Conceptual and Statistical Issues Regarding the Probability of Default and Modeling Default Risk

Emilia ȚIȚAN
Department of Statistics and Econometrics
Romanian Academy of Economic Studies
Romana Sq, no 6, 1st district, Bucharest, Romania
Email: emilia_titan@yahoo.com

Adela Ioana TUDOR
PhD Candidate, Romanian Academy of Economic Studies
Romana Sq, no 6, 1st district, Bucharest, Romania
Email: adela_lungu@yahoo.com

In today’s rapidly evolving financial markets, risk management offers different techniques in order to implement an efficient system against market risk. Probability of default (PD) is an essential part of business intelligence and customer relation management systems in the financial institutions. Recent studies indicates that underestimating this important component, and also the loss given default (LGD), might threaten the stability and smooth running of the financial markets. From the perspective of risk management, the result of predictive accuracy of the estimated probability of default is more valuable than the standard binary classification: credible or non credible clients. The Basle II Accord recognizes the methods of reducing credit risk and also PD and LGD as important components of advanced Internal Rating Based (IRB) approach.

Keywords: probability of default, stress test, PD buckets, pooled PDs, predictive analytics, data mining techniques, statistical methods, loss given default

1 Introduction
Exposure to financial markets affects most of the financial organizations because they are involved in the risk business: there is either a possibility of loss, or an opportunity for gain. Risk can be defined as the volatility of unexpected outcomes. The risks associated with the banking sector differ by the type of service rendered and they are classified into five types: systematic risk, credit risk, liquidity risk, operational risk and legal risk.

Systematic risk (market risk) is the risk of asset value change associated with systematic factors. We mention accordingly interest rate risk and the foreign exchange risk.

Credit risk arises when counterparties are unwilling or unable to fulfil their contractual obligations. Of course, there is a risk of involuntary default, when the borrower may not have enough money to pay the loan or strategic default, when the borrower may simply refuse to pay up. The effect is measured by the cost of replacing cash flows. The real risk from credit is the deviation of portfolio performance from its expected value.

Liquidity risk has two meanings. There is market liquidity when a transaction can not be performed at current market prices due to insufficient market activity, and also we can refer to funding risk, as the inability to meet cash flow obligations. In both cases, the liquidity risk can be managed by setting limits on certain markets, products or cash flow gaps as well.

The liquidity risk is an important counterparty variable, beyond credit rating. If an obligor has high liquidity, then the one-year PD will be lower because liquidity is more important in the short term. The all-important liquidity risk arises from a variety of sources and, if left
unchecked, it has the potential to damage a firm’s reputation.
Operational risk is associated with potential losses resulting from inadequate systems, management failure, faulty controls, fraud or human error.
Legal risks arise when new statutes, tax legislation, court opinions and regulations can put formerly transactions into contention. They include compliance and regulatory risks, which concern activities that might breach government regulations, such as market manipulation, insider trading and suitability restrictions.
The banking industry has long viewed the issue of risk management as the need of control the risks mentioned above, especially the credit risk. Understanding the various ways in which lenders manipulate and mitigate the default risk is the key to explaining some of the main features of credit markets.
Nonetheless, in a performant financial system, risk prediction is of great importance. This involves analytical processes and prediction models whose purpose is to use financial statements, customer transaction, repayment records and so on, in order to predict business performance or credit risk and to reduce the uncertainty and default. To forecast probability of default is a major challenge and it needs intense study.
In the next section, I have in view general considerations on default in the credit mechanism and several estimation methods. Section 3 refers to stress testing and stress probability of default. Section 4 is dedicated to data mining techniques which include statistical algorithms for PD evaluation and section 5 includes conclusions.

