LUND UNIVERSITY
School of Economics and Management

Determinants of Non-Performing Loans

A Panel Data Empirical Analysis for South European Countries

By

Tom Ghorbani

Viktor Jakobsson

May 29, 2019

Master's Program in Economics

Supervisor: Milda Norkute

Abstract

The paper analyzes the determinants of non-performing loans in 19 European banks in Portugal, Italy, Greece and Spain based on quarterly data between 2006–2018. We implement a panel data model with interactive effects, which accounts for unobserved heterogeneity and cross-sectional dependence. The estimation approach we employ uses an analytical and iterative bias correction based on the work of Bai (Panel Data Models with Interactive Fixed Effects, *Econometrica*, 77, pp. 1229–1279, 2009). To the best of our knowledge, this is the first time such an approach is used to analyze the determinants of non-performing loans, while simultaneously accounting for cross-sectional dependence. The results indicate that the unemployment rate has a positive relationship with non-performing loans. When investigating bank-specific determinants, this paper finds some evidence suggesting that bank profitability and capitalization has a negative relationship with non-performing loans.

Keywords: Non-performing loans, Bank-specific determinants, Macroeconomic determinants, Interactive Effects, Cross-sectional dependence

Acknowledgements

We want to thank our supervisor Milda Norkute, Researcher at the Department of Economics at Lund University, for valuable guidance and input during the completion of this thesis. Her door was always open whenever we ran into trouble or had a question regarding our research. Without her passionate participation and input, the thesis could not have been successfully conducted.

Table of Contents

1. Introduction

The global financial crisis in 2008 and the European debt crisis in 2010-2011 lead to tremendous economic and social losses across the European continent. In order to limit the impact of future crises, regulators and bank managers work proactively to develop new determinants of financial instability. Given the central role of banks in the economy to act as a financial intermediary, provide liquidity and facilitate economic growth, it is essential for policymakers to model bank health and soundness.

In the wake of the global financial crisis and the European debt crisis, the banking sector experienced a surge in non-performing loans (NPLs).[1] The low-interest rate environment, paired with earlier accumulated capital contributed to excessive investments and increasing debt. Furthermore, increasing capital flows facilitated by the abolishment of currency risk, increasing financial integration and regulatory harmonization further contributed to internal imbalances. Together these circumstances increased the dependence of foreign funds to service the rapidly escalating debt levels, particularly in the so-called PIGS countries (Portugal, Italy, Greece and Spain). However, the global financial crisis caused a chain reaction, which quickly spread to the European continent. As uncertainty emerged, confidence plummeted, and demand effectively stopped; the European financial system experienced a liquidity crunch that deteriorated banks' balance sheets and saw non-performing loans soar.

Today, many of these European banks still struggle with overdue loans, affecting both individual bank profitability, macroeconomic stability and overall economic growth. Non-performing loans have been highlighted as a potential precursor for banking crises (Reinhart and Rogoff, 2010), but are also commonly studied when analyzing bank liquidity, credit risk, and bank profitability. As such, researchers have attempted to identify which macroeconomic and bank-specific factors determine these loans, hoping that they can provide important information on the dynamics of non-performing loans.

The purpose of this paper is to examine several commonly used bank-specific variables and macroeconomic variables in determining non-performing loans. This paper uses a balanced panel data set that contains quarterly information for 19 commercial banks between 2006–2018.

[1] The ratio of bank's NPLs is defined as the value of non-performing loans divided by the gross value of total loans. The most commonly used definition states that default occurs when the obligor has failed to conduct any form of repayment for more than 90 days (Beck et al., 2015).

The sample countries include Portugal, Italy, Greece, and Spain, as these countries were at the center of the European debt crisis.

Previous researchers have primarily applied different adaptations of the GMM estimation approach when examining the determinants of non-performing loans. Although these models have several advantages, they fail to account for cross-sectional dependence, which has been shown to cause biased estimators (Andrews, 2005). As such, this paper attempts to fill the gap in the literature by applying a panel data model with interactive effects, as suggested by Bai (2009). This approach is designed to estimate a panel data with interactive effects where the factors are estimated following the principal component analysis, which allows for consistent estimators under the presence of cross-sectional dependence.

The paper contributes to the existing research field in two ways. First, we consider a panel data model with interactive effects following the approach suggested by Bai (2009) which is based on bias-correction and accounts for cross-sectional dependence and unobserved heterogeneity. To the best of our knowledge, this is the first time such a method is utilized in the non-performing loans literature. Second, this paper complements the previous literature as it considers the impact of both the recent global financial crisis and the European debt crisis.

The remainder of this thesis has the following disposition: Section 2 reviews the literature on non-performing loans and its determinants. Section 3 presents a thorough description of the econometric framework. Section 4 describes the data and variables. Section 5 showcase the results. Section 6 presents a discussion of the results. Section 7 covers our most important conclusions and offers some suggestions for future research.

2. Literature review

The literature on the determinants of credit risk is vast, stretching from theoretical credit risk models to modern forward-looking models. However, a surge in non-performing loans during the recent global financial crisis and the European debt crisis attracted the attention of researchers and policymakers which in turn lead to an extensive body of research. The research on non-performing loans can be divided into two core categories; macroeconomic determinants and bank-specific determinants. In this section, we will provide an outline of the current non-performing loans literature that has been particularly influential for this paper.[2]

The first category has exclusively focused on understanding how macroeconomic determinants influence NPLs. Beck et al. (2013) conducted an influential and comprehensive study, combining two data sets from the IMF and the World Bank to create a panel covering 75 countries between 2000–2010. The authors applied several econometric specifications; panel data model with fixed effects, two-step difference GMM (Arellano and Bond, 1991) and system GMM (Blundell and Bond, 1998). The authors found that the nominal exchange rate and lending interest rates have a positive effect on NPLs, while real GDP growth rate and stock prices have a negative effect on NPLs. Moreover, the authors found that the relative size of the country could explain the magnitude of the estimated coefficient. In other words, a fall in share prices has a more significant impact in countries with larger stock-markets relative to their economic mass.

The second category includes bank-specific variables, which are believed to contain additional information in explaining the development of NPLs. Louzis et al. (2012) use quarterly data from the nine largest Greek banks for 2003–2009. Their data has information on different types of loans; mortgage, corporate, and consumer loans. The authors implement a restricted GMM procedure developed by Judson and Owen (1999), which uses a limited number of lagged regressors and one bank-specific variable at a time, reducing the need for additional instruments. Investigating the hypothesis that both bank-specific variables and macroeconomic variables affect NPLs, the authors find that proxies for bank performance and inefficiency may serve as leading indicators for determining future NPLs. Furthermore, return on equity is found to have a negative relationship with NPLs. The authors note that adding bank-specific variables does not significantly affect the estimated results of macroeconomic determinants (GDP,

[2] For a more comprehensive review we refer to Table A1 in Appendix I.

inflation, unemployment and real interest rate). However, they conclude that it is challenging to find significance among bank-specific variables.

Klein (2013) investigates the predictive ability of macroeconomic conditions and bank-specific factors. Using data from Central, Eastern and South Eastern Europe for 1998–2011, the author focuses on countries which have faced high and rising levels of NPLs. The author implements three different approaches; fixed effects, difference- and system GMM. His findings suggest that the level of NPLs could be attributed to macroeconomic and bank-specific variables. However, similar to Louzis et al. (2012), the explanatory power of the bank-specific variables are found to be relatively low. The results suggest that higher quality of bank management, measured by profitability, leads to lower NPLs. Moreover, the author found that excessive lending, measured by the loans-to-assets ratio, is positively related to NPLs. On the macroeconomic level, the results show that increased levels of unemployment, inflation, the volatility index (VIX) and depreciation of currency contribute to higher NPLs. The author concludes that although robustness checks confirm the obtained results, one should treat these results with caution as they are subject to caveats, such as differences in the classification of NPLs across countries and varying data quality.

Baselga-Pascal et al. (2015) sought to understand the determinants of NPLs by examining data from a large sample of commercial banks in several European countries for 2001–2011. Studying both bank-specific and macroeconomic determinants, the authors utilize the system GMM that controls for unobserved heterogeneity and endogeneity. Their findings indicate that bank-specific variables such as profitability, efficiency, and capitalization are inversely and significantly related to NPLs. As for macroeconomic determinants, lower inflation, higher interest rates and higher GDP growth reduce NPLs. The authors apply several robustness checks using different model specifications. Furthermore, they divide their sample into two periods to examine possible differences due to the impact of the 2008 financial crisis and the sovereign debt crisis starting in 2010 in the European banking sector. These robustness checks confirm their findings.

Ghosh (2015) adds to the literature by examining regional economic factors determining NPLs. The data contains information from savings institutions and commercial banks across 50 US states for the period 1984–2013, on national, state and industry-level. The author argues that data on a state-level ought to provide more information compared to a national level. These determinants are inflation rate, state house price index (HPI), home ownership ratios and state-GDP. Following the same approach used by Baselga-Pascal et al. (2015), the author found that return on assets, home ownership and HPI have an inverse relationship with NPLs. Moreover, higher capital-to-assets, loans-to-assets and loan loss provisions ratio increases NPLs. On the macroeconomic level, the author's results lack robustness throughout various model specifications. Interestingly, the results differ slightly for commercial banks and savings institutions, where size was found to be positively related to NPLs solely for commercial banks.

