Estimating the distribution of total default losses on the Spanish financial system

Rubén García-Céspedes and Manuel Moreno*

September 29, 2014

Abstract

This paper quantifies the credit risk loss distribution of the Spanish financial system by introducing a general Monte Carlo importance sampling (IS) approach. We start obtaining all the required information for the standard credit risk model. Then we quantify the loss distribution under the standard IS method and allocate the total risk over the different institutions in the Spanish financial system. We extend the current IS framework to deal with more general assumptions like random recoveries and market valuation. We also study the variability of the risk measures over the business cycle and the possible variability due to the model parameters uncertainty. Our results show that this approach can be very useful for banking supervisors from a macroprudential point of view and that the risk allocation can vary considerably depending on the valuation model under analysis.

Keywords: Monte Carlo, importance sampling, credit risk, risk allocation, *VaR*, expected shortfall.

JEL classification: C15, C63, G21

*The authors are very grateful to Paul Glasserman for his many helpful suggestions. We also thank the comments received from Alberto Bueno-Guerrero and attendants at the XXI AEFIN Finance Forum at IE Business School. Rubén García-Céspedes is from BBVA, P^o Castellana, 81, 5^o, 28046 Madrid, Spain. E-mail: rubgac@yahoo.es. Phone: (+34) 687 69 8008. Manuel Moreno is from University of Castilla La-Mancha, Department of Economic Analysis and Finance, 45071 Toledo, Spain. E-mail: manuel.moreno@uclm.es. Phone: (+34) 925 26 8800 (Ext. 5133). Manuel Moreno gratefully acknowledges financial support by the grant ECO2012-34268. The usual caveat applies.

1 Introduction

This paper quantifies the credit risk loss distribution of the Spanish financial system under a general Monte Carlo importance sampling (IS) model. One of the main activities in financial institutions consists on financing investors and paying depositors. Under the Basel regulation the financial institutions are required to have a minimum level of own resources so that they will not go bankruptcy in the case that investors do not pay back their loans.

Micro-prudential financial regulation focuses on a one by one supervision of the financial institutions in order to ensure a maximum default probability of each institution, however a macro-prudential financial regulation focuses on the whole loss distribution of the financial system. In the past regulators did just a micro-prudential supervision (see Basel (2006)) however they have recently switched to a macro-prudential supervision (see Basel (2011)) that tries to capture the interconnectedness between the financial institutions, their size and the magnitude of the possible negative effects in the economy. Over the current economic crisis many financial institutions had to be rescued by the governments due to their size and potential negative effects in the economy, among others we have Fannie Mae, Freddie Mac, AIG, Northern Rock, RBS, Lloyds, Nordea, Dexia, ING, Fortis, IKB, Commerzbank, Hypo Real Estate or Bankia, CAM, CatalunyaCaixa, Novacaixagalicia (NCG), and Unnim in Spain and some have merged or been absorbed by others financial institutions. Therefore knowing the loss distribution of a whole financial system and being able to correctly allocate the risk of each institution is crucial for a good banking supervision and the financial system stability.

Over the previous years the measurement and allocation of the systemic risk of the financial system has been a very relevant topic. Many papers have analyzed this issue with a certain focus on the US financial system. Giesecke and Kim (2011) measured the systemic risk of this system by proposing a dynamic hazard model aiming to predict the number of future defaults conditional on a set of explanatory variables.

Other papers have focused on the risk allocation rather than on the total risk measurement, see Huang et al. (2009), Acharya et al. (2010), Adrian and Brunnermeier (2010), and Brownlees and Engle (2012), among others. All these papers use a risk allocation criterion based on market information and therefore produce point-in-time and risk-neutral estimates. Huang et al. (2009) use an allocation rule based on the CDS implied default rates and the equity implied asset correlations. Acharya et al. (2010) and Brownlees and Engle (2012) provide a risk allocation method based on the expected capitalization of the financial institutions on a stressed scenario. Finally, Adrian and Brunnermeier (2010) introduce the $CoVaR_i$ concept as the VaR of the portfolio conditional to the institution i being in distress and use this measure as the main driver for the risk allocation.

This paper estimates the loss distribution of the Spanish financial system under the model introduced in Vasicek (1987). This model is widely used in practice and is the starting point for the Basel Internal Rating Based capital charges (see Basel (2006)). As far as we know, Campos et al. (2007) is the only previous study that tried to measure the risk of the Spanish financial system. However, these authors a) did not take into account the diversification effect of the institutions that are not only based in Spain, b) used a base recovery value of 60% which, according to USA default data, is too low, and c) did not allocate the risk over the different financial institutions. Bennett (2002), Kuritzkes et al. (2002), and Cariboni et al. (2013) used a similar approach to that in Campos et al. (2007) to define an optimum deposits insurance fund in USA and Italy, respectively.

As we have said, Campos et al. (2007) considered a unique macroeconomic factor that links all the institutions in the economy. Our paper goes one step forward as we define as many factors as countries. We propose to use the public information of consolidated net interest income generated by the banking groups in the different countries (see BBVA (2009) and Santander (2009)) as a way to capture the risk exposure of the institutions to the different countries.

We use the Monte Carlo Importance Sampling (IS) technique introduced in Glasserman (2005) and Glasserman and Li (2005) to measure and allocate the total risk of a certain portfolio. One of the main advantages of this technique is that it can generate very accurate loss distributions and risk allocation at a low computational cost compared with that of the standard Monte Carlo method. In addition, compared with other approximate methods to obtain loss distributions like those in Pykhtin (2004) and Huang et al. (2007), its accuracy can be improved by increasing the number of simulations.

To address some criticism raised from the constant recoveries assumption we have used data of the deposits guarantee fund in United States (FDIC, Federal Deposit Insurance Corporation) to extend the IS model to deal with random recoveries. After testing several random recoveries models, our results show that the random recoveries impact on the risk allocation over the different institutions but not on the portfolio 99.9% probability loss. We have also extended the IS framework in Glasserman and Li (2005) to obtain the market valuation of the portfolio by using a model similar to that in

Grundke (2009). The impact of this valuation on the loss distribution can double that of the random recovery model.

This paper provides three major contributions. First, we measure and allocate the risk of the Spanish financial system under the IS method. Second, we extend the IS method to deal with more realistic assumptions such as random recoveries and market valuation. Third, we study the variability of the loss distribution over the business cycle and the variability of the loss distribution due to the uncertainty in the model inputs. We also highlight that a simple default mode model can seriously underestimate the possible losses and the risk allocation compared with a market mode model. We suggest not to focus only in one model but to test the impact of the different models to asses the solvency of a financial system and the impact of each financial institution. We believe that our approach goes one step forward in the current risk measurement methods applied by financial system supervisors and it can be a basic tool to identify Systemically Important Financial Institutions (SIFI) and to quantify the required capital surcharge for these institutions. As stated in Basel (2011), the Basel banking supervision Committee considers a number of global systemic banks and sets additional capital requirements using a score function that quantifies the effects of a default in one of these banks on the whole system. Among other variables, this score function considers the size, the cross-jurisdictional claims and liabilities, and positions (loans, liabilities) with other institutions. As we will see later, this interconnectedness among entities is captured in the Vasicek (1987) model through the macroeconomic common variables.

This paper is organized as follows. Section 2 reviews the main ideas regarding credit risk and the Vasicek (1987) model. Section 3 introduces the IS model proposed in Glasserman and Li (2005) and explains the optimal changes in the sampling distributions. Section 4 describes the main features of the Spanish financial system portfolio. Section 5 presents the IS results, loss distribution, and risk allocation for this financial system. Section 6 develops the random recoveries and market mode valuation extensions. Section 7 analyzes the variability of the the loss distribution over the business cycle and its variability due to the uncertainty in the model parameters estimates. Section 8 summarizes our main results and concludes.

2 The Vasicek (1987) Model

Vasicek (1987) introduced the most extended credit risk models assuming that the default behavior of a given client j (or counterparty) is driven by a set of macroeconomic factors $Z = \{z_1, z_2, \cdots, z_k\}$ and an idiosyncratic (client-specific) term ε_j. The factors $\{z_i\}_{i=1}^k$ and ε_j are independent and distributed as standard normal random variables.[1] Under these assumptions, default is modeled through the so called asset value of the client j, defined as

$$V_j = \sum_{f=1}^{k} \alpha_{f,j} z_f + \varepsilon_j \sqrt{1 - \sum_{f=1}^{k} \alpha_{f,j}^2} \tag{1}$$

This client defaults in her obligations if V_j falls below a given default threshold level k. As $V_j \sim N(0,1)$, we have that $k = \Phi^{-1}(PD_{j,C})$, where $\Phi(\cdot)$ denotes the normal distribution function and $PD_{j,C}$ denotes the historical average default rate of the client j over long enough periods.[2]

Given the specification (1) and conditional to the macroeconomic factors Z, the default probability of the client j is

$$Prob(D_j = 1|Z) = Prob(V_j \leq k|Z) = \Phi\left(\frac{\Phi^{-1}(PD_{j,C}) - \sum_{f=1}^{k} \alpha_{f,j} z_f}{\sqrt{1 - \sum_{f=1}^{k} \alpha_{f,j}^2}}\right)$$

Bank portfolios are composed of this kind of contracts. The total loss of a portfolio including M contracts or clients with individual losses x_j is given as $L = \sum_{j=1}^{M} x_j$. Under infinite granularity, the idiosyncratic risk of the different clients disappears and there is no uncertainty on the loss conditional to the macroeconomic scenario.

Under granular homogeneous single factor portfolios, the unconditional default rate distribution function is given as

$$\begin{aligned} Prob(DR_z \leq L) &= Prob\left(\Phi\left(\frac{\Phi^{-1}(PD_C) - \alpha z}{\sqrt{1-\alpha^2}}\right) \leq L\right) \\ &= \Phi\left(\frac{\Phi^{-1}(L)\sqrt{1-\alpha^2} - \Phi^{-1}(PD_C)}{\alpha}\right) \end{aligned}$$

Since the Basel II accord, the banking regulation uses the Vasicek (1987) asymptotic single factor model and forces the financial institutions to have an amount of own resources (equity and other assets with similar behavior to the equity) equal to the worst loss with a 99.9% probability.

[1] Dependent factors can always be orthogonalized.

[2] It might be more useful to think on the historical average default rates of clients similar to j rather than on the historical average default rates of the client j.

The estimation of the portfolio loss distribution requires estimating PD_C for the different portfolios. This can be done by using the historical default rates of the portfolios but another components are also needed:

1. EAD: Exposure at default, the amount of money owed by the investor when he defaults.
2. LGD: Loss given default, the final loss after all the recovery processes.[3]
3. α: Sensitivity of the asset value to the macroeconomic factors. The Basel accord provides standard α values for the different portfolios of a bank.

Then, the portfolio loss can be expressed as

$$L = \sum_{j=1}^{M} x_j = \sum_{j=1}^{M} EAD_j LGD_j \mathbf{1}(V_j \leq \Phi^{-1}(PD_{j,C}))$$

In the general case of non-granular, non-homogeneous and multi-factor portfolios, the loss distribution of a loan portfolio can be obtained by Monte Carlo methods or by approximated ones.

It should be noted that our objective is to know just some statistical measures of the accumulated loss distribution $F(L)$, being the most important the following ones:

1. Value at Risk: $VaR(q) = F^{-1}(q)$.[4]
2. Expected Shortfall or Tail VaR, that is, the expected loss given that a minimum loss level has been reached: $ES(q) = E(L|L \geq VaR(q))$.
3. Risk contributions of the client j. We can consider two alternatives:
   (a) Value at Risk contribution, $CVaR_j(q) = E(x_j|L = VaR(q))$.
   (b) Expected Shortfall contribution, $CES_j(q) = E(x_j|L \geq VaR(q))$.

[3]For a certain client j, both EAD_j and LGD_j are random variables although they are commonly assumed to be constant. Along the paper, we will indicate whether the LGD is in percentage terms of the EAD or in euros.

[4]The Basel regulation requires a bank to have an amount of own resources equal to the $VaR(99.9\%)$.

3 Importance sampling for credit risk

The importance sampling (IS) is a Monte Carlo simulation method that helps to estimate expectations of random variables through an smart change of the sampling distribution. As explained previously, the most general measure in credit risk is $Prob(L \geq l)$, directly related to the VaR at a given confidence level, or the maximum loss with a given probability. Then, to apply the IS method, we start transforming this probability into an expectation as follows

$$Prob(L \geq l) = E(\mathbf{1}(L \geq l)) = \int_{-\infty}^{\infty} \mathbf{1}(L \geq l) f(L) dL = \int_{-\infty}^{\infty} \mathbf{1}(L \geq l) \frac{f(L)}{g(L)} g(L) dL$$

One estimator of $Prob(L \geq l)$ is then given as

$$\widehat{Prob}(L \geq l) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}(L_i \geq l) \frac{f(L_i)}{g(L_i)}$$

where L_i is sampled from $g(L)$. As the simulated random variables are independent, the variance of this estimator is[5]

$$\begin{aligned} Var(\widehat{Prob}(L \geq l)) &= \frac{1}{N^2} \sum_{i=1}^{N} Var\left(\mathbf{1}(L_i \geq l) \frac{f(L_i)}{g(L_i)}\right) = \frac{1}{N} Var\left(\mathbf{1}(L_i \geq l) \frac{f(L_i)}{g(L_i)}\right) \\ &\approx \frac{1}{N}\left[\frac{1}{N} \sum_{i=1}^{N} \mathbf{1}(L_i \geq l) \frac{f^2(L_i)}{g^2(L_i)} - \left(\frac{1}{N} \sum_{i=1}^{N} \mathbf{1}(L_i \geq l) \frac{f(L_i)}{g(L_i)}\right)^2\right] \end{aligned}$$

where we have used sample statistics. Using this variance estimate and the central limit theorem we can get the confidence intervals of the probability estimates.

The expected shortfall (ES) is defined as

$$ES = E(L|L \geq l) = \int_{-\infty}^{\infty} L f(L|L \geq l) dL = \frac{\int_L^{\infty} L f(L) dL}{\int_L^{\infty} f(L) dL}$$

and can be estimated using the IS method as

$$\widehat{ES} = \frac{\sum_{i=1}^{N} L_i \mathbf{1}(L_i \geq l) \frac{f(L_i)}{g(L_i)}}{\sum_{i=1}^{N} \mathbf{1}(L_i \geq l) \frac{f(L_i)}{g(L_i)}}$$

[5] It can be noted that the variance of this estimator vanishes for the sampling distribution $g(L_i) \propto \mathbf{1}(L_i \geq l) f(L_i)$.

