

**MODELLING THE DISTRIBUTION
OF CREDIT LOSSES
WITH OBSERVABLE
AND LATENT FACTORS**

2007

Gabriel Jiménez and Javier Mencía

**Documentos de Trabajo
N.º 0709**

BANCO DE ESPAÑA
Eurosistema



**MODELLING THE DISTRIBUTION OF CREDIT LOSSES WITH OBSERVABLE
AND LATENT FACTORS**

MODELLING THE DISTRIBUTION OF CREDIT LOSSES WITH OBSERVABLE AND LATENT FACTORS (*)

Gabriel Jiménez and Javier Mencía

BANCO DE ESPAÑA

(*) This paper is the sole responsibility of its authors and the views represented here do not necessarily reflect those of the Bank of Spain. We are grateful to Jesús Saurina, for very helpful comments and suggestions. The suggestions of an anonymous referee have also substantially improved the paper. Of course, all remaining errors are entirely ours. Address for correspondence: Alcalá 48, 28014 Madrid, Spain, tel.: +34 91 338 5414, fax: +34 91 338 6102.

The Working Paper Series seeks to disseminate original research in economics and finance. All papers have been anonymously refereed. By publishing these papers, the Banco de España aims to contribute to economic analysis and, in particular, to knowledge of the Spanish economy and its international environment.

The opinions and analyses in the Working Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the INTERNET at the following website: <http://www.bde.es>.

Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

© BANCO DE ESPAÑA, Madrid, 2007

ISSN: 0213-2710 (print)

ISSN: 1579-8666 (on line)

Depósito legal: M.19559-2007

Imprenta del Banco de España

Abstract

This paper develops a flexible and computationally efficient model to estimate the credit loss distribution of the loans in a banking system. We consider a sectorial structure, where default frequencies and the total number of loans are allowed to depend on macroeconomic conditions as well as on unobservable credit risk factors, which can capture contagion effects between sectors. In addition, we also model the distributions of the Exposure at Default and the Loss Given Default. We apply our model to the Spanish credit market, where we find that sectorial default frequencies are affected by a persistent latent factor. Finally, we also identify the potentially riskier sectors and perform stress tests.

Keywords: Credit risk, Probability of default, Loss distribution, Stress test, Contagion.

JEL: G21, E32, E37

1 Introduction

During the last years, a more volatile and dynamic financial environment has caused an increasing concern about the stability of banking systems. In this sense, it is widely agreed that credit risk is one of the variables that are more directly related to financial stability. Indeed, the Basel II framework has put forward the need of measuring this type of risk accurately. As a consequence, there has been a number of papers that estimate the credit loss distributions for different countries. To cite a few examples, Boss (2002) has developed a credit risk model for Austria, Virolainen (2004) has considered the case of Finland, Drehmann (2005) and Drehmann, Patton, and Sorensen (2006) have studied the credit loss distribution in the U.K., while Pesaran, Schuermann, Treutler, and Weiner (2006) have considered an international credit risk model.

All of these papers emphasise the need of assessing the variability of credit risk across different sectors. In addition, since the early works of Wilson (1997a,b), most subsequent studies relate changes in the probabilities of default to changes in macroeconomic conditions. Specifically, it is usually assumed that, conditional on the macroeconomic explanatory variables, defaults are independent across sectors. However, this assumption might yield strongly biased results if a relevant factor is omitted. What is more important, on top of macroeconomic variables, there might exist some credit risk factors that induce contagion across sectors, but which we cannot directly observe. This issue has already been a cause of concern in the literature. For instance, Schuermann and Stiroh (2006) have found an important presence of “hidden risk factors” in U.S. banks stock returns. Hence, correlations across sectors will be underestimated if these factors are not considered. Furthermore, their effect might be even higher if they are persistent, as Duffie, Eckner, Horel, and Saita (2006) have noticed.

This paper proposes a model of credit risk that accounts for the presence of persistent unobserved factors. The model has four stochastic components: default frequencies, the size of the credit portfolio, the exposures at default and the losses given default. The importance of modelling the size of the credit portfolio has been traditionally neglected in credit risk models. However, it is necessary to model this variable if we want to study the total credit losses of a banking system, and not just those due to a fixed number of

loans. In this sense, we find that the size of the credit market grows more during economic upturns than during economic downturns. In consequence, it is important to account for this feature, specially to carry out realistic stress scenarios. For each sector, our model assumes that changes in the default frequencies and the total number of loans are a function of past observations of these variables, a set of macroeconomic characteristics, two unobserved potentially persistent common latent factors and one idiosyncratic component. The two latent factors are independent from market conditions, but they can be correlated with each other. However, to ensure the identification of the model, we assume that only one of the factors affects the sectorial default frequencies, whereas the other one only influences the growth of the credit portfolio. This model can be easily estimated by means of a linear Kalman filter. Conditional on default, the loss given default and the exposure at default are initially assumed to be independent of default rates and the size of the credit market, although they are allowed to have a different distributional shape for each sector. With the exception of Madan and Unal (2006) in the context of deposit insurance, the literature has paid little attention to the distribution of exposures. However, we believe that it is necessary to account for the variability of exposures within each sector in order to correctly describe the heterogeneity of loans. Specifically, we employ either the Inverse Gaussian or the Gamma distribution. Both are rather flexible distributions. Importantly for our purposes, we show that their statistical properties can be exploited to reduce by a considerable amount the computational time required to simulate our model. Additionally, we propose a generalisation in which these distributions can change as a function of the observable macroeconomic factors. Finally, we consider the usual Beta distribution for the loss given default (see e.g. Gupton and Stein, 2002).

We use our model to estimate the credit loss distribution of the Spanish banking system. We have quarterly loan data from 1984.Q4 to 2006.Q4, obtained from the Spanish Credit Register. This database contains information on every loan granted in Spain with an exposure above €6,000. Hence, we have high quality loan data at a frequency at which it is not usually available. We consider 10 corporate sectors plus one mortgage sector and a sector of consumption loans to individuals. We first estimate a simple model with GDP growth and changes in three-month interest rates as our macroeconomic factors. Then, we obtain the credit loss distribution by simulating losses from our model under

the current economic conditions and under some stressed scenarios. Interestingly, we are able to identify a persistent unobservable factor that generates dependence between sectorial default frequencies, and a correlated common effect that affects the growth of the number of loans. These factors remain significant when we reestimate our model with an augmented set of macroeconomic characteristics, such as an interest rate spread, sectorial gross value added changes and the unemployment rate. We also determine which sectors are riskier, and compare our model with simpler versions that have been previously implemented. In this sense, a series of specification tests on the correlations matrix of default frequencies show that latent factors are crucial to capture intersectorial correlations. In addition, we assess the out-of-sample stability of our model by means of Kolmogorov tests. Finally, we explore the relationship between exposures at default and macroeconomic conditions, where we find that they tend to be higher on average during recessions than during expansions. The pro-cyclicality of exposures at default is consistent with the results obtained by Jiménez, López, and Saurina (2007), who find, also for the Spanish loan market, that a higher usage rate of credit lines during recessions induces higher exposures at default in these periods.

In summary, we believe that our paper provides some important contributions to the literature. Firstly, to the best of our knowledge, this is the first paper that introduces unobservable common shocks in a credit risk model for loans. Secondly, the paper takes advantage of the use of a very rich dataset which contains precise information about almost all the loans granted in the Spanish economy. In particular, we are able to model the distribution of exposures at default, as well as the loan market dynamics. In addition, we consider an extensive sectorial structure that includes mortgage loans. Thirdly, our results show that expected losses, and especially value at risk, can be significantly underestimated if contagion effects between sectors are not allowed. Finally, we dramatically reduce the computational demands of our model by exploiting its statistical properties.

The rest of the paper is organised as follows. We describe our model in the next section, and discuss the estimation of its parameters in Section 3. In Section 4, we consider an empirical application to Spanish loan data. Finally, concluding remarks and directions for future research are suggested in Section 5. Auxiliary results are gathered in an appendix.

2 The credit risk model

We are interested in modelling credit risk in an economy with K sectors. We will consider a sample of T periods of data. Following the usual notation, the losses due to a loan i from sector k can be decomposed at any time period t as

$$L_{i,k,t} = D_{i,k,t} LGD_{k,t} EAD_{i,k,t},$$

where $D_{i,k,t}$ is a binary variable that equals 1 in case of default and 0 otherwise, while $LGD_{k,t}$ and $EAD_{i,k,t}$ are, respectively, the loss given default and the exposure at default. We will denote as p_{kt} the proportion of non-performing loans in sector k at time t , i.e. the ratio of the number of loans in default to the total number of loans in each sector. This variable is usually known as default frequency. Hence, the losses from sector k at time t can be expressed as

$$L_{k,t} = \sum_{i=1}^{n_{k,t}} L_{i,k,t} = LGD_{k,t} S_k(p_{kt} n_{k,t}),$$

where $n_{k,t}$ is the total number of loans in sector k and

$$S_k = \sum_{i=1}^{\lfloor p_{kt} n_{k,t} \rfloor} EAD_{i,k,t}. \quad (1)$$

where $\lfloor p_{kt} n_{k,t} \rfloor$ rounds $p_{kt} n_{k,t}$ to the nearest integer. If we assume that the probability of default is constant in each sector, p_{kt} will converge to the probability of default of sector k as $n_{k,t}$ grows to infinity. However, for small $n_{k,t}$, they will not necessarily coincide.

The main dynamic features of our model are introduced through p_{kt} and n_{kt} . For every sector, we define the growth of the number of loans as $\Delta n_{kt} = \log(n_{kt}) - \log(n_{kt-1})$, and transform p_{kt} into $y_{kt} = \Phi^{-1}(p_{kt}) - \Phi^{-1}(p_{kt-1})$, where $\Phi^{-1}(\cdot)$ is the inverse of the cumulative standard normal distribution function. This choice is consistent with a probit model, although a logit model could also be adopted. We propose the following vector autoregression for the evolution of Δn_{kt} and y_{kt} :

$$\Delta n_{kt} = \alpha_{1,k} + \sum_{j=1}^q \rho_{1,j} \Delta n_{kt-j} + \sum_{j=1}^r \gamma'_{1,j} \mathbf{x}_{t-j} + \beta_{1,k} f_{1,t} + u_{1,kt}, \quad (2)$$

$$y_{kt} = \alpha_{2,k} + \sum_{j=1}^q \rho_{2,j} y_{kt-j} + \sum_{j=1}^r \gamma'_{2,j} \mathbf{x}_{t-j} + \beta_{2,k} f_{2,t} + u_{2,kt}, \quad (3)$$

where \mathbf{x}_t are m macroeconomic variables whose dynamics are given by

$$\mathbf{x}_t = \delta_0 + \sum_{j=1}^s \mathbf{A}_j \mathbf{x}_{t-j} + \mathbf{v}_t, \quad (4)$$

and $f_{1,t}$ and $f_{2,t}$ are two unobservable common factors independent from the macroeconomic explanatory variables. If we define the vector $\mathbf{f}_t = (f_{1t}, f_{2t})'$, the dynamics of \mathbf{f}_t can be expressed in terms of the following VAR(1) model:

$$\mathbf{f}_t = \mathbf{R}\mathbf{f}_{t-1} + \mathbf{w}_t. \quad (5)$$

In consequence, the evolution of Δn_{kt} and y_{kt} depends on their previous history, a set of observable characteristics \mathbf{x}_t , two unobservable common factors \mathbf{f}_t and the idiosyncratic shocks $u_{1,kt} \sim N(0, \sigma_{1k}^2)$ and $u_{2,jt} \sim N(0, \sigma_{2k}^2)$, for $j, k = 1, \dots, K$. These idiosyncratic terms are assumed to be jointly Gaussian such that $\text{cov}(u_{1,kt}, u_{2,jt}) = 0$ for $k \neq j$. Thus, we only allow for correlation between the two idiosyncratic factors from the same sector. In addition, these terms are independent from the vectors of common shocks \mathbf{v}_t and \mathbf{w}_t , which are also multivariate Gaussian random vectors with $\text{cov}(\mathbf{v}_t, \mathbf{w}_t) = \mathbf{0}$, and $V(\mathbf{v}_t) = \mathbf{\Omega}$. Since \mathbf{f}_t is unobservable, we need to impose a restriction on its scale so that the model is identified. In particular, we will parametrise \mathbf{R} and $V(\mathbf{w}_t)$ as

$$\mathbf{R} = \begin{bmatrix} \phi_1 & 0 \\ 0 & \phi_2 \end{bmatrix}$$

and

$$V(\mathbf{w}_t) = \begin{bmatrix} 1 - \phi_1^2 & \rho \sqrt{(1 - \phi_1^2)(1 - \phi_2^2)} \\ \rho \sqrt{(1 - \phi_1^2)(1 - \phi_2^2)} & 1 - \phi_2^2 \end{bmatrix}.$$

Hence, ϕ_i is the first order autocorrelation of f_i , $i=1,2$, and ρ is the conditional correlation between $f_{1,t}$ and $f_{2,t}$.

