Estimating a financial distress rating system for Spanish firms with a simple hazard model

Christian E. Castro[1]

This draft 11th June, 2008

Abstract

Under the Basel II Internal Rating Based (IRB) approach banks should accurately discriminate among different grades of obligors in their credit portfolios taking into account not only the obligor-specific characteristics but also their sector and macro-economic environment. The final objective is the assignment of a credit rating to each of their exposures based upon a set of estimated probabilities of default. With this purpose in mind, the industry has been gradually moving to more advanced techniques in financial distress prediction, with survival models playing a prominent role in the last few years.

In this paper we apply a parametric proportional hazard model in its discrete version in order to predict probabilities of financial distress (PFDs) on a large dataset of non-listed private Spanish firms during 1994 and 2005. We include four financial dimensions in the analysis of the firm, jointly with controls by sector and size. In addition we include two common factors in order to study the possible effect of fluctuations in the macroeconomic environment. Next, we study the discriminatory power of the prediction model and we apply a set of rating techniques in order to calibrate a through-the-cycle (TTC) rating system. Finally, we examine the stability of the rating system across each year of our sample.

The period between 1994 and 2005 has been characterized by a notable expansion of the Spanish economy, with all along positive GDP growth rates and declining interest rates. Even when it has been a long expansionary phase, we find that some macroeconomic fluctuations during this time have generated interesting effects on the estimated firms' PFDs. These effects can be better understood if the firm's age is considered into the analysis. Additionally, we make some warnings about the construction of rating systems exclusively based on data from this particular period.

[1] I gratefully acknowledge the advice and useful comments from Prof. Vicente Salas y Fumás in this work. Of course, all remaining errors are my own.

christian.castro@uab.es

1. Introduction and main contributions

Basel II has clearly represented an improvement respect it already famous predecessor: the 1988 Basel Accord (Basel I). Basel I was a so successful financial standard that is has been explicitly adopted in more than 100 countries and has been widely recognized as a significant advance in establishing an internationally recognized language to analyze and compare capital across different jurisdictions (summarized in the concept of assets at risk) and in attempting to establish a leveled playing field for international bank competition (Balzarotti *et al*, 2004). Nevertheless, Basel I was too little risk-sensitive and it did not provide meaningful measures of risk for all the involved stakeholders in the financial activity, therefore it faced many critics. In particular, as banks' own models of credit risk measurement have become more sophisticated, this has driven a wedge between the concepts of 'regulatory' and 'economic' capital. The regulatory response to this growing wedge has been the set of new proposals embodied in Basel II that attempts to better align capital requirements and the way banks manage their actual risk.

The Basel II framework is fully compiled in the 2006 document "International Convergence of Capital Measurement and Capital Standards". The framework basically consists of 3 Pillars: (1) Minimum Capital Requirements, (2) Supervisory Review Process and (3) Market Discipline. Under the First Pillar, banks that have received supervisory approval to use the Internal Rating Based (IRB) approach may rely on their own internal estimations of a set of risk components to determine the capital requirement for a given exposure. This implies that banks should accurately assess the different risk profiles of each obligor in their credit portfolios. The banks portfolios are categorized in four classes of assets: (a) corporate, (b) sovereign, (c) bank, (d) retail, and (e) equity. Within the corporate asset, there are five sub-classes categorized as specialized lending (project finance, object finance, commodities finance, income-producing real estate, and high-volatility commercial real estate). In the retail case there are three sub-classes (exposures secured by residential properties, qualifying revolving retail exposures and all other retail exposures)[2]. In the present work we want to concentrate in the study of the corporate sector.

[2] In corporate and retail classes, purchased receivables may have a distinct treatment in certain conditions.

Banks under the IRB may apply different techniques and methodologies in order to assign a given credit rating to each of their exposures and gradually discriminate between different types of borrowers according to their risk. These credit ratings will be based on a set of estimated probabilities of default (PDs)[3]. PDs can be obtained by a variety of approaches in the spectrum of subjective to objective measures. Basel II argues in favor of more objective measures, essentially founded in econometric scoring type models. Besides, it also recognizes that "human judgment and human oversight is necessary to ensure that all relevant and material information, including that which is outside the scope of the model, is also taken into consideration, and that the model is used appropriately" (Basel Committee, 2005). Increasingly, for new clients, banks are moving towards more objective scoring techniques often based on a scorecard of particularly important variables with weights determined by an underlying econometric scoring modeling exercise. It is also desirable that the estimated PDs consider the effects of possible fluctuations in the overall conditions over different moments of the economic cycle. In general, given the difficulties in forecasting future events and the influence they will have on a particular borrower's financial condition, banks must also take a conservative view of the projected information.

The estimation of the PDs and the construction, validation and back-testing of the internal credit ratings are crucial issues in Basel II given that the regulatory capital requirement is set as a function of the PD in the Foundational IRB, and PD plus the rest of risk inputs in the Advanced version. In what respect to the calculation of these internal risk inputs the Basel Committee establishes a series of minimum requirements:

"391. The minimum requirements set out in this document apply to both foundation and advanced approaches unless noted otherwise. Generally, ***all IRB banks must produce their own estimates of PD and must adhere to the overall requirements for rating system design, operations, controls, and corporate governance, as well as the requisite requirements for estimation and validation of PD measures****. Banks wishing to use their own estimates of LGD and EAD must also meet the incremental minimum requirements for these risk factors included in paragraphs 468 to 489."*

(Basel Committee, 2006). Highlighted in bold was added to the original text.

[3] The risk components in the Foundational IRB only include the own measure of the probability of default (PD). In the Advanced IRB it also includes measures of the loss given default (LGD), the exposure at default (EAD), and effective maturity (M). For a description of the main approaches in the IRB approach, see Powell (2004).

Along these lines we find the main motivation for this paper: the application of a consistent and flexible methodology for the estimation of probabilities of financial distress (PFDs)[4] in a large sample of non-listed Spanish firms, jointly with the consideration of possible macro-economic effects during a particular period of the time in the Spanish history, and the final construction and validation of a financial distress rating system. The proposed econometric model and the different steps we follow in this paper can be transferred to the estimation and construction of credit ratings by financial institutions. Up to the best of our knowledge it is the first study that using a non-credit bureau database on Spanish firms applies a hazard model for the construction of a risk rating system following the set of requirements and techniques suggested by the Basel Committee.

We use a commercial database containing balance sheet information for Spanish corporate firms over the 1994-2005 period in the estimation of firms PDs and posterior construction of a risk rating system. In addition to firm-specific variables, sectorial and size controls, we also consider the effect of fluctuations in the macroeconomic environment. We surprisingly find that an increase in the real GDP growth rate can generate an increase in the frequency of defaults and estimated PDs. Nevertheless, this effect depends on the age of the firm, being especially important in the case of "young" firms. On its side, the level of interest rate level seem to be not a restriction in the early years of creation of the firms, but its effect becomes significant in the case of mature firms. A second contribution of our work is that it helps to understand the macroeconomic circumstances that may increase the risk of failure when firms are young. Systematic differences in the determinants of firm failure between firms that fail early in their life and those that fail after having successfully negotiated the early liabilities of newness and adolescence has been identified by several papers in the Resource-based view literature (see for example Thornhill and Amit, 2003). The "liability of the adolescence" or "honeymoon effect" feature has been also reported in the Spanish case by Lopez-Garcia and Puente (2006), but they do not explore the possible effect of the macroeconomic environment. We think that the study of these effects is particularly important in the explanation of the availability of resources during a long and lasting expansionary period.

[4] Even when our definition of PFDs not necessary imply a default by the firm, paragraph 453 in Basel II indicates that the circumstances we are considering as a trigger of financial distress are indications of the unlikeliness to pay and therefore it would imply a boost to the PD (though not necessarily a default itself). For this reason and to facilitate the exposition in what follows we will generally refer as PD to the object of our estimation.

The last paragraph is also related with the fact we find a better fit in the prediction model using a quadratic duration term. This parametric specification is also consistent with the findings of Lopez-Garcia and Puente (2006) where they report an inverted-U shape curve in the relation between firms' exit and its age. As these authors point out, this type of duration dependence has also been found in other countries such as the United States, United Kingdom, Italy or Germany. It responds to what it has been known as the "liability of the adolescence" or the "honeymoon effect" brought about by stock of the initial resources of the new firm. Those resources help the new firm go through the first years even if the firm results to be inefficient. Once the initial stock of resources is used up, if the firm is inefficient it will exit the market. The fact that we find that an increase in the GDP growth has a significant and positive effect in the PDs of young firms but it becomes less important in the case of mature firms and that at the same time, fluctuations in the interest rates seem not to be affecting young firms[5], would be giving some support to those lines of reasoning behind the "honeymoon effect" or "liability of adolescence".

A third contribution of our paper is that as opposite to other studies that concentrate on the difficulties of building a rating system in countries with large macroeconomic fluctuations, we study the main challenges in its construction when we use information exclusively based on a long expansionary period. In this sense, the Spanish economy has experienced a steady positive growth rate since 1994, not without fluctuations, but definitely consolidated as one of the fastest economic growth among the EU-15 economies. Under this scenario we build a quite stable rating system, but we acknowledge that we have to be careful in this case. As it was mentioned, in a clear contrast with other cases where an stable through-the-cycle (TCT)[6] rating can not be obtained given the existence of large economic fluctuations (see for example the case of Argentina between 2000 and 2005 in Valles, 2006), in our case we are building a rating based on a long expansionary period, which goes from 1994 up to the 2008 approximately. Even when ratings based on this time period would be probably fulfilling with most of the Basel Committee requirements, the particular case of the Spanish economy in the last 12 years makes necessary the adjustment of the resulting TCT in order to reflect potential fluctuations in the macro-economic fundamentals. Of course, including data from a contractionary period (the 1990-1994, for example) would be highly

[5] The interest rate is close to be significant at a 10% of confidence level of mature firms when all the financial ratios are included, and it is significant at the 5% level when only include the two macro-economic variables. In both cases it has a positive sign.
[6] We briefly explain the characteristics of a TTC rating in section 6.

desirable. Yet, it should be also noticed that even in this case it would be a quite hard task given the substantial growth of the Spanish financial systems, its production levels and number of firms in the last 20 years. In some way the described situation can be interpreted as an opposite challenge to the design of credit ratings in countries with large economic fluctuations. But maybe the lack of well stressed PDs under different potential bad scenarios could be even more dangerous and challenging in countries with long periods of uninterrupted expansion, than in other cases.

At the whole, we think that our database reflects some of the main characteristics of credit databases based on posterior years to 1992-1993. Probably these databases will count with relatively few defaults observations in relation with the total number of healthy firms. This situation increase the difficulty of discriminate between "good" and "bad" obligors[7]. Despite these difficulties, we have found the prediction model we use to estimate the PDs and the posterior risk rating we build to be very satisfactory for the normal range of practices in the field.

It is also worthy to notice that our sample is not biased by selection as it happens with those studies exclusively based on obligors' information from Credit Bureau and whom it has been already granted access to the banking lending. For this reason we think that it could serve as an helpful complement for those studies.

This paper is organized as follows: in Section 2 we summarize and describe the main minimum requirement established by the Basel Committee for the estimation of PDs, and the construction and validation of internal rating systems. We also comment some common practices in the field. In Section 3 we review some of the most relevant papers and techniques used in the prediction of bankruptcy/default by firms. We specially remark some specific studies closely related with our analysis. In Section 4 we introduce the methodology to estimate a Parametric Proportional Hazard model[8] in discrete time. In section 5 we describe the characteristics of our database and the construction sample that we use in our estimations. Next, in Section 6 we estimate conditional PDs for the whole sample of year-firms. Also in this section we comment some features of the baseline model we use in the PDs estimation and we analyze the effect of macroeconomic fluctuations. In

[7] In addition, estimations based on this long expansionary period most likely will not generate well stressed PDs
[8] One of the better known Proportional Hazard Model (PHM) is the Cox model.

Section 7 we build a rating system using cluster analysis on the estimated PDs and frequency of distresses. Then, we validate the rating system applying conventional tests of discrimination power and calibration. Finally, we extract some conclusions; we comment possible extensions and future work.

2. Building a rating system "*a la*" Basel II and beyond

2.1. IRB minimum requirements

The IRB framework establishes a set of minimum criteria to be followed by banking entities and supervised by the corresponding regulators. In our case we want to concentrate in those requirements explicitly related with PDs estimation process and posterior construction and validation of the resulting credit ratings. In appendix 1 we present a full summary of the Basel II minimum criteria grouped into four main issues:

1) General guidelines.
2) Estimation of PDs.
3) Credit rating construction.
4) Validation and tests.

In what respect to point 2, estimation of PDs, we propose a prediction model based on survival analysis. The use of hazard models in bankruptcy/default prediction permits to apply all the standards techniques presented in others estimation procedures such as discriminant analysis and logit/probit models. There is evidence that hazard model outperforms static models (of the logit/probit/discriminant analysis type) not only in the efficiency their estimations but also in others aspects. For example, a hazard model allows the analysis of multiple influences over time, avoiding the usual bias in the one-year-ahead estimation models. Several works have highlighted these types of inconsistencies inherited in static models (see for example, Duffie *et al*. 2007, Das *et al.* 2007, Shumway 2001, and Kiefer 1998). For the particular case we want to study, there is a relatively simple implementation methodology in discrete-time (Jensen, 2005) which is, from our point of view, a much more transparent, stable and flexible methodology than other dynamic techniques based on dynamic panel models with categorical dependent variables. Even when it is a relatively simple methodology, the hazard model takes advantage of the whole time span of the dataset, what makes it a very valuable strategy in the estimation and validation of PDs through the economic cycle.

