

Macroeconomic determinants of the Non-Performing Loans in Spain and Italy

Department of Economics
University of Leicester

Supervisor: Dr. Ludovic Renou

24 August 2012

Carlos Andres Olaya Bonilla¹

***Dissertation submitted to the University of Leicester in partial
fulfillment of the requirements for the award of the Masters of Science
(MSc) in Business Analysis and Finance degree***

¹ Beneficiario COLFUTURO 2011

Table of Contents

Introduction	1
Literature review.....	2
Macroeconomic review	2
Empirical Literature.....	5
Econometric Literature	8
Data and methodology	11
Explained variable.....	11
Explanatory variables.....	11
Econometric methodology.....	15
Empirical Results and Interpretation	20
Results for Spain	20
Results for Italy	25
Conclusions	30
References	32
Appendix 1 - Source of Macroeconomic Variables.....	35
Appendix 2 – Test of Assumptions of the Regression.....	36
Appendix 3 – Regression Outputs.....	43
Appendix 4 – Heteroskedasticity-consistent standard error estimators.....	49

Introduction

From an economic point of view, the bad loans have been considered as a key factor in banking crises. In recent years, the European debt crisis and the following recession, have increased loans defaults, specifically in Spain and Italy, causing significant losses for banks.

In this analysis the main macroeconomic determinants of Non-Performing Loans (NPL) will be studied over the period from January 2004 to March 2012. Separate analyses will be conducted for Spain and Italy considering that these countries have one of the largest amounts of bad loans in Europe, as well as deteriorated macroeconomic indicators during recent years.

The paper has been organised in the following way: in the chapter entitled 'literature review', the most important facts about the macroeconomic variables will be analysed focusing on NPL for Italy and Spain. I will move on to describe the main literature on determinants of Non-Performing Loans. Econometric literature will also be discussed, with special emphasis on time series procedures for macroeconomic variables. This section is important in order to determine the variables and to be aware of atypical data in the time series. Previous research in this field using the same or similar methodology additionally corroborates its efficiency and aptness for use in this paper.

Secondly, the variables used in this analysis will be explained bearing in mind the literature before mentioned. The econometric methodology and procedures used to obtain econometric results will then be presented in detail. Lastly, the results will be carefully analysed for Spain and Italy respectively, and conclusions regarding the determinants of Non-Performing Loans in these two countries will be drawn.

Literature review

Macroeconomic review

Since early 2010, when it became generally known that Greek debt had reached €300bn, which represented 113% of the Greek Gross Domestic Product (GDP) (nearly double the Eurozone limit of 60%²), the rating agencies started to downgrade Greek bank and government debt. In late 2010, concerns started to form about other overindebted economies: Portugal, Ireland, and Spain. Ever since, the European Union has provided measures such as capital raising, bailout programmes and, adjustment of taxes. In the case of Greek bonds, private banks even agreed to a 50% reduction of Greece's debt³.

Despite measures taken by the European Union, the instability spread towards the continent to the extent that, in early 2012, the credit rating agency Standard & Poor's downgraded France and eight other Eurozone countries. The failure of Eurozone leaders to appropriately deal with the debt crisis was given as a reason for this⁴.

Recent evidence suggests that the debt crisis of the European economies has a direct impact on indicators such as unemployment, GDP and inflation, amongst others⁵. Although this situation has been experienced by several European countries, this paper will only take the macroeconomic indicators of the Spanish and Italian economies into account.

-Facts for Spain

Compared to other European nations, the Spanish public debt as a percentage of the GDP (68% as of 2011) is not considerably high. However, Spain has one of the highest unemployment rates on the continent (23% as of 2011) which has grown constantly since 2005 and has affected the credit growth since 2009, as shown in table 1.

² The unfolding Euro crisis, June 2012, BBC News,

³ Eurocrisis, October 2011, The guardian.

⁴ Idem 1

⁵ Stability of Banking Systems and Interest-Rate Interventions (2008), Centre of Economic Research, p7

Table 1 – Macroeconomic variables for Spain (2005-2011)

	2005	2006	2007	2008	2009	2010	2011
GDP (€ billions)	909	986	1,053	1,088	1,048	1,051	1,073
Unemployment	8.7%	8.3%	8.8%	14.9%	19.2%	20.5%	23.2%
Inflation	3.4%	3.5%	2.8%	4.1%	-0.3%	1.8%	3.2%
Credit Growth (Year)	24.3%	22.8%	15.8%	6.1%	-1.9%	0.1%	-3.2%
Public Debt as % of GDP	43%	40%	36%	40%	54%	61%	68%

Sources: Bank of Spain, Euromonitor International from national statistics, International monetary Fund

Furthermore, the gross domestic product has shown a weak growth rate for the last three years. Recent news coverage further shows that the Spanish GDP shrank by 0.3% in the first three months of 2012 - the second consecutive decline - while the country's austerity measures and the wider economic slowdown affect growth⁶.

-Facts for Italy

The Italian level of debt has remained high since early 2000 to the extent that in 2011 their debt represented 122% of their GDP. Furthermore, there is a decrease in credits as a consequence of higher costs for loans.

Table 2 – Macroeconomic variables for Italy (2005-2011)

	2005	2006	2007	2008	2009	2010	2011
GDP (€ billions)	1,436	1,493	1,554	1,575	1,527	1,556	1,578
Unemployment	7.5%	6.2%	6.6%	7.1%	8.4%	8.2%	9.0%
Inflation	2.0%	2.1%	1.8%	3.3%	0.8%	1.5%	2.8%
Credit Growth	7%	11%	19%	8%	2%	-2%	2%
Public Debt as % of GDP	105%	106%	103%	106%	116%	118%	122%

Sources: Bank of Italy, Euromonitor International from national statistics, International monetary Fund

The National Bureau of Economic Research partly defines a recession as two consecutive quarters of GDP decline. Considering that Italy's GDP fell by 0.7% in the last quarter of 2011, following a 0.2% decline in the third quarter, Italy has technically entered into a recession. Moreover, its unemployment rate has grown by 10% from 2010 to 2011 while inflation remained moderate.

⁶ Eurozone crisis, April 2012, The guardian

-Non-performing Loans Evolution

A non-performing loan is considered a loan that is in default or close to being in default. Along with Germany and the United Kingdom, Spain and Italy are the nations with the largest amount of Non-performing loans in Europe (See table 3).

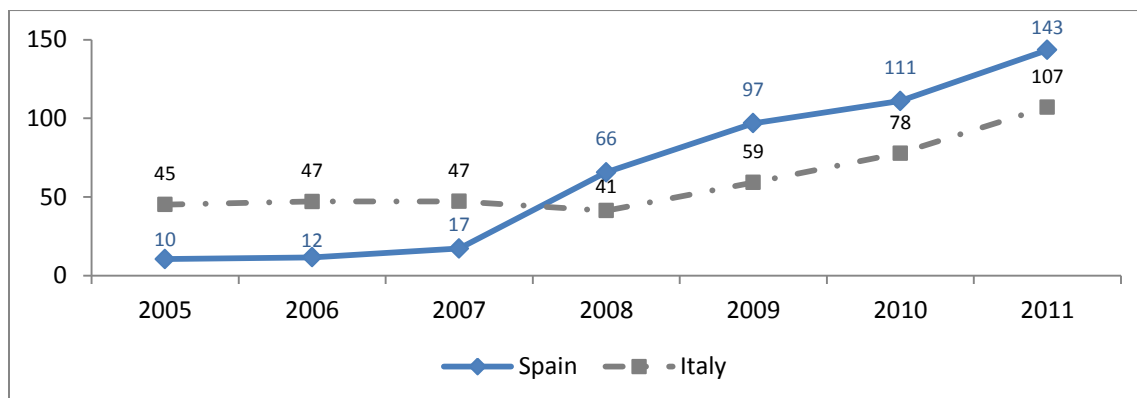
Table 3 - NPL Size in Europe

Figures in € billions

	2011	Source
Germany	225	Pwc London
United Kingdom	175	Pwc London
Spain	143	Bank of Spain
Italy	107	Bank of Italy
Greece	24	Bank of Greece
Estimated for Europe	850	Pwc London

As expected due to the deteriorated macroeconomic indicators illustrated before, there is a bullish tendency in the NPL market in the case of Spain and Italy (See graph 1). Nonetheless, additional factors have also contributed to the uptick in bad loans of both economies. In Spain, “regulatory changes implemented by the Bank of Spain have resulted in banks increasing their provisioning for bad debts significantly”⁷, while in Italy, the lengthy foreclosure procedures have constrained investors’ appetite for NPL portfolios sales that could decrease the amount of this variable.

Graph 1 - NPL tendency in Spain and Italy (Figures in € billions)



Source: Bank of Spain, Bank of Italy.

⁷ Bergaz, Jaime, (2011), European outlook for non-core and NPL portfolios, PWC, Issue 3, pp. 12, 16.

Empirical Literature

The uncertainty in macroeconomic indicators raises concerns about how the rate of the Non-performing loans in Spain and Italy will evolve. In this respect, the relationship between macroeconomic conditions and the banking system has been widely investigated. However, “such research has concentrated mostly on how to evaluate the risk of borrowers before the event and the risk of individual operations”⁸.

Fernandez, Martinez and Saurina (2000) analysed the cyclical behaviour of bank credit, loan losses and provisions for loan losses in Spain. Their research is to be considered in this paper given their findings in lags condition for NPL. According to their study, which was conducted with panel data, the low quality of loans will only become apparent with the ex post emergence of default problems, this will commonly appear during downturns with an estimated lag of approximately three years in the case of Spain.

Two years later, another study for Spain was performed by Salas and Saurina (2002) showing that the real GDP growth, credit growth and bank size explained the credit risk. In their study, the loan loss of Spanish banks with respect to macroeconomic factors and microeconomic variables from 1985 to 1997 is examined. Using Panel data, Salas and Saurina found that credit risk is significantly determined by microeconomic variables such as families' indebtedness and loan portfolio composition. Their research is important as a reference for this model when defining the explanatory variables.

Similarly, in Italy, Marcucci J. and Quagliariello M (2005) analysed the cyclical behaviour of the default rates for Italian bank borrowers over the period from 1990 to 2004. Vector autoregression (VAR) modelling was employed to assess the effects of business cycle conditions on bank customers' default rates. Their results confirm that the default rates follow a cyclical pattern, decreasing in good macroeconomic times and increasing during downturns. Furthermore, their findings suggest that, when capital

⁸ Salas, V., and J. Saurina, (2002), Credit Risk in Two Institutional Regimes: Spanish Commercial and Savings Banks, *Journal of Financial Services Research*, Vol. 22, pp. 203.

surpluses over the regulatory minimum are low, banks may reduce lending, which, in turn, negatively affects the output levels. Finally the authors do not find strong evidence of feedback effects from the soundness of banks' balance sheets to economic activity. For my research, it is important to bear in mind the robustness of the VAR for dealing with macroeconomic time series in Italy.

