# Loss Coverage and Stress Testing Mortgage Portfolios: A Non-Parametric Approach

Adolfo Rodríguez

**Banco de España (Bank of Spain)** [*]

Carlos Trucharte

**Banco de España (Bank of Spain)** [*]

## Abstract

In this paper we outline the development of a practical approach to simulating a credit loss distribution function and to implementing a stress test exercise. This approach focuses on what is currently one of the banks' key loan portfolios: the mortgage portfolio, in particular, the entire Spanish mortgage portfolio. Specifically, we first determine, via regression model, the main factors that explain why households fail to meet their mortgage payment commitments. This allows us to consistently assign individual borrowers' PDs and to estimate a rating system for the mortgage portfolio of the entire credit system. Then, we simulate the empirical distribution function of mortgage loss rates for a whole economic cycle using a Monte-Carlo resampling method, and compare the loss rates from this function with those provided by the Basel II IRB formulas. Finally, we assess, by running a stress exercise, the ability of banks to withstand certain adverse situations. The main result from this exercise is that, in general terms, Basel II IRB regulatory loss coverage offers fairly adequate protection for banks. All in all, this paper shows the usefulness and importance of stress and simulation tools for supervisory authorities to analyse and properly define banks' risk profile and loss protection adequacy measures to account for possible future losses.

Key words: loss coverage, stress test, mortgage rating structure, Monte-Carlo simulation, Basel II.

**1. Introduction**

The crucial role that banks play within the financial system (credit granting function) is certainly not free from risks. When these risks materialize they become effective losses. Consequently, a thorough measurement of loan losses and an adequate assessment of the features of their distribution to properly estimate protection and coverage against them seem desirable.

To assess the nature and impact of credit losses, stress tests are proving themselves as a highly useful tool in hands of the financial authorities. Their development and progressive implementation as a prudential technique complementing traditional supervisory practices are making them increasingly valuable to supervisory authorities in monitoring and protecting the stability of the financial environment[1]. In addition, stress tests have become similarly valuable for individual banks[2]. In their search for better ways to control their risks (management and measurement) and to define their risk profile more precisely, banks have started to subject their different business areas to stress tests. This initiative has as a main aim to determine, first, the possible impact of shocks liable to affect credit institutions' financial condition; second, how sensitive their various types of risk, or key portfolios, are to these adverse shocks; and third, whether banks' defences are adequate to absorb the losses entailed by those negative scenarios.

A more thoughtful analysis of credit loss distributions and a more widespread application of stress tools may make it possible for credit institutions to understand better the consequences of possible future events and, in particular, to assess their impact more accurately. In this way, banks will be able to anticipate and identify potential vulnerabilities in their core business, and consequently implement countermeasures that could help them to weather unfavourable economic situations that otherwise could severely affect their financial condition. In the medium term, this will also have implications for safeguarding the interests of shareholders and depositors, as banks will try to minimise the negative consequences of plausible negative scenarios. This could be achieved by either reducing their exposure to such scenarios (by putting in practice specific mitigation techniques or structured portfolio hedging strategies), or by building up enough solvency defences to withstand the potential losses that could emerge from a harsh economic context.

Therefore, a proper evaluation of loss protection measures and the implementation of stress tests should be seen as a way of approximating to the more important banking risks so that an accurate assessment of how prospective risks, whatever their origin, may influence and alter the stability and financial condition of credit institutions. Hence stress tests applied to portfolio loss distributions should add value to the internal control exercised by banks in the course of their risk management process, serve as a basis for

---

[1] A similar opinion can be found, for example, in Intrater (2002), Bunn et al. (2005), Caruana (2005), FSA (2005) and López (2005). For a general discussion on stress tests, see Blaschke et al. (2001).

[2] Evidence of that is presented in the survey carried out by the Committee on the Global Financial System, CGFS (2005).

fostering prudential techniques of protection against adverse situations, and facilitate prevention and early warning and response measures to deal with these adverse situations.

An additional consideration on stress exercises regards banks' solvency measures, in particular their adequacy. The revised structure for the international convergence of capital measurement and capital standards (more commonly known as Basel II)[3] proposes the performance of stress tests to determine whether minimum regulatory capital set under Pillar 1 can be considered adequate to banks' risks. Under this perspective, stress tests are considered as an additional prerequisite for banks that opt for the more advanced approaches. Moreover, stress exercises will take on particular importance, as supervisors will also get involved in assessing their design, calibration and impact, as stipulated in Pillar 2 of Basel II.

At first sight it seems clear that the aim of Basel II in the performance of stress tests under Pillar 2 is to determine whether, given a certain scenario, regulatory capital requirements set by the advanced approach (Internal Ratings-Based Approach, IRB) plus the capital buffer that banks may hold, constitute sufficient defences to absorb a possible adverse shock to a bank's credit exposure. However, Basel II also observes that it is advisable to carry out other more specific tests to assess the effect that a given scenario may have on the calculation of regulatory capital (Pillar 1), or, in other words, to stress the conditions and parameters that determine the required minimum solvency resources[4].

Bearing all the aforementioned in mind, this paper intends to obtain the distribution function of a particular credit portfolio and then, exemplify how a practical stress test exercise may be implemented and the consequences that can be extracted from it. We show, first, which the main determinants of credit risk drivers are. Second, how an empirical distribution of credit losses can be attained. Third, how this distribution function is modified when running a stress exercise, and, finally we address the adequacy in the amount of prudential resources required for banks to cover, with a certain probability, estimated credit losses. All this will be done by focusing on what is currently one of banks' key loan portfolios: the mortgage portfolio. In particular, the scope of application of our analysis will be the whole Spanish mortgage credit system.

In general, mortgage markets have been widely studied. Allen (2004) surveys this literature as well as more general Basel II issues (i.e. procyclicality, securitization, and capital arbitrage). Calem and LaCour-Little (2004) simulate economic capital for a mortgage portfolio with the aim of obtaining evidence of significant divergence between economic and regulatory capital. Whitley et al. (2004) estimate an empirical model of household arrears, reflecting the importance of both macro and micro (loan characteristics) variables as determinants of the going-into-arrears situation of borrowers. Dimou et al. (2005) calibrate a simple simulation model based on the relationship between the state of the economy,

---

[3] See Basel Committee on Banking Supervision, BCBS (2006).

[4] A detailed document (consultation paper under the supervisory review process) on stress testing for the assessment of capital adequacy has been produced by the Committee of European Banking Supervisors, CEBS (2006).

residential property values and the rate of loan default and suggest that Basel II IRB capital calculations may fall short of prudential capital requirements.

However, not much has been done to date on loss coverage analysis and stress tests of mortgage portfolios. In particular, Coleman et al. (2005) outline a case study for developing stress tests of housing loan portfolios particularised for the Australian case. Our paper diverges from it in various respects: first, the credit loss equation model. Second, the determinants of the credit risk parameters. Third, the simulation of a distribution of mortgage losses and, finally, in the implementation of the stress test to address whether or not regulatory protection measures (Basel II requirements) for credit losses can be considered to be adequate.

In this paper, we outline the development of a practical model for a systematic approach to stress testing: assessing potential risks and evaluating their impact on the financial condition of credit institutions. More precisely, the objectives of this paper are twofold: first, to determine, via regression model, the main factors, both systematic (macroeconomic variables) and idiosyncratic (mortgage loan characteristics) that explain why households fail to meet their mortgage payment commitments. Second, to simulate a loss distribution of mortgage loss rates and compare, under different economic conditions (normal and stressed), the capability of banks to withstand those adverse situations.

In short, we first estimate a regression model which, in turn, classifies borrowers according to their credit quality. It incorporates several of the elements that should be taken into account when measuring the creditworthiness of each obligor: both risk profile characteristics (idiosyncratic elements) and macroeconomic variables as a measure of the systematic component of the latent risk of a credit portfolio.