In addition, all the estimations in the Basell II framework must represent a conservative view given the likely of unpredictable errors (par. 451). That means that once calibrated the rating system, it is possible to introduce extra adjustments based on the results of stress-testing exercises. The list of requirements in the IRB also established that the estimated PD has to take appropriate account of the long-run experience in each grade and that the length of the underlying historical information must be at least of five years[9] (pars. 447, 448, 461, 462, 463). On this point, the database we are using goes from 1994 up to 2005, clearly fulfilling with the minimum history requisite. It also proposed a default definition (par. 452) and a set of indicators of the unlikeliness to pay (par. 453). As commented before, our trigger for financial distress can be included in this last group of indicators.

In order to build a rating based upon the estimated PDs (point 2) we apply cluster analysis techniques. We build a seven grades rating[10] system in line with the normal practice of two major rating agencies, S&P and Moody's, and as it is also required in the New Accord (par. 404). To each grade we associate the average PDs of all the year-firms in each bucket (par. 462). More specifically, we perform the k-mean cluster analysis jointly in the estimated PDs and the frequency of year-firms in order to avoid over-concentrations in specific grades (par. 403) and to make our results directly scalable to different sizes in the random sample of healthy firms we take. The fact that we are including sectorial controls and macro-economic variables allows us to explore the existence of systemic effects on the PDs. However, as we are not including a clear contractionary economic period we have to be careful with our results. On this point, according to Basel II a borrower rating must represent the bank's assessment of the borrower's ability and willingness to contractually perform despite adverse economic conditions or the occurrence of

[9] Internal and external sources can be combined, but at least one must be at least five years.
[10] There is an implicit grade 8^{th} grade, the bankruptcy stage. But as is quite common in the bankruptcy prediction literature we do not count with information beyond the time of declaration of bankruptcy or *concurso preventivo*.

unexpected events. The range of economic conditions that are considered when making assessments must be consistent with current conditions and those that are likely to occur over a business cycle within the respective industry/geographic region (par. 415). In conclusion, banks should not just rely on present estimations of the PD but should also calculate PDs in stress scenarios characterized by poor economic conditions.

Finally, in what respect to point 4, we apply a set of validation tests suggested by the Basel Committee (2004) and normally used in the field (see for example Duffie *et al* 2007, Fernandes 2005, Standard & Poor´s 2007, etc). In particular, we test the discriminatory power of the model by calculating the Accuracy Ratio (AR) from the cumulative accuracy profile (CAP) curve. After that, we analyze the rating structure, we calibrate it and we analyze its stability along each of the years of our sample. Even when we have to be careful in the comparison of AR from different samples, we find that the AR values we get are very satisfactory for the normal range of values. Moreover, we find a stable rating that is able to follow suit the observed pooled frequency of distresses.

During the estimation process and the ratings construction we notice that different groups of firms are harder to discriminate according to their risk in investment grades, than in speculative grades. In fact, there seems to be a clear discontinuity in the estimated PD in grades 1-5 respect 6-7. On this particular, Schuermann and Hanson (2004) remark that for ratings that introduce conditioning on the state of the business cycle, it is easier to distinguish adjacent PDs in recession than in expansion. They also find that speculative grades are more business cycle sensitive than investment grades, consistently with the rating agencies own view. Given the fact that we are not including a clear recessionary period, it has been harder to distinguish between the different types of investment grades in the estimation, what have introduced some inclination to a more dichotomic picture in our rating system before calibration.

3.2. Beyond Basel II

In the previous section we have summarized some of the most relevant points in the list of minimum requirements for the design and validation of internal credit ratings under the IRB framework. But despite its regulatory use in setting regulatory capital requirements, the industry has been using credit ratings as a traditional approach to credit risk assessments from many years now. Most of the industry-based credit ratings are based both on quantitative and qualitative measures. Rating systems are usually applied to non-financial corporations and special approaches are employed for banks and other financial institutions (Crouhy *et al*, 2001). Currently, the three major credit rating agencies are Standard & Poor's, Moody's and Fitch.

The subject of a credit rating might be a company issuing debt obligations; this is the case of an "issuer credit rating". For example, S&P's issuer credit rating is a current opinion of an obligor's overall financial capacity (its creditworthiness) to pay its financial obligations. This opinion focuses on the obligor's capacity and willingness to meet its financial commitments as they come to due. It does not apply to any specific financial obligation, as it does not take into account the nature and provisions of the obligation, by contrast, its standing in bankruptcy or liquidation, statutory preferences, or the legality and enforceability of the obligation. In addition, it does not take into account the creditworthiness of the guarantors, insurers, or other forms of credit enhancement on the obligation. On its side Moody's define the issuer credit ratings as opinions of the ability of entities to honor senior unsecured financial obligations and contracts. In the Moody's case, rating symbols for issuer ratings are identical to those used to indicate the credit quality of long-term obligations. In S&P issuer credit ratings can be either long term or short term. The issuer credit rating categories are counterparty ratings, corporate credit ratings, and sovereign credit ratings.

Another class of credit ratings is "issue-specific credit ratings", with the distinction between long-term and short-term credits also. In rating a specific issue, the attributes of the issuer, as well as the specific terms of the issue, the quality of the collateral and the creditworthiness of the guarantors, are taken into account.

Given the information we have available the rating we construct can be considered as an "issuer credit rating"[11].

[11] Of course we are not performing many of the steps usually followed by private agents. For a description of the whole procedures they normally follow, see Crouhy *et al* (2001)

Three levels of influence on the firm's performance are usually considered during the process of rating designing. First, at the macro or country level, a set of macro-economic variables or common factors are usually postulated. Second, industry-specific features, such as market concentration, classification by sector of activity, indicators of sector production, among others, are also used. Combined industry and macro-economic factors can be assessed to calculate the correlation between assets for the purpose of calculating portfolio effects[12] (Crouhy *et al*, 2001). Finally, the firm-level analysis is focused on variables that seem best at predicting individual defaults or bankruptcies. The election of variables is not necessary founded on theory and therefore there is a large list of commonly used financial ratios.

On the other hand, closer to the regulatory point of view, credit ratings can be point-in-time (PIT) or through-the-cycle (TTC)[13]. On this particular the Basel Committee does not provide an explicit definition but describes them in its 2005 document on validation as follows:

• A point-in-time (PIT) rating system it that one that uses all currently available obligor-specific and aggregate information to assign obligors to risk buckets. Obligors with the same PIT grade are likely to share similar unstressed PDs. An obligor's rating can be expected to change rapidly as its economic prospects change. Overall, PIT ratings will tend to fall during economic downturns and rise during economic expansions.

• A through-the-cycle (TTC) rating system is that one that uses static and dynamic obligor characteristics but tends not to adjust ratings in response to changes in macroeconomic conditions. Obligors with the same TTC grade are likely to share similar stressed PDs. An individual obligor's rating may change as its own dynamic characteristics change, but the distribution of ratings across obligors will not change significantly over the business cycle.

The PDs that incorporate stress scenarios of the business cycle are named "stressed PDs" and the PDs for a definite period of time are the "unstressed PDs". The unstressed PDs will change with

[12] The existence of correlation among exposures is an issue that can affect some of the tests we perform on the ratings. In particular, the binomial test and the Hesmer-Lemeshow test assume independence between defaults events. Even when the independence assumption is normally not appropriate for practical purposes, we have left the analysis of correlation (ex.-ante correlation and ex-post) for future work.

[13] As the Basel Committee acknowledges between point-in-time and through-the-cycle rating systems lie a range of hybrid rating systems. These systems may exhibit characteristics of both TTC and PIT rating philosophies.

economic conditions while stressed PDs will be relatively stable in economic cycles. The main idea is that stressed PDs are "cyclically neutral", they move as obligors' particular conditions change but they do not respond to changes in overall economic conditions (Valles, 2006).

PIT systems attempt to produce ratings that are responsive to changes in current business conditions while through-the-cycle systems attempt to produce ordinal rankings of obligors that remain relatively constant over the business cycle. PIT systems tend to focus on the current conditions of an obligor, while TCT systems tend to focus on an obligor's likely performance at the trough of a business cycle or during adverse business conditions (Basel Committee, 2005)

3. Literature review on prediction models

Financial distress prediction models (both default and bankruptcy prediction) provides fundamental inputs for standards structural models of default timing. For example, in the famous Black and Scholes (1973) model, a firm's conditional default probability is completely determined by its distance to default (i.e. the number of standard deviations on annual asset growth by which the asset level exceeds the firms liabilities). Many credit risk models (ex.: the KMV approach[14]) have been built upon these concepts.

Default prediction techniques have been extensively used both for credit risk and bankruptcy risk. In credit risk, the objective is the PD of individual borrowers during a given period of time. In the second case, the analysis is usually concentrated at the corporate level. Whereas some studies are centered in large and listed firms, others are focused on small and medium firms.

Duffie et al (2007) distinguish three generation of models for default/bankruptcy prediction: models based on discriminant analysis, logit/probit models and finally, duration models. In the first generation of models, two foundational studies in corporate default analysis have been Beaver (1966, 1968) and Altman (1968). The main methodology introduced by these papers is Discriminant Analysis (DA). Particularly, the pioneer work of Altman has produced one of the

[14] Crouhy *et al* (2001) provide a good description of the main approaches in portfolio credit risk models.

most famous risk measures in the field, the Altman's Z-score. A second generation of prediction models has been dominated by qualitative-response models of the logit/probit type. On this line, Ohlson (1980) propose an O-score in his year-ahead default prediction models. The list of empirical applications of this generation of models is really large, a review of them can be found in Johnsen and Melicher (1994) and in Sobehart and Keenan (2001).

The third generation of models is dominated by duration analysis and this is the type of methodology that we want to concentrate in this article. According to Shumway (2001) there are at least three reasons to prefer hazard models for bankruptcy prediction. First, it is important to recognize that firms usually enter in risk many years before they file for bankruptcy. Static models do not adjust for period at risk, but hazard models adjust for it automatically. So with hazard models is possible to avoid the inherit selection bias in static model, correcting for period at risk. Second, hazard models are able to incorporate time-varying covariates. This allows the financial data to reveal its gradual change in health. Generally, the firm's financial health has multiple influences over time, and not only in a specific and arbitrary point. Observing only the preceding year not only we lose very valuable information at different points, but it can also generate some trivial conclusions[15] and an "excessive" list of significant explanatory variables. Hazard models exploit each firm's time-series data by including annual observations as time-varying covariates (for example, the effect of macroeconomic variables can be also included). Hazard models can also account for potential duration dependence (i.e. the possibility that firm's age might be an important explanatory variable). Third, they generate more efficient out-of-sample forecasts by utilizing much more data than static models.

Beck et al (1998) also finds that single period models generally produce inconsistent coefficient estimates. In particular, they show that observations from time-series-cross-section data with a binary dependent variable are likely to violate the independence assumption of the ordinary logit/probit models.

Strictly on the empirical side, several articles show that hazard models can produce quite different statistical inferences than do static models. For example, Schumway (2001) find that half of the

[15] In general, highly risky firms are relatively easy to distinguish if we just observe the last year before bankruptcy.

accounting rations that have been used in previous models are not statistically significant. His work compares the estimation results of two traditional papers on the field, Altman (1968) and Zmiejewski (1984). On its side, Begley *et al* (1996) find that the goodness of fit in the hazard model they propose is better than its comparisons.

Chava and Jarrow (2004) using an expanded bankruptcy database for U.S. firms companies validate the superior prediction performance of Shumway´s (2001) model as opposed to Altman (1968) and Zmijewski (1984). They also find that the industry grouping has significant effects on the intercept and on the slope of coefficients in the forecasting equations. The work of these authors represents an attempt to solve a typical limitation in the application of duration analysis, the availability of data. In our case, even when it has been necessary to previously filter the data, the final database contain 1143 distresses observations which represent a clear advantage respect other comparable studies.

Pompe and Bilderbeek (2005) study the predictive power of different ratios categories during successive phases before bankruptcy and the relationship between the age of a firm and the predictability of bankruptcy. In line with our work they focused on non-listed firms, the majority of which are small and medium firms.

Hillegeist et al (2004) also use a discrete duration model in order to assess whether if the Altman's (1968) Z-score and Ohlson's (1980) O-score, effectively summarize publicly-available information about the probability of bankruptcy. They also emphasize two of the main deficiencies in single-period models: First, a sample selection bias that arises from using only one, non-randomly selected observation per bankrupt firm. And second, a failure to model time-varying changes in the underlying or baseline risk of bankruptcy that induces cross-sectional dependence in the data. Orbe et al (2001) analyze the duration of firms in bankruptcy in Chapter 11, using a flexible model that does not assume any distribution for the duration or the proportionality of the hazard functions.

Beck *et al* (1998) propose a simple estimation strategy for the analysis of time-series-cross-section data with a binary dependent. The solution is to add a series of dummy variables to the logit specification. These variables mark the number of periods (usually years) since the start of the sample period to the last observable year. The last period can correspond to the last observation available or due to an exit by financial distress "event". Then a standard statistical test on whether these dummy variables belong in the specification is a test of whether the observations are temporally independent. The addition of these dummy variables to the specification, if the test indicates they are needed, corrects for temporally dependent observations

In what respect to studies on Spanish firms, there are several papers that estimate default/bankruptcies probabilities for a wide variety of samples and purposes. Andreev (2007) presents a survey of the main empirical works in this field for the Spanish case. Closely related with the regulatory purposes, Corcostegui *et al* (2003) use a logit model to relate the default probability and a set of financial ratios in their analysis about the possible pro-cyclical effects of a credit rating based system. Benito *et al* (2004) construct a firm-level estimate of PD for a large sample of Spanish firms. They use a year-ahead probit model and find a broad set of significant financial indicators, maybe indicating the presence of some of the common failures in one-year-ahead models that we commented before. In what respect to the use of internal models in Basel II, García Baena *et al* (2005) highlight the importance of introducing qualitative judgements in the capital calculations. They recognize the difficulties inherited in the estimation of risk inputs when few defaults are observed and they analyze the impact of the economic cycle and correlation between PDs and LGDs in the construction of credit ratings. In relation with the estimation strategy we are proposing, Lopez-Garcia and Puente (2006) study the determinants of new firm survival using a discrete-time duration model analogous to one we apply. Our parametric specification for the baseline hazard function is consistent with their findings and further international empirical evidence.

