III MODELLING WEEK UCM Master in Mathematical Engineering - UCM Madrid, June 22-30, 2009

"Modelling default risk through macroeconomic factor evolution"



Participants: Carmen Guaza Daniel La Orden Jose Miguel Rodriguez Duccio Brogi Asgeir Birkisson Ana Mª Guerra

> Coordinator: Nacho Villanueva

Table of Contents

1.	Introducction. Description of the problem	3
2.	Definition and Values of Variables.	5
3.	Calculation of Loss and Value at Risk for Risk Units.	9
4.	Calculation of Value at Risk and Loss for the Combination of the two Risk Units	14
5.	Study the impact of the Macro Economic Factors	20
6.	Conclusions.	24

1. Introducction. Description of the problem.

Complubank has two businesses: credit cards and mortgages. Therefore, it has two risk unit (RU), each composed of a number of contracts. The information that Management Solution gives us are the following:

- The Loss Given Default (LGD) and the Exposure at Default (EAD) of each contract. Both the LGD and EAD are considered deterministic variables.
- The probability of default (PD) is an unknown variable but we are supposing that all contracts in a given RU have same PD, which corresponds to the probability of default within 1 year horizon. Furthermore, this probability should not be conditioned to any particular state of the economy realization.
- For each RU we have a time series of conditional PD^{Z} for a whole economic cycle (from 1990 to 2007). In each of these series, for every three month period, we have the PD of the RU conditioned to the macroeconomic situation in that moment.
- Also, we assume that the relation between quarterly PD^{Z} known and the unconditional PD (unknown) is given by the Vasicek function:

$$pd^{Z} = N\left(\frac{N^{-1}(pd) - \sqrt{\rho}z}{\sqrt{1-\rho}}\right)$$

where:

- Z is assumed to be a N(0,1), common to all contracts in a RU. It is called systemic factor, and represents the state of the economy in given time horizon. It is the fundamental risk driver of the loss.
- ρ is an unknown parameter between 0 and 1.
- Correlation Θ between the Z factors of risk units can be derived from the correlation between systemic factors of each RU.

• Assuming EAD and LGD are deterministic, in a **given horizon**, each RU has a loss distribution given by:

$$L = \sum_{i=1}^{N} Ber(PD_i^z) \cdot EAD_i \cdot LGD_i$$

where:

- *N* is the number of contratcs in the risk unit.
- *PD_i^Z* is the probability of default along the horizon, conditionned to a state of the economy. Such state of the economy is the risk factor that codifies the evolution of the losses.
- Loss distribution of *M risk units* of a bank will be given by:

$$L = L_1 + \ldots + L_M$$

In order to properly measure the risk, it is essential to take into account the **correlation** between contracts within the same risk unit and the correlation between different risk units.

2. Definition and Values of Variables.

a. Probability of Default.

PD values are the probability of default independent of the economic situation (i.e. unconditional probability of default).

The information that we have to calculate de PD are the following:

Year	Conditional PD A	Conditional PD B
1990	1.40%	7.84%
1991	2.16%	5.44%
1992	3.08%	6.38%
1993	4.56%	9.57%
1994	5.03%	5.56%
1995	1.88%	4.86%
1996	2.41%	3.21%
1997	1.88%	1.80%
1998	1.74%	0.99%
1999	1.52%	0.97%
2000	1.63%	0.87%
2001	1.38%	0.54%
2002	1.18%	0.44%
2003	1.04%	0.40%
2004	0.63%	0.47%
2005	0.52%	0.43%
2006	0.44%	0.44%
2007	0.89%	0.62%

Using the data above we have obtained PD's values as the average of the PD at each year for each risk unit.

Average (PD)	1,85%	2,82%

After calculating the PD, we could calculate a constant which is used to calculate when someone does a default as we can see below:

$$P(Default = 1) = P(X < c)$$

$$P(\sqrt{\rho}z + \sqrt{\rho - 1}\varepsilon < c) = N(c) \longrightarrow c = N^{-1}(PD)$$

With the information that we have and the formula above, we calculate a constant, c, for each unit risk.

	RU A	RU B
С	-2.5850	-1.97020

b. Relation between doing default and macro economic factors, ρ .

 ρ values are a measure of the independence of the debtors. If $\rho = 1$, this implies all debtors are 100% correlated (if one defaults, everybody will) but if $\rho = 0$ all debtors default independently.