Then, we calculate a distribution of credit losses using a Monte Carlo re-sampling method (following Carey, 1998 and Carey, 2001) which provides us with an acceptable non-parametric estimate of that distribution, and allows us to avoid any assumptions about certain risk parameters such as loss and asset correlations. Bearing in mind that we develop our model from actual banks' mortgage portfolios, default correlations are implicit in the data. Moreover, given the duration of the database used (1990 to 2004, including observations for both recession and boom periods) we can be sure that different realizations of the systematic factor take place which, to some extent, mitigates the problem of tail estimates when the effects of depression years are ignored. Additionally, we obtain insightful conclusions by stressing certain elements that determine the mortgage credit loss distribution. Finally, in light of the estimated loss distribution function, we run a stress exercise to assess the adequacy of regulatory loss protection measures required for credit institutions.

In this vein, the rest of the paper is structured as follows. In the next section, background information on the Spanish domestic credit market is provided, as well as some of the most relevant housing market and general economic indicators. The third section contains a brief description of the database used, while in the fourth section we present the methodology used to develop a model to estimate the probability of

default of households, including robustness and performance checks of the proposed model. Section five is dedicated to simulate the empirical distribution function of mortgage losses. Additionally, the main results and conclusions from both the simulation process and the stress exercise are also shown in this section. Finally, a summary of the main ideas presented in this paper concludes our work.

## 2. Background information on the Spanish credit and housing market

The Spanish economy has been experiencing an intensive economic upswing that began in the mid-1990s and still persists today. This process has been fuelled, among other things, by highly dynamic domestic demand which has been heavily dependent on consumption and construction investment. This impulse has been sustained by the persistence of very favourable monetary and economic conditions permitting important growth in financing to households and non-financial corporations. Private-sector borrowing expanded during the previous and current decades at a very high rate, showing signs of significant acceleration in the case of households. The buoyancy of lending for property activities has definitely played a major role in the expansion of the Spanish economy and has also contributed to sustaining continuous and significant house-price rises.

As a feedback effect, this growth in house prices has continued to reactivate and re-stimulate the increase in household financing. As can been seen in Table 1, in recent years, the pattern of bank credit exposures to individuals has been significantly altered, showing that the importance of credit to them has increased relative to that to firms.

### TABLE 1. DISTRIBUTION OF DOMESTIC CREDIT EXPOSURES

| Year | Total Credit Exposures | By type of borrower | | | Mortgage Loans | |
|------|------------------------|---------------|---------|---------|----------|---------------|
| | € million | % Individuals | % Firms | % Other | % Total | % Individuals |
| 1992 | 225,465 | 30.3% | 64.9% | 4.8% | 17.8% | 58.8% |
| 1993 | 232,368 | 34.2% | 61.8% | 4.0% | 21.2% | 61.9% |
| 1994 | 242,871 | 36.8% | 58.9% | 4.3% | 24.3% | 65.9% |
| 1995 | 259,084 | 38.4% | 58.0% | 3.6% | 25.8% | 67.3% |
| 1996 | 279,470 | 39.6% | 56.8% | 3.6% | 27.1% | 68.5% |
| 1997 | 319,871 | 41.6% | 55.9% | 2.5% | 29.7% | 71.4% |
| 1998 | 374,897 | 42.4% | 54.9% | 2.7% | 29.4% | 69.3% |
| 1999 | 448,139 | 44.1% | 53.5% | 2.5% | 30.3% | 68.8% |
| 2000 | 526,625 | 44.3% | 53.3% | 2.4% | 31.6% | 71.4% |
| 2001 | 586,010 | 45.9% | 52.0% | 2.0% | 33.1% | 72.1% |
| 2002 | 662,267 | 46.3% | 51.8% | 1.9% | 33.4% | 72.3% |
| 2003 | 761,927 | 46.7% | 51.0% | 2.4% | 33.9% | 72.7% |
| 2004 | 900,633 | 46.9% | 50.8% | 2.3% | 34.6% | 73.9% |
| 2005 | 1,147,480 | 48.2% | 49.9% | 1.9% | 36.5% | 75.6% |

Using banking data on domestic financing exposures to the private sector, Table 1 presents the distribution of bank credit in Spain. It shows the proportion of finance granted to the private sector corresponding to individuals and to firms. Clearly, the weight of the credit granted to firms has decreased

in recent years, from 65% to less than 50%, while that to individuals has significantly increased from 30% to almost the other 50%. This shift in portfolio composition reveals the increasing importance that household financing has been acquiring for credit institutions. Complementing the previous reasoning, the second to last column of Table 1 shows how the mortgage segment has gained momentum as a proportion of total credit exposure, especially in recent years. The final column reveals that 75% of total credit to households in 2005 was dedicated to the property market compared to less than 60% in 1992.

This rapid and important growth in housing credit has been subsequently accompanied by a sharp rise in households' indebtedness ratios (household credit to disposable income). Furthermore, as stated above, house prices have recorded historical highs in recent years propelled by all-time interest rate lows. Table 2 shows all these facts providing a clear idea of the context in which the Spanish economy, in particular the housing and the household sector, are currently immersed.

**TABLE 2. MAIN HOUSING MARKET ECONOMIC INDICATORS**

| Year | GDP growth rate | Interest rate | House prices growth rate | Households Indebtness ratio | Total Credit growth rate | Mortgage Loans growth rate | Mortgage Loans PD |
|------|------|------|------|------|------|------|------|
| 1992 | 0.9% | 13.3% | -1.2% | 45.3% | -- | -- | 4.9% |
| 1993 | -1.0% | 11.7% | -0.4% | 45.4% | 3.1% | 22.5% | 5.8% |
| 1994 | 2.4% | 8.0% | 0.7% | 45.7% | 4.5% | 19.9% | 4.1% |
| 1995 | 2.8% | 9.4% | 3.5% | 46.1% | 6.7% | 13.6% | 1.9% |
| 1996 | 2.4% | 7.5% | 1.9% | 46.9% | 7.9% | 13.3% | 1.6% |
| 1997 | 4.0% | 5.4% | 1.6% | 50.1% | 14.5% | 25.3% | 1.0% |
| 1998 | 4.3% | 4.3% | 4.6% | 55.6% | 17.2% | 15.9% | 0.7% |
| 1999 | 4.2% | 2.9% | 10.5% | 63.1% | 19.5% | 23.3% | 0.9% |
| 2000 | 4.4% | 4.4% | 14.5% | 69.2% | 17.5% | 22.6% | 0.9% |
| 2001 | 2.8% | 4.2% | 15.4% | 74.4% | 11.3% | 16.6% | 0.8% |
| 2002 | 2.2% | 3.3% | 16.6% | 80.3% | 13.0% | 14.1% | 0.7% |
| 2003 | 2.5% | 2.3% | 17.0% | 87.7% | 15.0% | 16.7% | 0.5% |
| 2004 | 2.7% | 2.1% | 17.5% | 99.0% | 18.2% | 20.7% | 0.6% |
| 2005 | 3.4% | 2.2% | 13.2% | 104.1% | 27.4% | 34.1% | 0.7%* |

* Provisional data

Accordingly, some of the risks to the sustainability of economic growth are based on the possibility of unexpected and unfavourable changes in employment, interest rates or property prices. Although no sign of deterioration in creditworthiness for mortgage borrowers has taken place so far (the last column in Table 2 reveals the low level that the default rate has for this specific portfolio), their growing credit exposure may pose a threat to the strength and soundness of the Spanish banking system, which is the main creditor of Spanish households, if certain negative events should coincide. These negative events would translate into an increase in both households' probability of default (unemployment and interest rate rises taking place) and in their losses in case of default (downward correction of house prices). This might jeopardise the financial condition of credit institutions ultimately affecting the stability of the Spanish economy.

5

This paper analyses the effects of an adverse scenario (unlikely but plausible) on the mortgage credit market, in particular its impact, possible consequences and the adequacy of loss coverage defences raised by credit institutions to weather such a negative situation. To achieve this goal stress tests appear to be one of the most appropriate tools at the disposal of financial authorities to explore, gauge and, finally, evaluate the latest implications of a stressed economic scenario within the banking system.