Finally, we will suppose that, conditional on default and the current macroeconomic conditions, $LGD_{k,t}$ are random Beta variates, while $EAD_{i,k,t}$ are independent Inverse Gaussian or Gamma variates.¹ We will first suppose that the parameters of these distributions are constant over time but possibly different for each sector. This implies that their distributions do not depend on the cycle. Later on, we will extend this model by

¹We have compared the empirical performance of these two distributions with other potential candidates. Our results show that the Gamma and the Weibull yield a similar empirical fit, while the shapes generated by the IG are similar to those of the log-normal. These results are available on request. However, we will not consider the Weibull and the Log-normal because they are not closed under aggregation.

allowing the mean of $EAD_{i,k,t}$ to depend on the macroeconomic factors. Specifically, if we denote the mean of the exposures at default in sector k and period t as μ_{kt} , we propose the following parametrisation:

$$\mu_{kt} = \mu_{kt-1} \exp \left[\eta_k + \boldsymbol{\varphi}'_k \mathbf{v}_{t-1} - \frac{1}{2} \boldsymbol{\varphi}'_k \boldsymbol{\Omega} \boldsymbol{\varphi}_k \right] \quad (6)$$

where η_k captures a time trend, \mathbf{v}_{t-1} is the lagged vector of innovations in equation (4) and $\boldsymbol{\Omega}$ is its covariance matrix. Thus, we allow μ_{kt} to be influenced by the same shocks that affect \mathbf{x}_t . Of course, if $\boldsymbol{\varphi}_k = \mathbf{0}$ we are back in the initial setting. The time trend component turns out to be important for estimation purposes. For example, in a context of historically decreasing exposures, this component will be negative. However, when we compute the credit loss distribution, we will assume no particular trend. In consequence, it is important to include the term $\boldsymbol{\varphi}'_k \boldsymbol{\Omega} \boldsymbol{\varphi}_k / 2$ in (6) to ensure that

$$E \left[\exp \left[\boldsymbol{\varphi}'_k \mathbf{v}_{t-1} - \frac{1}{2} \boldsymbol{\varphi}'_k \boldsymbol{\Omega} \boldsymbol{\varphi}_k \right] \right] = 1.$$

This result, which is a consequence of the normality of \mathbf{v}_t , ensures the constancy of the unconditional mean of (6) when η_k is set to zero. It is also possible to consider a dynamic parametrisation of the distribution of the loss given default (see Bruche and González-Aguado, 2006). However, due to lack of data in our application, we will not be able to explore this extension.

3 Estimation and simulation of the model

To estimate the parameters in (2) and (3), we need to use the Kalman filter to deal with the unobserved factors (see e.g. Hamilton, 1994, for a formal discussion). The intuition of this procedure is as follows. To evaluate the likelihood at each period t , we first compute the expected value of the factors given the information available up to time $t - 1$:

$$\mathbf{f}_{t|t-1} = E(\mathbf{f}_t | \{\Delta \mathbf{n}_s, \mathbf{y}_s, \mathbf{x}_s\}_{1 \leq s \leq t-1}),$$

where $\Delta \mathbf{n}_s = (\Delta n_{1,s}, \dots, \Delta n_{K,s})'$ and $\mathbf{y}_s = (y_{1,s}, \dots, y_{K,s})'$. In addition, since $\mathbf{f}_{t|t-1}$ is a noisy estimate of the true realisation \mathbf{f}_t , we also need to measure the uncertainty of this estimate:

$$\mathbf{P}_{t|t-1} = V[\mathbf{f}_t | \{\Delta \mathbf{n}_s, \mathbf{y}_s, \mathbf{x}_s\}_{1 \leq s \leq t-1}].$$

Finally, the estimation procedure consists basically in treating (2) and (3) as a pure vector autoregressive model, by using the series of $\mathbf{f}_{t|t-1}$ as if they were actually observed. However, we must adjust the variance of the model with $\mathbf{P}_{t|t-1}$ to account for the fact that $\mathbf{f}_{t|t-1}$ is not equivalent to the true realisation \mathbf{f}_t .

Interestingly, as new data arrives, we can update our previous estimates of the realisations of the factors, and obtain more accurate ones. For example, given the whole sample of data, we can estimate the evolution of the latent factors as:

$$\mathbf{f}_{t|T} = E(\mathbf{f}_t | \{\Delta \mathbf{n}_s, \mathbf{y}_s, \mathbf{x}_s\}_{1 \leq s \leq T}).$$

To identify the factors, we need at least two sectors. In fact, the more sectors we have, the more precise our estimates of \mathbf{f}_t will be. Hence, latent factors are particularly valuable in models with many sectors, since they allow for rich dynamics and correlation structures without requiring too many parameters.

As we have remarked, we consider two possible distributions for $EAD_{i,k,t}$: the Inverse Gaussian (IG) and the Gamma distribution. For each sector, we choose the one that best fits the data from the sector. Their parameters are estimated by maximum likelihood, where their density functions can be expressed as:

$$f_{IG}(EAD_{i,k,t} = x; \mu_k, \lambda_k) = \left(\frac{\lambda_k}{2\pi x^3} \right)^{1/2} \exp \left[-\frac{\lambda_k}{2\mu_k^2 x} (x - \mu_k)^2 \right] \quad (7)$$

$$f_{Gamma}(EAD_{i,k,t} = x; \nu_k, \tau_k) = \frac{(x/\tau_k)^{\nu_k/2-1}}{2^{\nu_k/2} \Gamma(\nu_k/2) \tau_k} \exp \left(\frac{-x}{2\tau_k} \right) \quad (8)$$

We will denote these distributions as $IG(\mu_k, \lambda_k)$ and $Gamma(\nu_k, \tau_k)$, respectively. In the IG case μ_k is the mean, and μ_k^3/λ_k is the variance, whereas for the Gamma distribution the mean is $\nu_k \tau_k$ and the variance $\nu_k \tau_k^2$. The subindices indicate that these parameters are sector specific. As we show in the empirical application, both distributions provide a good fit of the data, although the IG generally outperforms the Gamma. In addition, it can be shown that sums of *iid* IG or Gamma variates remain within the same family (see Johnson, Kotz, and Balakrishnan, 1994). Due to this property, we can express the distribution of S_{kt} in closed form for a given number of defaults $\lfloor p_{kt} n_{kt} \rfloor$. Specifically, it can be shown that the distribution of S_{kt} conditional on the number of defaults at t is a $IG[\lfloor p_{kt} n_{kt} \rfloor \mu_k, \lfloor p_{kt} n_{kt} \rfloor^2 \lambda_k]$ in the IG case, while it is a $Gamma(\lfloor p_{kt} n_{kt} \rfloor \nu_k, \tau_k)$ in the Gamma case. From this result, we can express the distribution of the sum of EAD's given

only the information known at $t - s$ by means of the following sum:

$$f(S_{kt} | I_{t-s}) = \sum_{i=0}^{\infty} g(S_{kt} | p_{kt}n_{kt} = i, I_{t-s}) \Pr(\lfloor p_{kt}n_{kt} \rfloor = i | I_{t-s}) \quad (9)$$

where $g(S_{kt} | \lfloor p_{kt}n_{kt} \rfloor = i, I_{t-s})$ denotes the conditional distribution of S_{kt} given i defaults occurring at t , while $f(S_{kt} | I_{t-s})$ is the distribution of S_{kt} given the information known at $t - s$, denoted as I_{t-s} . Finally, $\Pr(\lfloor p_{kt}n_{kt} \rfloor = i | I_{t-s})$ is the probability of i defaults occurring at t given I_{t-s} .

Unfortunately, we cannot compute (9) in closed form because it is extremely difficult to obtain the exact values of $\Pr(\lfloor p_{kt}n_{kt} \rfloor = i | I_{t-s})$ due to the dynamic features of the model followed by p_{kt} and n_{kt} . Moreover, when we consider the dynamic parametrisation (6) for the means of exposures at default, we will only be able to express $g(S_{kt} | p_{kt}n_{kt} = i, I_{t-s})$ in closed form for $s = 1$. Due to this complexity, we will have to compute the credit loss distribution by simulation. However, the IG and the Gamma distributions offer important computational advantages.² In particular, thanks to their properties, we do not need to simulate individual exposures at default, but just their sum S_{kt} , which will severely speed up the computation of the credit loss distribution.

4 Empirical application

We use loan data from the Credit Register of the Bank of Spain (CIR). This database records monthly information about all the loans granted by credit institutions in Spain (banks, savings banks, cooperatives and credit finance establishments) for a value above €6,000. Although the database offers a wider amount of information, we will focus on the particular details directly related to our application (see Jiménez and Saurina, 2004, and Jiménez, Salas, and Saurina, 2006, for a thorough description). In particular, the database reports the amount drawn and available for each loan, and whether its borrower is an individual or a company. In the latter case, the specific economic sector to which the borrower belongs is reported as well. There is also information available about the state of the loans. Every new loan is assigned a code which only changes if its situation deteriorates or if it matures. A loan that is expected to fail in the near future is classified as “doubtful”. If the loan eventually defaults, every month the database reports the time

²See in the Appendix how to simulate IG variates.

elapsed since its default. In particular, we will know whether it has been in default from 3 to 6, 6 to 12, 12 to 18, 18 to 21, or more than 21 months.

From the CIR, we have obtained quarterly series from 1984.Q4 to 2006.Q4 of sectorial default frequencies (p_{kt}), the total number of loans per sector (n_{kt}) and the exposures of the defaulted loans. Most papers usually focus on corporate loans. Typically, this is due to lack of available data on loans to individuals. However, we believe that loans to individuals, and specially mortgages, play an important role in the credit loss distribution of banks. In consequence, we consider 2 sectors for individuals and 10 corporate sectors. For individuals, we consider a group of mortgages and a group of consumption loans. For corporate loans, we define the following groups: (1) Agriculture, livestock and fishing; (2) Mining; (3) Manufacture; (4) Utilities; (5) Construction and real estate; (6) Commerce; (7) Hotels and restaurants; (8) Transport, storage and communications; (9) Renting, computer science and R&D. Finally, those companies that cannot be classified in any of the previous sectors are gathered in an additional sector denoted as Other Corporates (10). However, we remove from the database all the companies from the financial sector, because of their particular characteristics.

In each quarter, we compute the default rates as the ratio of the number of loans that have been in default from 3 to 6 months to the total number of loans in each sector. Thus, we are consistent with the Basel II framework. Those loans that have been in default for more than 6 months are left out because they were already considered in one of the previous quarters. Thus, only newly defaulted loans are considered at each period. Additionally, every quarter we also collect the exposures of those loans that have been in default from 3 to 6 months.

We have plotted the historical evolution of default frequencies in Figure 1 (a). For the sake of comparability, we have also plotted quarterly series of the Spanish GDP annual growth in Figure 1 (c), and 3-month real interest rates in Figure 1 (d).³ We can observe an increasing trend of default frequencies in all sectors from the end of the 1980s until almost the mid 1990s. This period coincides with a strong recession in the Spanish economy which had its trough in 1993, as we can observe in Figure 1 (c). In addition, interest rates also increased from 4% in 1988 to values above 8% in the first

³Following the methodology of Davidson and MacKinnon (1985), we have obtained real interest rates from the nominal rates and inflation.

half of the 1990's. Figure 1 (a) shows that loans to construction companies and hotels were more affected than the rest, with default frequencies peaking at 4%. From 1995 to the present, economic conditions have steadily improved, except for a brief period from 2000 to 2001. Interest rates have experienced a sharp decline in the last decade due to the convergence and integration in the European Monetary Union, and GDP growth has remained positive and less volatile than in the past (see Martín, Salas, and Saurina, 2005, for a more detailed analysis). As a consequence, during this expansionary period default frequencies have dropped to the lowest historical values in the sample. Under the current conditions, hotels and communications are the two sectors with higher default frequencies. In comparison, defaults in the construction sector are remarkably low at the moment.

In Figure 1 (b), we can observe the quarterly series of the total number of loans in each sector. The loan market has steadily grown in all sectors during the sample period under analysis. From this impressive growth it is not difficult to conclude that assuming a constant number of loans could yield inaccurate results. In addition, if we take a closer look at this figure, we can see that the rate of growth decreased for almost all sectors in the first half of the previous decade, that is, during the last recession. In consequence, the evolution of these variables seems to be correlated with the economic cycle. However, this conjecture will have to be confirmed with more formal results.

4.1 A simple model with two macroeconomic factors

We will start with a simple model that only considers two macroeconomic factors: quarterly change in real GDP growth and the variation of three-month real interest rates.⁴ We employ these two factors because they are generally regarded in the literature as the most important macroeconomic determinants of credit risk fluctuations. In addition, in this first set of estimations, we will assume that the parameters of the distribution of the exposures are constant over time.

Default frequency and market size growth. Let us consider the estimation of (2) and (3).⁵ We will introduce the lags 2, 3 and 4 of our two macroeconomic variables. To

⁴A similar analysis has been conducted with nominal interest rates yielding similar results, which are available on request.