Also in Italy but more recently, Bofondi and Ropele (2011) analysed the quality of loans to households and firms separately, on the grounds that macroeconomic variables may affect these two classes of borrowers differently. The Italian research covers quarterly data from the 1st quarter of 1990 to the first quarter of 2010 and concludes that i) the quality of lending to households and firms can be explained by a small number of macroeconomic variables (GDP annual growth, unemployment and short term interest rate); ii) changes in macroeconomic conditions generally affect loan quality with a lag (in variables such as interest rates the lag can be up to 4 quarters) and iii) the out-of-sample prediction accuracy of the models is quite satisfactory and proved to be robust to the recent financial crisis.

Two recent studies using a single-equation time series method show the efficiency of the methodological approach also chosen for my research. On the one hand, Arpa M. and Giulini I. (2001) assessed the effects of macroeconomic developments on risk provisions and bank earnings over the period from 1990 to 1999 in Austria.

They work with a single-equation time series model using the banks' risk provisions as the dependent variable and explanatory variables such as the growth rate of the gross domestic product, real estate price developments and real interest rates. Their main findings are that Austrian banks increase risk provisions in times of falling real GDP growth rates and in times of rising bank operating income or operating results. Net interest income appears to be uncorrelated with real GDP growth and interest rate developments. The only exception is that net interest income shrinks at very low interest rate levels and banks increase their risk provisions in times of declining real GDP growth rates.

On the other hand, with a similar procedure, Simons and Rolwes (2008) found a convincing negative relation between the default rate and GDP growth in the Netherlands. In their analysis, they also included the oil price, the interest rates and exchange rate as explanatory variables for the Dutch loans default, finding a significant relation in several sectors. Remarkably, for the analysed variables, not the change but the level of the variables turned out to be significant. However, the macroeconomic relations with the sector default rates are mostly unstable except for the oil price. According to this study, a reason for the instability is that results amongst sectors can differ according to the growth opportunities of the sector of economic activity to which firms belong. This cannot pose problems to my research as the default rate is expressed by the ratio of the total bad loans in the economy.

In a different approach on default rates, Segoviano M. (2006) modelled the probabilities of loan defaults from a list of companies as functions of identifiable macroeconomic and financial variables. The study which was conducted for the Norwegian and Mexican economy shows that increases of credit to GDP and asset prices have a significant explanatory power on the probability of defaults in both countries. It is also observed that when periods of combined strong increases in credit and real asset prices occur, there is an enhanced likelihood of stress occurring in the financial system (reflected by increased default probabilities), two to four years ahead. It is, therefore, important to bear in mind that the lags in the macroeconomic variables could skew the interpretation of the results the fact that in my analysis. A set of lags will be analysed in this research in order to improve the model.

Research outside of Europe has also reaffirmed the relationship between macroeconomic variables and the default loans. Recently, Saba I. and Kouser R. (2012) attempted to ascertain the determinants of NPL in the United States banking sector by using Pearson's correlation analysis and Ordinary Least Squares (OLS). Their result supports the view that macroeconomic factors, such as, Interest rate and Real GDP per capita stand in relation with the NPL rate. The authors consider that the outcome in such different variables depends on the situational factors which include country level

factors, bank level factors and the characteristics of the legal and regulatory framework. This analysis is an important reference because it relies on the NPL rate as the dependant variable rather than the default rate. By using a similar treatment for the variable NPL, my paper uses the NPL ratio for the Spanish and Italian economies.

In addition, my research aims to analyse the effects of two economies simultaneously as well as return to this topic under a more updated and wider time series. Moreover, I will consider more macroeconomic variables than the ones previously used in Spanish and Italian research based on the experience of external investigations. For instance, in order to include *Inflation* as a determinant, I consider the influence of this variable in NPL of the Czech Republic explained by Babouček and Jančar (2005). Accordingly, I will include the *unit labour costs* grounded on the benchmark models for the euro area of Peersman and Smets (2001) which consider the labour market variables as influential for the quality of the loans.

Econometric Literature

Chan-Lau J (2006) conducted one of the main researches regarding fundamental-based models for estimating default probabilities (which is similar to the concept of the NPL rate). The author distinguishes between four approaches within fundamental-based modelling for default probabilities: macroeconomic-based, accounting-based, rating-based and hybrid models. Under the macroeconomic model, it needs to be considered that the variables are typically cyclical indicators. It is therefore suggested to use more than one business cycle of data. Also, aggregated economic data usually reports lags, making it difficult to estimate models with updated information. However, some advantages in this type of model are that it is very suitable to design stress scenarios and it is able to conduct cross-country comparative studies.

The author also classifies the macroeconomic models into exogenous and endogenous models based on whether the model allows feedback between financial distress and the explanatory economic variables.

The first category assumes that the economic variables are exogenous and not affected by financial distress. The general approach to modelling this category is described by the following equation:

$$pd_t = g(x_1, x_2, \dots, x_n) + \varepsilon \quad (1)$$

Where pd is the probability of default over a given period t . The function $g(x_1, x_2, \dots, x_n)$ is a set of macroeconomic variables summarizing the state of the economy and the term ε is a random variable.

The second category of macroeconomic-based models assumes that the economic variables are endogenous and differ in times of financial distress. This category can be described with the following equation:

$$Z_{t+1} = \alpha_t + \sum_{j=1}^p \beta_j Z_{t+1-j} + \varepsilon_{t+1} \quad (2)$$

Where α_t is a constant vector, β_j are lagged coefficients matrices, ε_{t+1} is a vector of residuals, and Z is the vector of endogenous variables, which includes both default probabilities and aggregate economic variables associated with the state of the business cycle.

This analysis fits into the second category, provided that the state of the economy impacts the behaviour of the debtors and thus the bad loan rate. A typical econometric framework used in these models is the vector autoregressive (VAR) methodology. Generally, the interpretation in VAR models is susceptible to the selection of lags. Accordingly, if a short number of lags are included, important lag dependencies may be omitted while if lags occur in greater number, degrees of freedom are lost. Nonetheless, Kenneth F (1995) highlights that, the VAR models captures the linear interdependencies among multiple time series and is therefore suitable for macroeconomic analysis. Although this is true, some authors suggest that a large

number of parameters may be necessary for an adequate description of the data (See Lütkepohl 2004). Given limited sample data in some variables for Italy, the VAR model could result in low estimation precision for this research.

The linear time series model is another econometric framework for non-performing loans modelling. In this respect, after considering a large variety of models for macroeconomic variables for Italy, Marcellino M. (2007) concludes that “general linear time series models can hardly be beaten if they are carefully specified, and therefore still provide a good benchmark for theoretical models of growth and inflation”. This finding supports the use of linear regressions in my research.

Similarly, Segoviano, M. (2006), mentions that, in order to model the impact of macroeconomic development on probabilities of default, risk managers and regulators have commonly used ordinary least squares (OLS) estimation procedures. However, when attempting to do so, they usually face a challenging problem, since the number of observations on the time series of probability defaults frequently just exceeds the number of parameters to be estimated. Accordingly, this will be an issue for this research that will be addressed when running the regressions (see the econometric methodology).

Data and methodology

Explained variable

The dependent variable was initially set as the growth of the total bad loans and then as the change of the NPL ratio in both economies. However, the regression output did not satisfy the normality assumption and the strength of the model was not as high as expected. Consequently, in order to measure the Non-performing loans (NPL) in Spain and Italy, I use the monthly ratio of the bad loans with respect to the total loans. The data is based on reports of the Bank of Spain and the Bank of Italy covering the period from January 2004 to March 2012. This period is sufficient to analyse the time before the crisis, considering that the recent recession significantly worsened the quality of credit. For my purposes, the measure of this variable includes all the types of default credits for each economy.

Explanatory variables

The variables used to describe the non-performing loans include the same period as used for the explained variable (January 2004 to March 2012). Several sources were used to complete the data (See Appendix 1).

As indicators for the general state of the economy, the *Gross Domestic Product* (GDP), the credit growth (Credit) and the *Unemployment rate* (Unemplo) are used.

The GDP is expected to have a negative relationship with the bad loans, as an increase in the GDP would represent a source of liquidity in the market that would allow the borrowers to face their debts. As a consequence, the NPL ratio would at least remain stable under GDP growth. Since the GDP information is published quarterly, I will do a linear interpolation in order to estimate the GDP's annual change on a monthly basis. This methodology is based on Segoviano M., Goodhart C. and Hofmann B. (2006).

The credit growth is expressed as the monthly percentage of growth for the total loans borrowed in each economy. Conversely, this variable is expected to have a positive relation with the NPL ratio as the ratio depends mathematically on the total loans in the economy. Also, “a rapid credit expansion is considered one of the most important causes of problem loans”⁹ provided that under an expansionary credit policy a bank is willing to reduce the quality of their clients. For my model, this variable is computed as the monthly rate of change of the total loans’ stock. The data is based on the asset reports for credit entities in Spain and Italy.

Regarding the unemployment, it is rational to suppose that an increase in this variable would be associated with the number of borrowers who became unemployed during the same period, and so it was initially considered to analyse the monthly *change of the unemployment rate* (Unemplo). However, for this variable, not the change of the unemployment rate but the unemployment rate itself turned out to be more significant in Italy. Therefore, this variable is expressed as the monthly unemployment rate for Italy and as the monthly *change of the unemployment rate* for Spain.

Several studies investigating the relation between unemployment and NPL have found it to be positive (See Bofondi, M. and Ropele T, (2011) and Głogowski A. (2008)). Consequently, it is expected that an uptick in unemployment would affect the quantity of borrowers not paying off their debts. This effect is likely to be stronger in the Spanish economy given their rise in this variable.

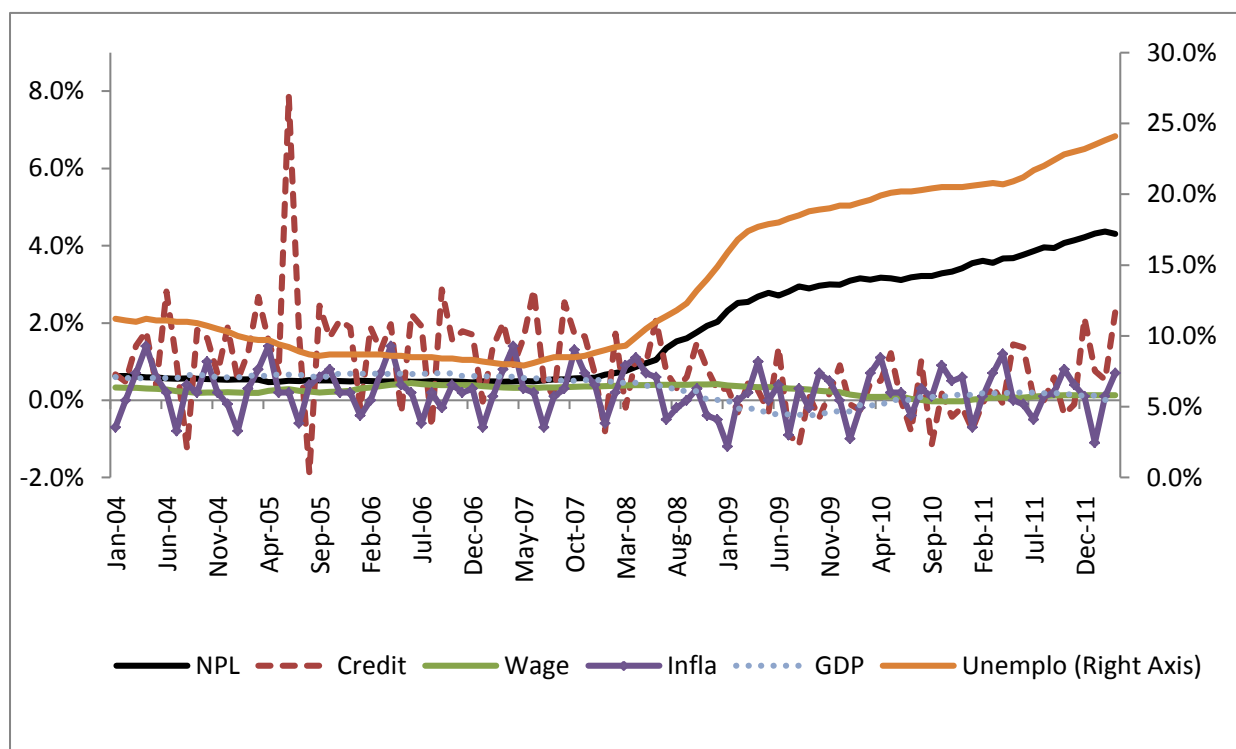
In order to reflect the price stability in the model, I include the inflation (Infla) as the general consumer prices rate with a monthly basis. For this variable, it is expected that, as inflation increases, the cost of borrowing gets more expensive, which could deteriorate the quality of the loan portfolios. However, other studies have found a negative relation between the inflation rate and non-performing loans.

