Modelling the distribution of credit losses with observable and latent factors [1]

Gabriel Jiménez
Bank of Spain
<gabriel.jimenez@bde.es>

Javier Mencía
Bank of Spain
<javier.mencia@bde.es>

Fist version: November 2006
This version: May 2007

[1]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 Max Bruche, Mark Flannery, Mark Flood, Albert Lee Chun, Ángel León, Miroslav Misina, Antonio Rubia, and Jesús Saurina, as well as seminar audiences at the University of Alicante, the Bank of Spain, CEMAF-ISCTE-Nova Conference on Credit Risk (Lisbon), the Third International Conference on Credit and Operational Risks (Montreal) and the ProBanker Symposium 2007 (Maastricht) for very helpful comments and suggestions. Of course, all remaining errors are entirely ours. Address for correspondence: Alcalá 48, E-28014 Madrid, Spain, tel: +34 91 338 5414, fax: +34 91 338 6102.

Abstract

This paper proposes a dynamic model to estimate the credit loss distribution of the aggregate portfolio of loans granted in a banking system. We consider a sectorial approach distinguishing between corporates and individuals. The evolution of their default frequencies and the size of the loans portfolio are expressed as functions of macroeconomic conditions as well as unobservable credit risk factors, which capture contagion effects between sectors. In addition, we model the distributions of the Exposures at Default and the Losses Given Default. We apply our framework to the Spanish banking system, where we find that sectorial default frequencies are not only affected by economic cycles but also by a persistent latent factor. Finally, we identify the 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 of the loans portfolios of different countries.[1]

These papers generally follow a top-down approach by analysing the banking sector as a whole. Most of them also 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 (see also Demchuk and Gibson, 2006). 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. Unfortunately, most of the empirical research has generally focused on either large corporates or publicly traded instruments, such as bonds or stock returns. For instance, Schuermann and Stiroh (2006) have found an important presence of "hidden risk factors" in U.S. banks stock returns, while Duffie, Eckner, Horel, and Saita (2006) have noticed that the effects of these factors on the correlation of defaults might be larger if they are persistent. However, much less is known about the presence of latent factors in the credit loss distribution of loans.

This paper proposes a credit risk model that allows for the presence of persistent latent factors. We express loans losses in terms of four stochastic components: default frequencies, the size of the loans portfolio, the exposures at default and the losses given

[1] To cite a few examples, Boss (2002) has developed a credit risk model for Austria, Virolainen (2004) has considered the case of Finland, Misina, Tessier, and Dey (2006) have analysed the Canadian loans portfolio, 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.

default. The importance of modelling the size of the loans portfolio has been traditionally neglected. However, it is necessary to take into account this variable if we want to study the total losses of a banking system, and not just those due to a fixed number of loans. For each of the economic sectors in which we arrange the loans, we assume that changes in the default frequencies and the total number of loans are a function of past observations of the dependent variables, a set of observable characteristics, some potentially persistent common latent factors and one idiosyncratic component. The effect of observable factors is to introduce correlation between different loans due to clearly identifiable shocks, such as a fall in GDP growth. In contrast, the latent components will generate contagion effects that are orthogonal to the observable events. 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 at default. 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 flexible distributions whose statistical properties can be exploited to reduce by a considerable amount the computational demands of 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 to describe 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. Since this threshold is very low, we can safely assume that we have data on virtually every loan granted in Spain. Hence, we use high quality loan data at a frequency at which it is not usually available. In this sense, it is worth remarking that we are able to obtain actual default rates from our database. In contrast, most of the literature usually relies on bankruptcy rates, which are imperfect proxies of defaults.[2] We consider 10 corporate sectors plus one group for mortgages and another one for consumption loans. We first estimate a simple model with changes in GDP growth and 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 an analogous effect on the growth of the number of loans. These factors remain significant when we reestimate our model with an augmented set of macroeconomic characteristics. We also determine which sectors are riskier, and compare our model with simpler versions that have been previously implemented. In this sense, we show that latent factors are crucial to capture the empirical correlations between sectorial default frequencies. In addition, we assess the out-of-sample stability of our model. 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. This result is consistent with the findings of 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, this paper introduces unobservable common shocks in a credit risk model of loans losses. 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 mortgages and consumption loans. Thirdly, our results show that 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

[2]See the discussion by Duffie, Eckner, Horel, and Saita (2006)

for future research are suggested in Section 5.

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. In this context, 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 $LGD_{k,t} \in (0,1)$ and $EAD_{i,k,t} > 0$ are, respectively, the loss given default and the exposure at default. We will denote the proportion of non-performing loans in sector k at time t as p_{kt}, 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

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

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

$$S_{kt} = \sum_{i=1}^{\lfloor p_{kt} n_{k,t} \rfloor} EAD_{i,k,t}. \tag{2}$$

where $\lfloor p_{kt} n_{k,t} \rfloor$ rounds $p_{kt} n_{kt}$ to the nearest integer. Without loss of generality, we have assumed that the first loans in the sum (1) are those that default. We have also supposed that the losses given default are homogeneous in each sector because this type of information is rarely available for loans at a more disaggregated level. 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_{kt} grows to infinity. However, for small n_{kt}, they will not necessarily coincide.

The main dynamic features of our model are introduced with a joint model for p_{kt} and n_{kt}. In order to work with variables with support on the whole real line, we transform the default frequencies by means of the probit functional form $y_{kt} = \Phi^{-1}(p_{kt})$, where $\Phi^{-1}(\cdot)$ is the inverse of the standard normal cumulative distribution function. Alternatively, a logit model could also be adopted. For every sector, we define the growth of the number of loans as $\Delta n_{kt} = \log(n_{kt}) - \log(n_{kt-1})$, while the changes in the transformed default frequencies are defined as $\Delta y_{kt} = y_{kt} - y_{kt-1}$.[3] We propose the following vector autoregression for these variables:

$$\begin{aligned}
\Delta n_{kt} &= \alpha_{1,k} + \sum_{j=1}^{q} \rho_{1,j} \Delta n_{kt-j} + \sum_{j=1}^{r} \boldsymbol{\gamma}'_{1,j} \mathbf{x}_{t-j} + \boldsymbol{\beta}_{1,k} f_{1,t} + u_{1,kt}, &(3)\\
\Delta y_{kt} &= \alpha_{2,k} + \sum_{j=1}^{q} \rho_{2,j} \Delta y_{kt-j} + \sum_{j=1}^{r} \boldsymbol{\gamma}'_{2,j} \mathbf{x}_{t-j} + \boldsymbol{\beta}_{2,k} f_{2,t} + u_{2,kt}. &(4)
\end{aligned}$$

In consequence, the evolution of Δn_{kt} and Δy_{kt} depends on their previous history, a set of m observable characteristics $\mathbf{x}_t$, two unobservable common factors, $f_{1,t}$ and $f_{2,t}$, and the idiosyncratic shocks $u_{1,kt} \sim N(0, \sigma_{1k}^2)$ and $u_{2,jt} \sim N(0, \sigma_{2k}^2)$, for $j, k = 1, \cdots, K$. These idiosyncratic terms are assumed to be *iid* jointly Gaussian and independent from the common shocks. In addition, we only allow for correlation between the two idiosyncratic terms from the same sector, i.e. $\text{cov}(u_{1,kt}, u_{2,jt}) = 0$ for $k \neq j$.

