



Universidad de Castilla La Mancha

TESIS DOCTORAL

**Three essays on computational finance and
credit risk measurement and management**

Autor:

Rubén García Céspedes

Director:

Manuel Moreno Fuentes

Facultad de Ciencias Jurídicas y Sociales

Departamento de Análisis Económico y Finanzas

Toledo 2014

To my family and friends

Contents

Acknowledgements	V
Resumen	VI
Summary	X
List of Figures	XIV
List of Tables	XVI

1 Estimating the distribution of total default losses on the Spanish financial system	1
1.1 Introduction	1
1.2 The Vasicek (1987) Model	3
1.3 Importance sampling for credit risk	5
1.3.1 Optimal conditional distribution	6
1.3.2 Optimal macroeconomic distribution	8
1.4 Portfolio data	8
1.4.1 Probability of default (<i>PD</i>)	9
1.4.2 Exposure at default (<i>EAD</i>)	9
1.4.3 Loss given default (<i>LGD</i>)	10
1.4.4 Factor correlation (α)	11
1.4.5 Portfolio expected loss and Basel loss distribution	11
1.5 Importance sampling results	12
1.6 Importance sampling modifications and extensions	13
1.6.1 Random loss given default	14
1.6.2 Market mode	18
1.7 Parameter variability	21

1.7.1	Pre-crisis analysis	21
1.7.2	Parameter uncertainty	22
1.8	Conclusions	25
Appendix	27
Multimodal distributions	27
Loop decoupling	28
Appendix of Tables	31
Appendix of Figures	36
2	Extended saddlepoint methods for credit risk measurement	45
2.1	Introduction	45
2.2	Saddlepoint methods	46
2.2.1	Density approximations	46
2.2.2	Cumulative distribution approximations	47
2.3	Credit risk and saddlepoint approximation	49
2.3.1	Saddlepoint method for credit risk	49
2.3.2	Risk contributions	51
2.3.3	Random <i>LGD</i> models	55
2.3.4	Market mode models	57
2.4	Empirical analysis	58
2.4.1	Data	59
2.4.2	Results	59
2.5	Conclusions	65
Appendix of Tables	68
Appendix of Figures	71
3	Taylor expansion based methods to measure credit risk	78
3.1	Introduction	78
3.2	The Vasicek (1987) model	79
3.3	The Pykhtin (2004) approximate model	80
3.4	Portfolio Results	84
3.4.1	Portfolio VaR and ES	85
3.4.2	Analytical VaR Contributions	85
3.5	Correlated random LGD	88
3.5.1	Simple random LGD model	88

3.5.2	Advanced random LGD model	89
3.6	Conclusions	93
	Appendix of Figures	94
	Conclusions	98
	Conclusiones	100
	References	102

Acknowledgements

This dissertation is the conclusion of many years of hard work. I have had the privilege to meet some wonderful and interesting people during these years.

First and foremost, I would like to mention my gratitude to my supervisor Manuel Moreno Fuentes. Over the past years he has guided me over the process to write this dissertation and helped me with it. Without his motivation and support this thesis could not have been completed. Being the advisor of a part-time PhD student can be a risky bet and I want to thank him for giving me the opportunity to work with him.

The first ideas for this dissertation took place during the AFI - Quantitative Finance Executive Master, I want to thank Jose Luis Fernandez and Paul MacManus, teachers at AFI, for their support on the first stages of this dissertation and Juan Antonio de Juan for suggesting the initial topic. What started as a game finished becoming a PhD dissertation.

I would like to thank professor Paul Glasserman, we met on the Financial Engineering Summer School 2012 and his suggestions and comments on the importance sampling chapter were extremely useful. The saddlepoint chapter also got very interesting comments from Cornelis Oosterlee, Xinzheng Huang and Richard Martin, I am very grateful to them for their suggestions.

I am also indebted to the attendants to the Spanish Finance Forum 2013 for their comments, and specially to Pedro Serrano and Alberto Bueno-Guerrero.

My family, “*ama*”, Juan Carlos and Mariló, and friends have always been there supporting me, I am very grateful to all of them, without their support this dissertation would not had been possible.

Thank you!

Resumen

El riesgo de crédito es el tipo de riesgo más relevante al que se enfrentan las entidades financieras puesto que los riesgos de mercado u operacional son generalmente menos relevantes para este tipo de empresas. De hecho, durante la crisis financiera anterior el riesgo de crédito fue la principal fuente de pérdidas en la cuenta de resultados de las entidades financieras. En este sentido, tanto los reguladores como las instituciones financieras están interesados en medir con precisión el riesgo de crédito de una cartera. Sin embargo, el problema al que se enfrentan los reguladores es diferente de aquel al que se enfrentan las instituciones financieras puesto que los primeros están interesados en la distribución de pérdidas de todo el sistema financiero en lugar de en la de una cartera concreta.

Los reguladores financieros miden de manera separada los riesgos de crédito, de mercado y operacional y exigen a las instituciones financieras que tengan un nivel de recursos propios suficiente para soportar escenarios extremos de pérdidas con muy baja probabilidad de ocurrencia. Esto se conoce como la regulación de Basilea que trata de asegurar la estabilidad del sistema financiero. En el caso del riesgo de crédito el escenario seleccionado tiene una probabilidad de ocurrencia del 0.1%.

Vasicek (1987) propuso el modelo para la medición del riesgo de crédito más extendido en la actualidad. Este modelo es utilizado tanto por los reguladores como por la industria bancaria para obtener la distribución de pérdidas de una cartera. El supuesto fundamental de este modelo es que el comportamiento de incumplimiento de un cliente j viene determinado por un conjunto de factores macroeconómicos $Z = \{z_1, \dots, z_k\}$ y un término idiosincrático ε_j . El valor de los activos, V_j , de un cliente se expresa como

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

El incumplimiento se produce si el valor de los activos cae por debajo de un determinado nivel k_j . En este modelo todos los clientes de la cartera comparten los mismo factores macroeconómicos pero distintos términos idiosincráticos. Por tanto, los factores macroeconómicos capturan la interrelación entre los clientes. Las pérdidas totales de una cartera formada por M clientes se pueden expresar

como

$$L = \sum_{j=1}^M EAD_j LGD_j \mathbf{1}(V_j \leq k_j)$$

donde EAD_j es la exposición en el momento del incumplimiento y LGD_j es la tasa de pérdida provocada por el incumplimiento, es decir, 1 menos la tasa de recuperación final. Bajo condiciones muy restrictivas el modelo de Vasicek (1987) proporciona una solución analítica para la distribución de pérdidas. Éste es el caso de carteras expuestas a un único factor macroeconómico y formadas por muchos clientes idénticos, esto se conoce como una cartera asintótica con un único factor de riesgo (ASRF). Bajo condiciones generales la distribución de pérdidas se suele estimar mediante métodos de simulación de Monte Carlo.

Sin embargo la estimación de escenarios de baja probabilidad empleando métodos de Monte Carlo puede ser una tarea que requiera mucho tiempo en el caso de carteras grandes y probabilidades de pérdidas muy bajas. Esta estimación requiere mucho más tiempo si queremos repartir el riesgo entre los distintos clientes de la cartera. Como consecuencia, han surgido métodos alternativos al “simple” Monte Carlo. Estas alternativas se pueden dividir en dos grandes grupos, métodos exactos y aproximados. El método propuesto en Glasserman and Li (2005) se incluye dentro de los métodos exactos mientras que dos de los métodos aproximados se pueden encontrar en Huang et al. (2007) y Pykhtin (2004).

En esta tesis exploramos las distintas alternativas para estimar y repartir el riesgo de crédito de una cartera y proponemos nuevos modelos y extensiones para los actualmente disponibles. Contrastamos empíricamente la precisión de los modelos empleando una cartera formada por todas las instituciones del sistema financiero español. Esta es una cartera muy concentrada en la que todos los clientes excepto dos (Santander y BBVA) están expuestos a una única geografía. Por lo tanto esta tesis presenta dos grandes contribuciones a la literatura sobre riesgo de crédito:

1. Por un lado exploramos y extendemos algunos de los métodos actualmente disponibles para medir el riesgo de crédito a la vez que proponemos otros nuevos. Más en detalle, proponemos considerar un modelo de LGD aleatoria y un modelo de valoración de mercado para los casos de muestreo por importancia (IS) y los métodos de *saddlepoint*. En el caso de los métodos de *saddlepoint* adicionalmente proponemos un nuevo método para repartir el riesgo que utiliza polinomios de Hermite. También obtenemos una fórmula analítica que permite considerar la correlación entre los incumplimientos y las recuperaciones basada en las aproximaciones de Taylor.
2. Por otro lado, medimos y repartimos el riesgo del sistema financiero español. Aunque en Campos et al. (2007) se trató de medir el riesgo de este sistema financiero, hemos mejorado su análisis mediante la utilización de estimaciones de LGD basadas en datos históricos, hemos

repartido el riesgo entre las distintas instituciones financieras y hemos permitido que las mismas estén expuestas a más de un único factor macroeconómico.

El Capítulo 1 comienza analizando el modelo propuesto en Glasserman and Li (2005). Primero introducimos el método del muestreo por importancia (IS) y obtenemos los parámetros del modelo para la cartera de instituciones financieras españolas en Diciembre 2010. Para cada una de las instituciones obtenemos una PD (probabilidad de incumplimiento) basada en su calificación (*rating*) externo, una LGD mediante el procedimiento descrito en Bennet (2002) y los datos históricos de recuperaciones de la Federal Deposit Insurance Corporation (FDIC), una EAD basada en la información pública del balance y, finalmente, la exposición a los factores macroeconómicos basada en la información pública anual de las instituciones financieras por país. A continuación obtenemos la distribución de pérdidas de la cartera y el reparto del riesgo basado en los criterios de *Value-at-Risk* (VaR) y *Expected Shortfall* (ES). Encontramos que ambos criterios de reparto pueden generar resultados muy distintos para las dos instituciones más grandes.

Seguidamente extendemos el modelo de IS para los casos de recuperaciones aleatorias y valoración de mercado. En el caso de recuperaciones aleatorias permitimos que las mismas dependan solo de factores macroeconómicos o, alternativamente, de factores macroeconómicos e idiosincráticos. Ambos casos generan resultados similares para la distribución de pérdidas pero no para el reparto del riesgo. De hecho el modelo mixto de recuperaciones macroeconómicas e idiosincráticas requiere un número elevado de simulaciones para generar repartos del riesgo con intervalos de confianza pequeños. Respecto al modelo de valoraciones de mercado, la distribución de pérdidas se ve considerablemente desplazada a la derecha ya que ahora existen nuevos estados que producen pérdidas elevadas.

Este primer capítulo finaliza midiendo el impacto de la variabilidad en los parámetros del modelo sobre la distribución de pérdidas. Primero medimos el efecto del ciclo económico obteniendo la distribución de pérdidas en Diciembre 2007, una fecha anterior a la crisis. De acuerdo a nuestros resultados la principal fuente de cambios en la distribución de pérdidas son los cambios de ratings. En segundo lugar, medimos los efectos de la incertidumbre en los parámetros calibrados del modelo, $\alpha_{f,j}$, PD y LGD . La incertidumbre en las estimaciones proviene del reducido número de incumplimientos disponibles para calibrar dichos parámetros. En este caso, de acuerdo a nuestros resultados, la incertidumbre en la LGD es la principal fuente de preocupación en la estimación de la distribución de pérdidas. Adicionalmente incluimos otras dos extensiones del método de IS que desarrollamos, llamadas “*distribuciones multimodales*” y “*desacople de bucles*”.

El Capítulo 2 estudia el método de *saddlepoint* presentado en Huang et al. (2007) y propone un nuevo método de reparto del riesgo junto con otras extensiones de los métodos de *saddlepoint* actualmente disponibles. Primero presentamos los métodos de *saddlepoint*. Seguidamente proponemos un nuevo método de reparto del riesgo cuya idea principal consiste en no emplear un *saddlepoint*

distinto para cada cliente sino emplear el de la cartera total. Mostramos los detalles para modificar el método clásico de aproximación de *saddlepoint*. De modo similar al Capítulo 1, este capítulo también extiende el modelo de *saddlepoint* para los casos de *LGD* aleatoria y valoración de mercado.

Todas estas extensiones son contrastadas empíricamente empleando la cartera de instituciones financieras españolas. El nuevo método de reparto propuesto reduce considerablemente el número de cálculos necesarios para realizar el reparto del riesgo y, comparado con el método aproximado propuesto en Martin and Thompson (2001), este nuevo método requiere un tiempo similar de computación pero los resultados están más cercanos a los exactos. También mostramos que los métodos de *saddlepoint* no son la mejor opción bajo recuperaciones mixtas, macroeconómicas e idiosincráticas, o valoración de mercado debido al alto número de cálculos necesarios.

Finalmente, en el Capítulo 3 se aplican las ideas de Pykhtin (2004) sobre aproximaciones basadas en expansiones de Taylor a nuestra cartera y se propone un nuevo modelo de *LGD* aleatoria. Comenzamos explicando el modelo de Pykhtin (2004) para aproximar la distribución de pérdidas de una cartera y los criterios de reparto del riesgo basados en *VaR* y *ES* presentados en Morone et al. (2012). Mostramos que los modelos basados en expansiones de Taylor no pueden capturar de manera precisa el perfil de concentración de nuestra cartera tanto en la distribución de pérdidas como en el reparto del riesgo. Este capítulo finaliza empleando las ideas de la expansión de Taylor para estimar la distribución de pérdidas de una cartera expuesta a un único factor que determina el incumplimiento bajo *LGD* aleatoria. Este nuevo modelo permite que la *LGD* tenga un componente idiosincrático y otro macroeconómico que puede estar correlacionado con el factor macroeconómico que determina el incumplimiento. Empleando las ideas basadas en la expansión de Taylor y una cartera de referencia obtenemos resultados muy satisfactorios para los casos de distribución normal y lognormal en la *LGD* y para distintos niveles de correlación. Creemos que este modelo puede ser implementado de manera sencilla bajo la regulación actual de Basilea para considerar las recuperaciones aleatorias mejor de como lo hace la *downturn LGD*.

Summary

The credit risk is the most relevant type of risk that financial institutions face as market or operational risk are usually less relevant for this type of firms. In fact, during the previous financial crisis credit risk has been the main source of P&L losses. In this sense both financial institutions and regulators are interested in accurately measuring the credit risk of a given portfolio. However the problem faced by regulators is different from that faced by financial institutions as they are concerned about the loss distribution of the whole financial system rather than with that of a given portfolio.

Financial regulators measure separately the credit, market, and operational risks and force financial institutions to have a level of own resources so that they can bear extreme scenarios with very low probability of occurrence. This is known as the Basel regulation and tries to ensure the stability of the financial system. In the case of the credit risk the target scenario is a 99.9% probability loss scenario.

Vasicek (1987) proposed the most widely used model to measure credit risk. This model is used by practitioners and regulators to obtain the loss distribution of a given portfolio. Its main assumption is that the default behavior of a counterparty j is driven by a set of macroeconomic factors $Z = \{z_1, \dots, z_k\}$ and an idiosyncratic term ϵ_j . The assets value, V_j , of this counterparty is expressed as

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

Default happens if the assets value drops below a given threshold level k_j . In this model all the counterparties in the portfolio share the same macroeconomic factors but different idiosyncratic ones, then the macroeconomic factors capture the interrelation between the counterparties. The total loss of a portfolio made up of M counterparties can be expressed as

$$L = \sum_{j=1}^M EAD_j LGD_j \mathbf{1}(V_j \leq k_j)$$

where EAD_j is the exposure at default and LGD_j is one minus the final recovery rate. Under very restrictive conditions the model in Vasicek (1987) has a closed form representation. This is

the case of a single macroeconomic factor and a portfolio made up of many identical counterparties, this is known as Asymptotic Single Risk Factor (ASRF) portfolio. Under general conditions the loss distribution is usually estimated through Monte Carlo simulation methods.

However estimating low probability scenarios with simple Monte Carlo methods can be very time demanding in the case of very low probability losses and big portfolios. This gets even more time demanding when we want to allocate the risk over the different counterparties in the portfolio. As a consequence, alternative methods to the pure Monte Carlo one have arisen. We can divide these alternatives in two broad groups, exact and approximate methods. The method proposed in Glasserman and Li (2005) is among the exacts methods and two approximate methods can be found in Huang et al. (2007) and Pykhtin (2004).

In this thesis we explore the different alternatives to estimate and allocate the credit risk of a portfolio and we propose new models and extensions to the currently available ones. We test empirically the accuracy of the models considering a portfolio that includes all the institutions in the Spanish financial system. This is a very concentrated portfolio where all the counterparties but two of them (Santander and BBVA) are exposed to only one geography. Therefore this thesis has two broad contributions to the literature:

1. On one hand we explore and extend some of the currently available methods to measure the credit risk and propose new ones. In more detail, we propose to consider a random *LGD* model and a market valuation model for the case of the importance sampling (IS) and the saddle-point methods. In the case of the saddlepoint we additionally propose a new risk allocation method that uses Hermite polynomials. We also obtain a closed-form formula that captures the correlation between default and recoveries based on Taylor expansion approximations.
2. On the other hand, we measure and allocate the risk of the Spanish financial system. Although Campos et al. (2007) tried to measure the risk of this system, we have improved their analysis using a historical data based *LGD*, allocating the risk over the financial institutions and allowing institutions to be exposed to more than one macroeconomic factor.

Chapter 1 starts analyzing the model proposed in Glasserman and Li (2005). First we introduce the importance sampling (IS) method and obtain the model parameters for the portfolio of the Spanish financial system at December 2010. For each institution we obtain a *PD* based on its external rating, a *LGD* based on the Federal Deposit Insurance Corporation (FDIC) historical recovery data and the procedure described in Bennet (2002), an *EAD* based on the public balance information, and, finally, the macroeconomic factor loadings based on the public country by country yearly reports information of the financial institutions. Then we obtain the loss distribution of the portfolio and the risk allocation based on the Value-at-Risk (*VaR*) and the Expected Shortfall (*ES*) criteria. We find

that both loss allocation criteria can generate very different results for the two biggest institutions.

Next we extend the IS model to work under random recoveries and market valuation. In the case of the random recoveries we allow for pure macroeconomic recoveries and for mixed idiosyncratic and macroeconomic recoveries. Both methods produce similar results for the loss distribution but not for the risk allocation. In fact the mixed idiosyncratic and macroeconomic recoveries model requires a very high number of simulations to generate risk allocation results with narrow confidence intervals. Regarding the market valuation model the loss distribution gets considerably shifted to the right as it allows for new rating states producing high losses.

This first chapter concludes measuring the impact of the variability in the model parameters on the estimated loss distribution. First we measure the effect of the business cycle obtaining the loss distribution at December 2007, a pre-crisis date. According to our results the main source of changes in the loss distribution are changes in the ratings. Secondly, we measure the effects of the uncertainty in the calibrated model parameters, $\alpha_{f,j}$, PD , and LGD . The uncertainty in the estimates comes from the reduced number of historical defaults used to calibrate these parameters. In this case, according to our results, the LGD uncertainty is the main source of concern in the estimation of the loss distribution. Additionally we include other two extensions of the IS method that we developed called, respectively, “*multimodal distributions*” and “*loop decoupling*”.

Chapter 2 studies the saddlepoint method introduced in Huang et al. (2007) and proposes a new risk allocation method and other extensions to the currently available saddlepoint methods. We first introduce the saddlepoint methods. Later we propose a new risk allocation method whose underlying idea is based on not using a different saddlepoint for each counterparty in the portfolio but the one of the whole portfolio. We show the detailed steps to modify the classical saddlepoint approximation. This chapter also extends the random LGD and market valuation framework exposed in Chapter 1 to work under the saddlepoint approximations.

All these extensions are tested using the portfolio of the Spanish financial system. Our new risk allocation method reduces considerably the number of calculations required to estimate the risk contributions and, compared with the approximate method presented in Martin and Thompson (2001), this new method requires a similar computation time but the results are closer to the exact ones. We also show that the saddlepoint methods are not the best alternative under mixed idiosyncratic and macroeconomic recoveries or market valuation due to the high number of calculations that are needed.

Finally, Chapter 3 applies the Taylor expansion based ideas in Pykhtin (2004) to our portfolio and proposes a new random LGD model. We start explaining the model presented in Pykhtin (2004) to approximate the loss distribution of a portfolio and the VaR and ES allocation criteria based on Morone et al. (2012). We show that the Taylor expansion models can not accurately capture

the concentration profile of our portfolio in the loss distribution nor do they in the risk allocation process. This chapter ends employing the Taylor expansion ideas to estimate the loss distribution of a portfolio exposed to a single default driving factor under random *LGD*. This new model allows the *LGD* to have an idiosyncratic and a macroeconomic term that may be correlated with the default driving macroeconomic factor. Using the Taylor expansion based ideas and a benchmark portfolio we obtain very good results for normal and lognormal *LGD* models and for different correlation levels. We consider that this model can be easily implemented under the current Basel regulation to account for random recoveries better than the *downturn LGD* does.

List of Figures

1.1	Assets and deposits share of the top twenty-five Spanish financial institutions.	36
1.2	Assets, Expected Loss, and Basel 99.9% loss share	36
1.3	Loss distribution	37
1.4	Expected Shortfall (<i>ES</i>)	37
1.5	Risk Allocation under constant <i>LGD</i>	38
1.6	Comparison of the random <i>LGD</i> and constant <i>LGD</i> loss distributions	38
1.7	Risk Allocation under macroeconomic random <i>LGD</i>	39
1.8	Comparison of the two random <i>LGD</i> models and constant <i>LGD</i>	39
1.9	Risk Allocation under mixed macroeconomic and idiosyncratic random <i>LGD</i>	40
1.10	Median 5Y CDS spread evolution	40
1.11	Loss distribution under the market mode	41
1.12	Risk Allocation under market valuation	41
1.13	Loss distribution of the portfolio at December 2007	42
1.14	Loss distribution of the portfolio under the different macroeconomic sensitivity parameter estimates	42
1.15	Loss distribution of the portfolio under <i>PD</i> uncertainty	43
1.16	Loss distribution of the portfolio under <i>LGD</i> uncertainty	43
1.17	Multimodal distributions	44
2.1	Loss distribution of the Spanish portfolio	71
2.2	Loss distribution of the portfolio under the <i>adaptive saddlepoint approximation</i>	71
2.3	Risk Allocation based on 100 simulations	72
2.4	Risk Allocation based on 1000 simulations	72
2.5	Risk Allocation Normal Macroeconomic <i>LGD</i>	73
2.6	Risk Allocation Lognormal Macroeconomic <i>LGD</i>	73
2.7	Risk Allocation Logit Normal Macroeconomic <i>LGD</i>	74
2.8	Risk Allocation Probit Normal Macroeconomic <i>LGD</i>	74

2.9	Risk Allocation Normal Square Macroeconomic <i>LGD</i>	75
2.10	Risk Allocation Basel Macroeconomic <i>LGD</i>	75
2.11	Loss distribution of the Spanish portfolio under market valuation	76
2.12	Loss on current value distribution of the Spanish portfolio	76
2.13	<i>VaR</i> and <i>ES</i> contributions for the market mode model	77
2.14	<i>VaR</i> and <i>ES</i> contributions on an equally weighted (<i>EAD</i>) portfolio	77
3.1	Loss distribution and expected shortfall for the Spanish financial institutions	94
3.2	Risk Allocation for the Spanish financial institutions	95
3.3	Simplified random LGD model	95
3.4	Advanced random LGD model	96
3.5	Advanced random LGD model under Normal LGD	96
3.6	Advanced random LGD model under Lognormal LGD	97

List of Tables

1.1	Spanish financial institutions involved in a merger / acquisition	31
1.2	<i>LGD</i> estimates for losses on deposits and losses on assets	32
1.3	BBVA and Santander country exposures	32
1.4	Comparison of the 99.9% probability loss levels under different random <i>LGD</i> models	33
1.5	Discount factors by rating grade	33
1.6	Average 1-year rating migration matrix	33
1.7	Comparison of the possible values of α	34
1.8	Average S&P default rates	34
1.9	PD uncertainty	35
2.1	Constant conditional <i>LGD</i> models	68
2.2	Random conditional <i>LGD</i> models	68
2.3	Calibrated parameters for the constant conditional <i>LGD</i> model	68
2.4	Tail probabilities for different loss levels for the pure macroeconomic <i>LGD</i>	69
2.5	Calibrated parameters for the random conditional <i>LGD</i> model	69
2.6	Tail probabilities for different loss levels for the mixed macroeconomic and idiosyncratic random <i>LGD</i>	70
2.7	Discount factors by rating grade	70
2.8	Average 1-year rating migration matrix	70

Chapter 1

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

1.1 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 interconnect-edness 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.

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. Bennet (2002), Kuritzkes et al. (2002), and Cariboni et al. (2011) used a similar approach to that in Campos et al. (2007) to define an optimum deposits insurance fund.

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 assess 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.

1.2 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 $Z = \{z_1, z_2, \dots, z_k\}$ 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

$$V_j = \sum_{f=1}^k \alpha_{f,j} z_f + \varepsilon_j \sqrt{1 - \sum_{f=1}^k \alpha_{f,j}^2} \quad (1.1)$$

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

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

¹Dependent factors can always be orthogonalized.

²It 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 .

Bank portfolios are composed of this kind of contracts. The total loss of a portfolio including M contracts or clients is given as $L = \sum_{j=1}^M x_j$, being x_j the individual loss of the client or contract j . Under homogeneous granular portfolios,³ the probability of default $Prob(D_j = 1|Z)$ and the observed default rates DR_z tend to be equal. This means that all the idiosyncratic risk of the different clients disappear 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.

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.⁴
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:

³This type of portfolios is made up of many identical contracts, with the same risk parameters.

⁴For 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.

1. Value at Risk: $VaR(q) = F^{-1}(q)$.⁵
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, $CES_j(q) = E(x_j|L \geq VaR(q))$.

1.3 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 is⁶

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

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

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

⁵The Basel regulation requires a bank to have an amount of own resources equal to the $VaR(99.9\%)$.

⁶It can be noted that the variance of this estimator vanishes for the sampling distribution $g(L_i) \propto \mathbf{1}(L_i \geq l) f(L_i)$.

and can be estimated using the IS method as

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

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

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

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

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

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

$$Var(\widehat{ES}) \approx N \frac{\sum_{i=1}^N (L_i - \widehat{ES})^2 \mathbf{1}(L_i \geq l) \left(\frac{f(L_i)}{g(L_i)}\right)^2}{\left(\sum_{i=1}^N \mathbf{1}(L_i \geq l) \frac{f(L_i)}{g(L_i)}\right)^2} \quad (1.2)$$

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 (1.2).

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.

1.3.1 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

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

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

$$\psi(\theta) = \sum_{j=1}^M \ln \left(1 + P_{j,Z} \left(e^{LGD_j EAD_j \theta} - 1 \right) \right) \quad (1.3)$$

Note that, conditional to the macroeconomic state Z , the losses of every client j are independent. Then, (1.3) 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 $\theta = 0$.⁷

1.3.2 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

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

It still remains to estimate $Prob(L \geq l|Z)$. To this aim, we decided to use a simple approach assuming that $L|Z \sim N(a, b^2)$ where⁸

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

1.4 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, 2010.⁹ This

⁷We 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 .

⁸Other alternatives such as the *constant approach* or the *tail bound approach* can be found in Glasserman and Li (2005).

⁹The 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.

fund was analyzed in Campos et al. (2007) by using a simple single factor model and 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.

1.4.1 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 agencies¹⁰ to infer a probability of default. This probability is obtained adjusting an exponential function to the default rates of the ratings up to B- and imposing a value of 0.3% for a rating AA, a commonly accepted feature. Entities with no external rating are assigned one notch less than the average rating of the portfolio with external rating.¹¹ This implies that banks without external rating are assigned a A- rating and the remaining institutions 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.¹² 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.

