Modelling Week 2012

DETERMINING THE INFLUENCE OF MACROECONOMIC FACTORS ON THE RUN-OFF OF BANKING DEPOSITS

Enrique Ayensa Aparicio

Javier Barba Rueda

Elena Tanarro Santamaría

Rodrigo Zubieta García

Miguel González Hidalgo

Problem proposed by

Management Solutions

CONTENS

1. INTRODUCTION

In the times of plenty that preceded the current financial crisis, there was a little or no regulation on the amount of liquid assets (i.e. cash and quick-to-sell instruments) that a bank should keep in case its clients suddenly claimed it.

This has led to several cases of dramatic situations where a large number of clients wished to withdraw their savings but the bank would not allow it due to lack of liquidity. Some of them, like the Northern Rock case in 2007 or the Argentinian "corralito" in 2001-2002 became famous as examples of bad liquidity management.

Therefore, it has been proven that liquidity is not an inexhaustible source and thus banks must keep a buffer of liquid assets to be ready to return the cash to its clients in a short time in case they should claim it.

To regulate this, the Basel Committee on Banking Supervision issued some recommendations in a paper that has been informally named "Basel III", and that has recently become mandatory in all European Union countries.

In this paper, among other prudential measures, banks are advised to monitor a liquidity coverage ratio (LCR), which essentially shows whether a bank will be able to satisfy its due payments over the next 30 days.

Its expression is:

$$\text{LCR} = \frac{\text{Stock of high quality liquid assets}}{\text{Net cash out flows over a 30-day time period}} \geq 100\ \%$$

The denominator includes several parameters, one of which is the run-off rate: the percentage of the amount of deposits that presumably will be withdrawn by customers in severe stress scenario.

Basel III estimates the run-off rate to be 5% for stable deposits and 10% for less stable deposits. However, it does not provide any mathematical or economical background for these rates.

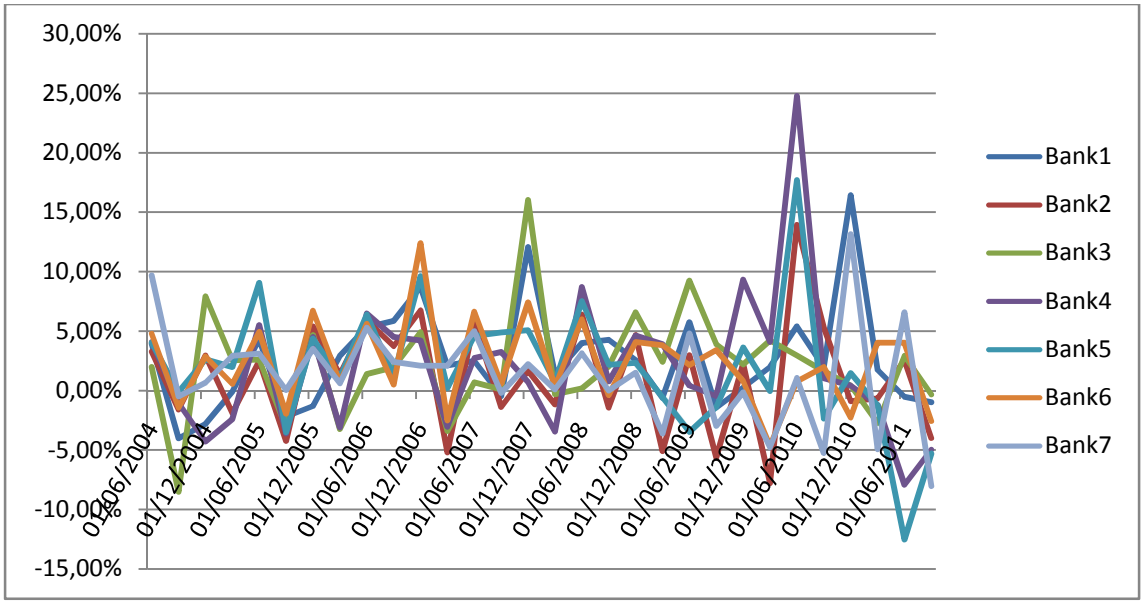
Our main goal in this problem is to challenge these rates through rigorous mathematical modeling, using real data from Spanish banks over a seven-year-period.

2. DATA INSIGHT

We've worked with real data from seven Spanish Banks over seven year period (from June 2004 to September 2011). The time interval between data is three months. So, for each bank we have an amount of 30 data.

In a first approximation of the problem, we plot the data received. We plot the total amount of current and term accounts by quarter in the same graph.

The graph obtained is as follows:



Within the aim of explaining the data, we have been studying the relationships between 19 macroeconomics variables and the amount and variations in retail deposits. Some of them are: Euribor, Exchange Rate, Unemployment Rate, Consumer Price Index, GDP, House Price, Stock Exchange, IBEX 35, Saving Rate, Available Rent, Legal Interest Rate, Treasury Bonds Volume, Balance of Assets in Pension Funds and Investment Funds Volume.