We consider the following vector autoregressive structure for the observable factors:

$$\mathbf{x}_t = \delta_0 + \sum_{j=1}^{s} \mathbf{A}_j \mathbf{x}_{t-j} + \mathbf{v}_t, \tag{5}$$

where $\mathbf{v}_t \sim N(\mathbf{0}, \boldsymbol{\Omega})$. To ensure the identification of the model, we assume that f_{1t} only affects (3), whereas f_{2t} can only influence default frequencies. However, we allow for correlation between these factors. In particular, 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:

$$\mathbf{f}_t = \mathbf{R}\mathbf{f}_{t-1} + \mathbf{w}_t. \tag{6}$$

where

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

and $\mathbf{w}_t$ is Gaussian with zero mean and

$$V(\mathbf{w}_t) = \begin{bmatrix} 1-\phi_1^2 & \rho\sqrt{(1-\phi_1^2)(1-\phi_2^2)} \\ \rho\sqrt{(1-\phi_1^2)(1-\phi_2^2)} & 1-\phi_2^2 \end{bmatrix}. \tag{7}$$

[3] We specify our model in first differences because the levels are usually nonstationary in this type of applications (see e.g. Boss, 2002, and our empirical application). However, it will be straightforward to rewrite our model in levels if necessary.

Hence, ϕ_i is the first order autocorrelation of $f_{i,t}$, for $i = 1, 2$, and ρ is the conditional correlation between $f_{1,t}$ and $f_{2,t}$. Since $\mathbf{f}_t$ is unobservable, we have to fix its scale to ensure the identification of the model. This is why we have parametrised (7) so that the latent factors have unit unconditional variances. In addition, we assume that $\text{cov}(\mathbf{v}_t, \mathbf{w}_t) = \mathbf{0}$, which implies that the latent factors are orthogonal to the observable characteristics. Hence, these unobservable components introduce a source of contagion between sectors that cannot be attributable to the observable shocks. Giesecke and Weber (2004) show that these effects may be caused by the interaction of firms with their business partners, while Kiyotaki and Moore (1997) argue that the relationship between credit limits and asset prices can create a transmission mechanism by which shocks will persist and spill over to other sectors. Nevertheless, our approach is focused on empirically assessing the existence of latent factors, without precluding or favouring any of these explanations.

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.[4] 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 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:

$$\mu_{kt} = \mu_{kt-1} \exp\left[\eta_k + \boldsymbol{\varphi}_k' \mathbf{v}_{t-1} - \frac{1}{2}\boldsymbol{\varphi}_k' \boldsymbol{\Omega} \boldsymbol{\varphi}_k\right] \tag{8}$$

where η_k captures a time trend, $\mathbf{v}_{t-1}$ is the lagged vector of innovations in equation (5) 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 static 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 by setting this parameter to zero. In consequence, it is important to include the term $\boldsymbol{\varphi}_k' \boldsymbol{\Omega} \boldsymbol{\varphi}_k/2$ in (8)

[4] We 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 nor the Log-normal because they are not closed under aggregation.

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 (8) 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 (3) and (4), we need to use the Kalman filter to deal with the unobserved factors. 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 $t-1$:

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

where $\Delta\mathbf{n}_s=(\Delta n_{1,s},\cdots,\Delta n_{K,s})'$ and $\Delta\mathbf{y}_s=(\Delta y_{1,s},\cdots,\Delta y_{K,s})'$. 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:

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

Finally, the estimation procedure consists basically in treating (3) and (4) 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 $\mathbf{f}_t$(see e.g. Hamilton, 1994, for a formal discussion).

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:

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

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:

$$f_{IG}\left(EAD_{i,k,t}=x;\mu_k,\lambda_k\right) = \left(\frac{\lambda_k}{2\pi x^3}\right)^{1/2}\exp\left[-\frac{\lambda_k}{2\mu_k^2 x}\left(x-\mu_k\right)^2\right] \quad (9)$$

$$f_{Gamma}(EAD_{i,k,t}=x;\nu_k,\tau_k) = \frac{(x/\tau_k)^{\nu_k/2-1}}{2^{\nu_k/2}\Gamma(\nu_k/2)\tau_k}\exp\left(\frac{-x}{2\tau_k}\right) \quad (10)$$

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:

$$f(S_{kt}|\, I_{t-s}) = \sum_{i=0}^{\infty} g(S_{kt}|\, p_{kt}n_{kt}=i, I_{t-s})\Pr(\lfloor p_{kt}n_{kt}\rfloor = i|\, I_{t-s}) \quad (11)$$

where $g(S_{kt}|\,\lfloor p_{kt}n_{kt}\rfloor = i, I_{t-s})$ is the conditional density function of S_{kt} given i defaults occurring at t, while I_{t-s} denotes the information known at $t-s$. 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 (11) 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 (8) 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. 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 (commercial banks, savings banks, credit 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 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 defaulting 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 one group of mortgages and another one for consumption loans. For corporate loans, we define the following economic sectors: (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 group 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. This definition is 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, we have also obtained the individual exposures of the non-performing loans for every quarter.

Figure 1 (a) shows the historical evolution of default frequencies. For the sake of comparability, we represent in Figures 1 (c) and 1 (d) the quarterly series of the Spanish GDP annual growth and the 3-month real interest rates, respectively.[5] 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 check in Figure 1 (c). In addition, interest rates also increased from 4% in 1988 to values above 8% in the first half of the 1990's. Loans to construction companies and hotels were more affected than the rest in this recession, with default frequencies peaking at 4%. In contrast, the default frequencies of mortgages reached 1.5% at the worst moment of the recession. 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.

Figure 1 (b) shows the quarterly series of the total number of loans in each sector. The loan market size has steadily grown in all sectors during the sample period under

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

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: the quarterly change in real GDP growth and the variation of three-month real interest rates.[6] 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 (3) and (4).[7] We will introduce the lags $2, 3$ and 4 of our two macroeconomic variables. To save parameters, we do not include the first lag, because we obtain insignificant estimates for this lag once the subsequent 3 lags are considered. The intuition of this result relies in the definition of default: not meeting the scheduled payments for at least one quarter. In consequence, the default frequencies of period t are related to borrowers who originally became insolvent in period $t-2$. In this sense, it seems reasonable that we do not obtain significant sensitivities with respect to the first lag of the observable factors. As for the autoregressive structure, we consider the effect of the first lag of the dependent variables, as well as a seasonal effect by means of the fourth lag. Finally, we consider three dummies whose values are 1 in 1988.Q1, 1988.Q4 and 1996.Q2, respectively, and zero otherwise.[8] These dummies are intended to capture the effects of historical exogenous changes in the

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

[7] Prior 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 rates and the total number of loans in first differences to ensure their stationarity.