## 3. Database and sample data

One of the main problems when using banking default data is the difficulty in getting access to accurate information so that the final results are as precise and correct as possible. The database used to estimate our mortgage default equation model, which constitutes the basis for obtaining the distribution of mortgage losses and for implementing the stress exercise, is the Bank of Spain's Credit Register (CR). This register records monthly information on all credit operations granted by credit institutions in Spain for a value of over €6,000. CR's data structure distinguishes between credits given to firms and those to individuals. Among the latter it is also possible to identify loans granted to households (either for consumption or real estate purposes).

The CR contains information on the main characteristics of each loan, including the following: type of instrument (trade credit, financial credit, leasing, etc.), currency denomination, maturity, existence or not of guarantees or collateral, type of guarantor, coverage of the guarantee, the amount drawn and undrawn of a credit commitment and, finally, but very importantly, whether the loan is current in payment or past due (distinguishing, in turn, between delinquency and default status). The CR also includes some information relating to the characteristics of borrowers (for individuals the province of residence). Unfortunately, there is no available information regarding the interest rate of the loan and the loan-to-value ratio at origination.

This paper focuses only on loans to individuals, in particular, mortgages to households. These are collateralized credits with maturity over five years. The in-sample period goes from 1990 to 2003 (training sample), using the year 2004 as the out-of sample set for evaluating results (validation purposes). This time span covers a whole business cycle for the Spanish economy, including the recession of the early nineties and the subsequent upturn during the mid-end nineties and the first years of the current decade.

The entire population of mortgages in Spain for the observed period is well over 30 million loans. This figure makes it impossible to estimate any possible regression model. Consequently, we had to turn the population into a manageable sample. We implemented a procedure based on a very simple rule that produces a stratified sample (roughly 10% of the original number of observations) whose main descriptive statistical properties perfectly match with those of the entire population. After applying the sampling method, we were left with more than two and a half million observations. In Table 3 we reproduce some descriptive statistics for the whole population of mortgages and those for the selected

sample. As said before, the main statistics of the sample reflect almost perfectly those of the population (e.g. loan size, problem loans ratio and default rate)[5].

It should be noted that even though the default ratio (proportion of defaulted borrowers) is relatively low, the amount of defaults included in the sample assures the required stability that an estimation process requires. Therefore, possible slight changes in this number (e.g. incorrect reporting from individual banks to the CR) that may take place would not significantly alter the sample statistical properties.

**TABLE 3. DESCRIPTIVE STATISTICS OF THE POPULATION AND SAMPLE USED**

| Year | Number of Total Operations | Total Exposure € Million | Average Loan € thousand | 90th Percentile € thousand | Problem loans Ratio | Default rate |
|---|---|---|---|---|---|---|
| 1990 | 1,095,881 | 17,900 | 16.4 | 30.0 | 6.7% | 5.8% |
| 1990 sample | 109,803 | 1,800 | 16.4 | 30.0 | 6.5% | 5.6% |
| | | | | | | |
| 1993 | 1,792,216 | 48,906 | 18.5 | 36.0 | 7.1% | 6.1% |
| 1993 sample | 179,814 | 4,910 | 18.5 | 36.0 | 7.1% | 6.1% |
| | | | | | | |
| 1997 | 3,967,016 | 95,126 | 20.7 | 42.0 | 1.7% | 1.5% |
| 1997 sample | 396,034 | 9,620 | 20.8 | 42.0 | 1.7% | 1.5% |
| | | | | | | |
| 2001 | 5,787,661 | 194,083 | 28.0 | 55.0 | 1.4% | 1.2% |
| 2001 sample | 578,260 | 19,400 | 28.0 | 55.0 | 1.3% | 1.2% |

## 4. Empirical Model

### 4.1. Model Design and Variables

As previously stated, one of the objectives of this paper is to find the determinants that explain the event causing an individual obligor to default. At the same time it also provides us with a system that classifies bank borrowers according to their creditworthiness. One of the most important requirements for these classification systems is that they should generate effective discrimination between good credit quality (non-defaulting) borrowers and bad credit quality (defaulting) ones. The statistical power (predictive power) of these systems mostly depends on their accuracy and capability of creating a consistent rating structure among obligors, as their main output is an estimate of each borrower's PD. This, in turn, is used as an input for determining the probability density function of the loss rate associated with a bank's credit portfolio.

Beyond any shadow of a doubt, it can be affirmed that the crucial parameter that determines the real capacity of an obligor to service his credit debts (and, in the last instance, his probability of default), is his income stream. Unfortunately, we, in general, lack information about borrowers' financial structure. Trying to overcome this caveat we will construct certain financial and loan variables with the aim of

---

[5] Even though only four years are presented, the commented similarity among sample and population statistics applies to the whole period (1990-2004).

substituting the informational content of the basic income characteristics of individuals. Banks, based on direct information from borrowers' income sources, discriminate among them and grant mortgage loans with significantly different characteristics. Some of these differential features are, directly or indirectly, contained in the information included in the CR and they constitute the basis for our approximation to replace the unavailable variables which characterize each obligor's financial structure.

Among others, the instruments to be used to try to explain a default event of a mortgage credit are the following. First, loan characteristics: maturity of the loan, its size, and its repayment flows. Second, obligors' risk profile variables which include both information from alternative credits (credit cards or consumer credit granted to each borrower), and from previous default or delinquency events in servicing either the mortgage loan or other type of consumer credits. Finally, we also pay attention to the information content of variables such as the number of banks a borrower asks for a loan, existence of repayment delay agreements, restructured contracts (refinancing agreements), and possible variations or combinations of the variables described above.

Turning to the estimation of the regression model, the first step to be taken is to represent the default event as a random variable. It is generally assumed that a borrower's default is determined by the value of his assets, such that if this value falls below that of his liabilities a default event is triggered[6]. This precise event may be thought to determine when an obligor becomes an explicit danger that may erode the quality of a bank's credit portfolio. The difference between assets and liabilities is generally assumed to be a continuous variable, yet that variable is not directly observable. Consequently, the approach to be followed is to use a binary variable that takes value one or zero based on the occurrence of an observable event.

At this stage we resort to the information contained in the Credit Register regarding the default status of each loan in order to determine whether or not a default event is considered to have occurred[7]. Based on this information the endogenous variable for the regression model is constructed[8] and the next step is to obtain its determinants from a group of eligible candidates, both macroeconomic and loan characteristic variables that, to some extent, try to approximate the financial structure of each borrower and also portray his risk profile.

---

[6] In the particular case of households, whenever the affordability ratio (household debt over disposable income) reaches a critical value (threshold limit), individuals are supposed to become unable to meet their credit obligations and a default event arises.

[7] The definition of defaulted loans included in the Credit Register is similar to that established in Basel II when referring to a default event, i.e. those obligors that are past due more than 90 days on any credit obligation, or those that can be considered highly unlikely to be able to pay their credit obligations. For more information see BCBS 2006, paragraphs 452-453, page 100.

[8] The endogenous variable, $PD_{it}$, is a dichotomous (zero-one) variable which takes value 1 if the borrower i defaults (according to the definition of default previously stated and contained in the information included in the CR) in year t and 0 otherwise.

The statistical model selected to establish the relationship between the default variable and the selected explanatory variables is the logistic one[9] and it is estimated by the standard maximum likelihood maximization process[10]. The regression-based estimates of individual PD's for mortgage loans depend on the following generic variables:

$$PD_{it} = F\{\beta_1 RISKPROFILE_{it} + \beta_2 LOAN\_TYPE_{ti} + \beta_3 MACRO\_VAR_t\} \qquad (1)$$

where, $PD_{it}$ is the predicted probability of default of borrower $i$ in year $t$ (1990-2003) and $F[\ ]$ is the cumulative standard logistic function.