4. The estimation strategy

Duration analysis has its origins in what is typically called survival analysis, where the duration of interest is survival time of a subject (Wooldridge, 2002). The concept of conditional probabilities is a central point in the methodology. Duration analysis will be useful to answer the following type of questions: What is the probability that a firm goes bankrupt next year given that it has already survived in the market for x number of years?[16] This probability will be the result of compute a sequence of events (i.e.: the probability of going bankrupt the first year, then the probability of going bankrupt the second year given the fact that firm has survived the first year, and so on). As Kiefer (1988) points out, conditional and unconditional probabilities are related, so the mathematical description of the process is the same in either case. But the conceptual difference is important in the economic modelling of duration data. The economic literature on duration has made use of a big number of statistical methods largely developed in industrial engineering and from biomedical sciences. In general, in social sciences the objective of analysis is an individual (ex: a borrower, a family, a firm etc) that begins in an initial state and is either observed to exit the state or is censored. In the specific case of economics, we can find a wide range of applications, which goes from the typical study of the lengths of spells of unemployment up to the study of the time in office of a given political position or a time to commit a crime after release, etc.

Given the common practices in the field, the previous theoretical and empirical literature on the subject, and the particular characteristics of the data, we have decided to use a Parametric Proportional Hazard (PPH) model in discrete time. Proportional Hazard models are also known as 'multiplicative hazard' models[17].

These kinds of models are characterised by a separability assumption:

$$\theta(t, X) = \theta_0(t) \exp(\beta' X) = \theta_0(t)\lambda$$

Where $\theta_0(t)$ is the 'baseline hazard function', which depends of the spell duration (t), but not from X[18]. It summarizes the pattern of "duration dependence", assumed to be common to all persons. $\lambda = \exp(\beta' X)$ is the firm-specific non-negative function of covariate X which scales the baseline

[16] On the other hand, the corresponding question to an unconditional probability would be: What is the probability that a firm goes bankrupt exactly in the 10th year?

[17] Jenkins (2005) presents an exhaustive and clear treatment of Survival Analysis models. He also provides helpful notes for the implementation of the models in STATA.

[18] The Accelerated Failure Time (AFT) model assumes a linear relationship between the log of (latent) survival time T and characteristics X.

hazard function to all people. The PH property implies that the absolute difference in X generates proportional differences in the hazard at each period t.

The specification we propose is based on the assumption that a firms goes bankruptcy (or also in our case the firm calls to a *concurso preventivo*) at time T, where this point in time is assumed to be a continuous random variable. In the case of annual data, the data is grouped into discrete time intervals (years), it is commonly assumed that the survival times occurred in any given month in which they have been grouped. We introduce time-varying and fixed explanatory variables describing the characteristics of the firm and the sectorial and macro-economic environment of the observation units.

Then, the probability that a firm goes bankruptcy in some period later than time t can be resumed by the following cumulative distribution function of the random variable T, called the Survival function:

Survivor function: $\Pr(T > t) = 1 - F(t) = S(t)$

Analogously, the Failure function at time T can be defined as:

Failure function: $\Pr(T \leq t) = F(t)$.

Thus, the density function at the time of bankruptcy can be expressed as:

$$f(t) = \frac{\partial F(t)}{t} = -\frac{\partial S(t)}{t} \geq 0$$

Using the distribution and density function we can define a core concept in survival analysis, the hazard rate:

Hazard rate: $\theta(\mathrm{t}) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} \quad \geq 0$

The Hazard rate shows the probability that a firm which has survived until time t will go bankruptcy in the next period[19]. Subsequently, the Parametric Proportional hazard model with time-variable covariates is of the form:

$$\theta(t,\ X_t) = \theta_0(t)\ \exp(\beta' X_t)$$

The next step consists in the construction of the Likelihood function that will be maximized.

In our case we observe a person i's spell from year k=1 through to the end of the j[th] year, at which point i´s spell is either complete (c_i =1) or right censored (c_i =0).

The discrete hazard can be defined as:

h_{ij} =Pr(T_i =j/T_i ≥ j)

Likelihood contribution for a right censored spell is given by:

$$L_i = \Pr(T_i > j) = S_i(j) = \prod_{k=1}^{j}(1-h_{ik})$$

Likelihood contribution for each completed spell is given by:

$$L_i = \Pr(T_i = j) = f_i(j) = h_{ij}S_i(j-1) = \frac{h_{ij}}{1-h_{ij}}\prod_{k=1}^{j}(1-h_{ik})$$

Likelihood for the whole sample:

$$L_i = \prod_{i=1}^{n}\left[\Pr(T_i = j)\right]^{c_i}\left[\Pr(T_i > j)\right]^{1-c_i} = \prod_{i=1}^{n}\left[\frac{h_{ij}}{1-h_{ij}}\prod_{k=1}^{j}(1-h_{ik})\right]^{c_i}\left[\prod_{k=1}^{j}(1-h_{ik})\right]^{1-c_i} =$$

$$= \prod_{i=1}^{n}\left[\left(\frac{h_{ij}}{1-h_{ij}}\right)^{c_i}\prod_{k=1}^{j}(1-h_{ik})\right]$$

$$\log L_i = \sum_{i=1}^{n} c_i \log\left(\frac{h_{ij}}{1-h_{ij}}\right) + \sum_{i=1}^{n}\sum_{k=1}^{j}\log(1-h_{ik})$$

[19] In continuous time it would be:

$$h(t) = \lim_{\Delta t \to t}\frac{\Pr(t < T < t+\Delta t \mid T > t)}{\Delta t} = \frac{S'(t)}{S(t)} = \frac{f(t)}{(1-F(t))}$$

Next, we define a new binary indicator variable $y_{ik} = 1,$ such as :

$c_i = 1 \Rightarrow y_{ik} = 1$ for $k = T_i$, $y_{ik} = 0$ otherwise
$c_i = 0 \Rightarrow y_{ik} = 0$ for all k

Hence, $\log L_i = \sum_{i=1}^{n} \sum_{k=1}^{j} y_{ik} \log\left(\frac{h_{ik}}{1-h_{ik}}\right) + \sum_{i=1}^{n} \sum_{k=1}^{j} \log(1-h_{ik}) = \sum_{i=1}^{n} \sum_{k=1}^{j} [y_{ik} \log h_{ik} + (1-y_{ik}) \log(1-h_{ik})]$

But this expression has exactly the same form as the standard likelihood function for a binary regression model in which y_{ik} is the dependent variable and in which the data structure has been reorganized from having one record per spell to having one record for each year that a person is at risk of transition from the state (person-period data).

On the empirical implementation of the preceding model, Jenkins (2005) presents an easy four-step estimation method that we will mainly follow through:

1. Reorganize data into firm-period format.
2. Create any time-varying covariates (at the very least this include a variable describing duration dependence in the hazard rate).
3. Choose the functional form for h_{ik}.
4. Estimate the model using any standard binary dependent regression package (ex: STATA).

For point 3, we choose a clog-log specification. The clog-log model is a form of generalized linear model with particular function for the hazard:

$$\log(-\log\left[1-h_j(X_t)\right]) = \beta' X_t + \gamma_j$$

or

$$h_j = 1 - \exp\left[1 - \exp(\beta' X_t + \gamma_j)\right]$$

where γ_j is the log of the difference between the integrated baseline hazard $\theta_0(t)$ evaluated at the end of the interval and the beginning of the interval.

When estimated using interval-censored survival data, it is obtained the regression coefficients β and γ_{j}. The β coefficients are the same ones as those characterizing the continuous time hazard rate $\theta(t,\ X_t) = \theta_0(t)\ \exp(\beta' X_t)$. γ_j summarizes the pattern of duration dependence in the discrete time hazard, but in order to identify the precise pattern of duration dependence in the continuous time hazard, further assumptions are required. Most analysts do not impose any restriction on how γ_j varies from interval to interval, instead the variation in γ_j is specified using a parametric functional form. Ex: $c\log\log\left[h(j, X)\right] = z_2 j^2 + \beta' X$ and z_2 is a parameter to be estimated together with β.

4. Database

The main source of information has been the SABE (System of Analysis of Spanish Balance Sheets) database, elaborated by Bureau Van Dijk. This database includes accounting and financial information from Spanish firms obtained from the annual financial statements deposited at the Registry of Companies. We extract balance sheet information and additional controls for a set of non-listed firms in the stock exchange that have announced a "*concurso preventivo*" or declared their bankruptcy in the period that goes from 1993 up to 2005[20] inclusive. These two situations constitute our definition of financial distress. We are able to indicate our "exit" event using two variables ("Close_year" and "Firm_Status") that indicate the year and change of status (from active to *concurso preventivo*/bankruptcy).

We also took a comparable large sample of "healthy" non-listed firms from the same source. We do not account for merger or acquisitions, so will have a binary dependent variable that equals to 1 only at the time a "*concurso preventivo*" or a bankruptcy is observed (the "exit" event to be targeted in the estimation). Given the fact that in many occasions the change of status is announced some years after the disruption of information by the firm to SABE, we consider the last year of available information as the last year previous distress, always controlling that the firm

[20] The number of firms included in the database has been increasing with time and unfortunately there are very few firms in the database before 1993 and especially very few *concursos preventivos* or bankruptcies with available balance sheet information.

effectively announces the bankruptcy or *concurso preventivo* in some subsequent year. In particular we consider all the distress announcements up to five years after the last year of observation.

In order to compare the evolution of firms that suffered financial distress we extract a random sample of comparable "healthy" firms from the universe of firms in the SABE database. We extract a 20% random sample of firms by decile of number of employees in each of the industrial sectors defined by the NACE[21] code up to two digits. We have tested our results to be robust to a simple random sample without considering the stratification by the number of employees or the industrial sector.

We clean the data from cases where no information is provided for a set of basic fields (for example, no information on total assets and/or total liabilities is presented for more than 3 consecutive years). Then, we construct a set of standard financial ratios usually used in the bankruptcy prediction literature and we drop from the sample those firms with extreme values in these set of financial indicators. The criteria we follow to filter the data is to eliminate those firms that fall out the following interval $[mean(x_i) - 3sd(x_i), mean(x_i) + 3sd(x_i)]$, where *sd* is the standard deviation for each of the financial indicators[22].

As Chava and Jarro (2004) recognize most of the bankruptcy prediction models fitted in the academic literature are based on a limited data set containing at most 300 bankruptcies. At this respect our sample is larger and presents more observations of financial distresses (1143) than some of the usually referred studies in the area, such as, Altman´s (1980) with 33 bankruptcies, Ohlson (1980) with 105 bankruptcies and Shumway (2001) with 300 bankruptcies, among others. Our database is similar in the number of observed exit events to the one used in Pompe and Bilderbeek (2005) where they have 1369 bankruptcies observations. Nevertheless it is worthy to notice that we are working with a relatively low proportion of distressed firms respect the total number of healthy ones, especially if we compare with other previous studies also cited in the present work (3% in Corcóstegui *et al* 2003, 6% in Duffie *et al* 2007, 17% in Valles 2006). On

[21] NACE stands for "Nomenclature générale des activités économiques dans les Communautés Européennes".
[22] An analogous strategy is followed by Aybar Arias e*t al* (2003).

this point, the proportion of distresses and healthy firms in our sample (0.85%) is close to the proportion of 0.97% found in Hillegeist *et al* (2004). We consider that our relatively low frequency of distresses could reflect the real situation in many Spanish entities with credit databases that usually contain very few default events. This situation may jeopardize the confidence in their estimations.

In table 1 we present a brief description of the number of healthy and distressed firms in each year of our sample. The SABE database, and by consequence the random sample that we extract, clearly expands the number of firms since 1993 onwards. The year 1993 has been dropped from the database given the no presence of distressed firms in that year.

[Insert table 1]

The selection of the specific ratios to consider into the regressions varies among the diverse papers in the literature; nevertheless we can summarize the most frequent financial ratios into four dimensions: profitability, productivity/activity, liquidity, and solvency/leverage. In addition to indicators in each these dimensions, other systemic or common factors (typically macro variables) are usually included, jointly with controls by sector and size. Pompe and Bilderbeek (2005) and Crouhy *et al* (2001) present an extensive list of commonly used financial ratios in bankruptcy/default analysis. In table 2 we present a set of financial ratios we have explored in each of the cited dimensions.

[Insert table 2]

Additionally we consider a set of controls by sector of activity. We have grouped the firms according their NACE codes into four sectors: Agricultural[23], Industrial[24], Construction, and Services. We also use two macroeconomic variables, the real GDP growth rate, and the short term deposit interest rate, this last one as proxy of the general interest rate level in the economy. Finally we study different specifications for the duration term.

[23] As it was not reported any "exit" event in the Agricultural sector and therefore it is not possible to run a hazard regression that includes this sector as control, we have dropped these observations (we dropped 2545 year-firms, 1.85% of the total sample).
[24] We have also included 460 year-firms from the Energy sector (0.33% of the total sample) into the Industrial sector.

[Insert table 3]

5. Preliminary statistics and estimation results

In this section we present the results we get using the hazard model previously described. It is worthy to notice that one key element in the estimation strategy we are following consist in the organization of the data into firm-period data. Next, as it was commented before is necessary to define the "exit" event, in our case bankruptcy or *concurso preventivo*. Those firms that disappear from the database but then declare their bankruptcy after 1, 2, 3, 4 or 5 years, have been also considered and the corresponding distress indicator has been assigned to the last period of available information.