⁹ Solttila and Vihriala (1994), after controlling for the composition of the bank loan portfolio, find evidence that past credit growth explains the current level of problem loans.

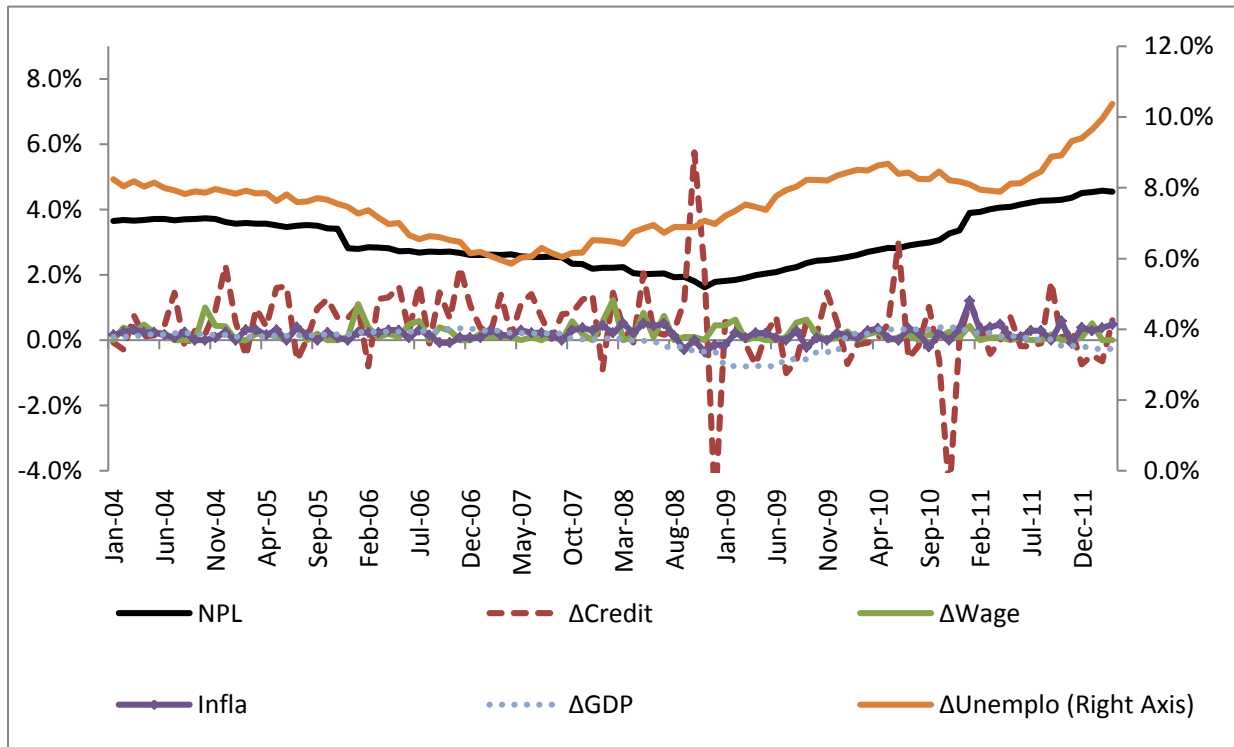
Finally, I include the monthly cost of labour (Wage) as a measure of the borrower income. This variable is the monthly estimation of the average change in the cost of an employee. I believe that an increase in the monthly wage reflects a flow of income for the borrowers while a decrease in this variable restricts the future purchasing power of the borrowers and potentially prevents them from supporting their debts.

Overall, five macroeconomic variables are used: Credit, Wage, Inflation, Unemployment and GDP. The data has 99 monthly periods that cover the economy state, the price stability and the borrower income. Graph 2 and 3 show the evolution of the main macroeconomic variables and the non-performing loans for Spain and Italy respectively.

Graph 2 – Evolution of NPL's and Macroeconomic Variables in Spain



Graph 3 – Evolution of NPL's and Macroeconomic Variables in Italy



Econometric methodology

-Econometric Model

In order to explain the determinants of the NPL in Spain and Italy, I use the ordinary least squares model (OLS). Under this approach it needs to be considered that the OLS's main assumption is that the errors must be uncorrelated. Considering that the model involves time series, this assumption could be violated since it is reasonable to think that the cyclical effect in the economy indicates positive autocorrelation. However, previous studies have shown that OLS is a suitable model to describe NPL time series (For instance, Espinoza R. and Prasad A. (2010) and Bofondi, M. and Ropele T, (2011)) as long as the existence of autocorrelation in residuals is investigated.

In a linear regression model the response variable is a linear function of the regressors:

$$y_t = \alpha + x_t' \beta + \varepsilon_i \quad (3)$$

Where y is the explained variable, α is a constant term, X is a $n \times p$ matrix of regressors and ε are unobserved errors. Drawing on the empirical literature (See Brooks, C. (2008)) we can use this equation to set the base model as:

$$NPL_t = \alpha + \beta_1 \Delta Credit + \beta_2 \Delta Wage + \beta_3 Infla + \beta_4 Unemplo + \beta_5 GDP + \varepsilon_i \quad (4)$$

Equation (4) will be estimated with a multivariate ordinary linear regression. The econometric software used to process the regression is SPSS version 17. Although several regressions are included in the results chapter (See tables 4, 5, 6 and 7), for procedure purposes, only the process done in the conclusive regression which turned out to be '6 months lag' for the Spanish and '12 months lag' for the Italian data is described.

-Econometric Procedure

Initially, I verified that the sample size is sufficiently large for a multiple linear regression. In this respect, Tabachnick, B.G. and Fidell, L.S. (2007) recommend that the required number of cases should be larger than eight times the number of independent variables plus 50, i.e. a sample size of 90 observations for this analysis. Although the sample size for linear regressions depends on several aspects, I ensured to have roughly 100 observations by checking the data with descriptive statistics from the SPSS. This tool allowed me to identify the quality of the data (See appendix 2, tables 8 and 9).

A large presence of outliers could slightly skew the errors distribution. Thus, the outliers were also removed before running the regressions. The boxplot utility of SPSS suggested that there was one potential outlier in the Spanish (observation 18) and three in the Italian data (observations 58, 60 and 83). (See appendix 2, graphs 6 and 7)

Accordingly, a standard linear regression was run requiring, in addition, an analysis of collinearity diagnostics. Likewise, the unstandardized predicted values and the unstandardized residuals were saved in order to assess the aptness of the model. Consequently, before interpreting the regression output, it is advisable to test the assumptions of the regression analysis.

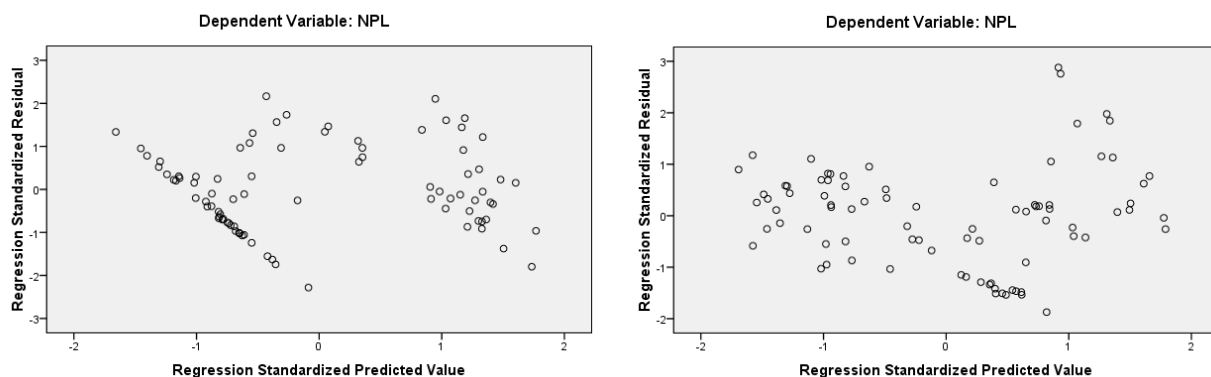
Regarding the *linearity*, I drew a scattered plot between the explanatory variables and the NPL, so that it was possible to check if there is a linear function between the variables. In all the cases, the points were symmetrically distributed around a diagonal or horizontal line. In spite of the fact that no clear linearity was reflected for the chart of the variable 'credit' for Italy, once the outliers were removed, the linear relationship improved (See appendix 2, graphs 8 and 9).

Another concern was the *Independence of the errors* bearing in mind that as multicollinearity increases, the regression model coefficient could be inflated. In order to analyse this issue, I used the collinearity diagnostics of SPSS. In both, Spanish and Italian data, the tolerance indicator was greater than 0.5 for most of the independent

variables. This implies that the variance of the independent variables is unique and not explained among each other. Similarly, the variance inflation factors (VIF) were within acceptable levels, i.e. smaller than 10 (See Appendix 2, table 13). Therefore, no presence of multicollinearity was found in this analysis.

On the other hand, in order to check for *homoscedasticity*, a graph of the unstandardized residuals with the unstandardized predicted values was plotted to determine whether there is a linear relationship (See graph 4).

Graph 4 - Unstandardized residual vs NPL*



*Notes: Spain left hand side, Italy Right hand side

Although the graph is not a statistical approach to examine homoscedasticity, it is shown that the residuals might have a relationship and get wider forming a funnel shape. This suggests a minor existence of heteroscedasticity. Accordingly, it is important to formally test this issue. However, SPSS does not include a test to statistically measure the significance of heteroscedasticity. In order to do so, a macro was included in the syntax menu of SPSS based on the methodology described by Gwilym Pryce (2002). In his paper, Pryce develops a macro to compute both the Breusch-Pagan and Koenker tests which check for heteroscedasticity. More attention was paid in the Koenker test considering that it is rigorous for small size time series as the ones used in this study (See Appendix 2, table 10).