It should be stated that the estimated PD's are conditional, requiring that a particular obligor defaulting in a certain year shall not have defaulted during the previous period. Therefore, the predicted PDs measure the likelihood that an obligor will default during a certain assessment horizon. This horizon is fixed to a one-year period. Additionally, it is common practice in credit risk analysis to aggregate all loans to a borrower into a single exposure. Consequently, if a borrower has several mortgages, failure to meet his payments in any one of them means that this borrower is in default.

Among the explanatory variables included in equation (1), $RISKPROFILE_{it}$ is a vector of variables that accounts for the main risk profile characteristics (for which information is available) of each borrower. In particular, in this array we include variables such as $DELINQ_{it}$ which informs on whether or not a certain borrower has been delinquent on his mortgage loan. Similarly, $HIST\_DELINQ_{it}$ accounts for the possibility that a borrower has been historically delinquent, that is, in previous periods (t-1, t-2...) to the current one. Accordingly, it represents borrowers historically overdue on their mortgage loans but who have finally met their financial obligations before the 90-day threshold, that is, before having become defaulted. It has to be remarked that most of the problems that are behind an overdue loan are "technical" ones, spanning only a few days as a result of mistakes or lack of monitoring of balances, incidental cash shortage, holidays, etc. This variable represents the expected behaviour of risk-averse borrowers who will always hold a minimum income buffer for unexpected events to avoid becoming overdue. This variable is weighted by the distance in time a borrower committed delinquency on his loan.

In the same vein, $HIST\_DEF_{it}$ is a another risk profile variable which informs whether a certain borrower has defaulted in any period previous to the one used to fix the one-year assessment period for which the

---

[9] The choice of the logistic model is based not only on the fact that the default status of an obligor can be reasonably described in probabilistic terms (using probabilities of default), but also on the fact that such status is determined by the asset value of the borrower. As commented before, this status is an unobservable latent variable so it can be approximated by a binary one, i.e. the endogenous variable in a logistic regression.

[10] An alternative estimation process can be seen in Alfó et al (2005). They extend the standard logistic transformation to another which includes random effects for possible predictive capability gains.

PD is calculated (i.e. possible failure in period t-2, t-3...)[11]. This variable is also weighted by the distance in time since the default of an obligor took place.

Another risk profile variable introduced in equation (1) is $DELINQ\_NOMORTGAGE_{it}$, whose meaning is exactly the same as for $DELINQ_{it}$ except that the delinquency status is recorded for each obligor on a non-mortgage credit, (i.e. consumer or credit card). The explanatory power of this variable is complemented with that of the following two. First, $DEF\_NOMORTGAGE_{it}$, which informs whether or not a borrower is currently servicing his debts in other loans than his mortgage. In second place, we use $HIST\_DEF\_NOMORTGAGE_{it}$ which reveals whether or not a borrower has historically experienced a default event in a credit different from his mortgage loan. These two variables provide evidence on the nature of household defaults. Defaults on other credits than mortgages may be used to anticipate future mortgage defaults. Additionally, it has also to be noted that the second variable is weighted by the distance in time since the default or the delinquency event of each obligor took place. Thus, the more distant in time the default took place, the less it counts, something that seems in line with banks' risk management practices in general to assess loan quality.

The $LOAN\_TYPE_{it}$ vector includes the following variables: $REPAYMENT_{it}$ is the proportion of the original risk (original credit amount granted) that has been amortized up to period t. With this variable we try to verify the hypothesis on mortgages that the higher the amount of a loan repaid, the more unlikely a borrower can be considered to default. A complement to the former variable is $AGE_{it}$ which measures the age of each loan. It usually coincides with the number of years a borrower has been reported to the Credit Register. The rationale for including this type of variable in the regression model is based on the commonly accepted premise which presupposes that there is a particular relationship between the age of each loan and its probability of default. In general, this assumption considers that higher rates of default take place during the first years after a mortgage is granted. After that period of time (three to four years), the rate of default decreases progressively with the age of the loan[12].

We also include in equation (1) the variable $NUMBANKS_{it}$, representing the number of banks with which an obligor has lending relationships (number of banks to which a borrower applies for a mortgage loan). It has to be noted that the data used for the regression analysis concerns individual borrowers, not banks. As a result, we hypothesize that the more banks a borrower is related to, the more constrained he may be in terms of liquidity. $CONSUMP_{it}$ indicates the proportion that other risks (consumer, credit card or both) represent with respect to the original mortgage risk of each borrower. We expect that the higher the proportion of alternative risks to be served, the higher the probability of default. Additionally, the variable $MATURITY_{it}$ is the maturity of each loan; $SIZE_{it}$ stands for the size of each loan (amount of the

---

[11] It is necessary to remark that the recorded failure in this variable had to take place either in t-2, or before so as to be consistent with the established definition of default: a borrower failing to meet his credit debts in period t given that he was in current payment in t-1.

[12] A similar behaviour of an inverted U-shaped curve between the age of a loan and its probability of default is found on credit card delinquencies by Gross and Souleles (2002).

mortgage loan at origination)[13] and $RESTRUCT_{it}$ is a dummy variable that represents whether or not the credit obligation of a borrower has been subject to restructuring either in amount, maturity, or other significant financial characteristic.

Finally, among the $LOAN\_TYPE_{it}$ variables, we have included two sets of dummies. One of them stands for the region where the borrower is granted the loan. The other one controls for the type of mortgage lender (type of financial institution): commercial bank, savings bank, credit cooperative, or credit finance establishment. With this variable we try to describe the different lending policies that distinct types of credit institutions may have and consequently, the different level of risk that they may incur based on this fact.

The macro variables (those that change in time but take the same value for all borrowers) included in (1) are the unemployment rate, $UNEMPLOYRATE_t$, and the variation (percentage point increase or decrease) of official mortgage interest rates, $INTRATE\_VAR_t$. With these variables we represent the systematic factors that affect credit risk. As can be observed in Table 2, there is a cyclical pattern in the percentage of defaults per year which reasonably resembles that of the business cycle[14]. In other words, it can be appreciated that the period with the highest percentage of defaulted obligors is associated with an economic downturn and that, as the Spanish economy started to recover (1995 onwards), the percentage of defaults started to decline, reaching its lowest point in 2003.

These macro variables should be interpreted as the key ones to be used for implementing the stress test exercise. As the regression model links macroeconomic factors to the credit quality of each borrower, we can quantify their impact on the process of assigning probabilities of default to individual borrowers and, in the last instance, on the distribution of credit losses. That is to say, by changing (stressing) the values that the macro variables take in equation (1), we are able to impact the probability of default of each borrower and finally, the average PD of the whole mortgage portfolio.

All the above-mentioned variables are only a subset of the total amount of variables used to fit the logistic model. We also tried different transformations of them (both continuous and discrete specifications). To select the variables with the highest explanatory power, we first ran univariate regressions (borrower mortgage default as a function of each variable at a time). From those regressions we were able to confirm the right sign (explicative capability) and the explanatory power (predictive capability) of each variable on a stand-alone basis. With the results of the univariate exercise in mind, we ran equation (1) using combinations of the selected variables. Based on forecasting capacity[15], the main guide for banks

---

[13] This variable has been transformed and enters equation (1) in logarithmic terms.

[14] It can be accepted that the latest Spanish business cycle runs from the early nineties to the first years of the current decade. The trough of the business cycle was reached in 1993 while the peak can be positioned near 2003.

[15] This predicting capacity has to be understood in terms of accurate classification of defaulted and non-defaulted borrowers.

when they develop their scoring systems, we finally decided the set of variables included in the final multivariate regression. The main results from it are presented in the next subsection.

### 4.2. Results

Table 4, column 1, shows the estimated coefficients of the variables included in equation (1)[16]. An indication of the goodness of fit of the model can be found in the signs (expected ones) with which all variables enter the equation and in the predictive power obtained both for the training and for the validation sample (Table 5).