3. Econometric framework

Examining the existing research field, one can quickly conclude that the differenced GMM and system GMM are the two most frequently used approaches for researchers analyzing the determinants of NPLs (e.g. Louzis et al., 2012; Klein, 2013; Baselga-Pascal et al., 2015). These models have proven to be reliable in dealing with samples with fixed T and large N, while also accounting for possible endogeneity by introducing lagged variables serving as instruments. However, these methods have been proven to be biased when using small samples, N (Hayakawa, 2007). Therefore, this paper seeks to contribute to the existing literature on non-performing loans by introducing a new methodology. The estimation approach developed by Bai (2009) offers the ability to estimate panel data models with interactive effects. This setup relies on an approach where the error term is characterized by a multi-factor structure. Specifically, in a regression equation, the error term is decomposed into an idiosyncratic error term and a common component, which is a linear combination of a finite number of unobserved common factors and the cross-section specific factor loadings (Totty, 2017).

Another puzzling finding which stands out when reviewing the existing literature is the fact that no paper discusses the possible existence of cross-sectional dependence. This issue has received increasing attention recently as it might lead to inconsistent estimators in data panels. Sarafidis and Robertson (2009) investigated the impact of cross-sectional bias in GMM dynamic panel estimators. Their findings suggest that using lagged variables as instruments in a GMM estimator setting generates inconsistent estimators as $N \to \infty$ for fixed T. Moreover, the authors highlighted the inadequacy of GMM estimation when accounting for endogeneity, showing that the method violates the moment's condition. This is one of the strengths of the panel data model with interactive effects used in this paper, as it accounts for the existence of cross-sectional dependence (Bai, 2009).

In this paper, we will first introduce a panel data model with fixed effect as it is commonly used in the existing literature. Second, we will employ a panel data model with interactive effects, developed by Bai (2009). Lastly, serving as a robustness check we use the Common Correlated Effects (CCE) estimator developed by Pesaran (2006), which similarly controls for cross-sectional dependence by using cross-sectional averages.

3.1 Panel Data Model with Fixed effects

Panel data models are often estimated using fixed effects to account for time-varying and fixed unobserved heterogeneity for individual observations. Furthermore, it addresses the omitted-variable bias problem (Ghosh, 2015). We follow the existing literature and utilize a fixed effects framework where the standard specification for estimating the effect of bank-specific and macroeconomic variables on NPLs is given by

$$Y_{it} = \beta Z_{it} + \Gamma X_{it} + \alpha_i + \tau_t + \varepsilon_{it} \tag{1}$$

Y_{it} denotes the logit transformation of NPLs for bank i at period t. Defined in equation (2), this procedure generates a dependent variable that spans over the interval $(-\infty, \infty)$ and is distributed symmetrically (cf. Klein, 2013; Ghosh, 2015). Z_{it} denotes a vector of bank-specific variables and X_{it} is a vector of macroeconomic variables, defined in section 4.2. Individual and period fixed effects are represented by α_i and τ_t, respectively.

$$f(Y_{it}) = \frac{e^{Y_{it}}}{e^{Y_{it}} + 1} = \ln\left(\frac{Y_{it}}{1 - Y_{it}}\right) \tag{2}$$

Equation (1) accounts for time- and bank-specific heterogeneity within the sample. However, it leaves a lot of information within the error term. Moreover, the simple fixed effects model fails to account for potential cross-sectional dependence bias.

3.2 Panel Data Model with Interactive Effects

The panel regression model of equation (1) assumes that ε_{it} is an idiosyncratic error term, implying that there are no missing variables that are correlated with non-performing loans. However, the panel data model with interactive effects allows for the possibility that there exist unobserved common factors in the error term, which in turn might be related to the explanatory variables (Totty, 2017). This is the crucial difference between a panel data model with fixed effects and a panel data model with interactive effects, meaning that the error term from equation (1) now has the following multi-factor error structure:

$$\varepsilon_{it} = \lambda_i' F_t + u_{it} \tag{3}$$

Where λ_i is a (k x 1) vector of factor loadings capturing bank-specific responses to common shocks and F_t is a (k x 1) vector of unobserved common factors; u_{it} represents idiosyncratic errors; λ_i, F_t and u_{it} are all unobserved and estimated by the model.

The inclusion of unobserved common factors, F_t establishes dependence between NPLs across banks and countries since they are affected by common unobserved shocks. Specifically, this approach provides a more general and flexible specification of bank and country-specific factors. For example, macroeconomic shocks and other unobservable policy changes might affect the aggregate economy, while their impact is likely heterogeneous across different banks and countries.[3] The presence of unobserved common factors can cause cross-sectional dependence across NPLs, which is problematic for inference (Andrews, 2005) and can lead to bias if the common factors are correlated with the explanatory variables. To test whether our panel has cross-sectional dependence, we employ general diagnostic tests developed by Pesaran (2004) and Pesaran (2015).

Pesaran (2006) and Bai (2009) provide methods to obtain consistent estimates under the presence of potential unobserved common factors and cross-sectional dependence. The advantage of Bai's (2009) method lies in the construction of the common shock, F_t, and the factor loadings, λ_i, which are determined by the data. By assuming a factor error structure, it is possible to solve the potential issues of having similarity among banks by allowing the factor loadings to identify cross-sectional correlation in the data. The panel data model with interactive effects also differs in its approach as it allows the data to determine the form of the unobserved heterogeneity (Totty, 2017).

3.2.1 Interactive Effects Estimator

Introduced by Bai (2009), the interactive effects estimator determines the common factors using principal component analysis. This approach is constructed using OLS estimation, given the estimated number of common factors. Given the coefficients obtained from the OLS regression, the factor structure is subtracted from the data and later estimated by principal component analysis. However, since both the factor structure and the OLS coefficients are unknown, Bai (2009) suggests an iterative procedure where the model iterates between estimating the factor structure following principal components on the OLS residuals and estimating the regression coefficients following OLS on the de-factored data. As such, the iterative procedure generates bias-corrected estimators by conducting initial guesses of the common factors and their loadings until the model converges. This occurs when the change in the sum of squared residuals is below a specific threshold. The model is unique in its ability to identify important unobserved aggregate factors while simultaneously measuring their heterogenous impact using, in our case, bank-specific factor loadings (Totty, 2017).

[3] For example, the financial crisis or the debt crisis might have had different impacts based on country- or bank-specifics, which are unobservable. Similarly, changes and shocks to regulations of the financial system might have led to heterogeneous impacts, while being somewhat aggregate in nature (e.g. European Union).

The panel data model with interactive effects is created by using equation (1) and substituting in the error structure from equation (3), resulting in the following specification

$$Y_{it} = \beta Z_{it} + \Gamma X_{it} + \lambda_i' F_t + \alpha_i + \tau_t + u_{it} \tag{4}$$

The interactive effects estimator estimates the factor structure, but the number of factors needs to be specified in advance. In this paper, we implement the methodology proposed by Glorfeld (1995) which is an improvement of Horn’s (1965) Parallel Analysis. It is recognized as one of the more accurate methods (cf. Heyton et al. 2004; Dinno, 2009). To examine the robustness of the results, we add one more factor component to the specification suggested by the test.

3.3 Common Correlated Effects

Another widely used method for linear factor model estimation is the Common Correlated Effects (CCE) estimator developed by Pesaran (2006). The model is constructed by OLS applied to regressions, where the dependent and explanatory variables are augmented with cross-sectional averages (Pesaran, 2006). As such, these averages work as proxies for the factors. Hence the estimated model has the following structure

$$Y_{it} = \beta Z_{it} + \Gamma X_{it} + \eta \bar{Y}_t + \phi \bar{Z}_t + \delta \bar{X}_t + \varepsilon_{it} \tag{5}$$

where $\bar{Y}_t$, $\bar{Z}_t$ and $\bar{X}_t$ denotes the cross-sectional averages of Y_{it}, Z_t and X_t, respectively.

4. Data and variables

4.1 Data

The panel data sample used in this paper consists of bank-specific and macroeconomic variables for 19 commercial banks operating in four Southern European countries from 2006 to 2018 with a quarterly frequency. The sample countries are Portugal, Italy, Greece and Spain. The data is collected from four different sources; The World Bank, Bloomberg Terminal, Thomson Reuters Datastream and S&P Capital IQ. Some variables are readily available for analysis while others need to be constructed by combining the different datasets.

The selection of banks is based on a couple of criterions. First, we examine what bank information is available in Bloomberg Terminal and Thomson Reuters Datastream. From these sources we retrieve a list of 28 banks. However, information on some of these banks is missing for our requested time period. This is an important concern as the panel data model with interactive effects requires a balanced dataset. Moreover, some banks lack information on a quarterly basis and are hence excluded. Lastly, we do not include subsidiaries of banks that have a country of domicile outside of our sample countries. Table A2 in Appendix II provides a detailed list of the banks that are used in this paper and their respective countries.