2. Default estimation and the role of PD as key factor
There is no standard definition of what ‘default’ means. Regulators and rating agencies define default as any of the following events: bankruptcy, write-down, 90 days past due loan or placement on internal non-accrual list. The obligor is considered defaulted as of the date of any of these accounting and financial failures. Originally, the Basel Committee suggested that, to ensure consistent estimation of credit risk across the banking industry and provide for data sources concerning default statistics, a default be defined as involving one or more of four criteria:

- It is determined that the obligor is unlikely to pay its debt obligations (principal, interest, or fees) in full.
- There is a charge-off.
- The obligor is overdue more than 90 days on any credit obligation.
- The obligor has filed for bankruptcy or similar protection from creditors.

Subsequently, these four criteria have been reduced to only two: more than 90 days overdue, and unlikely to pay in full.
The IRB method, according to Basel II, allows the banks to set the capital requirements for different exposures, using their own estimations for the credit risk components. The best estimate of exposure to the counterparty will depend on:

- Probability of Default (PD), which is the likelihood that a loan will not be repaid and fall into default. PDs are largely based on credit ratings, whether internal to the bank or by independent agencies; but there are also other factors. Liquidity risk and credit risk (and therefore PD) correlate. The PD is both influenced by and impacts on liquidity.
- Loss Given Default (LGD), which is the loss recorded by the bank (as a percentage of the exposure value) when the debtor is in default.
- Exposure at Default (EAD), which is the amount of money involved in the default process.
- Effective Maturity (M) of the credit instrument.
Using their own methodology for estimating these components of credit risk is subject to approval by the supervising authority, and in some cases, banks will have to use values provided by the supervisor.

A bank may use its own values for PD and/or LGD, only if a strict set of regulations are accomplished. It settles minimum requirements to be fulfilled in order to implement a risk management system based on credit ratings internally generated.

The principle underlying these requirements is that the rating and risk estimated systems and processes should provide a relevant assessment of the counterparty and transaction characteristics, a significant differentiation of risk and a reasonably and consistent accuracy of the quantitative estimates of risk.

In addition, the systems and processes must be consistent with internal use of these estimates.

Basel Committee, recognizing the differences between markets, rating methodologies, products and banking practices in various countries, let at the discretion of national supervising authorities the development of the necessary procedures for implementation of the internal rating system.

2.1. The calculation of minimum capital requirements

In the banking system, the main tasks of the capital are:

- Protection of the deponents in the event of bank insolvency and liquidation;
- Absorption of the unanticipated losses to maintain trust, so that under the stress conditions, the bank can continue to work;
- Purchasing of the buildings and equipment for operation;
- Serving as a limit for the undue expansion of assets.

The regulatory capital is associated to minimum capital requirements that banks are obliged to held under the regulation of surveillance from the perspective of regulatory institution. The aim of the capital requirements is to ensure the stability and viability of the banking system.

The minimum capital requirements consist of three elements:

1. The capital definition (unchanged towards Basel I Accord).
2. The definition of the weighted assets towards risk (RWA).
3. The ratio between the capital and RWA.

The bank must maintain capital equal to at least 8% of its risk-weighted assets. For example, if a bank has risk-weighted assets of $100 million, it is required to maintain capital of at least $8 million. So, the minimum capital requirements are calculated by multiplying the amount of the weighted assets depending on risk and the percentage of 8:

\[
\text{Capital} = \sum_{k} (\text{RWA}) \times 8\%
\]

RWA can be calculated based on two approaches: standard and internal rating. I will focus on the second method because it makes the purpose of the article. In this case, RWA is based on the four components mentioned in the beginning: probability of default, loss given default, exposure at default and effective maturity.

For the foundation IRB approach, only PD is calculated by the bank, the remaining components of risk being provided either by the Basel Committee on Banking Supervision or by national supervising institution. In case of advanced IRB approach, all four components of risk are calculated by the bank.

Based on these four keys, for each product, RWA is calculated. For a given exposure, RWA is as follows:

\[
\text{RWA} = 12.5 \times \text{EAD} \times \text{K}
\]
Conceptual and Statistical Issues Regarding the Probability of Default and Modeling Default Risk

where $K$ is the minimum capital for an exposure unit and it is calculated like:

$$K = \text{LGD} \left[ N \left( \frac{N^{-1}(PD) + \sqrt{R N^{-1}(0.999)}}{\sqrt{1-R}} \right) - PD \right] \text{MF}(M, PD),$$

where:

- $N()$ is the loss of the homogeneous portfolio with a probability of 99.9% and LGD of 100%. The loss is calculated based on a Merton method.