In what respect to the macro-variables we include a measure of the overall state of the economy, as it is the real GDP growth rate. We test the robustness of this variable using a measure of the output-gap reported by the IMF in the World Economic Outlook statistics (WEO). In addition, as proxy of the general level of interest rates in the economy we use the short-term deposit interest rate also extracted from WEO. It is worthy to notice that even when we are reporting the regression output for non-lagged explanatory variables, the fact that our last observation for distressed firms corresponds to a previous year to the disappearance of the firm in the sample, and therefore the *de facto* declaration of distress (our "exit" event), the whole set of variables can be considered as 1 year lagged. Despite this clarification we have also run a set of regressions with up to three lags in the real GDP growth rate and the short-term interest rate. The results remains qualitative the same.

In table 4 we presented the accumulated frequency of distresses by years of survival or age of the firm. It can be observed that more than half of the observed distresses correspond to firms with less than 5 years of age. We see a peak of distresses around the third year and then a continuous decreasing from that point on.

[Insert table 4]

In table 5 we describe the distribution of firms in each sector of activity. The highest frequency of default is found in the Industrial sector (1.2%), followed by Construction (1.0%) and finally by the Services sector (0.7%).

[Insert table 5]

For the set of financial ratios we are using as explanatory variables we have explore several specifications and combinations of them in each of the defined dimensions. After considering their correlations patterns and their level of individual and global significance in the model, we have selected the following six ratios for their use in our baseline model:

[Insert table 6]

Tables 7, 8 and 9 present a set of descriptive statistics for the explanatory variables. Table 3 shows how the main financial indicators reflect an evident average deterioration in distressed firms. The analysis of the correlation matrix was useful in the selection process of the financial indicators.

[Insert table 7]
[Insert table 8]
[Insert table 9]

As a previous step to the regression analysis, we present in table 8 the results of an ANOVA analysis comparing healthy *vs* distressed firms. We find significant differences among both groups in all the studied variables.

[Insert table 10]

We present in table 11 the results of the clog-log regression. The signs of the financial indicators are the ones we can intuitively expect *a priori*. The time parametric specification with better fit has been a quadratic one. Lopez-Garcia and Puente (2006) also find an inverted-U shape curve with a maximum at around 4 years in the firms' age. They also notice that the baseline hazard function does not vary much depending on the specification used (parametric or non-parametric). As these authors point out this type of duration dependence has also been found in other countries such as the United States, United Kingdom, Italy or Germany. It responds to what is known as the "liability of the adolescence" or "the honeymoon effect" brought about by stock of the initial resources of the new firm. Those resources help the new firm go through the first years even if the firm results to be inefficient. Once the initial stock of resources is used up, if the firm is inefficient it will exit the market. Consistently with these empirical findings, we also observed a pick in the smoothed baseline hazard function around 5 years in the case of our estimation (see figure 1).

[Insert table 11]

[Insert figure 1]

In what refers to the activity sectors it can be seen that both the Construction and the Services sectors has been relative safer than the Industrial Sector. The very high expansion of the Construction sector and the less pronounced but steadier growth in Services during the analyzed period gives some factual support to these findings. The average growth rate in the Gross Production Value (GPV) during 1994-2005 was around 5.1% in Construction, 3.5% in Services and a 3.1% in the Industrial sector. In spite of the fact that the Construction was the sector with highest GPV average growth rate, it was more volatile (3.0% std. deviation) than Services (1.1%) and closer to the industrial sector volatility (2.3%). In figure 2 we present a graph with the smoothed hazard estimate for different sectors. It can be observed how the Services sector has been the safest one in relative terms, followed by the Construction sector.

[Insert figure 2]

In the analysis of the macro environment, both variables are significant when they are the only covariates included, but with a counter-intuitive positive sign in the case of real GDP growth. In principle, one would normally expect that an improvement in the macro-economic conditions which is reflected in higher growth rates would be leading to a decrease in the probabilities of financial distress. We try to explain the sign that we find studying the possible effects of this highly expansionary period splitting the sample into "young" and "mature" firms. We explain our findings on this point a couple of paragraphs below.

On its side, in addition to model (1) the interest rate turns up significant only in models (2) and (3), where we exclude the duration term and the financial ratios. The coefficient shows a positive relation, reflecting that harsher financial conditions would increase the firms' probabilities of distress. The fact that it becomes non-significant when financial conditions are included in the regression would be indicating that an important part of its fluctuations directly affects some of the firms' financial indicators (take for example its connection with Financial Profits/Assets ratio).

In table 12 we present the outputs for a set of partial models where we have divided the sample into "young" firms (less or equal than 5 years) and "mature" firms (more than 5 years). We find that the short-term interest rate, as a proxy of the overall financial conditions, is not significant in any of the partial models in the young firms' case. This result could be indicating that during the analyzed period the financial conditions for starting and firms below the threshold of 5 years has been not very binding. Nevertheless, after a period of time, and as noticed by Lopez Garcia and Puente, when the resources are used up, the financial restrictions start to play a role as it is reflected in its sign and level of significance in the model (16). In the case of mature firms, when only macro-economic variables are included[25], the interest rate has a significant and positive sign, indicating a worsening in the financial conditions could lead to an increase in their estimated PDs. Aligned with this line of reasoning, the real GDP growth effect is significant but with a positive sign for young firms and not significant for mature firms. This could be interpreted as evidence that during the period analyzed, very favorable financial conditions (or at least, not binding) and high GDP growth rates have been encouraging the advent of start-ups firms that have grown inefficiently and have not been able to continue more years in the market and to successfully overcome the barrier of the 5 years.

[25] When all the financial ratios are included the interest rate is close to be significant at a 10% and maintain its sign. As it was indicated before, it is expected that part of its effect hits directly to some of the financial ratios we are including.

[Insert table 12]

As it was the case in models 1 to 10, the financial ratios present the expected sign and all of them remain significant in both samples, with the exception of WK/Sales in the case of young firms (model 15).

At the whole, considering the log-likelihood values, the individual and global levels of statistical significance, the correlations among variables, the ex-ante attempt of trying to keep represented the four dimensions usually present in the analysis of financial ratios, the introduction of controls by sector and macroeconomic variables, and the specification of a parametric baseline hazard function in line with other empirical findings, the model we use to forecast each year-firm PD is model 10.

6. Validation

6.1. Discriminatory power

There are several statistical methodologies for the assessment of the discriminatory power of the prediction model; a set of them are presented in the 2005 Basel Committee document on validation of internal rating systems. This document differentiates between two stages for the validation of PDs: validation of the discriminatory power of a rating system and validation of the accuracy of the PD quantification (calibration). The Basel Committee also recognizes that compared with the evaluation of the discriminatory power, methods for validating calibration are at a much earlier stage and that a major impediment to back-testing of PDs is the scarcity of data, mainly caused by the infrequency of default events and the possible impact of default correlation. Due to the limitations of using statistical tests to verify the accuracy of the calibration, benchmarking is also suggested as a complementary tool for the validation

One of the most famous techniques in the assessment of discriminatory power is the cumulative accuracy profile (CAP) and its summary index the accuracy ratio (AR). The CAP is also known as the “Gini curve”, “Power curve” or “Lorenz curve” and it is a graphical representation of the proportionality of a distribution. To build the CAP, the observations are ordered from the low PDs to the high ones. Then the cumulative frequency of firms is plot on the x-axis and the cumulative frequency of distress firms is plot on the y-axis. A perfect CAP curve would accumulate all the firms in distress first at the left. In that case the CAP is an increasing line from the origin up to 1 and then staying at 1. This perfect model line is illustrated in figure 3 with a dash line. A random model would not have any discriminatory power and therefore any fraction of firms will contain the same proportion of distressed and non distressed firms. The curve for the random model is represented by the diagonal in figure 3. The AR is the ratio of the area between the CAP curve and the diagonal, and the area between the CAP curve and the perfect model. So, the rating is better the closer the AR is to 1. The concavity of the CAP also permits to visualize the degree of discrimination and use of the information in the score function and estimated PDs. The shape of the CAP depends on the proportion of distresses and non-distresses firms; hence as it is explicitly stated by the Basel Committee, a direct comparison of CAPs across different portfolios could be misleading. In figure 4 we show the CAP curve for the one-year-ahead prediction up to four-years ahead, and the CAP for all the firm-years. The associated ARs are shown in table 13.

[Insert figure 3]

[Insert figure 4]

[Insert table 13]

In order to test the discriminatory power of the model out sample, we use the baseline model to produce the estimated PDs for the period 2005-2007, but only using the data available for 1994-2002. In this case the AR for 1-year-prediction has been of 52.1%.

6.2. Calibration

Given the satisfactory performance of the forecasted PDs using the proposed model, the next step consist in the mapping of the estimated PDs intro risk buckets or categories in order to get a rating system for corporate firms in Spain. In order to do so, we perform a k-mean cluster analysis on the predicted PDs and the cumulative frequency of the firms (see analogous applications in Fernandes 2005 in Portugal and Valles 2006, for the Argentinean case). We build the rating system using the whole data span of year-firms.

We perform the k-mean cluster analysis[26] on the PDs and the frequency of firms with two objectives in mind. First, because we follow the Basel II recommendation on avoiding large exposure concentrations in each grade of the rating system. And second, because in this way the rating we get is easily comparable respect any other random sample size of healthy firms that can be extracted from the universe of Spanish firms between 1994 and 2005. In this sense, it is important to remember that objective of the calibration analysis is to study the relative properties of the ratings, such as decreasing PDs for each rating class, the relation between the PDs associated at each bucket and the corresponding frequency of distresses and the overall consistency of the rating. We build the seven-grades rating in line with the normal practice of two major rating agencies, S&P and Moody's, and as the minimum required in the New Accord. Then, we associate the average PDs of all the year-firms in each bucket to each grade,

In table 14 we present the rating we obtained. In practice, bank's PD estimates for each bucket will differ from the default rate calculated for the global database. The calibration task consists in the analysis of the distance between the estimated PDs and the realized default rates. In our case we perform binomial tests[27] in order to detect large deviations from the pooled PDs in each grade. The binomial test can be applied to one rating category at a time only. As it can be seen, we only have one violation both at the 99% and 95% confidence levels (in grade 7) and two at 90% (grade 6 and 7), indicating the need to adjust our last and more risky grades. In fact, a quite common feature when large databases containing a relatively few number of defaults are used and particularly if they are not covering a clear recessionary period, is that the estimations are prone to over-estimate the PDs in the good grades and under-estimated them in the bad grades. This situation is illustrated in figure 5 where it can be seen how the curve of estimated PDs and the pooled frequency of observed distresses cross each other after grade 5. It can be also observed how there is an smoother classification according to the observed pooled frequency of distresses than in the estimated PDs. Schuermann and Hanson (2004) find that for ratings that introduce

[26] More details on the use of k-mean cluster analysis for classification purposes be found in Barholomew (2002).
[27] An explanation of the binomial test and the HL test can be found in Appendix2.

conditioning on the state of the business cycle, it is easier to distinguish adjacent PDs in recession than in expansion. Given the fact that we are not including a clear recessionary period, it is harder for our model to distinguish between the different types of investment grades, what introduce some tendency to a more dichotomic picture and a sharper change of slope when we approximate to more risky grades.

[Insert table 14]

[Insert figure 5]

One suggested calibration of the estimated PDs, according to the conservative view suggested by the Basel Committee, and in order to adjust the estimated PD curve to the observed characteristics of the pooled frequency of distresses is to re-scale corresponding average PDs for grades 5, 6, 7. This external adjustment generates the PDs curve that is presented in figure 6.

[Insert figure 6]

Despite we found average PDs for each of the grades that closely follows the observed pooled frequencies of distresses and have a good discriminatory power, we still have to analyze how stable are the observed pooled frequencies across the years of our sample. For this purpose we compare the observed pooled frequency of distresses with the observed frequency in each year and in each grade of our sample (see table 15).

[Insert table 15]

We then perform a Hoshmer-Lemeshow test on the observed frequency of distresses for the whole rating system by year. The HL test is a join test that permits the examination of all rating grades simultaneously. The null hypothesis in the case we are testing is that the pooled frequencies of distresses are the true frequencies in each of the year of the sample. The results of this test can be seen at the bottom of table 14. The test only would reject the null hypothesis in 1999, 2002 and 2005 at a 1% confidence level. In addition, if we observed the contribution in each bucket we do not observe systematic high contributions (for example, values higher than 5 are shaded in table

16). The average values for the whole period are in the range of acceptance, providing support on the stability of the TCC rating we have built. It is worthy to notice that the HL test is a quite strict test and with lower support than the binomial test. Both tests should be consider together in order to assess the overall stability of the rating system.

[Insert table 16]

7. Conclusions

During this work we have applied a set of standards techniques, measures and tests, suggested by the Basel Committee in the IRB framework and also normally used by practioneers in the field. The PDs estimation method based on a parametric proportional hazard model has proven to be a flexible and efficient tool for this purpose and with clear advantages respect first and second generation models. Based on the estimated PDs we have constructed a stable TTC rating system for Spanish non-listed corporate firms during 1994-2005. The overall discrimination power and calibration has been satisfactory.

During the analysis some interesting effects of a set of macro variables has been found. In particular that the effect of growth and interest rate on the estimated PDs depend on the maturity of the firm. It seems that that during the analyzed period, the emergence of young starting firms have been boosted in periods of high growth, with the interest rate playing a little role. Nevertheless, many of these firms were not able to survive and have suffered of financial distress experiences after a few years of entry. On the other hand, the financial position of mature firms have resulted lees sensible to fluctuations in the real GDP growth rates but they performance seems to be more exposed to fluctuations in the interest rate. This situation jointly with other domestic and international empirical evidence justifies the inclusion of an U-shaped baseline hazard function in our PDs prediction model. The analysis of macroeconomic distinct effect on young and mature firms can also contribute to give some extra support to the “liability of the adolescence” or “honeymoon effect” arguments.

In addition, along this document we have highlighted some difficulties that arise in the construction of TCT rating systems based on a period of long and persistent expansion. Estimations and ratings based on this period should be probably extra-adjusted in order to account for a possible deterioration in the macroeconomic environment. To achieve this objective in the Spanish case could be highly challenging given the scarcity of information before 1994 and the rapid growth of productive sectors and financial markets since then.