For the case of Spain, the Koenker test yields a CHI Square value of 12.5 and a significance level of 0.0276. Similarly, for Italy, the Koenker test yields a CHI Square value of 12.4 and a significance level of 0.0286. The null hypothesis of homoscedasticity is rejected based on the significance levels which are less than 0.05. Therefore the presence of heteroscedasticity is concluded for both sample data. The above suggests that the residuals get bigger across time, impacting the predictive capacity of the model over time. This could be explained by the volatility caused from the 2007 debt crisis in the Spanish and Italian data.

Notwithstanding, Berry and Feldman (1985) and Tabachnick and Fidell (1996) say that slight heteroscedasticity has little effect on significance tests; however, when heteroscedasticity is marked it can lead to serious distortion of findings and seriously weaken the analysis thus increasing the possibility of a Type I error. Although the Koenker test suggests only a slight heteroscedasticity for the data used in this paper, this issue will be preventively addressed in order to obtain more robust estimates.

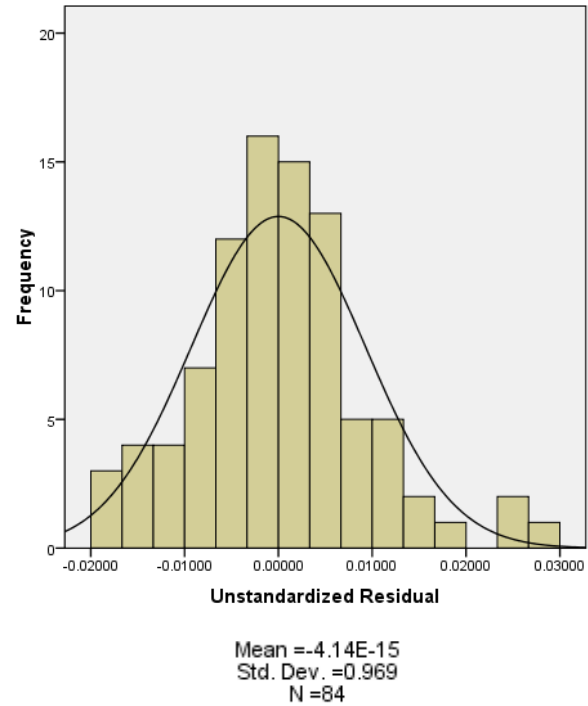
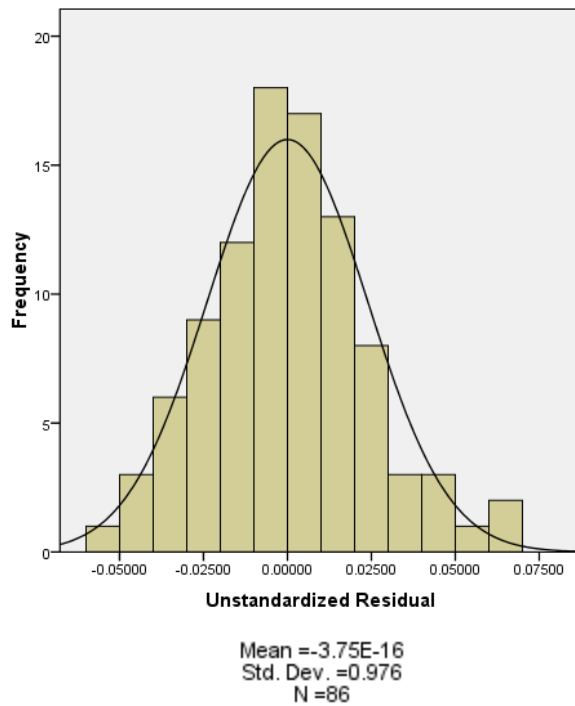
There are methods to solve heteroscedasticity issues, which include transforming the data, use of weighted least squares (WLS) regression and generalized least squares (GLS) estimation. In this paper however, it is adequate to use heteroscedasticity-consistent standard error (HCSE) estimators for OLS parameter estimates given the small sample size. (See Appendix 4)

Finally, the *normality* was tested by drawing a histogram of the unstandardized residuals as shown in graph 5. For this test, it was considered that after lagging the time series, some new outliers could appear. Nonetheless the casewise diagnostics did not report new atypical values and so, the normality test reflects appropriate observations.

The histograms suggest that the errors may be normally distributed. However, in order to formally test the distribution, I used the Kolmogorov-Smirnov test (See Appendix 2, tables 11 and 12). In both cases, the significance level was greater than 0.05

suggesting a normal distribution in the errors. Accordingly, the test shows the levels of Skewness and Kurtosis, which are close to zero for Spain and Italy.

Graph 5 - Unstandardized residual Histogram*



*Notes: Spain left hand side, Italy Right hand side

Empirical Results and Interpretation

Results for Spain

In the first place, the explanatory power of all the five variables was tested without considering any lag. Table 4 reports the results (Specification 1). In this model, the explanatory variables describe 85.0% of the variance of the Non-performing loans index in Spain. The Analysis of variance (ANOVA) tests the null hypothesis stating that there is not an overall relationship between the dependent variable and the independent variables. This hypothesis is rejected with a 95% level of confidence as the p-value is reported to be smaller than 5%.

Nonetheless, the association between the NPL index and the inflation is not statistically significant. The p value for this variable is greater than the 5% confidence level and I, therefore, exclude it in order to analyse how the model changes (see table 4, specification 2).

Table 4 – Estimation results Spain

Regressor	Specification 1		Specification 2	
	Coefficient	P-value	Coefficient	P-value
(Constant)	3.68	0.00	3.68	0.00
Credit	-0.18	0.04	-0.18	0.03
Wage	-5.45	0.00	-5.44	0.00
Infla	-0.02	0.84		
Unemplo	1.08	0.00	1.08	0.00
GDP	-1.76	0.00	-1.76	0.00
R Square	0.850		0.870	
Adjusted R Square	0.873		0.875	
F	134.7		170.1	
P-value	0.000		0.000	
N. Obs.	98		98	

After excluding the variable 'inflation' from the Spanish data, the r-square and the F value increased. Furthermore, the distribution of the residuals seemed to be more normal. Based on the above and the low correlation of the inflation with the NPL (only - 0.10), this variable will be excluded for further analyses.

In the next place, a set of lagged regressions was analysed in order to adjust the model. In order to do so, the independent variables were transformed using the menu "create variable" of SPSS. Table 5 reports the results:

Table 5 – Estimation results – Lag specifications for Spain

Regressor	Lag 3 months		Lag 6 months		Lag 9 months		Lag 12 months	
	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value	Coefficient	p-value
(Constant)	3.64	0.00	3.61	0.00	3.63	0.00	3.60	0.00
Credit	-0.19	0.02	-0.25	0.00	-0.26	0.01	-0.28	0.07
Wage	-5.20	0.00	-4.76	0.00	-4.46	0.00	-3.95	0.00
Unemplo	1.46	0.00	1.63	0.00	1.73	0.00	1.79	0.00
GDP	-1.73	0.00	-1.64	0.00	-1.66	0.00	-1.66	0.00
R Square	0.888		0.892		0.870		0.851	
Adjusted R Square	0.884		0.887		0.864		0.844	
F	179.3		179.7		140.6		116.1	
Sig	0.000		0.000		0.000		0.000	
N. Obs.	95		92		89		86	

The above table evidences that the model improves by adding lags. However, the inclusion of delays stops adding importance in the F value from the 'lag 9'. Similarly, the model starts losing explanatory power when more than 6 months of lag are added. Thus, the analysed macroeconomic variables have a higher impact in the nonperforming loans after six months for the case of Spain. The model 'Lag 6 months' explains the behaviour of the NPL index with more robustness and will be analysed in detail. Appendix 3 reports the complete output for this regression.

The conclusive regression: Lag 6 months for Spain

The correlation matrix corroborates my expectation of a negative correlation for the NPL with the *GDP* and the *WAGE*, which respectively report as -0.83 and -0.57 relationships. Consequently, the unemployment follows a positive correlation (0.5) with the NPL. Surprisingly, the *credit growth* has a strong negative relationship of -0.79 with the NPL.

The four independent variables included in the model explain 89.2% of the variance of the Non-performing loans index (See appendix 3, table 15, R-square). This is supported by the ANOVA analysis which reports an overall p-value of approximately zero, i.e., the results are statistically significant (See appendix 3, table 16, significance level). Likewise, the adjusted R-square which is more rigorous by taking into consideration the number of observation and the number of predicted variables is considerable high as well (88.7%).

On the other hand, the unstandardized betas and their statistical significances are reported in table 17 of Appendix 3:

The constant value represents the intercept of the model. In the Spanish example, when all the four independent variables are zero, the Non-performing loans index would start on average at 3.6%.

For the variable credit, the model indicates that if the credit grows by 1%, the Non-performing loans index decreases by 0.2%. The statistical significance for this variable is 0.003, i.e. it is substantially correlated with the Non-performing loans. The partial correlation table shows that the association between the credit growth and the NPL index is -0.3. This indicator together with the low Beta shows only a weak dependence of the NPL on the Credit.

The variable wage has a negative relationship as well. For every 1% increase in the monthly cost of labour, the NPL index would go down by 4.7%. Suspiciously, this is a large Beta, however, considering that the increases in the labour cost is tiny on a monthly basis; this relationship could make sense in a model based on monthly data. Consequently, the p-value for this variable is approximately zero, which makes it reliable.

The unemployment on the other has a positive partial correlation of 0.58. Based on the unstandardized Beta, for every 100 percentage points increased of the unemployment rate, the NPL index would increase by 163 percentage points. The p-value of 0.000 for this variable makes this relationship statistically significant. Note that for Spain the unemployment is expressed as the monthly change in the unemployment rate.

Lastly, the GDP's partial correlation of -0.61 shows a strong negative relationship with the NPL index. According to the unstandardized Beta, if the GDP grows by 1% in one month, the NPL index would decrease by 1.6%. This finding is statistically reliable based on the 0.000 p-value.

In a similar way, the standardised betas reported in table 17 represent the unique contribution of the variables as predictors of the NPL. These values are, however, strongly related to the unstandardized betas explained above.

Provided that the main assumptions in this model are partially fulfilled, the model 'lag 6 months' fits with statistically significant. The non-performing loans index in Spain is mostly explained by the wage, the unemployment and the GDP. The credit growth however does not seem to be a strong explanatory variable based on the partial correlation. Likewise, as explained earlier, the addition of the inflation as an explanatory variable, did not add power to the model.

The findings of this model are consistent with those explained by Salas and Saurina (2002) for the Spanish economy with data from 2001 and backwards. However, some

differences are evident: i) the low quality of the loans is affected faster in the current model, which includes updated data. Salas and Saurina found a significant impact of lags after 18 months, while the current model shows a strong evidence of the deterioration of the loans after 6 months. This difference could be an effect of including data that covers the ex and post debt crisis in Europe. ii) The credit growth lost explanatory power in the current model. This could be partially explained by the changes in the credit policy after the concerns spread in the overindebted European economies.