### TABLE 4. RESULTS OF THE REGRESSION MODEL

| Variable | Regression estimates Sample period (1990-2003) | | Regression estimates Sample period (1990-2002) | | Regression estimates Sample p. (1990-1995; 1998-2003) | |
|---|---|---|---|---|---|---|
| | Coefficient | Std.Dev. | Coefficient | Std.Dev. | Coefficient | Std.Dev. |
| UNEMPLOYRATE | 0.1124 | 0.0046 | 0.1084 | 0.0049 | 0.1016 | 0.0049 |
| INTRATE_VAR | 0.0513 | 0.0108 | 0.0445 | 0.0111 | 0.0409 | 0.0113 |
| HIST_DELINQ | 0.5620 | 0.0484 | 0.5554 | 0.0493 | 0.5366 | 0.0522 |
| HIST_DEF | 2.4531 | 0.0485 | 2.3503 | 0.0512 | 2.6012 | 0.0494 |
| HIST_DEF_NOMORTGAGE | 0.2236 | 0.0289 | 0.2304 | 0.0298 | 0.2348 | 0.0297 |
| DELINQ | 1.1598 | 0.0546 | 1.1432 | 0.0554 | 1.1355 | 0.0590 |
| DELINQ_NOMORTGAGE | 1.4684 | 0.0386 | 1.4115 | 0.0411 | 1.6190 | 0.0376 |
| DEF_NOMORTGAGE | 3.1168 | 0.0355 | 3.0188 | 0.0385 | 3.2632 | 0.0338 |
| AGE | -0.3715 | 0.0834 | -0.3464 | 0.0898 | -0.1966 | 0.0775 |
| REPAYMENT | -0.0263 | 0.0128 | -0.1983 | 0.0220 | -0.0522 | 0.0231 |
| MATURITY | 0.0010 | 0.0001 | -0.0001 | 0.0010 | 0.0015 | 0.0010 |
| SIZE | 0.3501 | 0.0197 | 0.4146 | 0.0206 | 0.3247 | 0.0195 |
| RESTRUCT | 0.2447 | 0.0215 | 0.2699 | 0.0224 | 0.1898 | 0.0216 |
| CONSUMP | 0.0061 | 0.0008 | 0.0061 | 0.0009 | 0.0061 | 0.0009 |
| NUMBANKS | 0.9046 | 0.0458 | 0.8525 | 0.0480 | 0.9137 | 0.0466 |
| Type_of_lender dummies: | | | | | | |
| Savings_banks | 0.5502 | 0.0585 | 0.5265 | 0.0245 | 0.5450 | 0.0232 |
| Credit_cooperatives | 0.4006 | 0.0234 | 0.3516 | 0.0506 | 0.4756 | 0.0450 |
| Credit_finance_est. | 0.5677 | 0.0474 | 0.5350 | 0.0528 | 0.6877 | 0.0522 |
| Regional dummies | | | | | | |
| Constant | -6.4879 | 0.1848 | -6.7421 | 0.1931 | -6.2406 | 0.1876 |

No. of observations: 2,221,197 (sampling period: 1990-2003).

With respect to the variables included in equation (1), in general terms, it may be said that the riskier the profile of a borrower, the higher his probability of default. Being overdue either in mortgage or in non-mortgage loans, may be considered as a signal of future mortgage defaults. Consequently, the hypothesis of previous defaults in other loans, either consumer or credit cards, anticipating mortgage defaults, is confirmed. Furthermore, given how these variables have been constructed, the closer in time a default in consumer credit, the more likely the mortgage default.

---

[16] The second column of Table 4 provides the standard error of the estimates in column 1. All variables are significant at the 95% confidence level.

Regarding other variables, we appreciate in Table 4 that the older a loan, the lower its probability of default (negative sign of the age variable)[17]. In terms of repayment, the negative sign of the repayment variable confirms the hypothesis that the higher the amortized amount of a loan, the lower its probability of default. With regard to maturity issues, longer maturity terms imply higher probability of default. This reasoning goes in line with that which establishes that riskier borrowers are granted longer maturities allowing for less strict repayment conditions (more time to pay back the loan and, possibly, lower monthly payments). With respect to the original size of a loan, large loans appear to be riskier as they are more prone to default than smaller ones. Furthermore, borrowers whose loans have undergone negotiated restructurings possess lower credit quality than the rest (positive sign in the RESTRUCT variable).

From Table 4 we can also infer that the more indebted a borrower is, the more risk he supports and, as a result, the lower his creditworthiness (positive sign in the CONSUMP variable). Additionally, the more lenders an individual resorts to, the higher the probability of default (positive sign in the variable number of banks). Thus, liquidity constraints also seem to play a role in mortgage defaults. Commercial banks (omitted variable of the lender-type dummy variable) reveal themselves as the type of credit institution whose credit risk policies seem to be the most conservative regarding mortgages, whereas credit finance establishments grant loans to borrowers with higher probability of default (positive and highest coefficient for this type of lender dummy). Finally, regional dummies (coefficients not shown in the table) account for structural risk that derive from the region a loan is granted.

The sign of the macro variables included in the regression are as expected. The sign of the unemployment rate and that of the interest rates variation are both positive. Generally speaking, this means that mortgage defaults increase during downturns and recessions (negative economic growth, high rates of unemployment and increases in interest rates), and vice versa during upturns.

To check for robustness and stability of the estimated model, we performed different tests. In particular, we omitted several years of the sample period over which we estimate equation (1). From the outcome thus obtained, we can conclude that, in general terms, results hold. The third through sixth columns of Table 4 show that estimates are little affected when the sampling period is modified. Column 3 presents the results of the estimated model when year 2003 is taken out of the sample. The stability of the coefficients can be observed in that the sign basically holds for all variables (with the exception of the maturity variable) and that variables continue to be significant at the 95 % confidence level. The same conclusion (Column 5 of Table 4) is also drawn when the omitted years are taken out from the middle of the sample, i.e. years 1996 and 1997. The main conclusion from this test is that stability is to be found not

---

[17] The univariate analysis confirmed the U-shape relationship between the PD and the age of each loan. This was done by separating loans according to their age and calculating their rate of default (descriptive statistics), and by means of dummy variables: a different dummy variable was constructed regarding the age of a loan (1-year-old loans, 2-year-old loans, etc.). Then, a single regression of the default rate against those dummies was run. The coefficients of the dummy variables provided the evidence of the hypothesized relationship between age and default rates.

by adding a large number of periods to the estimation process, but picking up stable relationships among financial and economic variables that only need a short number of periods to become evident.

Once the variables that determine an individual's possibility of defaulting have been established and their coefficients and signs within the multivariate model are known, it is advisable to establish certain performance measures for the estimated regression model to evaluate its predictive capability.

Table 5 shows the classification power of the estimated logistic model. From this table, it can be observed that the fitted model classifies correctly approximately 71% of the defaulted borrowers included in the sample and more than 83% of non-defaulted obligors. In terms of statistical power of the regression model, alternative performance measures confirm precision in prediction terms. In particular, the area under the ROC curve roughly reaches 82% which results in an Accuracy Ratio[18], AR, slightly higher than 64%.

**TABLE 5. LOGISTIC MODEL PERFORMANCE**

| | Classification Table In-sample model (1990-2003) | | Classification Table Out-of-sample model (2004) | |
| --- | --- | --- | --- | --- |
| | Predicted Defaults | Predicted Non-Defaults | Predicted Defaults | Predicted Non-Defaults |
| Observed defaults | 70.63% | 29.37% | 66.19% | 33.81% |
| Observed non-defaults | 16.61% | 83.39% | 13.63% | 86.37% |
| | Area under ROC curve = 0.823 Accuracy ratio = 64.6% | | Area under ROC curve = 0.813 Accuracy ratio = 62.6% | |

This classification power corresponds to the capacity to correctly predict borrowers´ status within the training sample (classification table in Table 5 for the in-sample model). However, in order to test the predictive performance of the model, a validation process is carried out evaluating results on an out-of-sample set (sub-sample of obligors not included in the estimation sample, i.e. using 2004 default data). The validation process simply consists of calculating the score of every borrower and comparing it with its observed default status. The results are fairly reasonable compared to those of the training sample (66% of defaulted individuals were correctly classified with an AR ratio of 62%) indicating a satisfactory classification power of the estimated model.