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

database (see Delgado and Saurina, 2004, for a formal justification).

The estimates of the default frequency model are shown 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 an 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 most sectors, with the first two lags being highly significant in many of them. Nevertheless, mining and utilities react less to the cycle, while some sectors seem to respond more slowly to aggregate shocks. For instance, we only observe a significant effect on R&D and mortgages two 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 for interest rates, higher values generally tend to increase default frequencies, with significant coefficients for agriculture, 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 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 -0.473 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.[9] Furthermore, the time series structure of these factors also deserves some attention. Table 3 shows the autoregressive structure of the observable and unobservable factors. As we can observe, f_{2t} has a significant first order autocorrelation of 0.198. 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 -0.193. 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 first column of Table 2 (b) shows the positive and highly significant intercept terms that we obtain for the market size growth, which are consistent with the expansion of the loan market already documented in Figure 1 (b). These intercepts are negative but statistically insignificant for default frequencies, as Table 1 (b) shows. The second column of Table 1 (b) shows that the marginal effect of lagged default frequencies from the previous quarter is negative, whereas the seasonal effect (third column) is positive when it is significant. In contrast, both terms are generally positive in the market size equation. Finally, we can observe in the last columns of both tables that the correlation between the idiosyncratic terms from the same sector are 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,

[9] Notice that the latent factors are independent from the observable factors by construction.

the absence of latent factors causes an increase in the absolute 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 static specifications of the IG and the Gamma distributions by maximum likelihood. Since we assume that these parameters remain constant over time, we focus on the current situation. Hence, we only use the exposures of the loans that defaulted in 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 for each sector the empirical fit at the right tail of the IG and the Gamma with a Kernel estimate of the empirical density. Except for mortgages, the IG 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.[10] 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, which is close to the values reported by Altman, Resti, and Sironi (2004).

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 (3) and (4), 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 (5) and (6), respectively. In these simulations, we set to zero the unconditional means of the changes of default frequencies, since a positive (negative) intercept would imply that default frequencies would tend to 1(0) in the long run. Thus, our restriction rules out

[10] Fifth Quantitative Impact Study of the Basel Committee on Banking Supervision.

these extreme cases. Finally, given the total number of defaults, we can generate random replications of (2) 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.[11] 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. If anything, it seems that the model without latent factors yields higher sectorial losses. However, as the last row of the table shows, total unexpected 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 unexpected loss 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 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.

[11] These 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.

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 three-month and six-year 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 series 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 ability of the three different specifications of the VAR model to fit the empirical correlations between default frequencies.[12]To do so, we compute the fitted residuals of the default frequencies in (4) for the three cases. That is, we compute $\boldsymbol{\varepsilon}_{kt}(\hat{\boldsymbol{\theta}}_T) = \Delta y_{kt} - E(\Delta y_{kt-1}|I_{t-1};\hat{\boldsymbol{\theta}}_T)$ for $k = 1, \cdots, K$, where the expectation is based on the information known at time $t-1$ and the maximum likelihood estimates of the parameters, denoted by the vector $\hat{\boldsymbol{\theta}}_T$. The specification that does not include latent factors assumes that these fitted residuals are uncorrelated because in this case intersectorial correlations are only captured by the observable common characteristics, which are part of the information set I_{t-1}. In contrast, the model with latent factors introduces a factorial structure for these correlations: $\text{cov}(\boldsymbol{\varepsilon}_{it}(\hat{\boldsymbol{\theta}}_T),\boldsymbol{\varepsilon}_{jt}(\hat{\boldsymbol{\theta}}_T))=\beta_{2,i}\beta_{2,j}$. We test in Table 10

[12]For 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.

whether the empirical correlations of the fitted residuals are equal to those hypothesised 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 residual 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 fitted residuals of (4), but in this case we will also consider those of (3). 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 the 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 the out-of-sample stability of our model.

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 (8) for μ_k in (9). Although we have also estimated an analogous model with the Gamma distribution, we do not report the results for this model due to its poorer 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 for 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.[13]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 (5)). 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

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

and interest rates are the only observable characteristics. We report in Table 13 the 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.[14] 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 5-7% under the interest rate shock. This result is a direct consequence of the much higher explanatory power of GDP in the VAR models of 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 for short horizons, we now obtain fatter tails at the five-year horizon, where VaR 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 correlation between default frequencies and exposures at default, since both of them are influenced by the same 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.

[14] Hence, we directly simulate from (8), by imposing $\eta_k = 0$, because we do not expect that the downward trend documented in Table 12 will persist in the future.

5 Conclusions

We develop a flexible model to estimate the credit loss distribution of the loans portfolio in a national banking system. We classify the loans in sectors, and model default frequencies, individual 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. In addition, we model the loss given default with a Beta distribution. Finally, we fit the distributions of the exposures at default with the Gamma and the Inverse Gaussian distributions, where we 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 an additional group for each of these categories. 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 of its 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, there is such a large number of loans in these groups 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 explanatory 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, stressed GDP 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 one for the interbank market, such as those 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

Altman, E., A. Resti, and A. Sironi (2004). Default recovery rates in credit risk modelling: a review of the literature and empirical evidence. *Economic Notes 33*, 183–208.

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.

Demchuk, A. and R. Gibson (2006). Stock market performance and the term structure of credit spreads. *Journal of Financial and Quantitative Analysis 41*, 863–887.

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.

Giesecke, K. and S. Weber (2004). Cyclical correlations, credit contagion and portfolio losses. *Journal of Banking and Finance 28*, 3009–3036.

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). Losscalc$^{\text{TM}}$: 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.

Kiyotaki, N. and J. Moore (1997). Credit cycles. *Journal of Political Economy 105*, 211–248.

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.

Misina, M., D. Tessier, and S. Dey (2006). Stress testing the corporate loans portfolio of the Canadian banking sector. Bank of Canada Working Paper No. 2006-47.

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.