In the Merton approach to modelling credit risk, it is assumed that a default happens if the value of an obligor’s assets falls short of the value of debt. This provides financial analysts with the ability to forecast future, or implied, credit risk using information available at the current time. On a firm-by-firm basis, this is an important component of modern credit risk.

- $\text{LGD}[N() – PD]$ is the unexpected loss of the same portfolio.

- $R$ is the correlation coefficient between the assets (loans) for the same portfolio. $R$ was estimated by the Basel Committee as:

$$R = 0.12 \left( \frac{1 - e^{-50PD}}{1 - e^{-50}} \right) + 0.24 \left( \frac{1 - e^{-50PD}}{1 - e^{-50}} \right).$$

$R$ represents a decreasing function of PD and it has values between 12% and 24%. The debtors with a superior financial situation have a superior systemic risk towards the inferior quality debtors.

- $\text{MF}$ is the maturity function and it is:

$$\text{MF}(M, PD) = \frac{1 + (M - 2.5) \cdot b(PD)}{1 - 1.5b(PD)}$$

where:

$$b(PD) = [0.11852 - 0.05478 \cdot \log(PD)]^2$$

The MF function was obtained by the Basel Committee on Banking Supervision by calibration and it equals 1 for one year maturity.

As shown by Kiefer and Larson (2007), in case of bank exposures, the risk weight curve is generally a concave function in PD (see the graph below).

![Risk Weight Function](image)

Source: Biases in Default Estimation and Capital Allocations under Basel II; Kiefer, Larson, 2007

Fig.1. The risk weight function

A calculation of the second derivative which is negative, proves the concavity and its meaning is explained in the next section.

The limits of PD are 0 and 1. The problem is often that ratings are not sufficiently responsive to changes in economic cycles, resulting in a certain overestimation or underestimation of likelihood of default over different periods. The value of 1 implies that the lender will recover all money in case of default by the counterparty, whereas the lender will recover nothing with 0. LGD is also bound between 0 and 1 and its measurement is not linear.

### 2.2. Estimating probability of default through probability of default buckets

The evaluation of the debtors is made through statistical models on an individual basis, by assigning individual PDs and/or individual scores. Obligors with similar
PDs/scores are then grouped into rating classes, or buckets (“PD-buckets”). Under Basel II, an IRB bank must assign obligors to risk buckets. Credit quality is the main principle that states at the basis of each bucket and it is defined by a mean value and a variance in credit standing. The Basel Accord then requires that all obligors falling into the same bucket be assigned the same “pooled” PD (which can be thought as the mean of individual PDs). In this case, capital charges are calculated. Related to the variance in credit standing, which is bound to exist with every pool, all banks have an interest in establishing a level of confidence at 99.9 per cent. Banks with experience in the implementation of Basel II rules suggest that, to apply the method of PD buckets properly, user organizations should provide themselves with the means to continue drawing a distinction between the concepts of:
- A default probability linked to an individual obligor, and
- The pooled PD assigned to a credit risk bucket.

A PD associated with an individual obligor is a metric of the probability that this obligor will default during a one-year credit assessment. By contrast, the pooled PD assigned to a risk bucket is a measure of the average value of the PDs of obligors in that bucket. Related questions have been addressed in the literature:
- How should pooled PDs be derived that reflect the PDs of obligors assigned to each risk bucket in an accurate manner?
- How should deviations from the pool’s mean value be accounted for and presented?
- How should PD bucket mean values and variances in credit risk, among individual pool members, be stress tested?