Overall, the model seems to be strong. However, the earlier mentioned presence of heteroscedasticity could slightly affect its robustness. Even though SPSS does not allow addressing this issue directly, another macro will be included in the Syntax menu of SPSS in order to estimate the heteroscedasticity-consistent standard error. This is based on the procedure done by Andrew F. Hayes (2007).

The macro develops a procedure to fit the four explanatory variables taking into account that the standard errors associated with the multiple regressions are not large enough. The results are presented in Appendix 4. The standard errors are estimated with a robust technic and therefore it computes more rigorous p-values. As expected, the moderate presence of heteroscedasticity does not affect the explanatory power of the model since the statistical significances of the variables are not importantly modified (See Appendix 4, table 26, P-value).

Results for Italy

Table 6 (Specification 1) depicts the first model, which includes the five explanatory variables without considering lags. In this model, the explanatory variables explain 62.9% of the variance of the Italian Non-performing loans index. The Analysis of variance (ANOVA) confirms the relationship between the dependent variable and the independent variables with a 95% level of confidence as the p-value is reported to be smaller than 0.05.

Nonetheless, the association between the NPL and both, the Credit Growth and the Wage is not statistically significant. The p value for these two variable is greater than the 5% confidence level, and, therefore, those variables are excluded in order to analyse how the model improves (see table 6, specification 2).

Table 6 – Estimation results Italy

Regressor	Specification 1		Specification 2	
	Coefficient	P-value	Coefficient	P-value
(Constant)	1.46	0.00	1.54	0.00
Credit	0.01	0.90		
Wage	-0.19	0.34		
Infla	0.49	0.05	0.44	0.07
Unemplo	0.58	0.00	0.58	0.00
GDP	1.05	0.00	1.06	0.00
R Square	0.629		0.625	
Adjusted R Square	0.608		0.612	
F	30.5		51.1	
Sig	0.000		0.000	
N. Obs.	96		96	

By excluding the variables 'Credit' and 'Wage' from the Italian data, the r-square is slightly reduced. There is a soft improvement in the adjusted r-square and the F statistical though. Nonetheless, the change in the model is not decisive and therefore, the five independent variables will be kept unchanged for further analysis.

A supplementary analysis includes regressions over a set of lags in order to find a model with a strongest specification. Table 7 reports the results.

Table 7 – Estimation results – Lag specifications for Italy

Regressor	Lag 3 months		Lag 6 months		Lag 9 months		Lag 12 months	
	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value	Coefficient	P-value
(Constant)	1.97	0.00	2.48	0.00	2.97	0.00	3.33	0.00
Credit	0.00	0.79	0.00	0.66	-0.02	0.48	-0.03	0.29
Wage	-0.30	0.11	-0.28	0.12	-0.23	0.16	-0.18	0.02
Infla	0.37	0.13	0.27	0.25	0.21	0.34	0.28	0.16
Unempho	0.65	0.00	0.73	0.00	0.79	0.00	0.84	0.00
GDP	1.01	0.00	0.94	0.00	0.82	0.00	0.56	0.00
R Square	0.671		0.711		0.766		0.809	
Adjusted R Square	0.652		0.694		0.752		0.797	
F	35.4		41.4		53.0		66.0	
Sig	0.000		0.000		0.000		0.000	
N. Obs.	93		90		87		84	

The above table gives evidence for the fact that the model improves by adding lags. However, the model starts losing explanatory power when adding more than 12 months of lags. Similarly, the inclusion of delays stops adding importance to the F-value after 1 year. Thus, the analysed macroeconomic variables have a higher impact on the nonperforming loans after twelve months in the case of Italy. The specification ‘Lag 12 months’ explains the performance of the NPL index with more strength.

It is surprising however, how the variable ‘Credit’ shifts the sign from positive to negative when adding lags. This finding could bear on the volatility of this variable (provided that the value of the loans has several upticks in the Bank of Italy data) or on an adjustment of the model by considering lags. The p-value for this variable improved when more lags were added. However, it is still not small enough and does not suggest statistically significance for the variable credit. The complete output of this regression will be analysed in order to exclude this variable from the model (See appendix 3)

The conclusive regression: Lag 12 months for Italy

The correlation matrix indicates a weak negative correlation with the NPL for the *Credit* and the *wage*, which respectively reports -0.23 and -0.12 relationships. Consequently, the unemployment follows a strong positive correlation (0.8). Contrary to the Spanish relationship, the GDP has a negative positive correlation of 0.18 with the NPL (See appendix 3, table 20).

Based on the R-square, the five independent variables included in the model, explains 80.9% of the variance of the Non-performing loans index (See appendix 3, table 21). This is supported by the ANOVA analysis which reports an overall p-value of approximately zero, i.e., the results are statistically significant (See appendix 3, table 22).

Similarly, the adjusted R-square which is more rigorous due to taking into consideration the number of observation and the number of predicted variables is considerably high as well (79.7%).

The unstandardized betas and its statistical significances are reported in the table 23 of Appendix 3:

The intercept for the Italian model, is very similar to the Spanish one. Based on the *constant* value, when all the five independent variables are zero, the Non-performing loans index would start on average at 3.3%.

For the variable 'credit', the model indicates that if the credit grows by 1%, the Non-performing loans index decreases by 0.03%. However, the p-value of 0.291 for this variable is greater than the 0.05 confidence level, which means that the credit is not significantly correlated with the Non-performing loans in Italy. This is supported by the low partial correlation index of -0.07. Conclusively, this variable is not considered as an explanatory variable and will be excluded from the Italian model.

This exclusion could partially be caused by the upticks registered in the total loans data from the bank of Italy¹⁰. However, it was impossible to locate another source to substitute this variable.

On the other hand, the variable wage is statistically significant based on its 0.02 p-value. In the same way as in Spain, this variable has a negative relationship. The beta indicates that for every 1% increase in the Italian Index of wage contract, the NPL index would go down by 4.7%.

The variable 'inflation' has a moderate positive relationship (beta of 0.28). However, the low partial correlation of 0.15 infers that this variable does not add power to the model. This is corroborated with the 0.16 p-value which suggest that this variable is not statistically significant in the model. Note that this variable was also excluded from the Spanish regression.

The unemployment, on the other hand, has a robust positive partial correlation of 0.88. Based on the unstandardized Beta, for every 1% increase in the unemployment rate in a given month, the NPL index would increase by 0.84%. The p-value of 0.000 for this variable makes this relationship statistically significant. This was also one of the strongest explanatory variables in the Spanish conclusive regression.

Remarkably, the GDP turned out to have a positive correlation different from what was expected. This relationship goes against the findings in the Spanish data. Nonetheless, the partial correlation of 0.45 and the p-value of approximately zero, show this explanatory variable as meaningful reliable. According to the unstandardized Beta, if the GDP grows by 1% in one month, the NPL index would increase by 0.56%. Logically, there is no explanation for the NPL to grow after an increase in the GDP. Empirical evidence, however, suggests that this could be an effect of the recent recession in Italy which causes an atypical bearish pattern in the GDP.

¹⁰ Some of these breaks correspond to an adjustment in the balance sheet of the Bank of Italy. See, REGULATION (EC) No 25/2009 OF THE EUROPEAN CENTRAL BANK of 19 December 2008 concerning the balance sheet of the monetary financial institutions sector, Official Journal of the European Union.

Similarly, the standardised betas reported in table 23 represent the unique contribution of the variables as predictors of the NPL. These values drop significantly with respect to the unstandardized betas (explained above) for the variables 'inflation' and 'GDP'.

Lastly, it is important to estimate the heteroscedasticity-consistent standard error in order to be sure of the model's robustness. After applying the same SPSS macro conducted for the Spanish regression, the results were reported in Appendix 4.

For this specification, the moderate presence of heteroscedasticity does not affect the explanatory power of the model since the statistical significance of the explanatory variables is not considerably modified (See appendix 4, table 27).

Conclusively, the non-performing loans index in Italy is mainly explained by the unemployment, the wage and the GDP based on the data covering the period from January 2004 to March 2012. The credit growth and the Inflation are not strong explanatory variables.

Conclusions

This research investigated the macroeconomic determinants of the non-performing loan indices in Spain and Italy for the period from January 2004 to March 2012. The NPL ratio was defined as the percentage of bad loans over the total loans. The macroeconomic variables were expressed as credit growth, wage, inflation, unemployment and GDP.

In both Spain and Italy, the macroeconomic variables are strong determinants of the Non-performing loans. However, of the five explanatory variables used, only unemployment, wage and GDP turned out to be statistically significant.

Another important finding of this paper is the influence of the lags. This research showed the strongest explanatory power to explain the NPL index when adding 6 months of lag for the Spanish economy and 12 months of lag for the Italian economy. Previous researches had found adding more than 12 to 18 months to be important for their models. Thus, under the updated time series, the bad loans are affected faster by changes in the economy. This reduction in the size of the lags could be caused by the volatility of the economy after the debt crisis.

The variable credit growth has a weak explanatory power in the Spanish model and it was excluded from the Italian model after finding it to be unreliable. Salas, V. and Saurina, J. (2002) had found this variable to be useful for explaining the increase in bad loans. However, the updated time series reduced the benefit of this variable in the model. This finding infers that after the debt crisis in Europe, the new Non-performing loans in the economy could be more affected by the existing loans than by new loans. This belief is supported by the new credit policies adopted by the banks after the debt crisis, which have affected the credit markets (see Chmelar A., (2012)).

Unemployment is a very strong variable in both countries. The partial correlation shows a defined positive relationship for this variable with the NPL index. This finding is

consistent with the researches reviewed in the literature review chapter. The analysed data suggests that a shift in unemployment has a faster impact on bad loans in the Spanish economy than in the Italian economy.

The variable 'Wage' is also explanatory in both Spain and Italy. Although the relationship was neutral, it is statistically significant. As far as I am concerned, this is a new finding for explaining the NPL in the mentioned economies. However, for further analyses, it is advisable to disaggregate this variable into different geographical categories within a country. This could lead to a stronger correlation between wage and the NPL index.

Certainly, inflation is not an explanatory variable of the NPL index, neither in Spain nor in Italy. Several regressions suggested the inclusion of this variable in the model not to be reliable under a statistical point of view. This, however, was a surprising result, provided that several papers had shown the inflation to be significant.

On the other hand, the GDP had a negative correlation for the Spanish data and a positive correlation for the Italian data. From these results it is difficult to determine a general relationship of this variable with the NPL. For further researches, however, it is advisable to analyse this variable on a quarterly basis in order to avoid interpolations and possible skewedness of the data.

Overall, the results of my research are satisfactory. The econometric tests required for multiple linear regressions were carefully considered in this dissertation. The assumptions were fulfilled and the tests done support the reliable outputs.

At a later academic stage, with more time available, it would be worth analysing the macroeconomic determinants of the Non-performing loans with econometric procedures such as the Vector Auto regressive and panel data models. These procedures are also adequate to analyse macroeconomic time series.