The estimators for the VaR and ES risk contributions of the client j are respectively

$$\widehat{CVaR_j} = \frac{\sum_{i=1}^{N} x_{j,i}\mathbf{1}(L_i = l)\dfrac{f(L_i)}{g(L_i)}}{\sum_{i=1}^{N} \mathbf{1}(L_i = l)\dfrac{f(L_i)}{g(L_i)}}, \quad \widehat{CES_j} = \frac{\sum_{i=1}^{N} x_{j,i}\mathbf{1}(L_i \geq l)\dfrac{f(L_i)}{g(L_i)}}{\sum_{i=1}^{N} \mathbf{1}(L_i \geq l)\dfrac{f(L_i)}{g(L_i)}}$$

As $\widehat{CVaR_j}$ can not be implemented computationally, the following modification is required:

$$\widehat{CVaR_j} = \frac{\sum_{i=1}^{N} x_{j,i}\mathbf{1}(l(1-R) \leq L_i \leq l(1+R))\dfrac{f(L_i)}{g(L_i)}}{\sum_{i=1}^{N} \mathbf{1}(l(1-R) \leq L_i \leq l(1+R))\dfrac{f(L_i)}{g(L_i)}}$$

where R is an interval defining parameter. From now on we will employ $R = 1\%$.

The confidence intervals of the expected shortfall and the risk contributions can be derived using Serfling (1980) to obtain that

$$Var(\widehat{ES}) \approx N \frac{\sum_{i=1}^{N} (L_i - \widehat{ES})^2 \mathbf{1}(L_i \geq l) \left(\frac{f(L_i)}{g(L_i)}\right)^2}{\left(\sum_{i=1}^{N} \mathbf{1}(L_i \geq l)\frac{f(L_i)}{g(L_i)}\right)^2} \tag{2}$$

This equation can be extended to provide estimators of the variance of the empirical estimates of the ES and VaR risk contributions just replacing $(L_i - \widehat{ES})$ by $(x_{j,i} - \widehat{ES})$ or $(x_{j,i} - \widehat{VaR})$ and $\mathbf{1}(L_i \geq l)$ by $\mathbf{1}(l(1-R) \leq L_i \leq l(1+R))$ in (2).

So far no functional form for the function $g(L)$ has been suggested. Glasserman and Li (2005) suggested to obtain $g(L)$ in two steps, changing a) the default probabilities conditional on the macroeconomic factors and b) the macroeconomics factors distribution, respectively.

3.1 Optimal conditional distribution

Conditional to the macroeconomic factors realization, the default probability of the client j is

$$PD_{j,Z} = \Phi\left(\frac{\Phi^{-1}(PD_{j,C}) - \sum_{f=1}^{k} \alpha_{f,j} z_f}{\sqrt{1 - \sum_{f=1}^{k} \alpha_{f,j}^2}}\right)$$

Glasserman and Li (2005) suggested to change the default probability by a new one using an exponential twist

$$PD_{j,Z,\theta} = \frac{PD_{j,Z}e^{LGD_j EAD_j \theta}}{1 + PD_{j,Z}(e^{LGD_j EAD_j \theta} - 1)}$$

The change in the default probability of a client depends only on his specific default parameters plus a parameter θ, common for all the clients. Under this twist, the weight to be assigned to every loss simulation i of the total portfolio is

$$W_{1,i} = \frac{f(D_{i,1}, \cdots, D_{i,M})}{g(D_{i,1}, \cdots, D_{i,M})} = \prod_{j=1}^{M} \left(\frac{PD_{j,Z}}{PD_{j,Z,\theta}} \right)^{D_{j,i}} \left(\frac{1 - PD_{i,Z}}{1 - PD_{j,Z,\theta}} \right)^{1-D_{j,i}}$$

where $D_{j,i}$ is the default indicator of the client j in the simulation i. A little algebra leads to

$$W_{1,i} = e^{-L_i\theta + \psi(\theta)}$$

where

$$\psi(\theta) = \sum_{j=1}^{M} \ln \left(1 + P_{j,Z} \left(e^{LGD_j EAD_j \theta} - 1 \right) \right) \tag{3}$$

Note that, conditional to the macroeconomic state Z, the losses of every client j are independent. Then, (3) implies that $\psi(\theta)$ is the cumulant generating function of the random variable $L(Z)$, with an important role in the saddlepoint approximation method.

Now the problem is to estimate the optimal value of θ that minimizes the variance of the estimator under the new distribution $g(L, \theta)$. Glasserman and Li (2005) proved that

$$Var_{g(L,\theta)} \left(\frac{1}{N} \sum_{i=1}^{N} \mathbf{1}(L_i \geq l) W_{1,i} \right) \leq e^{-2\theta L + 2\psi(\theta)}$$

Differentiating this upper bound and using the convexity of $\psi(\theta)$, the optimum shift θ_l satisfies $\psi'(\theta_l) = l$ if $l > \psi'(0)$ being null otherwise. Straightforward calculations lead to $\psi'(\theta) = \sum_{j=1}^{M} LGD_j EAD_j PD_{j,Z,\theta} = E_{g(L,\theta)}(L)$.

The intuition behind this result is that we aim to obtain high enough losses close to the loss value l. Under the current macroeconomic factor simulations, expected losses can be much lower than l and, then, the default probabilities are changed so that the new expected losses equate the desired loss level, this is done by using $\theta_l \geq 0$. However, if the actual expected losses are higher than the desired one (l), default probabilities are not changed at all. In this case θ_l should be negative to get an expected loss of l.

If the *VaR* based loss contributions (CVaR) are calculated, the default probabilities will always be shifted to the desired loss level l, so that many simulations will lay inside the interval $l(1 \pm R)$. According to our experience, the number of simulations in the *VaR* interval can be doubled from that obtained when forcing $\theta_l \geq 0$.

Another interesting property of the Glasserman and Li (2005) approach is that, as $\psi(\theta)$ equates the cumulant generating function, the optimization problem $\psi'(\theta) = l$ to be solved under the IS method coincides with that solved under a saddlepoint approach. The value θ_l is computed through a non-linear iterative process that departs from an initial estimate obtained by applying a third-order Taylor expansion to $\psi(\theta)$ around $\theta = 0$.[6]

3.2 Optimal macroeconomic distribution

As with the default probability it is possible to change the distribution of the macroeconomic factors to a new one that reduces the variance of the estimates. The probability we are interested in is

$$Prob(L \geq l) = \int_{-\infty}^{\infty} Prob(L \geq l|Z) f(Z) dZ \propto \int_{-\infty}^{\infty} Prob(L \geq l|Z) e^{-\frac{Z'Z}{2}} dZ$$

The optimal sampling distribution $g(Z)$ is proportional to $Prob(L \geq l|Z)e^{-(Z'Z)/2}$. Sampling from this distribution is complex but feasible through the Markov chain Monte Carlo technique using the Metropolis-Hasting algorithm. However, Glasserman and Li (2005) suggested sampling from a normal distribution with the same mode as the optimum distribution, that is, $g(Z) \sim N(\mu, I)$, where $\mu = \max_Z \left\{ Prob(L \geq l|Z)e^{-(Z'Z)/2} \right\}$. According to this, a new weight $W_{2,i} = e^{-\mu' Z + \mu'\mu/2}$ has to be applied and the IS estimators will be given by

$$\begin{aligned} \widehat{Prob}(L \geq l) &= \frac{1}{N} \sum_{i=1}^{N} \mathbf{1}(L_i \geq l) W_{1,i} W_{2,i} \\ \widehat{E}(L|L \geq l) &= \frac{\frac{1}{N}\sum_{i=1}^{N} L_i \mathbf{1}(L_i \geq l) W_{1,i} W_{2,i}}{\widehat{Prob}(L \geq l)} \end{aligned}$$

[6] We used this expansion to approximate the non-linear problem that has to be solved and defined a rule to choose among the three possible solutions. This approach generated initial estimates very close to the real value of θ_l.

It still remains to estimate $Prob(L \geq l|Z)$. To this aim, we decided to use a simple approach assuming that $L|Z \sim N(a, b^2)$ where[7]

$$\begin{aligned} a &= E(L|Z) = \sum_{j=1}^{M} PD_{j,Z} LGD_j EAD_j \\ b^2 &= Var(L|Z) = \sum_{j=1}^{M} Var(x_j|z) = \sum_{j=1}^{M} PD_{j,Z}(1 - PD_{j,Z}) LGD_j^2 EAD_j^2 \end{aligned}$$

4 Portfolio data

We evaluate alternative credit risk measures (loss distribution and risk contributions) considering the 157 financial entities covered by the Spanish deposit guarantee fund (FGD) at December, 2010.[8] This fund was analyzed in Campos et al. (2007) by using a simple single factor model and plain Monte Carlo simulations. These authors just tested a range of constant *LGDs* not directly linked to historical recovery rates and did not estimate any risk contribution measure. We will try to overcome these limitations and will assume that the two biggest institutions (BBVA and Santander) are exposed to other economies and, hence, to other macroeconomic factors.

4.1 Probability of default (PD)

We use the credit ratings available at December, 2010 for the Spanish financial institutions and the historical observed default rates reported by the rating agencies[9] to infer a probability of default. The ratings of the agencies for these institutions can include some implicit Government support that improves the risk profile. This support is relevant for the investors but from the point of view of the Government, we should remove its effect from the ratings. This task is not straightforward and we decided not to perform it and keep our framework as simple as possible.

The probability of default is obtained adjusting an exponential function to the default rates of the ratings up to B- while the rating AA- is provided a

[7]Other alternatives such as the *constant approach* or the *tail bound approach* can be found in Glasserman and Li (2005).

[8]The FGD is built up to help the financial system stability and includes the three previously existing funds (for banks, saving banks, and cooperative banks) that were merged in October, 14th, 2011 under the Real Decreto 16/2011.

[9]See Fitch (2009), Moody's (2009), and S&P (2009).

probability of 0.03%, a commonly accepted number. Entities without external rating are assigned one notch less than the average rating of the portfolio with external rating.[10] This implies that banks without external rating receive an A- and saving banks a BB+ rating, values that are consistent with Campos et al. (2007). Once a rating is recovered, a long-term default rate is assigned to each institution.

We obtain that the S&P and Moody's ratings have very similar historical default rates for the different rating letters while Fitch rating is very different from the other two.[11] Even though Fitch and S&P use the same letters to measure credit risk, the underlying default risk is different, specially for the very bad ratings. Luckily, no institution had this rating at the date of analysis and, then, we can still use the calibrated probabilities of default.

4.2 Exposure at default (EAD)

Details on assets, liabilities, and deposits for the FGD institutions are available in the AEB, CECA, and AECR webpages.[12] The FGD covers not only depositors but also any loss due to a Governmental intervention of a financial institution. Hence, our analysis focuses on total assets losses and not only on losses to depositors.

Balance information at December 2010 was used for the analysis. As many mergers took place during 2010 (see Table 1), we have summed all the information from the different institutions that belong to the same group.

[INSERT TABLE 1 AROUND HERE]

Figure 1 shows the assets and the deposits shares of the top 25 financial institutions. These entities account for 92.1% of the assets and 92.8% of the deposits in the financial system. The inverse of the Herfindahl index[13] H shows that there are only 10.8 and 13.7 effective counterparties (from both assets and deposits points of view). This means that the Spanish financial

[10] This average is computed weighting by assets and distinguishing between banks and saving banks.

[11] For the sake of brevity, these results are not reported here and are available upon request.

[12] AEB is the Spanish Bank Association, CECA is the Spanish Saving Bank Association, and AECR is the Spanish Credit Cooperatives Association. Other sources as Bankscope were tested, however the set of available institutions was smaller.

[13] The Herfindahl index is a measure of portfolio concentration and its inverse can be seen as the number of effective counterparties in the portfolio. See Allen et al. (2006), Hartmann et al. (2006), and Carbó et al. (2009) for further details on the Herfindahl index in the banking sector.

system has few players and is very concentrated, a common feature in most of the countries worldwide.

[INSERT FIGURE 1 AROUND HERE]

4.3 Loss given default (LGD)

Schuermann (2004) provided a review of the (academic and practitioner) literature on the LGD. In more detail, this author focused on the meaning of the LGD and its role in the internal ratings based (IRB) approach, described the main factors that can drive $LGDs$, and discussed several approaches that can be applied to model and estimate the LGD. See also Carey (1998), Altman and Suggitt (2000), Amihud et al. (2000), Thorburn (2000), Unal et al. (2003), and Altman et al. (2005), among others, for details on the LGD main characteristics.

As Schuermann (2004) stated in its Section 7, *"the factors (or drivers or explanatory variables) included in any LGD model will likely come from the set of factors we found to be important determinants for explaining the variation in LGD. They include factors such as place in the capital structure, presence and quality of collateral, industry and timing of the business cycle."* In practice, industry models such as LossCalc^{TM} use most of these factors, see Gupton and Stein (2002) for more details on this model.

Bennett (2002) computed the losses due to financial institutions default in the FDIC and showed that the average losses are bigger in the smallest banks for the period 1986-1998. We update this analysis for the period 1986-2009 using FDIC public data and the banks assets are updated using the USA CPI series aiming to have comparable asset sizes. We obtain an average LGD for deposits of 20.73% but this value may be biased as there are many observations in the initial and final years of the database. Hence, we decide to use $E(E(LGD_{j,t}|t))$ as an estimate of the real average LGD and obtain 18.35%, that is, 88.56% of the initial average LGD. Then, we estimate $E\left(LGD_{j,t}|\text{Asset Bucket}\right)$ and multiply it by the 88.56% adjustment factor. Finally, these $LGDs$ on deposits are transformed into $LGDs$ on assets using a multiplicative factor of 1.378.[14] Table 2 provides the $LGDs$ obtained in this way.

[INSERT TABLE 2 AROUND HERE]

[14] This factor is based on the numbers obtained in Bennett and Unal (2011) that used FDIC data for 1986-2007 and estimated an average depositors LGD of 24.4%, equivalent to a 29.95% total LGD over assets before the time effect and a 33.61% after the discount effect.