⁵Prior to estimation, we have conducted a series of unit root tests on the data (see Breitung and Pesaran, 2005, for a review of this literature). Our results have shown us that we need to model default

save parameters, we do not include the first lag, because we obtain insignificant estimates for this lag once the subsequent 3 lags are considered. As for the autoregressive structure, we consider the effect of the first lag of the dependent variables, as well as a seasonal effect of the dependent variables one year before. In our estimation, we set to zero the intercept parameters of the default frequency, i.e. α_{2k} , for $k = 1, \dots, 12$. This is due to the fact that a positive (negative) intercept would imply that default frequencies would tend to 1(0) in the long run. Thus, our restriction rules out these extreme cases. Finally, we consider three dummies whose values are 1 in 1988.Q1, 1988.Q4 and 1996.Q2, respectively, and zero otherwise.⁶ These dummies are intended to capture the effects of historical exogenous changes in the database (see Delgado and Saurina, 2004, for a formal justification).

The estimates of the default frequency model are reported in Table 1, whereas analogous results for the evolution of the size of the credit portfolio can be found in Table 2. Intuitively, an increase in GDP growth tends to reduce default frequencies and induce the expansion of the loan market. This is why we observe that GDP growth generally has a negative impact on the variation of default frequencies and a positive effect on the growth of the credit market. As Table 1(a) shows, the effect of GDP on default frequencies seems to be more important for agriculture, manufacture, construction, commerce, communications and consumption loans, with the first two lags being highly significant. Nevertheless, we also obtain significant coefficients in all the remaining sectors except for mining and utilities, which react less to the cycle. Some sectors seem to respond more slowly to aggregate shocks. For instance, we only observe a significant effect on R&D and mortgages three quarters after a shock to GDP has occurred. In Table 2(a), we can observe that the effect of GDP on the size of the credit market is smaller, although it is still significant for manufacture, construction, commerce, and R&D.

As regards interest rates, higher values generally tend to increase default frequencies, where significant coefficients are obtained for hotels and communications. However, the overall effect of higher interest rates on the size of the loan industry is less clear. In some cases, they may even strengthen its growth. Nevertheless, from a theoretical point of view, it is unclear how interest rates should affect the growth of the number of loans. On

rates and the total number of loans in first differences to ensure their stationarity.

⁶The first dummy only affect mortgages, the second dummy affects mortgages and consumption loans, whereas the third dummy affects all sectors.

the one hand, higher interest rates will reduce the demand of loans. On the other hand, on the supply side banks will have incentives to grant more loans if interest rates rise. Nevertheless, the effect of interest rates seems to be less important than the impact of GDP. This may well be due to the fact that, until very recently, most Spanish borrowers, either corporates or individuals, preferred fixed to variable interest rates. For instance, in 1992 only 26.11% of the credit granted in Spain was linked to variable interest rates. This proportion has steadily increased in subsequent years, reaching 55.02% in 2000, and 74.47% in 2005. However, the predominant fixed interest rates for most of our sampling period have surely weakened the impact of interest rates variations in our model.

The last column of Tables 1(a) and 2(a) report the loadings of the unobservable factors. Although we consider two latent factors, we have explained in section 2 that f_{2t} only affects default frequencies, whereas f_{1t} exclusively alters the size of the credit portfolio. As we can see, we obtain significant estimates for both factors in all sectors. In addition, we find a significant correlation of -0.477 between f_{1t} and f_{2t} (see Table 3). In consequence, a high value of f_{2t} in a given quarter will induce an increase in default frequencies in all sectors. Moreover, through the negative correlation with f_{1t} , it will tend to cause a reduction in the growth of the loan market. Likewise, a low (negative) value of f_{1t} would produce a similar effect. Hence, f_{1t} and f_{2t} are able to capture a presence of contagion between sectors that the observable factors cannot account for.⁷ Furthermore, the time series structure of these factors also deserves some attention. We report in Table 3 the autoregressive structure of the observable and unobservable factors. As we can observe, f_{2t} has a significant first order autocorrelation of 0.209. Hence, since shocks to f_{2t} tend to persist through time, their effect on default frequencies will die away slowly. In contrast, f_{1t} has a significant negative autocorrelation of -0.190 . In consequence, the effect of a shock to f_{2t} will tend to be reverted in the following periods. For the observable factors, we find a positive (first order) autocorrelation for interest rates, and a negative autocorrelation for GDP growth.

We report the remaining parameters of the model in the lower panels of Tables 1 and 2. The third column of Table 1 (a) shows that the marginal effect of lagged default frequencies from the previous quarter is negative, whereas the seasonal effect (fourth column) is positive when it is significant. In contrast, both terms are generally positive

⁷Notice that the latent factors are independent from the observable factors by construction.

in the market size equation, as Table 2 (b) shows. We also report in the first column of Table 2 (b) the positive intercept terms that we obtain for the market size growth, which are consistent with the growth of the loan market already documented in Figure 2. Finally, as shown in the last column, the correlation between the idiosyncratic terms $u_{1,kt}$ and $u_{2,kt}$ is generally negative in the significant cases. Hence, shocks that increase the growth of the number of loans in a particular sector tend to be correlated with declines in the rate of defaults from the same sector.

These results can be compared with the estimates reported in Tables 4 and 5, which correspond to a restricted version of our model, where no latent factors are considered. GDP and interest rates have a qualitatively similar impact in this model. However, the absence of latent factors causes an increase in the correlations between the idiosyncratic terms of default frequencies and loan market growth in each sector (see the last column of Tables 4 (b) and 5 (b)).

Exposure at default. For each sector, we estimate the parameters of the IG and the Gamma distributions by maximum likelihood. We assume that these parameters remain constant over time. Since we want to focus on the current situation, we only use the exposures of defaulted loans from 2006 to fit the parameters of these distributions. Prior to estimation, we have adjusted the data for inflationary effects. In Figures 2 and 3 we compare the empirical fit of both distributions at the right tail with a Kernel estimate of the density. Except for mortgages, the Inverse Gaussian distribution provides a better fit in all sectors. In consequence, we will model the exposures of non-performing mortgages with the Gamma distribution, and employ the Inverse Gaussian in the remaining cases.

Loss given default. Unfortunately, we do not have data on the loss given default of the loans in our database. However, Spanish banks have reported the historical average loss given default for corporate, consumption and mortgage loans to the QIS5.⁸ Using this data, we choose the parameters of the Beta distribution so that the mean loss given default is 35% for corporates, 25% for consumption loans and 15% for mortgages. Finally, we choose 20% as the standard deviation in the three cases.

⁸Fifth Quantitative Impact Study of the Basel Committee on Banking Supervision.

Credit loss distribution. We estimate the credit loss distribution by simulating losses from our model. For each quarter of the horizon that we consider, we first obtain draws of the total number of loans and the default rates per sector. In particular, we use (2) and (3), where we sample the idiosyncratic terms from their joint Gaussian distribution, and generate the draws of the observable and latent common factors by means of (4) and (5), respectively. Finally, given the total number of defaults, we can generate random replications of (1) and the loss given default from their respective distributions. To ensure the stability of our results, we obtain one million simulated losses from our model.

We report descriptive statistics of the credit loss distribution in Table 6 for the model with latent factors. Specifically, we focus on the expected loss, the Value at Risk (VaR) at the 99.9% level and the unexpected loss, defined as the difference between the first two measures. We consider three different time horizons: 1, 3 and 5 years.⁹ We can see that, due to higher uncertainty, the three measures increase more than proportionately as the horizon increases. In terms of expected losses, consumption loans is the riskiest group for short horizons, followed by construction and manufacture. However, for longer horizons mortgages and specially construction also have high expected losses. These three sectors are also the riskiest ones in terms of unexpected losses, specially for long horizons. Again, the VaR of the construction sector seems to grow relatively more with the horizon than in the other cases. This is due to the strong dependence of this sector on cyclical effects, as we already observed in Tables 1 and 2.

Table 7 reports analogous results for the model without latent factors. The differences between sectors are qualitatively similar in this model. For instance, construction and consumption loans are still the riskiest categories. In addition, if we view each sector individually, there are not large quantitative discrepancies between the two models. However, as the last row of the table shows, total losses are much lower in this model, specially for longer horizons. This is due to the fact that we are underestimating contagion effects across sectors when we do not consider the unobservable factors. For example, the value at risk at a three year horizon is about 15% larger in the model with latent factors than in the model with only observable explanatory variables. Graphically, we perform a similar

⁹These horizons start at the end of December 2006, because we are conditioning on the final date of our sample. For instance, three-year horizon losses add all losses that occur up to three years after the start date.

comparison in Figure 4, where we plot the total credit loss densities for the two models. Again, we can observe that the model that allows for unobservable factors has fatter tails.

4.2 Extensions and robustness checks

To begin with, we will determine whether we are still able to identify contagion through latent factors when we consider a richer set of observable explanatory variables. Specifically, we will consider, as an additional common factor, the spread between six-year and three-month interest rates. This variable, related to the slope of the term structure of interest rates, will affect all sectors. Moreover, we consider six additional variables that will only have an impact on those sectors that are more related to these characteristics. In particular, we allow the change in the unemployment rate to affect consumption loans and mortgages; gross value added of market services will affect communications, hotels and commerce; gross value added of industry will affect manufacture and mining; and the gross value added of agriculture, energy and construction will affect agriculture, utilities and construction, respectively. The coefficients obtained with this specification are displayed in Tables 8 and 9. We can observe some significant values for the impact of the spread variable, specially in the evolution of the growth of the number of loans. Specifically, a steepening of the term structure seems to induce an expansion of the number of loans in some sectors. Unfortunately, at least in terms of statistical significance, most of the sectorial factors yield somewhat unsatisfactory results. Nevertheless, in spite of the additional factors, we still obtain highly significant factor loadings for the unobservable effects.

We will now compare the performance of the three different specifications that we have considered in our VAR model. To do so, we will compute the residuals of the default frequency equation (3) for the three cases.¹⁰ The specification that does not include latent factors assumes that these residuals are uncorrelated because intersectoral correlations are only captured by the observable common characteristics in this case. In contrast, the model with latent factors introduces a factorial structure for these correlations. We test in Table 10 whether the empirical correlations of the residuals are equal to those hypoth-

¹⁰For the sake of brevity, we focus only on default frequencies. However, we have obtained similar results with the residuals of the equation for the number of loans, which are available upon request.

esised by each of these specifications. As we can observe in Panel (a), most correlations are not adequately captured when latent factors are neglected. In contrast, Panels (b) and (c) show that these unobservable effects are able to yield a very accurate fit of the empirical correlations. Although these results show the good in-sample performance of our model, we are also interested in assessing its out of sample reliability. We will consider the period from 2004.Q1 to 2006.Q4 for this analysis. Hence, we need to reestimate the three specifications of our VAR model using only data up to 2003.Q4. With these estimates, we again compute the residuals (3), but in this case we will also consider those of (2). We could use these residuals to compute tests analogous to those of Table 10. However, since we only have 12 periods, these tests will have low power. Thus, we prefer to follow a different approach in this case. In particular, we standardise the residuals with the inverse of the Cholesky factorisation of their hypothesised covariance matrices under each specification. The resulting values should be *iid* standard normal under a correct specification. We check this hypothesis in Table 11 by means of a Kolmogorov test. This table shows that the null can be easily rejected when we do not consider latent factors, but it can no longer be rejected once these factors are included. Hence, this result confirms that our model will still be reliable in out-of-sample exercises.

Finally, we will explore the linkages between aggregate macroeconomic shocks and the distribution of exposures at default. We have estimated by maximum likelihood the parameters of the IG distribution, substituting (6) for μ_k in (7). Although we have also estimated an analogous model with the Gamma distribution, we do not report the results for this model due to its poor empirical fit. For the sake of parsimony, we will only consider the effect of the innovations to GDP growth and real interest rate variations. The results are displayed in Table 12. As expected, the estimated means at the end of our sample period, displayed in the first column of Table 12, reflect the differences between the loan sizes across sectors. Specifically, loans to individuals, either mortgages or consumption loans, are characterised by small mean exposures when compared to the much larger sizes of loans to corporates. As regards corporates, the more capital intensive sectors have larger mean exposures. For instance, utilities is a sector with relatively few but very large loans. We can also observe in the second column that the time trend coefficients are generally negative though small in magnitude. Imposing $\eta_k = 0$ in these

estimations would have yielded unstable estimates of the factor loadings. Specifically, the interest rates would then be forced to capture the time effects, because of their decreasing historical trend (see Figure 1d). In the third column, we can observe that GDP generally has a negative and significant effect. In consequence, higher GDP growth will tend to reduce the magnitude of exposures at default on average. Conversely, these exposures will be higher during economic downturns. As for interest rates, we generally obtain positive coefficients. Hence, higher interest rates tend to increase the means of the exposures. These results are consistent with the use of credit lines as a liquidity management tool by firms, as Jiménez, López, and Saurina (2007) show. Moreover, the observed dependence of EAD on the business cycle can reinforce the pro-cyclicality of the Basel II framework. The impact of Basel II on pro-cyclicality has been extensively debated in the literature.¹¹ The main conclusion is that the minimum capital requirements computed under the Internal Ratings Based (IRB) approach will be more risk-sensitive under Basel II, increasing during recessions and falling as the economy enters expansions. Thus, this will make the lending decisions of banks more pro-cyclical, which, in turn, will amplify the economic cycle. In this sense, our results support the concerns of this literature about the strong relationship between economic cycles and credit risk. However, the global impact of Basel II on the financial stability of the banking system is an issue beyond the scope of this paper.

