*

The general approach chosen for the purpose of categorising firms in terms of credit default classes is (i) to obtain reference values for default probabilities from external sources, then (ii) to choose thresholds in terms of z-scores for the different credit classes, and finally (iii) to check *ex post* the observed in-sample default probabilities' consistency with the previously defined credit classes. We also provide a more detailed analysis of the transitions of firms across credit categories and to default.

We now turn to the question of defining credit quality classes. The horizontal dashed lines of Figure 2 represent upper limits of credit classes according to the Eurosystem credit quality system (see Table 1). For example, class 3 corresponds, in the standard framework of monetary policy, to the lowest-rated firms whose loans can still be posted as collateral by financial institutions for monetary refinancing operations with the Eurosystem. Instead of using the Logit curve to compute conditional probabilities—which is depicted as the dash-dot curve in the graph—we adopt a semiparametric approach and fit a smoothing spline to this set of points. Additional robustness exercises were performed but are not reported here in terms of the parameters of smoothing spline.

Comparing the semiparametric curve with the Logit curve in Figure 2, we see that for the lowest estimated default probabilities for which we have data in the sample the smoothing spline is more conservative in terms of credit class classification, while over the mid-range of z-scores the Logit is slightly more conservative. For higher estimated default rates, the two curves

are equivalent, and for the highest estimated default probabilities the Logit is again more conservative than the smoothing spline.

The strategy followed here will be to use the intersections of the smoothing spline with the upper limits of the credit classes as classification thresholds in terms of z-scores.<sup>7</sup> These values can be observed in Figure 3, where we also depict the upper value of the probability within the class.

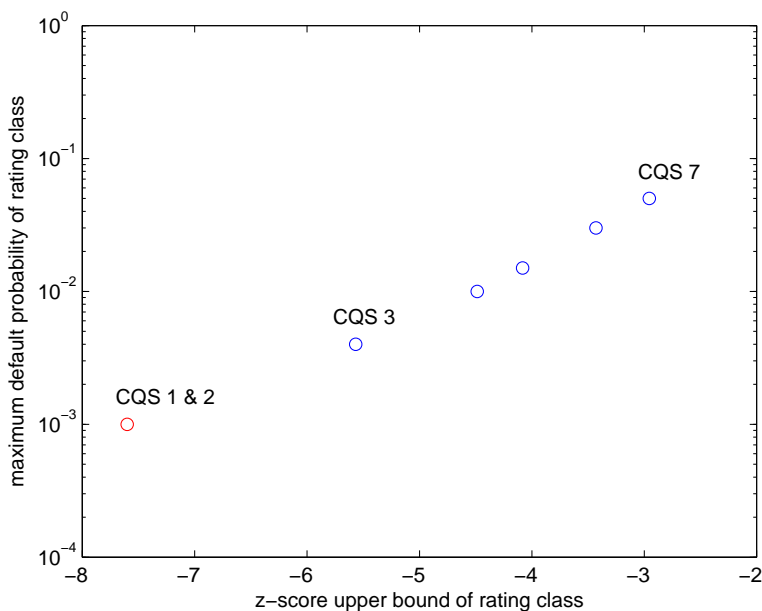


FIGURE 3: Thresholds in terms of z-scores defined according to the text.

Source: ECB, Banco de Portugal and authors' calculations.

Two observations are important at this point. First, it is clear that even with this strategy a post-classification evaluation of the method is warranted. This is because the thresholds define classes in terms of z-scores but if the observed default rates are too noisy they will have no discrimination power relative to adjacent classes. The fact that the dots represent a relatively smooth function of the probability of default with respect to the z-score gives us confidence about the capacity of the classification method to produce reasonable results.