The consideration of the access to banking lending is an important issue to take into account at the moment of study why young firms are or not sensible to fluctuations in interest rates. It could be the case that they have experimented a very easy access to external financial resources, especially from banks during the analyzed period, but on the opposite side it could be also the case that they have been in fact, financially constrained. An alternative explanation in this last case would be that without access to external financial access the decision of starting a new project could be mostly relying on own internal resources and depending more on the expected returns of the owners/managers than in the current level of interest rates. More research is deserved on these issues.

Finally, the estimation strategy we presented in this document could be also used to study the existence of ‘frailty’ or unobserved heterogeneity in the sample of firms. As it is shown in Das et al (2007) the existence of frailty/contagion can invalidate many of normal assumptions in credit risk models, in particular, the doubly stochastic assumption. The study of asset correlation and ex-ants and ex-post contagion is part of our future research agenda.

References

Altman, E. I., 1968, "Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy", Journal of Finance, 23, 58-609.

Andreev, A. Y., 2007, "Predicting financial distress: A multinomial logistic approach to Spanish firms", mimeo, Autonomous University of Barcelona.

Aybar Arias C., Casino Martinez A. and López García J., 2003, "Estrategia y estructura de capital en la PYME: una aproximación empírica", Estudios de Economía Aplicada, vol. 21, No. 001, April.

Bartholomew D.J., Steele F., Moustaki I. and Galbraith J.I., 2002. The analysis and interpretation of multivariate data for social scientists. Texts in Statistical Science. Chapman & Hall/CRC.

Balzarotti V. Castro C.E. and Powell A., 2004, "Reforming capital requirements in emerging countries: Calibrating Basel II using historical Argentine Credit Bureau data and CreditRisk+". Business School Working Paper, Universidad Torcuato Di Tella.

Basel Committee on Bank Supervision, 2006, "International Convergence of Capital Measurement and Capital Standards. A Revised Framework. A Revised Framework.
Comprehensive Version", Bank for International Settlements, June.

Basel Committee on Bank Supervision, 2005, "Studies on the validation of internal rating systems", Working Paper, 14, May.

Basel Committee on Bank Supervision, 2002, "Quantitative Impact Study 3. Technical Guidance", Bank for International Settlements, October.

Basel Committee on Bank Supervision, 2001, "The New Basel Capital Accord", Consultative Document, Bank for International Settlements, January.

Basel Committee on Bank Supervision, 1988, “The Basel Capital Accord”, Bank for International Settlements, March.

Beaver, B., 1966, “Financial ratios as predictors of failure”, Journal of Accounting Research, Autumn, Empirical Research in Accounting: Selected Studies, Supplement, 91–101.

Beaver, B., 1968, “Alternative accounting measures as predictors of failure”, Accounting Review, 113–122, January.

Beck N., Katz J.N. and Tucker R., 1998, “Taking time seriously: Time-series-cross-section analysis with a binary dependent variable”, American Journal of Political Science, Vol. 42, No.4, 1260-1288, October.

Begley, J., Ming J. and Watts, S., 1996, “Bankruptcy classification errors in the 1980s: An empirical analysis of Altman’s and Ohlson’s models”, Review of Accounting Studies, 1, 267-284.

Benito A., Delgado J.F. and Pages J.M., 2004, “A synthetic indicator of financial pressure for Spanish firms”, Banco de España, Documentos de Trabajo, No. 411, June.

Black, F. and Scholes M., 1973, “The pricing of options and corporate liabilities”, Journal of Political Economy, 81, 637-654.

Chava S. and Jarrow R. A., 2004, “Bankruptcy prediction with industry effects”, Review of Finance, 8, 537-569.

Corcóstegui C., Gonzalez-Mosquera L., Marcelo A. and Trucharte C., 2003, “Analisys of procyclical effects on capital requirements derived from a rating system”, Basel Committee Workshop, Banking and Financial Stability: A Workshop on Applied Banking Research, Banca d’Italia, Rome.

Crouhy M., Mark R. and Galai D., 2001. Risk Management. McGraw-Hill.

Crouhy, M. and Mark R., 1999, “A comparative analysis of current credit risk models”, Financial Stability Review, June, 110-112.

Das S.R., Duffie D., Kapadia N. and Saita L., 2007 “ Common failings: How corporate defaults are correlated”, The Journal of Finance, Vol. 62, No. 1, February.

Davis, M. H. and Lo V.,2001, “Infectious defaults”, Quantitative Finance 1, 305-308.

Dewatripont M. and Tirole J., 1994. The Prudential Regulation of Banks. Cambridge, Mass.:MIT Press.

Dichev I.D.,1998, “Is the risk of bankruptcy a systematic risk”, The Journal of Finance, Vol. 53, No. 3, June.

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

Fernandes, J. E., 2005, “Corporate credit risk modeling: Quantitative rating system and probability of default estimation”, mimeo.

Garcia Baena R., Gonzalez Mosquera L. and Oroz García M., 2005, “Aspectos críticos en la implantación y validación de modelos internos de riesgo de crédito”, ISSN 1579-2498, No. 9, 29-58.

Gordy, Michael B., 2000, “A comparative anatomy of credit risk models,” Journal of Banking and Finance, January, 24 (1-2), 119-149.

Hammes, W. and Shapiro M., 2001, “The implications of the New Capital Adequacy rules for portfolio management of credit assets.” Journal of Banking and Finance, January, 97–114.

Hillegeist S.A., Keating E. K, Cram D. P. and Lundstedt K.G., 2004, "Assessing the probability of bankruptcy", Review of Accounting Studies, 9, 5-34.

Jarrow, R. A.; Yu F., 2001, "Counterparty risk and the pricing of defaultable securities", Journal of Finance 61, 1765-1799.

Jenkins S.P., 2005, Course Notes, University of Essex, Institute for Social and Economic Research, July.

Jensen M. and Meckling W., 1976, "Theory of the firm: Managerial behavior, agency costs and ownership structure". Journal of Financial Economics, 3, 305-360.

Johnsen T. and Melicher R. W., 1994, "Predicting corporate bankruptcy and financial distress: Information value added by multinomial logit models", Journal of Economics and Business, 269-286.

Kiefer, N.M., 1998, "Economic duration data and hazard functions", Journal of Economic Literature, Vol. 26, No. 2, 646-679, June.

Lando, D., 1998, "On Cox processes and credit risky securities", Review of Derivatives Research 2, 99-120.

Li, D. X., 2000, "On default correlation: A copula function approach", Working Paper, The Risk Metrics Group.

López-García P. and Puente S., 2006, "Business demography in Spain: determinants of firms survival", Banco de España, Documentos de trabajo No. 0608.

Ohlson, J.S., 1980, "Financial ratios and the probabilistic prediction of bankruptcy", Journal of Accounting Research, 19, 109-131.

Orbe J., Ferreira E. and Núñez-Antón V., 2002, “Chapter 11 bankruptcy: a censored partial regression model”, Applied Economics, Vol. 34, Issue 15, 1949-1957, October.

Pompe P.P. and Bilderbeek J., 2005, “The prediction of bankruptcy of small- and medium-sized industrial firms”, Journal of Business Venturing, Vol. 20, issue 6, 847-868 Powell A., 2004, “Basel II and Developing Countries: Sailing through the Sea of Standards”, World Bank Policy Research Working Paper 3387, September.

Powell A., 2004, “Sailing through the sea of standards”, World Bank Policy Research Working Paper 3387.

Saidenberg M and Schuerman T., 2003, “The New Basel Capital Accord and questions for research”, Federal Reserve Bank of New York, May.

Saunders, A. and Allen L., 1999. Credit Risk Measurement: New Approaches to Value at Risk and Other Paradigms. John Wiley & Sons Inc. New York, 1999.

Schuerman T. and Hanson S., 2004, “Estimating probabilities of default”, Federal Reserve Bank of New York, Staff Report no. 190, July.

Shumway, T., 2001, “Forecasting bankruptcy more accurately: A simple hazard model”, The Journal of Business, Vol. 74, No. 1, 101-124, January.

Soberhart, J. and Keenan S. C., 2001, “A Practical review and test of default prediction models”, The RMA Journal, November.

Standard & Poor´s, 2007, “Default, transition, and recovery: 2007 Annual global corporate default study and rating transitions”, Standard&Poor´s Rating Direct, February.

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

Valles, V., 2006, “Stability of a ‘through-thecycle’ rating system during a financial crisis”, Financial Stability Institute, FSI Award 2006 Winning Paper, September.

Vasicek O., 1987, “Probability of loss on loan portfolio”, KMV Corporation (available at kmv.com)

Wooldridge, J. M., 2002. Economic analysis of cross section and panel data. MIT Press, Cambridge.

Zmijewski M. E., 1984, “Methodological issues related to the estimation of financial distress prediction models”, Journal of Accounting Research, 22, 59-82.

Appendix 1

Issue	Main topic	Basel II paragraph number	Summary of relevant requirements/criteria
General	General principle	389	The overarching principle behind the IRB minimum requirements is that rating and risk estimation systems and processes provide for a meaningful assessment of borrower and transaction characteristics; a meaningful differentiation of risk; and reasonably accurate and consistent quantitative estimates of risk. The systems and processes must be consistent with internal use of these estimates.
	Internal estimation of PDs	391	All IRB banks must produce their own estimates of PD and must adhere to the overall requirements for rating system design, operations, controls, and corporate governance, as well as the requisite requirements for estimation and validation of PD measures.
	The use of internal models	417	The requirements in this section apply to statistical models and other mechanical methods used to assign borrower or facility ratings or in estimation of PDs, LGDs, or EADs. Credit scoring models and other mechanical rating procedures generally use only a subset of available information. Although mechanical rating procedures may sometimes avoid some of the idiosyncratic errors made by rating systems in which human judgement plays a large role, mechanical use of limited information also is a source of rating errors. Credit scoring models and other mechanical procedures are permissible as the primary or partial basis of rating assignments, and may play a role in the estimation of loss characteristics. Sufficient human judgement and human oversight is necessary to ensure that all relevant and material information, including that which is outside the scope of the model, is also taken into consideration, and that the model is used appropriately.
			● The variables that are input to the model must form a reasonable set of predictors ● The model must be accurate on average across the range of borrowers or facilities to which the bank is exposed and there must be no known material biases. ● The bank must have in place a process for vetting data inputs into a statistical default or loss prediction model which includes an assessment of the accuracy, completeness and appropriateness of the data specific to the assignment of an approved rating. ● The bank must demonstrate that the data used to build the model are representative of the population of the bank's actual borrowers or facilities.
	Documentation	418	Banks must document in writing their rating systems' design and operational details. The documentation must evidence banks' compliance with the minimum standards. If the bank employs statistical models in the rating process, the bank must document their methodologies. This material must: ● Provide a detailed outline of the theory, assumptions and/or mathematical and empirical basis of the assignment of estimates to grades, individual obligors, exposures, or pools, and the data source(s) used to estimate the model; ● Establish a rigorous statistical process (including out-of-time and out-of-sample performance tests) for validating the model ● Indicate any circumstances under which the model does not work effectively.
PDs	Long-run average	447	PD estimates must be a long-run average of one-year default rates for borrowers in the grade, with the exception of retail exposures.
	Internal and external data	448	Internal estimates of PD, LGD, and EAD must incorporate all relevant, material and available data, information and methods. A bank may utilise internal data and data from external sources (including pooled data). Where internal or external data is used, the bank must demonstrate that its estimates are representative of long run experience.
	Empirical evidence	449	Estimates must be grounded in historical experience and empirical evidence, and not based purely on subjective or judgmental considerations. A bank's estimates must promptly reflect the implications of technical advances and new data and other information, as it becomes available. Banks must review their estimates on a yearly basis or more frequently.
	Matching	450	The population of exposures represented in the data used for estimation, and lending standards in use when the data were generated, and other relevant characteristics should be closely matched to or at least comparable with those of the bank's exposures and standards.
	Conservative view	451	In general, estimates of PDs, LGDs, and EADs are likely to involve unpredictable errors. In order to avoid over-optimism, a bank must add to its estimates a margin of conservatism that is related to the likely range of errors

	Definition of "default"	452	A default is considered to have occurred with regard to a particular obligor when either or both of the two following events have taken place. ● The bank considers that the obligor is unlikely to pay its credit obligations to the banking group in full, without recourse by the bank to actions such as realizing security (if held). ● The obligor is past due more than 90 days on any material credit obligation to the banking group. Overdrafts will be considered as being past due once the customer has breached an advised limit or been advised of a limit smaller than current outstandings.
	Indications of unlikeliness to pay	453	The elements to be taken as indications of unlikeliness to pay include: ● The bank puts the credit obligation on non-accrued status. ● The bank makes a charge-off or account-specific provision resulting from a significant perceived decline in credit quality subsequent to the bank taking on the exposure. ● The bank sells the credit obligation at a material credit-related economic loss. ● The bank consents to a distressed restructuring of the credit obligation where this is likely to result in a diminished financial obligation caused by the material forgiveness, or postponement, of principal, interest or (where relevant) fees.The bank has filed for the obligor's bankruptcy or a similar order in respect of the obligor's credit obligation to the banking group. ● The obligor has sought or has been placed in bankruptcy or similar protection where this would avoid or delay repayment of the credit obligation to the banking group.
	Combination of information and techniques I	461	Banks must use information and techniques that take appropriate account of the long-run experience when estimating the average PD for each rating grade. For example, banks may use one or more of the three specific techniques set out below: internal default experience, mapping to external data, and statistical default models.
	Combination of information and techniques II	462	Banks may have a primary technique and use others as a point of comparison and potential adjustment. Supervisors will not be satisfied by mechanical application of a technique without supporting analysis. Banks must recognise the importance of judgmental considerations in combining results of techniques and in making adjustments for limitations of techniques and information ● A bank may use data on internal default experience for the estimation of PD. The use of pooled data across institutions may also be recognised. ● Banks may associate or map their internal grades to the scale used by an external credit assessment institution or similar institution and then attribute the default rate observed for the external institution's grades to the bank's grades. ● A bank is allowed to use a simple average of default-probability estimates for individual borrowers in a given grade, where such estimates are drawn from statistical default prediction models. The bank's use of default probability models for this purpose must meet the standards specified in paragraph 417
	Historical observation period of at least 5 years	463	Irrespective of whether a bank is using external, internal, or pooled data sources, or a combination of the three, for its PD estimation, the length of the underlying historical observation period used must be at least five years for at least one source
Rating	Definition of "rating system	394	The term "rating system" comprises all of the methods, processes, controls, and data collection and IT systems that support the assessment of credit risk, the assignment of internal risk ratings, and the quantification of default and loss estimates.
	Customised ratings	395	Within each asset class, a bank may utilise multiple rating methodologies/systems. For example, a bank may have customised rating systems for specific industries or market segments (e.g. middle market, and large corporate)
	Dimensions	396	A qualifying IRB rating system must have two separate and distinct dimensions: (i) the risk of borrower default, and (ii) transaction-specific factors
	Risk of borrower default	397	The first dimension must be oriented to the risk of borrower default. Separate exposures to the same borrower must be assigned to the same borrower grade, irrespective of any differences in the nature of each specific transaction (there are two exception: in the case of country transfer risk and when the treatmen of associated guarantees to a facility may be reflected in an adjusted borrower grade)
	Transaction-specific factors	398	The second dimension must reflect transaction-specific factors, such as collateral, seniority, product type, etc.
	No excessive concentrations	403	A bank must have a meaningful distribution of exposures across grades with no excessive concentrations, on both its borrower-rating and its facility-rating scales.