4.3 Stress tests

We will end this empirical study by assessing the consequences of a strong shock to either GDP or interest rates. We follow the standard practice in stress testing exercises and introduce artificial shocks in the vector of innovations of the factors (see (4)). In particular, we stress our model with a 3-standard deviation shock that occurs in the first quarter of the period under study. We consider separate shocks to each of the two macroeconomic factors that we stress. The GDP shock will be negative, whereas the interest rate shock will be positive. Thus, these tests are designed to induce a recession in both cases.

As in the previous sections, we will start with our baseline model, in which GDP and interest rates are the only observable characteristics. We report in Table 13 the

¹¹See for instance Goodhart (2005), Goodhart and Taylor (2005), Gordy and Howels (2006), Kashyap and Stein (2004) and Ayuso, Pérez, and Saurina (2004)

percentage change in the expected loss and the VaR caused by these shocks. The effect of the GDP shock is similar for most sectors, although it is relatively larger for manufacture, construction and mortgages, and smaller for utilities. In contrast, due to its poorer explanatory power, the interest rate shock causes more heterogeneous responses. In Table 14, we compare these results with the ones obtained from our two extensions. In the first extension we assess the effect of including the augmented set of macroeconomic factors, while in the second one we analyse the impact of modelling the dynamics of the mean of the exposures at default. In both cases, we allow for the presence of latent factors, although in the latter extension we only consider our specification with two observable factors. In addition, we assume that the unconditional means of the exposures at default will remain constant over time.¹² The two models that use a static distribution for exposures at default yield fairly close results. Indeed, both seem to respond more to a GDP shock than to an interest rate shock. For example, at a three-year horizon, the expected loss and the value at risk increase by 17% under the GDP shock, but only by 7-8% under the interest rate shock. This result is a direct consequence of the much higher explanatory power of GDP (see Tables 1, 2 and 8).

In contrast, we find larger effects when we allow for time varying means of exposures at default. Although the expected loss and the VaR under normal conditions are similar, we now obtain fatter tails at the five-year horizon, where VaR almost reaches €50 billion. We also find a higher sensitivity to the GDP and interest rate shocks. These larger losses are mainly due to two sources. Firstly, exposures at default deteriorate as the economy worsens, whereas in the previous models they remained unaltered. Secondly, we have introduced contagion between default frequencies and exposures at default, since both of them are influenced by the macroeconomic factors. For instance, increments in default frequencies due to a lower GDP growth are reinforced with higher exposures at default. In consequence, the overall effect is fatter tails and larger responses to stress tests of the same magnitude.

¹²Hence, we directly simulate from (6), by imposing $\eta_k = 0$, because we do not expect that the downward trend of the last decade will persist in the future.

5 Conclusions

We develop a flexible model to estimate the credit loss distribution of a national banking system. We classify the credit market in sectors, and model default frequencies, exposures at default, losses given default and the total number of loans in each sector. This latter variable has not been previously considered in the literature. However, we believe that the growth of the credit industry may have important effects on total credit losses, specially for medium and long term horizons. We propose a dynamic model for default frequencies and the growth of the credit industry, using as explanatory variables a set of macroeconomic factors. As a distinguishing feature of our approach, we also allow for the presence of unobservable common factors. These factors are able to capture contagion effects between sectors, which are orthogonal to the observable macroeconomic conditions. Both observable and unobservable variables are modelled with a vector autoregressive structure. Finally, conditional on default, exposure at default and loss given default are assumed to be independent of the previous variables. Loss given default is modelled with a Beta distribution. As for the exposures at default, we use the Gamma and the Inverse Gaussian distributions, and propose a dynamic parametrisation that relates their expected values to macroeconomic shocks.

In the second part of the paper we apply our model to analyse the loss distribution of the total credit portfolio of Spanish banks. We use quarterly loan data from the Spanish Credit Register. Our database starts in 1984.Q4 and ends in 2006.Q4. It contains information on every loan granted in Spain with an exposure above €6,000. Hence, we are able to analyse the whole Spanish loan market. We consider 10 corporate sectors. Furthermore, we also investigate the role of consumption loans and mortgages in the credit loss distribution by including one additional sector for each of these groups. We first study a simple model that uses the quarterly changes in GDP growth and the variation in three-month real interest rates as the only macroeconomic explanatory variables. Exposures are modelled in a static setting for each sector with the Inverse Gaussian distribution, except for mortgages, where we employ the Gamma because it provides a better fit. We estimate the parameters by maximum likelihood and obtain the credit loss distribution for the 1, 3 and 5 year horizons by simulation. Despite the analytical complexity of our

model, we show that we can generate extremely fast simulations by exploiting the statistical properties of the Gamma and the Inverse Gaussian distributions. In particular, we compute for each sector the expected loss, the unexpected loss and the value at risk of credit losses. We also estimate the density function of losses. Our results show that credit losses in the Spanish economy are mainly due to the manufacture, construction, consumption loans and mortgages. The result for the latter two sectors should be interpreted in absolute terms. Despite the typically low losses given default and exposures at default in loans to individuals, the numbers of loans in these sectors are so large that they are one of the main sources of credit risk in Spain. At the other extreme, mining and utilities are the sectors with lower absolute risk in Spain. We compare our results with the losses generated by a simpler model that does not take into account the presence of “hidden” factors. Although the two models provide similar results for sectorial losses viewed separately, aggregate or total losses are larger in the more general setting, due to the higher correlation between sectors, introduced by the latent factors. In this sense, we show by means of in and out of sample specification tests that latent factors capture the intersectoral correlations very accurately, whereas a model with only observable variables misses important contagion effects. Furthermore, we are also able to find a significant impact of macroeconomic cycles on the distribution of exposures at default.

Finally, we perform two stress tests to assess the sensitivity of credit losses to macro shocks. In particular, we assess the separate effects of a sudden drop in GDP growth and a sharp increase in interest rates. Both shocks occur in just one quarter, and they have a magnitude of three standard deviations. Overall, the GDP stress test has a stronger effect than the interest rate shock. However, we obtain a higher sensitivity once we account for the dependence of exposures at default on the cycle.

A fruitful avenue for future research would be to integrate this credit risk model with market risk and operational risk models, as Rosenberg and Schuermann (2006) propose. It would also be interesting to combine our model with a model for the interbank market, such as the ones developed by Goodhart (2005) and Elsinger, Lehar, and Summer (2006). These types of general models could be extremely helpful in providing analytical systemic risk measures.

References

- Ayuso, J., D. Pérez, and J. Saurina (2004). Are capital buffers procyclical? evidence from spanish panel data. *Journal of Financial Intermediation* 13, 249–264.
- Boss, M. (2002). A macroeconomic credit risk model for stress testing the Austrian credit portfolio. *Financial Stability Report of the Austrian National Bank* 4, 64–82.
- Breitung, J. and M. H. Pesaran (2005). Unit roots and cointegration in panels. Forthcoming in L. Matyas and P. Sevestre “The Econometrics of Panel Data (Third Edition)”, Kluwer Academic Publishers.
- Bruche, M. and C. González-Aguado (2006). Recovery rates, default probabilities and the credit cycle. CEMFI Working Paper No. 0612.
- Davidson, R. and J. G. MacKinnon (1985). Testing linear and loglinear regressions against Box-Cox alternatives. *Canadian Journal of Economics* 18, 499–517.
- Delgado, J. and J. Saurina (2004). Riesgo de crédito y dotaciones a insolvencias. Un análisis con variables macroeconómicas. *Moneda y Crédito* 219, 11–41. Also available in English by request to the authors.
- Drehmann, M. (2005). A market based macro stress test for the corporate credit exposures of UK banks. Bank of England Working Paper.
- Drehmann, M., A. J. Patton, and S. Sorensen (2006). Corporate defaults and macroeconomic shocks: non-linearities and uncertainty. Bank of England Working Paper.
- Duffie, D., A. Eckner, G. Horel, and L. Saita (2006). Frailty correlated default. Working paper.
- Elsinger, H., A. Lehar, and M. Summer (2006). Risk assessment for banking systems. *Management Science* 52, 1301–1315.
- Goodhart, C. A. E. (2005). Financial regulation, credit risk and financial stability. *National Institute Economic Review* 192, 118–127.

- Goodhart, C. A. E. and A. Taylor (2005). Procyclicality and volatility in the financial system: The implementation of Basel II and IAS 39. In S. Gerlach and P. Gruenwald (Eds.), *Procyclicality of Financial Systems in Asia*. Palgrave Macmillan.
- Gordy, M. B. and B. Howels (2006). Procyclicality in Basel ii: can we treat the disease without killing the patient? *Journal of Financial Intermediation* 15, 395–417.
- Gupton, G. M. and R. M. Stein (2002). LosscaleTM: model for predicting Loss Given Default (LGD). Moody's KMV, New York.
- Hamilton, J. D. (1994). *Time series analysis*. Princeton: Princeton University Press.
- Jiménez, G., J. A. López, and J. Saurina (2007). Empirical analysis of corporate credit lines. Unpublished manuscript.
- Jiménez, G., V. Salas, and J. Saurina (2006). Determinants of collateral. *Journal of Financial Economics* 81, 255–281.
- Jiménez, G. and J. Saurina (2004). Collateral, type of lender and relationship banking as determinants of credit risk. *Journal of Banking and Finance* 28, 2191–2212.
- Johnson, N. L., S. Kotz, and N. Balakrishnan (1994). *Continuous univariate distributions*. New York: John Wiley and Sons.
- Kashyap, A. K. and J. C. Stein (2004). Cyclical implications of the Basel II capital standards. *Economic Perspectives* 1Q/2004, 18–31.
- Madan, D. B. and H. Unal (2006). Designing countercyclical and risk based aggregate deposit insurance premia. CFR Seminar Series Library, FDIC.
- Martín, V. Salas, and J. Saurina (2005). A test of the law of one price in retail banking. Bank of Spain Working Paper No. 0530.
- Pesaran, M. H., T. Schuermann, B. J. Treutler, and S. M. Weiner (2006). Macroeconomic dynamics and credit risk: a global perspective. *Journal of Money, Credit and Banking* 38, 1211–1261.

Rosenberg, J. V. and T. Schuermann (2006). A general approach to integrated risk management with, skewed, fat-tailed risks. *Journal of Financial Economics* 79, 569–614.

Schuermann, T. and K. J. Stiroh (2006). Visible and hidden risk factors for banks. Financial Institutions Center, Wharton School of Business. Working paper No. 06-10.

Virolainen, K. (2004). Macro stress testing with a macroeconomic credit risk model for Finland. Bank of Finland discussion paper No. 18.

Wilson, T. C. (1997a). Portfolio credit risk (I). *Risk* 10, 56–61.

Wilson, T. C. (1997b). Portfolio credit risk (II). *Risk* 10, 111–117.

Appendix

Simulation of Inverse Gaussian variates

Most commercial packages provide fast simulation algorithms to generate gamma random variates. However, it is less common to find generators of Inverse Gaussian variates in these packages. Nevertheless, it is fairly easy to generate random IG draws. Consider the following steps:

1. Generate u : a uniform random variable between 0 and 1.
2. Generate s : a chi-square random variable with 1 degree of freedom, independent of u .
3. Define:

$$\begin{aligned} r_1 &= \frac{\mu}{2\lambda} \left[2\lambda + \mu s - \sqrt{4\lambda\mu s + \mu^2 s^2} \right] \\ r_2 &= \frac{\mu^2}{r_1} \end{aligned}$$

4. Then

$$x = \begin{cases} r_1 & \text{if } u < \frac{\mu}{\mu+r_1} \\ r_2 & \text{if } u \geq \frac{\mu}{\mu+r_1} \end{cases}$$

is a random draw from the $IG(\mu, \lambda)$ distribution.