The time period 2006–2018 is primarily selected since it covers both the global financial crisis in 2008 and the preceding European debt crisis in 2010–2011. Ideally, we would have had access to information dating further back in history as it would have provided a longer pre-crisis period. However, this was not possible due to limited data availability. Moreover, the information from 2012–2018 is relevant as many European banks still have legacy issues with non-performing while others have weathered the crisis, managed non-performing loans and returned to normal operations. Lastly, since we are particularly interested in the crises periods, we use quarterly data containing higher frequency observations as it allows for more precise estimates.

It is relevant to highlight that the construction of our dataset has a few shortcomings. Comparing our sample to some of the existing literature (e.g. Baselga-Pascual et al., 2015; Ghosh, 2015), one can quickly conclude that their sample is often more extensive. Ideally, we would have had access to the Bankscope database maintained by Bureau Van Dijk, which has been frequently used by previous researchers. Instead, we opted to construct our own database by using information from different sources and combine these. This may have created minor inconsistencies in the data collection, which we do not believe to be a severe issue.

4.2 Variables

Table 1 displays our dependent variable non-performing loans and the explanatory variables used in this paper. The explanatory variables consist of five bank-specific variables and five macroeconomic variables. Their expected sign in relation to non-performing loans is shown below, which will receive further attention in section 4.2.1 and 4.2.2.

Table 1 – Description of variables

	Description	Source	Expected sign
Dependent variable			
Non-Performing Loans	Non-performing Loans-to-Total Loans (%)	BB, DS & CIQ	
Explanatory variables: Bank-specific variables			
Asset Structure	Total loans-to-total assets (%)	BB, DS & CIQ	(+)
Capitalization	Equity capital-to-total assets (%)	BB, DS & CIQ	(+/−)
Diversification	Non-interest income-to-total income (%)	BB, DS & CIQ	(+/−)
Bank Profitability	Return on assets (%)	BB, DS & CIQ	(−)
Investor Profitability	Return on equity (%)	BB, DS & CIQ	(+)
Explanatory variables: Macroeconomic variables			
Gross Domestic Product	Real change in GDP, year-over-year (%)	WB	(−)
Unemployment	Unemployed-to-labor force (%)	WB	(+)
Inflation	Change in CPI (%)	WB	(+)
Current Account	Current Account to GDP (%)	BB	(−)
Interest Rate	Nominal fixing rate (%)	ECB	(+)

Note: BB = Bloomberg Terminal, DS = Datastream, CIQ = Capital IQ, WB = World Bank, ECB = European Central Bank

4.2.1 Dependent variable

We retrieve bank-level data of non-performing loans from our four sample countries. This approach is in line with earlier studies (e.g. Louzis et al., 2012; Klein, 2013; Ghosh, 2015). This measure is defined as the value of non-performing loans to gross loans. This ratio is typically found on banks' balance sheets and is simple to compute. Figure 1 shows the average development of NPLs in each respective country for 2006–2018. We include information from Germany to serve as a benchmark.

A potential issue when creating the NPL ratio is that the definition might vary across countries. As such, one should exercise caution when comparing NPL ratios across countries. The most commonly used definition states that a loan is considered non-performing when an obligor has failed to conduct any form of repayment for more than 90 days (Beck et al., 2015). However, some banks report loans which are overdue by 31 days or 61 days. An alternative approach is to use aggregate data from the IMF, which provides a more consistent definition of NPLs.

Unfortunately, the IMF dataset is restricted to a yearly frequency at country level. This is problematic since it contains fewer observations, does not allow to analyze individual bank-level NPL ratios and does not cover the entire time period investigated.

Figure 1 – Historical development of Non-Performing Loans

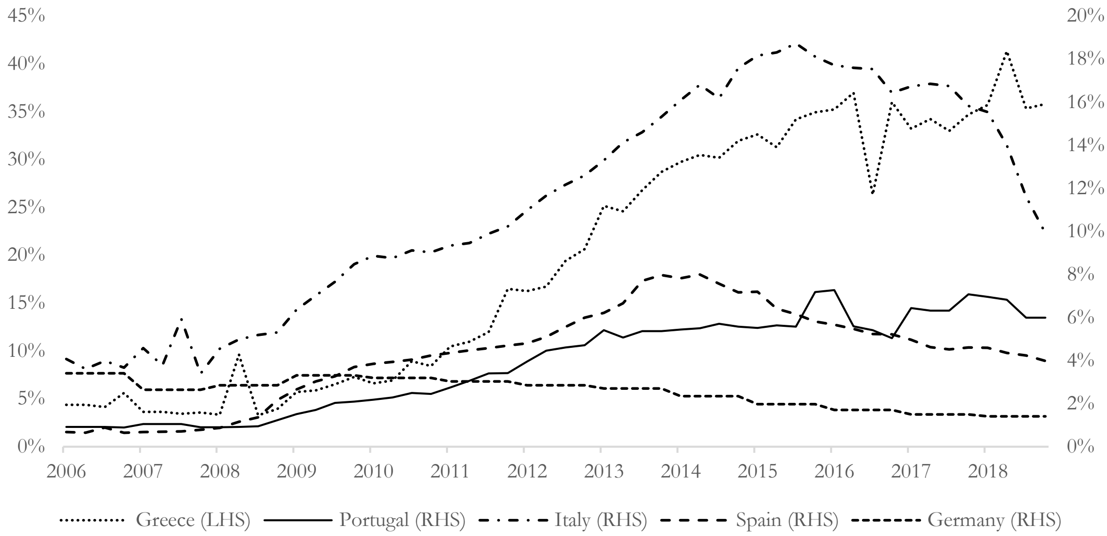


Figure 1 shows that while Germany has experienced a declining trend in NPLs, the PIGS countries suffered sharp increases in NPLs when the global financial crisis and the European debt crisis erupted. In contrast to Greece and Portugal, Italy and Spain have experienced some recovery from their respective peaks. However, the level of NPLs remains high, particularly in Greece and Italy.

4.2.2 Explanatory variables

Asset structure is defined as total loans over total assets and serves as a proxy for banks' general credit risk. The theory states that a relative increase in total loans compared to safer assets increases bank profitability, but introduces more liquidity risk (Trujillo-Ponce, 2013). Earlier studies have highlighted the relationship between asset structure and non-performing loans and other performance measures (Blasko and Sinkey, 2006 and Männasoo and Mayes, 2009). In line with earlier research (cf. Baselga-Pascual et al., 2015; Ghosh, 2015), we hypothesize that the asset structure will positively influence NPLs.

Capitalization is defined as common equity over total assets, capturing bank capital requirements from regulators to maintain financial soundness and decrease systemic risk. The impact of capitalization on non-performing loans is ambiguous. Some research finds that an increase in bank capitalization reduces bank-risk taking, risk of insolvency and probability of default (Furlong and Keeley, 1989; Gennotte and Pyle, 1991; Santos, 1999). Others conclude that imposed regulation forces banks into riskier investment strategies which might impact non-performing loans (Koehn and Santomero, 1980; Blum, 1999). Therefore, we do hypothesize that there exists an ambiguous relationship between capitalization and NPLs.

Diversification is defined as non-interest income over total income, capturing the level of risk-taking in the bank's business model. Bank income can be divided into interest and non-interest income. Interest income represents the traditional commercial banking income such as income from loans and other securities. Non-interest income consists of income earned from for example asset management, investment banking, trading, and derivatives activities as well as insurance underwriting. Previous research highlights a negative relationship between bank diversification and non-performing loans (Louzis et al., 2012), while others use size as a proxy for diversification, stating that a larger bank has better diversification possibilities which decrease non-performing loans (Salas and Saurina, 2002; Rajan and Dhal, 2003). Therefore, we do hypothesize that there exists an ambiguous relationship between diversification and NPLs.

Bank profitability is defined as return on assets, capturing the performance of banks. Theoretically, a profitable bank is less likely to have a higher ratio of non-performing loans (Berger and DeYoung, 1997; Poghosyan and Čihak, 2012). In line with previous research, we hypothesize a negative relationship between profitability and NPLs.

Investor profitability is defined as return on equity, capturing the performance of the bank's stock. A pure profit maximization strategy is likely to induce higher levels of risk-taking (García-Marco and Robles-Fernándes, 2008**;** Mehran and Mollineaux, 2012; Makri et al., 2014). In line with previous research, we hypothesize a positive relationship between investor profitability and NPLs.

Gross Domestic Product is defined as real GDP growth. The banking sector is in general procyclical, and an expanding economy should have beneficial effects for the banks and in extension decrease non-performing loans, something which previous research corroborates (Borio and Lowe, 2002; Klein, 2013; Beck et al., 2013). In line with previous literature, we expect a negative relationship between GDP and NPLs.

Unemployment is defined as the amount of unemployed over the total labor force and is another measure for the overall state of the economy. Worsening economic conditions and increasing unemployment has the potential to affect non-performing loans adversely (Bofondi and Ropele, 2013; Škarica, 2014). Hence, we hypothesize a positive relationship between unemployment and NPLs.

Inflation is defined as the percentage change in the consumer price index. Theory suggests that for unchanged nominal interest rates inflation should decrease the real value of debt and hence lower NPLs. However, higher inflation might be countered through rising nominal interest rates which reducers borrower's ability to service their debt as their real income decreases. As such, a rise in inflation should be accompanied by an increase in NPLs, given that income does not rise in line with inflation (Louzis et al., 2012; Baselga-Pascual et al., 2015; Ghosh, 2015). Hence, we hypothesize a positive relationship between inflation and NPLs.