4.4 Factor correlation (α)

We use the total factor sensitivities (α) stated in the Basel accord. These values are computed according to the formula $\sqrt{0.12\omega + 0.24\,(1-\omega)}$ where $\omega = \frac{1-e^{-50PD}}{1-e^{-50}}$ and, hence, range between $\sqrt{0.12}$ and $\sqrt{0.24}$.[15] Recently, the Basel III accord has increased the previous Basel II correlations by a factor of $\sqrt{1.25}$. In this way, we would generate correlations in the range of those used in Campos et al. (2007). In the following analysis we use the Basel III correlations.

We assume geographic macroeconomic factors and that all the financial institutions are exposed only to the Spanish factor except for BBVA and Santander that are exposed to additional geographies. This assumption seems reasonable and its motivation can be seen in Figure 1 which shows that, among the 25 biggest financial institutions, apart from these two entities, only Barclays is not a fully Spain based bank and its share is very small.

The exposure of BBVA and Santander to the macroeconomic factors is computed using the reported net interest income by geography obtained from the public 2010 annual reports. We think that this variable can be a good proxy of the risk faced by a financial institution and, then, it can indicate appropriately its exposure to the different countries in which the institution operates. Hence, an income based allocation method can be better than a method only based on exposures that would assign small weights to the non-Spanish geographies.

Finally, we assume that the correlation between the macroeconomic factors for different countries is equal to that between the GDP of the countries.[16]

Table 3 shows the exposure of BBVA and Santander to the different countries according to their net interest incomes. As these country factors are correlated, those exposures have to be standardized so that the total variance of the sum of each client's macroeconomic factors equates one.

[INSERT TABLE 3 AROUND HERE]

4.5 Portfolio expected loss and Basel loss distribution

The total assets, expected loss, and BIS 99.9% probability loss for the Spanish financial institutions are 2,921,504 MM €, 453 MM €, and 13,733 MM

[15] Kuritzkes et al. (2002) and Campos et al. (2007) use $\sqrt{0.15}$ and $\sqrt{0.30}$, respectively. In practice, most of the entities show sensitivities closer to $\sqrt{0.24}$.

[16] These correlations are available upon request.

€, respectively.

Figure 2 includes these numbers for the top 25 Spanish financial institutions. The left graph in this Figure shows the share of these variables for the biggest (ordered by assets) 25 financial institutions. For example, Santander represents 21% of the total assets, 7% of the total BIS 99.9% loss, and 4% of the expected loss of the Spanish financial system, approximately.

[INSERT FIGURE 2 AROUND HERE]

Two conclusions can be extracted from this Figure:

1. Expected loss and Basel 99.9% probability loss generate a very similar ordering.

2. The ordering according to the assets amount is very different from that based on expected or Basel losses.

The right graph in Figure 2 shows the expected loss and Basel 99.9% loss divided by the size of each institution. We find that the two biggest institutions (BBVA and Santander) share very low risk parameters.

We will introduce now the results obtained with the IS method as a way to deal with non-granular and multifactorial portfolios. The main ideas behind this modification of the asymptotic single factor model are a) BBVA and Santander have some diversification effects as they are exposed to more than one macroeconomic factor that reduces their risk and b) having non-granular portfolios increases the risk.

5 Importance sampling results

We start orthogonalizing the country factors by applying principal components analysis. As the correlation between the different economies is very high we end up having a very important common factor across all the financial institutions. When we obtain the optimum change in the factor mean for a target loss of 10 times the expected loss we get a 1.62 value in the main common factor and almost zero otherwise.

Figure 3 shows the loss distribution under IS and plain Monte Carlo simulations. According to the Basel model the loss level with 99.9% probability is 13,733 MM€. While under multifactorial non-granular portfolios this loss level is 32,102 MM€, 2.3 times more![17]

[17] All the figures in the paper are based on the IS results rather than on the plain Monte Carlo method.

[INSERT FIGURE 3 AROUND HERE]

Figure 4 shows the results for the expected shortfall. The *VaR* contributions are usually less stable as few simulations fall inside the interval. That is why it is quite common using the expected shortfall contributions at a loss level whose tail expectation equals the $VaR(99.9\%) = 32{,}102$ MM €. In this case this loss level is 16,274 MM €.

[INSERT FIGURE 4 AROUND HERE]

Figure 5 shows the risk allocation rule according to the *ES* and *VaR* contributions and the confidence intervals for the IS technique.[18] These intervals are quite thin after only 10,000 simulations, one of the main advantages of the IS method over the plain Monte Carlo simulations. Moreover, the IS method can generate many high loss simulations from a thin loss interval and, then, more accurate estimates at a lower computation time. The risk picture is completely different from that obtained using the simple expected loss or the Basel loss model. The main reason for this is that the non-granularity effect increases (decreases) the risk allocated to the biggest (smallest) institutions.

[INSERT FIGURE 5 AROUND HERE]

The main ideas that can be extracted from this Figure are the following:

1. The *LGDs* (in euros) for BBVA and Santander are higher than the $VaR(99.9\%)$. Then their *VaR* contributions are zero. This is a serious limitation of the VaR based risk allocation for concentrated portfolios with low default probabilities. In fact, this can induce a higher concentration profile as it promotes the growth of the biggest financial institutions. Alternatively, we can allocate the risk to Santander and BBVA considering the risk that would be allocated to two institutions identical to BBVA and Santander but with one euro exposure each.

2. The *LGD* of Bankia is 28,948 MM €, close to the $VaR(99.9\%)$ value. Then, this firm captures most of the risk under the *VaR* contribution allocation criterion.

3. The risk allocations of Caixabank and Unnim have big confidence intervals. This is due to the fact that the *LGD* of both entities together is close to the $VaR(99.9\%)$ and there are few simulations in which Caixabank and Unnim default.

[18] For the *VaR* contributions we have used a $\pm 1\%$ interval around the desired loss level.

4. The confidence intervals of the 99.9% probability loss ratio are bigger as the risk is adjusted by the institution size and Unnim has the biggest confidence intervals for the risk allocation.

6 Importance sampling modifications and extensions

This Section extends the classical IS framework to deal with random recoveries and market valuation. Other extensions were performed:[19]

1. We found that using the mode for the macroeconomic factor shifts may introduce a low sampled region problem and we developed a method based on the mean of the optimal distribution to overcome this problem.

2. For granular multifactorial portfolios, we found that the 99.9% probability losses of the Spanish financial are 13,478 MM €.

3. We also evaluated the suitability of the *simulation loop decoupling*, based on simulating N_{Macro} macroeconomic scenarios and $N_{Default}$ default scenarios for each (simulated) macroeconomic scenario. This modification is very interesting in terms of speed and accuracy for portfolios with few counterparties that are exposed to the same macroeconomic factor, as it is our case. The following IS results are based on this extension.

6.1 Random loss given default

So far the LGD has been considered as constant but it is a random variable with the same span as the default rates. Then, it seems natural to assume that the LGD follows a similar distribution to that of the default rate. Considering this, the simplest case assumes that the whole recovery risk comes from macroeconomic factors, for example, a single factor called z_{LGD}:

$$LGD_{j,Z} = \Phi\left(\frac{\Phi^{-1}(LGD_{j,C}) - \alpha_j z_{LGD}}{\sqrt{1-\alpha_j^2}}\right)$$

Under this specification the only parameters to be estimated are α_j and the correlation between z_{LGD} and the rest of the macroeconomic factors.

[19] For the sake of brevity, we just enumerate here these additional extensions and defer the details to a final Appendix.

This model also allows to have more macroeconomic factors but the idea is that no idiosyncratic risk is considered.

The previous formula has been widely studied[20] and some of their moments have a closed-form expression, for example

$$
\begin{array}{rcl}
E(LGD_{j,Z}) & = & LGD_{j,C} \\
E(LGD_{j,Z}^2) & = & \Phi_2\left(\Phi^{-1}\left(LGD_{j,C}\right), \Phi^{-1}\left(LGD_{j,C}\right), \alpha_j^2\right)
\end{array}
$$

where $\Phi_2(x, y, \rho)$ stands for the probability distribution function (evaluated at the point (x, y)) of a bivariate standard normal random variable with correlation parameter ρ.

We have shown previously that the LGD depends on the institution size and that most of the defaults in our sample correspond to institutions with less than 1,000 MM € in assets. To keep the database as clean as possible we will estimate the parameters using just the institutions with this assets size.

The above formulas and the historical recovery rates from the FDIC data imply $LGD_{j,C} = 19.13\%$, $E(LGD_{j,Z}^2) = 4.3178\%$ and, therefore, $\alpha_j = 29.26\%$. Using these estimates we recover the z_{PD} and z_{LGD} factors from the historical default series of the FDIC and obtain that the correlation between the default and recovery factors is 22.63%. The random LGD is introduced replicating the factor correlation of the PD for the LGD as follows:

$$
G = \left[\begin{array}{ccc|ccc}
 & & & 22.63\% & 0\% & \cdots \\
 & M_{PD} & & 0\% & \cdots & 0\% \\
 & & & \cdots & 0\% & 22.63\% \\
22.63\% & 0\% & \cdots & & & \\
0\% & \cdots & 0\% & & M_{LGD} & \\
 & \cdots & 22.63\% & & &
\end{array}\right]
$$

where $M_{PD} = M_{LGD}$ equates the GDP correlation matrix of the different countries.[21] Now not only $PD_{j,z}$ has to be estimated but also $LGD_{j,z}$ in every simulation step. The optimal exponential twist and the optimal change in the mean of the macroeconomic factors are obtained using $PD_{j,Z}$ and $LGD_{j,Z}$.

Figure 6 shows the comparison between the loss distributions of the portfolio under random and constant $LGDs$. The 99.9% probability loss is 36,970 MM €, that is, 1.15 times the loss level under constant LGD.

[20] See Gordy (2000) or Dullmann et al. (2010).

[21] For BBVA and Santander the weights of the LGD to the different LGD factors are the same as those defined before according to their net interest incomes.

The equivalent expected shortfall level is 19,326 MM €. Figure 7 shows the risk allocation under *VaR* and *ES* for the new 99.9% probability loss level. Comparing with Figure 5 we can see that this model assigns risk to all the institutions, even to Santander whose initial *LGD* was 53,146 MM €, much higher than the 99.9% probability loss. However, as now the *LGD* is random, there are some scenarios where Santander defaults and the total loss is close to 36,970 MM €.

[INSERT FIGURES 6 AND 7 AROUND HERE]

Compared with the constant *LGD* case, the random *LGD* provides the following facts:

1. The confidence intervals in the risk allocation are wider. Now, in the event of default, the losses have a bigger variability and, hence, the estimation of $E(X_i|L = VaR)$ is also more volatile.

2. The risk allocations based on the *VaR* and the *ES* are relatively "similar" and the risk is not concentrated in some institutions as in the case of constant *LGD*.

Under the Basel accord, the random *LGD* is considered under a very broad definition of a downturn *LGD*, defined as the *LGD* under a stress scenario. This constant downturn *LGD* tries to capture somehow the effect of the random *LGD*.

In the previous setup, two clients with the same $LGD_{j,C}$ and the same sensitivity to the macroeconomic variables will have the same $LGD_{j,Z}$. To avoid this possibility, an idiosyncratic term $\gamma_j \sim N(0,1)$ can be included in the previous formula:

$$LGD_{j,z,\gamma_j} = \Phi\left(\frac{\Phi^{-1}(LGD_{j,C}) - \alpha_j(rz_{LGD} + s\gamma_j)}{\sqrt{1-\alpha_j^2}}\right)$$

with $r^2 + s^2 = 1$. This second specification reduces the correlation between the *LGD* and the defaults as a new independent term is considered but it can increase the variability of the recoveries.

The parameter r controls the variability in $LGD_{j,Z}$ over the business cycle. According to our data, we obtain $E(Var(LGD_{j,z}|z)) = 1.338\%$,[22] implying a variability that is higher than the average *LGDs* of the big financial institutions. Intuitively, now, more institutions can generate high and

[22] Then, the *LGD* can change $\pm 11.56\%$ with respect to its mean.

low loss levels compared to the constant LGD case and the confidence intervals will be wider than under constant LGD and under fully macroeconomic random LGD. Calibration of the LGD data provides $r = 59.39\%$

Using these data, for every default and recovery observation in the FDIC database, we recover the value $rz_{LGD}+s\gamma_j$ using the previous formula. Then, for every year, we obtain empirically $E(rz_{LGD} + s\gamma_j)$ that equates rz_{LGD}. In this way we estimate z_{LGD} for every year and obtain that the correlation between z_{LGD} and the default driving macroeconomic factor z_{PD} is 19.02%.

This new specification causes some changes in the IS framework. For instance, the exponential twist of the default probabilities conditional to a given set of macroeconomic factors was defined as that generating an expected loss equal to the target loss level. Now, conditional to these factors, $LGD_{j,z}$ is not constant and we have two alternatives to find the optimum exponential twist:

1. To keep using the average loss given default $LGD_{j,C}$ regardless of the macroeconomic factors.
2. To estimate $E\left(LGD_{j,z} \mid z\right)$ and $E\left(LGD^2_{j,z} \mid z\right)$ for every macroeconomic factor simulation.

We use the second method given that $E(LGD_{j,z})$ has a closed-form expression given as

$$E(LGD_{j,z}) = Prob(V_{j,z} < \Phi^{-1}(LGD_{j,C})) = \Phi\left(\frac{\Phi^{-1}(LGD_{j,C}) - \alpha r z_{LGD}}{\sqrt{\alpha^2(s^2-1)+1}}\right)$$

Computing the optimal change in the mean of the factors is a bit more complex as it requires estimating $Var(LGD_{j,z}|z)$ or, equivalently, $E\left(LGD^2_{j,z}|z\right)$, this is,[23]

$$E\left(LGD^2_{j,z}|z\right) = \Phi_2\left(\begin{pmatrix} \Phi^{-1}(LGD_{j,C}) \\ \Phi^{-1}(LGD_{j,C}) \end{pmatrix}, M, \Sigma\right)$$

with

$$\begin{aligned} M &= \begin{pmatrix} \alpha r z_{LGD} \\ \alpha r z_{LGD} \end{pmatrix} \\ \Sigma &= \begin{pmatrix} \alpha^2 s^2 + (1-\alpha^2) & \alpha^2 s^2 \\ \alpha^2 s^2 & \alpha^2 s^2 + (1-\alpha^2) \end{pmatrix} \end{aligned}$$

[23] $\Phi_2(X, M, \Sigma)$ denotes the probability distribution function (evaluated at the point X) of a bivariate normal random variable with mean vector M and covariance matrix Σ.