Table 1
Model for default frequencies with GDP, interest rates and latent factors
(a) Explanatory variables

	GDP_{t-2}	GDP_{t-3}	GDP_{t-4}	INT_{t-2}	INT_{t-3}	INT_{t-4}	f_{2t}
Agriculture	-1.151**	-1.147**	-0.436	-0.257	1.458**	-0.288	3.368**
Mining	-1.194	-1.277	0.117	0.332	0.333	-1.018	5.885**
Manufacture	-1.532**	-1.758**	-0.869*	0.402	0.678	-0.434	4.483**
Utilities	-0.137	0.048	-0.504	0.129	0.645	-0.747	5.225**
Construction	-0.970**	-1.000*	-0.881**	0.710	0.102	0.279	3.434**
Commerce	-1.280**	-1.227**	-0.612	-0.185	0.721	-0.091	4.064**
Hotels	-1.316**	-0.839	-0.147	-0.094	1.860**	-0.327	4.060**
Communications	-0.965**	-1.065**	-0.861*	0.150	1.133**	-0.409	3.698**
R&D	-0.417	-1.436**	-1.491**	0.172	-0.181	-0.063	3.732**
Other Corp.	-0.350	-0.905*	-0.258	0.666	0.881*	-0.189	3.233**
Cons. loans	-0.853**	-1.040**	-0.527	0.045	0.607	0.263	3.303**
Mortgages	-0.809	-1.623**	-1.321**	0.378	0.055	0.078	1.736**

(b) Dynamics

	α	$y_{k,t-1}$	$y_{k,t-4}$	Correlation with $\Delta n_{k,t}$
Agriculture	0	-0.360**	0.222**	0.429**
Mining	0	-0.325**	-0.069	0.020
Manufacture	0	-0.329**	-0.012	0.083
Utilities	0	-0.374**	-0.131	0.058
Construction	0	-0.080	0.175**	-0.354**
Commerce	0	-0.238**	0.038	0.051
Hotels	0	-0.341**	-0.004	0.145
Communications	0	-0.317**	0.121*	0.320**
R&D	0	-0.160**	0.072	-0.116
Other Corp.	0	-0.215**	0.150**	-0.323**
Cons. loans	0	-0.273**	-0.022	-0.305**
Mortgages	0	0.050	0.062	-0.163

Notes: Two asterisks indicate significance at the 5% level, while one asterisk denotes significance at the 10% level. Prior to estimation, the dependent and the explanatory variables have been multiplied by 100. GDP_{t-i} and INT_{t-i} for $i = 2, 3, 4$ denote, respectively, the effect of lagged observations of changes of GDP growth and three-month real interest rates on the dependent variables. α is the intercept of the VAR model, and the columns labelled $y_{k,t-1}$ and $y_{k,t-4}$ denote the effect of lagged observations of the dependent variables. “Correlation with $\Delta n_{k,t}$ ” refers to the correlation between the two idiosyncratic residuals that affect the same sector.

Table 2

Model for the growth of the number of loans with GDP, interest rates and latent factors
(a) Explanatory variables

	GDP_{t-2}	GDP_{t-3}	GDP_{t-4}	INT_{t-2}	INT_{t-3}	INT_{t-4}	f_{1t}
Agriculture	0.251	0.171	0.189	-0.199	0.058	-0.077	1.257**
Mining	0.199	-0.248	0.038	-0.059	-0.065	0.222	1.376**
Manufacture	0.386**	0.064	0.120	-0.076	-0.075	0.085	1.601**
Utilities	0.248	-0.109	-0.097	-0.865**	0.561	-0.502	1.211**
Construction	0.323*	0.088	0.137	-0.243	0.067	-0.129	1.471**
Commerce	0.465**	0.129	0.072	0.082	-0.202	0.153	1.795**
Hotels	0.213	-0.068	0.063	0.018	0.026	-0.248	1.993**
Communications	0.130	0.541	0.425	0.615	-0.114	0.133	2.072**
R&D	0.626**	0.227	-0.058	-0.059	-0.097	-0.205	1.592**
Other Corp.	-0.898**	-0.802*	0.205	0.351	-0.261	0.533	1.022**
Cons. loans	0.032	0.061	0.522*	0.508	0.311	0.035	0.784**
Mortgages	0.156	0.039	0.115	0.754**	-0.518	-0.123	0.592*

(b) Dynamics

	α	$\Delta n_{k,t-1}$	$\Delta n_{k,t-4}$	Correlation with $y_{k,t}$
Agriculture	1.306**	0.308**	0.130	0.429**
Mining	0.861**	0.293**	0.081	0.020
Manufacture	0.593**	0.374**	0.186**	0.083
Utilities	1.158**	0.194*	-0.191*	0.058
Construction	0.950**	0.575**	0.249**	-0.354**
Commerce	0.771**	0.447**	0.290**	0.051
Hotels	1.212**	0.286**	0.488**	0.145
Communications	0.801**	0.514**	0.253**	0.320**
R&D	1.512**	0.313**	0.416**	-0.116
Other Corp.	1.523**	0.477**	0.095	-0.323**
Cons. loans	2.379**	0.094*	0.032	-0.305**
Mortgages	2.632**	-0.023	0.235**	-0.163

Notes: Two asterisks indicate significance at the 5% level, while one asterisk denotes significance at the 10% level. Prior to estimation, the dependent and the explanatory variables have been multiplied by 100. GDP_{t-i} and INT_{t-i} for $i = 2, 3, 4$ denote, respectively, the effect of lagged observations of changes of GDP growth and three-month real interest rates on the dependent variables. α is the intercept of the VAR model, and the columns labelled $y_{k,t-1}$ and $y_{k,t-4}$ denote the effect of lagged observations of the dependent variables. “Correlation with $y_{k,t}$ ” refers to the correlation between the two idiosyncratic residuals that affect the same sector.

Table 3
Dynamics of the explanatory variables

	Intercept	First lag	Second lag	Conditional covariance matrix			
				GDP	INT	f_{1t}	f_{2t}
GDP	0.035	-0.425**	-0.056	1.259**			
INT	-0.094	0.549**	-0.511**	-0.117	0.933**		
f_{1t}	0	-0.190*	0	0	0	1	
f_{2t}	0	0.209*	0	0	0	-0.477**	1

Notes: Two asterisks indicate significance at the 5% level, while one asterisk denotes significance at the 10% level. Prior to estimation, the dependent and the explanatory variables have been multiplied by 100. GDP and INT denote, respectively, the changes of GDP growth and three-month real interest rates.

Table 4
Model for default frequencies with GDP and interest rates
(a) Explanatory variables

	GDP _{t-2}	GDP _{t-3}	GDP _{t-4}	INT _{t-2}	INT _{t-3}	INT _{t-4}	f _{1t}
Agriculture	-1.067**	-1.114**	-0.324	-0.076	1.346**	-0.030	0
Mining	-1.012	-1.193	0.211	0.747	0.259	-0.860	0
Manufacture	-1.523**	-1.623**	-0.685	0.677	0.683	-0.395	0
Utilities	-0.121	0.031	-0.396	0.539	0.357	-0.358	0
Construction	-0.786*	-0.713	-0.770*	1.207**	-0.316	0.614	0
Commerce	-1.211**	-1.036**	-0.431	0.091	0.702	-0.045	0
Hotels	-1.283**	-0.693	-0.016	0.177	1.710**	-0.124	0
Communications	-0.750*	-0.804	-0.651	0.590	0.996*	-0.189	0
R&D	-0.211	-1.375**	-1.455**	0.439	-0.438	0.221	0
Other Corp.	-0.304	-0.850*	-0.187	0.756	0.756	0.042	0
Cons. loans	-0.652*	-0.897**	-0.410	0.340	0.462	0.494	0
Mortgages	-0.821	-1.672**	-1.427**	0.555	-0.183	0.177	0

(b) Dynamics

	α	$y_{k,t-1}$	$y_{k,t-4}$	Correlation with $\Delta n_{k,t}$
Agriculture	0	-0.326**	0.480**	0.063
Mining	0	-0.332**	0.008	-0.362**
Manufacture	0	-0.234**	0.154	-0.458**
Utilities	0	-0.349**	-0.043	-0.105
Construction	0	0.049	0.398**	-0.255**
Commerce	0	-0.130	0.260**	-0.432**
Hotels	0	-0.300**	0.122	-0.228**
Communications	0	-0.243**	0.388**	0.085
R&D	0	-0.121	0.270**	-0.104
Other Corp.	0	-0.197*	0.324**	-0.244**
Cons. loans	0	-0.231**	0.193*	-0.023
Mortgages	0	0.033	0.114	-0.142

Notes: Two asterisks indicate significance at the 5% level, while one asterisk denotes significance at the 10% level. Prior to estimation, the dependent and the explanatory variables have been multiplied by 100. GDP_{t-i} and INT_{t-i} for $i = 2, 3, 4$ denote, respectively, the effect of lagged observations of changes of GDP growth and three-month real interest rates on the dependent variables. α is the intercept of the VAR model, and the columns labelled $y_{k,t-1}$ and $y_{k,t-4}$ denote the effect of lagged observations of the dependent variables. “Correlation with $\Delta n_{k,t}$ ” refers to the correlation between the two idiosyncratic residuals that affect the same sector.

Table 5

Model for the growth of the number of loans with GDP and interest rates

(a) Explanatory variables

	GDP_{t-2}	GDP_{t-3}	GDP_{t-4}	INT_{t-2}	INT_{t-3}	INT_{t-4}	f_{2t}
Agriculture	0.282	0.173	0.146	-0.223	-0.008	-0.113	0
Mining	0.200	-0.210	-0.044	-0.091	-0.111	0.194	0
Manufacture	0.457**	0.168	0.095	-0.159	-0.111	-0.022	0
Utilities	0.243	-0.084	-0.112	-0.835**	0.487	-0.475	0
Construction	0.393**	0.125	0.122	-0.301	0.017	-0.245	0
Commerce	0.516**	0.210	0.023	0.007	-0.232	0.013	0
Hotels	0.212	-0.087	-0.022	-0.020	0.004	-0.350	0
Communications	0.219	0.712*	0.465	0.789*	-0.109	0.052	0
R&D	0.794**	0.460*	-0.052	-0.153	-0.045	-0.417	0
Other Corp.	-0.911**	-0.841*	0.152	0.322	-0.264	0.528	0
Cons. loans	0.012	0.021	0.531*	0.505	0.312	-0.023	0
Mortgages	0.162	0.042	0.120	0.728**	-0.466	-0.158	0

(b) Dynamics

	α	$\Delta n_{k,t-1}$	$\Delta n_{k,t-4}$	Correlation with $y_{k,t}$
Agriculture	1.207**	0.208*	0.293**	0.063
Mining	1.032**	0.064	0.175*	-0.362**
Manufacture	0.563**	0.160	0.415**	-0.458**
Utilities	1.297**	0.113	-0.191	-0.105
Construction	0.774**	0.461**	0.523**	-0.255**
Commerce	0.635**	0.262**	0.548**	-0.432**
Hotels	0.980**	0.171*	0.643**	-0.228**
Communications	0.830*	0.446**	0.410**	0.085
R&D	1.068**	0.115	0.685**	-0.104
Other Corp.	1.684**	0.442**	0.089	-0.244**
Cons. loans	2.377**	0.071	0.084	-0.023
Mortgages	2.608**	-0.033	0.251**	-0.142

Notes: Two asterisks indicate significance at the 5% level, while one asterisk denotes significance at the 10% level. Prior to estimation, the dependent and the explanatory variables have been multiplied by 100. GDP_{t-i} and INT_{t-i} for $i = 2, 3, 4$ denote, respectively, the effect of lagged observations of changes of GDP growth and three-month real interest rates on the dependent variables. α is the intercept of the VAR model, and the columns labelled $y_{k,t-1}$ and $y_{k,t-4}$ denote the effect of lagged observations of the dependent variables. "Correlation with $y_{k,t}$ " refers to the correlation between the two idiosyncratic residuals that affect the same sector.

Table 6
Descriptive statistics of the credit loss distribution.
Model with GDP, interest rates and latent factors

	Expected loss			VaR(99.9%)			Unexpected loss		
	1 year	3 years	5 years	1 year	3 years	5 years	1 year	3 years	5 years
Agriculture	37.99	125.30	234.86	124.11	530.18	1245.32	86.12	404.88	1010.46
Mining	5.61	19.26	36.86	26.01	117.74	287.73	20.40	98.48	250.87
Manufacture	281.24	899.27	1608.76	937.94	3785.70	8042.71	656.70	2886.43	6433.95
Utilities	4.08	14.10	27.00	26.42	93.63	218.66	22.34	79.52	191.66
Construction	312.44	1101.08	2165.23	1030.65	4904.66	12322.61	718.21	3803.58	10157.38
Commerce	187.28	618.10	1141.26	608.05	2485.46	5430.45	420.77	1867.36	4289.18
Hotels	31.51	109.82	216.10	108.00	483.11	1161.68	76.48	373.29	945.58
Communications	36.26	121.13	228.56	120.09	535.13	1298.12	83.83	414.01	1069.55
R&D	67.73	247.09	512.00	234.13	1141.23	3039.56	166.39	894.14	2527.56
Other Corp.	33.32	115.26	226.22	109.48	519.88	1330.16	76.15	404.62	1103.94
Cons. loans	403.45	1371.83	2592.96	1669.15	6690.09	14564.31	1265.70	5318.26	11971.35
Mortgages	255.78	947.58	2006.46	1553.64	8426.95	24269.16	1297.86	7479.36	22262.70
Total	1656.71	5689.82	10996.30	3818.47	17032.79	42160.61	2161.76	11342.97	31164.31

Notes: results in millions of euros. The unexpected loss is defined as the difference between the VaR(99.9%) and the expected loss. Statistics obtained from 1 million simulations of the credit risk model.