We use the following relation to obtain ρ values for both RU:

$$\rho = \frac{Varianza(N^{-1}(pd^{z}))}{\left(1 + Varianza(N^{-1}(pd^{z}))\right)}$$

giving

	RU A	RU B
ρ	6,67%	19,38%

The relation between ρ and PD^{z} is deduced in the following lines:

$$Varianza(N^{-1}(pd^{z})) = Varianza\left(\frac{N^{-1}(pd) - \sqrt{\rho z}}{\sqrt{1 - \rho}}\right) =$$
$$= \frac{1}{(\sqrt{1 - \rho})^{2}}Varianza\left(N^{-1}(pd) - \sqrt{\rho z}\right) =$$
$$= \frac{1}{1 - \rho}\left(Varianza\left(N^{-1}(pd)\right) + Varianza\left(-\sqrt{\rho z}\right)\right) =$$
$$= \frac{\rho}{1 - \rho}$$

$$Varianza(N^{-1}(pd^{z})) \cdot (1-\rho) = \rho$$
$$Varianza(N^{-1}(pd^{z})) = \rho \cdot (1+Varianza(N^{-1}(pd^{z})))$$
$$\rho = \frac{Varianza(N^{-1}(pd^{z}))}{(1+Varianza(N^{-1}(pd^{z})))}$$

c. Correlation between risk units, θ .

One of the last variables we have to estimate is Θ , which in our model represents the correlation between the probabilities of default of the two risk units, through the following formula:

$$\Theta = Co \text{ var } iance(Z_A, Z_B) = Correlation(Z_A, Z_B) =$$

$$= Correlation\left(\left(\frac{N^{-1}(pd_A) - \sqrt{\rho}z_A}{\sqrt{1-\rho}}\right), \left(\frac{N^{-1}(pd_B) - \sqrt{\rho}z_B}{\sqrt{1-\rho}}\right)\right) =$$

$$= Correlation\left(N^{-1}(pd_A^Z), N^{-1}(pd_B^Z)\right)$$

where pd_A^Z and pd_B^Z represents the contitioned probability, obtaining the value:

Θ=79,7%

3. Calculation of Loss and Value at Risk for Risk Units.

Once whe have calculated these values, we are interested in obtaining the loss distribution of each RU in order to measure the risk. To achive this we have simulated de stochastics variables that take part in the process as shown in this formula:

$$P(D=1/z_t=z) = P(\sqrt{\rho}z_t + \sqrt{\rho-1}\varepsilon < c$$

Being z the macroeconomic situation and epsilon the idiosicratic economical situation for each contract.

These are the steps we followed foe each RU:

- 1. Generate a random variable z (N(0,1)) to describe a macroeconomic situation.
- 2. Get the PD for this z

$$PD^{z} = P(\varepsilon < \frac{N^{-1}(PD) - \sqrt{\rho}z_{t}}{\sqrt{1 - \rho}}) = N(\frac{N^{-1}(PD) - \sqrt{\rho}z_{t}}{\sqrt{1 - \rho}})$$

- Generate a Bernoulli distributed random variable epsilon (N(0,1)) based on PD for each client like a individual feature of the client. If PD is higher than epsilon then the default event occurs, not in the other case.
- 4. The loss generated for this client is:

$$ClientLoss = Ber(PD^{z}) \cdot EAD \cdot LGD$$

5. The total loss of the portofolio is the sum of the loss of each client:

$$L = \sum_{i=1}^{n} Ber(PD_{i}^{z}) \cdot EAD_{i} \cdot LGD_{i}$$

where n denotes the number of clients in each unit.

6. We simulate the process 10000 times and get the distribution shown below:



The Value at Risk of the the distributions will give us a measure of the risk. For percentiles 99.9% and 99.97%, we get these results:

PORTFOLIO A	PORTFOLIO B
$VaR^{99.90} = 799.21$	$VaR^{99.90} = 1684.5$
$VaR^{99.97} = 891.1975$	$VaR^{99.97} = 1890.2$

Now, we analyze how the performing of VaR is affected by the variation of Rho in both portofolios.

Results for 100.000 simulations for portofolio A:



This also gives us the numerical derivative with respect to var around $Rho_A = 0.067$

$$\frac{\partial VaR^{99.9\%}}{\partial \rho A} = 8990 \not \rho A \qquad \qquad \frac{\partial VaR^{99.97\%}}{\partial \rho A} = 9810 \not \rho A$$

Results for 100.000 simulations for portofolio B:



This also gives us the numerical derivative with respect to var around $Rho_B = 0.1938$

$$\frac{\partial VaR^{99.9\%}}{\partial \rho B} = 9140 \, \cancel{\rho}B \qquad \qquad \frac{\partial VaR^{99.97\%}}{\partial \rho B} = 11040 \, \cancel{\rho}B$$

In both cases the increasing of rho increases VaR. This means that the more the default event depends on the macroeconomic factor, the more the risk increases.

4. Calculation of Value at Risk and Loss for the Combination of the two Risk Units.

The results we obtained before are valid for independent portofolios but since we have them aggregated and correlated, we may get different results.

To simulate the perform of correlated macroeconomic factors, we use the Cholesky factorization and these are the formulas we got:

$$Z_{PortfolioA} = random(N(0,1))$$
$$Z_{PortfolioB} = \Theta Z_{PortfolioA} + \sqrt{1 - \Theta^2} random(N(0,1))$$

Where Θ is the correlation between portfolios.