Current Account to GDP is typically used as an indicator for the country's level of international competitiveness and the overall health of the economy. This variable is the ratio of the external deficit or surplus to GDP. As for the relationship between current account and banking distress specifically, Laeven and Valencia (2012) showed that banking crises are most prevalent in countries suffering from large current account deficits. Moreover, Kauko (2012) found a negative relationship between the current account balance and non-performing loans. Hence, we hypothesize a negative relationship between current account and NPLs.

Interest rate is the European central bank's nominal fixing rate expressed in percentages. The impact of interest rates on bank risk-taking and non-performing loans is widely studied, and the majority of research finds a positive relationship between the two (Hoggarth et al., 2005; Nkusu, 2011). Hence, we hypothesize a positive relationship between interest rates and NPLs.

Table 2 visualizes the descriptive statistics of the panel data. The high values for NPLs are worth to notice, as they reflect the extreme situation that occurred in the PIGS countries. Moreover, by construction, most variables stationary as they are expressed as ratios.

Table 2: Descriptive statistics

	N	Mean	Std. Dev.	Min	Max
Non-Performing Loans	988	0.112	0.1078	0	0.597
Asset Structure	988	0.687	0.1011	0.436	1
Capitalization	988	0.074	0.0283	-0.049	0.185
Diversification	988	0.286	0.1516	-0.559	0.879
Bank Profitability	988	0.000	0.0177	-0.126	0.054
Investor Profitability	988	0.017	0.1947	-0.918	0.986
Gross Domestic Product	988	-0.001	0.0317	-0.102	0.067
Unemployment	988	0.138	0.0638	0.057	0.279
Inflation	988	0.015	0.0159	-0.024	0.055
Current Account	988	-0.013	0.0588	-0.149	0.070
Interest Rate	988	0.012	0.0138	0	0.043

4.3 Endogeneity

The variables utilized in our econometric approach might be prone to endogeneity which could potentially invalidate our estimated results. For instance, banks might have incentives to increase their profitability in the short-term to increase return on assets. Despite this, it is difficult to account for unobservable characteristics that affect bank risk (e.g., managerial ability). Given the characteristics of the macroeconomic variables, we follow the previous literature in arguing that these variables are strictly exogenous (e.g., Louzis et al., 2012; Baselga-Pascual, 2015; Ghosh, 2015). However, the exogeneity assumption would be too strong for the bank-specific variables. Similar to previous research, we argue that bank managers incorporate the future predicted levels of NPLs in their decision making. These decisions are not instantaneously implemented, creating a delay between inception and execution (Louzis et al., 2012). Moreover, management fails to account for future random shocks affecting NPLs since they are unpredictable. Hence, we assume a weak form of exogeneity for bank-specific variables in line with Louzis et al. (2012), which still allows us to estimate the determining factors of non-performing loans. Potential endogeneity concerns will receive additional attention in the discussion, see section 6.2.

5. Results

The results presented in this section are split into four sub-sections to carefully analyze the determinants of NPLs. First, section 5.1 presents the panel regression for our entire dataset. Second, section 5.2 investigates potential differences amongst countries. Third, section 5.3 introduces the panel data model with interactive effects and the common correlated effects estimator. Last, section 5.4 explores whether our results remain robust by changing the data sample and the number of factors.

5.1 Panel regression results

Initially, we run three baseline panel regressions by pooling over the banks for all countries to analyze the determinants of NPLs. These estimates allow us to have a first look at the results, even though the standard panel regressions fail in adjusting for earlier mentioned cross-sectional dependence. The three panel regressions are included in Table 3 and consist of the following; Model 1 with pooled data and no fixed effects, Model 2 with individual bank fixed effects (B) and Model 3 both individual bank and time fixed effects (B & T).

Table 3: Panel regression results for all countries

	(1)	(2)	(3)
Asset Structure	0.715	0.649	0.410
	(0.936)	(0.857)	(0.878)
Capitalization	2.357	2.067**	0.979
	(2.867)	(0.999)	(1.102)
Diversification	0.291	0.282*	0.349
	(0.256)	(0.166)	(0.236)
Bank Profitability	-1.080	-0.765	-0.416
	(3.485)	(1.908)	(2.485)
Investor Profitability	-0.168	-0.129	-0.035
	(0.367)	(0.182)	(0.180)
Gross Domestic Product	1.742	1.965**	2.661*
	(1.477)	(0.916)	(1.580)
Unemployment	6.727***	7.088***	6.784***
	(1.677)	(1.265)	(2.229)
Inflation	2.880	3.253***	2.665
	(2.391)	(1.180)	(4.062)
Current Account	3.109**	2.945**	2.954*
	(1.263)	(1.489)	(1.533)
Interest Rate	-31.281***	-31.470***	-49.646***
	(5.131)	(4.404)	(9.292)
Fixed Effects	No	B	B & T
N	988	988	988
Adj R^2	0.452	0.773	0.789
RMSE	0.435	0.425	0.409

*Note: * p < 0.10, ** p < .05, *** p < .01.* (Wild-Cluster Bootstrap std. errors).

When we include both individual bank and time fixed effects in Model 3, GDP, unemployment, current account, and interest rate are all significant. GDP has a positive impact at the 10%

significance level. Unemployment has a positive relationship at the 1% significance level, while interest rates have a negative relationship at the 1% significance level.

5.2 Individual panel regression results

Subsequently, we run three different panel regressions for each country to analyze the determinants of non-performing loans in each country separately. By dividing the sample into national sub-samples, we can analyze the bank-specific determinants at country levels while at the same time controlling for macroeconomic conditions. This allows us to analyze country-specific similarities and differences. However, the problems with cross-sectional dependence remain. The three OLS models are included in Table 4 and consist of the following; Model 4 with panel data models and no fixed effects, Model 5 with individual bank fixed effects and Model 6 with both individual bank and time fixed effects.

When we include both individual bank and time fixed in Model 6, we find significant bank-specific variables in the Spanish and the Italian sample. For Spain, capitalization has a negative impact on non-performing loans at the 5% significance level, and diversification has a positive effect on the non-performing loans ratio at the 5% significance level. For Italy, asset structure has a positive effect on the non-performing loans ratio at the 5% significance level. This implies that excessive lending leads to higher levels of NPLs. For macroeconomic variables, unemployment has a positive impact at the 1% level, current account has a positive impact at the 5% level while interest rates have a negative relationship at the 1% level.

These variables are significant over all three model specifications, which is promising but again, it is important to acknowledge that these results are likely to suffer from cross-sectional dependence bias. To investigate whether this is true, we conduct Pesaran (2004, 2015) tests for cross-sectional dependence. The results from this test, found in Table A3 in Appendix III, shows that we can reject the null hypothesis of cross-sectional independence in all cases.

Table 4: Individual panel regression results

Country	Variable	(4)	(5)	(6)
SPAIN	Asset Structure	0.640	2.891*	2.855
		(0.921)	(1.605)	(1.773)
	Capitalization	-21.349**	-19.265**	-24.267**
		(9.153)	(8.390)	(9.484)
	Diversification	1.094**	1.253**	1.522**
		(0.503)	(0.542)	(0.615)
	Bank Profitability	76.849	31.233	14.770
		(197.734)	(177.860)	(151.157)
	Investor Profitability	-4.850	-2.908	-2.063
		(9.884)	(8.787)	(7.291)
PORTUGAL	Asset Structure	-1.268	-0.474	-0.317
		(0.913)	(1.879)	(1.764)
	Capitalization	13.090	11.913	5.825
		(9.352)	(8.168)	(7.996)
	Diversification	0.183	0.184	0.295
		(0.308)	(0.423)	(0.629)
	Bank Profitability	-15.423	-18.964	-14.835
		(45.981)	(29.968)	(34.306)
	Investor Profitability	-0.304	-0.037	0.190
		(0.754)	(0.379)	(0.593)
GREECE	Asset Structure	0.414	-0.139	-0.841
		(0.658)	(1.530)	(1.393)
	Capitalization	3.222***	2.912***	1.468
		(0.781)	(1.086)	(1.218)
	Diversification	0.130	0.028	0.134
		(0.440)	(0.546)	(0.578)
	Bank Profitability	-1.460	-1.289	-0.227
		(2.128)	(2.206)	(2.593)
	Investor Profitability	0.145	0.093	0.160
		(0.145)	(0.116)	(0.197)
ITALY	Asset Structure	1.579	1.619**	1.975**
		(1.092)	(0.700)	(0.788)
	Capitalization	-2.433	-1.986	-3.573
		(8.792)	(3.031)	(3.780)
	Diversification	0.057	0.073	-0.041
		(0.509)	(0.173)	(0.194)
	Bank Profitability	-11.540	-11.120	-15.304
		(37.962)	(28.644)	(27.226)
	Investor Profitability	0.014	0.002	0.396
		(2.552)	(2.071)	(1.898)
ALL	Gross Domestic Product	2.061***	2.069***	0.760
		(0.717)	(0.688)	(1.119)
	Unemployment	8.174***	7.960***	7.423***
		(0.946)	(0.965)	(1.694)
	Inflation	2.179	2.298*	-0.914
		(1.400)	(1.206)	(2.534)
	Current Account	2.390*	2.892**	3.161**
		(1.417)	(1.264)	(1.312)
	Interest Rate	-28.285***	-28.899***	-47.889***
		(3.867)	(3.290)	(9.646)
	Fixed Effects	No	Bank	Bank & Time
	N	988	988	988
	Adj. R^2	0.763	0.883	0.900
	RMSE	0.411	0.412	0.391

*Note: * $p < 0.10$, ** $p < .05$, *** $p < .01$.* (Robust Std. Errors).