It is worthy to note that the optimal exponential twist is generated using $E\left(LGD_{j,Z,\gamma_j} \mid Z\right)$ rather than the simulated LGD_{j,Z,γ_j}. Then the weight $W_{1,i}$ must be obtained using $E\left(LGD_{j,Z,\gamma_j} \mid Z\right)$ rather than the realized LGD_{j,Z,γ_j}, that is, using $L_i^* = \sum_{j=1}^{M} D_{j,i} EAD_j E\left(LGD_j \mid Z\right)$ instead of L_i. This is

$$W_{1,i} = e^{-L_i^*\theta + \psi(\theta)}$$

where

$$\psi(\theta) = \sum_{j=1}^{M} \ln\left(1 + P_{j,Z}\left(e^{E(LGD_{j,Z,\gamma_j}|Z)EAD_j\theta} - 1\right)\right)$$

Under random conditional recoveries, the simulation process is as follows:

1. Get the change in the mean of the macroeconomic factors based on approximating $Prob(L \geq l|Z)$ by a normal distribution with parameters

$$\begin{aligned}
\mu &= \sum_{j=1}^{M} EAD_j PD_{j,Z,\theta} E(LGD_{j,Z,\gamma_j}|Z) \\
\sigma^2 &= \sum_{j=1}^{M} EAD_j^2 PD_{j,Z,\theta} E(LGD_{j,Z,\gamma_j}^2|Z) - \mu^2
\end{aligned}$$

2. We obtain a sample of macroeconomic factors Z based on the change in the mean estimated in the previous step.

3. For each macroeconomic scenario Z, we obtain the twist of the conditional default probabilities ($PD_{j,Z,\theta}$) such that

$$\sum_{j=1}^{M} EAD_j PD_{j,Z,\theta} E(LGD_{j,Z,\gamma_j}|Z) = l$$

4. Simulate losses for each macroeconomic scenario based on $PD_{j,Z,\theta}$ and a random drawn of LGD_{j,Z,γ_j}.

5. The change in the conditional default probabilities is not obtained from the simulated LGD (LGD_{j,Z,γ_j}) but from its expectation conditional to Z. Then, the weight $W_{1,i}$ must be computed using $E(LGD_{j,Z,\gamma_j}|Z)$ instead of LGD_{j,Z,γ_j}.

This simulation process is very interesting under the loop decoupling framework as, for each macroeconomic scenario, the twist of the conditional default probabilities remains constant and, then, the number of required calculations decreases.

Figure 8 provides the loss distributions under the three possible specifications: constant LGD, macroeconomic random LGD (LGD^C), and macroeconomic plus idiosyncratic random LGD (LGD^R). It can be seen that considering the idiosyncratic term adds some more risk to the 99.9% loss level.

[INSERT FIGURE 8 AROUND HERE]

The effect of the idiosyncratic risk is quite small in the loss distribution. Using the IS results, the 99.9% loss level under the idiosyncratic risk is 37,934 MM€, only 964 MM € more than that under the macroeconomic LGD model.[24] Hence, the impact of the idiosyncratic LGD on the loss distribution is small compared with that of the macroeconomic LGD. It can also be noted that, for small (large) loss levels, the idiosyncratic risk term reduces (increases) the chance of those losses.

Regarding the risk allocation, Figure 9 shows that, in this case, the (absolute and relative) risk allocation has even bigger confidence intervals than in the previous models. The reason is that previously highlighted: given default, the variability of the losses of the client j are wider under the idiosyncratic LGD model than under the pure macroeconomic LGD.

[INSERT FIGURE 9 AROUND HERE]

Other LGD distributions have been tested for the pure macroeconomic LGD model (LGD^C) and the mixed macroeconomic and idiosyncratic LGD model (LGD^R).[25]

Table 4 includes the resulting loss distributions using the IS method and shows that the results of the different random LGD models for the 99.9% loss level are quite similar in all the cases except for the Log-Normal one.

[INSERT TABLE 4 AROUND HERE]

To conclude this subsection, we want to mention that, in the random LGD framework, an alternative is to apply the IS method to the LGD distribution rather that to the default distribution. In fact with the IS ideas

[24] The ES equivalent loss is 19,473 MM €.

[25] Detailed results are not reported here and are available upon request.

we are interested in changing the conditional losses distribution so that the probability of high losses increases regardless we change the default probabilities or the *LGD* distribution. This way of thinking only applies to the case of random conditional *LGD*. In our case we have decided to maintain the ideas introduced in the previous sections and change just the default probabilities.

6.2 Market mode

This Subsection evaluates the portfolio risk under a market value model instead of a default mode one. Under this model the rating of the companies may change over the time and these changes affect the firm valuation. Then, it is more intuitive to talk about the portfolio value for a given scenario rather than about portfolio losses. To calibrate a discount factor we obtain the median CDS spread for a sample of European financial institutions ordered by ratings.[26] Figure 10 illustrates that the worse the ratings the higher the CDS spread and that the spread required by the market has increased considerably since 2008.

[INSERT FIGURE 10 AROUND HERE]

We have linearly extended the CDS values for the remaining ratings according to their average default probability and obtained the daily series of the median CDS spread level for each rating grade for the period 2008-2011. We assume that this is a representative spread to obtain a discount factor for the different ratings. However this spread assumes a *LGD* of 60% for bonds while we have an average *LGD* value of 18.35% x 1.378 = 25.28% over assets.[27] Hence we adjust linearly the spread. We assume an average maturity of 3 years for the assets in the portfolio; this is a mixture of the retail banking assets with longer maturity (like mortgages) and the corporate banking assets with shorter maturity. The average maturity of the assets is a key assumption in the model, the greater the maturity the higher the chance of high losses. Unluckily this information is not public for banks. Table 5 reports the 3-year discount factors obtained for each rating in this way.

[INSERT TABLE 5 AROUND HERE]

[26] These data correspond to 5-year senior CDS since 2008 and were obtained from Markit.

[27] As the financial institutions with available data in Markit have a high level of assets, it is quite possible that the *LGDs* of these entities will be smaller than 25.28% but this is a conservative assumption.

To simulate the rating transitions, we use an average rating transition matrix over the business cycle. We adjust the S&P public data in S&P (2010) to take into account the non-rated companies and we do impose the average probability of default previously adjusted. Table 6 includes the rating transition matrix employed.

[INSERT TABLE 6 AROUND HERE]

6.2.1 Migration rule

For a default mode model, the default probability of the client j conditional to a given macroeconomic scenario is

$$PD_{j,Z} = \Phi\left(\frac{\Phi^{-1}(PD_{j,C}) - \sum_{f=1}^{k} \alpha_{f,j} z_f}{\sqrt{1 - \sum_{f=1}^{k} \alpha_{f,j}^2}}\right)$$

This means that, to simulate the defaults, we can generate a random number $U_j \sim U(0,1)$ and the client defaults if $U_j \leq PD_{j,Z}$.

In the case of a market mode model a client can move from an initial rating to a new one. Let $MP_{j,C,IR,FR}$ denote the average probability (over the cycle) for the client j of migrating from an initial rating IR to FR, a final one. We can construct the accumulated probabilities $AccumMP_{j,C,IR,FR}$.[28]

Then, for a given macroeconomic state, we can calculate the point in time accumulated probability of migration between ratings, $AccumMP_{j,Z,IR,FR}$, as

$$AccumMP_{j,Z,IR,FR} = \Phi\left(\frac{\Phi^{-1}(AccumMP_{j,C,IR,FR}) - \sum_{f=1}^{k} \alpha_{f,j} z_f}{\sqrt{1 - \sum_{f=1}^{k} \alpha_{f,j}^2}}\right)$$

We generate a random number $U_j \sim U(0,1)$. Now, if $U_j \leq AccumMP_{j,Z,IR,D}$, the new rating of the client would be D. If $AccumMP_{j,Z,IR,CCC} \leq U_j \leq MP_{j,Z,IR,D}$, the new rating would be CCC and so on. For each possible final rating state the whole portfolio is evaluated.

6.2.2 Importance sampling

The IS framework must be modified in two ways: a) the exponential twisting rule should be extended to deal with more than two possible states and

[28] For example, $AccumMP_{j,C,IR,B-} = MP_{j,C,IR,B-} + MP_{j,C,IR,CCC} + MP_{j,C,IR,D}$.

b) the conditional portfolio value must be approximated to estimate the macroeconomic factor mean shift.

Given a macroeconomic scenario Z, the exponential twist of the migration probabilities MP of the client j from the rating state IR to FR can be extended as follows

$$MP_{j,Z,IR,FR,\theta} = \frac{MP_{j,Z,IR,FR}e^{V_{j,FR}\theta}}{\sum_{i=1}^{k} MP_{j,Z,IR,i}e^{V_{j,i}\theta}}$$

where $V_{j,i}$ is the loan value to the counterparty j given the rating state i, that is, $EAD_j \times DF_i$ where DF_i is the discount factor in the state i. Now, the natural extension of the default mode twist to the case of the mark to market valuation is

$$V_l = \sum_{j=1}^{M}\sum_{h=1}^{k} V_h \frac{MP_{j,Z,IR,h}e^{V_{j,h}\theta^*}}{\sum_{i=1}^{k} MP_{j,Z,IR,i}e^{V_{j,i}\theta^*}}$$

that is, the expected value of the portfolio equates the target value.

We use the normal approximation to change the mean of the factors. Under this approximation, conditional to the macroeconomic state Z, the portfolio value is distributed as $N(\mu_Z, \sigma_Z)$ with

$$\mu_Z = \sum_{j=1}^{M}\sum_{h=1}^{k} V_{j,h}MP_{j,Z,IR,h}, \quad \sigma_Z = \sqrt{\sum_{j=1}^{M}\sum_{h=1}^{k} V_{j,h}^2 MP_{j,Z,IR,h} - (\mu_Z)^2}$$

According to the ratings, the market value of the Spanish financial system is 2,842,499 MM €, representing a 2.7% discount with respect to the total assets. Applying the discounting factors to the migration probabilities, we get that the expected value of the portfolio is 2,839,535 MM €. Under a default mode model we focused on 4,528 MM € losses (ten times the expected loss) and, then, the equivalent market value is equal to 2,842,499 - 4,528 = 2,837,971 MM €. We will use this number as the target value for the IS method.

We will focus on value losses compared with the current market value rather than with total assets. The idea is that the difference between total assets and the current market value has been previously recognized through profit and losses statement and, hence, it does not represent a possible future loss. It means that debt holders and depositors should be concerned about the possible losses over the current market value and the amount of own resources that the institution has.

Figure 11 shows the loss distribution of the portfolio. For each simulation, losses are obtained as the market value minus the starting market value, 2,842,499 MM €. The 99.9% probability loss is 68,852 MM €, additional to the current market value loss, equal to 79,006 MM €. As the simulation speed is very sensitive to the number of possible states, it is very important to use only clearly different ratings.[29]

[INSERT FIGURE 11 AROUND HERE]

Regarding the *VaR* and *ES* based contributions we will allocate the 68,852 MM € loss over the current market value. Figure 12 provides the results and shows that the top contributor is Santander.

[INSERT FIGURE 12 AROUND HERE]

7 Parameter variability

The previous sections have analyzed the credit loss distribution of the Spanish financial system at December 2010 by considering different credit risk models. We study now the variability of the main parameters of the Vasicek (1987) model, namely, the *EAD*, *PD*, *LGD*, and the macroeconomic sensitivity α. We start analyzing the business cycle variability obtaining the loss distribution of the Spanish financial system at December 2007, a pre-crisis period. Later, we will study the impact of the variability of the risk parameters on the loss distribution by performing a sensitivity analysis.

7.1 Pre-crisis analysis

We estimate the loss distribution at December 2007 to asses the variability of the credit risk measures over the business cycle. The results can be different from those for December 2010 because of four possible reasons:

i) The ratings of the financial institutions may be different, therefore their *PD* may have changed.

ii) Many mergers took place after 2007, therefore the portfolio at December 2007 is more granular.

iii) The amount of assets of the institutions in the portfolio is different, as a consequence their *EAD* is different.

[29] The analysis has been performed using the rating scale considering modifiers but it could be done without these modifiers.

iv) As the LGD is assigned using asset buckets and the assets may have changed, the LGD may have also changed.

In the case of the portfolio at December 2007 the size of the institutions was similar to that at December 2010. However the ratings changed quite a lot between both dates, being this the main driver of the change in the loss distribution, followed by the change in the granularity of the portfolio.

Figure 13 includes the loss distribution of the Spanish financial portfolio at December 2007 under the default mode valuation and constant LGD. Compared with the loss distribution at December 2010, the loss distribution is shifted to the left assigning less probability to higher losses. This is mainly because the ratings deteriorated in the crisis period. In this case the 99.9% probability losses are only 13,995 MM €, a 44% of the estimate for December 2010. It can also be seen that even though many mergers had not taken place by December 2007 the loss distribution still has some discontinuities due to the presence of very big institutions.

[INSERT FIGURE 13 AROUND HERE]

Regulators should be aware of this kind of risk measurement variability if they want to use this type of models to quantify the risk of the financial system and to require a financial institution to have enough capital to make it safe. As suggested in Repullo et al. (2010), one way to deal with this issue can be to set a variable confidence level for capital requirements so that in periods with "high" ratings they can focus on more extreme probabilities while they may reduce the confidence levels in periods with "low" ratings.

7.2 Parameter uncertainty

The variability of the risk parameters can be related to the business cycle but also to some uncertainty in their estimates. The main reason for this uncertainty is that financial institutions do not default frequently and, then, the estimates of the risk parameters may not be very accurate. In this section we study the effects of this uncertainty on the loss distribution. Our analysis is based on several alternatives proposed in the literature for the three most important risk parameters in the default mode and constant LGD model.