---

[18] The AR measure is obtained from the Cumulative Accuracy Profile curve, CAP, and determines the performance enhancement over the random model of the model under evaluation (in our case, equation 1). Other common statistical measures for determining the discriminatory power of models are the area under the ROC (Receiver Operating Characteristic) curve, the use of contingency tables (particular cases of accuracy ratios), entropy measurements or others such as the Brier score or similar statistics. For references of performance power statistics see, among others, Sobehart, Keenan and Stein (2000), Sobehart and Keenan (2001) and Engelmann, Hayden and Tasche (2003).

**5. Mortgage credit loss distribution**

Despite the particular importance that the mortgage portfolio has for a majority of banks, relatively little is known about its credit loss distribution: loss rates at different percentiles, factors which this distribution depends on, and the effects and impact that a significant change in these factors may have on that portfolio. By a non-parametric estimate of the credit loss rate of mortgages, and based on the sensitivity of the probability of default of this type of loans to different macro variables and to other key parameters, we try to address and, eventually, provide an explanation to the above.

As mentioned in the introduction to this paper, we follow a non-parametric method for estimating the mortgage credit loss distribution to avoid any possible assumptions about correlations between defaults or specific functional forms for this distribution. To estimate it, a Monte-Carlo simulation is run. With this method we simulate the possible range of losses a given portfolio could face when assets are randomly selected from the universe of those available for investment. As already noted, this approach obviates the need to measure correlations because they are already embedded within the loss experience of the database (CR) from which, in this particular case, mortgage assets are drawn.

In our paper, the Monte-Carlo resampling method is applied to the entire banking system and simulates mortgage portfolios by randomly drawing mortgage assets (without replacement) from the Credit Register database. Nevertheless, for each simulated portfolio it is required that certain fixed specifications be met so as to ensure that the simulation conforms to the main properties observed within the universe of credits. Repeating the resampling process a sufficient number of times (say, 10,000), enables us to obtain the shape of the distribution together with its associated percentiles, since losses for each simulated portfolio are computed and their frequency finally becomes the estimate of the loss distribution.

The regression equation estimated in section 4 provides us with the main determinants of the default event of a mortgage loan, and gives us the necessary information to gauge how a change in them will alter each borrower's credit quality. This information is summarised in the main output obtained from the logistic model: an estimate of each borrower's PD. Accordingly, we can classify all obligors in different risk categories by means of a rating. This rating system reflects the risk structure (distribution of individuals in risk classes) that exists in the population of mortgages loans, and constitutes the basic characteristic that we impose on each simulated portfolio every time an iteration of the resampling method is run.

In particular, for each year of the sample period we simulate a portfolio establishing a target percentage of borrowers to fall in each rating category. Individuals for that year and for each rating class are randomly chosen from the database until the portfolio conforms to the latent risk structure that the population of mortgage loans possesses. As we separate for years, we distinguish among different realizations of the systematic factor that might otherwise result in undesirable combinations (good and bad economic years)

15

within the same portfolio, counterbalancing and (more importantly) understating the level of losses computed in a particular iteration.

Undoubtedly, the utilization of a resampling approach to obtain the distribution of losses crucially depends on several parameters: first, the mortgage default rate. Second, how borrowers are rated according to their creditworthiness. Third, the size of the simulated portfolio and, fourth, the number of iterations. The first parameter is clearly determined by the prevailing economic conditions of each year of the sample period. Recession years record more defaulted assets than good years and, as a result, there are more chances of picking them up in each iteration affecting the loss rate of the simulated portfolios and finally reflected in the loss distribution function.

Regarding the second parameter and as noted before, the individual PDs from the logistic model help us to create a rating system of mortgage borrowers. Certain widely accepted basic premises are required for building up a meaningful rating system: first, obligors should be reasonably and consistently assigned to risk grades[19]. Second, each rating category should include a minimum number of borrowers and, at the same time, avoid possible concentration of them both in number and volume of exposure. Lastly, the probability of default in each risk category should increase with respect to the previous one as we move from lower to higher credit risk classes. As can be appreciated, these common sense features do not impose special restriction on the way the classification system is created. Details on the PDs of each rating grade and on the rating average PD can be observed in Table 6.

Additionally, the results of the resampling method are relatively sensitive to the size of the simulated portfolios. The bulk of the discussion in this section is organized by reference to coverage of estimated credit losses. In particular, we assess the regulatory coverage stipulated by the IRB approach of Basel II. As a result, the size parameter of each simulation has been set up bearing this purpose in mind. Specifically, we establish that the size of the simulated portfolios will coincide with the average size of a mortgage portfolio of a Spanish bank that may opt for the Basel II IRB approach. Consequently, what we simulate is the credit loss distribution of a mortgage portfolio of an average IRB-eligible bank when taking into account the entire universe of mortgage assets that this particular type of bank could consider for investment. It should be noted that portfolio size for other types of bank is lower (less granular portfolios), thus increasing the possibility that, for simulated portfolios, defaulted assets could concentrate in them, and resulting in an empirical distribution characterized by higher loss rates[20].

---

[19] It is reasonable to admit that creditworthiness may be proportionally distributed in a certain fashion among the universe of credit borrowers. This means that we expect that there should be a low proportion of very good and very bad quality borrowers, whereas mid-credit-quality borrowers may be considered the most common type of obligors.

[20] It should be noted that the selection of an adequate size of the simulated portfolios is essential. To take an example: the choice of a very small portfolio size creates the possibility that in a particular iteration of the resampling method a high concentration of bad quality assets be drawn, giving rise to a high loss realization. On the other hand, for an extremely large size (e.g. equal to the number of loans that fall in a rating category) the estimated loss distribution will collapse to the average loss rate of that rating grade, since all obligors (defaulted and non-defaulted) of that rating grade are drawn in each iteration of the resampling method.

Finally, with respect to the number of iterations a sufficient number of them must be chosen so that convergence is attained. This convergence has to be understood as generating a sufficient number of combinations for possible losses so as not to understate the bad tail of the loss distribution.

Turning to the main aim of this section, i.e. deriving the distribution of mortgage credit losses, we look for evidence to assess whether or not banks have enough prudential defences to withstand unfavourable economic and financial conditions so that they may be safely protected against the actual risks they incur.

### TABLE 6. SPECIFICATIONS AND MAIN PARAMETERS OF THE SIMULATION PROCESS

| Simulation of the Base Case Parameter | Main characteristics |
|---|---|
| Number of iterations | 10,000. A higher number of iterations does not produce different results in estimating the bad tail of the loss distribution. |
| Sample period | 1990-2004; All years equally weighted. |
| Portfolio size | 20,000 loans; Estimated average mortgage portfolio size of an IRB-eligible Spanish bank. Basis year 2004. This average is corrected by the same parameter initially used to reduce the population of 30 million loans to a manageable size. Additionally and prior to 2004 this parameter is adjusted to changes in the actual portfolio size of the corresponding banks. |
| Rating structure | Seven rating grades. Grade PD estimates are cycle-average PDs, i.e. cycle-average of one-year default rates for borrowers in each grade. Cycle-average grade PDs: Rating 1 PD:0.27% Rating 2 PD:0.29% Rating 3 PD:0.43% Rating 4 PD:0.49% Rating 5 PD:0.87% Rating 6 PD:3.12% Rating 7 PD:12.15% Average-rating PD: average of all grade PD estimates weighted by the number of obligors falling in each grade. Weighted average- rating PD:1.46% |
| LGD | 100% (no correction of losses via LGD). Increases or decreases in this parameter move the estimated amount of losses proportionally up or down. |

Typically, capital and provisions protect banks against losses. Their adequacy may enhance the stability and soundness of credit institutions. In particular, provisions should protect banks against the natural credit risk inherent in ordinary banking business (prevention and coverage of expected losses), whereas capital is raised to avoid deeper solvency crises, in other words, to mitigate insolvency risk in banks (coverage of unexpected losses). However, our interest is only focused on loss rates and whether or not regulatory coverage accounts for them. The way this coverage or loss protection is split between either expected or unexpected losses may deserve a more profound and thoughtful discussion, and falls beyond the scope of this paper. In addition, in this paper we do not consider possible variations or different

specifications in LGD. Consequently, final results are not adjusted by this parameter and implicit loss rates are presented using a 100% LGD.