Before discussing about the model we would use to solve the problem, we decided to reject some of these variables, because of 19 variables are too many to try to explain a mathematical model and there are too big correlations among them.

To decide which variables we are going to consider we studied then those correlations and also the significant of each explanatory variable and our target variable (deposits variations average)

For example, we could see a table with the significance of variables three-month Euribor (EUR3M) and the available rent for each bank:

Corelaciones	EUR3M	AVAILABLE RENT
Bank 1	13,06%	46,40%
Bank 2	47,46%	59,09%
Bank 3	-19,91%	52,93%
Bank 4	3,26%	26,00%
Bank 5	14,55%	38,13%
Bank 6	6,28%	54,43%
Bank 7	15,59%	65,76%
Ponderate Bank	17,19%	70,67%

We observe that available rent has more influence than EUR3M.

3. DISCUSSION OF THE PROBLEM

To try to understand the problem we discussed about the following questions:

1) How to use the data

The variable to study is the quarterly change volume of deposits of each bank. As this data depends on the total volume of each bank, we would appreciate if we were to use the variable, or if we were going to perform some transformation so that the volume of each bank had no influence on the model.

2) A mathematical model against seven models (one for each bank)

With seven models would get better predictions because the macroeconomic variables impact differently on each bank. The disadvantage is that if we want to do the same study on a new bank, we have to do a new model for it.

Using a single model, we would lose accuracy in predictions, but in return we could extrapolate the model to other banks in future studies.

3) Data partition

We consider to do a model with the data before the start of the crisis and another model for the crisis period, because of banks' behavior depending on the scenario where we are. The main problem with this consideration is to establish on what date we place the beginning of the crisis.

4) Considering Variation Values against Absolute Values

We add the total volume of deposits to the data. With all the data together we build the target variable that we are going to use in the predictive models.

5) Economic proof for variables to include in the model

We do not select the variables that have a higher correlation in the model. Among these variables we must choose those with more economic sense.

6) Type of model to develop

We have different possibilities to develop our model:

Cluster analysis, linear regression, ARIMA models.

The features of the model developed are:

- A single model valid for all banks
- No data partition
- Target variable: Deposit variations average
- Given data: All variables in variations in order to use same magnitudes
- Regression model

4. DATA ANALYSIS

We follow the following steps to analyze the available data:

- Correlations between explanatory variables and the target variable: We want to include in the model the most significant variables according to their correlation with the target variable.

- Economic relationship between the data and the target variable: Among the variables previously selected, we make a second choice of the variables that are most relevant in the economic framework. We also consider delays in the variables.

- Correlations between the relevant variables: With the variables that we have chosen, we have to try that the correlation of the variables in the model are as small as possible, because if they have a high correlation, they would be giving the same information in the model and any of them do not add anything.

- Exclusion of the less relevant data: Our goal is to reduce as much as possible the number of explanatory variables following the previous steps.

After filtering the explanatory variables, the set of candidate variables to build the model are:

- Unemployment Rate Relative Variations
- GDP Relative Variations
- House Price Relative Variations
- IPC Relative Variations
- Saving Rate Relative Variations
- Available Rent Relative Variations
- Legal Interest Rate Relative Variations
- Treasury Bonds Volume Relative Variations
- Investment Funds Volume Relative Variations

5. REGRESSION MODEL

To build the model we used the SAS software. With its help we tested several combinations until the final model that we present below.

We test various combinations with all the explanatory variables until getting the final model.

We started with nine potential explanatory variables. For the development of a linear regression model, we thought that nine variables are too many. Since the objective is also to obtain a model as simple as possible, we decided that our regression model had only three explanatory variables.

First we tested with combinations of the most significant variables.

We also tested with the most significant variables delays.

In each model, we looked if the sign of the coefficient of the explanatory variables had economic sense.

After several tests, the explanatory variables to build the model were:

- Available Rent Relative Variations
- Funds Investment Relative Variations
- Saving Rate Relative Variations with one lag

Now we show the model equation:

$$Y = 0.28303 \cdot x_1 - 0.28531 \cdot x_2 + 0.00118 \cdot x_3$$

where,

Y: deposits variation average

x_1: available rent relative variations

x_2: funds investment relative variations

x_3: saving rate relative variations with one lag

Comments

We see that the coefficient associated to the variable x_3 is very small. This is because the saving rate is measured in millions of euros.

The coefficients of x_1 and x_3 are positive. It is consistent that the value of available rent and saving rate make deposits grow.

In the case of the x_2 we have a negative coefficient. This is also economically feasible, because if people that have funds investment, they have less money to save.