7.2.1 α uncertainty

Our previous results were based on the functional form of the parameter α proposed by the Basel committee. According to this, α_{BIS} varied between

38% and 54% depending on the *PD* of the counterparty. Few studies analyze possible values of α for financial companies. Most of these studies come from the Moody's corporation as they have a commercial software[30] to implement the Vasicek (1987) model. Examples of these studies are López (2004), Lee et al. (2009), Qibin et al. (2009), and Castro (2012).

López (2004) estimated α for general corporations ordered by *PD* and size buckets while Lee et al. (2009) estimated this parameter differentiating by financial-industrial sector, *PD* buckets, and size buckets. Qibin et al. (2009) obtained quarterly estimates for α and their percentiles considering several companies grouped by sector (financial vs. industrial) and by geography (Europe-USA). Finally, Castro (2012) estimated a mean value of α considering three different models. Table 7 shows the values of α from these papers and illustrates that the Basel Committee estimates[31] are close or lower than the results in all these papers but for two of the models in Castro (2012).

[INSERT TABLE 7 AROUND HERE]

Figure 14 reports the loss distributions obtained for the different macroeconomic sensitivity parameters for the portfolio of Spanish financial institutions at December 2010.

[INSERT FIGURE 14 AROUND HERE]

The 99.9% probability losses range between 29,674 MM € and 45,305 MM €. However the latest value is obtained under the 90% confidence level for the α estimate which is a very conservative assumption.

7.2.2 *PD* uncertainty

This uncertainty arises mainly because clients with high rating usually do not default. As with the macroeconomic sensitivity parameter we test a set of possible rating-*PD* calibrations and see the impact on the loss distribution. Table 8 shows the average historical default rates from S&P for several periods that start in 1981 and finish in 2007 or subsequent years up to 2012.[32] As it can be seen firms graded with the two highest ratings never default.

[30] This software is currently called RiskFrontier and it was previously known as KMV.

[31] These estimates will tend to be closer to 54% rather than to 38% because of the low *PD* in banks.

[32] Data at rating modifier level is not available for the periods 1981-2007 and 1981-2008.

[INSERT TABLE 8 AROUND HERE]

Few papers estimate *PDs* for rating grades and measure the uncertainty in the estimates, mainly due to the absence of public information. One alternative is to obtain confidence intervals using the observed defaults and the total population of firms ordered by rating grades for a long enough period.[33] However the yearly number of defaulted companies is not publicly available. Hanson and Schuermann (2006) estimated average default rates by rating grade for the period 1981-2002 using S&P data and analytical as well as parametric and non-parametric bootstrapping techniques to find the standard deviations and the corresponding confidence intervals of the *PD* estimates. Cantor et al. (2007) take a similar approach for the period 1970-2006 and Moody's internal data. Table 9 shows the ratio between the standard deviation of the estimated average *PD* and the estimated *PD* from both papers. As Cantor et al. (2007) uses a longer period the uncertainty of the estimates should be lower, however this is not always the case. There are two possible reasons for this: i) the default database is different and ii) the estimates uncertainty does not only depend on the number of observations but also on the estimated average *PD* level.

[INSERT TABLE 9 AROUND HERE]

As we are using a slightly different calibration period we prefer to keep our average *PD* estimates and apply the most conservative ratio in Table 9 to our *PD* estimates. To keep it simple we assume a normal distribution of the average estimates and a 95% confidence interval to stress our *PD* estimates. This approach imposes that all the estimates must be inside their 95% confidence interval at the same time. According to this methodology our 95% confidence level for the AAA estimate is greater than the 95% confidence level for the AA, therefore we bounded the *PDs* by that of the next rating. Figure 15 reports the loss distribution of the portfolio under this approach.

[INSERT FIGURE 15 AROUND HERE]

It can be seen that, under the *PD* uncertainty and with a 95% confidence level, the 99.9% probability losses are 36,021 MM €, a 12% higher than the initial estimate.

[33] These confidence intervals can be obtained analytically or numerically, for instance, using a binomial distribution or a bootstrapping technique.

7.2.3 *LGD* uncertainty

Regarding estimates of the *LGD* there is some information about bond *LGDs* in financial institutions (see Altman and Kishore (1996)) but few papers estimate the *LGD* on total assets of defaulted financial firms. James (1991) provided a first estimate of average losses on assets of 30.51% using data of US defaulted financial institutions over the period 1985-1988. This number is much higher than that used by us mainly because a) it is a point-in-time *LGD* estimate and b) most of the defaults in the sample were due to small institutions with higher *LGD*. Therefore shifting the mean estimates.

More recent papers have focused on the losses for the depositors or the deposits guarantee fund. Kuritzkes et al. (2002) analyzed the solvency of the FDIC and based their results on the historical losses suffered by the FDIC estimated in Bennett (2002) for the period 1986-1998. Kaufman (2004) gets similar results to those in Bennett (2002) but for the period 1980-2002.

Bennett (2002) provided a very detailed analysis of the total losses on total assets and for the FDIC due to bank failures but they do not perform an analysis by asset buckets. Then we decided to update the results of Bennett (2002) for losses to depositors and apply the ratio of losses to depositors to losses on total assets from Bennett and Unal (2011).

Two reasons can explain the statistical uncertainty in the *LGDs* estimates: i) the ratio of losses to depositors to losses on total assets may change over the asset buckets[34] and ii) the number of defaults is very low in the highest assets bucket; this may affect the average *LGD* estimate if the real *LGD* is not constant, as it is the case in the observed data. Regarding this issue we can test the effect of the *LGD* uncertainty using the 95% confidence level of the estimated *LGD*.[35] Figure 16 includes the loss distribution of the portfolio after considering these two sources of uncertainty.

[INSERT FIGURE 16 AROUND HERE]

Under the *LGD* uncertainty the 99.9% probability losses are 50,804 MM € with a 95% confidence level. The effect of the *LGD* uncertainty is much higher than that of α or *PD*. This is because the *LGD* has a linear impact on the portfolio losses and the uncertainty in the *LGD* estimates of the biggest institutions is very high due to the lack of historical defaults.

[34]Table 2 in Bennett and Unal (2011) shows that this ratio can be up to 1.47.

[35]If we have R defaults the LGD estimate is normally distributed with mean $\mu = \sum_{i=1}^{R} LGD_i/R$ and variance $\sum_{i=1}^{R} (LGD_i - \mu)^2/R$.

8 Conclusions

This paper has successfully extended the IS framework introduced by Glasserman and Li (2005) to the case of random recoveries and market mode models. We also tested the extensions of granular portfolios, simulation loop decoupling, and mean based macroeconomic factors shift.

Considering the *LGD* as a constant is an assumption that is not supported by the historical data, therefore this extension allows us to better capture the real behavior of the defaults. A similar conclusion can be drawn from the market mode valuation: real portfolios can be exposed to mark to market losses derived from rating changes. The simulation loop decoupling and mean based macroeconomic factors shift extensions allow for a faster and more accurate risk measurement. The loop decoupling is very interesting under unifactorial but not granular portfolios as it reduces the number of calculations required. On the other hand, mean based macroeconomic factors shift enables a better sampling process and therefore it also reduces the number of simulations required to obtain narrow confidence intervals of the estimates.

All these extensions allow to use this method inside financial institutions or for regulatory purposes. The extensions and modifications have been tested on a portfolio including Spanish financial institutions using plain Monte Carlo simulations as benchmark. Based on Bennett (2002), the *LGD* of the different institutions has been obtained and used to estimate the loss distribution of this financial system.

According to our results the 99.9% probability losses can range between 30,000 and 70,000 MM € depending on the *LGD* model and the valuation method employed. However, under a granular portfolio with constant *LGD*, the 99.9% probability losses would be only 13,478 MM €. The confidence intervals of the loss distribution obtained using the IS approach are very thin regardless of the *LGD* model or the valuation method used.

The confidence intervals of the risk allocation obtained using IS are much thinner than those obtained with the plain Monte Carlo method, specially for the *VaR* based risk allocation. In general, the risk allocation based on the *VaR* has wider confidence intervals than that based on the *ES*. More precisely, under constant *LGD*, the *VaR* based risk allocation has thin confidence intervals and requires a low number of simulations. However, as we move to a random *LGD* framework, the number of simulations required to obtain small confidence intervals in the risk allocation increases considerably. Hence, one possible way to deal with this issue is to use the IS method to estimate the risk allocation in the case of constant *LGD* and try to extend other methods such as those in Pykhtin (2004), Huang et al. (2007), and Voropaev (2011) to deal with the random *LGD* risk allocation.

Analyzing the suitability of the allocation criteria, we have found that the results can vary considerably. Probably the best approach is to obtain all the possible results and compare them. For example, under the CVaR, a given client may have a null risk allocation (as happened with BBVA and Santander in the constant *LGD* model) and, hence, provide a infinite risk adjusted return, but this would lead to a higher concentration.

Finally we have studied the variability of the estimates over the business cycle and the variability due to the uncertainty in the model parameters estimates. We have shown that the risk estimates can vary considerably over the business cycle. Regarding the parameters uncertainty we have shown that currently the main driver of uncertainty in the risk estimates is the *LGD*. This is due to the low number of historical defaults for the biggest financial institutions bucket.

This kind of analysis can provide a basic tool for regulators to analyze the solvency of the financial system and to study the relevance of the financial institutions in the economy. This last issue is specially interesting to establish the so called systematically important financial institutions surcharge in BIS III.

References

Acharya, V., Pedersen, L., Philippon, T., Richardson, M., 2010. Measuring systemic risk. Working Paper 1002, Federal Reserve Bank of Cleveland.

Adrian, T., Brunnermeier, M.K., 2010. CoVaR. Staff Report No. 348, Federal Reserve Bank of New York.

Allen, F., Bartiloro, L., Kowalewski, O., 2006. The financial system of the EU 25. In: Liebscher, K., Christl, J., Mooslechner, P. and Ritzberger-Grünwald D. (Eds.), Financial Development, Integration and Stability: Evidence from Central, Eastern and South-Eastern Europe, Edward Elgar, Cheltenham (chapter 7).

Altman, E.I., Brady, B., Resti, A., Sironi, A., 2005. The link between default and recovery rates: Theory, empirical evidence and implications. Journal of Business, 78 (6), 2203-2227.

Altman, E.I., Kishore, V.M., 1996. Almost everything you wanted to know about recoveries on defaulted bonds. Financial Analysts Journal, 52 (6), 57-64.

Altman, E.I, Suggitt, H., 2000. Default rates in the syndicated bank loan market: A mortality analysis. Journal of Banking and Finance, 24 (1-2), 229-253.

Amihud, Y., Garbade, K., Kahan, M., 2000. An institutional innovation to reduce agency costs of public corporate bonds. Journal of Applied Corporate Finance, 13 (1), 114-121.

Basel Committee on Bank Supervision, 2006. Basel II: International convergence of capital measurement and capital standards: A revised framework - Comprehensive Version. Available at http://www.bis.org.

Basel Committee on Bank Supervision, 2011. Global systemically important banks: Assessment methodology and the additional loss absorbency requirement. Available at http://www.bis.org.

BBVA Bank, 2009. 2008 Annual Report. Available at http://www.bbva.com.

Bennett, R.L., 2002. Evaluating the adequacy of the deposits insurance fund: A credit-risk modeling approach. Working Paper No. 2001-02, FDIC.

Bennett, R.L., Unal, H., 2011. The cost effectiveness of the private sector reorganization of failed banks. Working Paper No. 2009-11, FDIC.

Brownlees, C, Engle, R., 2012. Volatility, correlation and tails for systemic risk measurement (October 15, 2012). Available at http://ssrn.com/abstract=1611229.

Campos, P., Yagüe, M., Chinchetru, I., 2007. Un nuevo marco de seguro de depósitos para España. Estabilidad Financiera, 12, 93-110.

Cantor, R., Hamilton, D., Tennant, J., 2007. Confidence intervals for corporate default rates. Moody's Special Comment - Moody's Investors Service.

Carbó, S., Humphrey, D., Maudos, J., Molyneux, P., 2009. Cross-country comparisons of competition and pricing power in European banking. Journal of International Money and Finance, 28 (1), 115-134.

Carey, M., 1998. Credit risk in private debt portfolios. Journal of Finance, 53 (4), 1363-1387.

Cariboni, J., Maccaferri, S., Schoutens, W., 2013. Lévy processes and the financial crisis: can we design a more effective deposit protection? International Journal of Financial Research, 4 (1), 5-28.

Castro, C., 2012. Confidence sets for asset correlations in portfolio credit risk. Revista de Economia del Rosario, 15 (1), 19-58.

Dullmann, K., Kull, J., Kunisch M., 2010. Estimating asset correlations from stock prices or default rates which method is superior? Journal of Economic Dynamics and Control, 34 (11), 2341-2357.

Fitch, 2009. Fitch ratings global corporate finance 2008 transition and default study. Available at http://www.fitchratings.com under Credit Market Research.

Giesecke, K., Kim, B., 2011. Systemic risk: what defaults are telling us. Management Science, 57 (8), 1387-1405.

Glasserman, P., 2005. Measuring marginal risk contributions in credit portfolios. Journal of Computational Finance, 9 (2), 1-41.

Glasserman, P., Li, J., 2005. Importance sampling for portfolio credit risk. Management Science, 51 (11), 1643-1653.

Gordy, M., 2000. A comparative anatomy of credit risk models. Journal of Banking and Finance, 24 (1-2), 119-149.

Grundke, P., 2009. Importance sampling for integrated market and credit portfolio models. European Journal of Operational Research, 194 (1), 206-226.

Gupton, G.M., Stein, R.M., 2002. LossCalc^{TM}: Moody's model for predicting loss given default (LGD). Moody's KMV Company.

Hanson, S.G., Schuermann, T., 2006. Confidence intervals for probabilities of default. Journal of Banking and Finance, 30 (8), 2281-2301.

Hartmann, P., Ferrando, A., Fritzer, F., Heider, F., Lauro, B., Lo Duca, M., 2006. The performance of the European financial system, mimeo, European Central Bank, August.

Huang, X., Oosterlee, C.W., Van Der Weide, H., 2007. Higher order saddlepoint approximations on the Vasiceck portfolio credit loss model. Journal of Computational Finance, 11 (1), 93-113.