Specifically, in this paper we wish to address the adequacy of regulatory coverage measures by comparing them to the figure obtained when using an empirical distribution function of credit losses obtained by a Mote-Carlo simulation process.

Furthermore, we run a stress exercise to determine to what extent, under specific unfavourable conditions, regulatory minimum requirements of coverage of credit losses guarantee that losses from a stressed economic scenario are certainly accounted for.

To run the stress test exercise, we turn to the estimates of equation (1). By modifying the values of the macro variables and those of other structural variables (basically loan characteristics) that determine individual obligors' PDs, we create our stressed scenario. Accordingly, under such a predetermined scenario, we can exactly ascertain how mortgage losses are affected. Thanks to the output of the logistic model (PDs), we can identify defaults and measure the corresponding losses and see whether or not protection exists to cover them.

In short, we compare how the distribution of losses under a stressed economic environment changes with respect to that obtained for a base scenario (current values of each determinant of mortgage defaults). In this way, we can determine how sensitive mortgage losses are with respect to different economic contexts (each of them constituting a distinct stressed scenario), and, more importantly, whether or not banks may weather losses deriving from possible adverse scenarios with the aid of the regulatory elements available for them.

### 5.1. Results

In the first place, we obtain the distribution of mortgage losses (base case)[21] for the whole sample period, 1990-2004, to compare losses obtained from this empirical function with regulatory Basel II coverage measures for mortgage loans. The Basel II IRB mortgage formula is built on conditional PD values dependent on all possible different realizations of the common factor (representing both adverse and benign economic conditions). In the same vein, we calculate the non-parametric distribution of credit losses, as described in the former section, considering different realizations of the systematic factor in our sample. In particular, all available years are used and all are equally weighted.

This first comparison allows us to determine to what extent empirical mortgage losses obtained for a practically complete economic cycle may be said to be covered when using the formulas stipulated by Basel II. In Table 7 we can observe that for the 99.9[th] percentile, the simulated empirical credit loss rate (11.29%) is slightly higher than that obtained from the Basel II formula, 10.38%. This loss rate is

---

[21] The base case simulation conforms to the parameters and specifications shown in Table 6.

calculated by means of the estimated rating structure found in the data, that is, the cycle-average PD of each rating grade is plugged into the Basel II formula and the result multiplied by the mortgage exposure of that grade. However, if the weighted-average rating PD is used, then the loss rate provided by the Basel formula increases to 14.32%, thus covering empirical loss estimates. This difference is simply the consequence of employing a rating system and of applying it into a concave function as it happens to be that of Basel II. This rating system allows us to properly classify borrowers, and produces a more accurate assessment of their inherent risk and also highlights the importance that should be given to correctly determining the risk classes (number and composition) into which borrowers are classified, since regulatory coverage measures are calculated based on a consistent definition of those risk groups.

### TABLE 7. DISTRIBUTION OF CREDIT LOSSES (BASE CASE vs. BASEL II)

| | Loss rates (percentage of exposure) | | | |
| --- | --- | --- | --- | --- |
| | Mean | 99th percentile | 99.5th percentile | 99.9th percentile |
| Mortgage loss distribution | | | | |
| Base Case 1990-2004 | 2.94% | 9.59% | 10.14% | 11.29% |
| | | | | |
| Basel II IRB approach Rating Structure: Cycle-average Grade PD estimates | | 6.36% | 7.50% | 10.38% |
| | | | | |
| Basel II IRB approach Average-rating PD | | 8.27% | 9.99% | 14.32% |

It should also be noted that both for the 99$^{th}$ and for the 99.5$^{th}$ percentiles, Basel II falls short in covering simulated losses regardless of the type of PD used (i.e. either using individual rating grade PDs or the average rating one). Both Basel II tail rates are always below that of the base case up to a certain percentile. From that percentile onwards, the bad tail of our simulated credit loss distribution is flatter than that of Basel II (e.g. 99.9$^{th}$).

Once we have obtained the empirical mortgage loss distribution, we turn to the stress exercise. We first have to establish the portfolio for which we want to run the test and, at the same time, set the values to be taken by the macro and other structural variables in equation (1), which eventually determine the scenario for testing the loss rate of mortgage credits. A very simple approximation that may be used is to fix their values at their maximum historical variation (or the minimum where appropriate, i.e. depending on the sign in equation (1)). In short, we obtain the distribution of the observed annual variations and, based on that, we select the maximum observed variation and apply it to each variable.

Under these conditions, it can be argued that the largest value change observed during a sufficiently long period of time, especially in macro variables, would represent a case of stress. Nevertheless, this change has to be relativised to the levels that the variables take in 2004. This must be done because it may be the case that a given percentage increase in the level of a variable may still produce a relatively small value for this variable if its current level is very low. This usually happens for boom periods. As the value levels of most variables in equation (1) are at historical lows, we have added twice the value of the recorded

maximum variations to the current level of several variables to create a more severe scenario. This has been the case for the macro variables and the maturity variable. Regarding the risk profile variables, given their nature (zero-one variables weighted by the distance in time a default or a delinquency event takes place), we have substituted their 2004 value for their observed average value in the sample period. This has also been the case for the restructuring variable. The rest of the variables that enter equation (1) have not been modified.

An alternative method may be to define a scenario where the key variables to be stressed take their historically most adverse observed value. Such a situation might be considered as highly improbable, particularly if the current levels of the variables are very far from their historical highs. However, this worst-case scenario, which essentially represents losses associated with the worst year in the sample period, may be taken as an extreme case to test whether or not regulatory coverage measures are adequate for such an adverse situation.

Therefore, the stress exercise consists of establishing a reference portfolio (2004 mortgage portfolio) and then changing the value taken by the variables entering the estimated logistic model. As a consequence of the change in the value of the variables, the value of the estimated PD of each borrower changes, and so does the average PD of the whole portfolio. This variation in the PD estimates allows us to identify the defaulted loans under the stress scenario, and permits us to simulate a different distribution of losses under the new economic conditions. In other words, since the default rate for 2004 increases given the adverse conditions imposed by the stressed economic context, we simulate and obtain a new distribution of losses to account for such a change.

In Table 8 we present the results of the stress exercise. From this table we can appreciate the low level of losses for the actual portfolio of 2004 (0.95% for the 99.9[th] percentile). The prevailing economic conditions of this year provoke an extremely low level of simulated losses. Consequently, regulatory measures in force more than cover them. Furthermore, when running the stress exercise (change in value of variables according to maximum recorded variation), as the level of the variables that define the prevailing economic context is basically at historical lows, the simulated losses continue to be very far from those that would be covered by regulatory requirement (2.25% versus 10.38%).

However, the worst-case scenario implies a major difference. Table 8 shows that both mean and tail loss rates are far higher in a recession (worst-case) than in a good economic year (2004). The pattern observed in this table reflects that, in order to achieve loss coverage in a bad year, prudential elements in force should be more thirteen times higher than those needed in a good year (0.95% vs.12.69%).

As expected from the results presented in Table 7, the rate of losses in the bad tail for the worst year in the sample have to be higher than that of an average year (portfolio distribution for the base case where all sample years enter with an equal weight). Nevertheless, what we can appreciate with respect to Basel II regulatory measures in force for loss protection is that loss coverage using the rating structure found in

the data fails to provide protection in less than 20% of that needed for the worst case (12.69% vs. 10.38% for the 99.9[th] Basel II VaR).