5.3 Individual panel regression results using interactive effects

Finally, we run the panel data model with interactive effects for each individual country to analyze the determinants of non-performing loans. By doing so, we account for the previously mentioned problem with cross-sectional dependence. As the model requires a pre-determined number of factors, we conduct Glorfeld's (1995) improved approach for the Parallel Analysis. The model suggests a three-factor specification, shown in Table A4.

The resulting panel data model with interactive effects, using three factors, are found in Table 5. It displays the results of the following model specifications: Model 7 with interactive effects and no fixed effects, Model 8 with individual bank fixed effects, Model 9 with individual bank and time fixed effects. Model 10 is the panel data model with individual bank and time specific fixed effects, estimated using the common correlation effects estimator.

For Spain, asset structure is positive and significant at the 1% level for Model 8 and 9. Bank profitability and investor profitability are significant at the 1% level across Model 7, 8 and 9. Asset structure has a positive impact on non-performing loans. Bank profitability has a negative impact on non-performing loans, while investor profitability has a positive impact. The positive impact of asset structure and the negative impact of bank profitability on NPLs is in line with the findings of Baselga-Pascual et al., (2015). The positive impact of investor profitability corroborates the findings of García-Marco and Robles-Fernándes (2008). This result is of interest as the authors looked at the Spanish banking industry, and our results are in line with theirs.

For Portugal, asset structure, capitalization and bank profitability are all significant at the 1% level across Model 7, 8 and 9. Capitalization has a positive relationship with non-performing loans, while asset structure and bank profitability have a negative relationship. The negative impact of asset structure on NPLs contradicts the findings of earlier research (Baselga-Pascual et al., 2015; Ghosh, 2015). The positive impact of capitalization on NPLs opposes the findings of Baselga-Pascual et al. (2015). However, they hypothesized an ambiguous impact for capitalization as they found a negative relationship for non-performing loans but a positive relationship for an alternative risk measure (Z-score). Bank profitability has the same sign and the significance level for both Portugal and Spain, even though the effect is larger in Spain. These results will be further analyzed and discussed in Section 6.

Table 5: Results for the panel data model with interactive effects and CCE

Country	Variable	Interactive Effects – 3 Factors (7)	(8)	(9)	CCE (10)
SPAIN	Asset Structure.	0.482	3.138***	2.938***	0.052
		(0.684)	(0.583)	(0.659)	(0.048)
	Capitalization	-12.084***	0.112	-0.959	-2.126
		(3.521)	(3.259)	(3.559)	(1.791)
	Diversification	0.148	-0.397*	-0.211	0.041
		(0.214)	(0.211)	(0.245)	(0.056)
	Bank Profitability	-103.737***	-167.665***	-140.980***	17.420
		(35.129)	(32.830)	(37.126)	(17.661)
	Investor Profitability	6.273***	8.046***	7.214***	-0.655
		(1.982)	(1.727)	(1.990)	(1.051)
PORTUGAL	Asset Structure.	-1.518***	-1.289***	-1.231***	0.045
		(0.408)	(0.373)	(0.435)	(0.119)
	Capitalization	15.920***	7.619***	7.384***	0.976
		(2.219)	(2.066)	(1.843)	(0.943)
	Diversification	0.256	0.366**	0.253	-0.048
		(0.174)	(0.152)	(0.184)	(0.050)
	Bank Profitability	-35.876***	-22.229***	-16.602***	0.382
		(9.092)	(8.510)	(8.077)	(0.307)
	Investor Profitability	0.560	0.382	0.274	-0.096
		(0.342)	(0.310)	(0.291)	(0.085)
GREECE	Asset Structure.	0.770**	0.442	0.919**	0.473
		(0.323)	(0.277)	(0.372)	(0.541)
	Capitalization	5.301***	2.049**	1.279*	0.549
		(1.033)	(0.889)	(0.764)	(0.507)
	Diversification	0.090	0.198	0.170	0.043
		(0.265)	(0.266)	(0.264)	(0.114)
	Bank Profitability	-2.654*	-0.897	-0.537	-0.427
		(1.438)	(1.408)	(1.316)	(0.742)
	Investor Profitability	0.089	-0.060	-0.087	0.338
		(0.128)	(0.100)	(0.104)	(0.337)
ITALY	Asset Structure.	0.248	0.524	0.877	-0.123
		(0.367)	(0.676)	(0.651)	(0.695)
	Capitalization	-0.985	-1.588	-3.620	1.881
		(2.632)	(3.340)	(2.315)	(2.497)
	Diversification	-0.092	-0.057	0.002	0.004
		(0.139)	(0.158)	(0.174)	(0.056)
	Bank Profitability	-26.962**	-20.428	-19.198*	-65.485**
		(11.896)	(12.478)	(10.894)	(27.964)
	Investor Profitability	1.543**	1.034	0.899	4.787**
		(0.637)	(0.648)	(0.599)	(2.321)
ALL	Gross Domestic Product	1.169*	1.219**	0.493	0.277
		(0.658)	(0.579)	(0.639)	(0.994)
	Unemployment	9.257***	9.506***	11.562***	1.135
		(0.702)	(0.539)	(0.631)	(1.366)
	Inflation	0.411	0.543	5.826***	0.968
		(1.888)	(1.515)	(1.797)	(2.933)
	Current Account	5.306***	6.704***	7.068***	-2.180
		(1.052)	(0.970)	(0.822)	(2.553)
	Interest Rate	-3.523	13.733*	-4.119	1.942
		(6.418)	(7.290)	(6.519)	(12.885)
	Fixed Effects	No	B	B & T	B & T
	N	988	988	988	988
	Adj. R^2	N/A	N/A	N/A	0.559
	RMSE	0.234	0.211	0.208	0.496

Note: * *p < 0.10*, ** *p < .05*, *** *p < .01*. (Robust Std. Errors).

For Greece, asset structure is significant and positive at the 5% level for Model 7 and 9. Capitalization is significant and positive across Model 7, 8 and 9 decreasing in statistical significance level from 1% to 5% to 10%. For Italy, no bank-specific variables are significant over the three models.

For the macroeconomic variables, unemployment and current account are both significant at the 1% level for Model 7, 8 and 9. Inflation is significantly positive at the 1% level for Model 9, in line with previous research (cf. Louzis, 2012; Baselga-Pascual et al., 2015). Unemployment has a positive impact on NPLs, corroborating previous studies which find the same relationship (cf. Louzis et al., 2012; Škarica, 2014; Ghosh, 2015). Current account has a positive impact on NPLs, which goes against the findings of Kauko (2012), a puzzling fact that will receive further attention in Section 6.

As for Model 10 specifically, it uses the common correlated effects estimator instead of following the approach developed by Bai (2009). However, this method produces poor estimates for all country sub-samples. This is not surprising as a large number of common factors used as cross-sectional averages often reduce the significance in estimated coefficients. The only significant variables are bank profitability and investor profitability for Italy. Bank profitability is significant at the 5% level with a negative relationship with NPLs. Investor profitability is significant at the 5% level as well but with a negative relationship to NPLs.

5.4 Robustness

To check the robustness of our results, we use a shorter sample between 2010-2018, as such excluding the global financial crisis in 2008. We run the panel regressions and the factor model regressions with both principal component analysis and common correlated effects method for the individual countries. The estimated results are shown in Table A6 and A7 in Appendix VI.

6. Discussion

In this section we interpret and discuss the empirical results displayed in Table 5 in order to assess the determinants of non-performing loans for our sample of Southern European banks for the period 2006-2018. In addition, we discuss potential criticism of our econometric model and general drawbacks.

The macroeconomic variables unemployment, inflation and current account are significant for some of the model specifications. Unemployment is found to be significant and positive across Model 7, 8 and 9, which indicates that an increase in unemployment increases non-performing loans. This finding is in line with macroeconomic theory and confirms our earlier stated hypothesis. A rise in unemployment results in debt servicing problems for borrowers. Our results highlight a positive relationship between unemployment and non-performing loans which corroborates the findings of earlier studies (Louzis, 2012; Škarica, 2014; Ghosh, 2015). Similarly, inflation is found to be significant and positive for Model 9, suggesting that an increase in inflation increases non-performing loans. This result is in line with the earlier stated hypothesis and previous findings (Louzis et al., 2012; Baselga-Pascual et al., 2015; Ghosh, 2015).

However, the positive and significant sign of the current account coefficient across Model 7, 8 and 9 is surprising. These results imply that an increase in the current account balance increases non-performing loans. Intuitively, an increase in the value of a country's balance of trade generally implies an improvement in its' economic health. Therefore, non-performing loans are expected to decrease as a result of such an improvement. However, our results suggest that we have a positive relationship and cannot find support for our earlier stated hypothesis suggesting a negative relationship. Similarly, our findings undermine the negative relationship found in a previous study conducted by Kauko (2012). A possible cause for this finding might be outliers in the sample, impacting the results. Additionally, the effect current account has on the economic environment might be subject to lagging behaviour. Furthermore, the United States has been running a current account deficit since the early 1990s and has still managed to generate persistent economic growth. Thus, using the current account balance as a determinant for non-performing loans and the general economic environment might be inadequate.