	Min. number of grades	404	To meet this objective, a bank must have a minimum of seven borrower grades for non-defaulted borrowers and one for those that have defaulted. Supervisors may require banks, which lend to borrowers of diverse credit quality, to have a greater number of borrower grades.
	PDs to each grade	405	A borrower grade is defined as an assessment of borrower risk on the basis of a specified and distinct set of rating criteria, from which estimates of PD are derived.
	Horizon	414	Although the time horizon used in PD estimation is one year (as described in paragraph 447), banks are expected to use a longer time horizon in assigning ratings.
	Consideration of systemic effects	415	A borrower rating must represent the bank's assessment of the borrower's ability and willingness to contractually perform despite adverse economic conditions or the occurrence of unexpected events. For example, a bank may base rating assignments on specific, appropriate stress scenarios. Alternatively, a bank may take into account borrower characteristics that are reflective of the borrower's vulnerability to adverse economic conditions or unexpected events, without explicitly specifying a stress scenario. The range of economic conditions that are considered when making assessments must be consistent with current conditions and those that are likely to occur over a business cycle within the respective industry/geographic region.
	Conservative view	416	Given the difficulties in forecasting future events and the influence they will have on a particular borrower's financial condition, a bank must take a conservative view of projected information. Where limited data are available, a bank must adopt a conservative bias to its analysis.
	Review	424	Rating assignments and periodic rating reviews must be completed or approved by a party that does not directly stand to benefit from the extension of credit
	Rating refreshment	425	Borrowers and facilities must have their ratings refreshed at least on an annual basis.
	Rating history	430	Banks must maintain rating histories on borrowers and recognised guarantors, including the rating since the borrower/guarantor was assigned an internal grade, the dates the ratings were assigned, the methodology and key data used to derive the rating and the person/model responsible
	Internal relevance	444	Internal ratings and default and loss estimates must play an essential role in the credit approval, risk management, internal capital allocations, and corporate governance functions of banks using the IRB approach
	Track record of at least 3 years in the use of internal ratings	445	A bank must have a credible track record in the use of internal ratings information. Thus, the bank must demonstrate that it has been using a rating system that was broadly in line with the minimum requirements for at least the three years prior to qualification
	Adjustment by guarantee	480	When a bank uses its own estimates of LGD, it may reflect the risk-mitigating effect of guarantees through an adjustment to PD or LGD estimates
	Floor to the adjustment by guarantee	482	In no case can the bank assign the guaranteed exposure an adjusted PD or LGD such that the adjusted risk weight would be lower than that of a comparable, direct exposure to the guarantor.
Validation / tests	Stress-testing process	434	An IRB bank must have in place sound stress testing processes for use in the assessment of capital adequacy. Stress testing must involve identifying possible events or future changes in economic conditions that could have unfavourable effects on a bank's credit exposures and assessment of the bank's ability to withstand such changes. Examples of scenarios that could be used are (i) economic or industry downturns; (ii) market-risk events; and (iii) liquidity conditions.
	Stress test of specific conditions. Mild recession scenarios.	435	In addition to the more general tests described above, the bank must perform a credit risk stress test to assess the effect of certain specific conditions on its IRB regulatory capital requirements. The test to be employed must be meaningful and reasonably conservative. Individual banks may develop different approaches to undertaking this stress test requirement, depending on their circumstances. For this purpose, the objective is not to require banks to consider worst-case scenarios. The bank's stress test in this context should, however, consider at least the effect of mild recession scenarios. In this case, one example might be to use two consecutive quarters of zero growth to assess the effect on the bank's PDs, LGDs and EADs, taking account - on a conservative basis - of the bank's international diversification
	Sources of information	436	Whatever method is used, the bank must include a consideration of the following sources of information. First, a bank's own data should allow estimation of the ratings migration of at least some of its exposures. Second, banks should consider information about the impact of smaller deterioration in the credit environment on a bank's ratings, giving some information on the likely effect of bigger, stress circumstances. Third, banks should evaluate evidence of ratings migration in external ratings. This would include the bank broadly matching its buckets to rating categories

	Validation system	500	Banks must have a robust system in place to validate the accuracy and consistency of rating systems, processes, and the estimation of all relevant risk components.
	Calibration I (internal)	501	Banks must regularly compare realised default rates with estimated PDs for each grade and be able to demonstrate that the realised default rates are within the expected range for that grade.
	Calibration II (external)	502	Banks must also use other quantitative validation tools and comparisons with relevant external data sources
	Economic cycle	503	Banks must demonstrate that quantitative testing methods and other validation methods do not vary systematically with the economic cycle.
	Deviations	504	Banks must have well-articulated internal standards for situations where deviations in realised PDs, LGDs and EADs from expectations become significant enough to call the validity of the estimates into question. These standards must take account of business cycles and similar systematic variability in default experiences. Where realised values continue to be higher than expected values, banks must revise estimates upward to reflect their default and loss experience.

Appendix 2

(Extracted from Basel Committee on Banking Supervision 2005)

Binomial test

The binomial test can be applied to one rating category at a time only. If (say) twenty categories are tested, at 5% significance level one erroneous rejection of the null hypothesis "correct forecast" has to be expected. It is based on the assumption that defaults are independent events.

The binomial test is a natural possibility for the validation of *PD* estimates banks have to provide for each rating category of their internal rating systems. Its construction relies on an assumption of the default events in the rating category under consideration being independent.

The binomial test works as follows:

null hypothesis H0:	the *PD* of a rating category is correct
alternative hypothesis H1:	the *PD* of a rating category is underestimated

Given a confidence level q (e.g. 99%) the null hypothesis is rejected if the number of defaulters k in this rating category is greater than or equal to a critical value k^* which is defined as

$$k^* = \min\left\{k \mid \sum_{i=k}^{n}\binom{n}{i} PD^i (1-PD)^{n-i} \leq 1-q\right\},$$

where n is the number of debtors in the rating category. The critical value k^* can be approximated by an application of the central limit theorem to the above formula. This approximation results in

$$k^* = \Phi^{-1}(q)\sqrt{nPD(1-PD)} + nPD,$$

where Φ^{-1} denotes the inverse function of the standard normal distribution. If it is preferred to express it in terms of an observed default rate p^* that is allowed at maximum

$$p^* \approx \Phi^{-1}(q)\sqrt{\frac{PD(1-PD)}{n}} + PD.$$

Therefore both concepts are roughly equivalent. In the sequel we concentrate on the binomial test. However, all results apply also to the normal approximation to the binomial test.

Hosmer-Lemeshow test

The Hosmer-Lemeshow or chi-square allows to test to check several rating categories simultaneously. This test is based on the assumption of independence and a normal approximation.

Let $p_0, \ldots, p_k$ denote the forecasted default probabilities of debtors in the rating categories $0, 1, \ldots, k$. Define the statistic

$$T_k = \sum_{i=0}^{k} \frac{(n_i\, p_i - \theta_i)^2}{n_i\, p_i\, (1 - p_i)}$$

with $n_i =$ number of debtors with rating i and $\theta_i =$ number of defaulted debtors with rating i. By the central limit theorem, when $n_i \to \infty$ simultaneously for all i, the distribution of T_k will converge in distribution towards a χ^2_{k+1}-distribution if all the p_i are the true default probabilities. Note however, that this convergence is subject to an assumption of independent default events within categories and between categories.

The p-value of a χ^2_{k+1}-test could serve as a measure of the accuracy of the estimated default probabilities: the closer the p-value is to zero, the worse the estimation is. However, there is a further caveat: If the p_i are very small the rate of convergence to the χ^2_{k+1}-distribution may be low. But note that relying on the p-value makes possible a direct comparison of forecasts with different numbers of rating categories.

Tables and figures

Year	*Distressed*	*Healthy*	*Total*
1994	*2*	*613*	*615*
1995	*9*	*1533*	*1542*
1996	*40*	*3011*	*3051*
1997	*43*	*3951*	*3994*
1998	*62*	*6023*	*6085*
1999	*106*	*8322*	*8428*
2000	*131*	*10776*	*10907*
2001	*172*	*14286*	*14458*
2002	*173*	*17485*	*17658*
2003	*136*	*20090*	*20226*
2004	*160*	*23235*	*23395*
2005	*109*	*24688*	*24797*
	1143	***134,013***	***135,156***

Table 1

Category	*Financial ratios*
Activity	· *Sales/Total Assets*
	· *Fixed Assets/Number of employees*
Leverage	· *Total Liabilities/Total Assets*
	· *Shareholders equity/Total Assets*
	· *Shareholders equity/Total Liabilities*
	· *Current Liabilities/Working Capital (1)*
Liquidity	· *Working Capital/Total Assets*
	· *Working Capital/Sales*
	· *Current assets/Current Liabilities*
	· *Operational Profit/Working Capital*
	· *Current assets/Current Liabilities*
	· *Current Liabilities/Total Liabilities*
Profitability	· *EBIT/Total Assets*
	· *EBIT/Sales*
	· *Financial Profit/Total Assets*
	· *Cash Flow (2) /Shareholder equity*
	· *Cash Flow/Sales*

(1)Working Capital = Stocks + Debtors - Creditors

(2)Cash Flow = Profit - Loss for the period + Depreciatiation

Table 2

Variable	*Description*
Macroeconomic [1]	· *Real GDP growth*
	· *Output gap*
	· *Short-term deposit interest rate*
Sector	· *Agricultural*
	· *Construction*
	· *Industry*
	· *Services*
Size	· *Number of employees*
	· *Log(Number of employees)*
	· *Number of subsidiaries*
	· *Log(Number of subsidiaries)*
Baseline hazard function (alternative parametric specifications)	· *Ln (t)*
	· t^2
	· t^3

(1) From the IMF's World Economic Outook (WEO)

Table 3

Sector	Distressed	Healthy	Freq. distresses
CONS	251	25,671	0.97%
IND	294	25,405	1.14%
SERV	598	82,937	0.72%
Total	1,143	134,013	0.85%

Table 4

Age	Distresses	Acc. Freq.
1	69	6.0%
2	145	18.7%
3	185	34.9%
4	159	48.8%
5	140	61.1%
6	123	71.8%
7	93	80.0%
8	78	86.8%
9	51	91.3%
10	46	95.3%
11	29	97.8%
12	22	99.7%
13	3	100.0%

Table 5

Category	*Financial ratios*
Activity	· *Sales/Total Assets*
Leverage	· *Total Liabilities/Total Assets*
Liquidity	· *Working Capital/Sales* · *Current Liabilities/Total Liabilities*
Profitability	· *EBIT/Total Assets* · *Financial Profit/Total Assets*

Table 6

Variable	*Mean*	*Median*	*Sd*	*Min*	*Max*
EBIT_TA	*0.06*	*0.05*	*0.16*	*-1.86*	*1.91*
FinPr_Assets	*-0.02*	*-0.01*	*0.03*	*-0.46*	*0.42*
TL_TA	*0.79*	*0.83*	*0.27*	*0.00*	*5.63*
S_TA	*2.20*	*1.81*	*1.69*	*0.00*	*13.48*
WK_Sales	*0.14*	*0.03*	*3.75*	*-198.40*	*194.06*
CurrL_TL	*0.81*	*0.91*	*0.23*	*0.10*	*1.00*
Employees	*20.29*	*8.00*	*126.04*	*1*	*9968*
Growth	*3.47*	*3.43*	*0.73*	*2.34*	*5.04*
Short_dep_r	*3.36*	*3.28*	*1.28*	*2.04*	*8.90*
t	*5.88*	*6.00*	*3.08*	*1*	*13*

Table 7

	Healthy		*Distressed*	
Variable	*Mean*	*Sd*	*Mean*	*Sd*
EBIT_TA	*0.07*	*0.16*	*-0.02*	*0.24*
FinPr_Assets	*-0.02*	*0.03*	*-0.04*	*0.05*
TL_TA	*0.79*	*0.27*	*0.98*	*0.35*
S_TA	*2.20*	*1.69*	*2.19*	*1.83*
WK_Sales	*0.15*	*3.75*	*-0.21*	*3.55*
CurrL_TL	*0.81*	*0.23*	*0.85*	*0.20*
Employees	*20.25*	*127.53*	*21.49*	*49.07*
Growth	*3.47*	*0.72*	*3.75*	*0.88*
Short_dep_r	*3.34*	*1.26*	*4.10*	*1.71*
t	*5.92*	*3.08*	*4.48*	*2.56*