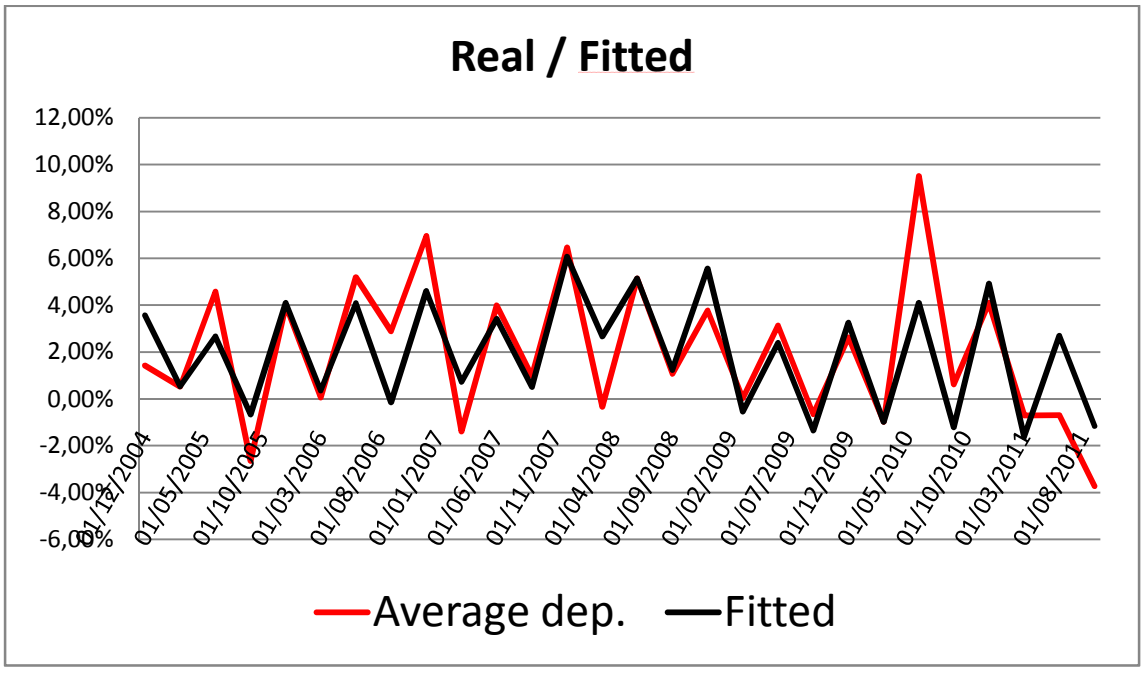
Results of the model

$R^2 = 0.7189$

Significance level at 95%

Although we could think that the value of R^2 is not too good, considering the simplicity of our model (we only use three of the nineteen explanatory variables) we can conclude that the results are acceptable.

Finally we will see the in the same graphic the model against real data to visually if the predictions are close to real data:



6. STRESS SCENARIOS

We work with the previous model.

Now, we have tried to see the changes in bank deposits with different stress scenarios.

To build a stress scenario, first we have to observe what are the extreme values of the variables we want to stress, and second manipulate the variable values to simulate a possible stress scenario.

We did several simulations and we compared the results with the percentage of provisions requested by "Basel III".

On several situations we observed large variations in deposits banks for severe stress scenarios.

For example,

We observe negative variations in deposits for the following cases:

- A big decrease of the available rent
- A big decrease of the saving rate
- A big increase of the investment funds

7. CONCLUSIONS

Although the problem is complex, the model for its study is very simple and also get a good approximation of the target variable.

The general model is not suitable for approximating a particular bank. That is, if we get more accurate results for a particular bank, we have to do a model for this bank. For example, if we want to study a particular bank with more detail.

We have seen variations of more than 5% in deposits in stress scenarios. In these cases, the percentage of provisions suggested by Regulators appears not to be enough.

Liquidity risk is not a systemic risk.

General review of the problem:

With the aim to really understand and model variations in deposits, we would need to consider variables representing the own behavior and structure of each entity.

Macroeconomic variables explain the situation of the country and also part of the situation of the entities but are not enough to deeply get into the dynamic system we have been trying to model.

We are talking here about liquidity risk so we would like to consider the need of liquidity an entity has. If you look into the “Liabilities and Assets structure”, you could have an idea about how big is the time gap among loans they need to return and loans others need to return to them.

Another interesting aspect to model how strong would be the struggle for retail deposits would be the difference between the interbank financing costs and the deposits profitability. A possible indicator of this could be the comparison of the entity’s credit spread and the profit they are giving in retail deposits.

Different approaches to the problem:

We have been able to observe that the total volume of cash in deposits has had a regular positive tendency during the period we are studying.

The question is then how entities share this “cake”. We can easily explain the total volume using macroeconomic variables; modeling the piece of cake each entity gets could be approached from different angles, as a dynamic system, statistics or graphs theory, either way using variables relative to each entity, like the ones we suggested, and possibly other ones.

8. IDEAS FOR NEW MODELS

- The total volume of cash in deposits has a regular positive tendencies so we could model this total, which can be understood as a cake banks are sharing.

- After modeling total volume using macroeconomic variables, we can see the rest of the problem as a dynamic system.

- Another approaches can be statistics or graphs theory, either way using variables relative to each entity.

- Need of liquidity ("Liabilities and Assets structure")

- Retail Deposit Profit vs Interbank Financing (Credit spread)