1.4.2 Exposure at default (*EAD*)

Details on assets, liabilities, and deposits for the FGD institutions are available in the AEB, CECA, and AECR webpages.¹³ 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 1.1), we have summed all the information from the different institutions that belong to the same group.

[INSERT TABLE 1.1 AROUND HERE]

¹⁰See S&P (2009), Moody's (2009), and Fitch (2009).

¹¹This average is computed weighting by assets and distinguishing between banks and saving banks.

¹²For the sake of brevity, these results are not reported here and are available upon request.

¹³AEB 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.

Figure 1.1 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 index¹⁴ 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 system has few players and is very concentrated, a common feature in most of the countries worldwide.

[INSERT FIGURE 1.1 AROUND HERE]

1.4.3 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 LossCalcTM use most of these factors, see Gupton and Stein (2002) for more details on this model.

Bennet (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 $E(LGD_{j,t}|Asset\ Bucket)$ 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 1.378.¹⁵ Table 1.2 provides the $LGDs$ obtained in this way.

¹⁴The 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.

¹⁵This factor is based on the numbers obtained in Bennet 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.

[INSERT TABLE 1.2 AROUND HERE]

1.4.4 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 $\sqrt{0.24}$.¹⁶ 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 1.1 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.¹⁷

Table 1.3 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 1.3 AROUND HERE]

1.4.5 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 €, respectively.

Figure 1.2 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

¹⁶Kuritzkes 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 $\sqrt{0.24}$.

¹⁷These correlations are available upon request.

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 1.2 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 1.2 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.

1.5 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 1.3 shows the loss distribution under IS and 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!¹⁸

[INSERT FIGURE 1.3 AROUND HERE]

Figure 1.4 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 €.

¹⁸All the figures in the paper are based on the IS results rather than on the Monte Carlo method.

[INSERT FIGURE 1.4 AROUND HERE]

Figure 1.5 shows the risk allocation rule according to the *ES* and *VaR* contributions and the confidence intervals for the IS technique.¹⁹ These intervals are quite thin after only 10,000 simulations, one of the main advantages of the IS method over the 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 1.5 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.
2. The *LGD* of Bankia is 28,948 MM €, close to the *VaR*(99.9%) value. Then, this firm copes 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.
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.

1.6 Importance sampling modifications and extensions

This Section extends the classical IS framework to deal with random recoveries and market valuation. Other extensions were performed:²⁰

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 €.

¹⁹For the *VaR* contributions we have used a $\pm 1\%$ interval around the desired loss level

²⁰For the sake of brevity, we just enumerate here these additional extensions and defer the details to a final Appendix.

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.

1.6.1 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. 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 studied²¹ and some of their moments have a closed-form expression, for example

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

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

²¹See Gordy (2000) or Dullmann et al. (2008).

factor correlation of the PD for the LGD as follows:

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

where $M_{PD} = M_{LGD}$ equates the GDP correlation matrix of the different countries.²² 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 1.6 shows the comparison between the loss distributions of the portfolio under random and constant LGD s. The 99.9% probability loss is 36,970 MM €, that is, 1.15 times the loss level under constant LGD . The equivalent expected shortfall level is 19,326 MM €. Figure 1.7 shows the risk allocation under VaR and ES for the new 99.9% probability loss level. Comparing with Figure 1.5 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 1.6 AND 1.7 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 .

²²For 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.

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 $E(Var(LGD_{j,z}|z)) = 1.338\%$,²³ implying a variability that is higher than the average $LGDs$ of the big financial institutions. Intuitively, now, more institutions can generate high and 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 $E(LGD_{j,z}|z)$ and $E(LGD_{j,z}^2|z)$ 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, $E(LGD_{j,z}^2|z)$, this is,²⁴

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

²³Then, the LGD can change $\pm 11.56\%$ with respect to its mean.

²⁴ $\Phi_2(X, M, \Sigma)$ denotes the probability distribution function (evaluated at the point X) of a bivariate normal random variable with mean vector M and covariance matrix Σ .

with

$$M = \begin{pmatrix} \alpha r z_{LGD} \\ \alpha r z_{LGD} \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \alpha^2 s^2 + (1 - \alpha^2) & \alpha^2 s^2 \\ \alpha^2 s^2 & \alpha^2 s^2 + (1 - \alpha^2) \end{pmatrix}$$

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

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

where

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

Figure 1.8 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 1.8 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.²⁵ 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 1.9 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 1.9 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).²⁶

Table 1.4 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.

²⁵The ES equivalent loss is 19,473 MM €.

²⁶Detailed results are not reported here and are available upon request.

[INSERT TABLE 1.4 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 than to the default distribution.²⁷ In fact with the IS ideas 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.

1.6.2 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.²⁸ Figure 1.10 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 1.10 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 $18.35\% \times 1.378 = 25.28\%$ over assets.²⁹ 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 2.7 reports the 3-year discount factors obtained for each rating in this way.

[INSERT TABLE 2.7 AROUND HERE]

²⁷We acknowledge one of the referees for highlighting this alternative.

²⁸These data correspond to 5-year senior CDS since 2008 and were obtained from Markit.

²⁹As 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 2.8 includes the rating transition matrix employed.

[INSERT TABLE 2.8 AROUND HERE]

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 $AccumMP_{j,C,IR,FR}$.³⁰

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.

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 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

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

³⁰For example, $AccumMP_{j,C,IR,B-} = MP_{j,C,IR,B-} + MP_{j,C,IR,CCC} + MP_{j,C,IR,D}$.

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 1.11 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.³¹

[INSERT FIGURE 1.11 AROUND HERE]

Regarding the *VaR* and *ES* based contributions we will allocate the 68,852 MM € loss over the current market value. Figure 1.12 provides the results and shows that the top contributor is Santander.

[INSERT FIGURE 1.12 AROUND HERE]

³¹The analysis has been performed using the rating scale considering modifiers but it could be done without these modifiers.

1.7 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.

1.7.1 Pre-crisis analysis

We estimate the loss distribution at December 2007 to assess 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.
- 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 1.13 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 1.13 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.

1.7.2 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.

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 software³² to implement the Vasicek (1987) model. Examples of these studies are López (2002), Lee et al. (2009), Qibin et al. (2009), and Castro (2012).

López (2002) 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 1.7 shows the values of α from these papers and illustrates that the Basel Committee estimates³³ are close or lower than the results in all these papers but for two of the models in Castro (2012).

[INSERT TABLE 1.7 AROUND HERE]

Figure 1.14 reports the loss distributions obtained for the different macroeconomic sensitivity parameters for the portfolio of Spanish financial institutions at December 2010.

³²This software is currently called RiskFrontier and it was previously known as KMV.

³³These estimates will tend to be closer to 54% rather than to 38% because of the low *PD* in banks.

[INSERT FIGURE 1.14 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.

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 1.8 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 2012.³⁴ As it can be seen firms graded with the two highest ratings never default.

[INSERT TABLE 1.8 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.³⁵ 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 1.9 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 1.9 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 1.9 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

³⁴Data at rating modifier level is not available for the periods 1981-2007 and 1981-2008.

³⁵These confidence intervals can be obtained analytically or numerically, for instance, using a binomial distribution or a bootstrapping technique.

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 1.15 reports the loss distribution of the portfolio under this approach.

[INSERT FIGURE 1.15 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.

***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 Bennet (2002) for the period 1986-1998. Kaufman (2004) gets similar results to those in Bennet (2002) but for the period 1980-2002.

Bennet (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 Bennet (2002) for losses to depositors and apply the ratio of losses to depositors to losses on total assets from Bennet 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 buckets³⁶ 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*.³⁷ Figure 1.16 includes the loss distribution of the portfolio after considering these two sources of uncertainty.

[INSERT FIGURE 1.16 AROUND HERE]

³⁶Table 2 in Bennet and Unal (2011) shows that this ratio can be up to 1.47.

³⁷If we have R defaults the *LGD* estimate is normally distributed with mean $\mu = \sum_{i=1}^R LGD_i/R$ and variance $\sum_{i=1}^R (LGD_i - \mu)^2/R$.

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.

1.8 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 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 Monte Carlo simulations as benchmark. Based on Bennet (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 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 Huang et al. (2007), Pykhtin (2004) 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.

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 1.17 provides the function $g(Z) = Prob(L \geq l|Z)e^{-\frac{z'l}{2}}$ for a simple case.³⁸ 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 1.17 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 optimum 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.³⁹ Trying to solve this problem, we decided to estimate the mean or the median of the optimum distribution $g(Z)$. Reitan and Aas (2010) proposed estimating the mean using Markov Chain Monte Carlo (MCMC)⁴⁰ and the Metropolis-Hasting

³⁸We 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.

³⁹It is easier to find numerically the maximum of a multivariate distribution than obtaining random samples from it.

⁴⁰However, they did not highlight any bias related to the use of the mode.

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 are⁴¹

$$\begin{aligned}\mu_{g(Z)} &= \frac{1}{N} \sum_{i=1}^N Z_j c \text{Prob}(L_i > l | Z_i) e^{-\frac{z_i z'_i}{2}} \frac{1}{\phi(Z_i)_{\mu, \Omega}} \\ \sigma_{g(Z)}^2 &= \frac{1}{N} \sum_{i=1}^N Z_i^2 c \text{Prob}(L_i > l | Z_i) e^{-\frac{z_i z'_i}{2}} \frac{1}{\phi(Z_i)_{\mu, \Omega}} - \mu_{g(Z)}^2\end{aligned}$$

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,

$$1 = \frac{1}{N} \sum_{i=1}^N c \text{Prob}(L_i > l | Z_i) e^{-\frac{z_i z'_i}{2}} \frac{1}{\phi(Z_i)_{\mu, \Omega}}$$

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 $Z = \{z_1, \dots, z_k\}$. 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 $\frac{1}{N} c \text{Prob}(L_i > l | Z_i) e^{-\frac{z_i z'_i}{2}} \frac{1}{\phi(Z_i)_{\mu, \Omega}}$ until the value 50% is obtained, that is,

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

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

⁴¹After all the experiments, the best results were obtained using as variance the maximum between 1 and the optimum variance.

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:

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

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)

$$\begin{aligned}
X_{n1} &= \frac{1}{N} \sum_{i=1}^{N_e} \sum_{k=1}^{N_i} L_{i,k} \mathbf{1}(L_{i,k} \geq l) \frac{f(L_{i,k})}{g(L_{i,k})} = \frac{1}{N} \sum_{i=1}^{N_e} S_i \\
X_{n2} &= \frac{1}{N} \sum_{i=1}^{N_e} \sum_{k=1}^{N_i} \mathbf{1}(L_{i,k} \geq l) \frac{f(L_{i,k})}{g(L_{i,k})} = \frac{1}{N} \sum_{i=1}^{N_e} R_i \\
\text{Var}(\widehat{ES}) &= \frac{\text{Var}(X_{n1} - \widehat{ES}X_{n2})}{X_{n2}^2} = \frac{\frac{1}{N^2} \sum_{i=1}^{N_e} \text{Var}(S_i - \widehat{ES}R_i)}{X_{n2}^2} \\
&\approx N \frac{\sum_{i=1}^{N_e} (S_i - \widehat{ES}R_i)^2}{\sum_{i=1}^{N_e} R_i} \tag{1.4}
\end{aligned}$$

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 (1.4) 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 (1.4) 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 1.1: 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 1.2: *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 1.3: 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 1.4: 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 macroeconomic 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 1.5: 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 1.6: 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 1.7: Comparison of the values of α in López (2002), Lee et al. (2009), Qibin et al. (2009), and Castro (2012). α_{Lo} denotes the estimate in López (2002) 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 (2002)	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 1.8: 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 1.9: 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

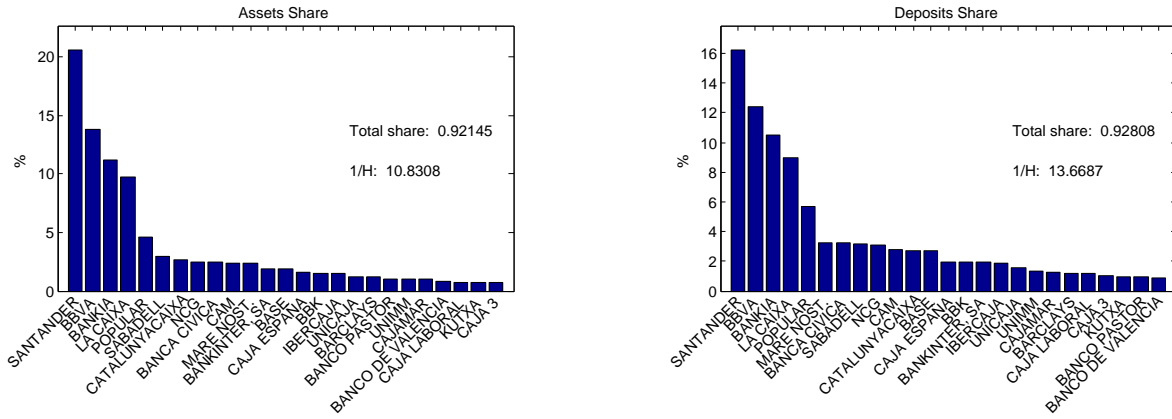


Figure 1.1: Assets and deposits share of the top twenty-five Spanish financial institutions.

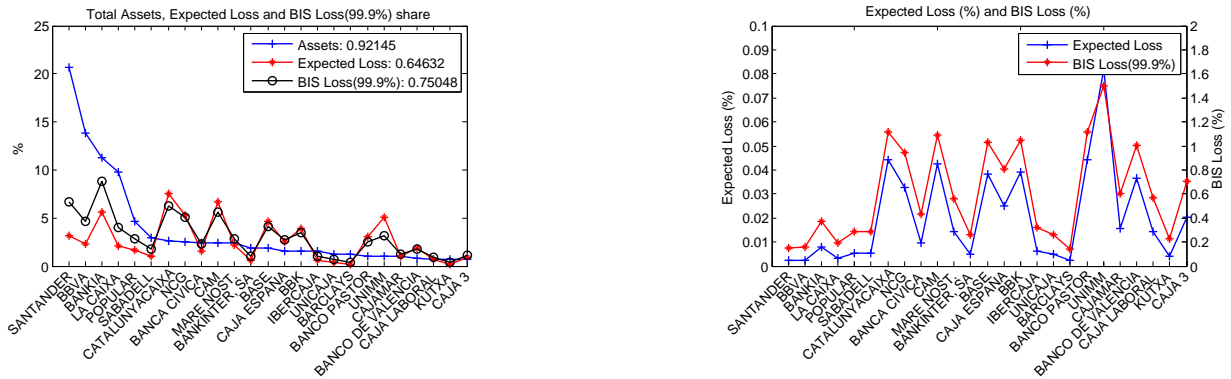


Figure 1.2: Assets, Expected Loss, and Basel 99.9% loss share of the top 25 Spanish financial institutions. Left and right graphs show, respectively, the amount allocation and the allocated amount relative to the institution size.

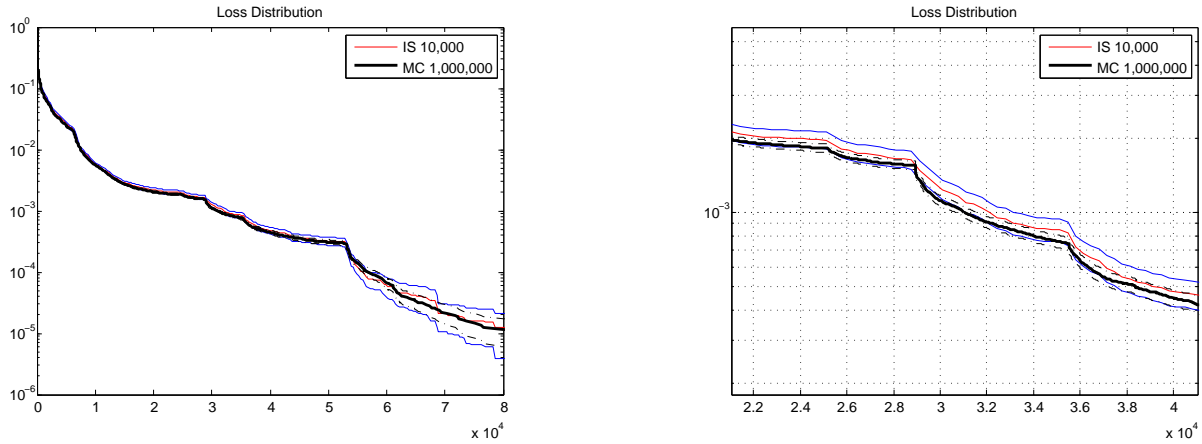


Figure 1.3: Loss distribution using 10,000 importance sampling (IS) and 1,000,000 Monte Carlo (MC) simulations. The black and red lines show, respectively, the Monte Carlo and IS results while the blue lines indicate the 5%-95% confidence interval of the IS estimates. Left and right graphs show, respectively, the tail distribution and its detail in the neighborhood of the 99.9% probability loss level.

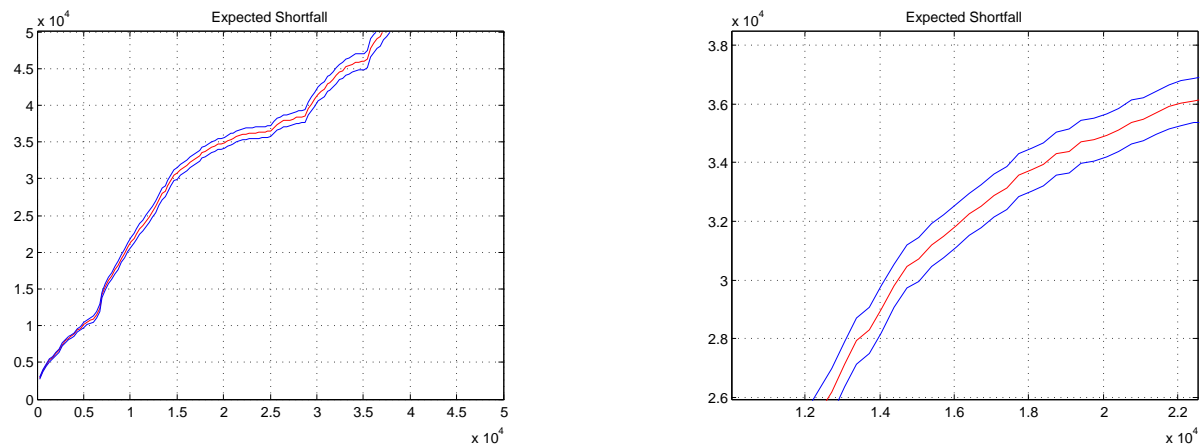


Figure 1.4: Expected Shortfall (*ES*) using 10,000 importance sampling (IS) simulations. The red and blue lines show, respectively, the IS results for the expected shortfall estimate and its 5%-95% confidence intervals.

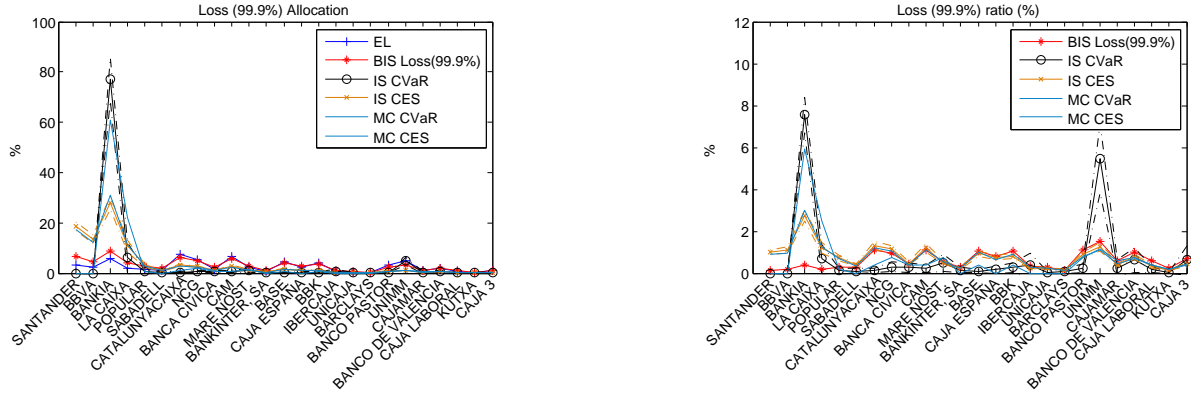


Figure 1.5: Risk allocation under constant *LGD* based on expected loss (EL), Basel loss 99.9% (BIS), contributions to *VaR* (CVaR) and *ES* (CES) both using importance sampling (IS) and Monte Carlo (MC) criteria. Left and right graphs show, respectively, the total risk allocation and the allocated risk relative to the institution size.

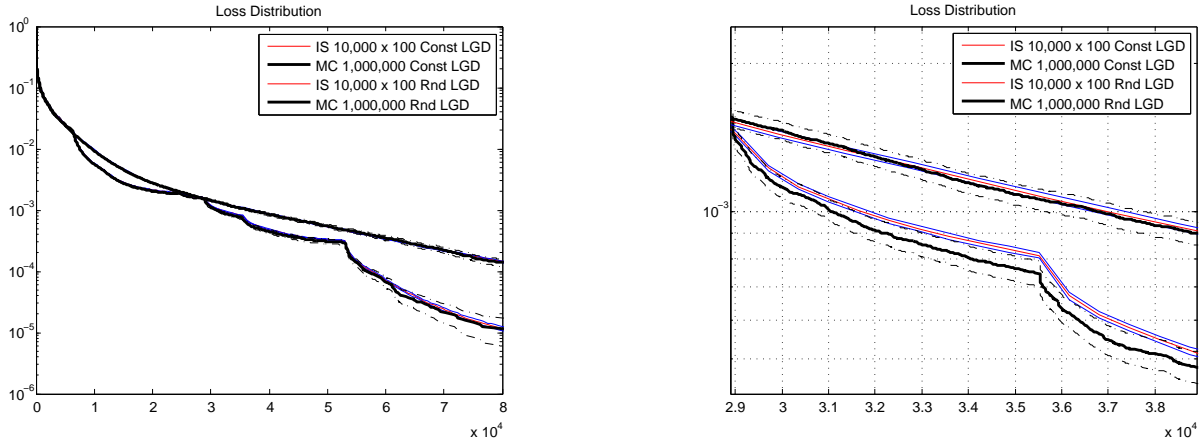


Figure 1.6: Comparison of the random *LGD* (Rnd LGD) and constant *LGD* (Const LGD) loss distributions. Black lines show the results of the Monte Carlo (MC) method using 1,000,000 simulations. The red and blue lines show, respectively, the importance sampling (IS) estimates and their 5%-95% confidence 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.

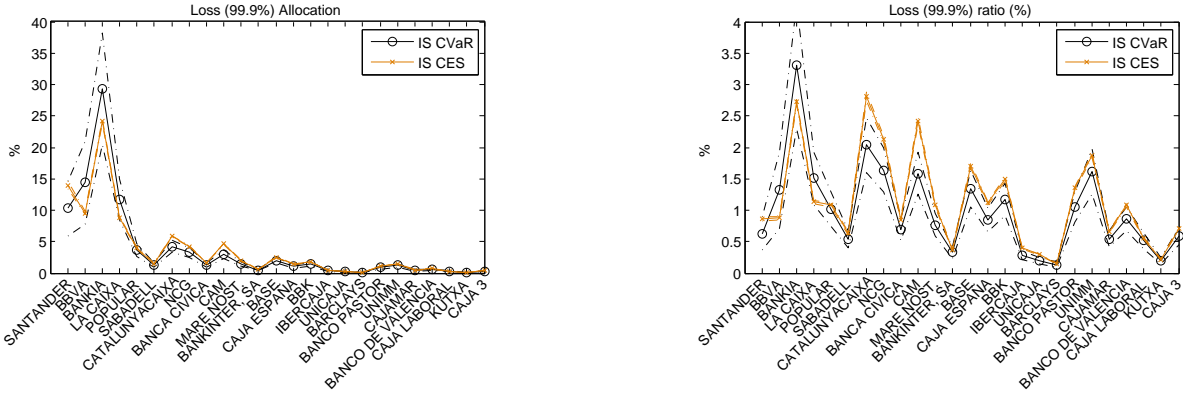


Figure 1.7: Risk allocation under macroeconomic random LGD (LGD^C) for the VaR ($CVaR$) and the ES (CES) criteria. Continuous and dashed lines represent, respectively, the IS estimates and the 5%-95% confidence intervals. Left and right graphs show, respectively, the total risk allocation and the allocated risk relative to the institution size.

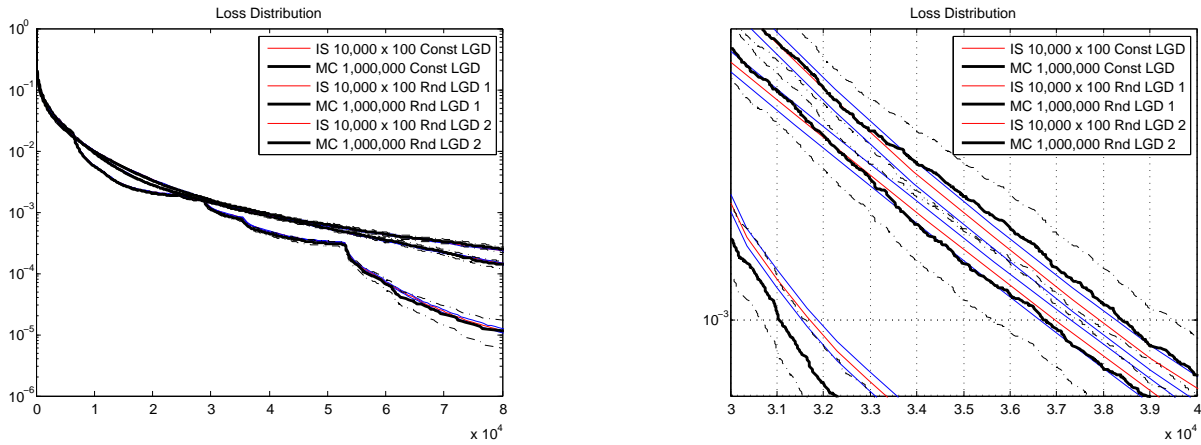


Figure 1.8: Comparison of the two random LGD models ($Rnd\ LGD^C / Rnd\ LGD^R$) and constant LGD ($Const\ LGD$) loss distributions. Black lines show the results of the Monte Carlo (MC) method using 1,000,000 simulations. The red and blue lines show, respectively, the importance sampling (IS) estimates 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.

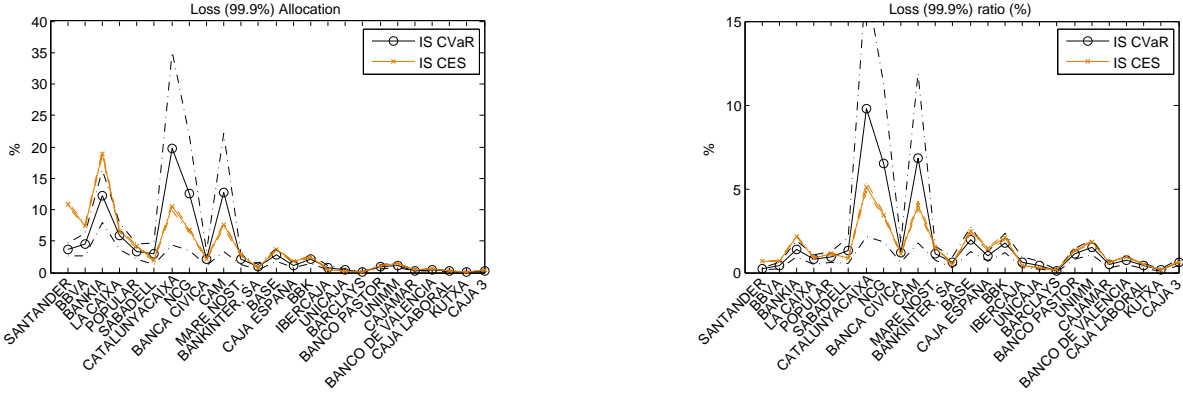


Figure 1.9: Risk allocation under mixed macroeconomic and idiosyncratic random LGD (LGD^R) for the VaR (CVaR) and ES (CES) criteria. Continuous and dashed lines represent, respectively, the IS estimates and the 5%-95% confidence intervals. Left and right graphs show, respectively, the total risk allocation and the allocated risk relative to the institution size.