There is no clear guideline on how this pooled PD could or should be stress tested. (Dimitris N. Chorafas, 2007)

PD is a continuous variable, taking values between 0 and 1, so there are infinitely many possible ways to partition the 0-1 interval into a set of discrete intervals (the PD-buckets). The choice of the “optimal” buckets (sometimes referred to as “PD bucketing”) is rarely reached analytically by banks. Most of the times, banks offer a defining label of the rating buckets like “very good” or “AAA” and a set of rating criteria which help their analysts to sort obligors into different classes.

Therefore, the buckets should be chosen carefully, since all obligors falling into a given rating class will eventually be assigned the same PD. As individual PDs within a bucket are expected to be similar, but not exactly equal to one another, replacing them with a pooled PD obviously causes a loss of precision in the rating system.

On the other hand, the buckets must contain a high number of observations in order to have a precise assessment. In this case, “the concavity of the risk weight curve (see fig. no. 1) means that, if two rating classes (having pooled PDs of \( p-k \) and \( p+k \)), containing an equal number of obligors, are pooled together into a single bucket with an average PD of \( p \), the new capital charge \( C(p) \) will be more than the sum of the capital charges on the two separate classes:

\[
C(p) > \frac{1}{2} C(p-k) + \frac{1}{2} C(p+k)
\]

Second, the use of pooled PDs instead of individual ones could cause opportunistic behaviour and adverse selection phenomena among a bank’s customers. Indeed, if a single rating bucket were to include a very wide array of individual PDs, all replaced by the same pooled PD and treated as equally risky for pricing and risk management purposes, then the best customers in the bucket would feel they can get substantially lower lending rates elsewhere, while the worst customers in the bucket would stay as they feel their credit risk is significantly underpriced.” (T. Krink, 2008)
Basel Committee argues that the default probability assigned to each debtor depends strongly on the type of rating methodology and quantification techniques being used. There are several important approaches for quantifying pooled PDs. The significant ones include the historical default method, statistical model approach and external mapping.

The actuarial approach or the historical default method consists of recording the default events over several years assigned to the specific bucket. The algorithm is:

$$\text{DF}_t = \frac{D_t}{N_t}$$  , where  \( \text{DF}_t \) = default frequency;
\( D_t \) = number of defaults observed for a bucket over year t;
\( N_t \) = number of debtors assigned to that bucket at the beginning of year t.

In order to estimate a default probability for each obligor assigned to a bucket, there are used predictive statistical methods. Therefore, the bucket’s pooled PD is then calculated as the median of obligor PD. This approach to individually quantifying pooled PDs can produce accurate estimates of credit exposure, gaining an important advantage over the historical default alternative.

Within the external mapping, a bank simply establishes a connection between its internal rating system and an external scale such as that of big rating agencies, calculates a pooled PD for each external grade using an external reference dataset, and then assigns the pooled PD for the external grade to its internal grade by means of the mapping. Despite its apparent simplicity, this approach poses some difficult validation challenges for risk managers. To validate the accuracy of a bank’s pooled PDs, supervisors and risk managers must first confirm the accuracy of the pooled PDs associated with the external rating scale. They must then validate the accuracy of the bank’s mapping between internal and external grades. Quantifying pooled PDs for an external rating system poses the same estimation problems as quantifying pooled PDs for a bank’s internal rating system. If a historical default experience approach is used, supervisors and risk managers must check to ensure that each bucket’s pooled PD can be expected to approach its long-run default frequency over time. If a statistical models approach is used, supervisors and risk managers must validate the reliability of the underlying default prediction model. The main benefit of quantifying PDs using external ratings is that more data are likely to be available for calculating long-run default frequencies and/or estimating statistical default prediction models.

### 3. Stress testing

Stress tests are an important risk management tool that has been used for a number of years now, both by banks as part of their internal risk management practices and by supervisors to assess the resilience of banks and of financial systems in general to possible shocks (European Central Bank, 2010). This method is also called scenario analysis and it consists of specific scenarios of interest in order to assess possible changes in the value of the portfolio.