Second, it is not possible to classify firms with credit rating classes with default probabilities below a certain value, that is, above a certain credit rating. The reason for this is the scarcity of observations classified in lower risk classes. For example, the upper limit of the default probability admissible for a

7. For class 1 & 2, the intersection was extrapolated. More on this below.



firm with a Credit Quality Step 1 would be<sup>8</sup> about 0.03% during one year. This means that we need approximately 67 thousand observations classified with that rating to expect observing 20 defaults.<sup>9</sup> If we cannot classify this number of firms with such rating in our sample, we also cannot be sure that those firms really have a probability of default compatible with the step 1 rating. Even if we are willing to lower the number of expected default events to, say, 5, we still need 17 thousand observations. In practice, for our data set we found that thresholds up to class 2 are possible: this is one class above the highest credit class for which it is possible to consistently estimate default rates. This point can be made by noting that, using the notation previously introduced,  $\hat{d}_1^{23} = \frac{23}{11486} = 0.002$ , that is, the first 23 defaults occur for the best 11,486 z-scores. This default rate is significantly lower than the upper limit of credit class 3, and above the upper limit of credit class 2.<sup>10</sup> Using the fitted curve of Figure 2 to extrapolate one class above (in terms of rating) class 3 seems reasonable. For this reason we lumped Credit Quality Steps 1 and 2 into the class labeled “1 & 2”. In Figure 4 we have depicted observed default rates for each class using the thresholds shown in Figure 3. Also represented are the upper default probability limits of each credit class. Since we are using a conservative approach in defining the thresholds, we see that, for all classes except class 1 & 2, the observed default rates are lower than the upper limit of each class. Moreover, assuming within-class binomial distribution<sup>11</sup> the lower bound of the 90% confidence interval of the default rate lies above the upper limit of the class immediately to its left (that is, with better credit quality) and the upper bound lies below the upper limit of the class.

### *Classes with few observations*

Class 1 & 2 merits a special reference. Out of a sample of more than 740 thousand firm-year observations spanning the period 2005–2014, the above methodology allows us to classify 1177 observations in class 1 & 2. Out of these observations only two were defaults. This means that the statistical significance of the empirical default rate is low: one more or one less default would change considerably the observed default rate of the class. In Figure 4, this can be seen by the wide 90% confidence interval, whose lower limit is 0 and higher limit is 0.35%, assuming a binomial distribution of defaults within

---

8. This would be roughly equivalent to ratings of AA- and above (Fitch and Standard & Poors) or Aa3 and above (Moody's).

9. That is,  $20 \times \frac{1}{0.0003} \approx 67,000$  observations.

10. Assuming a binomial distribution, the lower and upper limits of the 90% confidence interval of  $\hat{d}_1^{23}$  are 0.13% and 0.27%, respectively.

11. Under the binomial distribution, the observed default rate of a given class is the maximum likelihood estimator of the default rate.

the class. This also means that we do not reject the null hypothesis that, under a binomial distribution, the actual probability of default is lower than 0.1%.

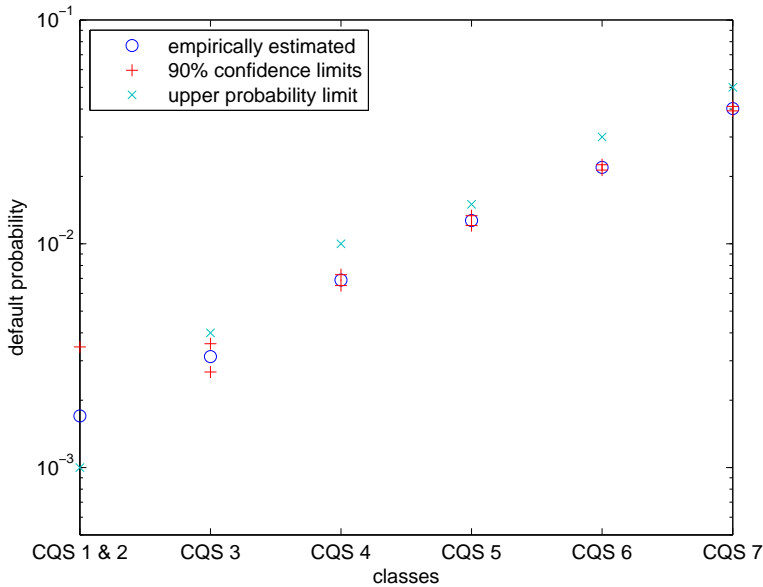


FIGURE 4: Observed default probabilities across classes using the thresholds in terms of z-scores defined according to the text. Confidence intervals are estimated assuming that within each class the default event follows a binomial distribution. Upper limits for default probabilities of each Credit Quality Step as defined by the Eurosystem also depicted as dashed horizontal lines.