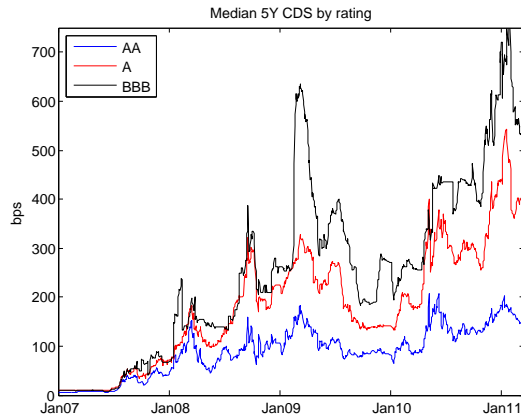


Figure 1.10: Median 5Y CDS spread evolution for a set of European financial institutions ordered by rating grades over the period 2007-2010.

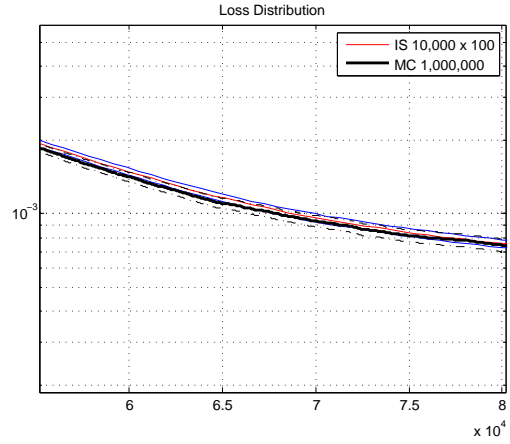
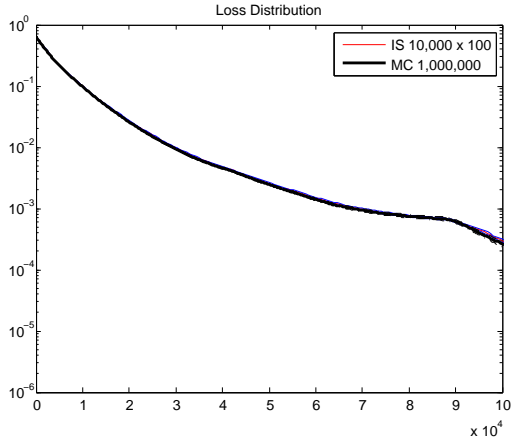


Figure 1.11: Loss distribution under the market mode. The black line shows the results of the Monte Carlo (MC) method using 1,000,000 simulations. The red and blue lines show, respectively, the importance sampling (IS) estimates and their 5%-95% confidence intervals using 10,000 macroeconomic scenarios and 100 default simulations on each macroeconomic scenario.

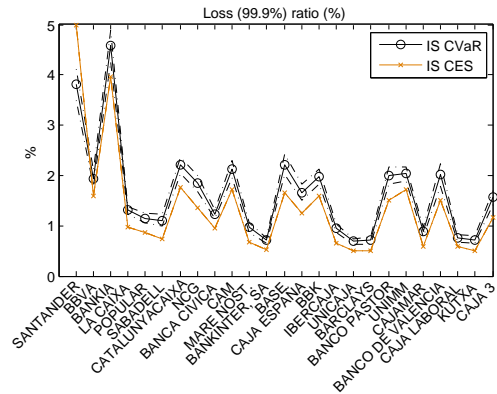
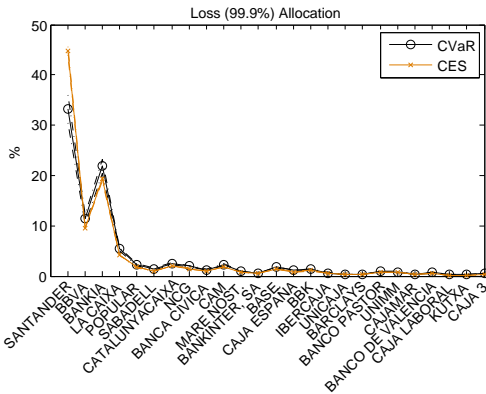


Figure 1.12: Risk allocation under market valuation for the VaR (CVaR) and ES (CES) criteria. Continuous and dashed lines represent, respectively, the IS estimates and the 5%-95% confidence intervals. Left and right graphs show, respectively, the total risk allocation and the allocated risk relative to the institution size.

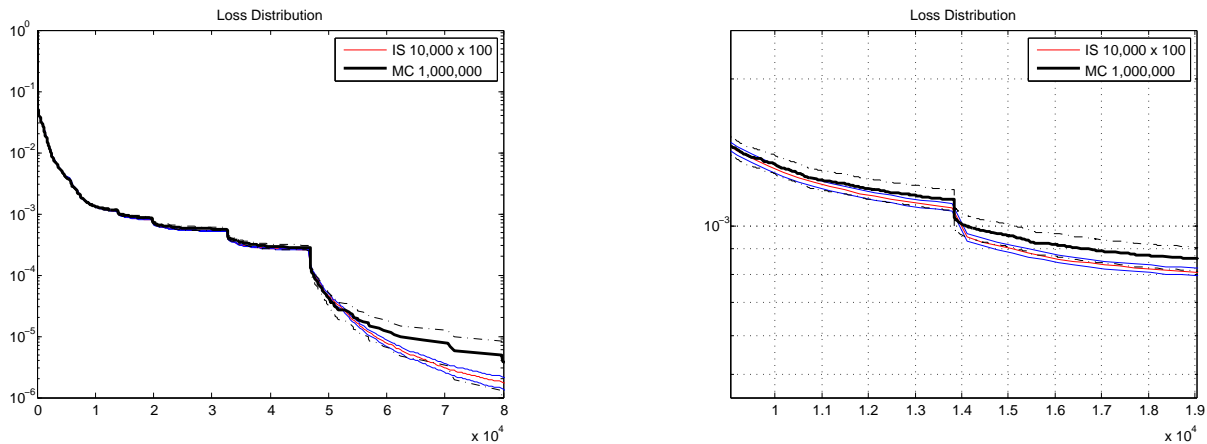


Figure 1.13: Loss distribution of the portfolio at December 2007 using 10,000 x 100 importance sampling (IS) and 1,000,000 Monte Carlo (MC) simulations. The black and red lines show, respectively, the Monte Carlo and IS results while the blue lines indicate the 5%-95% confidence interval of the IS estimates. Left and right graphs show, respectively, the tail distribution and its detail in the neighborhood of the 99.9% probability loss level.

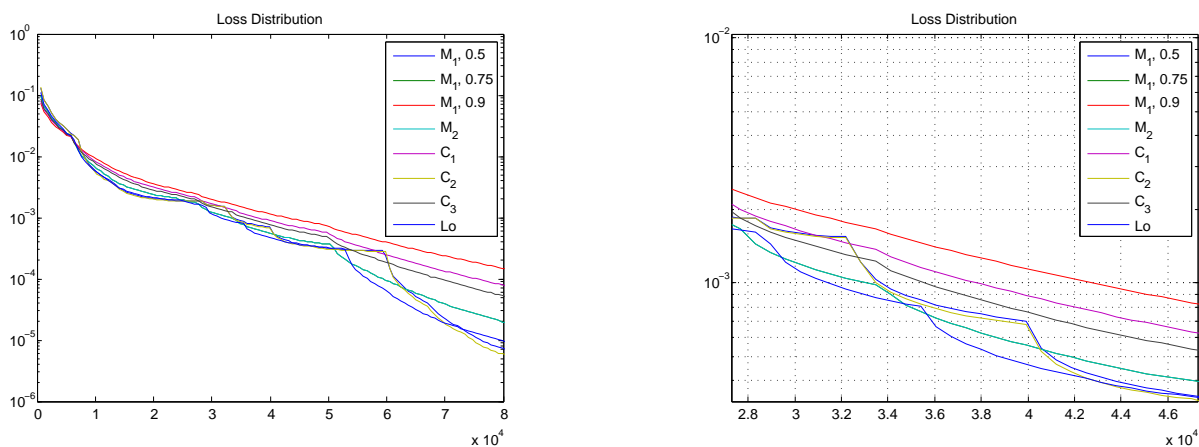


Figure 1.14: Loss distribution of the portfolio at December 2010 under the different macroeconomic sensitivity parameter estimates and using 10,000 x 100 importance sampling (IS) and 1,000,000 Monte Carlo (MC) simulations. Left and right graphs show, respectively, the tail distribution and its detail in the neighborhood of the 99.9% probability loss level.

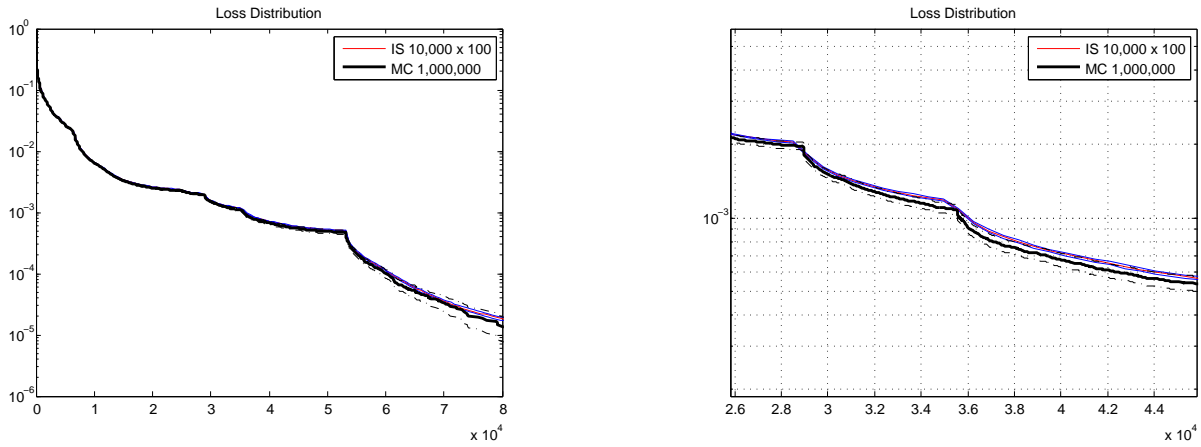


Figure 1.15: Loss distribution of the portfolio at December 2010 under *PD* uncertainty and using 10,000 x 100 importance sampling (IS) and 1,000,000 Monte Carlo (MC) simulations. The black and red lines show, respectively, the Monte Carlo and IS results while the blue lines indicate the 5%-95% confidence 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.

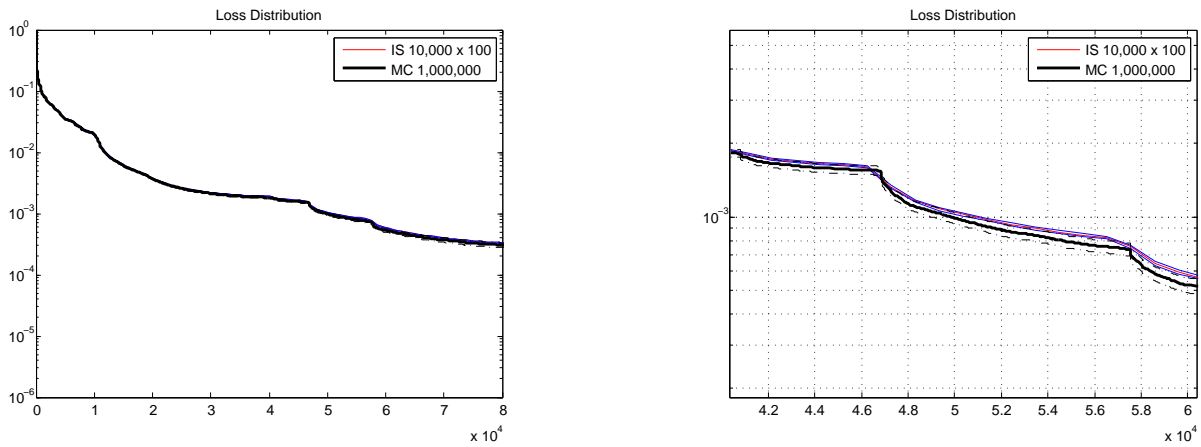


Figure 1.16: Loss distribution of the portfolio at December 2010 under *LGD* uncertainty and using 10,000 x 100 importance sampling (IS) and 1,000,000 Monte Carlo (MC) simulations. The black and red lines show, respectively, the Monte Carlo and IS results while the blue lines indicate the 5%-95% confidence 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.

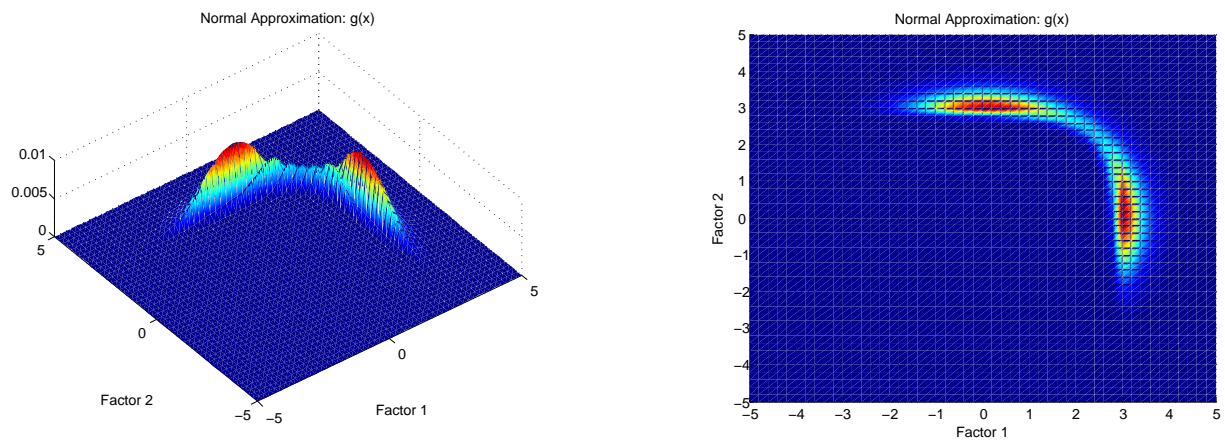


Figure 1.17: Function $g(x)$ provided by the normal approximation considering a portfolio of 1,000 counterparties and the parameters $PD = 1\%$, $LGD = 40\%$, $EAD = 1,000$, and $\alpha = 55\%$. Each half of the portfolio is exposed to a certain macroeconomic factor.

Chapter 2

Extended saddlepoint methods for credit risk measurement

2.1 Introduction

This paper reviews and extends the saddlepoint methods currently available to measure credit risk. We propose an approximate saddlepoint based method to allocate credit risk that, according to our results, performs very well and even better than other alternatives currently available as, for instance, Martin and Thompson (2001). We do also modify the default mode valuation saddlepoint method to deal with random recoveries and market valuation. All these modifications are tested for a portfolio including the Spanish financial institutions and compared with the results of a Monte Carlo importance sampling method (IS) as a benchmark.

The credit risk loss distribution of a portfolio can be obtained in many ways being the Monte Carlo simulation the simplest one, although it can be quite time consuming. As a consequence, methods such as the importance sampling introduced in Glasserman and Li (2005) to speed up the calculations have arisen. Other authors have developed approximate methods to estimate the loss distribution of a portfolio, see Pykhtin (2004) or Voropaev (2011). The saddlepoint approach is among these approximate alternatives. In more detail, this approach aims to approximate the inverse Laplace transform of a certain function and, then, can be applied in many fields as statistics and finance. In statistics it has been mainly used to approximate the distribution of the sum of independent random variables, see Daniels (1954, 1980, 1987) or Reid (1988). More recently, the saddlepoint method has been used to estimate the loss distribution of a loan portfolio, see Martin and Thompson (2001) and Huang et al. (2007), among others.

Pure or importance sampling based Monte Carlo methods rely on simulations to estimate the loss distribution and, hence, are exposed to sampling noise. In some cases, such as random idiosyncratic recoveries, the confidence intervals of the Monte Carlo estimates get reduced very slowly as the number of simulations increases (see García-Céspedes and Moreno (2014)). The saddlepoint method proposed in Huang et al. (2007) is a semi-analytic technique that can be a reasonable candidate to

generate more stable estimates than a pure simulation one.

We show that the classic saddlepoint methods can be modified in a very intuitive way to deal with pure macroeconomic random recoveries and market valuation. However in the case of idiosyncratic random recoveries the saddlepoint methods can not be extended in a simple way and other methods to obtain the inverse Laplace transform are required, see Abate and Whitt (1995).

This paper provides two major computational contributions to the literature. First, we propose a risk allocation method based on the saddlepoint approach that, according to our results, is faster and more accurate than other approaches previously analyzed in the literature. Second, we extend the classic default mode saddlepoint model to deal with more realistic assumptions such as random recoveries and market valuation.

Finally, for illustrative purposes, we consider the portfolio of the Spanish financial institutions to test the accuracy of our extensions. In this way we quantify the risk of this financial system as a whole and allocate it over the different financial entities. According to our results, our risk allocation approximate method is more accurate than other alternatives such as that in Martin and Thompson (2001) and requires a similar computational time. Compared with an exact risk allocation method, we generate similar results but at a time cost that is independent on the number of counterparties in the portfolio. The results for the random recoveries and the market valuation extensions are also very close to those obtained with the Monte Carlo benchmark. However in this last case the number of calculations increases considerably and the IS method can be more appropriate.

This paper is organized as follows. Section 2 introduces the first ideas of density and cumulative distribution approximations using saddlepoint methods. Section 3 describes how to use saddlepoint methods to measure credit risk and presents the new method to allocate risk based on the Hermite polynomials and the random loss given default and market valuation extensions. Section 4 applies the saddlepoint methods and our extensions to the portfolio of the Spanish financial institutions. This Section shows how to calibrate the model parameters and how to obtain the portfolio loss distributions and risk allocations. Finally, Section 5 summarizes our main results and concludes.

2.2 Saddlepoint methods

2.2.1 Density approximations

Let us define the moment and cumulant generating functions of a random variable x as $M_x(t) = E(e^{xt})$ and $K_x(t) = \ln(M_x(t))$, respectively. According to this, the moment generating function of the sum of random variables $L = \sum_{j=1}^M x_j$ is $M_L(t) = E(e^{Lt})$. For independent random variables we have that $M_L(t) = \prod_{j=1}^M M_{x_j}(t)$ and $K_L(t) = \sum_{j=1}^M K_{x_j}(t)$.

Replicating the steps in Daniels (1987) let us define $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$. Using the cumulant gener-

ating function of x , $K_x(t)$, we get the density function as the inverse Laplace transform

$$f_{\bar{x}}(\bar{x}) = \frac{n}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{nK_x(t)-nt\bar{x}} dt$$

Expanding the exponent in the integrand around a point \hat{t} such that $K'_x(\hat{t}) = \bar{x}$, we have that

$$f_{\bar{x}}(\bar{x}) = \frac{ne^{n(K_x(\hat{t})-\hat{t}\bar{x})}}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{\sum_{q=2}^{\infty} \frac{n}{q!} K_{qx}(\hat{t})(t-\hat{t})^q} dt$$

where $K_{qx}(t) = \frac{d^q K_x(t)}{dt^q}$. Defining $V = \sqrt{nK_{2x}(\hat{t})(t-\hat{t})}$, some computations lead to

$$f_{\bar{x}}(\bar{x}) = \frac{ne^{n(K_x(\hat{t})-\hat{t}\bar{x})}}{2\pi i} \int_{-i\infty}^{+i\infty} e^{\frac{V^2}{2}} \frac{A_1}{\sqrt{nK_{2x}(\hat{t})}} dV$$

where

$$A_1 = \sum_{i=0}^{\infty} \frac{1}{i!} \left(\sum_{j=3}^{\infty} \frac{1}{n^{j/2-1} j!} \lambda_j V^j \right)^i$$

with $\lambda_j = K_{jx}(\hat{t})(K_{2x}(\hat{t}))^{-j/2}$. Using that $\int_{-\infty}^{+\infty} i^k V^k e^{-\frac{V^2}{2}} dV = 0$ for k odd and applying some algebra, we get

$$f_{\bar{x}}(\bar{x}) = \frac{ne^{n(K_x(\hat{t})-\hat{t}\bar{x})}}{2\pi} \int_{-\infty}^{+\infty} e^{-\frac{V^2}{2}} \left(1 + \frac{B_1}{n} + \frac{C_1}{n^2} + \dots \right) \frac{1}{\sqrt{nK_{2x}(\hat{t})}} dV$$

where

$$\begin{aligned} B_1 &= \frac{1}{24} \left(\lambda_4 V^4 - \frac{\lambda_3^2 V^6}{3} \right) \\ C_1 &= \frac{1}{144} \left(-\frac{\lambda_6 V^6}{5} + \frac{\lambda_4^2 V^8}{8} - \frac{\lambda_3^2 \lambda_4 V^{10}}{12} \right) \end{aligned}$$

Previous integrals can be obtained using the statistical moments of a standard normal variable. Then, considering the terms up to n^{-2} , we get¹

$$f_{\bar{x}}(\bar{x}) \approx \frac{\sqrt{n} e^{n(K_x(\hat{t})-\hat{t}\bar{x})}}{\sqrt{2\pi K_{2x}(\hat{t})}} \left[1 + \frac{1}{8n} \left(\lambda_4 - \frac{5\lambda_3^2}{3} \right) + \frac{5}{48n^2} \left(-\frac{\lambda_6}{5} + \frac{7\lambda_4^2}{8} - \frac{21\lambda_3^2 \lambda_4}{4} \right) \right] \quad (2.1)$$

2.2.2 Cumulative distribution approximations

One way to obtain the cumulative probability distribution $F_x(x)$ is to integrate the saddlepoint approximation of $f_x(x)$ in (2.1). However this is not optimum because the density function $f_x(x)$

¹Daniels (1987) only shows the approximation up to terms in n^{-1} .

must be integrated over a region but the approximation (2.1) is accurate just around a given point x . The cumulative probability function can be obtained as

$$1 - F_x(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{e^{K_x(t)-tx}}{t} dt, \quad c > 0 \quad (2.2)$$

In a similar way to the previous Subsection, Daniels (1987) applies a Taylor expansion to expression (2.2) and approximates the cumulative distribution by

$$P(\bar{X} > \bar{x}) = \frac{e^{n(K_x(\hat{t})-\hat{t}\bar{x})}}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{\frac{n}{2}K_{2x}(\hat{t})(t-\hat{t})^2} \frac{A_2}{t} dt$$

with

$$A_2 = \sum_{i=0}^{\infty} \frac{n^i}{i!} \left(\sum_{j=3}^{\infty} \frac{1}{j!} K_{jx}(\hat{t})(t-\hat{t})^j \right)^i$$

Defining $z = t\sqrt{nK_{2x}(\hat{t})}$ and $\hat{z} = \hat{t}\sqrt{nK_{2x}(\hat{t})}$ provides

$$P(\bar{X} > \bar{x}) = \frac{e^{n(K_x(\hat{t})-\hat{t}\bar{x})+\frac{\hat{z}^2}{2}}}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{\frac{1}{2}z^2-\hat{z}z} \frac{A_3}{z} dz$$

with

$$A_3 = \sum_{i=0}^{\infty} \frac{1}{i!} \left(\sum_{j=3}^{\infty} \frac{1}{n^{j/2-1}j!} \lambda_j (z-\hat{z})^j \right)^i$$

Reordering we have

$$\begin{aligned} P(\bar{X} > \bar{x}) \approx & e^{n(K_x(\hat{t})-\hat{t}\bar{x})+\frac{\hat{z}^2}{2}} \left[I_0 + \frac{\lambda_3 I_3}{6n^{0.5}} + \frac{1}{n} \left(\frac{\lambda_4 I_4}{24} + \frac{\lambda_3^2 I_6}{72} \right) \right. \\ & \left. + \frac{1}{n^{1.5}} \left(\frac{\lambda_5 I_5}{5!} + \frac{\lambda_3^3 I_9}{3!^4} \right) + \frac{1}{n^2} \left(\frac{\lambda_6 I_6}{6!} + \frac{\lambda_4^2 I_8}{2!4!^2} + 3 \frac{\lambda_4 \lambda_3^2 I_{10}}{4!3!^3} \right) \right] \end{aligned}$$

where

$$I_r = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{(z-\hat{z})^r}{z} e^{\frac{1}{2}z^2-\hat{z}z} dz, \quad r = 0, 1, 2, \dots$$

We have that $I_0 = P(z > \hat{z}) = 1 - \Phi(\hat{z})$ where $\Phi(\cdot)$ indicates the distribution function of a standard normal variable.² Integration by parts provides the recursive relation $I_r = i^{r-1} \phi(\hat{z}) M_{r-1} - \hat{z} I_{r-1}$ where $\phi(\cdot)$ and M_i denote, respectively, the density function and the i -th moment of a standard normal variable. Using this relation we have the next approximation for the cumulative probability

²The cumulant generating function of a normal random variable $x \sim N(\mu, \sigma)$ is $K_x(t) = \mu t + \frac{1}{2} \sigma^2 t^2$.

distribution up to the term n^{-1} :

$$\begin{aligned}
P(\bar{X} > \bar{x}) &\approx e^{n(K_x(\hat{t}) - \hat{t}\bar{x}) + \frac{\hat{z}^2}{2}} \left[I_0 + \frac{\lambda_3 I_3}{6\sqrt{n}} + \frac{1}{n} \left(\frac{\lambda_4 I_4}{24} + \frac{\lambda_3^2 I_6}{72} \right) \right] \\
&= e^{n(K_x(\hat{t}) - \hat{t}\bar{x}) + \frac{\hat{z}^2}{2}} \left[(1 - \Phi(\hat{z})) \left(1 - \frac{\lambda_3 \hat{z}^3}{6\sqrt{n}} + \frac{1}{n} \left(\frac{\lambda_4 \hat{z}^4}{24} + \frac{\lambda_3^2 \hat{z}^6}{72} \right) \right) \right. \\
&\quad \left. + \phi(\hat{z}) \left(\frac{\lambda_3(\hat{z}^2 - 1)}{6\sqrt{n}} - \frac{1}{n} \left(\frac{\lambda_4(\hat{z}^3 - \hat{z})}{24} + \frac{\lambda_3^2(\hat{z}^5 - \hat{z}^3 + 3\hat{z})}{72} \right) \right) \right]
\end{aligned}$$

Due to some technicalities the final saddlepoint approximation is

$$P(X > x) \approx \begin{cases} e^{K_x(\hat{t}) - \hat{t}x + \frac{\hat{z}^2}{2}} [(1 - \Phi(\hat{z}))B_2 + \phi(\hat{z})C_2] & X > E(x) \\ \frac{1}{2} & X = E(x) \\ 1 - e^{K_x(\hat{t}) - \hat{t}x + \frac{\hat{z}^2}{2}} [(1 - \Phi(\hat{z}))B_2 + \phi(\hat{z})C_2] & X < E(x) \end{cases} \quad (2.3)$$

with

$$\begin{aligned}
B_2 &= 1 - \frac{\lambda_3 \hat{z}^3}{6} + \frac{\lambda_4 \hat{z}^4}{24} + \frac{\lambda_3^2 \hat{z}^6}{72} \\
C_2 &= \frac{\lambda_3(\hat{z}^2 - 1)}{6} - \frac{\lambda_4(\hat{z}^3 - \hat{z})}{24} - \frac{\lambda_3^2(\hat{z}^5 - \hat{z}^3 + 3\hat{z})}{72}
\end{aligned}$$

2.3 Credit risk and saddlepoint approximation

2.3.1 Saddlepoint method for credit risk

According to the Vasicek (1987) model the default behavior of the client j is driven by a set of random macroeconomic factors $Z = \{z_1, z_2, \dots, z_k\}$ and an idiosyncratic random term ε_j . This idiosyncratic term is a client specific random variable. The factors $\{z_i\}_{i=1}^k$ and ε_j are independent and distributed as standard normal random variables. Given this specification, the default of this client is modeled through the asset value variable V_j , defined as

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

where the terms $\alpha_{f,j}$ capture the macroeconomic sensitivity of the asset value. Then, the client j defaults if V_j falls below a threshold level k . By construction $V_j \sim N(0, 1)$ and k equates $\Phi^{-1}(PD_{j,C})$, where $PD_{j,C}$ denotes the historical average default probability of the client j .