Table 7
Descriptive statistics of the credit loss distribution.
Model with GDP and interest rates

	Expected loss			VaR(99.9%)			Unexpected loss		
	1 year	3 years	5 years	1 year	3 years	5 years	1 year	3 years	5 years
Agriculture	38.78	127.01	239.40	123.49	548.68	1424.32	84.71	421.67	1184.92
Mining	5.79	19.38	36.55	25.74	110.79	264.44	19.96	91.42	227.89
Manufacture	289.56	918.97	1638.13	940.92	3791.18	8250.55	651.36	2872.21	6612.42
Utilities	4.16	14.18	27.03	26.68	92.51	218.73	22.52	78.33	191.70
Construction	330.79	1213.63	2470.78	1075.92	5836.65	16746.27	745.14	4623.03	14275.50
Commerce	192.43	629.34	1157.27	605.88	2502.99	5651.97	413.45	1873.65	4494.70
Hotels	32.26	111.88	219.00	108.48	483.21	1183.03	76.23	371.32	964.03
Communications	36.07	119.51	231.18	118.57	583.80	1653.85	82.50	464.29	1422.67
R&D	68.88	250.12	524.28	232.59	1198.43	3415.74	163.72	948.32	2891.46
Other Corp.	33.96	117.22	230.81	108.93	523.95	1406.92	74.98	406.73	1176.11
Cons. loans	420.34	1448.00	2755.12	1716.47	7177.07	16384.08	1296.13	5729.07	13628.96
Mortgages	258.10	955.72	2028.67	1546.89	8333.75	24301.48	1288.79	7378.03	22272.82
Total	1711.10	5924.94	11558.21	3642.97	15651.14	38863.36	1931.88	9726.20	27305.14

Notes: results in millions of euros. The unexpected loss is defined as the difference between the VaR(99.9%) and the expected loss. Statistics obtained from 1 million simulations of the credit risk model.

Table 8
Model with latent factors, GDP, interest rates, spread and six sectorial effects

(a) Default frequencies													
	GDP_{t-2}	GDP_{t-3}	GDP_{t-4}	INT_{t-2}	INT_{t-3}	INT_{t-4}	SPR_{t-2}	SPR_{t-3}	SPR_{t-4}	SEC_{t-2}	SEC_{t-3}	SEC_{t-4}	f_{2t}
Agriculture	-0.938**	-1.118**	-0.480	0.640	0.301	0.501	0.680	-1.032	0.783	0.038	0.035	-0.036	3.350**
Mining	-0.907	-0.893	0.628	0.941	-1.018	-1.213	0.954	-2.049	-0.313	0.008	-0.543	-0.582*	5.233**
Manufacture	-1.366**	-1.480**	-0.600	0.702	-0.005	-0.841	0.274	-1.174	-0.873	-0.126	-0.209*	-0.186*	4.067**
Utilities	-0.442	0.356	-0.737	-1.344	2.358	-2.948**	-1.114	1.234	-2.424**	-0.140	0.209	-0.307	5.041**
Construction	-0.804*	-0.552	-0.861*	0.786	0.345	-0.292	0.267	0.158	-0.741	-0.190*	-0.002	-0.075	3.182**
Commerce	-1.173**	-1.221**	-0.915**	0.352	-0.043	-0.082	0.661	-0.858	-0.116	-0.086	0.452*	0.077	3.882**
Hotels	-1.220**	-0.529	-0.342	0.256	1.821*	-0.007	0.054	0.048	0.414	-0.688	0.256	0.402	4.140**
Communications	-0.835*	-1.150**	-1.140**	1.208	-0.163	0.223	1.052	-1.116	0.607	-0.020	0.344	0.282	3.682**
R&D	-0.470	-1.311**	-1.337**	-0.768	0.670	-0.574	-1.240	0.629	-0.732	-	-	-	3.675**
Other Corp.	-0.329	-0.900*	-0.187	0.088	1.101	0.251	-1.261*	0.105	0.331	-	-	-	3.317**
Cons. loans	-0.865**	-1.022**	-0.575	0.234	0.730	-0.188	0.484	0.172	-0.503	0.021	0.222	-0.573	3.241**
Mortgages	-0.881	-1.766**	-1.498**	1.825**	0.116	-0.615	2.756**	0.154	-0.460	-0.895	-0.352	0.191	1.926**

(b) Growth of the number of loans													
	GDP_{t-2}	GDP_{t-3}	GDP_{t-4}	INT_{t-2}	INT_{t-3}	INT_{t-4}	SPR_{t-2}	SPR_{t-3}	SPR_{t-4}	SEC_{t-2}	SEC_{t-3}	SEC_{t-4}	f_{1t}
Agriculture	0.189	0.176	0.199	-0.198	0.390	-0.341	0.190	0.367	-0.200	0.028	0.023	0.017	1.246**
Mining	0.120	-0.300	0.023	-0.047	0.306	-0.009	0.238	0.420	-0.033	0.072	-0.051	-0.036	1.303**
Manufacture	0.286*	0.034	0.107	-0.115	0.396	-0.308	0.245	0.508**	-0.235	0.055*	-0.038	-0.013	1.517**
Utilities	0.280	-0.101	-0.211	0.160	0.131	-0.863*	1.590**	-0.355	-0.428	0.076	0.034	-0.011	1.095**
Construction	0.216	0.031	0.112	-0.257	0.385	-0.489*	0.187	0.350	-0.302	0.012	0.000	0.070**	1.364**
Commerce	0.354**	0.177	0.178	-0.081	0.454	-0.357	0.056	0.712**	-0.447**	0.003	-0.162**	0.052	1.713**
Hotels	0.065	0.101	-0.080	-0.301	1.135**	-1.411**	0.225	1.070**	-0.999**	-0.285	0.213	0.063	1.777**
Communications	0.028	0.600	0.299	1.099*	-0.432	-0.332	0.916	-0.309	-0.586	-0.156	0.127	0.524*	1.954**
R&D	0.587**	0.241	-0.085	0.055	0.359	-0.435	0.344	0.630*	0.013	-	-	-	1.495**
Other Corp.	-0.838**	-0.850*	0.079	1.323**	-0.971	0.607	1.406**	-0.524	0.132	-	-	-	1.060**
Cons. loans	-0.008	0.071	0.503*	0.539	0.533	-0.209	0.218	0.258	-0.198	0.006	0.008	-0.046	0.724**
Mortgages	0.177	0.025	0.064	0.846*	-1.344**	0.253	-0.253	-0.798*	0.192	-0.221	0.078	0.206	0.782**

Notes: Two asterisks indicate significance at the 5% level, while one asterisk denotes significance at the 10% level. Prior to estimation, the dependent and the explanatory variables have been multiplied by 100. GDP_{t-i} , INT_{t-i} , SPR_{t-i} for $i = 2, 3, 4$ denote, respectively, the effect of lagged observations of GDP growth, the variation of three-month real interest rates, and the spread between six-year and three-month interest rates on the dependent variables. Except for R&D and Other Corp., each sector is additionally allowed to depend on an additional sectorial variable, whose effects are reported in the columns SEC_{t-i} . SEC denotes gross value added by sector for corporates and the unemployment rate for consumption loans and mortgages.

Table 9
Dynamics of the latent factors
Model with latent factors, GDP, interest rates, spread and six sectorial effects

	Intercept	First lag	Second lag								
GDP	0.029	-0.430**	0.017								
INT	-0.096	0.534**	-0.559**								
SPR	0.018	-0.068	-0.165*								
GVA _{Agriculture}	0.290	0.161*	0.057								
GVA _{Industry}	0.094	-0.214**	-0.030								
GVA _{Energy}	0.077	0.491**	0.145								
GVA _{Construction}	0.306	0.154*	-0.076								
GVA _{Services}	0.058	-0.281**	-0.052								
Unemployment	-0.016	0.255**	0.124								
f_{1t}	0.000	-0.287**	0.000								
f_{2t}	0.000	0.147	0.000								

	Conditional covariance matrix										
	GDP	INT	SPR	GVA _{Agr.}	GVA _{Ind.}	GVA _{Ene.}	GVA _{Con.}	GVA _{Ser.}	UNP	f_{1t}	f_{2t}
GDP	1.268**										
INT	-0.105	0.935**									
SPR	0.262*	0.135	1.367**								
GVA _{Agriculture}	1.038	0.430	0.217	32.089**							
GVA _{Industry}	0.502	-0.283	0.629*	-9.211**	8.087**						
GVA _{Energy}	-0.376	-0.675*	1.327**	-0.167	2.436**	14.685**					
GVA _{Construction}	0.196	-0.471	0.638	-4.605**	3.708**	5.285**	12.588**				
GVA _{Services}	0.103	0.009	0.325**	0.128	0.897**	1.465**	1.782**	1.617**			
Unemployment	-0.095	0.088	-0.048	0.323	-0.246	0.033	-0.373	-0.058	0.361		
f_{1t}	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.000	
f_{2t}	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.474**	1.000

Notes: Two asterisks indicate significance at the 5% level, while one asterisk denotes significance at the 10% level. GDP, INT, SPR and GVA denote, respectively, GDP growth, the variation of three-month real interest rates, the spread between six-year and three-month interest rates, and gross value added.

Table 10

P-values of specification tests of the correlation matrix of default frequencies

		(a) Model with GDP and Interest rates										
		1	2	3	4	5	6	7	8	9	10	11
Agriculture	1											
Mining	2	0.00										
Manufacture	3	0.00	0.00									
Utilities	4	0.00	0.00	0.00								
Construction	5	0.00	0.00	0.00	0.00							
Commerce	6	0.00	0.00	0.00	0.00	0.00						
Hotels	7	0.00	0.00	0.00	0.00	0.00	0.00					
Communications	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
R&D	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Other Corp.	10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
Cons. loans	11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Mortgages	12	0.03	0.16	0.06	0.06	0.01	0.11	0.27	0.10	0.01	0.00	0.00

		(b) Model with GDP, Interest rates and latent factors										
		1	2	3	4	5	6	7	8	9	10	11
Agriculture	1											
Mining	2	0.32										
Manufacture	3	0.68	0.03									
Utilities	4	0.80	0.70	0.94								
Construction	5	0.75	0.23	0.57	0.15							
Commerce	6	0.68	0.98	0.70	0.62	0.37						
Hotels	7	0.43	0.27	0.50	0.39	0.73	0.98					
Communications	8	0.69	0.55	0.86	0.96	0.46	1.00	0.75				
R&D	9	0.41	0.14	0.13	0.10	0.43	0.33	0.98	0.48			
Other Corp.	10	0.71	0.66	0.52	0.15	0.31	0.37	0.72	0.31	0.01		
Cons. loans	11	0.36	0.30	0.17	0.99	0.45	0.24	0.71	0.21	0.29	0.83	
Mortgages	12	0.79	0.68	0.56	0.55	0.47	0.46	0.55	0.57	0.42	0.41	0.19

		(c) Model with GDP, Interest rates, spread, six sectorial effects and latent factors										
		1	2	3	4	5	6	7	8	9	10	11
Agriculture	1											
Mining	2	0.37										
Manufacture	3	0.92	0.07									
Utilities	4	0.55	0.77	0.87								
Construction	5	0.73	0.15	0.44	0.30							
Commerce	6	0.94	0.98	0.62	0.96	0.73						
Hotels	7	0.74	0.40	0.56	0.65	0.83	0.92					
Communications	8	0.72	0.61	0.88	0.89	0.54	0.93	0.83				
R&D	9	0.35	0.36	0.20	0.38	0.65	0.34	0.79	0.68			
Other Corp.	10	0.59	0.99	0.63	0.61	0.59	0.39	0.49	0.40	0.11		
Cons. loans	11	0.22	0.37	0.19	0.48	0.30	0.27	0.71	0.28	0.50	0.97	
Mortgages	12	0.62	0.65	0.76	0.21	0.38	0.39	0.60	0.38	0.12	0.16	0.26

Notes: in each cell the null hypothesis is that the empirical correlation between the corresponding sectorial default frequencies equals the one hypothesised by the model. The p-values below 5% are expressed in bold.

Table 11

Kolmogorov specification tests of the out-of-sample distribution of the standardised residuals of the model of default frequencies and number of loans

Factors	Kolmogorov test	P-value
GDP, INT	0.115	0.001
GDP, INT, f_t	0.053	0.389
GDP, INT, SPR, SEC, f_t	0.041	0.712

Notes: The model has been estimated with data from 1984.Q4 to 2003.Q4. The test studies whether the orthogonalised residuals from 2004.Q1 to 2006Q4, a total number of 288 values, are independent standard normal. INT, SPR and SEC denote, respectively, real interest rates, interest rate effects and sectorial factors.