**TABLE 8. DISTRIBUTION OF CREDIT LOSSES (STRESS TEST EXERCISE)**

| | Mean | 99th percentile | 99.5th percentile | 99.9th percentile |
|---|---|---|---|---|
| | Loss rates (percentage of exposure) | | | |
| | Stress test exercise | | | |
| Mortgage loss distribution Base Case: year 2004 | 0.67% | 0.87% | 0.89% | 0.95% |
| Mortgage loss distribution Stress scenario: maximum variation | 1.82% | 2.13% | 2.17% | 2.25% |
| Mortgage loss distribution Stress scenario: worst case | 8.10% | 11.44% | 11.77% | 12.69% |
| Basel II IRB approach Rating Structure: Cycle-average Grade PD estimates | | 6.36% | 7.50% | 10.38% |

However, it should also be noted that we have considered losses at the 99.9[th] percentile in the stress scenarios when comparing with Basel II regulatory loss protection measures. Nevertheless, the scenario implied by a stress test exercise should be considered as instantaneous and, as a result, we should concentrate our attention on terms of likelihood, that is, the most likely result of that scenario (expected losses derived from it). In our particular case, regulatory loss protection, 10.38%, covers the most likely result, 8.10%. However, bad-tail events are indeed possible at any time, in particular in distress situations, and that is why we have paid attention to extreme percentiles of the stress exercise. If supervisory authorities are interested in soundness and protection of the banking system, comparison of extreme loss rates has to take place, even if they are very unlikely to occur.

In addition, we have kept rating grade and average-rating PD estimates constant when carrying out the stress exercise (Table 7 and 8). Basel II requires to calculate cycle-average PDs to account for loss protection. As commented above, stress tests evaluate the impact that instantaneous shocks may have on the financial system. Under this premise, cycle-average PDs should not be altered by those shocks and regulatory coverage measures should not be changed as a result. A different question will be a distress situation which implies a structural break. If a shock may be considered to be long-lasting, that is, it is expected to create a structural change in the economy, then cycle-averages should be modified (normally upwards) and a different loss coverage (higher) should be considered for stress test purposes.

## 6. Conclusions

Banks run risks when carrying out their ordinary business. When these risks become certain, losses appear. Their correct measurement and proper estimation emerges as one of the most important issues for supervisory authorities. This paper addresses this question by simulating an empirical distribution

function of credit losses, and by running a stress exercise to test whether or not banks possess adequate regulatory loss protection under certain adverse circumstances.

To achieve this goal, this paper focuses on one of the banks' most important credit portfolios: mortgage loan portfolio. We use the Bank of Spain's Credit Register to obtain default data from mortgage borrowers to estimate a model by which we characterize the event causing an individual borrower to default. We adjust a logistic regression model, and we find that macroeconomic variables (unemployment rate and interest rates) and other structural variables (obligors' risk profile and loan type variables) reasonably fit and predict the default event of mortgage borrowers.

Using the main output of a logistic model (individual obligors' probabilities of default), and by means of a rating system which reflects the inherent risk structure of the whole population of mortgage borrowers, we simulate the empirical distribution of the loss rate of mortgage credits. In particular, we use a resampling Monte-Carlo simulation method and estimate this credit loss distribution function to compare loss rate percentiles with regulatory loss protection and coverage measures provided by IRB Basel II formulas, in order to determine their adequacy. The main results obtained for a whole economic cycle show that, in general terms, empirical loss estimates at the 99.9[th] percentile are covered by Basel II regulatory loss protection measures.

Furthermore, we run a stress exercise to test banks' ability to withstand unfavourable economic conditions. In particular, we test whether regulatory loss protection seems adequate for banks to weather adverse situations. Two types of stress scenarios are presented: one based on the maximum historical variation taken by certain variables included in the logistic model (macro variables and risk profile ones), and another that we define as a worst-case scenario (worst recorded year in the sample used).

The main result from this exercise is that Basel II IRB regulatory loss coverage offers fairly adequate protection for banks. Specifically, when loss protection is calculated using cycle-average grade PDs from the borrowers' rating system, Basel II covers estimated losses for the most likely result, although it fails to cover tail rates at the 99.9[th] percentile. However, if the weighted average-rating PD is used, Basel II more than covers empirical stressed losses, even at the 99.9[th] percentile.

In light of the aforementioned, the final purpose of this paper is to show the usefulness that simulation and stress test exercises may have for supervisory authorities to obtain a better understanding of the risk profile of a given credit portfolio and explain its sensitivity to certain risk factors. At the same time, we have also tried to address the importance of several of the assumptions in the methodology that we have employed to quantify credit loss rates.

**Bibliography**

Alfó, M., Caiazza, S., Trovato G., 2005, Extending a Logistic Approach to Risk Modeling Through Semiparametric Mixing. Journal of Financial Services Research 28, 163-176.

Allen, L., 2004, The Basel capital accords and International mortgage markets: A survey of the literature. Financial Markets, Institutions & Instruments, Vol. 13, no. 2, May.

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

Bunn P., Cunningham, A. and M.Drehmann, 2005, Stress Testing as a Tool for Assessing Systemic Risks, Bank of England Financial Stability Review, June.

Blaschke W., Jones, M.T., Majnoni, G., Martínez Peria, S., 2001, Stress Testing of Financial Systems: an Overview of Issues, Methodologies, and FSAP Experiences, IMF Working Paper, June.

Calem, P.S., LaCour-Little, M., 2004, Risk-based capital requirements for mortgage loans, Journal of Banking & Finance 28, 647-672.

Carey M., 1998, Credit Risk in Private Debt Portfolios, Journal of Finance 53 (4), 1363-1387.

Carey M., 2001, Dimensions of Credit Risk and Their Relationship to Economic Capital Requirements, in Frederic S. Mishkin, ed., Prudential Supervision: Why Is It Important and What Are the Issues. Chicago: University of Chicago Press and NBER.

Caruana J., 2005, Monetary Policy, Financial Stability and Asset Prices, Occasional Paper. No. 0507. Banco de España.

Coleman A., Esho, N., Sellathurai I., Thavabalan N., 2005, Stress Testing Housing Loan Portfolios: a Regulatory Case Study, Australian Prudential Regulation Authority, Working Paper 2005-01.

Committee of European Banking Supervisors, CEBS, 2006, Consultation Paper on Stress Testing under the Supervisory Review, June.

Committee on the Global Financial System, 2005, Stress Testing at Major FinancialInstitutions: Survey Results and Practice, Bank for International Settlements.

Dimou, P., Lawrence, C., Milne, A., 2005, Skewness of Returns, Capital Adequacy, and Mortgage Lending, Journal of Financial Services Research, 28: 1/2/3, 135-161.

Engelmann, B., Hayden, E., Tasche, D., 2003, Measuring the Discriminative Power of Rating Systems, Discussion paper Series 2: Banking and Financial Supervision. No. 01/2003. Deutsche Bundesbank.

Financial Services Authority, FSA, 2005, Stress Testing, Discussion paper 05/02. May.

Gross, D., Souleles N.S., 2002, An Empirical Analysis of Personal Bankruptcy and Delinquency, The Review of Financial Studies, Spring 2002 Vol. 15, 319-347.

Intrater, M., 2002, Stress testing the commercial loan portfolio, Risk Management Association Journal. Volume 8, May.

López, J.A., 2005, Stress Tests: Useful Complements to Financial Risk Models, Federal Reserve Bank of San Francisco Economic Letter, No. 2005-14, June.

Sobehart, J., Keenan, S., Stein, R., 2000, Benchmarking Quantitative Default Risk Models: A Validation Methodology, Moody's Investor Service.

Sobehart, J., Keenan, S., 2001, Measuring Default Accurately, Credit Risk Special Report, Risk 14.

Whitley J., Windram, R., Cox, P., 2004, An empirical model of household arrears, Bank of England Working Paper no.214.