The total loss of a portfolio including M contracts or clients is given as $L = \sum_{j=1}^M x_j$, where x_j is the individual loss of the client or contract j . According to this the portfolio loss can be expressed using the exposure at default (EAD_j) and the loss given default (LGD_j) parameters of each client 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 case of default mode models with constant LGD , the cumulants of the loss distribution conditional to a macroeconomic scenario can be obtained in a very simple way. Conditional to a given macroeconomic scenario, the defaults are independent and the cumulants of the sum of independent random variables are given by the sum of the cumulants of each random variable. Then, assuming that the portfolio is made up of M clients with default probability $PD_j(Z)$ conditional to the scenario Z , the cumulant generating function of the loss distribution $L = \sum_{j=1}^M x_j$ is given by $K_x(t) = \sum_{j=1}^M k_x(j, t)$ with $k_x(j, t) = \ln(1 - PD_j(Z) + PD_j(Z)e^{tw_j})$ where $w_j = EAD_j \times LGD_j$. The derivatives of $K_x(t)$ can be easily computed up to any order by applying a recursive relation.³

Based on these cumulants, an approximation for $P\left(\sum_{j=1}^M x_j > l \mid Z\right)$ can be obtained for different values of Z and then empirically estimate $E\left(P\left(\sum_{j=1}^M x_j > l \mid Z\right)\right)$. Under the usual credit risk framework, the scenarios Z are distributed as $N(0, \Sigma)$ and can be simulated using Monte Carlo techniques.

For heavily concentrated portfolios the usual saddlepoint approximation is not accurate and a modification is required. This modification is called *adaptative saddlepoint approximation* and was proposed in Huang et al. (2007). Consider a portfolio where the firms A and B are the two biggest counterparties in terms of the final loss given default ($EAD \times LGD$). Let x_A and x_B be the $LGDs$ (in currency units) of these firms and consider $l \in (x_B, x_A)$. Then, $P(L > x|Z)$ can be obtained as

$$\begin{aligned}
P\left(\sum_{j=1}^M x_j \geq l \mid Z\right) &= 1 - P\left(\sum_{j=1}^M x_j < l \mid Z\right) \\
&= 1 - (1 - PD_A(Z))P\left(\sum_{j \neq A}^M x_j < l \mid Z\right) \\
&= 1 - (1 - PD_A(Z))\left(1 - P\left(\sum_{j \neq A}^M x_j \geq l \mid Z\right)\right) \\
&= 1 - (1 - PD_A(Z))(1 - Sad(L_A, l, Z))
\end{aligned} \tag{2.4}$$

where L_A represents the loss of the portfolio excluding the client A and $Sad(L_A, l, Z)$ stands for the saddlepoint approximation of the cumulative probability of L_A for a loss level l and conditional to a macroeconomic scenario Z . If the loss level is just below x_B , expression (2.4) becomes

$$P\left(\sum_{j=1}^M x_j \geq l \mid Z\right) = 1 - (1 - PD_A(Z))(1 - PD_B(Z))(1 - Sad(L_{A,B}, l, Z))$$

In a similar way, the density distribution can be written as $f_L(l|Z) = (1 - PD_A)Sad_f(L_A, l, Z)$.

³Details are available upon request. It can be observed that the first derivative is equivalent to the exponential twist of the conditional default probabilities proposed in the IS framework by Glasserman and Li (2005).

2.3.2 Risk contributions

The two most extended ways to allocate the risk of a portfolio among its different counterparties are the Value-at-Risk (VaR) contribution ($CVaR$) and the expected shortfall (ES) contribution (CES). Both measures indicate, respectively, the contribution of a client to a given percentile and to a given conditional moment of the loss distribution. The main reason for considering a moment rather than a percentile is that, under the usual Monte Carlo framework, the stability of the CES is much higher than that of the $CVaR$.

According to the $CVaR$ criterion, the risk contribution of the client j is

$$\begin{aligned}
 CVaR_j &= E \left(x_j \left| \sum_{j=1}^M x_j = VaR \right. \right) = \frac{E(x_j \mathbf{1}(\sum_{j=1}^M x_j = VaR))}{f(\sum_{j=1}^M x_j = VaR)} \\
 &= w_j \frac{E(f(D_j = 1, L = VaR|Z))}{E(f(L = VaR|Z))} \\
 &= w_j \frac{E(f(L_j = VaR - w_j|Z)PD_j(Z))}{E(f(L = VaR|Z))} \tag{2.5}
 \end{aligned}$$

where $L_j = \sum_{k \neq j}^M x_k$ is the loss of the total portfolio excluding the client j .

For the ES contribution, a similar formula can be derived

$$CES_j = w_j \frac{E(P(L_j > VaR - w_j|Z)PD_j(Z))}{E(P(L_j > VaR|Z))} \tag{2.6}$$

Then if we want to use saddlepoint methods we need to approximate $f(L_j = VaR - w_j|Z)$ and $P(L_j > VaR - w_j|Z)$ for each client j . This requires to estimate a different saddlepoint for each client and macroeconomic scenario, a very time demanding procedure.

To obtain the $CVaR$ of each client, an alternative way that avoids this client by client estimation is proposed in Martin and Thompson (2001) who suggested a formula that depends on the derivatives of the cumulant generating function of the whole portfolio and that, after some algebra, becomes

$$CVaR_j \approx \frac{E \left(f(L = VaR|Z) w_j \frac{PD_j(Z) e^{w_j \hat{t}}}{1 - PD_j(Z) + PD_j(Z) e^{w_j \hat{t}}} \right)}{E(f(L = VaR|Z))}$$

However, some papers as Huang et al. (2007) have suggested that this approximate formula may be not accurate enough. Then, we have decided to take a different approach. From now on, we consider that the saddlepoint of the whole portfolio loss L in the point $l = VaR$ conditional to a given macroeconomic scenario, \hat{t}_Z , should be close to $\hat{t}_{j,Z}$, the saddlepoint of the portfolio once the client j is removed in the point $l = VaR - w_j$.⁴

⁴This is a reasonable assumption if the portfolio is not very concentrated.

As a result we try to approximate the density and distribution functions of L_j around $l = VaR - w_j$ using the value \hat{t}_Z rather than $\hat{t}_{j,Z}$. Of course, as \hat{t}_Z is not the real saddlepoint of the portfolio loss L_j in $l = VaR - w_j$, some adjustments will be required in formulas (2.1) and (2.3).

Tsao (1995) provided the detailed steps to obtain the adjustments to extend expression (2.1) and approximate $f(L_j = VaR - w_j|Z)$ around a general point \hat{t}_Z . Using a Taylor expansion of the exponent in the integral, we know that

$$\begin{aligned}
f_x(x) &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{K_x(t)-tx} dt \\
&= \frac{e^{K_x(t_Z)-t_Z x}}{2\pi} \int_{-\infty}^{\infty} e^{(K'_x(t_Z)-x)it+1/2K_{2x}(t_Z)(it)^2+1/6K_{3x}(t_Z)(it)^3+\dots} dt \\
&= \frac{e^{K_x(t_Z)-t_Z x}}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{K_{2x}(t_Z)}{2}(t-ci)^2-\frac{(K'_x(t_Z)-x)^2}{2K_{2x}(t_Z)}+1/6K_{3x}(t_Z)(it)^3+\dots} dt \\
&= \frac{e^{K_x(t_Z)-t_Z x-\frac{(K'_x(t_Z)-x)^2}{2K_{2x}(t_Z)}}}{2\pi} \int_{-\infty}^{\infty} e^{-\frac{K_{2x}(t_Z)}{2}(t-ci)^2+1/6K_{3x}(t_Z)(it)^3+\dots} dt
\end{aligned}$$

where $c = \frac{K'_x(t_Z)-x}{K_{2x}(t_Z)}$. Defining $V = \sqrt{K_{2x}(t_Z)}(t-ci)$, expanding the exponential function, and working with the imaginary terms provides

$$\begin{aligned}
f_x(x) &= \frac{e^{K_x(t_Z)-t_Z x-\frac{(K'_x(t_Z)-x)^2}{2K_{2x}(t_Z)}}}{\sqrt{2\pi K_{2x}(t_Z)}} \int_{-\infty-ic\sqrt{K''_x(t_Z)}}^{\infty-ic\sqrt{K_{2x}(t_Z)}} \phi(V) \\
&\quad \times \left(1 + \sum_{j=3}^{\infty} \frac{\lambda_{jx}(t_Z)}{j!} (iV - c\sqrt{K_{2x}(t_Z)})^j + \frac{1}{2} \left(\sum_{j=3}^{\infty} \frac{\lambda_{jx}(t_Z)}{j!} (iV - c\sqrt{K_{2x}(t_Z)})^j \right)^2 + \dots \right) dV
\end{aligned}$$

As the integrand is an analytic function, the integration limits can be changed to $\pm\infty$. Moreover, we have

$$\int_{-\infty}^{\infty} \phi(V)(iV - x)^k dV = (-1)^k H_k(x)$$

where $H_k(x)$ is the k -th order Hermite polynomial.⁵

Finally we have

$$\begin{aligned}
f_x(x) &= \frac{e^{K_x(t_Z)-t_Z x-\frac{(K'_x(t_Z)-x)^2}{2K_{2x}(t_Z)}}}{\sqrt{2\pi K_{2x}(t_Z)}} \left[1 - \frac{\lambda_{3x}(t_Z)}{6} H_3 \left(\frac{K'(t_Z) - x}{\sqrt{K''(t_Z)}} \right) \right. \\
&\quad \left. + \frac{\lambda_{4x}(t_Z)}{24} H_4 \left(\frac{K'(t_Z) - x}{\sqrt{K''(t_Z)}} \right) + \frac{\lambda_{3x}(t_Z)^2}{72} H_6 \left(\frac{K'(t_Z) - x}{\sqrt{K''(t_Z)}} \right) + \dots \right] \quad (2.7)
\end{aligned}$$

⁵See Abramowitz and Stegun (1964) for details.

If t_Z is such that $K'_x(t_Z) = x$ then expression (2.7) equates (2.1). Moreover, if $t_Z = 0$, we get the (well-known) Edgeworth expansion

$$\begin{aligned} f(x) &= \frac{e^{-\frac{(K'_x(0)-x)^2}{2K_{2x}(0)}}}{\sqrt{2\pi K_{2x}(0)}} \left(1 - \frac{\lambda_{3x}(0)}{6} H_3 \left(\frac{K'(0) - x}{\sqrt{K''(0)}} \right) \right. \\ &\quad \left. + \frac{\lambda_{4x}(0)}{24} H_4 \left(\frac{K'(0) - x}{\sqrt{K''(0)}} \right) + \frac{\lambda_{3x}(0)^2}{72} H_6 \left(\frac{K'(0) - x}{\sqrt{K''(0)}} \right) + \dots \right) \\ &= \phi_{\mu,\sigma}(x) \left(1 + \frac{\lambda_3}{6} H_3 \left(\frac{x - \mu}{\sigma} \right) + \frac{\lambda_4}{24} H_4 \left(\frac{x - \mu}{\sigma} \right) + \frac{\lambda_3^2}{72} H_6 \left(\frac{x - \mu}{\sigma} \right) + \dots \right) \end{aligned}$$

Expression (2.7) can be obtained in a more intuitive way using the Edgeworth expansion instead of a Taylor expansion. The saddlepoint approximation of the density function can be seen as an Edgeworth expansion of a transformed distribution $g(x) = f(x)e^{-K_x(t_Z)+t_Zx}$, this transformation is known as exponential tilting.⁶ The Edgeworth approximation of $g(x)$ is

$$\begin{aligned} g(x) &= \frac{e^{-\frac{(K'_x(t_Z)-x)^2}{2K_{2x}(t_Z)}}}{\sqrt{2\pi K_{2x}(t_Z)}} \left[1 - \frac{\lambda_{3x}(t_Z)}{6} H_3 \left(\frac{K'(t_Z) - x}{\sqrt{K''(t_Z)}} \right) \right. \\ &\quad \left. + \frac{\lambda_{4x}(t_Z)}{24} H_4 \left(\frac{K'(t_Z) - x}{\sqrt{K''(t_Z)}} \right) + \frac{\lambda_{3x}(t_Z)^2}{72} H_6 \left(\frac{K'(t_Z) - x}{\sqrt{K''(t_Z)}} \right) + \dots \right] \end{aligned}$$

The Edgeworth approximation of $g(x)$ has the same structure as that of $f(x)$ but uses the cumulants of the original function $f(x)$ evaluated at a different point.⁷ Hence, $f(x)$ can be recovered as

$$\begin{aligned} f(x) &= e^{K_x(t_Z)-t_Zx} \frac{e^{-\frac{(K'_x(t_Z)-x)^2}{2K_{2x}(t_Z)}}}{\sqrt{2\pi K_{2x}(t_Z)}} \left[1 - \frac{\lambda_{3x}(t_Z)}{6} H_3 \left(\frac{K'(t_Z) - x}{\sqrt{K''(t_Z)}} \right) \right. \\ &\quad \left. + \frac{\lambda_{4x}(t_Z)}{24} H_4 \left(\frac{K'(t_Z) - x}{\sqrt{K''(t_Z)}} \right) + \frac{\lambda_{3x}(t_Z)^2}{72} H_6 \left(\frac{K'(t_Z) - x}{\sqrt{K''(t_Z)}} \right) + \dots \right] \end{aligned}$$

that equates expression (2.7).

For the *CES* we need to approximate $F(x)$ (or $1 - F(x)$) around a point $t = t_Z$. To this aim, Daniels (1987) uses the Edgeworth expansion of an exponentially tilted distribution. Reproducing

⁶The distribution $g(x)$ integrates one with mean and variance given as $K'_x(t_Z)$ and $K_{2x}(t_Z)$, respectively. See Reid (1988) for details.

⁷It can be said that the point $t = 0$ under $f(x)$ is equivalent to the point $t = \hat{t}$ under $g(x)$.

his steps we have that

$$\begin{aligned}
1 - F(x) &= \int_x^\infty f(y) dy \\
&= e^{K(t_Z) - t_Z x} \int_x^{+\infty} e^{-t_Z(y-x)} f(y) e^{t_Z y - K(t_Z)} dy \\
&= e^{K(t_Z) - t_Z x} \int_x^{+\infty} \frac{e^{-t_Z(y-x)}}{\sqrt{K''(t_Z)}} \phi\left(\frac{K'(t_Z) - y}{\sqrt{K''(t_Z)}}\right) \left[1 - \frac{\lambda_{3x}(t_Z)}{6} H_3\left(\frac{K'(t_Z) - y}{\sqrt{K''(t_Z)}}\right) \right. \\
&\quad \left. + \frac{\lambda_{4x}(t_Z)}{24} H_4\left(\frac{K'(t_Z) - y}{\sqrt{K''(t_Z)}}\right) + \frac{\lambda_{3x}(t_Z)^2}{72} H_6\left(\frac{K'(t_Z) - y}{\sqrt{K''(t_Z)}}\right) + \dots \right] dy \\
&= e^{K(t_Z) - t_Z x} \int_{-\infty}^v e^{-t_Z(K'(t_Z) - x - V\sqrt{K''(t_Z)})} \phi(V) \\
&\quad \times \left[1 - \frac{\lambda_{3x}(t_Z)}{6} H_3(V) + \frac{\lambda_{4x}(t_Z)}{24} H_4(V) + \frac{\lambda_{3x}(t_Z)^2}{72} H_6(V) + \dots \right] dV \\
&= e^{K(t_Z) - t_Z K'(t_Z)} \int_{-\infty}^v e^{t_Z(V\sqrt{K''(t_Z)})} \phi(V) \\
&\quad \times \left[1 - \frac{\lambda_{3x}(t_Z)}{6} H_3(V) + \frac{\lambda_{4x}(t_Z)}{24} H_4(V) + \frac{\lambda_{3x}(t_Z)^2}{72} H_6(V) + \dots \right] dV \\
&= e^{K(t_Z) - t_Z K'(t_Z) + \frac{t_Z^2 K''(t_Z)}{2}} \\
&\quad \times \left[I_0(v) - \frac{\lambda_{3x}(t_Z)}{6} I_3(v) + \frac{\lambda_{4x}(t_Z)}{24} I_4(v) + \frac{\lambda_{3x}(t_Z)^2}{72} I_6(v) + \dots \right] \tag{2.8}
\end{aligned}$$

with

$$\begin{aligned}
v &= \frac{K'(t_Z) - x}{\sqrt{K''(t_Z)}} \\
I_i(v) &= \int_{-\infty}^v e^{-\frac{t_Z^2 K''(t_Z)}{2} + t_Z V \sqrt{K''(t_Z)}} \phi(V) H_i(V) dV = \int_{-\infty}^v \phi(V - a) H_i(V) dV
\end{aligned}$$

where $a = t_Z \sqrt{K''(t_Z)}$. As in the previous Section we have that $I_0(v) = \Phi(v - a)$ and integrating I_i by parts leads to the recursive relation $I_i(v) = -\phi(v - a) H_{i-1}(v) + a I_{i-1}(v)$ that can be very easily implemented in a computer.

Truncating equation (2.8) at $I_3(v)$ and using a point t_Z such that $K'(t_Z) = x$ provides the saddlepoint approximation in expression (2.3)

$$1 - F(x) \approx e^{(K(t_Z) - t_Z K'(t_Z)) + \frac{t_Z^2 K''(t_Z)}{2}} \left[(1 - \Phi(a)) \left(1 - \frac{a^3 \lambda_{3x}(t_Z)}{6}\right) + \phi(a) \frac{(a^2 - 1) \lambda_{3x}(t_Z)}{6} \right]$$

As before, if $t_Z = 0$, then $a = 0$, $I_0(v) = \Phi(v)$, and $I_i(v) = -\phi(v) H_{i-1}(v)$. Then, we get the Edgeworth expansion

$$1 - F(x) = \Phi(v) - \phi(v) \left(-\frac{\lambda_3}{6} H_2(v) + \frac{\lambda_4}{24} H_3(v) + \frac{\lambda_3^2}{72} H_5(v) + \dots \right)$$

The approximations in (2.7) and (2.8) are named Hermite polynomials based saddlepoint approximations (Sad-Her). Using these formulas the estimation of the risk contributions requires much less time as, for each macroeconomic scenario, we do not need to estimate a saddlepoint for each client and we can use a unique value \hat{t}_Z . Later, the empirical Section will show that these formulas provide a much better accuracy than the Martin's approximation and with similar results to the exact method.

2.3.3 Random *LGD* models

The loss given default (*LGD*) is usually considered to be constant although it can be a random variable. The two simplest ways to model *LGD* as a random variable are, respectively, constant or random *LGD* conditional to a given macroeconomic scenario, Z . We discuss now these two alternatives.

1. In the first case (LGD^C), the saddlepoint approximation method does not need any modification and, for each macroeconomic scenario, we have to use $LGD_i(Z)$ instead of LGD_i . This way to introduce the random *LGD* behavior implies no restriction on the functional form of the *LGD* that can be used. Table 2.1 includes several possible pure macroeconomic *LGD* models ($LGD^C|z$), where z represents a set of macroeconomic factors driving the recoveries behavior. A usual assumption is that these factors follow a standard normal distribution.

[INSERT TABLE 2.1 AROUND HERE]

Calibrating these pure macroeconomic *LGD* models is not complex, assuming a single macroeconomic *LGD* factor z_{LGD} with a standard normal distribution we can obtain expressions for the unconditional mean and variance that depend only on the parameters a and b in Table 2.1.⁸ Then using a time series of average *LGDs* ($LGD_{t_1}, LGD_{t_2}, \dots, LGD_{t_y}$) we can calibrate the parameters that best fit these two moments.

After estimating the parameters a and b we can recover the time series of the *LGD* and default driving macroeconomic factors, $\{z_{LGD,t_1}, z_{LGD,t_2}, \dots, z_{LGD,t_y}\}$ and $\{z_{PD,t_1}, z_{PD,t_2}, \dots, z_{PD,t_y}\}$ and, then, we can calibrate the correlation between both factors, $\rho_{z_{PD}, z_{LGD}}$.

2. In the case of the random conditional *LGD* (LGD^R), the saddlepoint approximation requires to consider the moment generating function of the random variable $LGD_i(Z)$ of each client in

⁸This can be done for all the distributions in Table 2.1 but for the logit and probit distributions. For these two cases a Monte Carlo method was implemented to calibrate the parameters.

the portfolio ($M_{LGD_i|Z}(EAD_i t)$). Now, the cumulants of the portfolio are

$$K_x(t) = \sum_{j=1}^M \ln[1 - PD_j(Z) + PD_j(Z)M_{LGD_j|Z}(EAD_j t)]$$

Then, not all the distributions are suitable to characterize the *LGD* behavior. The most common continuous random variables with known moment generating function are the normal, gamma, and beta distributions.⁹ Table 2.2 includes the different models that will be analyzed in the empirical Section.

[INSERT TABLE 2.2 AROUND HERE]

Under random conditional recoveries the calibration of the *LGD* model parameters is much more complex. For the gamma and beta distributions we sample numerically from these distributions to obtain the parameters a , b , and c in Table 2.2 that provide the best fit to the values $E(LGD^R)$, $Var(LGD^R)$, and $E(Var(LGD^R|z_{LGD}))$ of the real recoveries sample.¹⁰ As in the previous case, after estimating these parameters, we can recover a time series of $\{z_{LGD,t_1}, z_{LGD,t_2}, \dots, z_{LGD,t_y}\}$ and estimate $\rho_{z_{PD}, z_{LGD}}$, the correlation between the macroeconomic default driving factor and the recoveries driving factor.

The extension of the *adaptative saddlepoint approximation* gets slightly more complex. Now, for $l \in (x_B, x_A)$, the probability $P(L > l|Z)$ is given as

$$P\left(\sum_{j=1}^M x_j \geq l \mid Z\right) = 1 - \int_0^l f_A(y|Z)P\left(\sum_{j \neq A}^M x_j < l - y \mid Z\right) dy$$

where $f_A(y)$ stands for the loss distribution of a portfolio including just the counterparty A . This integral can be obtained approximating the probability inside the integrand using the saddlepoint method and applying numerical integration. However this method gets complex and slow and we think that other methods to obtain the inverse Laplace transform can provide better results. Reproducing the ideas in Abate and Whitt (1995) and Glasserman and Ruiz-Mata (2006) we have

$$1 - F_x(x) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{e^{K_x(t)-tx}}{t} dt = \frac{2e^{-cx}}{\pi} \int_0^\infty Re\left(\frac{e^{K_x(c+ui)}}{c+ui}\right) \cos(ux) du$$

⁹See the Appendix for further details.

¹⁰The calibration of the normal idiosyncratic *LGD* can be done without the Monte Carlo sampling process as its conditional and unconditional moments are straightforward.

The numerical approximation for this integral provides

$$1 - F_x(x) \approx \frac{e^{-cx}\Delta u}{\pi} \left[\operatorname{Re} \left(\frac{e^{K_x(c)}}{c} \right) + 2 \sum_{k=1}^N \operatorname{Re} \left(\frac{e^{K_x(c+ik\Delta u)}}{c + ik\Delta u} \right) \cos(xk\Delta u) \right] \quad (2.9)$$

Later, the empirical Section will show that this method to obtain the inverse Laplace transform is very accurate once the model parameters have been carefully chosen. However, the time required to obtain the numerical integral is much higher than that of the saddlepoint methods.

In the case of the risk allocation under random conditional recoveries, the number of calculations needed to implement the *VaR* and *ES* based risk contribution formulas increases considerably and the time advantage of the approximate methods completely disappears. Then, we decide to stop here the methodological analysis of the saddlepoint and inverse Laplace transform methods for the default mode models. The next Subsection will extend the saddlepoint method to the case of market valuation.

2.3.4 Market mode models

Under a market valuation model the rating of a firm may change over time and these changes affect the firm value and can generate losses. Then we need to compute $V_{i,h}$, the value of the client i under any of the possible rating states h .

To build the state migration rule, let $MP_{j,C,IR,FR}$ denote the average (over the cycle) migration probability from an initial rating IR to a final rating state FR for the client j . Then we can compute $AccumMP_{j,Z,IR,FR}$, the accumulated probabilities for all the possible final rating states $FR = AAA, AA+, \dots, CCC, D$.¹¹

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

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

From these conditional accumulated probabilities we can recover the non accumulated ones. Under market valuation the cumulant generating function of the portfolio value function is given by

$$K_x(t) = \sum_{j=1}^M \ln \left(\sum_{h=1}^s MP_{j,Z,IR,h} e^{tV_{j,h}} \right) \quad (2.10)$$

We can work with the market mode value $V_{j,h}$ or, inversely, $x_{M,j,h}$, the losses of each client j at each state h . As we did in the default mode, using the cumulant generating function (2.10) we can

¹¹For example, $AccumMP_{j,Z,IR,B-} = MP_{j,Z,IR,B-} + MP_{j,Z,IR,CCC} + MP_{j,Z,IR,D}$.

approximate the loss distribution of the portfolio and the risk contributions. We need to extend the *adaptative saddlepoint approximation*. Focusing on the losses of the portfolio, we get

$$P\left(\sum_{j=1}^M x_{M,j,h} \geq l \mid Z\right) = 1 - P(x_{M,j,h} < l \mid Z) P\left(\sum_{j=1}^M x_{M,j,h} < l \mid x_{M,j,h} < l, Z\right) \quad (2.11)$$

As it can be seen, we have to condition on all the possible final states h where the losses of the client j do not exceed l . Recall that, under the default mode, for each client there was only one final state that could generate losses greater than l . Now, under the market valuation, there may exist several states that generate these high losses. We must also note that the expression (2.11) is true as long as losses are positive.¹²

Regarding the risk allocation, the *VaR* and *ES* contributions of the client j are given by

$$\begin{aligned} CVaR_i &= E\left(x_{M,j,h} \mid \sum_{j=1}^M x_{M,j,h} = VaR\right) = \frac{E(\sum_{h=1}^s x_{M,j,h} f(h_j|z) f(L_{M,j} = VaR - x_{M,j,h}|z))}{E(f(L_M = VaR|Z))} \\ CES_j &= \frac{E(\sum_{h=1}^s x_{M,j,h} f(h_j|z) f(L_{M,j} > VaR - x_{M,j,h}|z))}{E(P(L_M > VaR|Z))} \end{aligned}$$

with $L_M = \sum_{j=1}^M x_{M,j,h}$ and $L_{M,j} = \sum_{k \neq j}^M x_{M,k,h}$.

Note that the expression of *CVaR* and *CES* under the market valuation are very similar to those for the default mode but, in the numerator, we sum over all the possible final rating states and, hence, we have to estimate a saddlepoint for each client, final rating, and macroeconomic scenario. Therefore, the use of the Hermite polynomials based saddlepoint approximations becomes a very interesting tool as we must estimate just one saddlepoint \hat{t}_Z for each macroeconomic scenario and maintain it for all the different clients and final rating states.

2.4 Empirical analysis

This Section illustrates empirically the theoretical framework and the extensions that we have proposed in previous Sections. In more detail, we consider a portfolio that includes the Spanish financial institutions at December 2010 and obtain the loss distribution and the risk contribution of each entity. We start introducing this portfolio and describing the data sources that we have employed to obtain the main risk drivers of the portfolio. Later, we obtain the loss distribution and the risk allocation of these institutions for the case of constant and random *LGDs* and for market valuation losses. In all the cases, the results from a Monte Carlo importance sampling method are used as benchmark. Finally, we do also test the Hermite polynomials based saddlepoint method that we have proposed to allocate the risk.