Huang, X., Zhou, H., Zhu, H., 2009. A framework for assessing the systemic risk of major financial institutions. Journal of Banking and Finance, 33 (11), 2036-2049.

James, C., 1991. The loss realized in banks failures. Journal of Finance, 46 (4), 1223-1242.

Kaufman, G., 2004. FDIC Losses in bank failures: Has FDICIA made a difference? Economic Perspectives (Federal Reserve Bank of Chicago), 28 (3), 13-25.

Kuritzkes, A., Schuermann, T., Weiner, S., 2002. Deposit insurance and risk management of the U.S. banking system: How much? how safe? who pays? Working Paper 02-02, Wharton Financial Institutions Center.

Lee, J., Wang, J., Zhang, J., 2009. The relationship between average asset correlation and default probability. Moody's KMV Company.

López, J., 2004. The empirical relationship between average asset correlation, firm probability of default and asset size. Journal of Financial Intermediation, 13 (2), 265-283.

Moody's, 2005. Default, losses and rating transitions on bonds issued by financial institutions: 1983-2004. Moody's Special Comment - Moody's Investors Service.

Moody's, 2009. Corporate default and recovery rates 1920-2008. Moody's Special Comment - Moody's Investors Service.

Pykhtin, M., 2004. Multi-factor adjustment. Risk Magazine, 17 (3), 85-90.

Qibin, C., Amnon, L., Nihil, P., 2009. Understanding asset correlation dynamics from stress testing. Moody's KMV Company, available at http://www.moodysanalytics.com/Insight/Quantitative-Research/Portfolio-Modeling.

Reitan, T., Aas, K., 2010. A new robust importance-sampling method for measuring value-at-risk and expected shortfall allocations for credit portfolios. Journal of Credit Risk, 6 (4), 113-149.

Repullo, R., Saurina, J., Trucharte, C., 2010. Mitigating the pro-cyclicality of Basel II. Economic Policy, 25 (64), 659-702.

Santander Bank, 2009. 2008 Annual Report. Available at http://www.santander.com.

Schuermann, T., 2004. What do we know about loss given default? In: Shimko, D. (Ed.), Credit Risk: Models and Management, second ed. Risk Books, London, UK (chapter 9).

Serfling, R.J., 1980. Approximation theorems of mathematical statistics. John Wiley & Sons.

S&P, 2009. 2008 Annual global corporate default study and rating transitions. S&P Corporation Ratings Direct.

S&P, 2010. 2009 Annual global corporate default study and rating transitions. S&P Corporation Ratings Direct.

Thorburn, K., 2000. Bankruptcy auctions: costs, debt recovery and firm survival. Journal of Financial Economics, 58 (3), 337-368.

Unal, H., Madan, D., Guntay, L., 2003. Pricing the risk of recovery in default with APR violations. Journal of Banking and Finance, 27 (6), 1001-1025.

Vasicek, O.A., 1987. Probability of loss on loan portfolio. Working Paper, KMV Corporation.

Voropaev, M., 2011. An analytical framework for credit portfolio risk measures. Risk Magazine, 24 (5), 72-78.

Appendix

Multimodal distributions

We studied the behavior of the importance sampling algorithm on a symmetric portfolio made up of clients with the same PD_i, LGD_i, EAD_i, and macroeconomic factor sensitivity. These clients were split in two halves that are sensitive to two different macroeconomic factors. Surprisingly, we obtained a mode of the optimum sampling distribution that was not symmetric although the problem was completely symmetric.

To understand this issue, Figure 17 provides the function $g(Z) = Prob(L \geq l|Z)e^{-\frac{Z'Z}{2}}$ for a simple case.[36] Two modes can clearly be seen. These modes have a direct impact on the estimated risk and on the risk contributions estimates and generate a bias. The intuition is that using one of the modes will simulate normal macroeconomic factors with a mean that is very close to zero for one of the factors and positive for the other one. Then, most of the simulations will generate large losses on half of the portfolio and almost no losses on the other half. Hence, the importance sampling algorithm will generate two effects:

- The loss distribution will be underestimated because only half of the portfolio defaults in the simulations and the estimation confidence intervals will be very big.
- Even though the portfolio is symmetric, the half of the portfolio that does not default on the simulations will have very low risk contributions.

[INSERT FIGURE 17 AROUND HERE]

This bimodal characteristic should generate large confidence intervals for the estimates. However, this is only the case if the two modes are not close. To our knowledge, this is the first time that this bias has been detected in the literature.

Glasserman and Li (2005) proposed using the mode of the optimun sampling distribution to change the mean of the macroeconomic factors because it is easier to be estimated than other statistical moments as the mean or the median.[37] Trying to solve this problem, we decided to estimate the mean or

[36] We use the normal approximation, 1,000 counterparties, parameters *PD*=1%, *LGD*=40%, *EAD*=1,000, $\alpha = 55\%$, and target loss equal to 10 times the expected loss.

[37] It is easier to find numerically the maximum of a multivariate distribution than obtaining random samples from it.

the median of the optimum distribution $g(Z)$. Reitan and Aas (2010) proposed estimating the mean using Markov Chain Monte Carlo (MCMC)[38] and the Metropolis-Hasting algorithm, a very suitable procedure as it does not require having a proper density function that integrates one as is our case.

However the MCMC method is not very fast and we propose a method based again on importance sampling to estimate the mean and variance of $g(Z)$. We will sample from a normal distribution and then use weights to estimate these two moments. In more detail, these estimators are[39]

$$\begin{aligned}\mu_{g(Z)} &= \frac{1}{N}\sum_{i=1}^{N} Z_j cProb(L_i > l|Z_i)e^{-\frac{Z_i Z_i'}{2}}\frac{1}{\phi(Z_i)_{\mu,\Omega}} \\ \sigma^2_{g(Z)} &= \frac{1}{N}\sum_{i=1}^{N} Z_i^2 cProb(L_i > l|Z_i)e^{-\frac{Z_i Z_i'}{2}}\frac{1}{\phi(Z_i)_{\mu,\Omega}} - \mu^2_{g(Z)}\end{aligned}$$

where Z_i is obtained from a multivariate normal random variable $\phi(Z)_{\mu,\Omega}$. After many trials, the best results are obtained when the parameter μ is set to zero and the variance matrix Ω is the identity one.

The constant c ensures that we are working with a probability distribution, that is,

$$1 = \frac{1}{N}\sum_{i=1}^{N} cProb(L_i > l|Z_i)e^{-\frac{Z_i Z_i'}{2}}\frac{1}{\phi(Z_i)_{\mu,\Omega}}$$

This method is much faster than the MCMC and, at the same time, generates accurate results. Alternatively to the mean, the median of the optimum $g(Z)$ can also be used. This median is estimated for every macroeconomic factor of the set $Z = \{z_1, \cdots, z_k\}$. In the case of the component k, the median is obtained ordering the simulations 1 to N according to the values of z_k and then adding the weight $\frac{1}{N}cProb(L_i > l|Z_i)e^{-\frac{Z_i Z_i'}{2}}\frac{1}{\phi(Z_i)_{\mu,\Omega}}$ until the value 50% is obtained, that is,

$$\text{median}_{g(Z_i)} = \min\left(z_{i,n} \;\middle|\; \frac{1}{N}\sum_{i=1}^{n}\left[cProb(L_j > l|Z_i)e^{-\frac{Z_i Z_i'}{2}}\frac{1}{\phi(Z_i)_{\mu,\sigma}}\right] = 50\%\right)$$

[38] However, they did not highlight any bias related to the use of the mode.

[39] After all the experiments, the best results where obtained using as variance the maximum between 1 and the optimum variance.

Loop decoupling

The importance sampling (IS) framework explained in Glasserman and Li (2005) assumed that, for every macroeconomic factor simulation, an optimal exponential twist is calculated and one default simulation is performed. However, we can also generate several default simulations for every macroeconomic factor simulation. This is interesting when dealing with almost unifactorial portfolios that are not granular as the number of required optimizations gets reduced.

Let N_e and N_i denote the number of macroeconomic scenarios and default simulations conditional to a macroeconomic scenario, respectively. Then, $N = N_e N_i$. Again the confidence intervals for the estimations can be estimated as explained before but, now, the defaults are not totally independent as some of them share macroeconomic scenarios. Hence, the confidence interval formulas must be slightly modified:

$$\begin{aligned} Var(\widehat{Prob(L \geq l)}) &= \frac{1}{N^2}\sum_{i=1}^{N_e} Var\left(\sum_{k=1}^{N_i} \mathbf{1}(L_{i,k} \geq l)\frac{f(L_{i,k})}{g(L_{i,k})}\right) \\ &= \frac{N_e}{N^2} Var\left(\sum_{k=1}^{N_i} \mathbf{1}(L_{i,k} \geq l)\frac{f(L_{i,k})}{g(L_{i,k})}\right) \\ &= \frac{N_e}{N^2} Var(R_i) \approx \frac{N_e}{N^2}\left(\frac{1}{N_e}\sum_{i=1}^{N_e} R_i^2 - \left(\frac{1}{N_e}\sum_{i=1}^{N_e} R_i\right)^2\right) \end{aligned}$$

where $L_{i,k}$ stands for the loss on the external simulation i and the internal simulation k.

A similar result can be obtained for the expected shortfall (ES)

$$\begin{aligned} X_{n1} &= \frac{1}{N}\sum_{i=1}^{N_e}\sum_{k=1}^{N_i} L_{i,k}\mathbf{1}(L_{i,k} \geq l)\frac{f(L_{i,k})}{g(L_{i,k})} = \frac{1}{N}\sum_{i=1}^{N_e} S_i \\ X_{n2} &= \frac{1}{N}\sum_{i=1}^{N_e}\sum_{k=1}^{N_i} \mathbf{1}(L_{i,k} \geq l)\frac{f(L_{i,k})}{g(L_{i,k})} = \frac{1}{N}\sum_{i=1}^{N_e} R_i \\ Var(\widehat{ES}) &= \frac{Var(X_{n1} - \widehat{ES}X_{n2})}{X_{n2}^2} = \frac{\frac{1}{N^2}\sum_{i=1}^{N_e} Var(S_i - \widehat{ES}R_i)}{X_{n2}^2} \\ &\approx N\frac{\sum_{i=1}^{N_e}(S_i - \widehat{ES}R_i)^2}{\sum_{i=1}^{N_e} R_i} \end{aligned} \tag{4}$$

The previous formula can be used to obtain the variance of the risk contributions. The variance of the expected shortfall contributions of client

j can be obtained replacing S_i in (4) by $S_{i,j}$

$$S_{i,j} = \sum_{k=1}^{N_i} x_{i,k,j} \mathbf{1}(L_{i,k} \geq l) \frac{f(L_{i,k})}{g(L_{i,k})}$$

The variance of the VaR contribution estimates of the client j is obtained changing $\widehat{ES}$ by $\widehat{VaR}$ in (4) and redefining S_i and R_i by $S_{i,j}$ and $R_{i,j}$, respectively

$$\begin{aligned} S_{i,j} &= \sum_{k=1}^{N_i} x_{i,k,j} \mathbf{1}(l(1-R) \leq L_{i,k} \leq l(1+R)) \frac{f(L_{i,k})}{g(L_{i,k})} \\ R_{i,j} &= \sum_{k=1}^{N_i} \mathbf{1}(l(1-R) \leq L_{i,k} \leq l(1+R)) \frac{f(L_{i,k})}{g(L_{i,k})} \end{aligned}$$

As it can be seen, the sums of the default simulations for every macroeconomic factor simulation have to be performed. This forces to keep all the default data for one macroeconomic simulation. However, this is a tractable problem as, in general, $1,000 \leq N_e \leq 10,000$ and $100 \leq N_i \leq 1,000$.

In the case of the financial institutions there is a big non-granular effect and, then, it may be interesting to decouple the simulation loops. In this case we generate the loss distribution using, for example, 1,000 x 100 simulations. This method generates very accurate results and, at the same time, is much faster than the general one on the case of very non-granular portfolios as happens in the Spanish financial system.

Appendix of Tables

Table 1: Spanish financial institutions involved in a merger / acquisition process or belonging to the same corporation at December, 2010.

New Entity	Original Institutions
Banca Civica	Caja Municipal de Burgos, Caja Navarra, Caja Canarias, CajaSol, Caja Guadalajara
Banco Base	Caja Asturias, Banco de Castilla La Mancha, Caja Cantabria, Caja Extremadura
Banco Mare Nostrum	Caja Murcia, Caixa Penedés, Caja Granada, Caja Sa Nostra
Banco Popular	Banco Popular, Banco Popular Hipotecario, Banco Popular-e, Popular banca privada
Bankia	Caja Madrid, Bancaja, Caixa Laietana, Caja Avila, Caja Segovia, Caja Rioja, Caja Insular
BBK	BBK, Cajasur
BBVA	BBVA, Finanzia, Banco Depositario BBVA, UNO-E Bank
Caixabank	La Caixa, Caixa Girona, Microbank
Caja 3	Caja Inmaculada, Caja Burgos CCO, Caja Badajoz
Caja España de Inversiones	Caja España, Caja Duero
Catalunya Caixa	Caixa Cataluña, Caixa Tarragona, Caixa Manresa
Novacaixagalicia	Caja Galicia, Caixanova
Santander	Banco Santander, Banesto, Santander Investment, Openbank, Banif, Santander Consumer Finance
Unicaja	Unicaja, Caja Jaén
Unnim	Caixa Sabadell, Caixa Terrassa, Caixa Manlleu

Table 2: *LGD* estimates for losses on deposits and losses on assets for the period 1986-2009 obtained from the FDIC public data by institution size.

Assets (in \$bn)	Count	Mean (deposits)	Mean (assets)
< 1	1148	18.61%	25.65%
1 - 5	49	15.50%	21.37%
5 - 15	7	9.95%	13.72%
> 15	8	6.39%	8.82%

Table 3: BBVA and Santander country exposures obtained according to the net interest income data published in their 2010 Annual Reports.