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	-1.133**	-1.129**	-0.432	-0.281	1.453**	-0.336	3.335**
Mining	-1.162	-1.248	0.122	0.291	0.316	-1.094	5.791**
Manufacture	-1.515**	-1.740**	-0.862*	0.383	0.668	-0.469	4.447**
Utilities	-0.097	0.087	-0.494	0.073	0.647	-0.847	5.129**
Construction	-0.958**	-0.988*	-0.875**	0.702	0.093	0.259	3.411**
Commerce	-1.267**	-1.213**	-0.606	-0.198	0.712	-0.119	4.038**
Hotels	-1.304**	-0.826	-0.141	-0.101	1.849**	-0.348	4.038**
Communications	-0.953**	-1.053**	-0.857*	0.138	1.125**	-0.435	3.673**
R&D	-0.403	-1.421**	-1.486**	0.156	-0.187	-0.096	3.697**
Other Corp.	-0.331	-0.888*	-0.256	0.644	0.881*	-0.242	3.191**
Cons. loans	-0.840**	-1.026**	-0.526	0.020	0.604	0.219	3.261**
Mortgages	-0.805	-1.608**	-1.329**	0.364	0.022	0.029	1.668**

(b) Dynamics

	α	$\Delta y_{k,t-1}$	$\Delta y_{k,t-4}$	$\mathrm{corr}(u_{1k,t}, u_{2k,t})$
Agriculture	-0.605	-0.362**	0.215**	0.429**
Mining	-1.080	-0.327**	-0.074	0.017
Manufacture	-0.554	-0.329**	-0.013	0.084
Utilities	-1.122	-0.377**	-0.135	0.058
Construction	-0.368	-0.079	0.176**	-0.354**
Commerce	-0.459	-0.237**	0.038	0.052
Hotels	-0.395	-0.340**	-0.003	0.145
Communications	-0.420	-0.317**	0.120*	0.319**
R&D	-0.494	-0.160**	0.070	-0.116
Other Corp.	-0.625	-0.219**	0.141*	-0.322**
Cons. loans	-0.594	-0.277**	-0.030	-0.304**
Mortgages	-0.520	0.049	0.058	-0.162

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 $\Delta y_{k,t-1}$ and $\Delta y_{k,t-4}$ denote the effect of lagged observations of the dependent variables. "$\mathrm{corr}(u_{1k,t}, u_{2k,t})$" 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	0.250	0.171	0.189	-0.200	0.059	-0.078	1.258**
Mining	0.197	-0.249	0.038	-0.056	-0.064	0.226	1.375**
Manufacture	0.383**	0.062	0.120	-0.072	-0.074	0.090	1.600**
Utilities	0.246	-0.110	-0.097	-0.863**	0.562	-0.499	1.211**
Construction	0.321*	0.086	0.137	-0.240	0.068	-0.126	1.470**
Commerce	0.463**	0.127	0.072	0.086	-0.201	0.158	1.793**
Hotels	0.210	-0.070	0.063	0.023	0.027	-0.242	1.991**
Communications	0.126	0.537	0.424	0.621	-0.113	0.141	2.069**
R&D	0.623**	0.225	-0.059	-0.055	-0.096	-0.201	1.591**
Other Corp.	-0.902**	-0.805*	0.205	0.359	-0.261	0.544	1.019**
Cons. loans	0.029	0.058	0.522*	0.514	0.311	0.042	0.781**
Mortgages	0.155	0.038	0.116	0.756**	-0.516	-0.118	0.589*

(b) Dynamics

	α	$\Delta n_{k,t-1}$	$\Delta n_{k,t-4}$	$corr(u_{1k,t}, u_{2k,t})$
Agriculture	1.309**	0.308**	0.130	0.429**
Mining	0.917**	0.293**	0.081	0.017
Manufacture	0.659**	0.374**	0.186**	0.084
Utilities	1.199**	0.194*	-0.191*	0.058
Construction	1.002**	0.575**	0.249**	-0.354**
Commerce	0.846**	0.447**	0.289**	0.052
Hotels	1.303**	0.286**	0.488**	0.145
Communications	0.908**	0.514**	0.252**	0.319**
R&D	1.579**	0.314**	0.416**	-0.116
Other Corp.	1.649**	0.477**	0.094	-0.322**
Cons. loans	2.465**	0.094*	0.033	-0.304**
Mortgages	2.681**	-0.023	0.235**	-0.162

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. "$corr(u_{1k,t}, u_{2k,t})$" refers to the correlation between the two idiosyncratic residuals that affect the same sector.

Table 3
Dynamics of the factors

	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	-0.193*	0	0	0	1	
f_{2t}	0	0.198*	0	0	0	-0.473**	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	-1.058**	-1.105**	-0.326	-0.096	1.349**	-0.067	0.000
Mining	-0.984	-1.171	0.205	0.685	0.251	-0.949	0.000
Manufacture	-1.509**	-1.613**	-0.686	0.646	0.681	-0.430	0.000
Utilities	-0.076	0.071	-0.394	0.451	0.390	-0.491	0.000
Construction	-0.783*	-0.712	-0.770*	1.190**	-0.308	0.593	0.000
Commerce	-1.203**	-1.029**	-0.431	0.069	0.702	-0.073	0.000
Hotels	-1.273**	-0.688	-0.017	0.155	1.714**	-0.156	0.000
Communications	-0.745*	-0.800	-0.652	0.567	0.999*	-0.218	0.000
R&D	-0.207	-1.364**	-1.454**	0.412	-0.428	0.178	0.000
Other Corp.	-0.290	-0.840*	-0.192	0.736	0.766	-0.013	0.000
Cons. loans	-0.650*	-0.893**	-0.418	0.308	0.472	0.452	0.000
Mortgages	-0.825	-1.654**	-1.440**	0.530	-0.224	0.103	0.000

(b) Dynamics

	α	$\Delta y_{k,t-1}$	$\Delta y_{k,t-4}$	$\text{corr}(u_{1k,t}, u_{2k,t})$
Agriculture	-0.311	-0.329**	0.467**	0.061
Mining	-0.985	-0.338**	-0.002	-0.360**
Manufacture	-0.375	-0.237**	0.146	-0.458**
Utilities	-1.010	-0.357**	-0.053	-0.103
Construction	-0.156	0.047	0.393**	-0.256**
Commerce	-0.278	-0.131	0.253**	-0.431**
Hotels	-0.287	-0.301**	0.118	-0.227**
Communications	-0.254	-0.244**	0.382**	0.083
R&D	-0.352	-0.125	0.264**	-0.103
Other Corp.	-0.450	-0.203*	0.306**	-0.242**
Cons. loans	-0.405	-0.239**	0.174	-0.025
Mortgages	-0.553	0.034	0.105	-0.141

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 $\Delta y_{k,t-1}$ and $\Delta y_{k,t-4}$ denote the effect of lagged observations of the dependent variables. "$\text{corr}(u_{1k,t}, u_{2k,t})$" 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	0.174	0.146	-0.223	-0.008	-0.114	0.000
Mining	0.198	-0.212	-0.044	-0.086	-0.111	0.201	0.000
Manufacture	0.455**	0.166	0.095	-0.155	-0.111	-0.016	0.000
Utilities	0.242	-0.085	-0.112	-0.832**	0.486	-0.471	0.000
Construction	0.392**	0.124	0.122	-0.299	0.017	-0.243	0.000
Commerce	0.514**	0.208	0.022	0.011	-0.232	0.019	0.000
Hotels	0.211	-0.088	-0.023	-0.018	0.004	-0.347	0.000
Communications	0.220	0.712*	0.465	0.787*	-0.109	0.050	0.000
R&D	0.794**	0.460*	-0.052	-0.152	-0.045	-0.415	0.000
Other Corp.	-0.913**	-0.843*	0.152	0.328	-0.265	0.538	0.000
Cons. loans	0.012	0.021	0.531*	0.505	0.312	-0.023	0.000
Mortgages	0.162	0.041	0.121	0.730**	-0.463	-0.153	0.000