¹²This is the case for losses on total notional or on best state valuation but not necessarily for losses on current state valuation.

2.4.1 Data

We analyze a portfolio that contains the 157 Spanish financial institutions that are covered by the Spanish deposit guarantee fund (FGD) at December, 2010. To obtain the loss distribution we first need to estimate the probability of default (PD), exposure at default (EAD), loss given default (LGD), and the macroeconomic factor sensitivity (α) of each institution. These parameters are estimated as follows:¹³

1. **PD** : We use the credit ratings available at December, 2010 for these institutions and the historical observed default rates reported by the rating agencies to infer a probability of default. Entities with no external rating are assigned one notch less than the average rating of the portfolio with external rating.
2. **EAD** : Details on assets, liabilities, and deposits for the FGD institutions are available in the AEB, CECA and AECR webpages.¹⁴ Balance information at December 2010 was used for the analysis. As many mergers took place during 2010, we have summed all the information from the different institutions that belong to the same group.
3. **LGD** : Bennet (2002) computed the losses due to financial institutions default in the FDIC (Federal Deposits Insurance Corporation) and showed that the average losses are bigger in the smallest banks for the period 1986-1998. We update this analysis up to 2009 using FDIC public data and the banks assets are updated using the USA CPI series aiming to have comparable asset sizes. We also use the individual historical data of defaulted financial institutions from the FDIC to calibrate the LGD models.
4. **α** : We consider geographic macroeconomic factors and use the sensitivities to these factors stated in the Basel accord. We assume that all the financial institutions are exposed to the Spanish factor and that the two biggest banks (BBVA and Santander) are exposed to additional geographies. The exposure of both banks to all these factors is computed using the reported net interest income by geography obtained from the 2010 public annual reports.

2.4.2 Results

- **Loss distribution of the portfolio**

Figure 2.1 shows the confidence intervals for the loss distribution of the Spanish financial system under pure Monte Carlo simulations, Monte Carlo IS, and saddlepoint methods.

¹³See García-Céspedes and Moreno (2014) for further details.

¹⁴AEB is the Spanish Bank Association, CECA is the Spanish Saving Bank Association and AECR is the Spanish Credit Cooperatives Association.

[INSERT FIGURE 2.1 AROUND HERE]

It can be seen that the accuracy of the saddlepoint approximation is quite bad for loss levels below 53,146 MM €, a value that indicates the *LGD* of Santander, the biggest institution in our sample. This result is expected as we are not using the *adaptive saddlepoint approximation*. Figure 2.2 shows the loss distribution considering this modification and illustrates that the accuracy of the saddlepoint approximation increases considerably. As in the previous Figure, we provide the results from the saddlepoint approximation with 1,000 macroeconomic scenarios¹⁵, a pure Monte Carlo method with 1,000,000 simulations, and a Monte Carlo IS method with 1,000 x 100 simulations. The 99.9% probability loss level is 31,352 MM€, a value that will be used in the next issues.

[INSERT FIGURE 2.2 AROUND HERE]

- **Risk contribution of each financial institution**

We test the accuracy of the Hermite polynomials based saddlepoint approximations (Sad-Her) (see equations (2.7)-(2.8)) to estimate the risk contributions (see equations (2.5)-(2.6)) for a loss level of 31,352 MM€. The results of these approximations are compared with those obtained using a pure Monte Carlo method (MC), the Monte Carlo IS method (IS), the exact saddlepoint approximation (Sad), and the method proposed in Martin and Thompson (2001) (Sad-Mar). Figures 2.3 and 2.4 show the results for the *VaR* and *ES* risk allocation simulating, respectively, 100 and 1,000 macroeconomic scenarios.

[INSERT FIGURES 2.3 AND 2.4 AROUND HERE]

Figure 2.3 shows that, for the *VaR* based risk allocation, all the methods except the Martin's one produce very similar risk allocations. This is the case even though the IS and the saddlepoint methods use a very small number of macroeconomic scenarios. It is worth to mention that the risk allocated to BBVA and Santander is null. This is because their *LGDs* are higher than 31,352 MM€ and, then, they can not have defaulted conditional to scenarios that generate a total portfolio loss of 31,352 MM€. For the *ES* based risk allocation, the results for the IS and the saddlepoint methods are very close to each other¹⁶ but very far from those of the

¹⁵The mean of the macroeconomic scenarios is changed and a weighting function is introduced according to the ideas in Glasserman and Li (2005). However, the shift in the mean of the macroeconomic variables is done according to the mean (and not to the mode) of the optimum distribution. See García-Céspedes and Moreno (2014) for further details.

¹⁶Both methods share the same macroeconomic scenarios in the simulation process.

MC method. This is because using just 100 simulations is not enough to estimate correctly the risk contributions.

Figure 2.4 provides the results once we simulate 1,000 macroeconomic scenarios and illustrates that the *ES* risk allocation is very similar along the different methods while maintaining the conclusions for the *VaR* allocation obtained from Figure 2.3.

Finally, using the Hermite polynomials based saddlepoint approximations, the speed increases up to ten times compared with the pure saddlepoint expressions that require to estimate a different saddlepoint for each client and the results of both methods are very similar. According to the portfolio analyzed, the results from the Martin's formula are not very accurate and the computational time employed is very similar to that required by the Hermite polynomials method. Hence, we suggest to use the pure saddlepoint method only when estimating few risk contributions but, for a large number of clients, the Hermite polynomial based methods have shown to be the best alternative.

- **Random *LGD* models**

Table 2.3 provides the parameters calibrated for the pure macro *LGD* models using the FDIC information.

[INSERT TABLE 2.3 AROUND HERE]

Table 2.4 reports the tail probabilities for different loss levels under the saddlepoint and IS methods for the pure macroeconomic *LGD* models. Two issues can be mentioned: a) the saddlepoint method generates very accurate results for the different loss levels and b) the different models provide very similar results but for the lognormal recoveries case which assigns lower probability to high losses.

[INSERT TABLE 2.4 AROUND HERE]

Figures 2.5 to 2.10 provide the risk allocations of the 99.9% probability loss levels for the pure macroeconomic *LGD* models (*LGD^C*).

[INSERT FIGURES 2.5 TO 2.10 AROUND HERE]

Several conclusions arise from the risk allocation under pure macroeconomic random recoveries:

- The IS framework requires a very high number of simulations to obtain stable risk allocation values as it can be seen in the dashed lines.

- The introduction of the random *LGD* increases the risk allocated to the two biggest financial institutions.
- Even for a small number of simulations, the pure saddlepoint method generates very accurate risk allocation results although it is very time consuming.
- Compared to the pure saddlepoint approach, the Hermite polynomials based saddlepoint method generates very close results and is much faster.

Table 2.5 shows the parameters calibrated for the mixed macro and idiosyncratic *LGD* models using the FDIC information.

[INSERT TABLE 2.5 AROUND HERE]

To get the loss distribution under idiosyncratic recoveries we use the method to obtain the inverse Laplace transform described in Subsection 3.3. Following Abate and Whitt (1995), the Euler summation method can be used to increase the convergence speed.¹⁷ We will use $c = \hat{t}$ where $K'_x(\hat{t}) = x$ and $\Delta u = \frac{\hat{t}}{10}$ aiming to guarantee that $e^{K_x(c+ik\Delta u)}$ can be computed and that the distance between $Re\left(\frac{e^{K_x(c+i(k-1)\Delta u)}}{c+i(k-1)\Delta u}\right)$ and $Re\left(\frac{e^{K_x(c+ik\Delta u)}}{c+ik\Delta u}\right)$ is reasonably small.

To obtain the loss distribution we first define a vector $V_l = (l_1, l_2, \dots, l_N)$ of loss levels whose mean \bar{l} is considered a representative loss value. Then, we obtain $\hat{t}(\bar{l}, Z)$ for each macroeconomic scenario and estimate $1 - F_x(V_l|Z)$ using $\hat{t}(\bar{l}, Z)$ and equation (2.9). Finally we compute the weighted average of the estimates of $1 - F_x(V_l|Z)$ over all the simulated scenarios Z .

This method to estimate the loss distribution is faster than that in Abate and Whitt (1995) as, now, the values c and Δu do not depend on the loss level l_i . Hence, $Re\left(\frac{e^{K_x(c+2ik\Delta u)}}{a+2ik\Delta u}\right)$ remains constant for every loss value in V_l and the number of calculations required to estimate the loss distribution is reduced.

Table 2.6 provides the tail probabilities for different loss levels for the different mixed macroeconomic and idiosyncratic *LGD* models under saddlepoint and IS methods. It can be observed that the Laplace inversion method generates very accurate results.

[INSERT TABLE 2.6 AROUND HERE]

Comparing Tables 2.4 and 2.6 we can see that the loss distributions under the mixed idiosyncratic and macroeconomic *LGD* model are very close to those under the pure macro *LGD*

¹⁷This method approximates the sum of a series using the first N_d terms plus a weighted sum of N_l terms that uses the first N_d terms.

model. Then, we can say that, at least for the portfolio under analysis, modeling the idiosyncratic behavior of the recoveries is not important to measure the loss distribution.

Finally, using the inverse Laplace transform to recover the loss distribution requires much more time than the saddlepoint method. This gets even worse for the beta *LGD* distribution as we must evaluate numerically the moment generating function of the beta distribution.¹⁸ Compared with a pure Monte Carlo or an IS method, the time required to solve the idiosyncratic *LGD* case under the inverse Laplace transform approach is several times higher and the results are very similar if not equal.

- **Market mode models**

We conclude this Section analyzing the market mode losses. For the market mode valuation we use a discount factor by rating grade obtained from the average historical credit spreads. Table 2.7 shows the discount factors that we have used while Table 2.8 includes the migration matrix employed for the migration rule.

[INSERT TABLES 2.7 AND 2.8 AROUND HERE]

Figure 2.11 shows the results from applying equation (2.11) to the losses on a) total assets and on b) best state value (AAA valuation). We can see that these results are not very satisfactory and that the definition of the losses affects the results.

[INSERT FIGURE 2.11 AROUND HERE]

Next we test the case of losses on the current state valuation, in which equation (2.11) is not valid as we may have negative losses (that is, earnings). Let S (G) denote the set of the clients-states with losses smaller (greater) than l . Let S_n and G_n be the number of elements in S and G , respectively. Then we have that

$$P\left(\sum_{i=1}^M x_{M,i,h} \geq l \mid Z\right) = 1 - \sum_{i=1}^{\infty} T_i \quad (2.12)$$

¹⁸See the Appendix for further details.

where

$$\begin{aligned}
T_1 &= P(S|Z) P\left(\sum_{i=1}^M x_{M,i,h} < l \mid S, Z\right) \\
T_2 &= \sum_{i \in G} P(i \in G|Z) P\left(x_{M,i,h} + \sum_{j \neq i} x_{M,j,h} < l \mid i \in G, j \in S, Z\right) \\
T_3 &= \sum_{i \in G} \sum_{j \in G} P(i \in G|Z) P(j \in G|Z) P\left(x_{M,i,h} + x_{M,j,h} + \sum_{k \neq i,j} x_{M,k,h} < l \mid i \in G, j \in G, k \in S, Z\right) \\
&\vdots
\end{aligned}$$

The terms T_i represent the probability of extreme losses in $i - 1$ clients while total losses are smaller than l . The probability terms inside T_i can be approximated using saddlepoint methods. It is interesting to note that T_1 only requires one saddlepoint approximation, T_2 requires G_n saddlepoint approximations and so on. Hence, for clients-states with many possible high losses, the process becomes very time consuming if the pure saddlepoint method is used to estimate all the probabilities. Therefore we explore the utilization of a simple normal approximation to estimate the terms in T_2 and T_3 .¹⁹

Figure 2.12 provides the results using just the term T_1 in equation (2.12).²⁰ It can be seen that the accuracy of the saddlepoint method is very good and the 99.9% probability losses over the current state market value is 69,940 MM €.

[INSERT FIGURE 2.12 AROUND HERE]

Figure 2.13 shows the *CVaR* and *CES* risk allocations using the saddlepoint and the IS methods. In the case of the *CVaR* risk allocation the pure saddlepoint method performs very well with just 100 macroeconomic scenarios although it is much slower than the Hermite polynomials based saddlepoint method which performs poorer. For the *CES* risk allocation none of the saddlepoint methods performs well, probably due to the high concentration in the portfolio.

[INSERT FIGURE 2.13 AROUND HERE]

To show this, Figure 2.14 provides the risk allocation of the 99.9% probability losses if all the institutions in the portfolio would have the same *EAD*. We can see that the pure saddlepoint

¹⁹The Hermite polynomials based approximation can also be tested.

²⁰Using the normal approximation, our results show that the effect of T_2 is negligible.

method produces results very close to those of the IS and so do the Hermite polynomials based results. Then, we can conclude that the pure saddlepoint and the Hermite based methods work well in non-concentrated portfolios although the latter method is much faster.

[INSERT FIGURE 2.14 AROUND HERE]

2.5 Conclusions

This paper has introduced a modified saddlepoint approximation trying to improve the estimation process of the credit risk contributions for both *VaR* and *ES* based credit risk allocations. We have also extended the saddlepoint methods to deal with random *LGDs* and market valuation. The new risk contribution method is based on Hermite polynomials and requires as many calculations as the method in Martin and Thompson (2001) and, then, it is much faster than the exact method.

We have extended the saddlepoint technique to deal with random *LGDs* and market mode models. Under idiosyncratic random *LGD* the saddlepoint method can be applied just if the moment generating function of the *LGD* random variable has a closed-form representation. We have also shown that it may not work properly if the portfolio contains big clients with low default probability under the idiosyncratic random *LGD* model. In this case we have suggested a numerical inverse Laplace transform method. For the market mode models the proposed risk allocation does not require to estimate a different saddlepoint for each final rating state and therefore speeds up the calculations.

All these extensions have been empirically illustrated with the portfolio of Spanish financial institutions in order to analyze the accuracy and possible drawbacks of the method using Monte Carlo simulations as benchmark. We have shown that our risk allocation method is as accurate and less time demanding than the exact one. Compared with the method suggested in Martin and Thompson (2001) our results are more accurate with similar computational time. We have analyzed several random *LGD* models providing the detailed steps to calibrate the parameters. One of the main conclusions is that modeling a mixed macroeconomic and idiosyncratic *LGD* does not generate higher risk numbers than a pure macroeconomic *LGD* model. In the case of market valuation and concentrated portfolios, we find that the saddlepoint method may not provide good results at an acceptable time and a IS method may be more suitable.

To conclude with the suitability of the saddlepoint methods for credit risk measurement we want to mention that, according to our empirical analysis, these methods have shown to be very fast and accurate to estimate the credit risk loss distribution for constant (or conditionally constant) *LGD* and default mode. In the case of idiosyncratic random *LGD* or market mode models, the

Monte Carlo method may produce more reliable results at a smaller computational time than the saddlepoint approach. Similar conclusions apply to the risk allocation.

Appendix

This Appendix includes the moment generating functions for the normal, gamma, and beta distributions.

- **Normal distribution**

$$M_{N(\mu,\sigma)}(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$$

- **Gamma distribution**

$$M_{\Gamma(k,\theta)}(t) = (1 - t\theta)^{-k}$$

- **Beta distribution**

The moment generating function of the beta distribution is related to the Kummer function

$$M_{B(\alpha,\beta)}(t) = \sum_{n=0}^{\infty} \frac{(\alpha)_n}{(\alpha + \beta)_n} \frac{t^n}{n!}$$

with $(x)_n = \prod_{i=1}^n (x + i - 1)$.

The derivatives of the Kummer function are given by

$$\frac{\partial^n M_{B(\alpha,\beta)}(t)}{\partial t^n} = \frac{(\alpha)_n}{(\alpha + \beta)_n} M_{B(\alpha+n,\alpha+\beta+n)}(t)$$

For large values of $|t|$ ($|t| > 5.5 \min\{|\alpha|, |\alpha + \beta|\}$), the Kummer function at a given point is evaluated by applying an asymptotic expansion that provides²¹

$$M_{B(\alpha,\beta)}(t) = e^{\pm i\pi\alpha} t^{-\alpha} \frac{\Gamma(\alpha + \beta)}{\Gamma(\beta)} \sum_{n=0}^{R-1} (-t)^{-n} \frac{(\alpha)_n (1 - \beta)_n}{n!} + e^t t^{-\beta} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)} \sum_{n=0}^{S-1} t^{-n} \frac{(\beta)_n (1 - \alpha)_n}{n!}$$

We set the maximum number of terms used in the expansion at 500 and implement a stopping rule when the relative weight of an additional term is smaller than 10^{-15} .

²¹See expression (13.5.1) in Abramowitz and Stegun (1964).

Appendix of Tables

Table 2.1: Constant conditional *LGD* models (LGD^C).

Model	Normal	Lognormal	Logit Normal	Probit Normal	Normal Square	Basel
Distribution ($LGD^C z$)	$a + bz$	e^{a+bz}	$\frac{e^{a+bz}}{1+e^{a+bz}}$	$\Phi(a + bz)$	$(a + bz)^2$	$\Phi\left(\frac{\Phi^{-1}(a)-bz}{\sqrt{1-b^2}}\right)$

Table 2.2: Random conditional *LGD* models (LGD^R).

Model	Distribution	$E(LGD^R z)$	$Var(LGD^R z)$
Normal	$LGD^R z = a + b\left(cz + \gamma\sqrt{1-c^2}\right)$	$a + bcz$	$b^2(1-c^2)$
Gamma	$LGD^R z \approx \Gamma(c^{-1}Logit(a + bz), c)$	$Logit(a + bz)$	$cLogit(a + bz)$
Beta	$LGD^R z \approx B(cLogit(a + bz), c(1 - Logit(a + bz)))$	$Logit(a + bz)$	$\frac{Logit(a+bz)(1-Logit(a+bz))}{c+1}$

Table 2.3: Calibrated parameters for the LGD^C model.

LGD^C Model	a	b
Normal	0.192	0.081
Lognormal	-1.736	0.405
Logit Normal	-1.526	0.535
Probit Normal	-0.913	0.307
Normal Square	0.427	0.094
Basel	0.190	0.292

Table 2.4: Tail probabilities (in percentage) for different loss levels (in MM €) for the pure macroeconomic *LGD* (LGD^C) models. We have used 100 x 100 simulations for the importance sampling method (IS) and 100 simulations for the saddlepoint method (Sad).

Model	Loss levels						
	10,000	20,000	30,000	40,000	50,000	60,000	70,000
Normal (IS)	0.8968	0.2803	0.1599	0.1033	0.0579	0.0326	0.0185
Normal (Sad)	0.9934	0.272	0.1533	0.0898	0.0512	0.0297	0.0178
Lognormal (IS)	0.6262	0.2002	0.1009	0.0542	0.0346	0.0191	0.013
Lognormal (Sad)	0.6832	0.2086	0.1006	0.0563	0.0333	0.0192	0.0118
Logit Normal (IS)	0.8837	0.2819	0.1556	0.097	0.056	0.0328	0.0186
Logit Normal (Sad)	0.9825	0.2777	0.1505	0.0863	0.0511	0.0302	0.0179
Probit Normal (IS)	0.8725	0.2622	0.1494	0.0941	0.0552	0.0323	0.0177
Probit Normal (Sad)	0.9672	0.2675	0.1456	0.0846	0.0504	0.0301	0.0175
Normal Square (IS)	0.8871	0.2826	0.1575	0.0982	0.058	0.0332	0.0185
Normal Square (Sad)	0.9845	0.2766	0.1516	0.0876	0.0509	0.0302	0.0178
Basel (IS)	0.9700	0.2947	0.1496	0.0888	0.0569	0.0352	0.0245
Basel (Sad)	1.0214	0.3016	0.1529	0.0919	0.0582	0.0366	0.0251

Table 2.5: Calibrated parameters for the LGD^R model.

LGD^R Model	a	b	c
Normal	0.193	0.143	0.585
Gamma	-1.549	0.538	0.070
Beta	-1.549	0.538	10.040

Table 2.6: Tail probabilities (in percentage) for different loss levels (in MM €) for the mixed macroeconomic and idiosyncratic random LGD (LGD^R) models. We have used 100 x 100 simulations for the importance sampling (IS) method and 100 simulations for the Laplace inversion method (Lap).

Loss levels							
Model	10,000	20,000	30,000	40,000	50,000	60,000	70,000
Normal (IS)	1.0061	0.2752	0.1429	0.0947	0.0612	0.0373	0.0246
Normal (Lap)	1.0492	0.2502	0.136	0.0838	0.0509	0.0335	0.0233
Gamma (IS)	1.074	0.3102	0.1395	0.0787	0.0479	0.0356	0.0266
Gamma (Lap)	1.0106	0.3242	0.1454	0.0857	0.0572	0.0386	0.0287
Beta (IS)	1.0296	0.3498	0.1583	0.0969	0.0633	0.0491	0.0376
Beta (Lap)	1.0203	0.3224	0.1424	0.0811	0.0529	0.0357	0.0242

Table 2.7: 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 2.8: 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/C	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%

Appendix of Figures

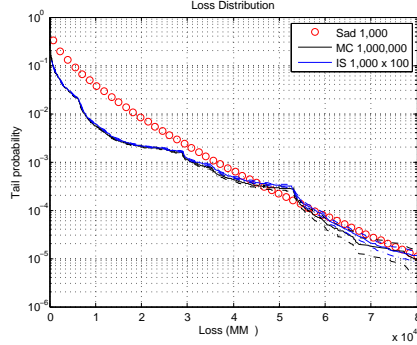


Figure 2.1: Loss distribution of the Spanish portfolio. We use the saddlepoint method (Sad) with 1,000 simulations, Monte Carlo (MC) with 1,000,000 simulations, and the importance sampling method (IS) with 1,000 x 100 simulations. The red circles indicate the saddlepoint estimates. The black and blue lines show, respectively, the Monte Carlo and IS results and their 5%-95% confidence intervals (dashed lines).

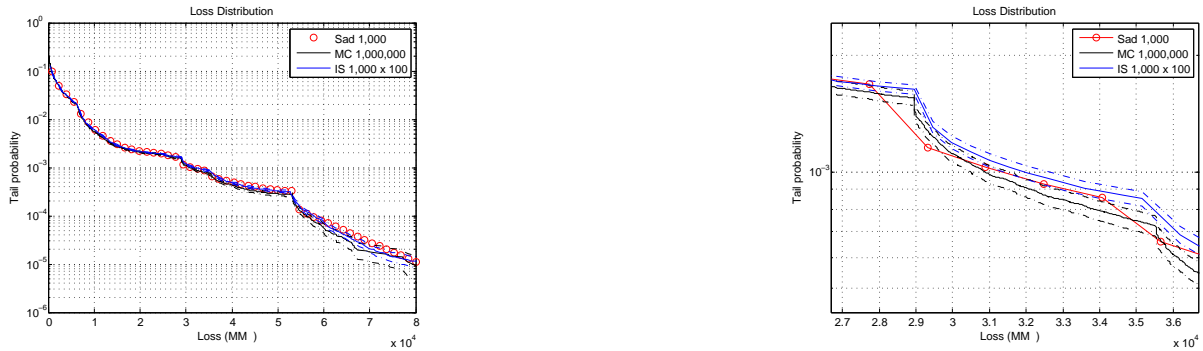


Figure 2.2: Loss distribution of the Spanish portfolio considering the *adaptive saddlepoint approximation*. Left and right graphs show, respectively, the tail distribution and the detail of the distribution in the neighborhood of the 99.9% probability loss level. We use the saddlepoint method (Sad) with 1,000 simulations, Monte Carlo (MC) with 1,000,000 simulations, and the importance sampling method (IS) with 1,000 x 100 simulations. The red circles indicate the saddlepoint estimates. The black and blue lines show, respectively, the Monte Carlo and IS results and their 5%-95% confidence intervals (dashed lines).

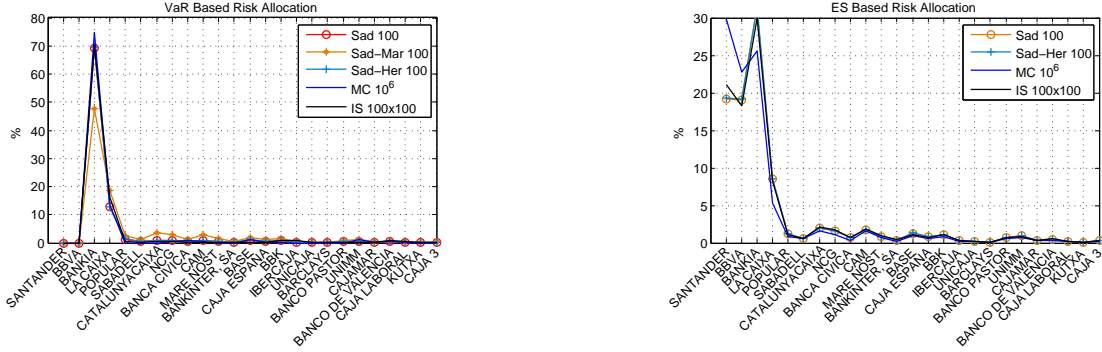


Figure 2.3: Risk allocation. Left and right graphs show, respectively, the risk allocations based on the *VaR* and the *ES*. We use the pure saddlepoint method (Sad), the saddlepoint based Martin approximation (Sad-Mar), and the Hermite polynomials based saddlepoint approximation (Sad-Her) with 100 simulations, Monte Carlo (MC) with 1,000,000 simulations, and the importance sampling method (IS) with 100 x 100 simulations.

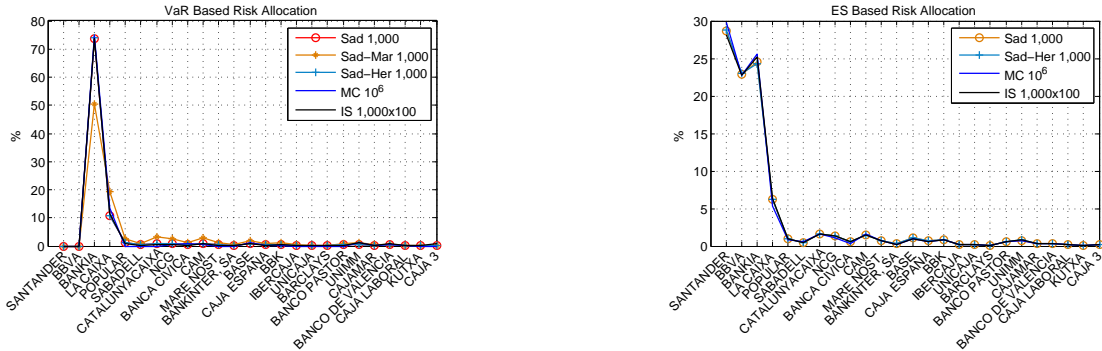


Figure 2.4: Risk allocation. Left and right graphs show, respectively, the risk allocations based on the *VaR* and the *ES*. We use the pure saddlepoint method (Sad), the saddlepoint based Martin approximation (Sad-Mar), and the Hermite polynomials based saddlepoint approximation (Sad-Her) with 1,000 simulations, Monte Carlo (MC) with 1,000,000 simulations, and the importance sampling method (IS) with 1,000 x 100 simulations.