References

Brooks, C. (2008). *Introductory Econometrics for Finance*, 2nd Edition, Cambridge: Cambridge University Press.

Tabachnick, B.G. and Fidell, L.S. (2007), *Using Multivariate Statistics*, 5th edition, Boston: Allyn and Bacon.

Kenneth, F. W. (1995). *Time series analysis and macroeconomic modelling*, 1st Edition, Aldershot: Edward Elgar Publishing Limited

Fernández, S., Martínez, J. and Saurina, J. (2000). Credit Growth, Problem Loans and Credit Risk Provisioning In Spain, Bank of Spain, Working paper N. 0018

Salas, V. and Saurina, J. (2002). Credit Risk in Two Institutional Regimes: Spanish Commercial and Savings Banks, *Journal of Financial Services Research*, Vol. 22, pp. 203–224.

Bergaz, J. (2011). European outlook for non-core and non-performing loan portfolios, *Pricewaterhouse coopers*, Issue 3, pp. 12.

Bofondi, M. and Ropele, T. (2011). Macroeconomic determinants of bad loans: evidence from Italian banks, *Banca D'Italia*, Number 89

Babouček, I. and Jančar, M. (2005). Effects of Macroeconomic Shocks to the Quality of the Aggregate Loan Portfolio, *Czech National Bank*, Series 1

Mojon, B. and Peersman, G. (2001). A VAR Description of the Effects of Monetary Policy in the Individual Countries of the Euro Area, *European Central Bank*, Paper 92

Espinoza, R. and Prasad, A. (2010). Nonperforming Loans in the GCC Banking System and their Macroeconomic Effects, *International Monetary Fund*, Working paper 10/224

Segoviano, M., Goodhart, C. and Hofmann, B. (2006). Default, Credit Growth and Asset Prices, *International monetary fund*, Working paper 06/223

Glogowski, A. (2008). Macroeconomic determinants of Polish banks' loan losses, *National Bank of Poland*, Working paper 53

Soltila, H. and Vihriälä, V. (1994). Finnish Bank's Problem Assets: Result of Unfortunate Asset Structure or too Rapid Growth, *Bank of Finland*, Discussion Papers, Paper 23/94.

Saba, I. and Kouser, R. (2012). Determinants of Non-Performing Loans: Case of US Banking Sector, *the Romanian Economic Journal*, Paper 44

Marcellino M. (2007). A comparison of time series models for forecasting GDP growth and inflation, Università Bocconi, IGIER and CEPR, <http://www.eui.eu/Personal/Marcellino/1.pdf>

Marcucci, J. and Quagliariello, M. (2005). Is bank portfolio riskiness procyclical? Evidence from Italy using a vector autoregression, *University of York*, Paper number 2005/09

Arpa, M. and Giulini, I. (2001). The influence of macroeconomic developments on Austrian banks: implications for banking supervision, *Bank for International Settlements (BIS)*, Paper No. 1

Simons, D. and Rolwes, F. (2008). Macroeconomic default modelling and stress testing, *De Nederlandsche Bank*, Presentation for the BIS's stress testing workshop 2008, <http://www.bis.org/bcbs/events/rtf08simonsrolwes.pdf>

Segoviano, M. (2006). Conditional Probability of Default Methodology, *International Monetary Fund*. Paper 03/2006

Chan-Law, J. (2006). Fundamentals-Based Estimation of Default Probabilities: A Survey, *International Monetary Fund*, Working Paper 06/149

Lütkepohl, H. (2004). Forecasting with VARMA Models, *Department of Economics, European University Institute*, <http://www.eui.eu/Personal/Luetkepohl/VARMA-r1.pdf>

Pryce, G. (2002). Heteroscedasticity: Testing and Correcting in SPSS, *University of Glasgow*,
<http://reocities.com/heartland/4205/SPSS/HeteroscedasticityTestingAndCorrectingInSPSS1.pdf>

Berry, W. D., and Feldman, S. (1985). Multiple Regression in Practice, *Sage University*, Paper Series on Quantitative Applications in the Social Sciences, series no. 07-050

Tabachnick, B. G., and Fidell, L. S. (1996). *Using Multivariate Statistics*, 3rd edition, New York: Harper Collins.

Hayes, A. F., (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation, *Behavior Research Methods*, series no. 2007, 39 (4), 709-722

Chmelar A., (2012). ECRI Statistical Package: Debt crisis hits Europe's retail credit markets, *European credit research institute*, http://www.ceps.be/publication_series

Appendix 1 - Source of Macroeconomic Variables

Variable	Acronym	Source for Spain	Source for Italy
Total Loans	NPL	Bank of Spain, Report: be04 - Entidades de crédito Credit entities reports - Assets	Bank of Italy, Money and Banking reports, Balance Sheet Of Banks Resident In Italy: Assets
Bad Loans	Credit	Bank of Spain, Report: be04 - Entidades de crédito Credit entities reports - Assets	Bank of Italy, Money and Banking reports, Balance Sheet Of Banks Resident In Italy: Assets
Cost of labor	Wage	Bank of Spain, Report: ie04 - Mercado Laboral Total labor cost per worker	Italian National Institute of Statistics, Labour reports Index of wage contract by contract
Inflation	Infla	Bank of Spain, Report: ie05 - Precios Monthly inflation index	Italian National Institute of Statistics, Prices reports Consumer price index for the whole nation
Unemployment	Unemplo	Eurostat, Unemployment statistics, Seasonally adjusted data	Italian National Institute of Statistics, Work reports Unemployment rate, seasonally adjusted data
Gross Domestic Product	GDP	Spanish National Statistics Institute Economy reports, PIB - Current prices	Italian National Institute of Statistics, National accounts reports Gross domestic product - quarterly data

Appendix 2 – Test of Assumptions of the Regression

*Note: Assumptions for definitive regressions (Spain: 6 months lag / Italy: 12 months lag)

A. Quality of the Data*

Table 8 - Descriptive Statistics Spain

	N	Minimum	Maximum	Mean	Std. Deviation
NPL	98	.4619	4.3691	1.739395	1.3891555
Credit	98	-1.1217	2.8288	.749519	.9908084
Wage	98	-.0250	.4417	.252891	.1297113
Infla	98	-1.2000	1.4000	.212245	.5887721
Unemplo	98	-.3000	1.0000	.132653	.2693159
GDP	98	-.3815	.7049	.322438	.3402094
Valid N (listwise)	98				

Graph 6 - Outliers for Spain

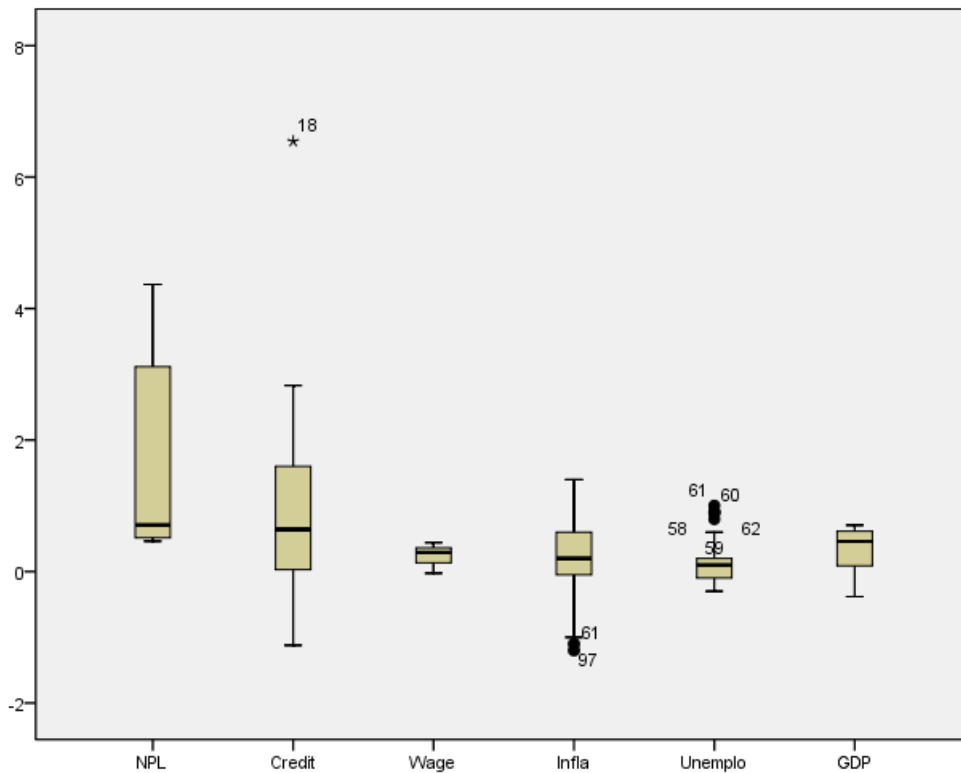
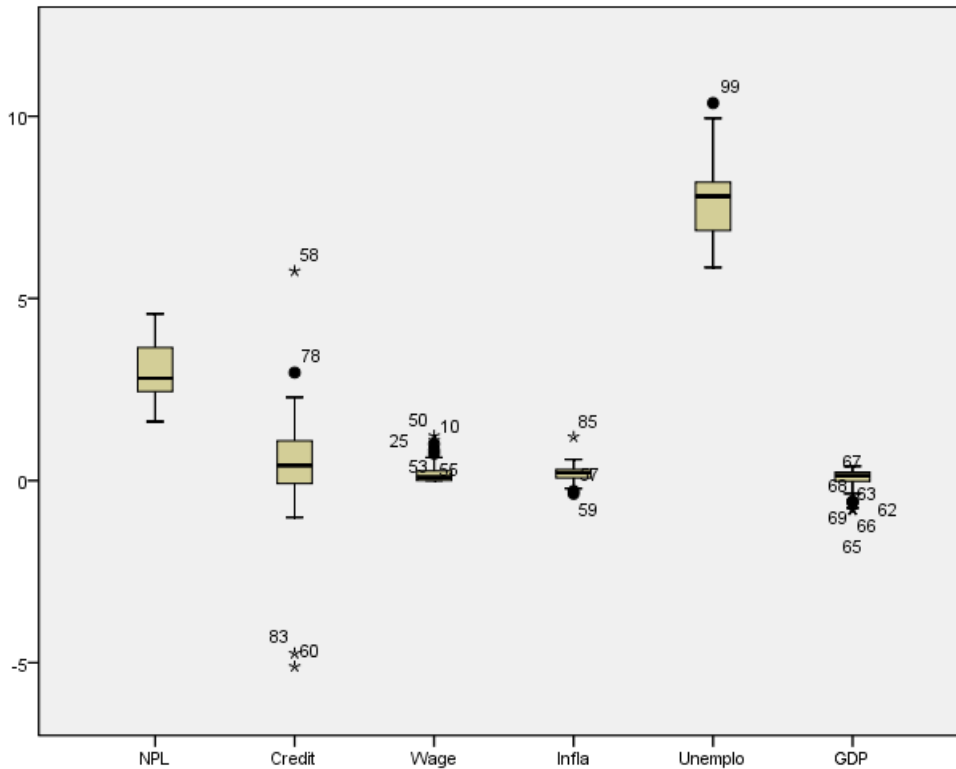