(b) Dynamics

	α	$\Delta n_{k,t-1}$	$\Delta n_{k,t-4}$	$\text{corr}(u_{1k,t}, u_{2k,t})$
Agriculture	1.197**	0.208*	0.293**	0.061
Mining	1.103**	0.063	0.173*	-0.360**
Manufacture	0.622**	0.159	0.413**	-0.458**
Utilities	1.332**	0.112	-0.191	-0.103
Construction	0.791**	0.461**	0.522**	-0.256**
Commerce	0.688**	0.261**	0.547**	-0.431**
Hotels	1.010**	0.171*	0.643**	-0.227**
Communications	0.813*	0.446**	0.410**	0.083
R&D	1.085**	0.115	0.685**	-0.103
Other Corp.	1.782**	0.443**	0.088	-0.242**
Cons. loans	2.383**	0.071	0.084	-0.025
Mortgages	2.648**	-0.033	0.251**	-0.141

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. "$\text{corr}(u_{1k,t}, u_{2k,t})$" 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	38.32	128.39	244.97	124.45	534.22	1281.41	86.13	405.83	1036.43
Mining	5.65	19.41	37.26	26.04	116.48	284.98	20.39	97.07	247.72
Manufacture	285.39	929.76	1697.45	948.88	3863.20	8379.99	663.49	2933.44	6682.54
Utilities	4.09	14.13	27.09	26.43	91.85	217.93	22.34	77.71	190.84
Construction	318.00	1153.53	2344.96	1051.21	5113.96	13162.55	733.21	3960.43	10817.59
Commerce	189.87	638.15	1203.09	609.49	2522.91	5661.09	419.62	1884.76	4458.00
Hotels	32.10	114.78	232.83	110.14	497.27	1232.32	78.03	382.49	999.49
Communications	36.88	126.52	246.08	122.23	560.86	1399.85	85.35	434.33	1153.78
R&D	68.39	252.79	532.46	235.98	1167.62	3140.44	167.59	914.82	2607.99
Other Corp.	33.97	121.02	245.51	111.12	536.74	1411.08	77.14	415.72	1165.57
Cons. loans	408.74	1412.72	2719.85	1692.62	6870.98	15186.85	1283.88	5458.26	12466.99
Mortgages	257.88	975.90	2116.68	1568.66	8599.30	25027.81	1310.78	7623.41	22911.12
Total	1679.29	5887.11	11648.23	3889.04	17443.22	43715.87	2209.75	11556.11	32067.64

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	39.11	131.01	254.02	124.82	566.39	1505.67	85.71	435.38	1251.65
Mining	5.81	19.59	37.21	25.68	110.27	263.80	19.87	90.68	226.59
Manufacture	293.13	952.26	1742.00	948.98	3944.37	8769.62	655.85	2992.11	7027.62
Utilities	4.16	14.24	27.26	26.55	92.13	215.66	22.39	77.89	188.40
Construction	335.68	1276.94	2720.49	1092.19	6099.12	18175.05	756.50	4822.18	15454.56
Commerce	194.19	649.47	1226.39	610.49	2585.41	6025.71	416.31	1935.93	4799.32
Hotels	32.65	116.01	233.56	110.08	500.78	1253.15	77.43	384.77	1019.59
Communications	36.49	123.76	246.24	120.14	604.35	1750.34	83.65	480.59	1504.09
R&D	69.25	255.10	543.98	234.30	1219.38	3532.81	165.05	964.28	2988.83
Other Corp.	34.49	122.97	251.52	111.14	547.53	1507.93	76.66	424.56	1256.41
Cons. loans	423.79	1487.22	2890.92	1719.56	7269.91	16734.72	1295.77	5782.69	13843.80
Mortgages	259.55	983.23	2144.15	1555.19	8743.74	25720.14	1295.64	7760.51	23575.99
Total	1728.28	6131.82	12317.74	3661.07	16058.23	41024.64	1932.79	9926.42	28706.91

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	-0.927**	-1.098**	-0.473	0.597	0.316	0.421	0.672	-1.025	0.731	0.038	0.035	-0.036	3.320**
Mining	-0.882	-0.843	0.647	0.835	-0.969	-1.407	0.935	-2.019	-0.436	0.004	-0.551	-0.586*	5.121**
Manufacture	-1.353**	-1.458**	-0.593	0.652	0.008	-0.931	0.267	-1.169	-0.932	-0.125	-0.208*	-0.185*	4.029**
Utilities	-0.408	0.406	-0.712	-1.536	2.411	-3.211**	-1.191	1.235	-2.566**	-0.142	0.191	-0.313	4.918**
Construction	-0.794*	-0.533	-0.852*	0.760	0.351	-0.345	0.262	0.159	-0.778	-0.192*	-0.003	-0.075	3.160**
Commerce	-1.161**	-1.199**	-0.904**	0.315	-0.032	-0.155	0.656	-0.854	-0.165	-0.089	0.445*	0.073	3.856**
Hotels	-1.208**	-0.506	-0.331	0.233	1.824*	-0.063	0.054	0.047	0.374	-0.694	0.247	0.393	4.122**
Communications	-0.824*	-1.132**	-1.130**	1.183	-0.156	0.168	1.049	-1.114	0.569	-0.023	0.335	0.275	3.665**
R&D	-0.460	-1.289**	-1.329**	-0.818	0.685	-0.666	-1.248	0.630	-0.791	-	-	-	3.632**
Other Corp.	-0.317	-0.877*	-0.181	0.029	1.134	0.140	-1.277*	0.120	0.265	-	-	-	3.270**
Cons. loans	-0.857**	-1.001**	-0.573	0.172	0.756	-0.292	0.472	0.180	-0.572	0.021	0.216	-0.580	3.193**
Mortgages	-0.882	-1.753**	-1.506**	1.781*	0.094	-0.694	2.734**	0.143	-0.509	-0.911	-0.347	0.166	1.860**