In my opinion, the key role of the stress tests is to draw attention of how much capital might be needed to absorb losses in case of a financial crisis or other shocks and therefore increase the banks resistance in recession times. The importance of these tests is bigger in a stable economy because, due to the fact that there are no special risks, the banks might not be aware of the major impact of a financial crisis upon their stability. Practically, stress testing forces management to consider events that they might otherwise ignore.
Fig. 2 shows a stressed and unstressed probability of default and it is based on the study by the Basel Committee on Banking Supervision in February 2005. In boom times, the stressed probability of default acts like a prevention of the unstressed probability of default in case of financial crisis. This way, it would be much easier to determine the value of the collateral that should be asked for.

It is quite obviously that SPDs tend to remain relatively stable over a business cycle compared with unstressed PDs. During economic expansion the unstressed PD declines and the obligor receives a higher rating; but during economic recession the unstressed PD increases, closely approximating stressed PD, and the obligor receives a lower rating. Considering the stress scenarios associated with obligor-specific PDs, I believe that the economic environment is not sufficient in order to offer a pertinent result regarding the creditworthiness of the debtor. The tests should also contain both information relevant to assessing the obligor’s ability and willingness to repay its debts, and macroeconomic variables (interest rate levels, market liquidity, inflation rates etc).

Basel Committee gives the following definitions for the classic PD and SPD:

- An *unstressed PD* is an unbiased estimate of the likelihood that an obligor will default over the next year, given all currently available information, including static and dynamic.

- A *stressed PD* (SPD) measures the likelihood that an obligor will default over the next year, using all available information, but assuming adverse economic and lender-specific conditions for the stress scenario.

![Fig 2. Unstressed and stressed probability of default, over time (time is expressed in years). Macroeconomic variables include a growth or downturn in gross domestic product, exchange rates and market psychology. Source: based on a study by the Basel Committee on Banking Supervision](image)

But, the choice of scenarios may be affected by the portfolio position itself. For instance, one month the portfolio may be invested in a national fixed-income market; the scenario will then focus on interest rate shifts in this market. The following month, the portfolio may be invested mainly in currencies. If scenarios change over time, measures of risk will change just because of these changes. Also, stress testing does not specify the likelihood of worst case scenarios. Expected risk should be a function not only of the losses but also of the probability of such losses to occur.

The stress tests implemented by the banks have registered some deficiencies lately. The amplitude and the severe current financial crisis has determined many banking institutions and supervising authorities ask if the stress tests used before this crisis were quite efficient and helped the banking sector to face this real challenge. The financial crisis showed several lacks in the stress tests systems of the banks.
especially regarding the crisis scenarios and the methodologies used for crisis simulation. In many banks, the stress tests were done only for specific activities or risks, without being considered an aggregation of results on the overall bank. Another issue is that most of the risk management methods, including stress simulations, use statistic data in order to assess the future exposures at risk. These data are based on long periods of economic stability and are not sufficient to identify a crisis. The banks underestimated the strong correlation between the lack of liquidities on the market and the financing pressure. Therefore, it is crucial to treat correctly the dependencies between different risks and integrate them on the overall financial group or bank.

4. Data mining techniques used for predictive default probability

In our days, data mining is an indispensable tool in decision supporting system and it is defined as “the process that uses statistical, mathematical, artificial intelligence and machine-learning techniques to extract and identify useful information and subsequently gain knowledge from large databases” (Turban, Aronson, 2007). Practically, data mining is a technique for extracting knowledge from information. This analysis process has a significant role in probability of default estimation, credit scoring, customer services, fraud detection and market segmentation.

The most important techniques are: discriminant analysis, logistic regression, artificial neural networks and K-nearest neighbour model.

4.1. Discriminant analysis (DA)

DA or Fisher’s rule is is a classification method that projects n-dimensional data onto a line, and performs classification in this one dimensional space. The projection is chosen so as to maximize the between-class mean, and minimize the within-class variance (R. Khemchandani, 2009).

In DA, a group of observations are used to measure parameter estimates of a discriminant function by minimizing the group misclassifications. This method is used in the decisional situations. For instance, DA provides data regarding the possibility of a loan application to default.