Table 8

	EBIT_TA	FinPr_Assets	TL_TA	S_TA	WK_Sales	CurrL_TL	Employees	Growth	Short_dep_r	t
EBIT_TA	1									
FinPr_Assets	-0.070	1								
TL_TA	-0.466	-0.222	1							
S_TA	0.091	-0.112	0.062	1						
WK_Sales	0.015	0.020	-0.053	-0.039	1					
CurrL_TL	0.045	0.126	-0.063	0.279	-0.068	1				
Employees	-0.003	0.011	-0.025	-0.025	-0.007	-0.021	1			
Growth	0.016	-0.041	0.028	0.017	-0.004	0.021	-0.006	1		
Short_dep_r	-0.004	-0.128	0.041	0.031	-0.009	0.050	-0.014	0.203	1	
t	0.034	0.049	-0.183	-0.056	0.012	-0.057	0.015	-0.177	-0.300	1

Table 9

	F	Prob>F
Model	490.49	0.0000
EBIT_TA	427.39	0.0000
FinPr_Assets	1147.34	0.0000
TL_TA	320.82	0.0000
S_TA	88.48	0.0000
WK_Sales	5.78	0.0163
CurrL_TL	4.66	0.0308
Employees	269.61	0.0000
Growth	212.77	0.0000
Short_dep_r	480.76	0.0000
t	122.84	0.0000

Table 10

Table 11

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
EBIT_TA					*0.19909****	*0.19679****	*0.19393****	*0.20730****	*0.20561****	*0.20452****
					(-15.71)	*(-15.61)*	*(-15.71)*	*(-14.96)*	*(-14.99)*	*(-15.05)*
FinPr_Assets					*0.00002****	*0.00003****	*0.00003****	*0.00003****	*0.00004****	*0.00004****
					(-22.68)	*(-22.25)*	*(-21.31)*	*(-21.02)*	*(-20.24)*	*(-20.32)*
TL_TA					*2.23606****	*2.19218****	*2.18542****	*2.19764****	*2.24175****	*2.23931****
					(12.65)	*(12.41)*	*(12.22)*	*(11.90)*	*(12.10)*	*(12.08)*
Sales_TA					*0.81997****	*0.82049****	*0.82003****	*0.81818****	*0.82426****	*0.82434****
					(-9.381)	*(-9.181)*	*(-9.178)*	*(-9.274)*	*(-8.859)*	*(-8.848)*
WK_Sales					*0.98252****	*0.98129****	*0.98208****	*0.98258****	*0.98244****	*0.98245****
					(-3.614)	*(-3.737)*	*(-3.567)*	*(-3.423)*	*(-3.385)*	*(-3.385)*
CurrL_TL					*3.60608****	*3.76086****	*3.76008****	*3.65311****	*3.66074****	*3.64585****
					(8.691)	*(8.882)*	*(8.870)*	*(8.700)*	*(8.637)*	*(8.612)*
Log_Employees		*1.24387****	*1.27396****	*1.24966****		*1.23858****	*1.24625****	*1.27006****	*1.25427****	*1.25522****
		(8.875)	*(10.05)*	*(8.995)*		*(8.658)*	*(8.877)*	*(9.703)*	*(9.017)*	*(9.052)*
Growth	*1.18764****	*1.19154****	*1.12540****	*1.12101****			*1.19815****	*1.15292****	*1.14516****	*1.14001****
	(4.429)	*(4.500)*	*(3.062)*	*(2.961)*			*(4.627)*	*(3.646)*	*(3.474)*	*(3.393)*
Short_dep_r	*1.07305****	*1.08707****	*1.03002*	*1.02507*			*1.01738*	*0.98544*	*0.98001*	
	(3.204)	*(3.760)*	*(1.313)*	*(1.099)*			*(0.774)*	*(-0.647)*	*(-0.890)*	
t2			*0.99194****	*0.99200****				*0.99477****	*0.99475****	*0.99494****
			(-8.665)	*(-8.590)*				*(-5.678)*	*(-5.684)*	*(-5.637)*
Constr				*0.85294**					*0.78495****	*0.78747****
				(-1.844)					*(-2.757)*	*(-2.722)*
Serv				*0.69409****					*0.72237****	*0.72512****
				(-5.061)					*(-4.450)*	*(-4.405)*
Constant	*0.00364****	*0.00212****	*0.00404****	*0.00552****	*0.00168****	*0.00103****	*0.00051****	*0.00079****	*0.00103****	*0.00098****
	(-37.61)	*(-37.77)*	*(-31.95)*	*(-28.01)*	*(-45.83)*	*(-45.05)*	*(-35.77)*	*(-32.19)*	*(-29.84)*	*(-31.40)*
Observations	*135156*	*135156*	*135156*	*135156*	*135156*	*135156*	*135156*	*135156*	*135156*	*135156*
Log likelihood	*-6575.95*	*-6539.00*	*-6496.86*	*-6483.69*	*-5985.45*	*-5949.78*	*-5938.04*	*-5920.69*	*-5911.11*	*-5911.52*
Failures	*1143*	*1143*	*1143*	*1143*	*1143*	*1143*	*1143*	*1143*	*1143*	*1143*
LR chi2(#)		*73.89*	*84.28*	*26.33*		*71.34*	*23.46*	*34.72*	*19.14*	
Prob > chi2		*0.0000*	*0.0000*	*0.0000*		*0.0000*	*0.0000*	*0.0000*	*0.0001*	

z statistics in parentheses
**** p<0.01, ** p<0.05, * p<0.1*

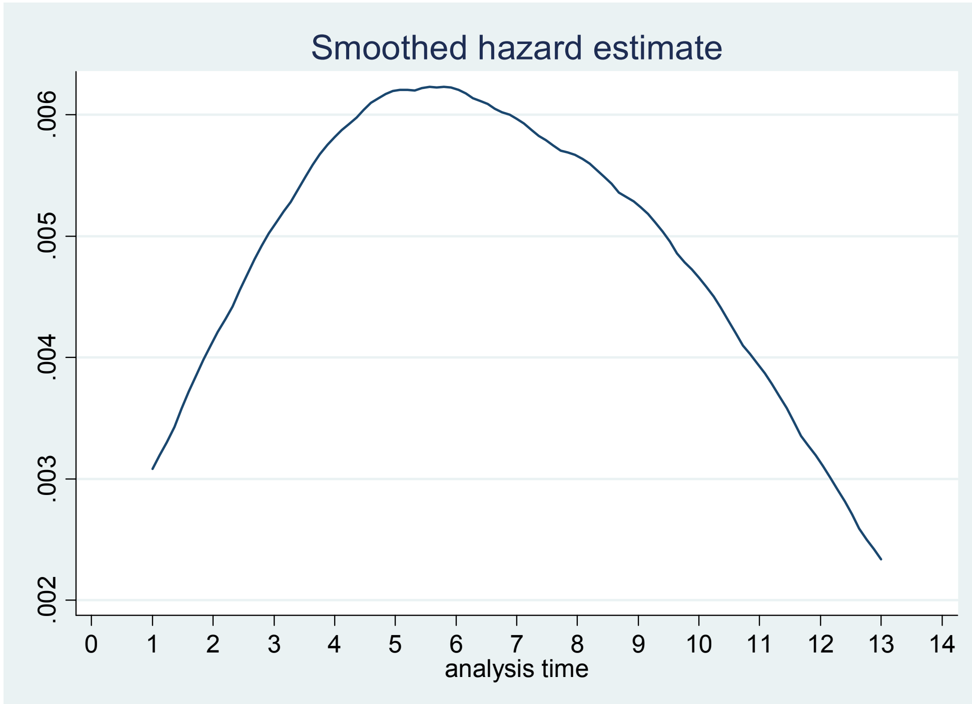

Figure 1

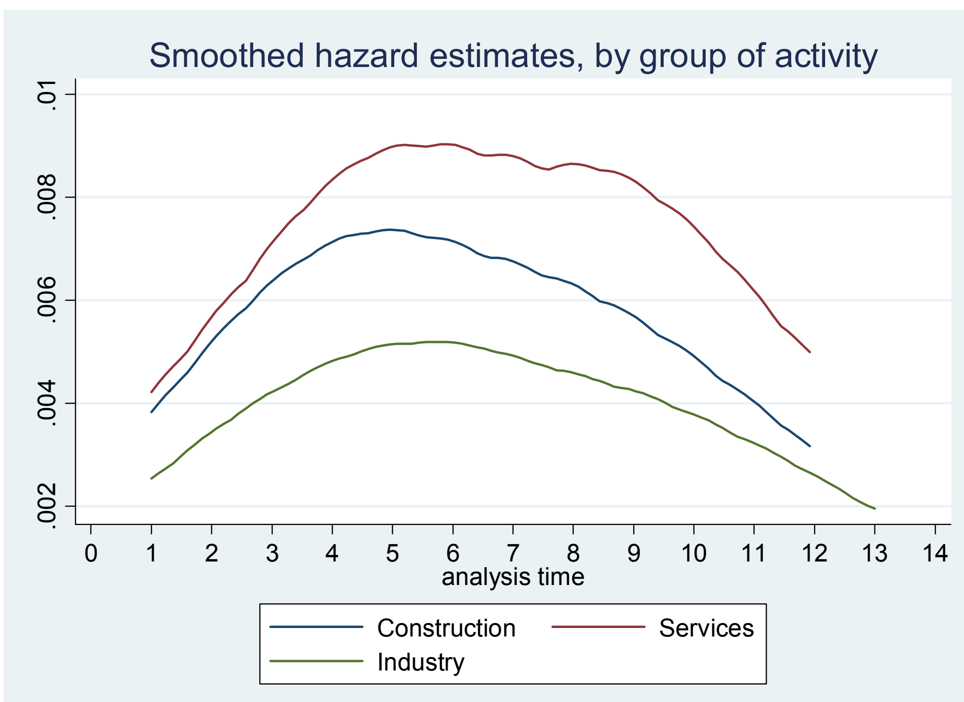

Figure 2

Table 12

	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
		t<=5					t>5				
EBIT_TA	0.20452***					0.25625***					0.10076***
	(-15.05)					(-9.708)					(-12.94)
FinPr_Assets	0.00004***					0.00008***					0.00001***
	(-20.32)					(-14.26)					(-14.81)
TL_TA	2.23931***					2.19835***					1.98353***
	(12.08)					(7.927)					(7.365)
Sales_TA	0.82434***					0.83408***					0.80182***
	(-8.848)					(-6.695)					(-5.834)
WK_Sales	0.98245***					0.98875					0.96857***
	(-3.385)					(-1.412)					(-4.778)
CurrL_TL	3.64585***					3.30310***					3.68459***
	(8.612)					(6.233)					(5.333)
Log_Employees	1.25522***		1.29235***	1.30118***	1.28310***	1.29446***		1.22484***	1.23188***	1.19747***	1.15801***
	(9.052)		(8.607)	(8.842)	(8.155)	(8.292)		(4.897)	(5.075)	(4.241)	(3.392)
Growth	1.14001***	1.14791***	1.14486***	1.15061***	1.14762***	1.15150***	1.09957	1.10478	1.03190	1.02730	1.07465
	(3.393)	(2.977)	(2.915)	(3.024)	(2.964)	(3.029)	(1.200)	(1.258)	(0.400)	(0.344)	(0.897)
Short_dep_r		1.01431	1.02248	1.01429	1.01195	0.97297	1.17740**	1.17939**	1.14884*	1.14516*	1.12500
		(0.587)	(0.912)	(0.579)	(0.483)	(-1.091)	(2.275)	(2.296)	(1.957)	(1.914)	(1.632)
t2	0.99494***			0.98995**	0.99047**	0.99867			0.99372***	0.99374***	0.99566***
	(-5.637)			(-2.158)	(-2.043)	(-0.282)			(-4.057)	(-4.040)	(-2.756)
Constr	0.78747***				0.95610	0.86246				0.71311**	0.76734*
	(-2.722)				(-0.404)	(-1.302)				(-2.429)	(-1.874)
Serv	0.72512***				0.78550**	0.81649**				0.58266***	0.61471***
	(-4.405)				(-2.526)	(-2.097)				(-4.882)	(-4.318)
Constant	0.00098***	0.00608***	0.00349***	0.00394***	0.00475***	0.00099***	0.00284***	0.00173***	0.00365***	0.00578***	0.00110***
	(-31.40)	(-26.76)	(-27.76)	(-26.33)	(-23.21)	(-23.55)	(-22.54)	(-22.58)	(-17.14)	(-15.08)	(-16.57)
Observations	135156	66485	66485	66485	66485	66485	68671	68671	68671	68671	68671
Log likelihood	-5911.52	-3870.10	-3836.25	-3833.90	-3829.82	-3544.68	-2679.93	-2668.57	-2659.73	-2648.41	-2339.99
Failures	1143	698	698	698	698	698	445	445	445	445	445
LR chi2(#)			67.69	4.70	8.17	570.26		22.71	17.69	22.63	616.84
Prob > chi2			0.0000	0.0302	0.0168	0.0000		0.0000	0.0000	0.0000	0.0000