(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	0.187	0.175	0.199	-0.194	0.390	-0.336	0.191	0.369	-0.197	0.028	0.023	0.017	1.246**
Mining	0.118	-0.303	0.023	-0.041	0.305	0.000	0.238	0.422	-0.028	0.072	-0.051	-0.036	1.302**
Manufacture	0.284*	0.032	0.106	-0.108	0.396	-0.299	0.246	0.510**	-0.231	0.055*	-0.038	-0.013	1.516**
Utilities	0.279	-0.103	-0.211	0.159	0.133	-0.862*	1.588**	-0.353	-0.427	0.076	0.034	-0.011	1.095**
Construction	0.214	0.029	0.111	-0.253	0.386	-0.483*	0.187	0.352	-0.299	0.012	0.000	0.070**	1.364**
Commerce	0.351**	0.174	0.177	-0.073	0.455	-0.346	0.056	0.714**	-0.441**	0.003	-0.162**	0.052	1.711**
Hotels	0.063	0.098	-0.081	-0.293	1.134**	-1.399**	0.226	1.073**	-0.992**	-0.285	0.212	0.063	1.774**
Communications	0.024	0.595	0.298	1.112*	-0.434	-0.311	0.918	-0.307	-0.574	-0.156	0.125	0.523*	1.950**
R&D	0.585**	0.238	-0.085	0.062	0.359	-0.426	0.345	0.632*	0.018	-	-	-	1.494**
Other Corp.	-0.840**	-0.854*	0.078	1.334**	-0.975	0.626	1.407**	-0.525	0.143	-	-	-	1.056**
Cons. loans	-0.010	0.067	0.503*	0.551	0.527	-0.191	0.221	0.257	-0.186	0.010	0.009	-0.041	0.720**
Mortgages	0.177	0.024	0.064	0.851*	-1.343**	0.260	-0.251	-0.797*	0.196	-0.220	0.078	0.208	0.779**

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 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	-0.291**	0.000
f_{2t}	0.000	0.136	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	-0.468**	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**	0.17	0.06	0.07	**0.01**	0.11	0.28	0.11	**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	0.30										
Manufacture	3	0.67	**0.03**									
Utilities	4	0.85	0.74	0.89								
Construction	5	0.76	0.24	0.57	0.14							
Commerce	6	0.67	0.95	0.69	0.59	0.36						
Hotels	7	0.43	0.27	0.50	0.38	0.72	0.99					
Communications	8	0.67	0.52	0.88	0.93	0.44	0.99	0.72				
R&D	9	0.44	0.15	0.15	0.09	0.40	0.35	0.94	0.51			
Other Corp.	10	0.76	0.71	0.57	0.13	0.27	0.41	0.77	0.35	**0.00**		
Individuals	11	0.39	0.34	0.20	0.92	0.52	0.28	0.64	0.25	0.24	0.78	
Mortgages	12	0.73	0.72	0.60	0.51	0.43	0.51	0.58	0.62	0.39	0.40	0.18

(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	0.33										
Manufacture	3	0.88	0.06									
Utilities	4	0.62	0.85	0.94								
Construction	5	0.75	0.16	0.44	0.29							
Commerce	6	0.91	0.94	0.60	0.98	0.71						
Hotels	7	0.73	0.41	0.55	0.65	0.83	0.90					
Communications	8	0.74	0.57	0.87	0.87	0.53	0.94	0.82				
R&D	9	0.39	0.40	0.22	0.34	0.59	0.37	0.84	0.73			
Other Corp.	10	0.65	0.93	0.69	0.55	0.53	0.44	0.53	0.46	0.10		
Individuals	11	0.26	0.41	0.22	0.53	0.36	0.32	0.63	0.33	0.45	0.92	
Mortgages	12	0.67	0.68	0.80	0.19	0.34	0.43	0.63	0.41	0.11	0.15	0.24

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 fitted residuals of the model of default frequencies and number of loans

Factors	Kolmogorov test	P-value
GDP, INT	0.103	0.004
GDP, INT, f_t	0.051	0.446
GDP, INT, SPR, SEC, f_t	0.046	0.573

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	7	12	13	2	5	5	2	5	4
Mining	8	10	10	7	10	10	2	-2	-2	2	-1	-2
Manufacture	10	17	18	10	15	17	3	3	3	3	4	4
Utilities	0	2	2	0	2	2	1	-2	-3	1	-1	-2
Construction	6	16	17	6	14	16	4	7	7	4	6	7
Commerce	7	13	14	8	13	13	1	3	3	1	3	2
Hotels	7	10	10	7	10	10	4	7	7	5	6	6
Communications	6	10	11	6	10	10	5	10	11	5	9	12
R&D	4	14	16	4	14	13	0	-2	-3	0	-3	-3
Other Corp.	7	16	18	8	16	18	6	13	15	7	13	15
Individuals	6	11	11	6	10	10	3	8	8	3	8	8
Mortgages	10	29	32	9	28	30	4	5	5	3	4	5
Total	7	16	18	7	18	21	3	5	6	3	5	5

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.

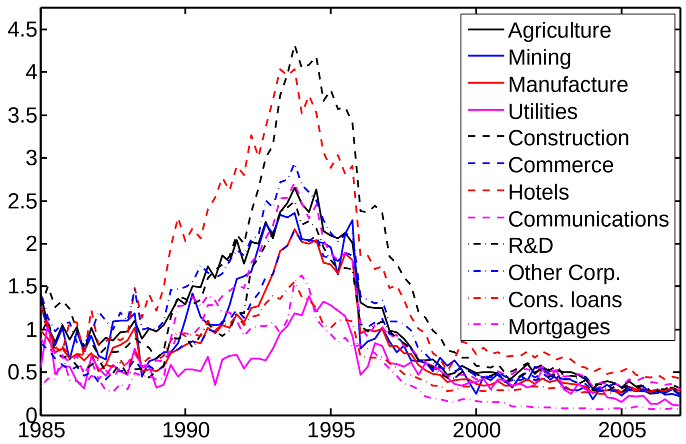
Table 14
Comparison of credit loss distributions

	(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	1679	1671	1486
3 years	5887	5769	5288
5 years	11648	11335	10647
VaR (99.9%)			
1 year	3889	3821	3501
3 years	17443	16693	17811
5 years	43716	40708	50076
Change due to -3 s.d. GDP shock (%)			
Expected loss			
1 year	7	6	20
3 years	16	16	32
5 years	18	18	35
VaR (99.9%)			
1 year	7	7	17
3 years	18	17	33
5 years	21	20	37
Change due to +3 s.d. Interest rate shock (%)			
Expected loss			
1 year	3	6	10
3 years	5	6	14
5 years	6	6	15
VaR (99.9%)			
1 year	3	7	10
3 years	5	7	14
5 years	5	7	15