Country	BBVA	Santander
Spain	37.7%	18.8%
Mexico	33.5%	5.9%
United States	9.6%	6.8%
Argentina	2.8%	0%
Chile	4.0%	5.3%
Colombia	4.0%	0%
Peru	4.8%	0%
Venezuela, RB	3.7%	0%
Portugal	0%	2.6%
United Kingdom	0%	14.5%
Brazil	0%	36.8%
Italy	0%	0.7%
Finland	0%	0.8%
Germany	0%	7.5%
Total	100%	100%

Table 4: Comparison of the 99.9% probability loss levels under different random LGD models. We consider a pure macroeconomic LGD (LGD^C), based on transformations of a random normal macroeconomic variable z_{LGD}, the random LGD conditional to the macroecomic variable z_{LGD} (LGD^R), and the case of $LGD|z_{LGD}$ with Beta and Gamma distributions.

Model	Loss (MM €)	Model	Loss (MM €)
Normal LGD^C	37,160	Probit Normal LGD^C	35,999
Normal LGD^R	38,131	Probit Normal LGD^R	35,318
Lognormal LGD^C	29,309	Normal^2 LGD^C	36,826
Lognormal LGD^R	36,139	Normal^2 LGD^R	36,587
Logit Normal LGD^C	35,909	Beta LGD^R	37,616
Logit Normal LGD^R	34,997	Gamma LGD^R	37,578

Table 5: Discount factors by rating grade based on the average CDS spread and 3-year average maturity.

Investment Grade										
Rating	AAA	AA+	AA	AA-	A+	A	A-	BBB+	BBB	BBB-
Discount Factor	98.71%	98.69%	98.6%	98.37%	97.93%	97.14%	96.96%	96.63%	96.02%	95.94%
Speculative Grade										
Rating	BB+	BB	BB-	B+	B	B-	CCC/C			
Discount Factor	95.79%	95.51%	94.99%	94.04%	92.33%	89.26%	83.93%			

Table 6: Average 1-year rating migration matrix from S&P (2010).

	AAA	AA+	AA	AA-	A+	A	A-	BBB+	BBB	BBB-	BB+	BB	BB-	B+	B	B-	CCC	D
AAA	91%	4%	3%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.0249%
AA+	2%	79%	12%	4%	1%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.0253%
AA	1%	1%	84%	8%	3%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.027%
AA-	0%	0%	5%	80%	10%	3%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.03%
A+	0%	0%	1%	5%	81%	9%	3%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0.0356%
A	0%	0%	0%	1%	5%	81%	7%	3%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0.0459%
A-	0%	0%	0%	0%	1%	7%	79%	8%	3%	1%	0%	0%	0%	0%	0%	0%	0%	0.0651%
BBB+	0%	0%	0%	0%	0%	1%	7%	78%	9%	2%	0%	0%	0%	0%	0%	0%	0%	0.1005%
BBB	0%	0%	0%	0%	0%	1%	1%	7%	80%	6%	2%	1%	0%	0%	0%	0%	0%	0.1659%
BBB-	0%	0%	0%	0%	0%	0%	0%	2%	9%	76%	6%	3%	1%	1%	0%	0%	0%	0.2871%
BB+	0%	0%	0%	0%	0%	0%	0%	1%	2%	13%	69%	7%	4%	1%	1%	0%	1%	0.5112%
BB	0%	0%	0%	0%	0%	0%	0%	0%	1%	3%	9%	71%	9%	3%	2%	1%	1%	0.9257%
BB-	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	2%	9%	71%	9%	4%	1%	1%	1.6925%
B+	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	7%	73%	9%	3%	2%	3.111%
B	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	9%	66%	9%	7%	5.735%
B-	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	3%	10%	60%	14%	10.5892%
CCC/C	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	1%	2%	3%	11%	63%	19.5689%

Table 7: Comparison of the values of α in López (2004), Lee et al. (2009), Qibin et al. (2009), and Castro (2012). α_{Lo} denotes the estimate in López (2004) for the bucket of companies with the biggest assets, α_{M_2} indicates the value for the buckets of smallest PD and biggest assets α_{M_2} in Lee et al. (2009), and $\alpha_{M_1,50\%}$, $\alpha_{M_1,75\%}$, and $\alpha_{M_1,90\%}$ are the values for the percentiles 50%, 75% and 90% in Qibin et al. (2009). Finally, $\alpha_{C,i},\ i = 1, 2, 3$ denote the mean estimates for the three models tested in Castro (2012).

	López (2004)	Lee et al. (2009)	Qibin et al. (2009)			Castro (2012)		
Parameter	α_{Lo}	α_{M_2}	$\alpha_{M_1,50\%}$	$\alpha_{M_1,75\%}$	$\alpha_{M_1,90\%}$	$\alpha_{C,1}$	$\alpha_{C,2}$	$\alpha_{C,3}$
Value	[47%-57%]	59%	45%	59%	74%	69%	43%	66%

Table 8: Average S&P default rates for the periods 1981-2007, 1981-2008, 1981-2009, 1981-2010, 1981-2011, and 1981-2012.

Rating	2012	2011	2010	2009	2008	2007
AAA	0%	0%	0%	0%	0%	0%
AA+	0%	0%	0%	0%	0%	0%
AA	0.01%	0.01%	0.01%	0.02%	0.03%	0.01%
AA-	0.02%	0.02%	0.03%	0.03%	0.05%	0.02%
A+	0.05%	0.05%	0.05%	0.06%	0.07%	0.05%
A	0.06%	0.07%	0.07%	0.07%	0.08%	0.06%
A-	0.07%	0.07%	0.07%	0.08%	0.09%	0.07%
BBB+	0.15%	0.16%	0.16%	0.17%	0.15%	0.14%
BBB	0.24%	0.25%	0.26%	0.27%	0.24%	0.23%
BBB-	0.3%	0.3%	0.31%	0.32%	0.28%	0.27%
BB+	0.61%	0.63%	0.67%	0.66%	0.73%	0.73%
BB	0.83%	0.86%	0.88%	0.9%	0.99%	1%
BB-	1.4%	1.42%	1.47%	1.5%	1.65%	1.67%
B+	2.36%	2.41%	2.47%	2.55%	2.32%	2.35%
B	6.81%	6.98%	7.17%	7.37%	6.7%	6.79%
B-	9.6%	9.8%	9.99%	10.23%	9.3%	9.43%
CCC/C	23.53%	23.41%	23.56%	23.61%	25.67%	25.59%

Table 9: Ratio of standard deviation of the *PD* estimate and the *PD* estimate.

Rating	Cantor et al. (2007)	Schuermann (2004)
AAA	-	-
AA	71.19%	62.78%
A	38.27%	23.12%
BBB	14.24%	23.52%
BB	6.64%	9.87%
B	3.51%	4.37%
CCC	3.84%	4.36%

Appendix of Figures

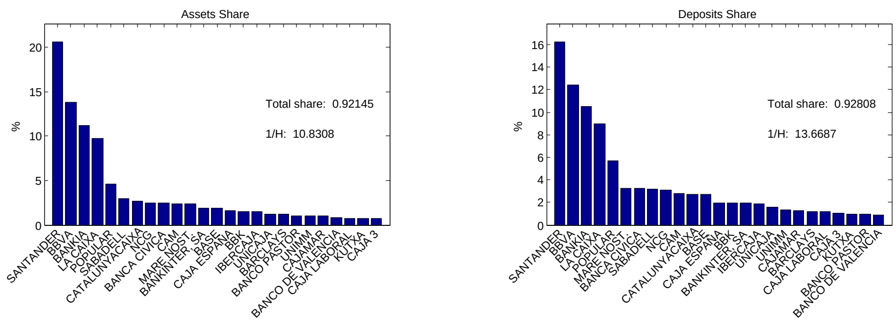

Figure 1: Assets and deposits share of the top twenty-five Spanish financial institutions.

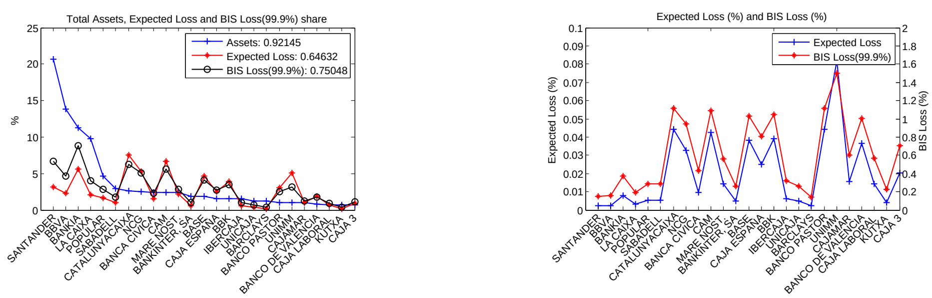

Figure 2: Assets, Expected Loss, and Basel 99.9% loss share of the top 25 Spanish financial institutions. Left and right graphs show, respectively, the amount allocation and the allocated amount relative to the institution size.

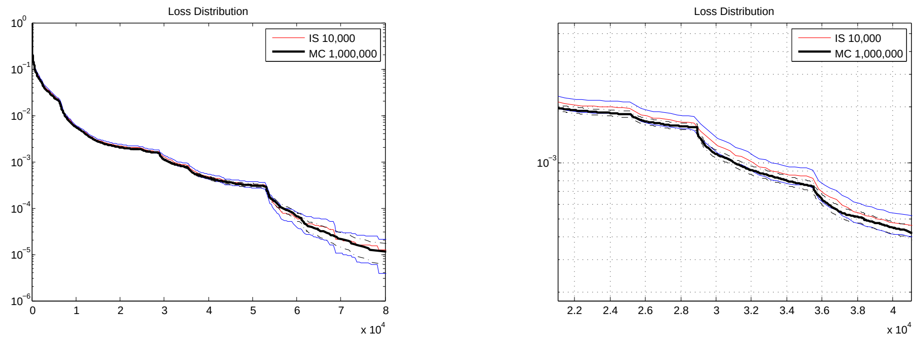

Figure 3: Loss distribution using 10,000 importance sampling (IS) and 1,000,000 plain Monte Carlo (MC) simulations. The black and red lines show, respectively, the plain Monte Carlo and IS results while the blue lines indicate the 5%-95% confidence interval of the IS estimates. Left and right graphs show, respectively, the tail distribution and its detail in the neighborhood of the 99.9% probability loss level.

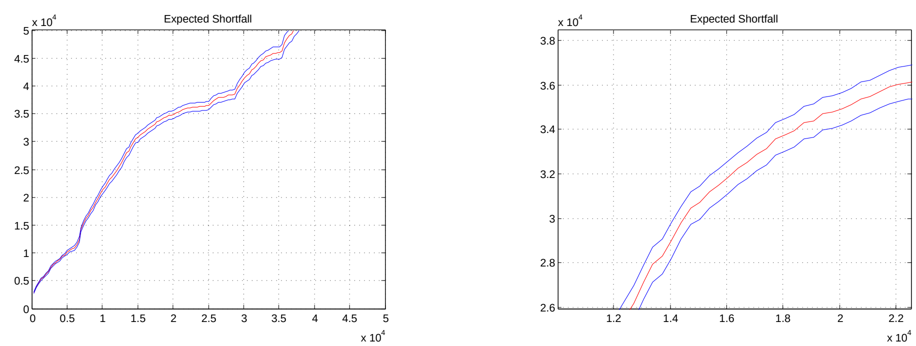

Figure 4: Expected Shortfall (*ES*) using 10,000 importance sampling (IS) simulations. The red and blue lines show, respectively, the IS results for the expected shortfall estimate and its 5%-95% confidence intervals.

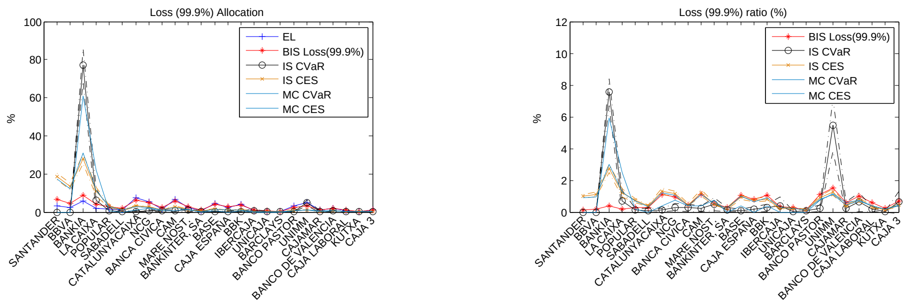

Figure 5: Risk allocation under constant LGD based on expected loss (EL), Basel loss 99.9% (BIS), contributions to VaR (CVaR) and ES (CES) both using importance sampling (IS) and plain Monte Carlo (MC) criteria. Left and right graphs show, respectively, the total risk allocation and the allocated risk relative to the institution size.

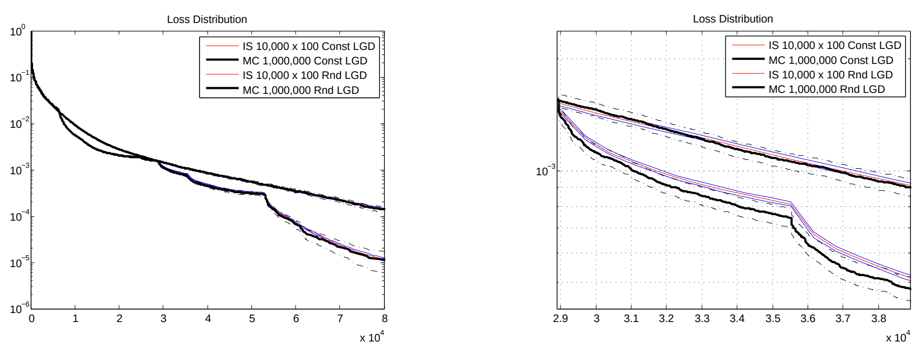

Figure 6: Comparison of the random LGD (Rnd LGD) and constant LGD (Const LGD) loss distributions. Black lines show the results of the plain Monte Carlo (MC) method using 1,000,000 simulations. The red and blue lines show, respectively, the importance sampling (IS) estimates and their 5%-95% confidence intervals using 10,000 macroeconomic scenarios and 100 default simulations on each macroeconomic scenario. Left and right graphs show, respectively, the tail distribution and its detail in the neighborhood of the 99.9% probability loss level.