Source: ECB, Banco de Portugal and authors' calculations.

All in all, one would assume that the model should be able to reliably distinguish firms in terms of all credit categories, with the best class being a residual class that lumps all high credit quality observations. The discriminating power of the model is limited by the number of observations in each class; we deem it reasonable to classify firms up to class 2. In the next section we perform an analysis of transitions of firms across classes and to default.

## Some results

We now present some of the results of the rating system applied to our data. The results are consistent with the observation from Figure 2 that the z-scores seem to be effective in distinguishing firms in terms of their propensity to default.

### *Credit risk dynamics*

Transition tables are a useful way to characterise the dynamics of firms across rating classes and to default. These tables typically contain the probability of moving to a specific credit rating class or to default, conditional on the current rating class. Table 6 contains some general statistics of our sample, including the observed default rates conditional on rating class and also exits from the sample.

Overall, we see that the default rates across classes vary considerably but are close to both their model-predicted values and the upper limit of the respective class, as seen in Figure 4. Class 8 is the most prevalent, while unsurprisingly the least numerous one is class 1 & 2, which accounts for about 0.16% of the sample. Applying the Spiegelhalter (1986) test within each class allows us not to reject (with the exception of class 8) the null that all model-estimated default forecasts match the true but unknown probability of default of the firm.<sup>12</sup>

As for exits without default from the sample, values vary between 11% and 18%, with an overall mean of 13.8%. These transitions are defined as permanent exits from the sample due to any of the following situations, all of them without any registered default: (i) exit from activity by merger, acquisition or formal extinction; (ii) the firm's loans are fully amortised; (iii) at least one of the regressors selected in the Logit model is not reported by the firm. Defaults can always be detected even if the firm ceases to report to CB because banks still have to report any non-performing loans by legally existing firms. These numbers compare favourably with similar measures found in the literature. For example, Figlewski *et al.* (2012) reports that, out of a sample of about 13,000 observations, the withdrawal rate was 33%.

Over time, the model-estimated default probabilities follow reasonably well the observed default rates. A notable exception is 2009, when observed default rates were considerably higher than what the respective credit risk class would suggest. This was a widespread phenomenon. See, for example, Chart 14 in Vazza and Kraemer (2015). In Table 7 this can be assessed by the differences in observed default rates in year  $t$  and the predicted default rates in year  $t - 1$  for year  $t$ . We see that most of the variation is due to the highest risk class, where the construction and real estate industry and the micro firms are over-represented (see Table 9 below).

Table 8 reports the overall transition matrix, which contains the share of firms migrating from one risk class to another in the subsequent year, conditional on non default and non exit. The table shows that in 3 out of 7 classes the majority of firms remained in the same risk class. It is also seen that

---

12. For class 8 we indeed reject the null at 5% significance. The average model-estimated default rate is 10.0% while the observed value is 10.3%. See Table 6.

the large majority of firms either stayed in the same category or moved only one category up or down. In addition, notice that, conditional on non default and non exit, firms were more likely to be downgraded than to be upgraded, except class 8 for obvious reasons.

The Markovian structure of the matrix allows us to compute a long-run distribution across credit classes (called the “ergodic” distribution). This would be the distribution prevailing in a year in the distant future if the rate at which firms entered and left the data set were those observed in the sample. It turns out that such distribution is remarkably similar to the actual shares of firms observed in Table 4. This suggests that the sample is a reasonable representation of the long-run dynamics of firms across credit rating classes.

One thing that is important to note is the relatively low persistence of credit class categories that emerges with this tool. The average persistence of a firm in the same class is much smaller than the persistence observed by ratings from rating agencies. For example, Vazza and Kraemer (2015) document that, out of 7 credit risk categories, the average fraction of firms staying in the same credit category is 87%; the comparable number in our sample is 45%. There are at least two reasons for this.