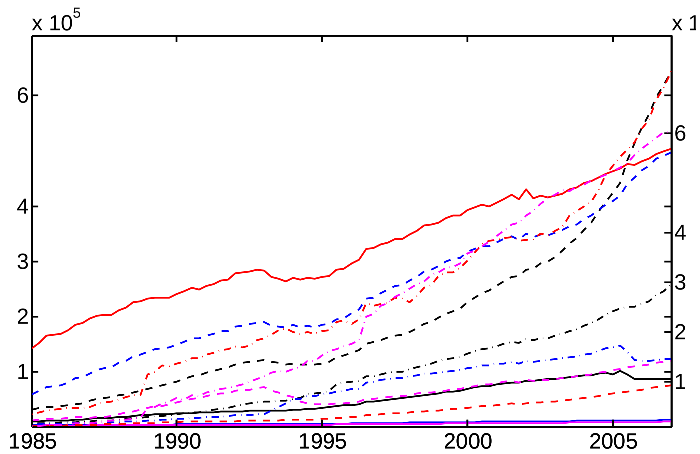
Notes: results in millions of euros. "Spread" denotes the difference between six-year and three-month interest rates. "GVA's" denotes gross value added factors, namely: agriculture, industry, energy, construction and market services. Statistics obtained from 1 million simulations of the credit risk model.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

Figure 1:

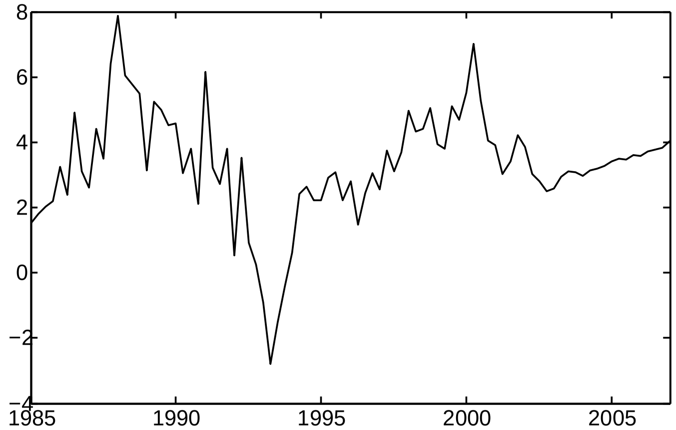
(a) Historical default frequencies in the Spanish Economy (%)



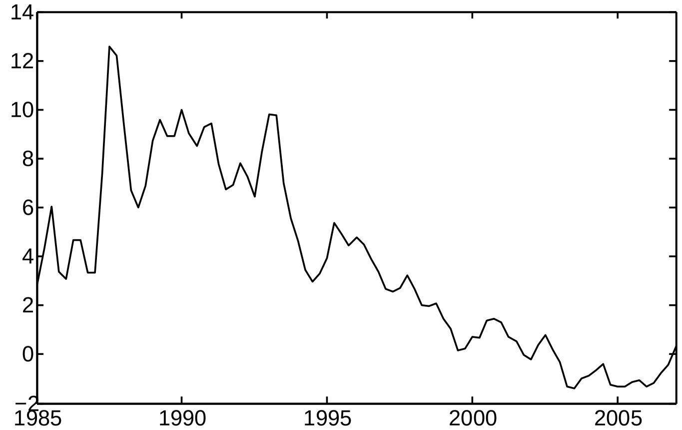
(b) Historical evolution of the total number of loans in the Spanish loan market



(c) Annual Spanish GDP growth (%)

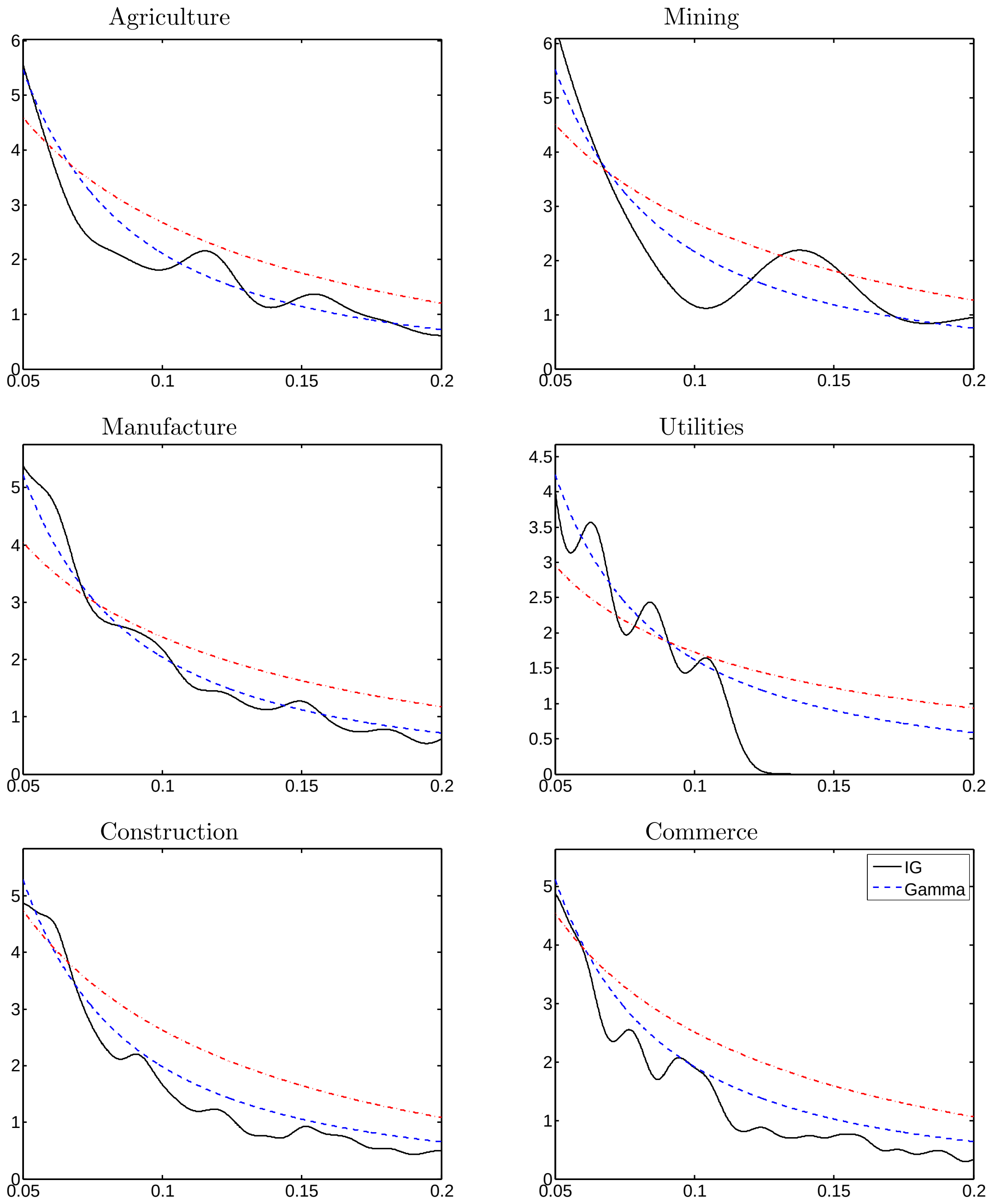


(d) Three-month real interest rates in Spain (%)