Table 9 - Descriptive Statistics Italy

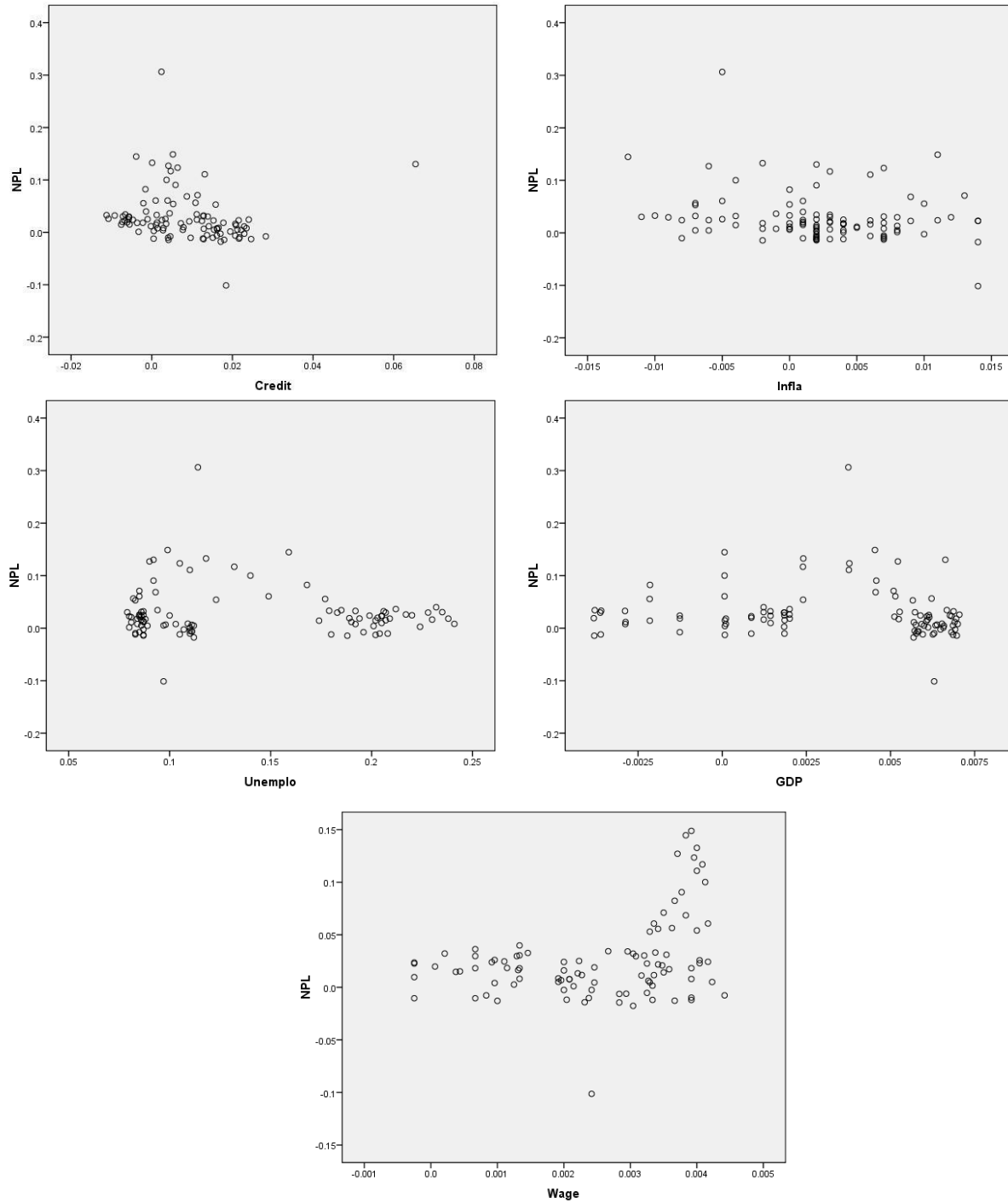
	N	Minimum	Maximum	Mean	Std. Deviation
NPL	96	1.6173	4.5754	3.019937	.7699607
Wage	96	.0000	1.2276	.197602	.2601972
Infla	96	-.3634	1.2000	.196402	.2078737
Unemplo	96	5.8536	10.3648	7.587750	.9480295
GDP	96	-.8023	.3944	.042982	.3082220
Valid N (listwise)	96				

Graph 7 - Outliers for Spain

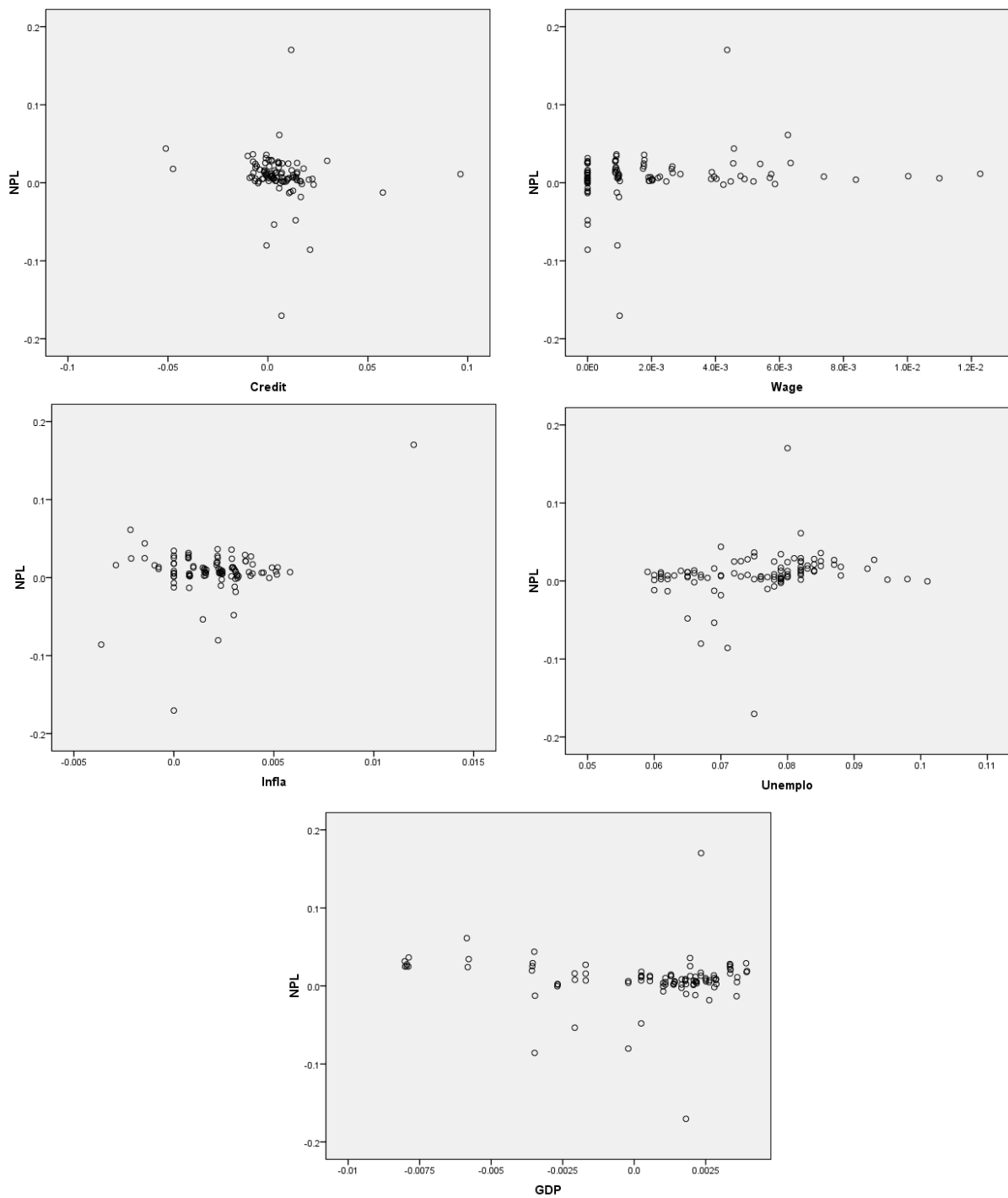


B. Linearity

Graph 8 - Linearity for Spanish Data



Graph 9 - Linearity for Italian Data



C. Heteroscedasticity

Table 10 - Breusch-Pagan and Koenker Test

Spain	Italy
<p>BP&K TESTS</p> <p>Sample size (N): 92</p> <p>Number of predictors (P): 4</p> <p>Breusch-Pagan test for Heteroscedasticity (CHI-SQUARE df=P) 9.971</p> <p>Significance level of Chi-square df=P (H0:homoscedasticity) .007</p> <p>Koenker test for Heteroscedasticity (CHI-SQUARE df=P) 12.588</p> <p>Significance level of Chi-square df=P (H0:homoscedasticity) .0276</p>	<p>BP&K TESTS</p> <p>Sample size (N): 84</p> <p>Number of predictors (P): 5</p> <p>Breusch-Pagan test for Heteroscedasticity (CHI-SQUARE df=P) 14.966</p> <p>Significance level of Chi-square df=P (H0:homoscedasticity) .0105</p> <p>Koenker test for Heteroscedasticity (CHI-SQUARE df=P) 12.498</p> <p>Significance level of Chi-square df=P (H0:homoscedasticity) .0286</p>

D. Normality

Table 11- Tests of Normality for Spain

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.063	92	.200 [*]	.982	86	.271

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

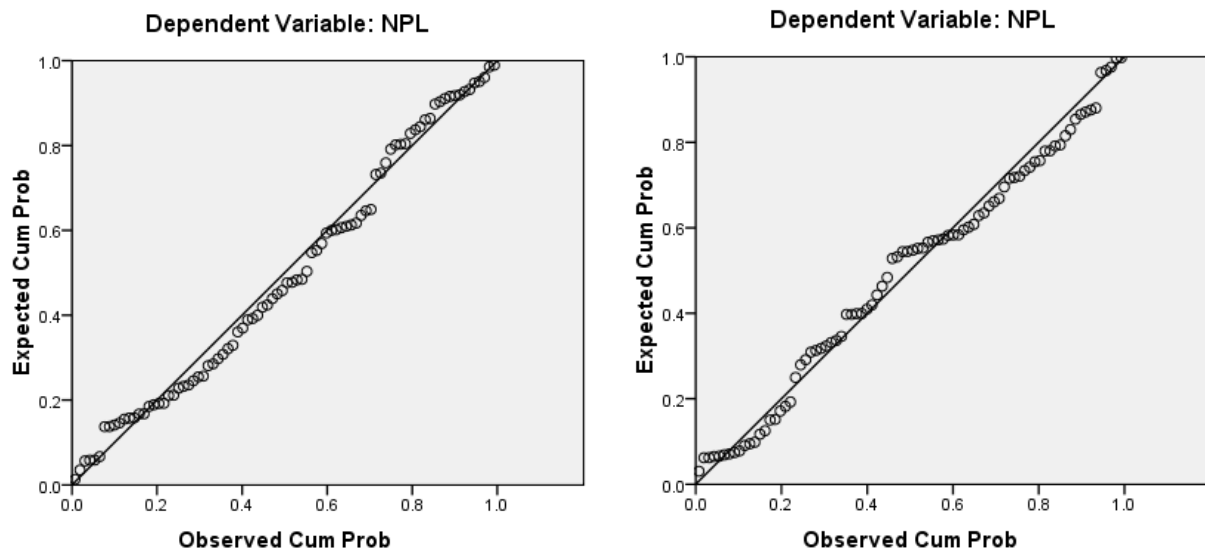
Table 12- Tests of Normality for Italy

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Unstandardized Residual	.077	84	.200 [*]	.968	84	.033

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Graph 10 - Normal P-P Plot of Standarized Residual*



*Notes: Spain left hand side, Italy Right hand side

E. Independence of errors

Table 13 – Collinearity Statistics

Model	Collinearity Statistics		Collinearity Statistics	
	SPAIN		ITALY	
	Tolerance	VIF*	Tolerance	VIF*
1 (Constant)				
Credit	.381	2.623	.877	1.140
Wage	.621	1.609	.948	1.054
Infla	N/A	N/A	.892	1.121
Unemplo	.539	1.856	.927	1.079
GDP	.392	2.551	.917	1.090

a. Dependent Variable: NPL

*Notes: A VIF smaller than 10 indicates no multicollinearity.