First, rating agencies typically produce ratings for relatively large corporations that have strong incentives to be rated, while in our case all firms are *ex ante* included in the sample. Moreover, several strategic considerations could bias the persistence values. While typically credit rating agencies follow firms even when they are no longer rated to detect potential defaults, firms that are currently rated might have an incentive to withdraw the rating if they suspect they will be downgraded. The other two possibilities—rating unchanged or upgrade—do not induce such a powerful incentive. This strong selection bias of the static pools of rating agencies, while not affecting the transitions to default—as ratings are conditional on the actual balance sheet of firms—would tend to produce much more persistent ratings than a rating tool that potentially includes all firms.

Second, ratings agencies and also other rating systems (such as Banco de Portugal’s ICAS, currently applied to mostly large Portuguese corporations) typically involve dedicated analysts which have some latitude in adjusting the ratings coming from the statistical models underlying the system. This could also be a origin of more persistent ratings as the analyst would be reluctant to change the rating if, for example, the newly computed probability of default were marginally outside the range of the previous rating. No such adjustments are done here and even minor changes in the model-estimated default probabilities could entail changes in credit risk category.

Table 9 presents the model-estimated probabilities of default versus the empirical probabilities of default separately for each industry group and for each size category, as well as the share in terms of observations of each risk class in the group. When compared to the other sectors, the table shows that the construction and real estate sectors (industry 2) have a particularly high

average default probability. This result is observed both in the comparison of estimated and empirical default probabilities and in the shares of each class. Class 8 is more than twice as large as any other risk class in this specific industry group.

Relatively risky are also micro-sized firms (size 1), none of which is considered to be in class 1 & 2 while about 74% of them are concentrated in the three worst risk classes. In contrast, about 57% of larger firms (size 2) are in the three worst risk classes.

The table shows that the five industries are generally skewed to riskier classes, particularly classes 6 and 8.

### *Additional validation*

It is outside the scope of this article to present a detailed characterization of the method's performance out-of-sample and validation exercises. For a simple approach to this issue, the interested reader is reported to, for example, Wu (2008). Aussenegg *et al.* (2011) and Coppens *et al.* (2016) and references therein provide more advanced material.

### **Conclusion**

The aim of this article is to present a method to assess the creditworthiness of the Portuguese non-financial firms by estimating the probability that any given firm will have a significant default episode vis-à-vis the banking system during the following year. The outcome of the model is then mapped into a masterscale where companies are grouped into homogeneous risk classes, originating a synthetic indicator of the firm's ability to fulfill its financial commitments.

By merging balance sheet information from 2005 until 2014 with credit register information from 2002 until 2015 we were able to estimate ten different models with good explanatory power in terms of the default risk of a firm. With the exception of class 8, the model-estimated default probabilities are not statistically different from the observed default probabilities.

The results also show how firms are mostly allocated to higher risk classes, with some industries and firm size classifications not represented in the lowest risk class. As expected, micro-sized firms have, on average, estimated and observed default probability higher than larger firms. The same can be seen for the construction and real estate sectors when compared to the rest of the industry sectors.

With respect to the dynamics in the transition tables presented, we can see that, from one year to the next, most firms remain in the same risk class or move to an adjacent class. Moreover, the overall transition table also seems

to indicate that our model is a fairly good representation of the long-run risk distribution of the Portuguese non-financial sector.

Finally, it should be stressed that the available data do not allow us to classify firms beyond a certain credit quality. This is due to the scarcity of observations for the lower risk classes. For a finer classification among high ratings it is necessary to include professional analysts in the process and, perhaps, resort to more structural models of default as opposed to statistical approaches like the one followed here.