The bank-specific variables asset structure, capitalization and bank profitability are significant for some countries and model specifications. The positive sign of the asset structure coefficient for Spain in Model 8 and 9 implies that an increase in total loans to assets increases non-performing loans. These findings are intuitive and are in line with previous findings (Baselga-Pascual et al., 2015; Ghosh 2015). Conversely, the negative sign of the asset structure coefficient for Portugal implies that an increase in total loans to assets decreases non-performing loans. However, this goes against earlier findings and is unintuitive from a bank perspective. An increase in total loans in relation to assets means a higher loan stock which naturally affects the amount of non-performing loans. In addition, more loans over safer assets introduce more liquidity risk. Therefore, it seems puzzling that we have a negative asset structure variable coefficient for Portugal. We attribute this sign to the small number of Portuguese banks in the sample (only two). In addition, the positive and significant sign of the asset structure coefficient for the Greek sample for Model 7 and 9 further contribute to the initial hypothesis which stated a positive relationship between asset structure and NPLs.

The positive sign of the capitalization coefficient suggests that an increase in common equity over total assets increases non-performing loans. As stated earlier, the previous literature is ambiguous regarding the impact of capitalization on non-performing loans. Again, the motivation for bank capitalization is regulatory, ensuring financial soundness. Nevertheless, the positive relationship for our sample seems to suggest the opposite, that capital requirements force banks into riskier investment strategies. This is a potential explanation which is in line with previous research (Koehn and Santomero, 1980; Blum, 1999). One interesting point to note is the capitalization coefficient for Portugal and Greece. They are both positive and significant but differs in magnitude. The reasons for these differences are hard to discern but might be due to country and bank-specific differences or differences in data.

The negative sign of the bank profitability coefficient indicates that as profitability increases, non-performing loans shrinks. These results are in line with previous findings and are sensible from a bank-specific perspective. Higher bank profitability will have the potential to limit the number of non-performing loans in several ways, ranging from lower probabilities of future financial distress to more rigorous screening and credit evaluations (Poghosyan and Čihak, 2012; Louzis et al., 2012). Another interesting point to note is the similarities of bank profitability coefficient for Spain and Portugal. Again, the sign of the coefficients is the same but the magnitude differs, possibly due to differences mentioned in the section above.

After scrutinizing the results obtained in this paper one can conclude that only a hand-full of our results are significant. This may not be surprising when comparing to the previous literature, where some papers conclude that it remains a challenge to find significance among bank-specific variables (cf. Louzis et al., 2012; Klein, 2013). This could offer some explanation to why much of our results remain insignificant. Moreover, the quality of our data and the differences in the classification of NPLs across countries might offer additional explanation. Another potential explanation might be found in the fact that this paper uses information on a national level instead of a regional level. Ghosh (2015) was able to retrieve an abundance of significant results in using information from state level. Allowing for a larger dataset could have potentially increased the significance and robustness of our results. Lastly, even though we are of the firm belief that our panel data model with interactive effects is well suited for modelling the determinants of non-performing loans as it deals with cross-sectional dependence, we cannot fully rule out potential misspecification of the model.

As mentioned in section 4.3, our bank-specific variables might suffer from weak exogeneity. Even though the model proposed by Bai (2009) accounts for cross-sectional dependence, it fails to provide a solution for possible weak exogeneity. Hence, this is the main weakness of the model employed. Since the panel data model with interactive effects does not consider possible endogeneity, one must proceed with caution when interpreting the estimated results since the issue of endogeneity could potentially invalidate them. Unfortunately, to the best of our knowledge, there does not exist a suitable econometric model that accounts for unobserved heterogeneity, endogeneity and cross-sectional dependence in a panel data setting. This leads to a predicament, forcing researchers to consider the tradeoff between weak exogeneity and cross-sectional dependence.

The previous literature on non-performing loans has favored econometric approaches accounting for endogeneity, as such there exists a gap in the literature. Researchers have failed to discuss the potential biases arising from cross-sectional dependence. Given the argued weak exogeneity and the important implications of cross-sectional dependence, this paper has attempted to fill this gap by accounting for said cross-sectional bias. Moreover, the interactive effects estimator with factors used in this paper allow the model to capture additional unobservable information that otherwise would remain unexplained. As such, this paper primarily attempts to solve the problem with cross-sectional dependence in the existing literature by utilizing the estimation approach of the panel data model with interactive effects developed by Bai (2009).

7. Concluding remarks

Learning more about the determinants of non-performing loans is relevant as the European Union has confronted significant challenges related to a weakened economic environment, increased credit risk and reduced profitability. This paper empirically analyzed the determinants of non-performing loans in Portugal, Italy, Greece and Spain from 2006 to 2018 using a balanced panel data set. The selected time span enabled us to consider the impact of the recent global financial crisis in 2008 and the European debt crisis in 2010–2011.

Our results indicate that a higher level of unemployment and current account has a positive impact on non-performing loans. These results remain robust throughout different model specifications. The fact that the current account has a positive impact on non-performing loans is surprising and goes against our earlier stated hypotheses. In investigating the bank-specific variables, this paper finds some evidence suggesting that bank profitability and capitalization has a negative relationship with non-performing loans. The results found in this paper varies among bank-specific variables and countries, researchers should therefore take caution not to extrapolate too widely the conclusions drawn here.

Our main contribution lies in treating the bias created by cross-sectional dependence using a panel data model with interactive effects, as suggested by Bai (2009). Previous researchers in this field have neglected the treatment of this important property which leads to inference problems. In doing so, we conclude there does not exist a suitable econometric model that accounts for endogeneity, unobserved heterogeneity, and cross-sectional dependence. As such, we have attempted to fill the gap in the literature by accounting for the latter two properties.

For future research it would be interesting to analyze financial stability by using forward-looking empirical models such as early warning systems and include non-performing loans as a bank-specific explanatory variable. Another potential approach to analyze credit risk specifically would be to use financial risk management measures such as value at risk (VaR) or expected shortfall (ES). Lastly, an introduction of an econometric model that is able to treat unobserved heterogeneity, cross-sectional dependence and endogeneity could contribute to beneficial developments in the non-performing loans literature.

References

Andrews, D. (2005) Cross-section regression with common shocks, *Econometrica*, 73, 1551–1585.

Arellano, M. and Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations, *Review of Economic Studies*, 58(2), 277–297.

Baboucek, I. and Jancar, M. (2005). A VAR analysis of the effects to macroeconomic shocks to the quality of the aggregate loan portfolio of the Czech-banking sector. Working Paper 1, Czech National Bank.

Bai, J. (2009). Panel Data Models with Interactive Fixed Effects, *Econometrica*, 77, 1229–1279.

Baselga-Pascual, L., Trujillo-Poncem, A. and Cardone-Riportella, C. (2015). Factors influencing bank risk in Europe: Evidence from the financial crisis, *North American Journal of Economics and Finance*, 34 (1), 138–166.

Beck, R., Jakubík P. and Piloiu. A. (2013). Non-performing loans: What matters in addition to the economic cycle? Working Paper No. 1515, European Central Bank.

Blasko, M., and Sinkey, J. (2006). Bank asset structure, real-estate lending, and risk-taking. *Quarterly Review of Economics and Finance*, 46, 53–81.

Blum, J. (1999). Do capital adequacy requirements reduce risks in banking? *Journal of Banking and Finance*, 23, 755–771.

Blundell, R., and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models, *Journal of Econometrics,* 87(1), 115–143.

Bofondi, M., and Ropele, T., (2011). Macroeconomic Determinants of Bad Loans: Evidence from Italian Banks, Occasional Paper No. 89, Bank of Italy.

Borio, C., and Lowe, P., (2002). Assessing the risk of banking crises, BIS Quarterly Review, 43–54.

Dinno, A. (2009). Implementing Horn's parallel analysis for principal component analysis and factor analysis, *The Stata Journal*, 9(2), 291–298.

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

Furlong, F. and Keeley, M. (1989). Capital Regulation and Bank Risk-Taking: A Note. *Journal of Banking and Finance* 13(6): 883–891.

García-Marco, T., & Robles-Fernández, D. (2008). Risk taking behavior and ownership in the banking industry: The Spanish evidence, *Journal of Economics and Business*, 60(4), 332–354.

Gennotte, G. and Pyle, D. (1991). Capital controls and bank risk, *Journal of Banking & Finance*, 15(5), 805–824.

Ghosh A. (2015). Banking-industry specific and regional economic determinants of non-performing loans: Evidence from US states, *Journal of Financial Stability*, 20 (1), 93–104.

Glorfeld, L. (1995). An Improvement on Horn's Parallel Analysis Methodology for Selecting the Correct Number of Factors to Retain, *Educaitional and Psychological Measurment*. 55 (3), 377–393.

Hayakawa, K. (2007). Small sample bias properties of the system GMM estimator in dynamic panel data models, *Economic Letters,* 95(1), 32–38.