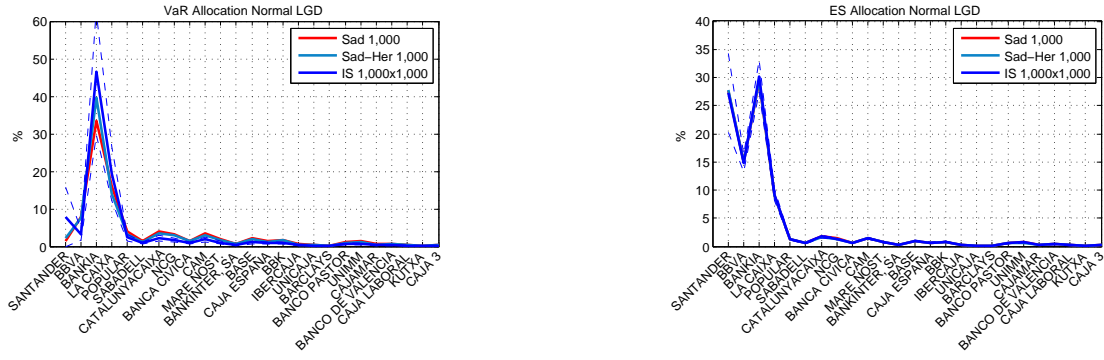


Figure 2.5: Risk Allocation Normal Macroeconomic *LGD*. We use the pure saddlepoint method (Sad) and the Hermite polynomials based saddlepoint approximation (Sad-Her) with 1,000 simulations and the importance sampling method (IS) with 1,000 x 1,000 simulations.

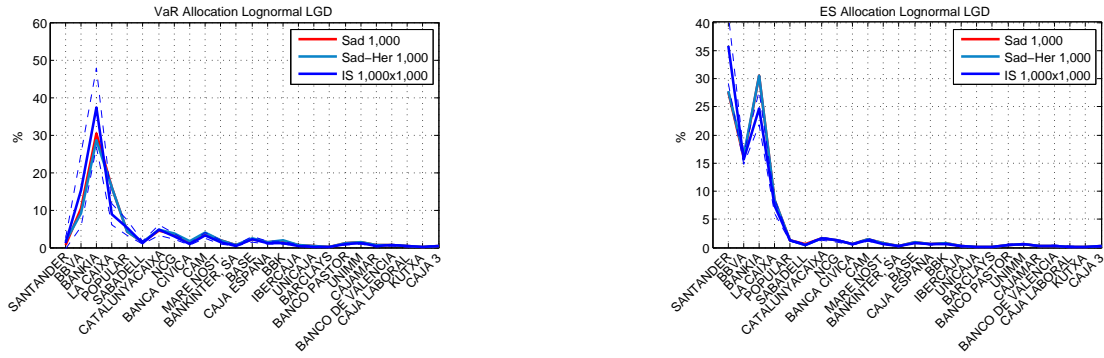


Figure 2.6: Risk Allocation Lognormal Macroeconomic *LGD*. We use the pure saddlepoint method (Sad) and the Hermite polynomials based saddlepoint approximation (Sad-Her) with 1,000 simulations and the importance sampling method (IS) with 1,000 x 1,000 simulations.

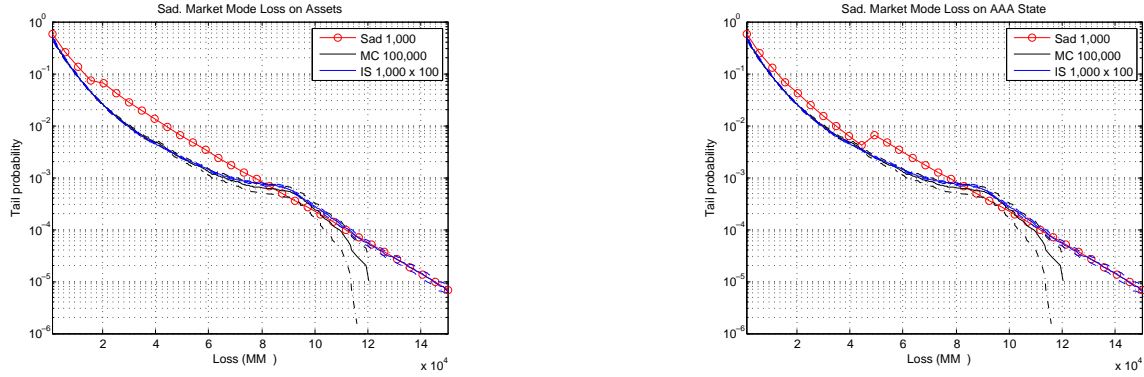


Figure 2.11: Loss distribution of the Spanish portfolio under market valuation considering the *adaptive saddlepoint approximation*. Left and right graphs show, respectively, the results obtained for losses on assets and losses on AAA state valuation. We use the saddlepoint method (Sad) with 1,000 simulations, Monte Carlo (MC) with 100,000 simulations, and the importance sampling method (IS) with 1,000 x 100 simulations. The red circles indicate the saddlepoint estimates. The black and blue lines show, respectively, the Monte Carlo and IS results and their 5%-95% confidence intervals (dashed lines).

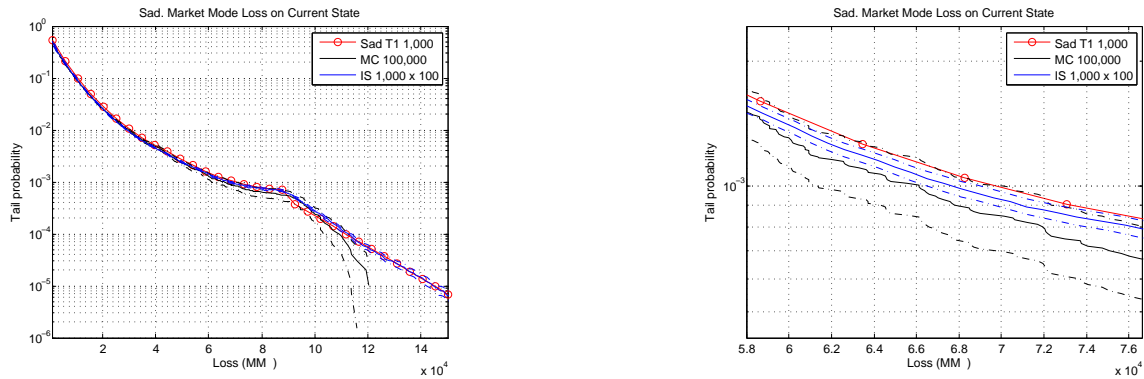


Figure 2.12: Loss on current value distribution of the Spanish portfolio under market valuation considering the *adaptive saddlepoint approximation*. Left and right graphs show, respectively, the tail distribution and the detail of the distribution in the neighborhood of the 99.9% probability loss level. We use the saddlepoint method (Sad) with 1,000 simulation, Monte Carlo (MC) with 100,000 simulations, and the importance sampling method (IS) with 1,000 x 100 simulations. The red circles indicate the saddlepoint estimates. The black and blue lines show, respectively, the Monte Carlo and IS results and their 5%-95% confidence intervals (dashed lines).

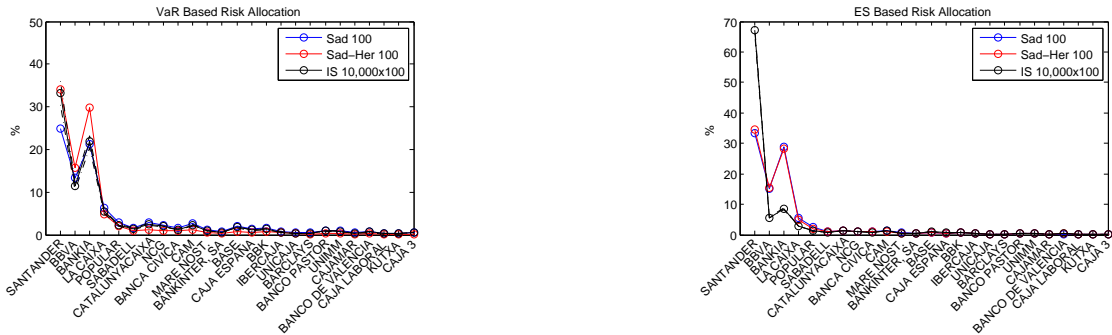


Figure 2.13: *VaR* and *ES* contributions for the market mode model. Left and right graphs show, respectively, the risk allocations based on the *VaR* and the *ES*. We use the pure saddlepoint method (Sad) and the Hermite polynomials based saddlepoint approximation (Sad-Her) with 100 simulations and the importance sampling method (IS) with 10,000 x 100 simulations.

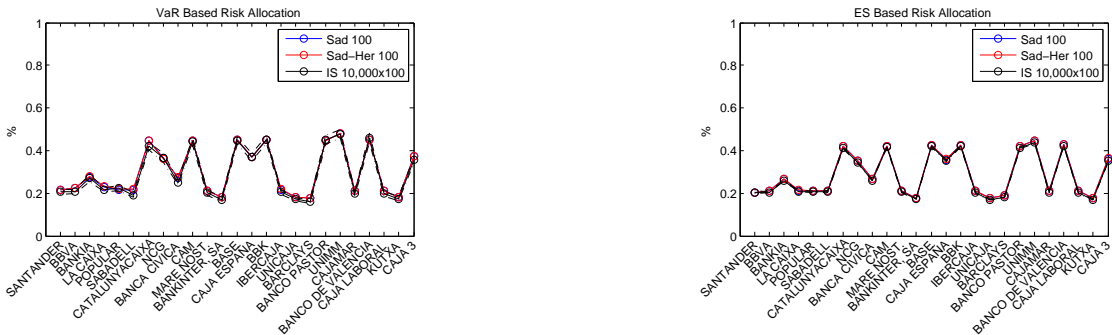


Figure 2.14: *VaR* and *ES* contributions on an equally weighted (*EAD*) portfolio for the market mode model. Left and right graphs show, respectively, the risk allocations based on the *VaR* and the *ES*. We use the pure saddlepoint method (Sad) and the Hermite polynomials based saddlepoint approximation (Sad-Her) with 100 simulations and the importance sampling method (IS) with 10,000 x 100 simulations.

Chapter 3

Taylor expansion based methods to measure credit risk

3.1 Introduction

This paper uses the Taylor expansion based method in Pykhtin (2004) to obtain an analytical estimation of the credit risk loss distribution of the Spanish financial system. We compare the results of the method with those obtained using Monte Carlo simulations. We also propose an approximate method to calculate the loss distribution of a portfolio under random correlated recoveries. We consider that this kind of approximations can be easily implemented under the Basel capital charge regulation.

In previous papers (see García-Céspedes and Moreno (2014) and García-Céspedes and Moreno (2014b)) we studied the loss distribution of the Spanish financial system using importance sampling techniques and the saddlepoint methods, respectively. Both methods are subject to random variability due to the Monte Carlo simulation, this noise is more intense in the risk allocation process. One of the objectives of this paper is to analyze the suitability of the Taylor expansion based method to measure the credit risk of a portfolio. In the Taylor expansion based method introduced in Pykhtin (2004), the real loss distribution is approximated using the loss distribution of a similar portfolio plus several adjustment terms. This type of expansion is related with the Taylor expansion of the real loss distribution and the adjustment terms are related with the derivatives of the moments of the portfolio loss.

This paper provides two major contributions to the literature. First, we use the approximation in Pykhtin (2004) to measure the risk of the Spanish financial system portfolio built in García-Céspedes and Moreno (2014) and we allocate it over the different financial institutions. To allocate the risk we use some of the results in Morone et al. (2012). Second, we develop two simple correlated recoveries models and we use the Taylor expansion ideas to approximate the loss distribution under random correlated recoveries.

According to our results, the Pykhtin model does not perform well in the case of concentrated

portfolios, it is not able to capture the sudden jumps in the loss distribution, neither it is able to properly allocate the risk over the counterparties. However we get very satisfactory results when we apply the Taylor expansion ideas to approximate the loss distribution of a portfolio under random and correlated recoveries.

This paper is organized as follows. Sections 2 and 3 introduce, respectively, the Vasicek (1987) and the Pykhtin (2004) models. Section 4 describes the Spanish financial institutions portfolio and obtains its loss distribution and risk allocation applying the Pykhtin (2004) model. In Section 5 we develop two correlated recoveries models and use the Taylor expansion ideas to approximate the loss distribution. Finally, Section 6 summarizes our main results and concludes.

3.2 The Vasicek (1987) model

We will start remembering the credit risk model developed in Vasicek (1987). This model states that the value of a counterparty j , V_j , is driven by a own macroeconomic normal factor Y_j and an idiosyncratic independent normal term ξ_j . The own macroeconomic factor Y_j is the linear combination of some more general macroeconomic independent factors z_f . Then we can write

$$\begin{aligned} V_j &= r_j Y_j + \sqrt{1 - r_j^2} \xi_j = r_j \sum_{f=1}^k \alpha_{f,j} z_f + \sqrt{1 - r_j^2} \xi_j \\ &= \sum_{f=1}^k \beta_{f,j} z_f + \sqrt{1 - \sum_{f=1}^k (\beta_{f,j})^2} \xi_j \end{aligned} \quad (3.1)$$

The client j defaults in his obligations if the assets value V_j falls below a given level k . As $V_j \sim N(0, 1)$ there exist a direct relation between the default threshold k and the historical average default rates of client j , $PD_{j,C}$ (or over the cycle default rate), this is $k = \Phi^{-1}(PD_{j,C})$, where $\Phi^{-1}(\cdot)$ denotes the inverse normal distribution. The total losses L of a portfolio made up of M clients can be obtained adding the individual ones, that is,

$$L = \sum_{j=1}^M x_j(Y_j, \xi_j) = \sum_{j=1}^M EAD_j LGD_j D_j(Y_j, \xi_j)$$

where D_j is a dummy variable that indicates if the clients defaults, EAD_j is the exposure at default (the amount owed by the client j , or by the clients in the subportfolio j , in the default moment), and LGD_j is the loss given default (the percentage of the final loss after all the recovery process relative to the exposure at default). For the sake of simplicity we will use the notation $EAD_j LGD_j = g_j$.

Given the specification in (3.1) and conditional to the macroeconomic factors $Z = \{z_1, z_2, \dots, z_k\}$, the default probability of the client j is

$$\begin{aligned} p_j(Z) &= \text{Prob}(D_j = 1|Z) = \text{Prob}(V_j \leq k) = \text{Prob}\left(\xi_j \leq \frac{k - r_j \sum_{f=1}^k \alpha_f z_f}{\sqrt{1 - r_j^2}}\right) \\ &= \Phi\left(\frac{\Phi^{-1}(PD_{j,C}) - r_j \sum_{f=1}^k \alpha_f z_f}{\sqrt{1 - r_j^2}}\right) \end{aligned}$$

In a granular portfolio (many identical clients) made up of M different subportfolios, the losses conditional to a macroeconomic scenario are

$$L|z = \sum_{j=1}^M g_j \Phi\left(\frac{\Phi^{-1}(PD_{j,C}) - r_j \sum_{f=1}^k \alpha_f z_f}{\sqrt{1 - r_j^2}}\right) = \sum_{j=1}^M g_j p_j(Z)$$

If all the clients in the portfolio are exposed to the same macroeconomic factor, the loss at a given probability level q can be obtained just replacing z by $\Phi^{-1}(q)$ in the previous formula. But in the general case of multi-factor and non-granular portfolios, closed-form expressions are not available and Monte Carlo methods or approximate formulas are needed.¹

3.3 The Pykhtin (2004) approximate model

Given a certain confidence level q , Pykhtin (2004) suggests to estimate the value at risk ($VaR(q)$) and the expected shortfall ($ES(q)$) by approximating the loss distribution of the real portfolio through a Taylor expansion. This author considers a granular and unifactorial portfolio as starting point and then adds some adjustment terms to capture the non-granularity and the multifactoriality of the real portfolio. Following Pykhtin (2004), we consider the random variable \bar{L} and define $L_\epsilon = \bar{L} + \epsilon U$, $\epsilon \in \mathbb{R}$, with $U = L - \bar{L}$.² Let $t_q(L)$ denote the percentile q of the random variable L . Then, a Taylor expansion of $t_q(L_{\epsilon=1})$ around $\epsilon = 0$ leads to

$$t_q(L) = t_q(L_{\epsilon=1}) = t_q(\bar{L}) + \left. \frac{dt_q(L_\epsilon)}{d\epsilon} \right|_{\epsilon=0} + \frac{1}{2} \left. \frac{d^2 t_q(L_\epsilon)}{d\epsilon^2} \right|_{\epsilon=0} + \dots \quad (3.2)$$

Gourieroux et al. (2000) and Martin and Wilde (2002) provide an analytical expression for the previous two derivatives

$$\left. \frac{dt_q(L_\epsilon)}{d\epsilon} \right|_{\epsilon=0} = E(U|\bar{L} = t_q(\bar{L})) \quad (3.3)$$

$$\left. \frac{d^2 t_q(L_\epsilon)}{d\epsilon^2} \right|_{\epsilon=0} = - \left. \frac{1}{f_{\bar{L}}(l)} \frac{d}{dl} (f_{\bar{L}}(l) \text{var}(U|\bar{L} = l)) \right|_{l=t_q(\bar{L})} \quad (3.4)$$

¹See, for instance, Pykhtin (2004), Glasserman and Li (2005) or Voropaev (2011).

²It is straightforward to see that $L_{\epsilon=1} = L$.

where $var()$ stands for the variance of a random variable. Pykhtin (2004) suggests to define \bar{L} as the losses of a granular and unifactorial portfolio that depends on a single macroeconomic factor $\bar{Y} = \sum_{f=1}^k b_f z_f$ where $\sum_{f=1}^k b_f^2 = 1$. Intuitively this unique macroeconomic factor should try to capture most of the influence of the different z_f on the losses of the real portfolio. Then, we have

$$\bar{L}(\bar{Y}) = \sum_{j=1}^M g_j \Phi \left(\frac{\Phi^{-1}(PD_{j,C}) - a_j \bar{Y}}{\sqrt{1 - a_j^2}} \right) = \sum_{j=1}^M g_j \hat{p}_j(\bar{Y}) \quad (3.5)$$

The terms a_j are the sensitivities of the clients in the granular portfolio to the single macroeconomic factor \bar{Y} .³

According to the previous definitions of \bar{L} and \bar{Y} , we can rewrite the macroeconomic factor Y_j and the asset value of each client V_j as

$$\begin{aligned} Y_j &= \bar{Y} \sum_{f=1}^k \alpha_{j,f} b_f + \sqrt{1 - \left(\sum_{f=1}^k \alpha_{j,f} b_f \right)^2} \gamma_j \\ V_j &= r_j \bar{Y} \sum_{f=1}^k \alpha_{j,f} b_f + \sqrt{1 - \left(r_j \sum_{f=1}^k \alpha_{j,f} b_f \right)^2} \psi_j \end{aligned} \quad (3.6)$$

where the random variables γ_j and ψ_j are independent of \bar{Y} but γ_j and γ_i are correlated between clients, and so are ψ_j and ψ_i .

Now we can estimate $E(L|\bar{L})$. Looking at equation (3.6), we have that

$$E(L|\bar{L}) = E(L|\bar{Y}) = \sum_{j=1}^M g_j \Phi \left(\frac{\Phi^{-1}(PD_{j,C}) - r_j \bar{Y} \sum_{f=1}^k \alpha_{j,f} b_f}{\sqrt{1 - \left(r_j \sum_{f=1}^k \alpha_{j,f} b_f \right)^2}} \right) \quad (3.7)$$

Comparing equations (3.5) and (3.7) we can see that if we define $a_j = r_j \sum_{f=1}^k \alpha_{j,f} b_f = \sum_{f=1}^k \beta_{j,f} b_f$ then equation (3.3) equals to zero and the first derivative term in the Taylor expansion vanishes. It is also important to note that, given the previous definition of a_j , we can rewrite equation (3.6) in a much shorter way

$$V_j = a_j \bar{Y} + \sqrt{1 - a_j^2} \psi_j \quad (3.8)$$

Then, V_j conditional to \bar{Y} is distributed as $N(a_j \bar{Y}, \sqrt{1 - a_j^2})$. The correlation between V_j and V_i conditional to \bar{Y} will be required later in order to obtain equation (3.4) and it is equal to that between ψ_j and ψ_i . Using equations (3.1) and (3.8), we can get

$$E(\psi_j \psi_i) = \frac{\sum_{f=1}^k \beta_{j,f} \beta_{i,f} - a_i a_j}{\sqrt{(1 - a_j^2)(1 - a_i^2)}} \quad (3.9)$$

³In the next paragraphs we will show a criterion to set a_j such that the first derivative in equation (3.2) vanishes.

On the other hand, Pykhtin (2004) suggests to set the coefficients b_f that define the single macroeconomic factor \bar{Y} as

$$b_f(q) = \frac{\sum_{j=1}^M \alpha_{j,f} g_j p_j(Y_j)}{\sqrt{\sum_{f=1}^k \left(\sum_{j=1}^M \alpha_{j,f} g_j p_j(Y_j) \right)^2}} \Bigg|_{Y_j = \Phi^{-1}(q)}$$

This is a weighted sum of the expected losses of every client in the portfolio given that each of the own macroeconomic factors is in its q percentile.

To estimate the second derivative in equation (3.4) we first note that conditioning to \bar{L} is the same as conditioning to \bar{Y} . As $\bar{Y} \sim N(0, 1)$, we have $\phi'(y) = -y\phi(y)$. Then, equation (3.4) becomes

$$\begin{aligned} \frac{d^2 t_q(L_\epsilon)}{d\epsilon^2} \Bigg|_{\epsilon=0} &= - \frac{1}{\phi(\bar{Y})} \frac{d}{d\bar{Y}} \left(\frac{\phi(\bar{Y})}{\bar{L}'(\bar{Y})} \text{var}(U|\bar{Y}) \right) \Bigg|_{\bar{Y} = \Phi^{-1}(q)} \\ &= \frac{1}{\bar{L}'(\bar{Y})} \left(-v'(\bar{Y}) + v(\bar{Y}) \left(\bar{Y} + \frac{\bar{L}''(\bar{Y})}{\bar{L}'(\bar{Y})} \right) \right) \Bigg|_{\bar{Y} = \Phi^{-1}(q)} \end{aligned} \quad (3.10)$$

where $v(\bar{Y}) = \text{var}(U|\bar{Y})$. Equation (3.10) requires the following inputs

$$\begin{aligned} \bar{L}'(\bar{Y}) &= - \sum_{j=1}^M g_j \frac{a_j}{\sqrt{1-a_j^2}} \phi \left(\frac{\Phi^{-1}(PD_{j,C}) - a_j \bar{Y}}{\sqrt{1-a_j^2}} \right) \\ \bar{L}''(\bar{Y}) &= - \sum_{j=1}^M g_j \frac{a_j^2}{1-a_j^2} \frac{\Phi^{-1}(PD_{j,C}) - a_j \bar{Y}}{\sqrt{1-a_j^2}} \phi \left(\frac{\Phi^{-1}(PD_{j,C}) - a_j \bar{Y}}{\sqrt{1-a_j^2}} \right) \\ v(\bar{Y}) &= \text{var}(U|\bar{Y}) = \text{var}(L|\bar{Y}) = \text{var} \left(\sum_{j=1}^M g_j D_j \Bigg| \bar{Y} \right) \\ &= E \left(\left(\sum_{j=1}^M g_j D_j \right)^2 \Bigg| \bar{Y} \right) - \left(E \left(\sum_{j=1}^M g_j D_j \Bigg| \bar{Y} \right) \right)^2 \\ &= \sum_{j=1}^M \sum_{i=1}^M g_j g_i E(D_i D_j | \bar{Y}) - \left(E \left(\sum_{j=1}^M g_j D_j \Bigg| \bar{Y} \right) \right)^2 \\ &= \sum_{j=1}^M \sum_{i \neq j}^M g_j g_i E(D_i D_j | \bar{Y}) + \sum_{j=1}^M g_j^2 E(D_j | \bar{Y}) - \left(\sum_{j=1}^M g_j E(D_j | \bar{Y}) \right)^2 \\ &= \sum_{j=1}^M \sum_{i \neq j}^M g_j g_i \Phi_2 \left(\Phi^{-1}(\hat{p}_j(\bar{Y})), \Phi^{-1}(\hat{p}_i(\bar{Y})), \rho_{i,j} \right) + \sum_{j=1}^M g_j^2 \hat{p}_j(\bar{Y}) \\ &\quad - \left(\sum_{j=1}^M g_j \hat{p}_j(\bar{Y}) \right)^2 \end{aligned} \quad (3.11)$$

where $\Phi_2(\cdot, \cdot, \cdot)$ denotes the bivariate cumulative standard normal distribution and $\rho_{i,j}$ is given by equation (3.9). We explored different algorithms to implement this bivariate distribution⁴ and obtained the best results in terms of accuracy, speed, and possibility to be vectorized using the method proposed in Genz (2004).

Equation (3.10) also requires to compute the first derivative of $v(\bar{Y})$, $v'(\bar{Y})$. Setting $Z_j(\bar{Y}) = \Phi^{-1}(\hat{p}_j(\bar{Y}))$ we have that $Z_j'(\bar{Y}) = \frac{\hat{p}_j'(\bar{Y})}{\phi(Z_j(\bar{Y}))}$ and some algebra leads to

$$\begin{aligned} \frac{d}{d\bar{Y}} [\Phi_2(Z_j(\bar{Y}), Z_i(\bar{Y}), \rho_{i,j})] &= \hat{p}_j'(\bar{Y}) \Phi \left(\frac{Z_i(\bar{Y}) - \rho_{i,j} Z_j(\bar{Y})}{\sqrt{1 - \rho_{i,j}^2}} \right) \\ &\quad + \hat{p}_i'(\bar{Y}) \Phi \left(\frac{Z_j(\bar{Y}) - \rho_{i,j} Z_i(\bar{Y})}{\sqrt{1 - \rho_{i,j}^2}} \right) \end{aligned}$$

Therefore we get that

$$\begin{aligned} v'(\bar{Y}) &= 2 \sum_{j=1}^M \sum_{i \neq j}^M g_j g_i \hat{p}_i'(\bar{Y}) \Phi \left(\frac{Z_j(\bar{Y}) - \rho_{i,j} Z_i(\bar{Y})}{\sqrt{1 - \rho_{i,j}^2}} \right) + \sum_{j=1}^M g_j^2 \hat{p}_j'(\bar{Y}) \\ &\quad - 2 \sum_{j=1}^M g_j \hat{p}_j(\bar{Y}) \sum_{j=1}^M g_j \hat{p}_j'(\bar{Y}) \end{aligned} \quad (3.12)$$

Pykhtin (2004) also obtains an analytical approximation for the expected shortfall (ES) given by

$$\begin{aligned} ES_q(L) &= \frac{1}{1-q} \int_q^1 t_s(L) ds = \frac{1}{1-q} \int_q^1 [t_s(\bar{L}) + \Delta t_s(L)] ds \\ &= ES_q(\bar{L}) + \frac{1}{1-q} \int_q^1 \Delta t_s(L) ds = ES_q(\bar{L}) + \Delta ES_q(L) \end{aligned}$$

where

$$\begin{aligned} ES_q(\bar{L}) &= \frac{1}{1-q} \int_{-\infty}^{\Phi^{-1}(1-q)} \bar{L}(\bar{Y}) \phi(\bar{Y}) d\bar{Y} \\ &= \frac{1}{1-q} \sum_{j=1}^M g_j \Phi_2(\Phi^{-1}(PD_{j,C}), \Phi^{-1}(1-q), a_j) \\ \Delta ES_q(L) &= -\frac{1}{2(1-q)} \int_{-\infty}^{\Phi^{-1}(1-q)} \frac{1}{\phi(\bar{Y})} \frac{d}{d\bar{Y}} \left(\frac{\phi(\bar{Y})}{\bar{L}'(\bar{Y})} v(\bar{Y}) \right) \phi(\bar{Y}) d\bar{Y} \\ &= -\frac{v(\bar{Y})}{2(1-q)} \frac{\phi(\bar{Y})}{\bar{L}'(\bar{Y})} \Big|_{\bar{Y}=\Phi^{-1}(1-q)} \end{aligned}$$

⁴See Chance and Agca (2003) for a general review of these alternatives or Owen (1956), Vasicek (1996), Genz (2004), and Hull (2011) for a more detailed explanation.