## References

- Aussenegg, Wolfgang, Florian Resch, and Gerhard Winkler (2011). "Pitfalls and remedies in testing the calibration quality of rating systems." *Journal of Banking and Finance*, 35(3), 698–708.
- Coppens, François, Fernando González, and Gerhard Winkler (2007). "The performance of credit rating systems in the assessment of collateral used in Eurosystem monetary policy operations." Working Paper 118, National Bank of Belgium.
- Coppens, François, Manuel Mayer, Laurent Millischer, Florian Resch, Stephan Sauer, and Klaas Schulze (2016). "Advances in multivariate back-testing for credit risk underestimation." Working Paper 1885, European Central Bank.
- ECB (2015). "Questions and answers on the compilation of the static pool for performance monitoring in the Eurosystem Credit Assessment Framework (ECAAF)." Available online.
- Figlewski, Stephen, Halina Frydman, and Weijian Liang (2012). "Modeling the effect of macroeconomic factors on corporate default and credit rating transitions." *International Review of Economics and Finance*, 21(1), 87–105.
- Imbens, Guido W. and Donald B. Rubin (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.
- Lingo, Manuel and Gerhard Winkler (2008). "Discriminatory power: an obsolete validation criterion?" *Journal of Risk Model Validation*, 2(1), 45–72.
- Martinho, Ricardo and António Antunes (2012). "A Scoring Model For Portuguese Non-Financial Enterprises." In *Financial Stability Report*, pp. 111–124. Banco de Portugal.
- Plosser, Matthew C. and João Santos (2014). "Banks' incentives and the quality of internal risk models." Staff Report 704, Federal Reserve Bank of New York.
- Santos, João and Andrew Winton (2015). "Bank capital, borrower power, and loan rates." Mimeo.
- Spiegelhalter, David J. (1986). "Probabilistic prediction in patient management and clinical trials." *Statistics in Medicine*, 5, 421–433.
- Vazza, Diane and Nick W. Kraemer (2015). "2014 Annual Global Corporate Default Study And Rating Transitions." Ratings direct, Standard & Poor's

Rating Services.

Wu, Xuezhen (2008). *Credit Scoring Model Validation*. Master's thesis, University of Amsterdam, Korteweg-de Vries Institute for Mathematics.

CQS	Withdrawn	Default rate		Upper limit	Share of total sample
		Observed	Estimated		
1 & 2	16.4	0.17	0.10	0.10	0.16
3	11.1	0.31	0.28	0.40	5.5
4	11.6	0.69	0.69	1.00	16.7
5	11.8	1.27	1.24	1.50	11.1
6	12.4	2.20	2.17	3.00	21.8
7	13.1	4.02	3.91	5.00	16.0
8	17.6	10.3	10.00	100	28.8
Full sample	13.8	4.36	4.25	<i>n.a.</i>	100

TABLE 6. Observed and model-estimated default rates and rate of exits from the sample without default, by rating class. Model default rates estimated using the semiparametric methodology. All values in percentage. Model-estimated default rate for CQS 1 & 2 set to the upper limit of the class.

Source: Banco de Portugal and authors' calculations.

CQS	Default rate	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Total
1 & 2	Estimated	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
	Observed	0.00	1.75	0.00	0.00	0.72	0.00	0.00	0.00	0.00	0.00	0.17
3	Estimated	0.29	0.28	0.29	0.28	0.28	0.28	0.28	0.28	0.28	0.28	0.28
	Observed	0.16	0.30	0.33	0.78	0.17	0.27	0.42	0.22	0.26	0.39	0.31
4	Estimated	0.70	0.70	0.70	0.70	0.69	0.69	0.69	0.69	0.69	0.69	0.69
	Observed	0.48	0.51	0.64	0.87	0.42	0.77	1.13	0.77	0.70	0.46	0.69
5	Estimated	1.24	1.24	1.24	1.24	1.24	1.24	1.24	1.24	1.24	1.24	1.24
	Observed	0.82	1.00	1.46	1.82	1.05	1.59	1.89	1.34	1.02	0.66	1.27
6	Estimated	2.17	2.17	2.18	2.18	2.18	2.17	2.17	2.17	2.16	2.16	2.17
	Observed	1.35	1.84	2.41	3.33	1.70	2.54	3.40	2.21	1.68	1.42	2.20
7	Estimated	3.90	3.90	3.91	3.91	3.91	3.91	3.90	3.92	3.90	3.89	3.91
	Observed	2.61	3.56	4.64	6.09	2.99	4.51	5.86	3.99	3.30	2.35	4.02
8	Estimated	9.06	9.20	9.32	9.52	10.30	10.25	10.62	10.95	10.20	9.78	10.04
	Observed	6.57	7.99	10.43	14.44	8.09	11.00	15.29	11.32	8.59	6.42	10.31
Total	Estimated	3.77	3.91	4.03	4.30	4.64	4.25	4.59	4.82	4.13	3.75	4.25
	Observed	2.63	3.40	4.54	6.53	3.62	4.68	6.74	4.98	3.47	2.43	4.36

TABLE 7. Observed and model-estimated default rates over time, by rating class. Model default rates estimated using the semiparametric methodology. All values in percentage. Model-estimated default rate for CQS 1 & 2 set to the upper limit of the class.