Table 12

Effect of macroeconomic factors on the expected exposures at default

	Mean in 2006.Q4	η_k	GDP_{t-1}	INT_{t-1}
Agriculture	0.107	-0.002	-0.054 **	0.131 **
Mining	0.089	-0.018 **	-0.011	0.059 *
Manufacture	0.096	-0.010 **	-0.029 **	0.041 **
Utilities	0.178	0.028	-0.150 **	-0.218 **
Construction	0.092	-0.021 **	-0.076 **	0.051 **
Commerce	0.090	-0.007 **	-0.043 **	0.024 **
Hotels	0.062	-0.023 **	-0.115 **	-0.026 *
Communications	0.054	-0.018 **	-0.061 **	-0.021 **
R&D	0.057	-0.014 **	-0.111 **	0.002
Other Corp.	0.094	-0.015 **	-0.029 **	-0.002
Cons. loans	0.016	-0.018 **	0.017 **	0.018 **
Mortgages	0.062	0.004 **	-0.042 **	0.022 **

Notes: Two asterisks indicate significance at the 5% level. Means in millions of euros. GDP and INT denote, respectively, GDP growth and the variation of three-month real interest rates. Data sample for the estimation: 1989.Q4 - 2006.Q4.

Table 13
Changes in the credit loss distribution caused by macroeconomic stress tests (3 standard-deviation shocks)
Model with GDP, interest rates and latent factors

	GDP shock						Interest rate shock					
	Expected loss			VaR(99.9%)			Expected loss			VaR(99.9%)		
	1 year	3 years	5 years	1 year	3 years	5 years	1 year	3 years	5 years	1 year	3 years	5 years
Agriculture	7	13	15	6	12	13	1	5	6	2	5	6
Mining	8	11	11	7	10	10	2	-1	-1	3	-1	-1
Manufacture	10	18	19	10	15	17	3	4	4	4	3	4
Utilities	0	2	2	1	2	2	1	-2	-2	1	-3	0
Construction	6	16	17	6	14	14	4	7	8	4	7	6
Commerce	7	13	14	6	12	13	1	3	3	0	3	3
Hotels	7	10	11	8	10	10	4	8	8	4	7	6
Communications	6	11	11	6	10	11	5	10	11	5	10	11
R&D	4	14	16	4	15	14	0	-2	-2	0	-2	-4
Other Corp.	7	16	18	7	16	18	7	14	15	7	13	14
Cons. loans	6	11	12	7	12	12	3	8	9	4	8	9
Mortgages	9	28	31	11	26	26	3	5	6	4	5	5
Total	7	16	18	8	17	19	3	6	6	4	7	6

Notes: percentage changes with respect to the normal scenario. The unexpected loss is defined as the difference between the VaR(99.9%) and the expected loss. Statistics obtained from 1 million simulations of the credit risk model. GDP is stressed with a negative 3 standard deviation shock, whereas interest rates are stressed with a positive shock of the same magnitude.

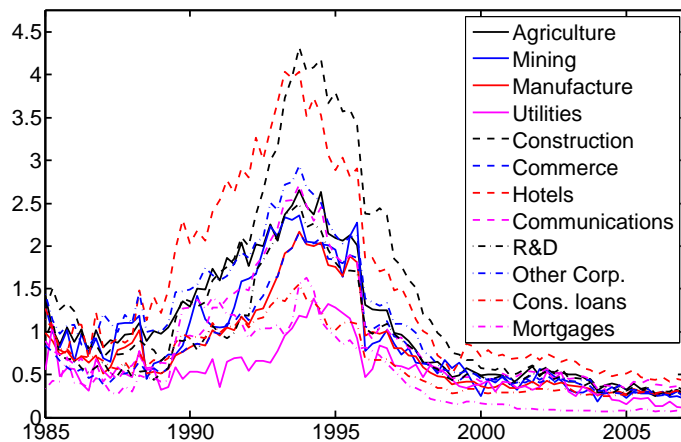
Table 14
Comparison of credit loss distributions

	(1)	(2)	(3)
Characteristics			
Included Factors			
-GDP, Interest rates	✓	✓	✓
-Spread, GVA's, Unemployment		✓	
Model of the distribution of exposures	Static	Static	Dynamic
Normal Scenario			
Expected loss			
1 year	1657	1652	1464
3 years	5690	5588	5103
5 years	10996	10711	10030
VaR (99.9%)			
1 year	3818	3797	3466
3 years	17033	16360	17344
5 years	42161	39104	47605
Change due to -3 s.d. GDP shock (%)			
Expected loss			
1 year	7	6	20
3 years	16	16	33
5 years	18	18	36
VaR (99.9%)			
1 year	8	7	17
3 years	17	17	33
5 years	19	20	37
Change due to +3 s.d. Interest rate shock (%)			
Expected loss			
1 year	3	6	10
3 years	6	6	15
5 years	6	6	16
VaR (99.9%)			
1 year	4	7	10
3 years	7	8	15
5 years	6	8	15

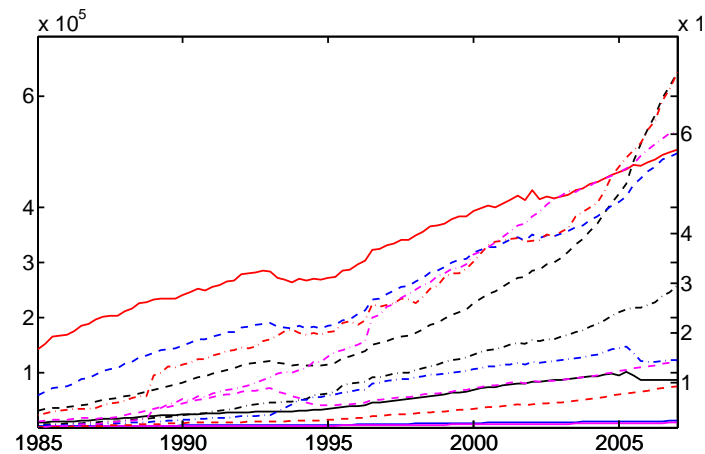
Notes: results in millions of euros. "Spread" denotes the difference between six-year and three-month interest rates. "GVA's" denotes gross value added factors, namely: agriculture, industry, energy, construction and market services. Statistics obtained from 1 million simulations of the credit risk model. All models include latent factors.

Figure 1:

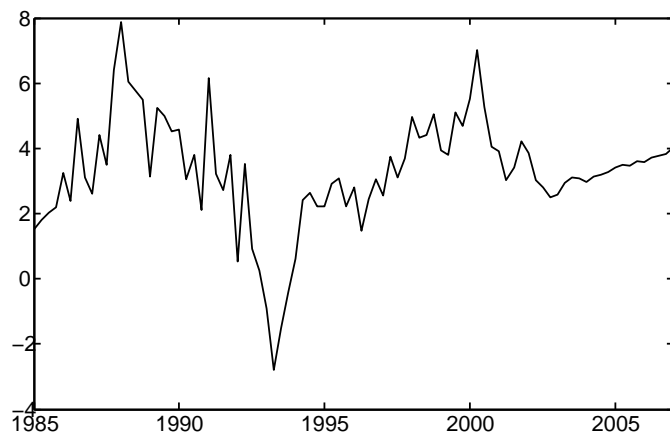
(a) Historical default frequencies in the Spanish Economy (%)



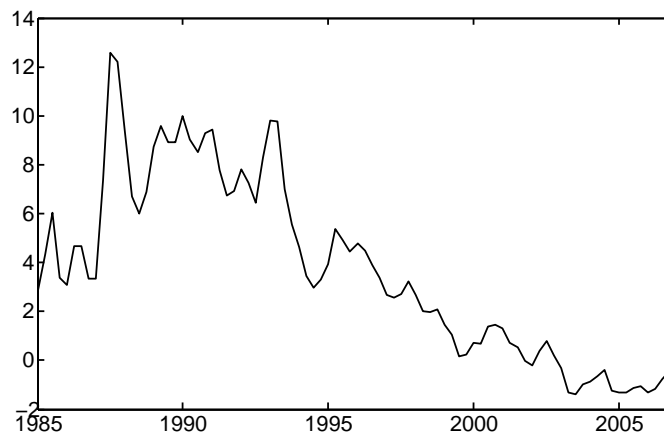
(b) Historical evolution of the total number of loans in the Spanish loan market



(c) Annual Spanish GDP growth (%)



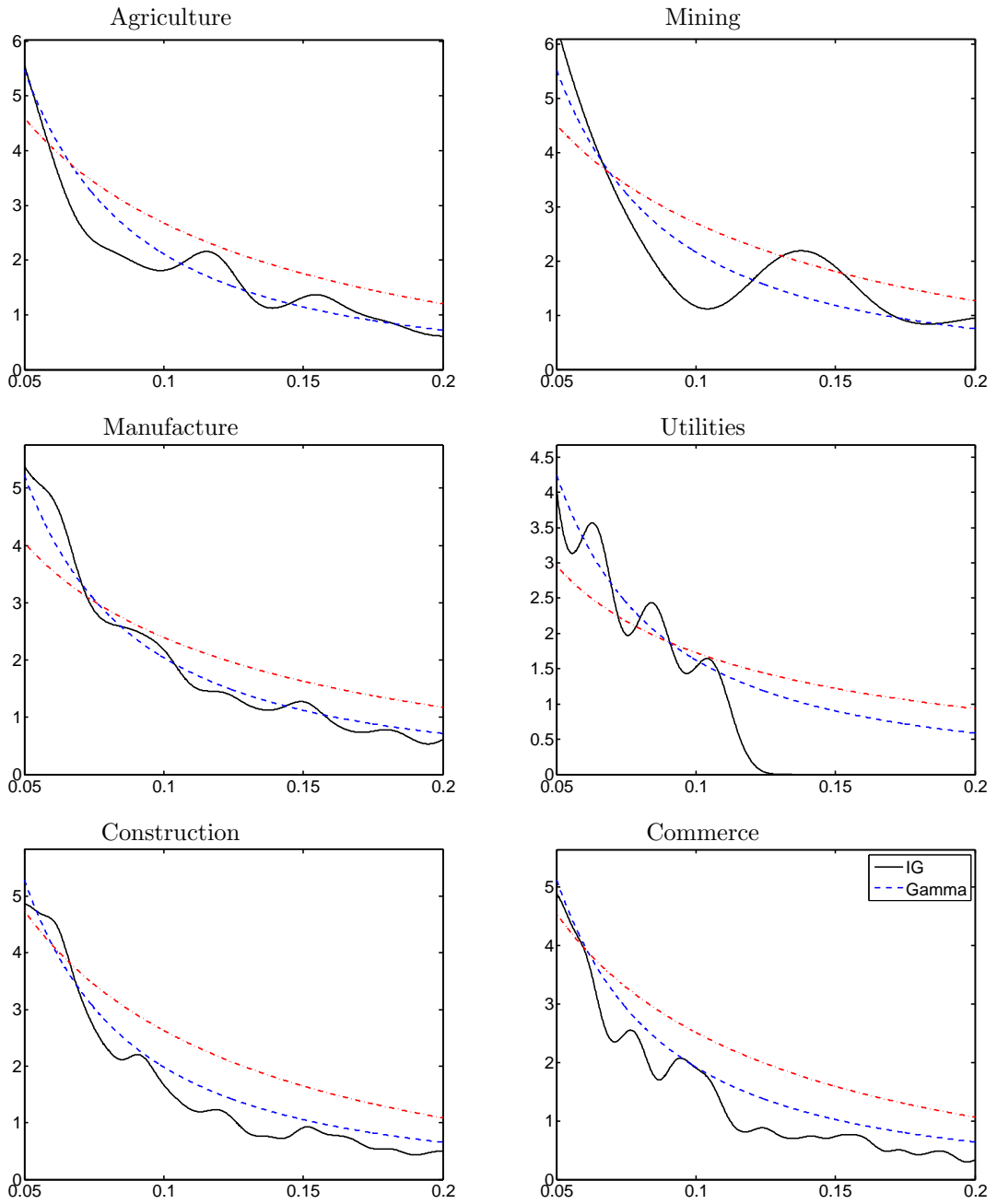
(d) Three-month real interest rates in Spain (%)



Notes: (a) and (b) share the same legends. The right scale on the y axis of figure (b) correspond to consumption loans and mortgages, whereas the left axis corresponds to the remaining cases.