4.2. Logistic regression (LR)

A LR model specifies that an appropriate function of the fitted probability of the event is a linear function of the observed values of the available explanatory variables. The major advantage of this approach is that it can produce a simple probabilistic formula of classification. (I-Chang Yeh, 2009)

This is a special case of linear regression models.

In logistic regression, there is no definition of the coefficient of determination ($R^2$) which is frequently used in the general linear model. $R^2$ has the desired interpretability as the proportion of variation of the dependent variable, which can be explained by the predictor variables of a given regression model.

4.3. Artificial neural networks

Artificial neural networks (ANN) are flexible computing frameworks and universal approximations that can be applied to a wide range of time series forecasting problems offering solutions in many fields, such as control and pattern recognition. ANN use non-linear mathematical equations in order to develop adequate correlations between input and output variables.

One of the major developments in neural networks over the last decade is the model combining or ensemble modelling. The basic idea of this multi-model approach is the use of each component model’s unique capability to better capture different patterns in the data. Both theoretical and empirical findings have suggested that combining different models can be an
effective way to improve the predictive performance of each individual model, especially when the models in the ensemble are quite different (Baxt, 1992; Zhang, 2007).

4.4. K-nearest neighbor model (KNN)
Nearest-neighbour (NN) techniques are non-parametric classification systems based on learning by analogy. Given an unknown sample, a KNN classifier searches the pattern space for the KNN that are closest to the unknown sample. This means finding out the shortest distance. In learning systems, generalisation performance is affected by a trade-off between the number of training examples and the capacity (e.g. the number of parameters) of the learning machine. The major advantage is that it is not required to establish predictive model before classification.

Empirical studies in literature outline that in the predictive accuracy of probability of default, artificial neural networks show the best performance based on $R^2$, regression intercept and regression coefficient. Therefore, ANN should be employed to score clients instead of other data mining techniques, such as logistic regression. (I-Cheng Yeh, 2009).

5. Conclusions
The management of financial risks has many dimensions and involves many types of decisions. The importance of this article comes from the complex issue of credit risk management in order to assure financial stability. Credit risk is considered the most dangerous category of banking risk and in order to prevent it, banks must meet a series of regulations. Recent studies show that default probabilities and average recovery rates are negatively correlated (see e.g. Altman et al. (2005); Acharya et al. (2007)). Both variables also seem to be driven by the same common factor that is persistent over time and clearly related to the business cycle: in recessions or industry downturns, default rates are high and recovery rates are low. Although the actual researches in this area are very elaborated, both in Romanian and foreign literature, through empirical and theoretical studies for loss predicting in case of default or sofisticated credit scoring models, at present the international financial crisis has revealed serious shortcomings and limitations in managing credit risk. Therefore, I believe a progressive research is required in terms of effects generated by the current international crisis.

Acknowledgements
This article is a result of the project POSDRU/88/1.5./S/55287 „Doctoral Programme in Economics at European Knowledge Standards (DOESEC)“. This project is co-funded by the European Social Fund through The Sectorial Operational Programme for Human Resources Development 2007-2013, coordinated by The Bucharest Academy of Economic Studies in partnership with West University of Timisoara.

References

Mrs. Emilia Ţiţan received the PhD title in 2005 and at present, she is professor at the Department of Statistics and Econometrics, in the Academy of Economic Studies, Bucharest. Since 2008, she has been a deputy dean of Cybernetics, Statistics and Economic Informatics faculty of the Academy of Economic Studies. Her research activity is found in over 50 articles and 20 books.

Adela Ioana Tudor has graduated the Faculty of Cybernetics, Statistics and Economic Informatics of the Bucharest Academy of Economic Studies in 2002. Two years later she graduated the MA in Financial Management and Capital Markets – DAFI, Faculty of Finance, Insurance, Banking and Stock Exchange Market. At present, she is a PhD student in the field of Economic Cybernetics and Statistics at the Academy of Economic Studies.