z statistics in parentheses

**** p<0.01, ** p<0.05, * p<0.1*

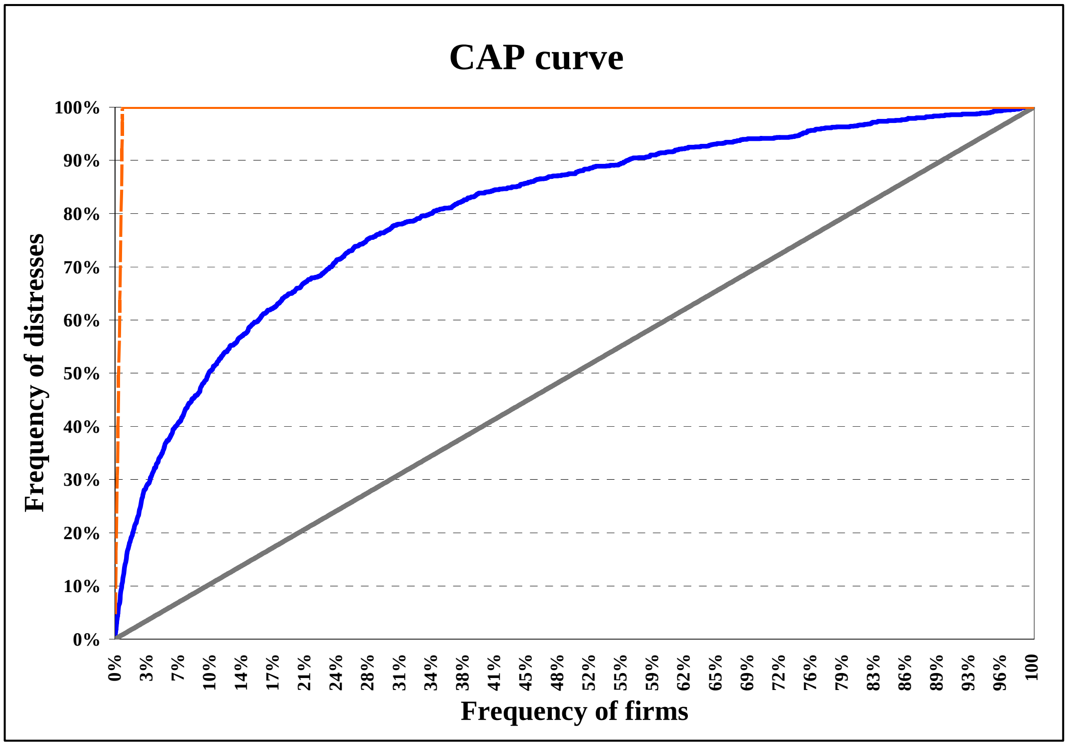

Figure 3

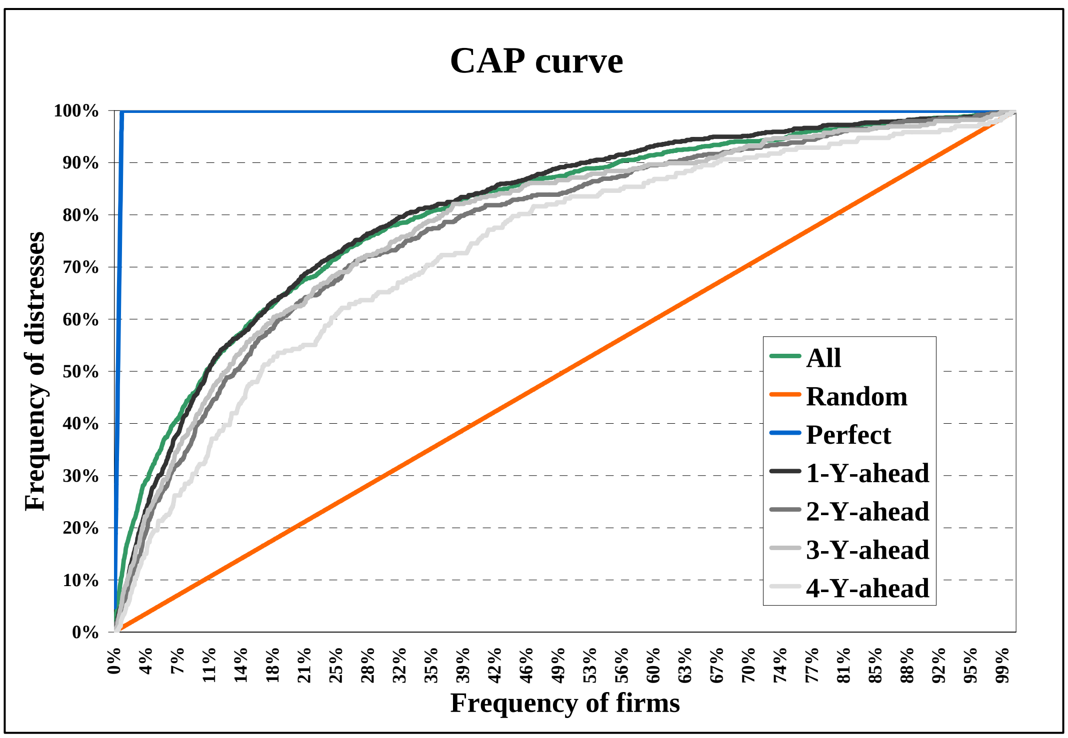

Figure 4

Lags	AR
All	61.2%
Lag1	63.4%
Lag2	56.1%
Lag3	57.6%
Lag4	47.4%

Table 13

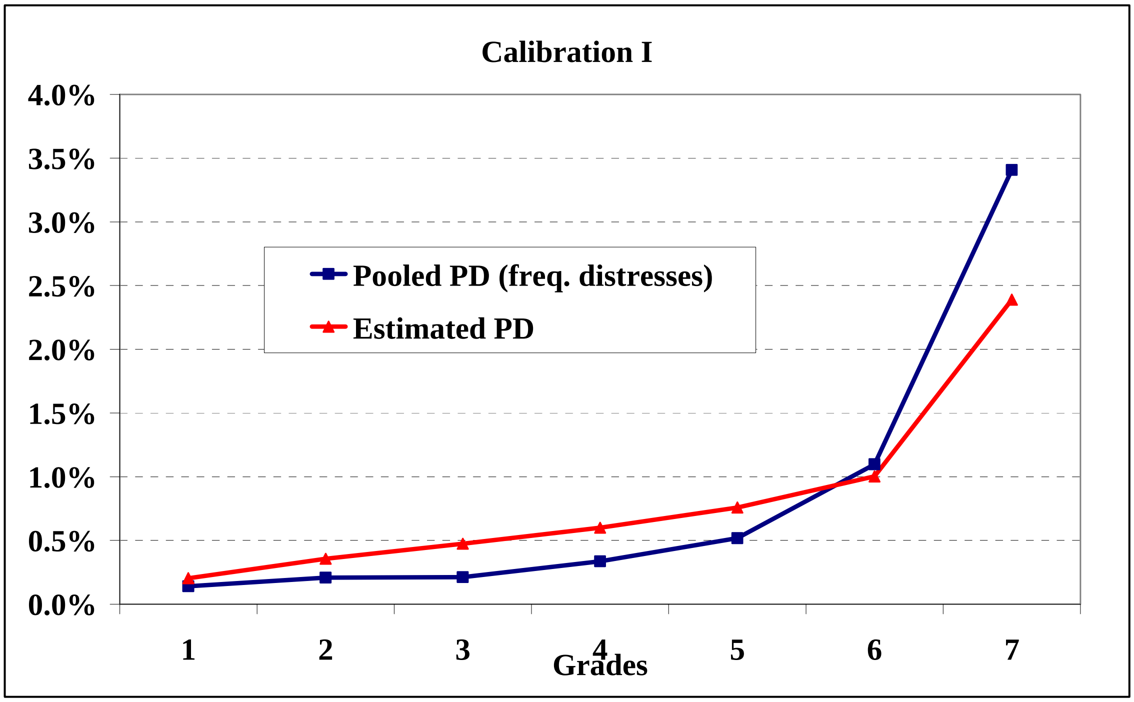

Figure 5

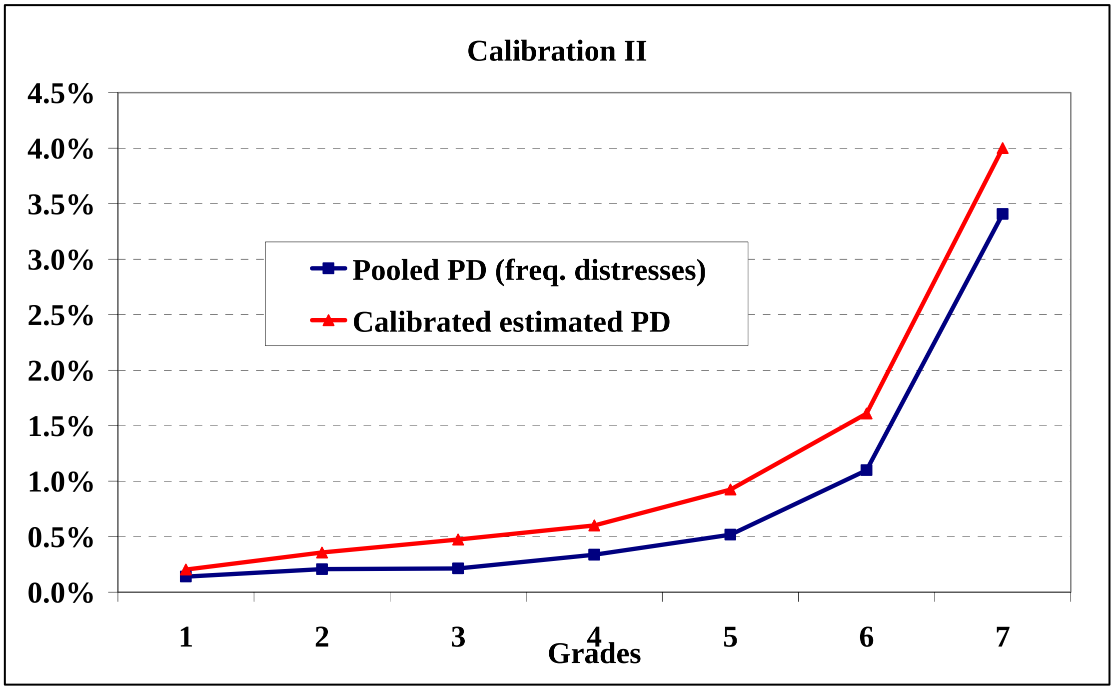

Figure 6

								Binomial Test					
								0.99		0.95		0.90	
Grade	Number of firms	Min PD	Max PD	Number of distresses	Pooled PD (freq. distresses)	PD	Pooled PD / PD	K		K		K	
1	19,304	0.00%	0.29%	27	0.14%	0.20%	0.69	54	OK	50	OK	47	OK
2	19,306	0.29%	0.42%	40	0.21%	0.36%	0.58	88	OK	82	OK	79	OK
3	19,308	0.42%	0.53%	41	0.21%	0.47%	0.45	114	OK	107	OK	104	OK
4	19,308	0.53%	0.67%	65	0.34%	0.60%	0.56	141	OK	134	OK	130	OK
5	19,309	0.67%	0.86%	100	0.52%	0.76%	0.68	175	OK	166	OK	162	OK
6	19,310	0.86%	1.19%	212	1.10%	1.00%	1.10	226	OK	216	OK	211	violation
7	19,311	1.19%	99.49%	658	3.41%	2.39%	1.43	511	violation	496	violation	489	violation

Table 14

Grade	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	**1994-2005**
1	0.00%	0.00%	0.73%	0.80%	0.56%	0.40%	0.51%	0.20%	0.04%	0.18%	0.14%	0.00%	**0.14%**
2	0.00%	0.00%	0.65%	0.00%	0.00%	1.16%	0.24%	0.16%	0.25%	0.09%	0.14%	0.23%	**0.21%**
3	0.00%	0.00%	0.29%	0.54%	0.15%	0.32%	0.37%	0.35%	0.49%	0.09%	0.08%	0.08%	**0.21%**
4	1.27%	0.63%	0.47%	0.38%	0.00%	0.45%	0.74%	0.27%	0.37%	0.33%	0.35%	0.14%	**0.34%**
5	0.00%	0.00%	0.63%	0.31%	0.38%	0.50%	0.60%	0.56%	0.66%	0.65%	0.65%	0.20%	**0.52%**
6	1.03%	0.94%	1.99%	1.20%	0.89%	0.95%	0.84%	1.35%	1.65%	1.13%	0.92%	0.81%	**1.10%**
7	0.00%	0.96%	2.82%	2.28%	2.78%	3.10%	2.86%	4.88%	4.75%	3.42%	4.31%	2.88%	**3.41%**
Average	0.33%	0.36%	1.08%	0.79%	0.68%	0.98%	0.88%	1.11%	1.17%	0.84%	0.94%	0.62%	**0.85%**

Table 15

Grade	Contributions												Average
	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	
1	0.08	0.13	6.79	7.80	4.45	2.48	5.76	0.40	2.14	0.34	0.00	7.21	*3.13*
2	0.12	0.19	2.91	0.60	0.87	26.59	0.03	0.20	0.24	2.30	0.85	0.08	*2.91*
3	0.18	0.30	0.09	1.89	0.13	0.51	1.33	1.72	10.36	2.27	2.92	3.33	*2.08*
4	2.03	0.40	0.22	0.03	2.58	0.41	7.33	0.27	0.09	0.00	0.03	3.80	*1.43*
5	0.62	1.15	0.11	0.51	0.38	0.01	0.27	0.07	0.98	0.91	1.08	5.89	*1.00*
6	0.00	0.08	3.67	0.08	0.50	0.35	1.44	1.32	6.26	0.03	0.79	2.07	*1.38*
7	4.41	9.48	0.75	4.21	1.87	0.60	2.58	14.70	10.20	0.00	5.15	1.82	*4.65*
T	*7.44*	*11.72*	*14.55*	*15.12*	*10.78*	*30.94*	*18.73*	*18.67*	*30.27*	*5.85*	*10.82*	*24.20*	*13.46*
P-value	0.4903	0.1639	0.0685	0.0568	0.2143	0.0001	0.0164	0.0167	0.0002	0.6638	0.2118	0.0021	

Table 16