Figure 2:

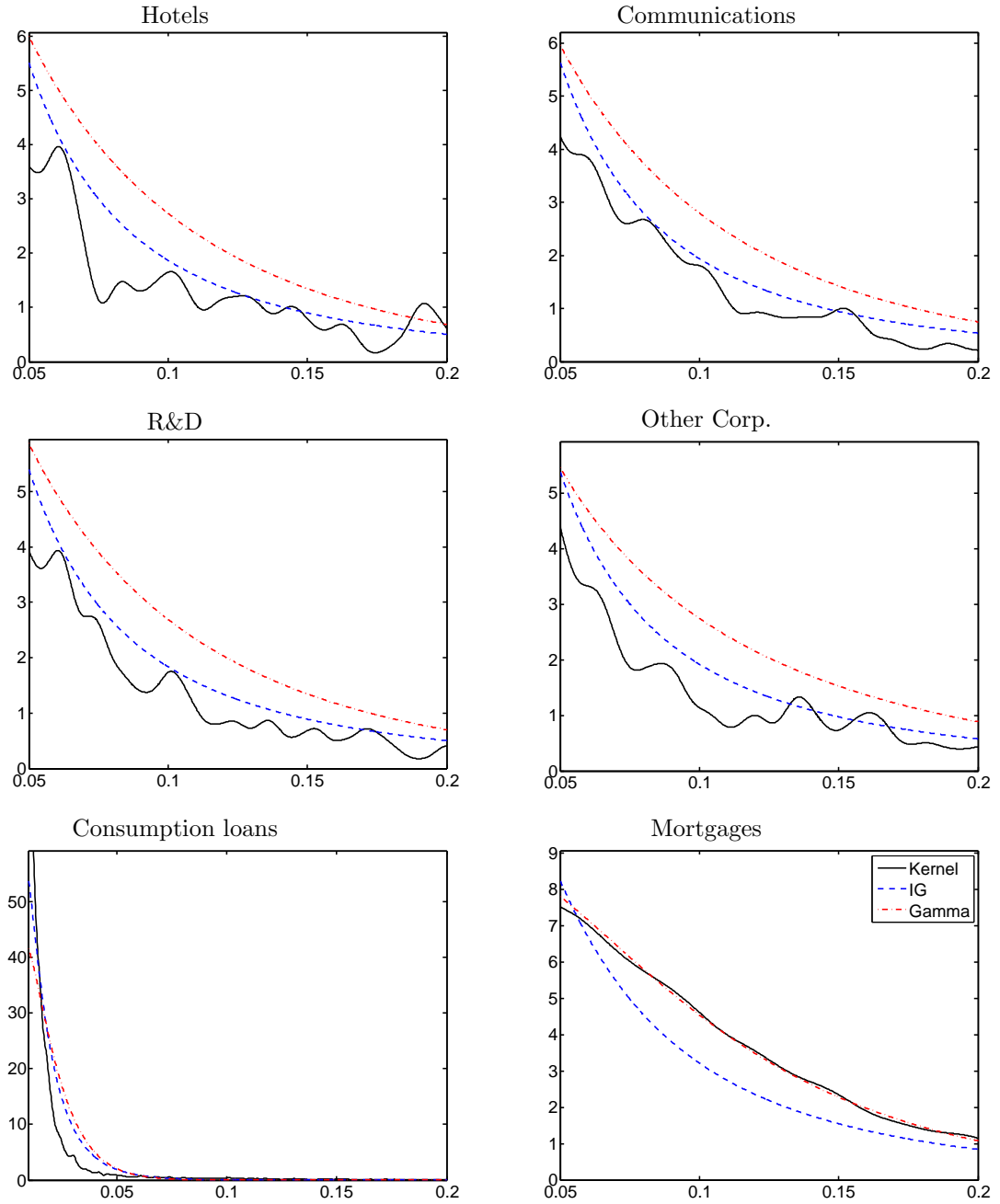
Kernel estimate and fitted densities of the right tail of the distribution of exposures at default



Notes: the x-axis is expressed in millions of euros. Both the kernel and the fitted densities are based on exposure data from 2001 to 2006.

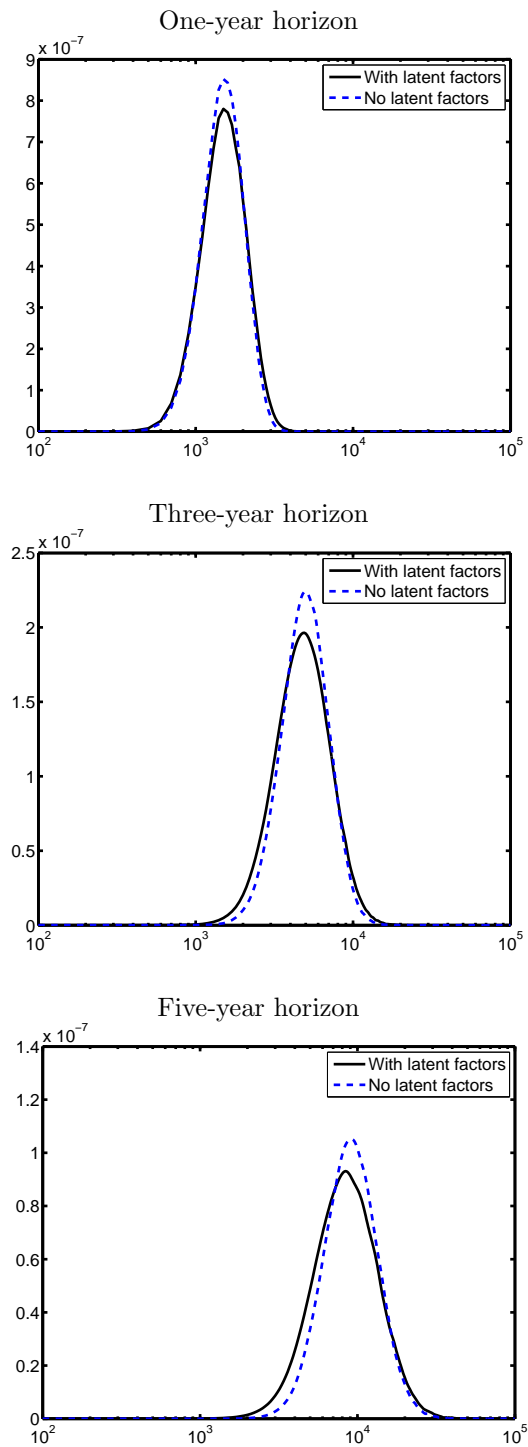
Figure 3:

Kernel estimate and fitted densities of the right tail of the distribution of exposures at default



Notes: the x-axis is expressed in millions of euros. Both the kernel and the fitted densities are based on exposure data from 2001 to 2006.

Figure 4:
Kernel estimates of the total credit loss distribution



Note: the x-axis is expressed in millions of euros, where a log-scale is employed. Estimates based on 100,000 simulations.

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS¹

- 0601 ARTURO GALINDO, ALEJANDRO IZQUIERDO AND JOSÉ MANUEL MONTERO: Real exchange rates, dollarization and industrial employment in Latin America.
- 0602 JUAN A. ROJAS AND CARLOS URRUTIA: Social security reform with uninsurable income risk and endogenous borrowing constraints.
- 0603 CRISTINA BARCELÓ: Housing tenure and labour mobility: a comparison across European countries.
- 0604 FRANCISCO DE CASTRO AND PABLO HERNÁNDEZ DE COS: The economic effects of exogenous fiscal shocks in Spain: a SVAR approach.
- 0605 RICARDO GIMENO AND CARMEN MARTÍNEZ-CARRASCAL: The interaction between house prices and loans for house purchase. The Spanish case.
- 0606 JAVIER DELGADO, VICENTE SALAS AND JESÚS SAURINA: The joint size and ownership specialization in banks' lending.
- 0607 ÓSCAR J. ARCE: Speculative hyperinflations: When can we rule them out?
- 0608 PALOMA LÓPEZ-GARCÍA AND SERGIO PUENTE: Business demography in Spain: determinants of firm survival.
- 0609 JUAN AYUSO AND FERNANDO RESTOY: House prices and rents in Spain: Does the discount factor matter?
- 0610 ÓSCAR J. ARCE AND J. DAVID LÓPEZ-SALIDO: House prices, rents, and interest rates under collateral constraints.
- 0611 ENRIQUE ALBEROLA AND JOSÉ MANUEL MONTERO: Debt sustainability and procyclical fiscal policies in Latin America.
- 0612 GABRIEL JIMÉNEZ, VICENTE SALAS AND JESÚS SAURINA: Credit market competition, collateral and firms' finance.
- 0613 ÁNGEL GAVILÁN: Wage inequality, segregation by skill and the price of capital in an assignment model.
- 0614 DANIEL PÉREZ, VICENTE SALAS AND JESÚS SAURINA: Earnings and capital management in alternative loan loss provision regulatory regimes.
- 0615 MARIO IZQUIERDO AND AITOR LACUESTA: Wage inequality in Spain: Recent developments.
- 0616 K. C. FUNG, ALICIA GARCÍA-HERRERO, HITOMI IIZAKA AND ALAN SU: Hard or soft? Institutional reforms and infrastructure spending as determinants of foreign direct investment in China.
- 0617 JAVIER DÍAZ-CASSOU, ALICIA GARCÍA-HERRERO AND LUIS MOLINA: What kind of capital flows does the IMF catalyze and when?
- 0618 SERGIO PUENTE: Dynamic stability in repeated games.
- 0619 FEDERICO RAVENNA: Vector autoregressions and reduced form representations of DSGE models.
- 0620 AITOR LACUESTA: Emigration and human capital: Who leaves, who comes back and what difference does it make?
- 0621 ENRIQUE ALBEROLA AND RODRIGO CÉSAR SALVADO: Banks, remittances and financial deepening in receiving countries. A model.
- 0622 SONIA RUANO-PARDO AND VICENTE SALAS-FUMÁS: Morosidad de la deuda empresarial bancaria en España, 1992-2003. Modelos de la probabilidad de entrar en mora, del volumen de deuda en mora y del total de deuda bancaria, a partir de datos individuales de empresa.
- 0623 JUAN AYUSO AND JORGE MARTÍNEZ: Assessing banking competition: an application to the Spanish market for (quality-changing) deposits.
- 0624 IGNACIO HERNANDO AND MARÍA J. NIETO: Is the Internet delivery channel changing banks' performance? The case of Spanish banks.
- 0625 JUAN F. JIMENO, ESTHER MORAL AND LORENA SAIZ: Structural breaks in labor productivity growth: The United States vs. the European Union.
- 0626 CRISTINA BARCELÓ: A Q-model of labour demand.
- 0627 JOSEP M. VILARRUBIA: Neighborhood effects in economic growth.
- 0628 NUNO MARTINS AND ERNESTO VILLANUEVA: Does limited access to mortgage debt explain why young adults live with their parents?
- 0629 LUIS J. ÁLVAREZ AND IGNACIO HERNANDO: Competition and price adjustment in the euro area.
- 0630 FRANCISCO ALONSO, ROBERTO BLANCO AND GONZALO RUBIO: Option-implied preferences adjustments, density forecasts, and the equity risk premium.

1. Previously published Working Papers are listed in the Banco de España publications catalogue.

- 0631 JAVIER ANDRÉS, PABLO BURRIEL AND ÁNGEL ESTRADA: BEMOD: A dsge model for the Spanish economy and the rest of the Euro area.
- 0632 JAMES COSTAIN AND MARCEL JANSEN: Employment fluctuations with downward wage rigidity: The role of moral hazard.
- 0633 RUBÉN SEGURA-CAYUELA: Inefficient policies, inefficient institutions and trade.
- 0634 RICARDO GIMENO AND JUAN M. NAVE: Genetic algorithm estimation of interest rate term structure.
- 0635 JOSÉ MANUEL CAMPA, JOSÉ M. GONZÁLEZ-MÍNGUEZ AND MARÍA SEBASTIÁ-BARRIEL: Non-linear adjustment of import prices in the European Union.
- 0636 AITOR ERCE-DOMÍNGUEZ: Using standstills to manage sovereign debt crises.
- 0637 ANTON NAKOV: Optimal and simple monetary policy rules with zero floor on the nominal interest rate.
- 0638 JOSÉ MANUEL CAMPA AND ÁNGEL GAVILÁN: Current accounts in the euro area: An intertemporal approach.
- 0639 FRANCISCO ALONSO, SANTIAGO FORTE AND JOSÉ MANUEL MARQUÉS: Implied default barrier in credit default swap premia. (The Spanish original of this publication has the same number.)
- 0701 PRAVEEN KUJAL AND JUAN RUIZ: Cost effectiveness of R&D and strategic trade policy.
- 0702 MARÍA J. NIETO AND LARRY D. WALL: Preconditions for a successful implementation of supervisors' prompt corrective action: Is there a case for a banking standard in the EU?
- 0703 PHILIP VERMEULEN, DANIEL DIAS, MAARTEN DOSSCHE, ERWAN GAUTIER, IGNACIO HERNANDO, ROBERTO SABBATINI AND HARALD STAHL: Price setting in the euro area: Some stylised facts from individual producer price data.
- 0704 ROBERTO BLANCO AND FERNANDO RESTOY: Have real interest rates really fallen that much in Spain?
- 0705 OLYMPIA BOVER AND JUAN F. JIMENO: House prices and employment reallocation: International evidence.
- 0706 ENRIQUE ALBEROLA AND JOSÉ M.ª SERENA: Global financial integration, monetary policy and reserve accumulation. Assessing the limits in emerging economies.
- 0707 ÁNGEL LEÓN, JAVIER MENCÍA AND ENRIQUE SENTANA: Parametric properties of semi-nonparametric distributions, with applications to option valuation.
- 0708 DANIEL NAVIA AND ENRIQUE ALBEROLA: Equilibrium exchange rates in new EU members: external imbalances vs. real convergence.
- 0709 GABRIEL JIMÉNEZ AND JAVIER MENCÍA: Modelling the distribution of credit losses with observable and latent factors.