As it can be observed all the information required to obtain the expected shortfall has been obtained previously to estimate the VaR, this is a big advantage of this method.

Now let us focus on two extreme cases:

1. For a single factor and non-granular portfolio, we have $\rho_{i,j} = 0$ and equations (3.11)-(3.12) simplify to

$$\begin{aligned}
v(\bar{Y}) &= \sum_{j=1}^M \sum_{i \neq j}^M g_j g_i \hat{p}_j(\bar{Y}) \hat{p}_i(\bar{Y}) + \sum_{j=1}^M g_j^2 \hat{p}_j(\bar{Y}) - \left(\sum_{j=1}^M g_j \hat{p}_j(\bar{Y}) \right)^2 \\
v'(\bar{Y}) &= 2 \sum_{j=1}^M \sum_{i \neq j}^M g_j g_i \hat{p}'_i(\bar{Y}) \hat{p}_j(\bar{Y}) + \sum_{j=1}^M g_j^2 \hat{p}'_j(\bar{Y}) \\
&\quad - 2 \sum_{j=1}^M g_j \hat{p}_j(\bar{Y}) \sum_{j=1}^M g_j \hat{p}'_j(\bar{Y})
\end{aligned}$$

2. In the case of multifactorial and granular portfolio, equations (3.11)-(3.12) simplify to

$$\begin{aligned}
v(\bar{Y}) &= \sum_{j=1}^M \sum_{i=1}^M g_j g_i \Phi_2(\Phi^{-1}(\hat{p}_j(\bar{Y})), \Phi^{-1}(\hat{p}_i(\bar{Y})), \rho_{i,j}) \\
&\quad - \sum_{j=1}^M \sum_{i=1}^M g_j g_i \hat{p}_i(\bar{Y}) \hat{p}_j(\bar{Y}) \\
v'(\bar{Y}) &= 2 \sum_{j=1}^M \sum_{i=1}^M g_j g_i \hat{p}'_i(\bar{Y}) \left(\Phi \left(\frac{Z_j(\bar{Y}) - \rho_{i,j} Z_i(\bar{Y})}{\sqrt{1 - \rho_{i,j}^2}} \right) - \hat{p}_j(\bar{Y}) \right)
\end{aligned}$$

3.4 Portfolio Results

We study the portfolio of financial institutions covered by the Spanish deposit guaranty fund (FGD) at December, 2010. To obtain the risk measures of the portfolio first we need to have an estimate of the probability of default (PD), exposure at default (EAD), loss given default (LGD) and the macroeconomic factor sensitivity (α) of each institution.

We estimate the EAD based on the information on assets and liabilities for the Spanish financial institutions at December 2010. This information can be obtained from the AEB, CECA, and AECR webpages.⁵ During year 2010 many mergers took place, therefore we sum all the balance information from the different institutions that belong to the same group.

For the PD we use the public credit ratings available at December 2010 and the historical observed default rates reported by the rating agencies (see S&P (2009), Moody's (2009) and Fitch

⁵AEB is the Spanish Bank Association, CECA is the Spanish Saving Bank Association, and AECR is the Spanish Credit Cooperatives Association.

(2009)). Based on those data we infer a default probability for each institution. For those financial institutions with no external rating we assign one notch less than the average rating of the portfolio of financial institutions with external rating.

We extend the LGD results for financial institutions in Bennet (2002) to the period 1986-2009 using the FDIC (deposits guarantee fund in United States) public data of default recoveries.

The macroeconomic factor sensitivity, α , is set as the one in the Basel accord. Additionally all the financial institutions in the portfolio are exposed to a single macroeconomic factor, the Spanish factor. This is the case for all the institutions but for BBVA and Santander that are exposed to additional geographies. The exposure of those two institutions to the geographic factors is computed according to their net interest income by geography obtained from the public 2010 annual reports. As we have more than one macroeconomic factor we need a factor correlation matrix, we obtain this matrix using the GDP series of the different countries. Finally we compute orthogonalized factors so that we can apply Pykhtin model.

3.4.1 Portfolio VaR and ES

We have tested the accuracy of the approximate formulas in Pykhtin (2004) using the portfolio of the Spanish financial system at December 2010. The left graph in Figure 3.1 shows the loss distribution obtained using a simple Monte Carlo method (MC) and the Pykhtin approximation. We show the results of the Pykhtin model when only the first term or the first two terms of the approximation are considered (Pykhtin 1, Pykhtin 2). It can be seen that the Pykhtin method underestimates the probability of high losses for our portfolio. Another issue to be noted is that this method generates an approximation of the percentiles of the loss distribution which is smooth (continuous derivatives) and can not capture the sudden jumps of the portfolio loss distribution.

[INSERT FIGURE 3.1 AROUND HERE]

The right graph in Figure 3.1 provides the expected shortfall estimates using the Pykhtin (2004) method (Pykhtin 1, 2) and the exact ones obtained using the Monte Carlo method (MC). The conclusions are similar to those obtained from the loss distribution approximation.

3.4.2 Analytical VaR Contributions

The VaR contribution of a client i can be defined as the derivative of the VaR with respect to the current exposure share of the client i , this is

$$CVaR_i = \left. \frac{\partial VaR(w_1, \dots, w_m)}{\partial w_i} \right|_{w_i=1} \quad (3.13)$$

where the portfolio loss is defined as $L = \sum_{j=1}^M [w_j EAD_j] LGD_j D_j$. Under the Pykhtin approximation approach, we need to obtain the derivative of (3.2) with respect to w_i . This can be done numerically by computing the values $t_q(L, w_1, \dots, w_i, \dots, w_M)$ and $t_q(L, w_1, \dots, w_i + \lambda, \dots, w_M)$ with λ being small enough and using these values to approximate the derivative in equation (3.13). Unfortunately, this method must be repeated for all the clients as the information used to obtain the derivative of the client j can not be used for another client with the corresponding lack of computational synergies.

Alternatively, we can try to derivate equation (3.2) analytically. However, this derivative involves those of $a_j(w_1, \dots, w_i, \dots, w_M)$ with respect to w_i and this increases the complexity of the analytical derivation. This derivation is considerably simplified if we assume $a_j(w_1, \dots, w_i, \dots, w_M)$ to be constant, a (naive) case considered in Morone et al. (2012). We will obtain the terms of the analytical approximation of the VaR contributions under this assumption and will compare the results against a numerical derivation rule and against a Monte Carlo risk allocation rule. The analytical derivation rule has a big computational advantage, namely, most of the terms required in the derivation process have already being obtained for the VaR calculation.

Therefore, the VaR contribution of the client i can be obtained as

$$CVaR_i \approx \left. \frac{\partial t_q(\bar{L})}{\partial w_i} \right|_{w_i=1} + \left. \frac{\partial}{\partial w_i} \left(\left. \frac{d^2 t_q(L_\epsilon)}{d\epsilon^2} \right|_{\epsilon=0} \right) \right|_{w_i=1} \quad (3.14)$$

Under the assumption that $a_j(w_1, \dots, w_i, \dots, w_M)$ is constant, obtaining the first term of the right-hand side of this equation is straightforward considering equation (3.5). Looking at (3.10), the second term in equation (3.14) is given by

$$\begin{aligned} \frac{1}{2(\bar{L}'(\bar{Y}))^2} &\times \left[-\bar{L}'(\bar{Y}) \frac{\partial v'(\bar{Y})}{\partial w_i} + v'(\bar{Y}) \frac{\partial \bar{L}'(\bar{Y})}{\partial w_i} \right. \\ &+ \left(\bar{Y} + \frac{\bar{L}''(\bar{Y})}{\bar{L}'(\bar{Y})} \right) \left(\bar{L}'(\bar{Y}) \frac{\partial v(\bar{Y})}{\partial w_i} - v(\bar{Y}) \frac{\partial \bar{L}'(\bar{Y})}{\partial w_i} \right) \\ &\left. + v(\bar{Y}) \bar{L}(\bar{Y}) \frac{\partial}{\partial w_i} \left(\bar{Y} + \frac{\bar{L}''(\bar{Y})}{\bar{L}'(\bar{Y})} \right) \right] \Big|_{\bar{Y}=\Phi^{-1}(1-q), w_i=1} \end{aligned}$$

where

$$\begin{aligned}
\left. \frac{\partial \bar{L}'(\bar{Y})}{\partial w_i} \right|_{w_i=1} &= g_i \hat{p}'_i(\bar{Y}) \\
\left. \frac{\partial \bar{L}''(\bar{Y})}{\partial w_i} \right|_{w_i=1} &= g_i \hat{p}''_i(\bar{Y}) \\
\left. \frac{\partial v(\bar{Y})}{\partial w_i} \right|_{w_i=1} &= 2 \sum_{j \neq i} g_j g_i \Phi_2(\Phi^{-1}(\hat{p}_j(\bar{Y})), \Phi^{-1}(\hat{p}_i(\bar{Y})), \rho_{i,j}) + 2g_i^2 \hat{p}_i(\bar{Y}) \\
&\quad - 2g_i \hat{p}_i(\bar{Y}) \sum_{j=1}^M g_j \hat{p}_j(\bar{Y}) \\
\left. \frac{\partial v'(\bar{Y})}{\partial w_i} \right|_{w_i=1} &= 2 \sum_{j \neq i} g_i g_j \hat{p}'_i(\bar{Y}) \Phi \left(\frac{Z_j(\bar{Y}) - \rho_{i,j} Z_i(\bar{Y})}{\sqrt{1 - \rho_{i,j}^2}} \right) \\
&\quad + 2 \sum_{j \neq i} g_i g_j \hat{p}'_j(\bar{Y}) \Phi \left(\frac{Z_i(\bar{Y}) - \rho_{i,j} Z_j(\bar{Y})}{\sqrt{1 - \rho_{i,j}^2}} \right) \\
&\quad + 2g_i^2 \hat{p}'_i(\bar{Y}) - 2g_i \hat{p}_i(\bar{Y}) \sum_{j=1}^M g_j \hat{p}'_j(\bar{Y}) - 2g_i \hat{p}'_i(\bar{Y}) \sum_{j=1}^M g_j \hat{p}_j(\bar{Y})
\end{aligned}$$

The left graph in Figure 3.2 shows the loss allocation of the 99.9% probability loss level⁶ based on the VaR contribution. We show the results for the risk allocation using Monte Carlo method (MC) and three alternatives in the Pykhtin approximation derivatives: analytical derivation formula (Pykhtin Anal), and two types of numerical derivatives, one that varies the values of a_j and b_f as λ varies on the derivation process (Pykhtin Num. Der.) and another one that maintains them constant (Pykhtin Num. Der. Const.). It can be seen that the results from the analytical derivation rule coincide with those from the numerical derivation rules. We also observe that the Pykhtin rule based method does not provide accurate results for our portfolio. This is obvious for BBVA and Santander as their contributions should be zero.

[INSERT FIGURE 3.2 AROUND HERE]

In the case of the expected shortfall we have a similar formula

$$\begin{aligned}
CES_i &= \left. \frac{\partial ES(w_1, \dots, w_m)}{\partial w_i} \right|_{w_i=1} \\
&\approx \frac{g_i}{1-q} \Phi_2(\Phi^{-1}(PD_{i,C}), \Phi^{-1}(1-q), a_i) \\
&\quad - \frac{\phi(\bar{Y})}{2(1-q)(\bar{L}'(\bar{Y}))^2} \left[\frac{\partial v(\bar{Y})}{\partial w_i} \bar{L}'(\bar{Y}) - \frac{\partial \bar{L}'(\bar{Y})}{\partial w_i} v(\bar{Y}) \right]
\end{aligned}$$

⁶25,645 MM€ and 31,053 MM€ for the Pykhtin approximation and the Monte Carlo method, respectively.

The results for the expected shortfall based risk allocation can be observed in the right graph in Figure 3.2. We can see that the results obtained using the Pykhtin approximation are quite close to those provided by the Monte Carlo method.

We finish here our study of the Taylor expansions for the Vasicek (1987) model. An obvious further research is to extend the Pykhtin model for the case of market valuation. In this case, the different clients in a portfolio can migrate to several discrete states (not only to the default one) and, on each state, the loans have a different valuation or loss. Therefore, the estimation of $v(y)$ requires to consider all the possible combinations of states of clients i and j . This problem is quadratic in the number of clients and in the number of states. For instance, considering 156 clients and a rating scale with 17 grades, we must estimate $156 \times 156 \times 17 \times 17$ bivariate cumulative standard normal probabilities. This makes the utilization of the Pykhtin method less interesting in the case of market valuation. In the next Section we will apply the Taylor expansion ideas to approximate the loss distribution in the case of random recoveries in the Vasicek (1987) model.

3.5 Correlated random LGD

So far we have introduced the Taylor approximation method proposed in Pykhtin (2004) and used it to estimate the loss distribution of the portfolio of Spanish financial institutions and its risk contributions. We believe that this Taylor based approximation can be also used to approximate the case of random correlated recoveries. In the next sections we will introduce a simple model of random LGD and use the Taylor expansion ideas to approximate a more general model.

3.5.1 Simple random LGD model

We have developed a very simple random LGD model based on a single macroeconomic factor z that drives both default and recoveries in a granular portfolio. To keep it simple we assume that the recoveries follow a Bernoulli random variable that takes the values $\{0, 1\}$ depending on the value of the random variable $V_{j,LGD}$. According to this we have that

$$L = \sum_{j=1}^M x_j = \sum_{j=1}^M g_j D_{PD,j}(V_{j,PD}) D_{LGD,j}(V_{j,LGD})$$

Each client has a default value function ($V_{j,PD}$) and a LGD value function ($V_{j,LGD}$) given by

$$\begin{aligned} V_{j,PD} &= \rho_{j,PD}z + \epsilon_j \sqrt{1 - \rho_{j,PD}^2} \\ V_{j,LGD} &= \rho_{j,LGD}z + \psi_j \sqrt{1 - \rho_{j,LGD}^2} \end{aligned}$$

As usual, if $V_{j,PD} < \Phi^{-1}(PD_{j,C})$, the client defaults but now we also have that if $V_{j,LGD} < \Phi^{-1}(LGD_{j,C})$ the LGD is 100%. Conditional to a macroeconomic scenario, the loss distribution of

each subportfolio x_j is given by

$$x_j|z = \begin{cases} EAD_j & \text{with prob. } p = \Phi\left(\frac{\Phi^{-1}(PD_{j,C}) - \rho_{PD}z}{\sqrt{1 - \rho_{j,PD}^2}}\right) \Phi\left(\frac{\Phi^{-1}(LGD_{j,C}) - \rho_{LGD}z}{\sqrt{1 - \rho_{j,LGD}^2}}\right) \\ 0 & \text{with prob. } 1 - p \end{cases}$$

As the portfolio is made up of many identical clients, the loss distribution of the whole portfolio conditional to a macroeconomic scenario gets reduced to a single value

$$L|z = \sum_{j=1}^M EAD_j \Phi\left(\frac{\Phi^{-1}(PD_{j,C}) - \rho_{PD}z}{\sqrt{1 - \rho_{j,PD}^2}}\right) \Phi\left(\frac{\Phi^{-1}(LGD_{j,C}) - \rho_{LGD}z}{\sqrt{1 - \rho_{j,LGD}^2}}\right) \quad (3.15)$$

According to this, the 99.9% probability losses can be obtained just replacing z by $\Phi^{-1}(0.999)$ in equation (3.15).

The left graph in Figure 3.3 includes the results of this simplified model, using the Vasicek constant LGD model (Const.) as benchmark.⁷ We can observe that under this simplified model the random LGD always increases the risk as $\rho_{j,LGD}$ increases, in fact the case $\rho_{j,LGD} = 0\%$ is equivalent to the Vasicek constant LGD model. Intuitively, the effect of the random LGD gets completely diversified for a granular portfolio with $\rho_{LGD} = 0\%$.

[INSERT FIGURE 3.3 AROUND HERE]

The right graph of this Figure recovers the constant LGD level required to obtain a given 99.9% probability loss under the simplified random LGD model. In the worst case the average LGD of 40% should be multiplied by 2.5 generating a LGD of 100%. This means that the random LGD model with $\rho_{LGD} = 100\%$ is equivalent to a constant LGD model with LGD of 100%.

The Basel capital accord (see Basel (2006)) tries to capture the effect of the random LGD using what it is called the downturn LGD which is a stressed LGD. This simplified model would avoid the introduction of an arbitrary downturn LGD. The parameters $\rho_{j,PD}$ and $\rho_{j,LGD}$ can be calibrated using the historical variability of default and recovery rates. However, one would expect that the random independent LGD could generate some diversification effect that may reduce the risk in some situations.⁸ This is not the case in this simplified model because of the single macroeconomic factor assumption. The next Subsection considers a more complex model that allows for risk diversification.

3.5.2 Advanced random LGD model

We extend now the previous model to deal with non-granular portfolios and to consider two different macroeconomic factors, one driving the defaults and another one driving the recoveries. Hence, we

⁷We use the parameters $PD_{j,C} = 1\%$, $LGD_{j,C} = 40\%$, $\rho_{j,PD} = \sqrt{24\%}$, and $\rho_{j,LGD}$ varying in $[0, 1]$.

⁸The simplified model does not allow for risk diversification once ρ_{PD} and ρ_{LGD} are set.

have

$$\begin{aligned} V_{j,PD} &= \rho_{j,PD} z_{PD} + \epsilon_j \sqrt{1 - \rho_{j,PD}^2} \\ V_{j,LGD} &= \rho_{j,LGD} z_{LGD} + \psi_j \sqrt{1 - \rho_{j,LGD}^2} \end{aligned}$$

where the macroeconomic factors are correlated with $\text{corr}(z_{PD}, z_{LGD}) = \rho_{z_{PD}, z_{LGD}}$. Then, $V_{j,LGD}|Z_{PD} \sim N\left(\rho_{j,LGD}\rho_{z_{PD}, z_{LGD}}z_{PD}, \sqrt{1 - \rho_{j,LGD}^2\rho_{z_{PD}, z_{LGD}}^2}\right)$.

To apply the ideas in Pykhtin (2004) we can define \bar{L} as the losses of a constant LGD portfolio with $LGD_{j,C} = LGD_j^*$, that is,

$$\bar{L} = \sum_{j=1}^M EAD_j LGD_j^* \Phi\left(\frac{\Phi^{-1}(PD_{j,C}) - \rho_{j,PD}z_{PD}}{\sqrt{1 - \rho_{j,PD}^2}}\right)$$

Then, we have that

$$\begin{aligned} E(L|\bar{L}) &= E(L|z_{PD}) \\ &= \sum_{j=1}^M EAD_j \left[\Phi\left(\frac{\Phi^{-1}(LGD_{j,C}) - z_{PD} \rho_{j,LGD} \rho_{z_{PD}, z_{LGD}}}{\sqrt{1 - \rho_{j,LGD}^2 \rho_{z_{PD}, z_{LGD}}^2}}\right) \right. \\ &\quad \left. \times \Phi\left(\frac{\Phi^{-1}(PD_{j,C}) - z_{PD} \rho_{j,PD}}{\sqrt{1 - \rho_{j,PD}^2}}\right) \right] \\ &= \sum_{j=1}^M EAD_j \Phi(G_j(z_{PD})) p_j(z_{PD}). \end{aligned}$$

Therefore, if we set $LGD_j^* = \Phi(G_j(z_{PD}))$, we get that $E(L - \bar{L}|\bar{L}) = 0$ and equation (3.3) vanishes. To estimate equations (3.4) and (3.10) we have that

$$\begin{aligned} E(L^2|z_{PD}) &= \sum_{j=1}^M \sum_{i \neq j}^M EAD_j EAD_i p_j(z_{PD}) p_i(z_{PD}) \Phi_2(G_j(z_{PD}), G_i(z_{PD}), \rho_{i,j}) \\ &\quad + \sum_{j=1}^M EAD_j^2 \Phi(G_j(z_{PD})) p_j(z_{PD}) \end{aligned}$$

with

$$\rho_{i,j} = \frac{\rho_{i,LGD} \rho_{j,LGD} (1 - \rho_{z_{PD}, z_{LGD}}^2)}{\sqrt{1 - \rho_{i,LGD}^2 \rho_{z_{PD}, z_{LGD}}^2} \sqrt{1 - \rho_{j,LGD}^2 \rho_{z_{PD}, z_{LGD}}^2}}$$

Then

$$\begin{aligned}
v(z_{PD}) &= \sum_{j=1}^M \sum_{i \neq j}^M EAD_j EAD_i p_j(z_{PD}) p_i(z_{PD}) \Phi_2(G_j(z_{PD}), G_i(z_{PD}), \rho_{i,j}) \\
&\quad + \sum_{j=1}^M EAD_j^2 \Phi(G_j(z_{PD})) p_j(z_{PD}) \\
&\quad - \left(\sum_{j=1}^M EAD_j \Phi(G_j(z_{PD})) p_j(z_{PD}) \right)^2
\end{aligned} \tag{3.16}$$

In the case of granular portfolios, equation (3.16) gets simplified to

$$\begin{aligned}
v(z_{PD}) &= \sum_{j=1}^M \sum_{i=1}^M EAD_j EAD_i p_j(z_{PD}) p_i(z_{PD}) \Phi_2(G_j(z_{PD}), G_i(z_{PD}), \rho_{i,j}) \\
&\quad - \left(\sum_{j=1}^M EAD_j \Phi(G_j(z_{PD})) p_j(z_{PD}) \right)^2
\end{aligned}$$

Straightforward algebra from equation (3.16) allows us to compute the term $v'(z_{PD})$.⁹

Using expression (3.16) and its derivative, we can approximate the loss distribution of a given portfolio for different values of the parameter $\rho_{z_{PD}, z_{LGD}}$. The left graph in Figure 3.4 provides the loss distribution of a portfolio made up of 100 identical clients with $PD_{j,C} = 1\%$, $LGD_{j,C} = 40\%$, $\rho_{PD} = \sqrt{0.24}$, $\rho_{LGD} = \sqrt{0.24}$, and $\rho_{z_{PD}, z_{LGD}}$ varying in $[0, 1]$. It can be observed that the approximate method generates results that are very close to the exact ones.

[INSERT FIGURE 3.4 AROUND HERE]

The right graph in Figure 3.4 provides the loss distribution of the previous portfolio if we had a granular portfolio instead of 100 clients and illustrates the high accuracy of the approximation for this case.

We have tested the approximate method with other alternative LGD distributions. If $LGD_j = a_j + b_j z_{LGD} + c_j \psi_j$, we have that

$$\begin{aligned}
E(LGD_j | z_{pd}) &= a_j + \rho_{z_{PD}, z_{LGD}} b_j z_{PD} \\
E(LGD_j^2 | z_{pd}) &= a_j^2 + b_j^2 (1 - \rho_{z_{PD}, z_{LGD}}^2) + c_j^2 + 2a_j b_j \rho_{z_{PD}, z_{LGD}} z_{PD} \\
E(LGD_j LGD_i | z_{pd}) &= a_i a_j + (a_i b_j + a_j b_i) \rho_{z_{PD}, z_{LGD}} z_{PD} \\
&\quad + b_i b_j (1 - \rho_{z_{PD}, z_{LGD}}^2 + \rho_{z_{PD}, z_{LGD}}^2 z_{PD})
\end{aligned}$$

⁹However, our examples are based on a numerical derivative estimation of $v'(z_{PD})$ rather than on the analytical alternative due to the complexity of the formula. Additional results are available upon request.

As before we set $LGD_j^* = E(LGD_j|z_{pd})$, then we have

$$\begin{aligned}
E(L|\bar{L}) &= \sum_{j=1}^M EAD_j E(LGD_j|z_{pd}) p_j(z_{PD}) \\
v(z_{PD}) &= \sum_{j=1}^M \sum_{i \neq j}^M EAD_j EAD_i p_j(z_{PD}) p_i(z_{PD}) E(LGD_j LGD_i|z_{pd}) \\
&\quad + \sum_{j=1}^M EAD_j^2 E(LGD_j^2|z_{pd}) p_j(z_{PD}) \\
&\quad - \left(\sum_{j=1}^M EAD_j E(LGD_j|z_{pd}) p_j(z_{PD}) \right)^2
\end{aligned}$$

In Figure 3.5 we show the result of the approximate method and the Monte Carlo method under Gaussian recoveries.¹⁰ The correlation between the two macroeconomic factors varies between 0 and -1, where the value -1 implies the maximum positive correlation between default and loss given default. It can be observed that the model accuracy is very high.

[INSERT FIGURE 3.5 AROUND HERE]

Finally, for $LGD_j = e^{a_j + b_j z_{LGD} + c_j \psi_j}$, we get

$$E\left((LGD_j)^k \middle| z_{pd}\right) = e^{k\mu_1 + \frac{1}{2}(k\sigma_1)^2}, \quad k = 1, 2$$

with

$$\begin{aligned}
\mu_1 &= a_j + \rho_{z_{PD}, z_{LGD}} b_j z_{PD} \\
\sigma_1^2 &= b_j^2 (1 - \rho_{z_{PD}, z_{LGD}}^2) + c_j^2
\end{aligned}$$

Moreover, we obtain

$$E(LGD_j LGD_i | z_{pd}) = e^{\mu_2 + \frac{1}{2}\sigma_2^2}$$

with

$$\begin{aligned}
\mu_2 &= (a_i + a_j) + (b_i + b_j) \rho_{z_{PD}, z_{LGD}} z_{PD} \\
\sigma_2^2 &= c_i^2 + c_j^2 + (1 - \rho_{z_{PD}, z_{LGD}}^2) (b_i + b_j)^2
\end{aligned}$$

Figure 3.6 shows the results for the lognormal recoveries¹¹ and illustrates that the model performance is very similar to that of Gaussian recoveries.

[INSERT FIGURE 3.6 AROUND HERE]

¹⁰We use $PD_{j,C} = 1\%$, $\rho_{PD} = \sqrt{0.24}$, $a_j = 0.19$, $b_j = 0.08$, and $c_j = 0.11$.

¹¹We use $PD_{j,C} = 1\%$, $\rho_{PD} = \sqrt{0.24}$, $a_j = -1.88$, $b_j = 0.4$, and $c_j = 0.51$.

3.6 Conclusions

This paper has described the Taylor expansion based method introduced in Pykhtin (2004) and applied it to measure the credit risk of the Spanish financial system. We have shown that in the case of concentrated portfolios the approximation does not perform very well and it can underestimate the probability of high losses. We have also shown that the risk allocation based on this method does not generate accurate results for our portfolio. This is the case for the VaR and the ES based risk allocation. Additionally we have evaluated different ways to allocate the risk based on three different derivation rules.

We have also introduced a simple model to consider the effect of correlated random recoveries on the Basel capital requirements formula. The main drawback of this model is that it does not allow for any diversification effect as it assumes that the defaults and the recoveries are driven by the same macroeconomic factor.

Finally we have built a more flexible model that considers correlated random recoveries and we have used the Taylor expansion method to approximate the loss distribution. Our results suggest that this a robust method that generates accurate results and that can handle many different recoveries distributions. We believe that this method can be easily implemented under the Basel capital charge framework in order to consider the correlated random recoveries.