Appendix 3 – Regression Outputs

A. Regression Spain

Table 14 – Correlations

		NPL	LAGS(Credit,6)	LAGS(Wage,6)	LAGS(Unemplo,6)	LAGS(GDP,6)
Pearson Correlation	NPL	1.000	-.796	-.571	.513	-.836
	LAGS (Credit,6)	-.796	1.000	.402	-.453	.744
	LAGS (Wage,6)	-.571	.402	1.000	.209	.314
	LAGS (Unemplo,6)	.513	-.453	.209	1.000	-.512
	LAGS (GDP,6)	-.836	.744	.314	-.512	1.000
Sig. (1-tailed)	NPL	.	.000	.000	.000	.000
	LAGS (Credit,6)	.000	.	.000	.000	.000
	LAGS (Wage,6)	.000	.000	.	.023	.001
	LAGS (Unemplo,6)	.000	.000	.023	.	.000
	LAGS (GDP,6)	.000	.000	.001	.000	.
N	NPL	92	92	92	92	92
	LAGS (Credit,6)	92	92	92	92	92
	LAGS (Wage,6)	92	92	92	92	92
	LAGS (Unemplo,6)	92	92	92	92	92
	LAGS (GDP,6)	92	92	92	92	92

Table 15 – Model Summary^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.944 ^a	.892	.887	.4712987

a. Predictors: (Constant), LAGS(GDP,6), LAGS(Wage,6), LAGS(Unemplo,6), LAGS(Credit,6)

b. Dependent Variable: NPL

Table 16 – ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	159.629	4	39.907	179.663	.000 ^a
	Residual	19.325	87	.222		
	Total	178.954	91			

a. Predictors: (Constant), LAGS(GDP,6), LAGS(Wage,6), LAGS(Unemplo,6), LAGS(Credit,6)

b. Dependent Variable: NPL

Table 17 – Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	3.606	.114		31.663	.000			
	LAGS (Credit,6)	-.246	.081	-.173	-3.032	.003	-.796	-.309	-.107
	LAGS (Wage,6)	-4.760	.482	-.441	-9.877	.000	-.571	-.727	-.348
	LAGS (Unemplo,6)	1.634	.245	.320	6.666	.000	.513	.581	.235
	LAGS (GDP,6)	-1.640	.228	-.404	-7.186	.000	-.836	-.610	-.253

a. Dependent Variable: NPL

Table 18 – Collinearity Diagnostics^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions				
				(Constant)	LAGS(Credit,6)	LAGS(Wage,6)	LAGS(Unemplo,6)	LAGS(GDP,6)
1	1	3.328	1.000	.01	.01	.01	.00	.01
	2	1.236	1.641	.01	.02	.00	.25	.02
	3	.218	3.910	.37	.36	.01	.42	.00
	4	.135	4.965	.03	.37	.05	.15	.96
	5	.083	6.337	.58	.23	.93	.17	.00

a. Dependent Variable: NPL

Table 19 – Statistics^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	-.472228	3.982489	1.813399	1.3244496	92
Residual	-1.0457309	1.1093434	.0000000	.4608241	92
Std. Predicted Value	-1.726	1.638	.000	1.000	92
Std. Residual	-2.219	2.354	.000	.978	92

a. Dependent Variable: NPL

B. Regression Italy

Table 20 – Correlations

		NPL	LAGS(Credit,12)	LAGS(Wage,12)	LAGS(Infla,12)	LAGS(Unemplo,12)	LAGS(GDP,12)
Pearson Correlation	NPL	1.000	-.234	-.128	.084	.860	.179
	LAGS (Credit,12)	-.234	1.000	.039	-.149	-.254	.157
	LAGS (Wage,12)	-.128	.039	1.000	.174	-.072	-.071
	LAGS (Infla,12)	.084	-.149	.174	1.000	-.033	.183
	LAGS (Unemplo,12)	.860	-.254	-.072	-.033	1.000	-.075
	LAGS (GDP,12)	.179	.157	-.071	.183	-.075	1.000
Sig. (1-tailed)	NPL	.	.016	.123	.224	.000	.052
	LAGS (Credit,12)	.016	.	.362	.087	.010	.077
	LAGS (Wage,12)	.123	.362	.	.057	.259	.261
	LAGS (Infla,12)	.224	.087	.057	.	.384	.047
	LAGS (Unemplo,12)	.000	.010	.259	.384	.	.247
	LAGS (GDP,12)	.052	.077	.261	.047	.247	.
N	NPL	84	84	84	84	84	84
	LAGS (Credit,12)	84	84	84	84	84	84
	LAGS (Wage,12)	84	84	84	84	84	84
	LAGS (Infla,12)	84	84	84	84	84	84
	LAGS (Unemplo,12)	84	84	84	84	84	84
	LAGS (GDP,12)	84	84	84	84	84	84

Table 21 – Model Summary ^b

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.899 ^a	.809	.797	.3508129

a. Predictors: (Constant), LAGS(GDP,12), LAGS(Wage,12), LAGS(Unemplo,12), LAGS(Infla,12), LAGS(Credit,12)

b. Dependent Variable: NPL

Table 22 – ANOVA ^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	40.621	5	8.124	66.013	.000 ^a
	Residual	9.599	78	.123		
	Total	50.220	83			

a. Predictors: (Constant), LAGS(GDP,12), LAGS(Wage,12), LAGS(Unemplo,12), LAGS(Infla,12), LAGS(Credit,12)

b. Dependent Variable: NPL

Table 23 – Coefficients ^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	3.334	.387		-8.625	.000			
	LAGS (Credit,12)	-.035	.050	-.037	-.691	.291	-.234	-.078	-.034
	LAGS (Wage,12)	-.177	.147	-.061	-1.200	.023	-.128	-.135	-.059
	LAGS (Infla,12)	.281	.197	.075	1.425	.158	.084	.159	.071
	LAGS (Unemplo,12)	.843	.050	.866	16.845	.000	.860	.886	.834
	LAGS (GDP,12)	.562	.125	.232	4.490	.000	.179	.453	.222

a. Dependent Variable: NPL

Table 24 – Collinearity Diagnostics ^a

Model	Dimension	Eigenvalue	Condition Index	Variance Proportions					
				(Constant)	LAGS(Credit,12)	LAGS(Wage,12)	LAGS(Infla,12)	LAGS(Unemplo,12)	LAGS(GDP,12)
1	1	3.481	1.000	.00	.02	.03	.02	.00	.01
	2	.991	1.874	.00	.02	.04	.00	.00	.78
	3	.719	2.201	.00	.53	.01	.21	.00	.03
	4	.475	2.706	.00	.00	.90	.04	.00	.06
	5	.330	3.249	.00	.34	.01	.71	.00	.12
	6	.005	26.067	.99	.08	.01	.01	.99	.00

a. Dependent Variable: NPL

Table 25 – Residuals Statistics ^a

	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	1.740554	4.176358	2.924762	.6995768	84
Residual	-.6570631	1.0104660	.0000000	.3400822	84
Std. Predicted Value	-1.693	1.789	.000	1.000	84
Std. Residual	-1.873	2.880	.000	.969	84

a. Dependent Variable: NPL

Appendix 4 – Heteroskedasticity-consistent standard error estimators

Table 26 – Model Fit for Spain

```

HC Method
3

Criterion Variable
NPL

Model Fit:
      R-sq      F      df1      df2      p
      .8920    188.1535    4.0000    87.0000    .0000

Heteroskedasticity-Consistent Regression Results
      Coeff      SE(HC)      t      P>|t|
Constant      3.6065      .1252    28.8138    .0000
Credit        -.2459      .1035    -2.3750    .0197
Wage          -4.7603      .4944    -9.6286    .0000
Unemplo       1.6341      .2591     6.3058    .0000
GDP           -1.6403      .2296    -7.1444    .0000

Covariance Matrix of Parameter Estimates
      Constant      Credit      Wage      Unemplo      GDP
Constant      .0157      -.0017      -.0489      .0026      .0049
Credit        -.0017      .0107      -.0100      .0077      -.0181
Wage          -.0489      -.0100      .2444      -.0635      -.0189
Unemplo       .0026      .0077      -.0635      .0672      .0075
GDP           .0049      -.0181      -.0189      .0075      .0527

Setwise Hypothesis Test
      F      df1      df2      p
      51.0418    1.0000    87.0000    .0000

Variables in Set:
GDP

----- END MATRIX -----

```

Table 27 – Model Fit for Italy

HC Method
3

Criterion Variable
NPL

Model Fit:

R-sq	F	df1	df2	p
.8089	93.4245	5.0000	78.0000	.0000

Heteroskedasticity-Consistent Regression Results

	Coeff	SE (HC)	t	P> t
Constant	3.3340	.4006	-8.3228	.0000
Credit	-.0349	.0624	-.5591	.5777
Wage	-.1766	.1132	-1.5604	.0227
Infla	.2812	.3816	.7370	.4633
Unemplo	.8428	.0491	17.1623	.0000
GDP	.5624	.1084	5.1889	.0000

Covariance Matrix of Parameter Estimates

	Constant	Credit	Wage	Infla	Unemplo	GDP
Constant	.1605	-.0157	.0056	-.0882	-.0193	-.0006
Credit	-.0157	.0039	-.0005	.0103	.0017	-.0019
Wage	.0056	-.0005	.0128	-.0093	-.0010	-.0001
Infla	-.0882	.0103	-.0093	.1456	.0085	-.0119
Unemplo	-.0193	.0017	-.0010	.0085	.0024	.0005
GDP	-.0006	-.0019	-.0001	-.0119	.0005	.0117

Setwise Hypothesis Test

F	df1	df2	p
26.9251	1.0000	78.0000	.0000

Variables in Set:

GDP

----- END MATRIX -----