Hayton, J., Allen, D. and Scarpello, V. (2004). Factor Retention Decisions in Exploratory Factor Analysis: a Tutorial on Parallel Analysis, *Organizational Research Methods*. 7(2), 191–205.

Hoggarth, G., Sorensen, S., and Zicchino, L. (2005). Stress tests of UK banks using a VAR approach, Working Paper 282, Bank of England.

Judson, R, and Owen, L. (1999). Estimating dynamic panel data models: a guide for macroeconomists, *Economics Letters,* 65(1), 9–15.

Kauko, K., (2012). External deficits and non-performing loans in the recent financial crisis, *Economics Letters*, 115(2), 196–199.

Klein, N., (2013). Non-Performing Loans in CESEE: Determinants and Impact on Macroeconomic Performance, IMF Working Paper, WP/13/72.

Koehn, M., and Santomero, A. (1980). Regulation of bank capital and portfolio risk, *Journal of Finance*, 43 (4), 1235–1244.

Furlong, F. and Keeley, M. (1989). Capital regulation and bank risk-taking: A note, *Journal of Banking & Finance*, 13(6), 883–891.

Laeven, L., and Valencia, F. (2012). Systemic Banking Crises Database: An Update, IMF Working Paper No. 12/163, International Monetary Fund.

Louzis, D., Vouldis, A. and Metaxas, V., (2012). Macroeconomic and Bank-Specific Determinants of Non-Performing Loans in Greece: A Comparative Study of Mortgage, Business and Consumer Loan Portfolios, *Journal of Banking and Finance*, 36(4), 1012–1027.

Makri, V., Tsagkanos, A., and Bellas, A., (2014). Determinants of non-performing loans: the case of eurozone. *Panoeconomicus*, 2(1), 193–206.

Mehran H, and Mollineaux L. (2012). Corporate governance of financial institutions, *Annual Review of Financial Economics*, 4(1), 215–232.

Männasoo, K., and Mayes, D. G. (2009). Explaining bank distress in Eastern European transition economies, *Journal of Banking and Finance*, 33, 244–253.

Nkusu, M., (2011). Non-performing loans and Macrofinancial Vulnerabilities in Advanced Economies, IMF Working Papers, WP/11/161, International Monetary Fund.

Pesaran, M. (2006). Estimation and inference in large heterogeneous panels with multifactor error structure, *Econometrica,* 74, 967–1012.

Pesaran, M. (2015) Testing Weak Cross-Sectional Dependence in Large Panels, *Econometric Reviews*. 34(6), 1089–1117.

Poghosyan, T., and Čihak, M. (2011). Determinants of bank distress in Europe: Evidence from a new data set. *Journal of Financial Services Research*, 40, 163–184.

Rajan, R., and Dhal, S. (2003). Non-performing loans and terms of credit of public sector banks in India: an empirical assessment, Occasional Paper 24, Reserve Bank of India, 81–121.

Reinhart, C. and Rogoff, K (2010). Growth in a Time of Debt, *American Economic Review: Papers & Proceedings,* 100(1), 573–578.

Salas, V., and Saurina, J. (2002). Credit risk in two institutional regimes: Spanish commercial and savings banks. *Journal of Financial Services Research,* 22, 203–224.

Santos, J. (1999). Bank Capital and Equity Investment Regulations, *Journal of Banking and Finance*, 23, 1095–1120.

Sarafidis, V. and Robertson, D. (2009). On the impact of error cross-sectional dependence in short dynamic panel estimation, *The Econometrics Journal*, 12(2), 62–81.

Škarica, B. (2014). Determinants of non-performing loans in Central and Eastern European countries, *Finance Theory and Practice*, 38(1), 37–59.

Totty, E. (2017). The Effect of Minimum Wages on Employment: A Factor Model Approach, *Economic Inquiry*, 55(4), 1712–1737.

Trujillo-Ponce, A. (2013). What determines the profitability of banks? Evidence from Spain, *Accounting and Finance*, 53, 561–586.

Appendices

Appendix I: Overview of the Literature

Table A1: Summary of the most relevant previous literature

Authors	Dependent	Explanatory	Data	Method	Key Findings
Espinoza and Prasad (2010)	NPLs & Z-Score	Macro, region and bank-specific	80 banks from GCC-region, 1995 – 2008	Fixed Effects, Difference- & System GMM	The authors found an inverse relationship between NPLs and economic growth. Further, the NPL ratio rises when interest rates and volatility (VIX) increases. The authors conclude that it is vital for policymakers to account for macroeconomic conditions when trying to identify potential shocks.
Bofondi and Ro-pele (2011)	NPLs	Macroeconomic	Italy, 1990 – 2010	Panel OLS	Household loans has a positive relationship with NPLs, interest rates, unemployment and a negative relationship with real estate prices and GDP growth.
Kauko (2012)	NPLs	Macroeconomic	34 countries, 2002 - 2006	Panel OLS	Current account deficits and credit growth predict the progress of NPL. Credit growth did not contribute to issues unless it was combined with a current account deficit.
Nkusu (2011)	NPLs	Macroeconomic	26 Countries, 1998–2009	Fixed Effects, Panel VAR	Increased levels of NPLs leads to long-run effects, hampering macroeconomic development.
Louzis et al. (2012)	NPLs	Macro and bank-specific	9 banks, Greece, 2003–2009	Restricted GMM	Performance and inefficiency are leading indicators. GDP has a negative relationship with NPLs, while unemployment and interest rates have a positive relationship. Finding robust and significant estimators for bank-specific variables proved challenging.
Beck et al. (2013)	NPLs	Macroeconomic	75 Countries, 2000 –2010	Fixed Effects & Two-Step Difference GMM	Exchange rate and interest rates have a positive effect on NPLs. GDP growth and stock prices have a negative effect on NPLs. Bank size has a magnifying effect.
Škarica (2013)	NPLs	Macroeconomic	7 EU-countries, 2007–2012	Fixed Effects	The author finds that the primary cause for high levels of NPLs is GDP growth, unemployment and the inflation rate.
Klein (2013)	NPLs	Macro and bank-specific	10 banks from 16 EU countries, 1998–2011	Fixed Effects, Difference- & System GMM	Higher levels of profitability results in lower NPLs. Increased lending is positively related to NPLs. Inflation, unemployment, volatility index and depreciation result in lower NPLs.
Baselga-Pascal et al. (2015)	NPLs & Z-Score	Macro and bank-specific	204 banks from 14 EU countries, 2001 - 2012	Two-Step System GMM	Profitability, efficiency, capitalization, and liquidity are significantly and inversely related to NPLs. Higher inflation, lower interest rates, and a falling GDP increase NPLs.
Ghosh (2015)	NPLs	Macro, state and bank-specific	50 states in USA, 1984 – 2013	Fixed Effects & System GMM	The author's estimates show that greater capitalization and cost inefficiency increase NPLs, while greater bank profitability lowers NPLs.

Appendix II: List of banks used

Table A2: Summary of all banks used in this paper, sorted by country.

Country	Bank
Greece	Alpha Bank Attica Bank Eurobank Ergasias SA National Bank of Greece Piraeus Bank
Italy	Banca Monte dei Paschi di Siena Banco BPM BPER Banca Credito Emiliano Intesa Sanpaolo Mediobanca UBI Banca UniCredit SPA
Portugal	Banco BPI Banco Comercial Portugues
Spain	Banco de Sabadell Banco Santander. S.A. Bankinter BBVA

Total: 19 Banks

Appendix III: Cross-sectional dependence (CD) test

Table A3: Pesaran (2004, 2015) test for cross-section dependence in panel time-series data.

	Pesaran (2004)		Pesaran (2015)	
	p–value	*t*–statistic	*p*–value	*t*–statistic
Non-Performing Loans	0.000	79.215	0.000	89.973
Asset Structure	0.000	9.190	0.000	93.802
Capitalization	0.000	16.574	0.000	89.020
Diversification	0.000	14.708	0.000	80.670
Bank Profitability	0.000	49.919	0.000	37.477
Investor Profitability	0.000	53.818	0.000	39.574
Gross Domestic Product	0.000	71.673	0.000	67.443
Unemployment	0.000	80.185	0.000	92.836
Inflation	0.000	79.767	0.000	84.949
Current Account	0.000	18.643	0.000	5.178
Interest Rate	0.000	94.297	0.000	94.297

Note: Under the null hypothesis of cross-section independence, t-stat $\sim N(0,1)$