Source: Banco de Portugal and authors' calculations.



CQS in year t	CQS in year t+1						
	1 & 2	3	4	5	6	7	8
1 & 2	36.5	<b>55.9</b>	5.9	0.7	0.8		0.1
3	1.5	<b>56.5</b>	32.0	4.5	3.6	1.1	0.8
4	0.0	10.7	<b>51.3</b>	17.3	13.7	4.1	2.8
5	0.0	2.0	25.8	26.1	<b>30.6</b>	9.3	6.2
6	0.0	0.8	9.4	14.4	<b>40.2</b>	20.5	14.7
7		0.3	3.5	5.3	24.6	31.8	<b>34.4</b>
8		0.1	1.4	2.2	9.1	16.0	<b>71.2</b>

TABLE 8. Transition matrix between credit rating classes, conditional on firms being in the sample in two consecutive years and not defaulting. Rows add up to 100 percent. All values in percentage.

Source: Banco de Portugal and authors' calculations.

CQS	Statistic	Industry					Size		Total
		1	2	3	4	5	1	2	
1 & 2	Estimated def. rate	0.10		0.10	0.10	0.10		0.10	0.10
	Observed def. rate	0.00		0.18	0.00	0.00		0.17	0.17
	Share of obs.	0.02		0.40	0.70	0.00		0.36	0.16
3	Estimated def. rate	0.29	0.34	0.27	0.26	0.31	0.33	0.27	0.28
	Observed def. rate	0.40	1.38	0.30	0.00	0.19	0.29	0.32	0.31
	Share of obs.	5.89	0.45	8.61	12.56	3.56	1.33	10.79	5.52
4	Estimated def. rate	0.69	0.74	0.68	0.68	0.70	0.72	0.67	0.69
	Observed def. rate	0.68	0.94	0.75	0.79	0.56	0.70	0.68	0.69
	Share of obs.	17.45	6.48	19.12	19.33	18.41	13.10	21.20	16.69
5	Estimated def. rate	1.24	1.25	1.24	1.24	1.24	1.24	1.24	1.24
	Observed def. rate	1.44	1.45	1.25	0.56	1.14	1.24	1.31	1.27
	Share of obs.	10.81	7.72	11.44	10.75	12.72	11.24	10.88	11.08
6	Estimated def. rate	2.17	2.22	2.16	2.16	2.16	2.18	2.16	2.17
	Observed def. rate	2.24	2.25	2.10	2.22	2.26	2.21	2.17	2.20
	Share of obs.	21.02	21.73	21.19	18.45	23.39	24.15	18.83	21.79
7	Estimated def. rate	3.91	3.94	3.89	3.91	3.89	3.91	3.90	3.91
	Observed def. rate	3.89	3.76	3.98	5.28	4.32	4.11	3.86	4.02
	Share of obs.	15.52	20.40	14.67	12.65	15.86	18.54	12.82	16.00
8	Estimated def. rate	10.15	10.47	10.12	9.83	9.45	9.54	10.83	10.00
	Observed def. rate	10.37	10.80	10.47	9.75	9.59	9.71	11.26	10.31
	Share of obs.	29.29	43.22	24.56	25.55	26.06	31.64	25.12	28.75
Total	Estimated def. rate	4.30	5.96	3.81	3.70	3.88	4.51	3.93	4.24
	Observed def. rate	4.41	6.10	3.91	3.78	3.97	4.60	4.05	4.36
	Share of obs.	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

TABLE 9. Model-estimated and observed default rate for selected groups of firms. Model default rates estimated using the semiparametric methodology. All values in percentage. Model-estimated default rate for CQS 1 & 2 set to the upper limit of the class.

Source: Banco de Portugal and authors' calculations.