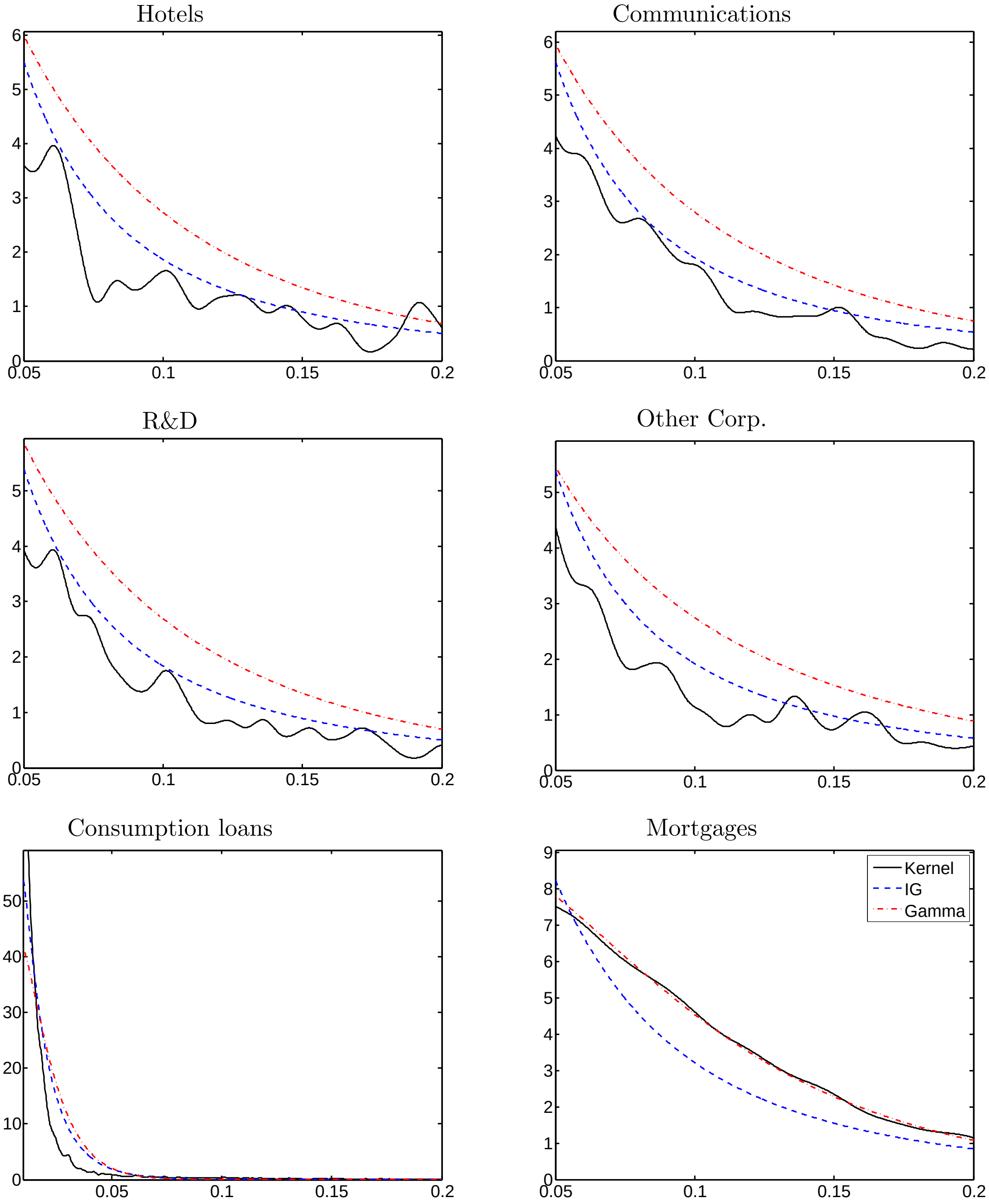
Notes: (a) and (b) share the same legends. The right scale on the y axis of figure (b) correspond to consumption loans and mortgages, whereas the left axis corresponds to the remaining cases.

Figure 2:
Kernel estimate and fitted densities of the right tail of the distribution of exposures at default



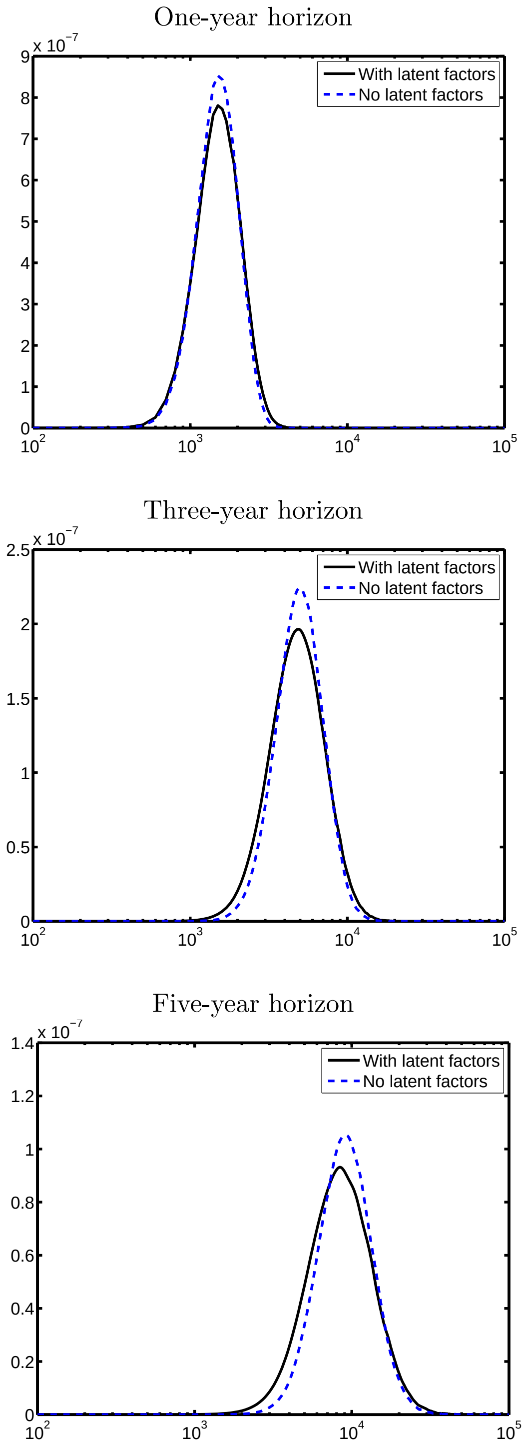
Notes: the x-axis is expressed in millions of euros. Both the kernel and the fitted densities are based on exposure data from 2001 to 2006.

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.