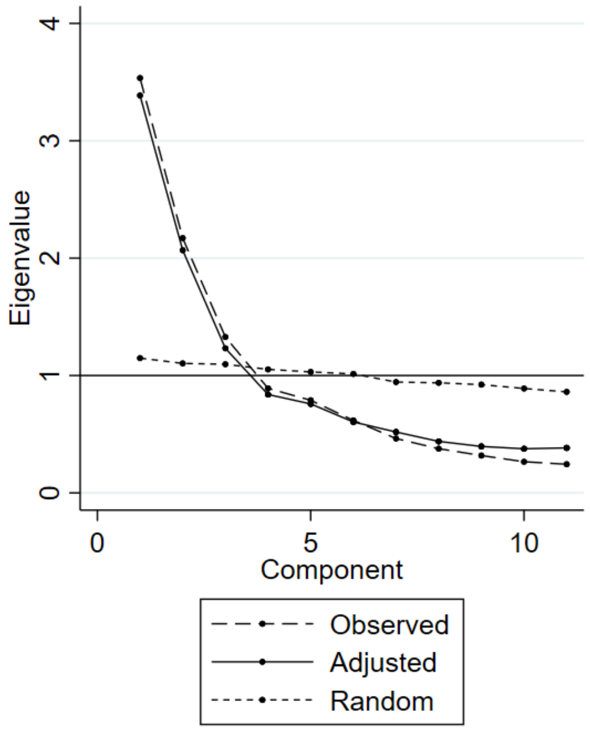
Appendix IV – Determining factors

Table A4: Results of Horn's Parallel Analysis for principal components

Components (Factors)	Adjusted Eigenvalue	Unadjusted Eigenvalue	Estimated Bias
1	3.319	3.535	0.216
2	1.992	2.171	0.179
3	1.222	1.328	0.106
4	0.851	0.891	0.040
5	0.772	0.789	0.017
6	0.626	0.618	-0.007
7	0.485	0.463	-0.021
8	0.458	0.377	-0.081
9	0.443	0.319	-0.125
10	0.408	0.265	-0.142
11	0.425	0.244	-0.181

Note: 10000 iterations, using the p95 estimate

Figure A1: Shows the adjusted, unadjusted and random eigenvalues



Appendix V – Panel data model with Interactive Effects

Table A5: Results for the panel data model with interactive effects, pooling over banks for all countries

	3 Factors			4 Factors		
	(1)	(2)	(3)	(4)	(5)	(6)
Asset Structure	0.013	0.468	0.785**	0.510**	-1.093***	0.320
	(0.269)	(0.339)	(0.371)	(0.254)	(0.267)	(0.279)
Capitalization	2.866**	-0.574	-1.287	1.573	0.781	0.051
	(1.117)	(1.148)	(1.069)	(1.352)	(0.988)	(0.817)
Diversification	0.035	0.003	0.040	0.078	0.074	0.134
	(0.103)	(0.119)	(0.128)	(0.104)	(0.103)	(0.123)
Bank Profitability	-1.416	0.814	0.486	-1.602	-0.506	0.030
	(1.791)	(1.480)	(1.359)	(1.736)	(1.561)	(1.374)
Investor Profitability	0.118	-0.208**	-0.177*	0.012	-0.038	-0.099
	(0.115)	(0.096)	(0.093)	(0.102)	(0.099)	(0.098)
GDP	-2.262***	-0.550	-0.273	-1.494*	-0.537	-0.700
	(0.778)	(0.627)	(0.672)	(0.771)	(0.583)	(0.668)
Unemployment	3.044***	10.943***	13.167***	5.957***	8.370***	13.405***
	(0.494)	(0.504)	(0.577)	(0.610)	(0.797)	(0.717)
Inflation	0.495	-0.312	8.246***	-1.595	-1.772	7.900***
	(1.356)	(1.470)	(1.936)	(1.612)	(1.161)	(1.957)
Current Account	5.486***	8.286***	8.668***	7.786***	-0.487	8.283***
	(0.396)	(0.928)	(0.834)	(0.669)	(1.143)	(0.827)
Interest Rate	-12.327***	31.905***	10.210	12.301	1.428	3.142
	(3.166)	(11.219)	(6.911)	(7.553)	(5.341)	(5.512)
Fixed Effects	No	B	B & T	No	B	B & T
N	988	988	988	988	988	988
RMSE	0.246	0.220	0.217	0.212	0.192	0.188

*Note: * p < 0.10, ** p < .05, *** p < .01.* (Robust Std. Errors).

Appendix VI – Robustness Tests

Table A6: Robustness results for all models, sample period 2010 – 2018.

	XTREG	FACTOR MODEL		CCE
		3 Factors	4 Factors	
Asset Structure.	0.747	0.362	0.193	1.298
	(1.083)	(0.241)	(0.258)	(1.830)
Capitalization	2.618***	0.371	-0.498	-2.447
	(0.942)	(0.925)	(0.754)	(3.750)
Diversification	0.032	0.124	0.008	0.161
	(0.184)	(0.087)	(0.079)	(0.319)
Bank Profitability	-0.687	-0.014	1.328	87.764
	(1.958)	(1.402)	(0.951)	(133.695)
Investor Profitability	0.005	-0.062	-0.014	-2.566
	(0.122)	(0.091)	(0.075)	(10.623)
Gross Domestic Product	2.105	-0.946	-0.053	6.443*
	(1.371)	(0.685)	(0.804)	(3.797)
Unemployment	5.304***	6.948***	6.020***	2.514
	(1.744)	(0.881)	(1.004)	(2.944)
Inflation	-0.101	0.624	-0.594	-3.847
	(2.962)	(1.635)	(1.772)	(4.296)
Current Account	1.895	4.863***	7.234***	7.847
	(1.452)	(0.818)	(0.967)	(4.768)
Interest Rate	-54.208***	-25.122***	-27.612***	24.036
	(15.658)	(6.122)	(6.311)	(24.023)
Fixed Effects	B&T	B&T	B&T	B&T
N	684	684	684	684
Adj. R^2	0.855			-1.061
RMSE	0.382	0.129	0.112	0.190

*Note: * $p < 0.10$, ** $p < .05$, *** $p < .01$.* (Wild-Cluster Bootstrap std. errors).

Table A7: Robustness results for all models, individually, sample period 2010 – 2018.

Country	Variable	XTREG	Bai (2009) 3 Factors	Bai (2009) 4 Factors	CCE
		(12)	(13)	(14)	(15)
SPAIN	Asset Structure.	4.186*	0.666	-3.271	0.234
		(2.236)	(0.491)	(2.687)	(0.284)
	Capitalization	-13.523	-4.722*	0.194	-1.450
		(16.162)	(2.610)	(0.230)	(0.873)
	Diversification	0.343	0.148	-65.498***	0.003
		(0.967)	(0.255)	(24.411)	(0.037)
	Bank Profitability	-88.500	-72.728***	-3.833**	-11.411
		(174.667)	(24.440)	(1.344)	(11.069)
	Investor Profitability	4.930	4.297***	-1.648**	0.472
		(9.191)	(1.289)	(0.748)	(0.392)
PORTUGAL	Asset Structure.	-0.074	-1.069	-6.764**	-0.180
		(2.203)	(0.727)	(2.677)	(0.144)
	Capitalization	4.753	6.105***	-0.078	1.078
		(4.818)	(1.924)	(0.224)	(0.751)
	Diversification	0.048	-0.129	-16.171*	-0.027
		(0.522)	(0.205)	(8.556)	(0.065)
	Bank Profitability	7.134	-20.018***	0.460	0.382
		(22.530)	(7.309)	(0.292)	(1.406)
	Investor Profitability	-0.404	0.530**	0.879*	0.036
		(0.755)	(0.261)	(0.478)	(0.112)
GREECE	Asset Structure.	-0.041	0.772*	-0.454	0.130
		(2.050)	(0.461)	(0.879)	(0.171)
	Capitalization	2.186*	-0.044	0.308	0.283
		(1.181)	(0.971)	(0.200)	(0.315)
	Diversification	0.098	0.364	1.462	-0.061
		(0.507)	(0.249)	(1.022)	(0.120)
	Bank Profitability	-0.496	0.176	-0.014	0.515
		(2.522)	(1.457)	(0.090)	(0.509)
	Investor Profitability	0.138	-0.069	-1.304***	0.044
		(0.188)	(0.101)	(0.409)	(0.099)
ITALY	Asset Structure.	2.109	-1.153**	2.474	0.517
		(1.417)	(0.525)	(1.574)	(0.620)
	Capitalization	-0.886	8.367***	-0.095	0.367
		(4.789)	(2.059)	(0.093)	(1.175)
	Diversification	-0.318	0.019	2.910	0.051
		(0.221)	(0.111)	(3.017)	(0.080)
	Bank Profitability	-12.876	4.809	-0.174	-6.646
		(31.102)	(4.766)	(0.188)	(19.772)
	Investor Profitability	0.268	-0.620**	-0.165	0.998
		(2.233)	(0.298)	(0.805)	(2.034)
ALL	Gross Domestic Product	1.856*	-0.456	5.769***	-2.469
		(1.039)	(0.672)	(0.864)	(1.555)
	Unemployment	5.124**	6.333***	0.418	-1.697
		(2.119)	(0.799)	(1.805)	(1.874)
	Inflation	-2.138	1.440	7.225***	-2.959
		(3.114)	(1.647)	(0.904)	(2.169)
	Current Account	1.398	5.167***	4.871***	-2.536
		(1.378)	(0.730)	(0.781)	(4.202)
	Interest Rate	-65.857***	-58.141***	0.395	-1.335
		(19.753)	(16.111)	(0.438)	(7.320)
	Fixed Effects	B&T	B&T	B&T	B&T
	N	684	684	684	684
	Adj. R^2	0.864	N/A	N/A	0.745
	RMSE	0.375	0.124	0.109	-1.321

*Note: * p < 0.10, ** p < .05, *** p < .01*. (Robust Std. Errors).