Like we did for a single units we can do now for both combinated, and we get the following Loss distribution:



Comparing the results for var we can appreciate that combining the units the risk is reduced.

	Var99.9	Var99.97
RU A	810.5599	1050.1
RU B	1902.7	2297.5
RU AB	2532.9	2982.9

a. Variation of VaR respect $\mathsf{Rho}_{\mathsf{A}}$ and $\mathsf{Rho}_{\mathsf{B}}$

The graphs bellow represent each of these parameters influeces the variation of VaR for the aggregated loss distribution.

As we got before, the increase of both Rho_A and Rho_B affects in the same way to VaR, but the first one is the more influent.





b. Variation of rho respect the whole portfolio

We now estimate the first derivate of VaR (99.97% and 99.9%) respect of Θ . The plot obtained is shown below:





This also gives us the numerical derivative with respect to $\,\Theta\approx 0.8$

$$\frac{\partial VaR^{99.9\%}}{\partial \Theta} = 501 \oint \Theta \qquad \qquad \frac{\partial VaR^{99.97\%}}{\partial \Theta} = 698 \oint \Theta$$

In this case, the more correlated are the portofolios, the more coincidence in their performing in the extreme events, whis is what VaR measures, and as result of that the losses grow.

c. Variation of VaR respect PD





As we expected PD and VaR have a direct relation.

5. Study the impact of the Macro Economic Factors

The aim is to establish a relationship between the macro economic factors and the systemic factors for each risk unit. For this purpose we decided to use a linear regression model after a long study of the problem. We calculated the systemic factors and then normalized booth the systemic and the macro economic factors. Using Excel, we calculate the regression coefficients and the errors that are involved since we need them in our further work. Our model therefore is:

$$z^{j} = \sum_{i=1}^{6} \beta_{i}^{j} F_{i} + \beta_{\varepsilon}^{j} \varepsilon^{j}, j \in \{A, B\}$$

where

	Risk Unit A	Risk Unit B
b1	0,266197489	-0,077627772
b2	-0,79495506	-0,621544539
b3	0,08264054	-0,12181275
b4	0,048293754	-0,564335208
b5	0,193449818	-0,039757321
b6	-0,021242591	-0,073168947
b_epsilon	0,362743864	0,15134132

So we conclude that the systemic factors depend of the values of all our macro economic factors. However this influence depends on the magnitude of the correspondent beta.

Also since $F_i, \varepsilon^j \forall i = 1, 2, 3, 4, 5, 6 \quad j \in \{A, B\}$ has a Normal (0, 1) as distribution function we know that Z also has a normal distribution.

Now to see the impact we have in the VaR for the different factors, find the necessity to fix one of the factors in the previous formula and get:

$$z^{j} = \beta_{1}^{j} F_{1} + \sum_{i=2}^{6} \beta_{i}^{j} F_{i} + \beta_{\varepsilon}^{j} \varepsilon^{j}, j \in \{A, B\}$$

Where Z does not follow a Normal (0, 1) any more. The only thing to be noticed is that to simulate the Z is needed the Cholesky matrix of the correlation matrix shown below.

	GDP	Unemployment	House_Price	E3M	IBEX	Inflation
GDP	1	-0,37297	-0,30282	-0,40533	-0,11382	-0,02879
Unemployment	-0,37297	1	-0,31707	0,57914	-0,06146	-0,39557
House_Price	-0,30282	-0,31707	1	0,26396	-0,11747	0,58063
E3M	-0,40533	0,57914	0,26396	1	-0,38748	-0,11351
IBEX	-0,11382	-0,06146	-0,11747	-0,38748	1	-0,19195
Inflation	-0,02879	-0,39557	0,58063	-0,11351	-0,19195	1

Given the formula we use it to see the variation around some values of GDP. In this case the use of the standardized variables is needed in order to use the formula. For the 6 variables the results obtained for the simulation are:





Var99.9% for RU AB as a function of IBX Var99.97% for RU AB as a function of IBX 3000



The derivate for the variables around the media of each variable are:

	VAR 99,90	VAR 99.97
GDP	-162.5861	-175.563
HOUSE PRIZE	-57.2944	-59.2271
UNEMPLOYMENT	64.6763	70.9143
EURIBOR 3m	118.0559	133.8181
IBX	-69.7972	-81.2009
INFLATION	-82.7152	-94.3398

6. Conclusions.

The results obtained show that the most influenced factor is the unemployment. However the variation of this factor does not show that this factor gives the greatest variation of VaR.

It is important to underline that the variation of VaR99.97 is more sensible at changes than the VaR99.90.

However the most relevant factor is the relation between variables, due to the fact that we are looking to extreme results. Furthermore it could be important to see the relation between VAR and profit which we have not studied.