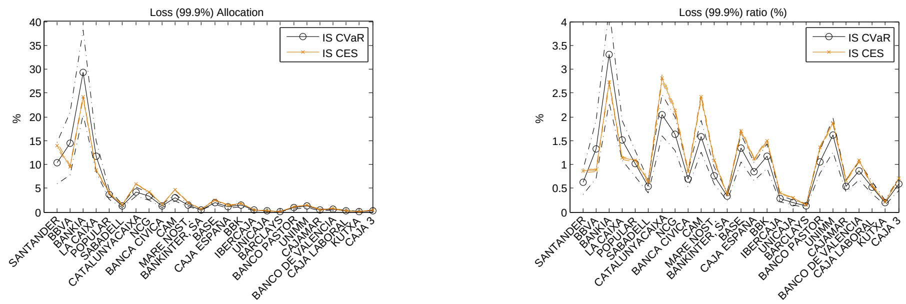

Figure 7: Risk allocation under macroeconomic random LGD (LGD^C) for the VaR (CVaR) and the ES (CES) criteria. Continuous and dashed lines represent, respectively, the IS estimates and the 5%-95% confidence intervals. Left and right graphs show, respectively, the total risk allocation and the allocated risk relative to the institution size.

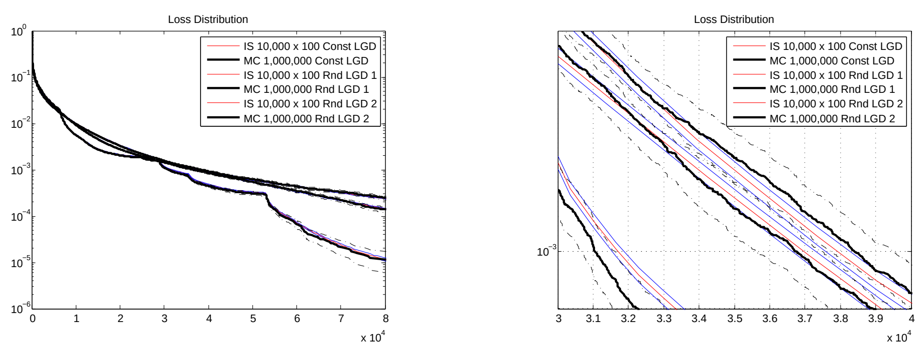

Figure 8: Comparison of the two random LGD models (Rnd LGD^C / Rnd LGD^R) and constant LGD (Const LGD) loss distributions. Black lines show the results of the plain Monte Carlo (MC) method using 1,000,000 simulations. The red and blue lines show, respectively, the importance sampling (IS) estimates and their 5%-95% confidence intervals using 10,000 macroeconomic scenarios and 100 default simulations on each macroeconomic scenario. Left and right graphs show, respectively, the tail distribution and its detail in the neighborhood of the 99.9% probability loss level.

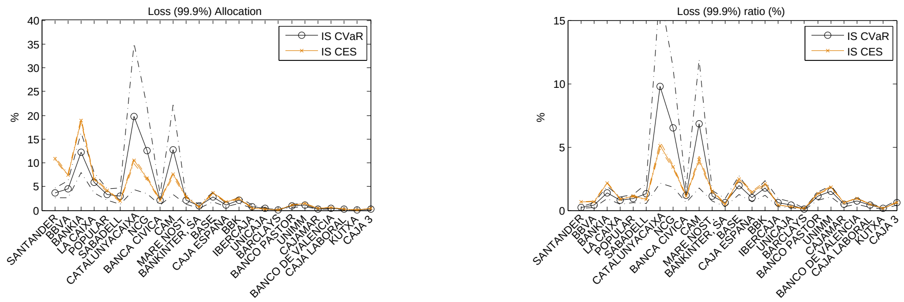

Figure 9: Risk allocation under mixed macroeconomic and idiosyncratic random LGD (LGD^R) for the VaR (CVaR) and ES (CES) criteria. Continuous and dashed lines represent, respectively, the IS estimates and the 5%-95% confidence intervals. Left and right graphs show, respectively, the total risk allocation and the allocated risk relative to the institution size.

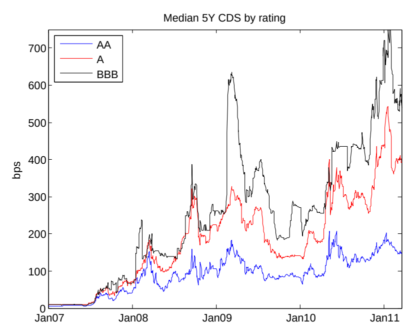

Figure 10: Median 5Y CDS spread evolution for a set of European financial institutions ordered by rating grades over the period 2007-2010.

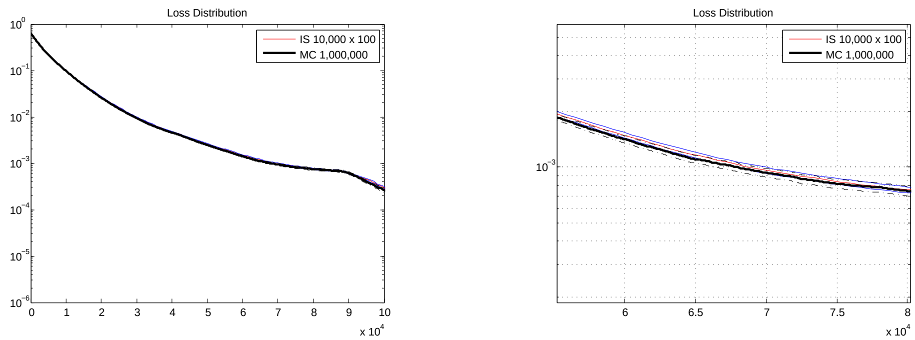

Figure 11: Loss distribution under the market mode. The black line shows the results of the plain Monte Carlo (MC) method using 1,000,000 simulations. The red and blue lines show, respectively, the importance sampling (IS) estimates and their 5%-95% confidence intervals using 10,000 macroeconomic scenarios and 100 default simulations on each macroeconomic scenario.

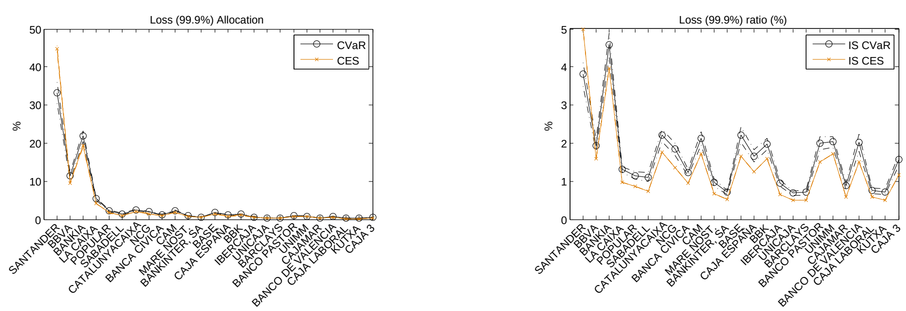

Figure 12: Risk allocation under market valuation for the *VaR* (CVaR) and *ES* (CES) criteria. Continuous and dashed lines represent, respectively, the IS estimates and the 5%-95% confidence intervals. Left and right graphs show, respectively, the total risk allocation and the allocated risk relative to the institution size.

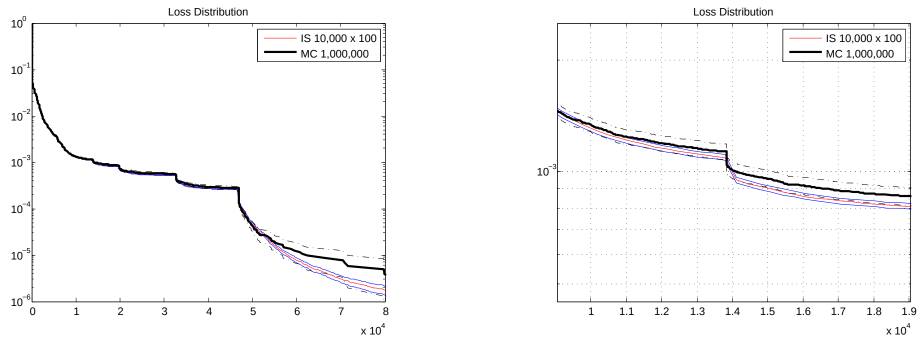

Figure 13: Loss distribution of the portfolio at December 2007 using 10,000 x 100 importance sampling (IS) and 1,000,000 plain Monte Carlo (MC) simulations. The black and red lines show, respectively, the plain Monte Carlo and IS results while the blue lines indicate the 5%-95% confidence interval of the IS estimates. Left and right graphs show, respectively, the tail distribution and its detail in the neighborhood of the 99.9% probability loss level.

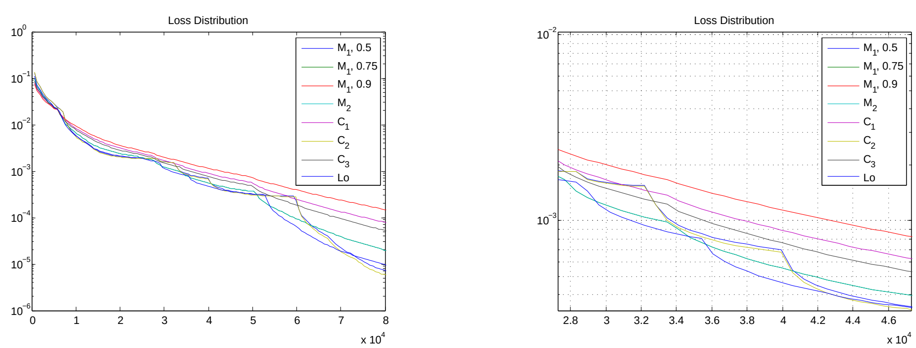

Figure 14: Loss distribution of the portfolio at December 2010 under the different macroeconomic sensitivity parameter estimates and using 10,000 x 100 importance sampling (IS) and 1,000,000 plain Monte Carlo (MC) simulations. Left and right graphs show, respectively, the tail distribution and its detail in the neighborhood of the 99.9% probability loss level.

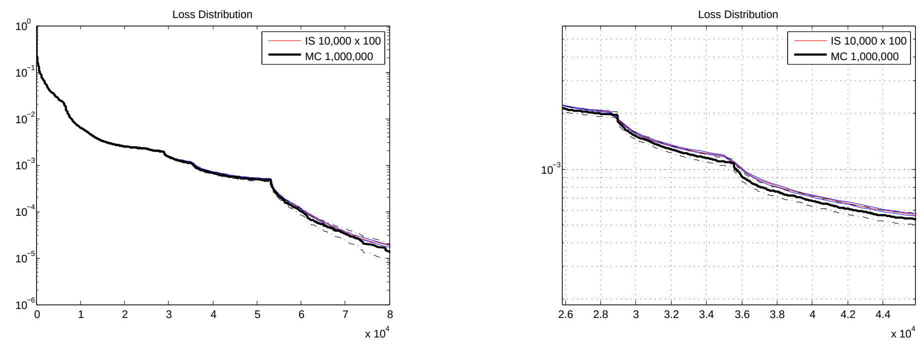

Figure 15: Loss distribution of the portfolio at December 2010 under *PD* uncertainty and using 10,000 x 100 importance sampling (IS) and 1,000,000 plain Monte Carlo (MC) simulations. The black and red lines show, respectively, the plain Monte Carlo and IS results while the blue lines indicate the 5%-95% confidence interval of the IS estimates. Left and right graphs show, respectively, the tail distribution and its detail in the neighborhood of the 99.9% probability loss level.

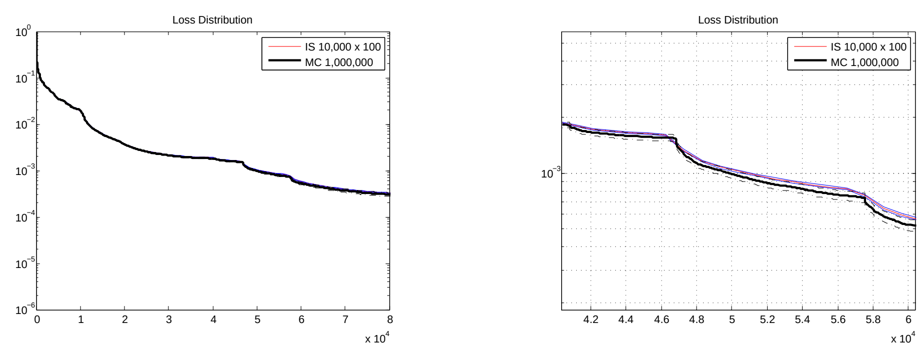

Figure 16: Loss distribution of the portfolio at December 2010 under *LGD* uncertainty and using 10,000 x 100 importance sampling (IS) and 1,000,000 plain Monte Carlo (MC) simulations. The black and red lines show, respectively, the plain Monte Carlo and IS results while the blue lines indicate the 5%-95% confidence interval of the IS estimates. Left and right graphs show, respectively, the tail distribution and its detail in the neighborhood of the 99.9% probability loss level.

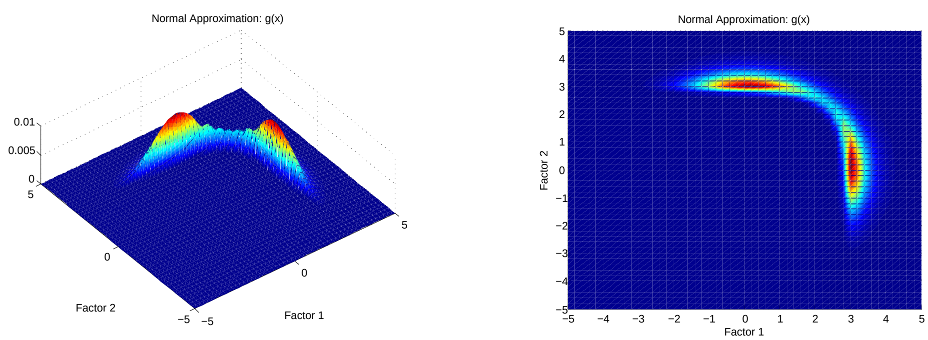

Figure 17: Function $g(x)$ provided by the normal approximation considering a portfolio of 1,000 counterparties and the parameters $PD = 1\%$, $LGD = 40\%$, $EAD = 1,000$, and $\alpha = 55\%$. Each half of the portfolio is exposed to a certain macroeconomic factor.