Appendix of Figures

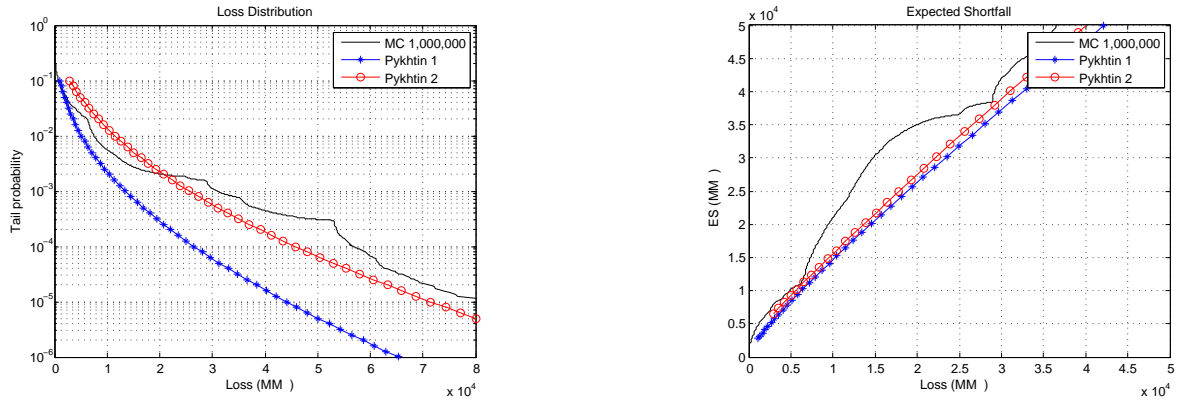


Figure 3.1: Loss distribution and expected shortfall for the Spanish financial institutions. These results are obtained using a Monte Carlo method with 1,000,000 simulations (MC) and the Pykhtin approximation when only the first term or both terms are considered (Pykhtin 1, Pykhtin 2). The black line indicates the MC results and the blue crosses and the red circles are the Pykhtin 1 and Pykhtin 2 results, respectively. Left and right graphs show the loss distribution and the expected shortfall results, respectively.

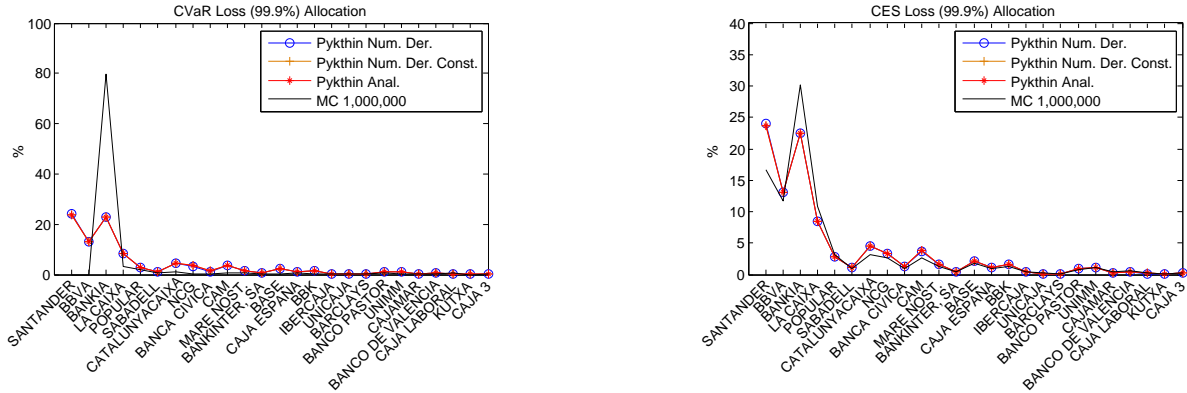


Figure 3.2: Risk allocation for the Spanish financial institutions. These results are obtained using a Monte Carlo method with 1,000,000 simulations (MC) and the Pykthin approximation derivatives, using a numerical derivation method and the analytical derivation formula (Pykthin Anal.). We use two types of numerical derivations, one that varies the values of a_j and b_f as λ varies on the derivation process (Pykthin Num. Der.) and another one that maintains them constant (Pykthin Num. Der. Const.). Left and right graphs show the loss allocation result based on the VaR and the expected shortfall criteria, respectively.

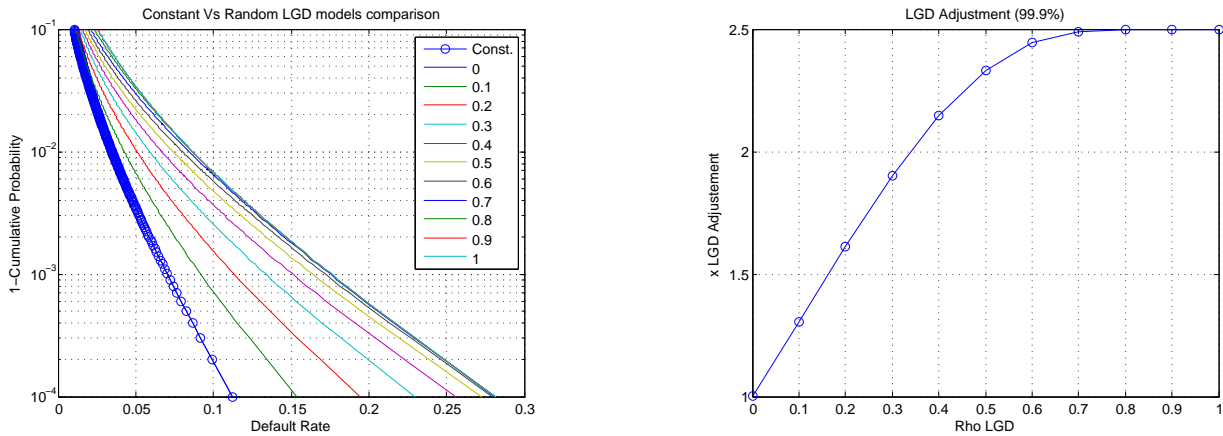


Figure 3.3: Simplified random LGD model. The left graph shows the tail loss for the cases of constant LGD (Const.) and for ρ_{LGD} varying in $[0, 1]$. The right graph provides the constant LGD multiplier required to achieve the random LGD 99.9% probability loss as the parameter ρ_{LGD} increases.

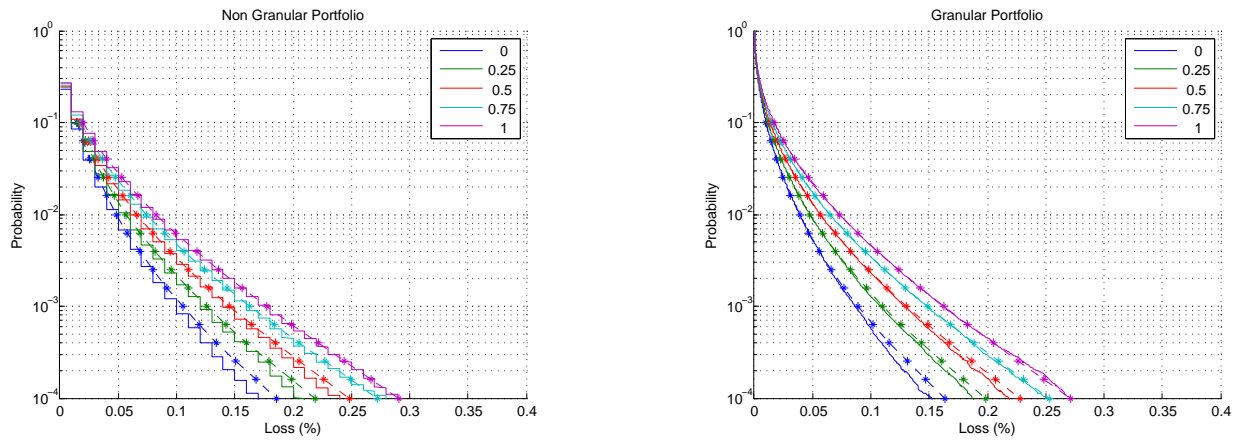


Figure 3.4: Advanced random LGD model. The left graph shows the results obtained for a portfolio with 100 equal clients for different values of the correlation between the two macroeconomic factors. The continuous lines represent the estimation of the loss distribution obtained using 1,000,000 Monte Carlo simulations and the dashed lines include the results obtained with the Taylor expansion based approximations. The right graph provides the results for a granular portfolio.

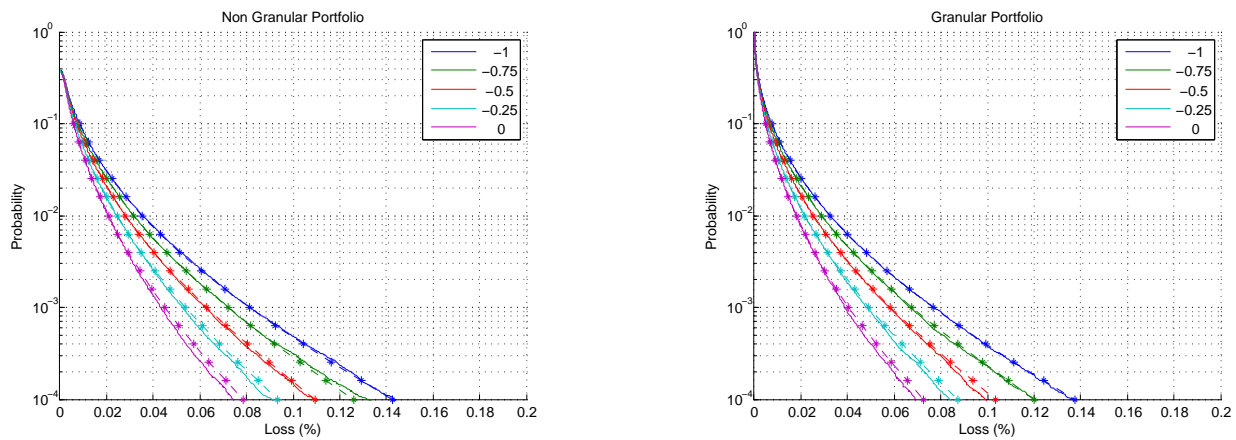


Figure 3.5: Advanced random LGD model under Normal LGD. The left graph shows the results obtained for a portfolio with 100 equal clients for different values of the correlation between the two macroeconomic factors. The continuous lines represent the estimation of the loss distribution obtained using 1,000,000 Monte Carlo simulations and the dashed lines include the results obtained with the Taylor expansion based approximations. The right graph provides the results for a granular portfolio.

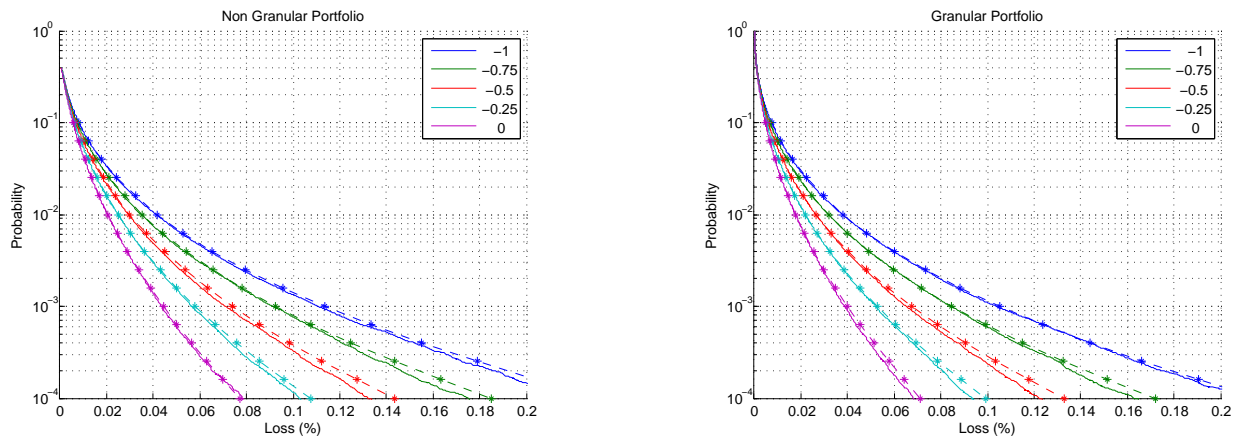


Figure 3.6: Advanced random LGD model under Lognormal LGD. The left graph shows the results obtained for a portfolio with 100 equal clients for different values of the correlation between the two macroeconomic factors. The continuous lines represent the estimation of the loss distribution obtained using 1,000,000 Monte Carlo simulations and the dashed lines include the results obtained with the Taylor expansion based approximations. The right graph provides the results for a granular portfolio.

Conclusions

The credit risk model proposed in Vasicek (1987) is the most common method to measure credit risk. However, under this model, estimating the loss distribution of a given portfolio usually requires Monte Carlo simulations. Estimating and allocating the credit risk of a portfolio using this type of simulations can be a very time demanding procedure. As a consequence alternative techniques have arisen to optimize the risk measurement process. This thesis has provided several contributions to the literature on credit risk from both methodological and empirical points of view.

From a methodological point of view, we have studied and extended the methods to estimate the credit risk of a portfolio presented in Pykhtin (2004), Glasserman and Li (2005), and Huang et al. (2007). We have modified the importance sampling and saddlepoint methods to work under random recoveries and market valuation. Additionally in the case of the saddlepoint method we have proposed a new risk allocation method that uses the Hermite polynomials. This new method has performed better than other approximate methods currently available (see Martin and Thompson (2001)) and, compared with an exact method, it has produced similar results but with fewer calculations.

Using the Taylor expansion based approximate methods we have developed a random *LGD* model that has allowed to consider the correlation between the default and the recoveries. Compared with a Monte Carlo method the results of this approximate method are very similar but this method has the advantage of providing a closed-form representation. Other methodological contributions that can be found in this thesis are the “*multimodal distributions*” and “*loop decoupling*” that allowed for a faster and more accurate estimation of the credit risk under the importance sampling method.

From an empirical point of view, we have obtained the loss distribution of the Spanish financial system based on the models and extensions previously mentioned. Additionally we have allocated the risk over the financial institutions in the financial system. We have used the data of the Spanish financial system at December 2010 and obtained the loss distribution under constant *LGD* for the three credit risk models analyzed. Compared with Campos et al. (2007) we have provided the detailed steps used to calibrate all the model parameters, even the *LGD*, and allowed institutions to be exposed to more than one macroeconomic factor.

We have also quantified the variability of the loss distribution by means of two different analysis: i) first, we have obtained the loss distribution at December 2007, a pre-crisis year and ii) secondly, we have studied the uncertainty in the model parameters, $\alpha_{f,j}$, PD , and LGD . This uncertainty comes from the lack of long enough historical time series to calibrate these parameters. The main qualitative conclusion of this analysis is that the LGD is the main source of uncertainty in the loss distribution estimation.

Under constant LGD the importance sampling and the saddlepoint methods have provided very accurate results; however the Taylor expansion based method introduced in Pykhtin (2004) has not been able to capture accurately the concentration profile of our portfolio. The same has happened for the risk allocation.

For the importance sampling and saddlepoint methods we have also obtained the loss distribution of the portfolio under random LGD and market valuation. We have provided all the detailed steps to calibrate the model parameters. Under random LGD the loss distribution is slightly shifted to the right compared with the constant LGD case and the results for the different random LGD models have been very similar, even between the pure macroeconomic models and the mixed idiosyncratic and macroeconomic models. When we have moved to the market valuation, the loss distribution gets considerably shifted to the right as we have more states that produce high losses.

Regarding the risk allocation we have shown that the risk allocated is very sensitive to the allocation criteria and to the loss definition. In the extreme, under constant LGD and VaR based risk allocation, Santander and BBVA have been allocated a null risk. In the case of the Hermite polynomials based risk allocation new method that we have proposed, the results for the portfolio have been very satisfactory. The method has shown to be more accurate than other approximate methods (see Martin and Thompson (2001)) while requiring a similar number of calculations.

Given our results for the portfolio we suggest not to focus on a unique model or risk allocation criteria but to test several of them. We consider that this thesis has contributed to the previous literature with new methodological issues and empirical results related to the measurement of credit risk. The methods we have proposed can be used directly by practitioners and regulators to monitor the credit risk of given portfolios. Finally, in the case of regulators some of the ideas proposed and developed in this thesis can serve as a reference to identify and monitor the so called systemically important financial institutions (SIFIs).

Conclusiones

El modelo de riesgo de crédito propuesto en Vasicek (1987) es el método más habitual para medir el riesgo de crédito. Sin embargo, estimar la distribución de pérdidas de una cartera bajo este modelo suele requerir la utilización de simulaciones Monte Carlo. La estimación y reparto del riesgo de crédito de una cartera empleando este tipo de método de simulación puede ser un proceso que requiera un elevado consumo de tiempo. Como consecuencia han surgido técnicas alternativas para optimizar el proceso de medición del riesgo. Esta tesis proporciona diversas contribuciones a la literatura del riesgo de crédito tanto desde un punto de vista metodológico como empírico.

Desde un punto de vista metodológico, hemos estudiado y extendido los métodos para cuantificar el riesgo de crédito de una cartera presentados en Pykhtin (2004), Glasserman and Li (2005) y Huang et al. (2007). Hemos modificado los métodos de muestreo por importancia y de *saddlepoint* para los casos de recuperaciones aleatorias y valoración de mercado. Adicionalmente, en el caso de los métodos de *saddlepoint*, hemos propuesto un nuevo método para repartir el riesgo que emplea polinomios de Hermite. Este nuevo método se comporta mejor que otros métodos aproximados disponibles actualmente (ver Martin and Thompson (2001)) y, comparado con un método exacto, produce resultados similares pero con un número menor de cálculos.

Empleando los métodos aproximados basados en expansiones de Taylor hemos desarrollado un modelo de *LGD* aleatoria que nos permite considerar la correlación entre el incumplimiento y las recuperaciones. Comparado con el método Monte Carlo los resultados de este método aproximado son muy similares pero este método tiene la ventaja de proporcionar una representación analítica. Otras contribuciones metodológicas que se encuentran en esta tesis son las llamadas “*distribuciones multimodales*” y el “*desacople de bucles*” que permiten una estimación del riesgo de crédito más rápida y precisa cuando se utiliza el método del muestreo por importancia.

Desde un punto de vista empírico, hemos obtenido la distribución de pérdidas del sistema financiero español empleando los modelos y extensiones mencionadas previamente. Adicionalmente hemos repartido el riesgo entre las distintas instituciones financieras. Hemos empleado datos del sistema financiero español correspondientes a Diciembre de 2010 y obtenido la distribución de pérdidas bajo *LGD* constante para los tres modelos de riesgo de crédito analizados. Comparado con Cam-

pos et al. (2007) presentamos los pasos detallados seguidos para calibrar todos los parámetros del modelo, incluida la *LGD*, y permitimos que las instituciones financieras estén expuestas a más de un factor macroeconómico.

También hemos cuantificado la variabilidad de la distribución de pérdidas realizando dos tipos de análisis: i) primero, hemos obtenido la distribución de pérdidas en Diciembre de 2007, una fecha anterior a la crisis y ii) segundo, hemos estudiado la incertidumbre en los parámetros $\alpha_{f,j}$, *PD* y *LGD* del modelo. Esta incertidumbre proviene de la ausencia de series temporales suficientemente largas para calibrar estos parámetros. La principal conclusión cualitativa de este análisis es que la *LGD* es la principal fuente de incertidumbre en la estimación de la distribución de pérdidas.

Considerando *LGD* constante, los métodos de muestreo por importancia y de *saddlepoint* han proporcionado resultados muy precisos; sin embargo el método basado en la expansión de Taylor propuesto en Pykhtin (2004) no ha sido capaz de capturar adecuadamente el perfil de concentración de nuestra cartera. La misma conclusión se obtiene para el reparto del riesgo.

Adicionalmente, para los métodos de muestreo por importancia y de *saddlepoint*, hemos obtenido la distribución de pérdidas de la cartera considerando *LGD* aleatoria y valoración de mercado. Hemos proporcionado todos los pasos detallados para calibrar los parámetros del modelo. Bajo *LGD* aleatoria la distribución de pérdidas se desplaza ligeramente a la derecha comparada con el caso de *LGD* constante y los resultados para los distintos modelos de *LGD* aleatoria han sido similares, incluso entre los modelos de recuperaciones macroeconómicas y los modelos de recuperaciones macroeconómicas e idiosincráticas. Cuando consideramos la valoración de mercado, la distribución de pérdidas se desplaza considerablemente a la derecha pues, en este caso, tenemos más estados que producen pérdidas elevadas.

Hemos mostrado que el reparto del riesgo es muy sensible al criterio de reparto y a la definición de pérdidas empleada. En el caso extremo de *LGD* constante y reparto de riesgo basado en el *VaR*, se asigna un riesgo nulo a Santander y BBVA. En el caso del nuevo método de reparto basado en polinomios de Hermite, los resultados para la cartera han sido muy satisfactorios. El método ha generado una mayor precisión que otros métodos aproximados (ver Martin and Thompson (2001)) mientras que requiere un número similar de cálculos.

De acuerdo a nuestros resultados empíricos sugerimos contrastar varios modelos o criterios de reparto. Consideramos que esta tesis ha contribuido a la literatura previa sobre la medición del riesgo de crédito con nuevas propuestas metodológicas y resultados empíricos novedosos y útiles a nivel práctico. Los métodos que hemos propuesto pueden ser empleados directamente por reguladores y la industria bancaria para monitorizar las carteras. Finalmente, en el caso de los reguladores, algunas de las ideas propuestas y desarrolladas en esta tesis pueden servir como referencia para identificar y monitorizar las conocidas como instituciones financieras sistémicamente importantes (SIFIs).

References

- Abate, J. and Whitt, W. (1995). Numerical inversion of Laplace transforms. *ORSA Journal on Computing*, 7 (1), 36-43.
- Abramowitz, M. and Stegun, I.A. (1964). *Handbook of mathematical functions with formulas, graphs and mathematical tables*, Courier Dover Publications.
- Allen, F., Bartiloro, L. and Kowalewski, O. (2006). The Financial System of the EU 25, In *Financial Development, Integration and Stability: Evidence from Central, Eastern and South-Eastern Europe*, edited by K. Liebscher, J. Christl, P. Mooslechner and D. Ritzberger-Grünwald, Edward Elgar, Cheltenham.
- Altman, E.I., Brady, B., Resti, A. and 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. and 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. and 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. and 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>.

- Bennet, R.L. (2002). Evaluating the adequacy of the deposits insurance fund: A Credit-Risk modeling approach. *FDIC Working Papers*, No. 2001-2002.
- Bennet, R.L. and Unal, H. (2011). The cost effectiveness of the private sector reorganization of failed banks. *FDIC Working Papers*, No. 2009-2011.
- Campos, P., Yagiüe, M. and Chinchetru, I. (2007). Un nuevo marco de seguro de depositos para España. Banco de España *Estabilidad Financiera*, 12, 93-110.
- Cantor, R., Hamilton, D. and Tennant, J. (2007). Confidence Intervals for Corporate Default Rates. *Moody's Special Comment*, Moody's Investors Service.
- Carbó, S., Humphrey, D., Maudos, J. and 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. and Schoutens, W. (2011). Applying Credit Risk Techniques to Design an Effective Deposit Guarantee Schemes Funds. *International Risk Management Conference working Paper*.
- Castro, C. (2012). Confidence sets for asset correlations in portfolio credit risk. *Revista de Economía del Rosario*, 15 (1), 19-58.
- Chance, D. M. and Agca, S. (2003). Speed and Accuracy Comparisons of Bivariate Normal Probability Distribution Approximations for Option Pricing. *Journal of Computational Finance*, 6, 61-96.
- Daniels, H.E. (1954). Saddlepoint approximations in statistics. *The Annals of Mathematical Statistics*, 25 (4), 631-650.
- Daniels, H.E. (1980). Exact saddlepoint approximations. *Biometrika*, 67 (1), 59-63.
- Daniels, H.E. (1987). Tail probability approximations. *International Statistical Review*, 55 (1), 37-48.
- Dullmann, K., Kull, J. and Kunisch M. (2008). Estimating asset correlations from stock prices or default rates which method is superior?. *Deutsche Bundesbank Banking and Financial Studies*, No 04/2008.
- Fitch (2009). Global Corporate Finance 2008 Transition and Default Study. *Credit Market Research*, Fitch Corporation.

- García-Céspedes, R. and Moreno, M. (2014). Estimating the distribution of total default losses on the Spanish Financial system, mimeo, BBVA and University of Castilla-La Mancha.
- García-Céspedes, R. and Moreno, M. (2014b). Extended saddlepoint methods for credit risk measurement, mimeo, BBVA and University of Castilla-La Mancha.
- Genz, A. (2004). Numerical Computation of Rectangular Bivariate and Trivariate Normal and t Probabilities. *Statistics and Computing*, 14, 151-160.
- Glasserman, P. (2005). Measuring Marginal Risk Contributions in Credit Portfolios. *Journal of Computational Finance*, 9 (2), 1-41.
- Glasserman, P. and Li, J. (2005). Importance Sampling for Portfolio Credit Risk. *Management Science*, 51 (11), 1643-1653.
- Glasserman, P. and Ruíz-Mata, J. (2006). Computing the credit loss distribution in the Gaussian copula model: a comparison of methods. *Journal of Credit Risk*, 2 (4), 33-66.
- Gordy, M. (2000). A comparative anatomy of credit risk models. *Journal of Banking and Finance*, 24, 119-149.
- Gourieroux, C., Laurent, J.P. and Scaillet, O. (2000). Sensitivity Analysis of Values at Risk. *Journal of Empirical Finance*, 7 (3-4), 225-245.
- Grundke, P. (2009). Importance sampling for integrated market and credit portfolio models. *European Journal of Operational Research*, 194 (1), 206-226.
- Gupton, G.M. and Stein, R.M. (2002). LossCalcTM: Moody's model for predicting loss given default (LGD). *Moody's Investor Service*, February.
- Hanson, S.G. and 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. and Lo Duca, M. (2006). The Performance of the European Financial System, mimeo, European Central Bank, August.
- Huang, X., Oosterlee, C.W. and Van Der Weide, H. (2007). Higher Order Saddlepoint Approximations on the Vasicek Portfolio Credit Loss Model. *Journal of Computational Finance*, 11, 93-113.
- Hull, J. (2011). *Options, Futures, and Other Derivatives*, eighth edn. Prentice Hall.
- James, C. (1991). The loss realized in banks failures. *Journal of Finance*, 46, 1223-1242.

- Kaufman, G. (2004). FDIC Losses in Bank Failures: Has FDICIA made a difference?. *Federal Reserve Bank of Chicago working paper*.
- Kuritzkes, A., Schuermann, T. and Weiner, S. (2002). Deposit Insurance and Risk Management of the U.S. Banking System: How Much? How Safe? Who Pays?. *Wharton School Center for Financial Institutions Working Papers*, 02-02.
- Lee, J., Wang, J. and Zhang, J. (2009). The relationship between average asset correlation and default probability. *Moody's KMV Company*.
- López, J. (2002). The Empirical Relationship between Average Asset Correlation, Firm Probability of Default and Asset Size. *Federal Reserve Bank of San Francisco*.
- Martin, R. and Thompson, K. (2001). VaR: who contributes and how much?. *Risk Magazine*, 14 (8), 99-103.
- Martin, R. and Wilde, T. (2002). Unsystematic Credit Risk. *Risk Magazine*, 15 (11), 123-128.
- Moody's (2005). Default, Losses and Rating Transitions on Bonds Issued by Financial Institutions: 1983-2004. *Moody's Investor service*.
- Moody's (2009). Corporate Default and Recovery rates 1920-2008. *Moody's Investor service*.
- Morone, M., Cornaglia, A. and Mignola, G. (2012) Determining marginal contributions of the economic capital of credit risk portfolio: an analytical approach, mimeo, Intesa Sanpaolo.
- Owen, D. B. (1956). Tables for computing bi-variate normal probabilities. *Annals of Mathematical Statistics*, 27(4), 1075-1090.
- Pykhtin, M. (2004). Multi-factor adjustment. *Risk Magazine*, 17 (3), 85-90.
- Qibin, C., Amnon, L. and Nihil, P. (2009). Understanding asset correlation dynamics from stress testing. *Moody's KMV Company*.
- Reid, N. (1988). Saddlepoint methods and statistical inference. *Statistical Science*, 3 (2), 213-238.
- Reitan, T. and 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. and 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.
- Tsao, M. (1995). Asymptotic and numerical methods for approximating distributions. *PhD Thesis*, Simon Fraser University.
- Unal, H., Madan, D. and 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*.
- Vasicek, O.A. (1996). A Series Expansion for the Bivariate Normal Integral. Working Paper, *KMV Corporation*.
- Voropaev, M. (2011). An analytical framework for credit portfolio risk measures. *Risk Magazine*, 24 (5), 72-78.