# Charles University in Prague

Faculty of Social Sciences
Institute of Economic Studies


MASTER THESIS

# Determinants of NPLs at the aggregate level: A comparative approach for middle and high income countries


Author: **Violeta Sandrovschi**

Supervisor: **PhDr. Ing. Petr Jakubík, Ph.D.**


Academic Year: **2013/2014**

# Declaration of Authorship

I hereby declare that I compiled this thesis independently, using only the listed resources and literature.

I grant to Charles University permission to reproduce and to distribute copies of this thesis document in whole or in part.

Prague, July 14, 2014

Signature

# Acknowledgments

I am grateful to my supervisor, PhDr. Ing. Petr Jakubík, Ph.D., for his professional recommendations and remarks, when consultations were imperative. I wish to extend my deep gratitude to Ph.Dr. Ing. Martin Janíčko Ph.D. and Mgr. Victoria Donu for their objective, useful criticisms and insightful suggestions for the present thesis. Furthermore, thanks go to Mgr. Nadzeya Laurentsyeva and Mgr. Adelina Hajzeraj for their inspiration, and guidance with technical tools, provided during this work.

I also acknowledge the financial support offered by the Czech Government to study in Prague and the infinite moral understanding by my family throughout my years of studies.

# Abstract

This thesis investigates the key determinants of the Non-performing loans (NPLs) comparing two groups of countries from Southeastern and Western Europe, with two different levels of economic development. We try to find empirical evidence and estimate whether the determinants of NPL ratio are different for the middle and high income countries. Applying panel data models for 14 countries overall, and using the regressions of subsampled countries, we analyze the importance of the determinants at the aggregate level. The final results show that all variables considered are significant, except inflation rate under all specifications and FDI when the subsampled dummy variables are used. As for the specifications of the exchange rate determinant, we conclude that the NPL ratio is negatively and significantly influenced in the export dominant middle income economies. An additional non-economic variable, such as the educational index, constructed at the national level, is found to increase the NPL ratio. Concerning the institutional quality index, averaging all six institutional indicators, this determinant does not show a consistent result across different data sample specifications.

| | |
|---|---|
| **JEL Classification** | G21, G28, F10, O43 |
| **Keywords** | Non-performing loans, subsamples, dynamic model, macroeconomic determinants, index |

| | |
|---|---|
| **Author's e-mail** | v.sandrovschi@gmail.com |
| **Supervisor's e-mail** | petrjakubik@seznam.cz |

# Contents

**LIST OF TABLES** ..... **VII**

**LIST OF FIGURES** ..... **VIII**

**ACRONYMS** ..... **IX**

**I. INTRODUCTION** ..... **1**

**II. IMPLICATIONS OF THE TOPIC** ..... **3**

2.1. DEFINITIONS AND REMARKS OF NPLS ..... 3
2.2. LITERATURE REVIEW ..... 5
2.3. NPL EVOLUTION AND ITS INTERCONNECTION WITH THE ECONOMIC CONDITIONS ..... 11
2.4. HYPOTHESES DEVELOPMENT ..... 18

**III. EMPIRICAL ANALYSIS** ..... **23**

3.1. DATA DESCRIPTION ..... 23
*3.1.1. Variables and data sources* ..... *23*
*3.1.2. Descriptive statistics* ..... *28*
*3.1.3. Stationarity and variable specifications* ..... *29*
*3.1.4. Stylized facts* ..... *33*
3.2. ECONOMETRIC MODELS ..... 34
*3.2.1. Pooled OLS* ..... *34*
*3.2.2. Multicollinearity* ..... *35*
*3.2.3. Robust standard errors* ..... *36*
*3.2.4. Structural break* ..... *36*
*3.2.5. Fixed Effect model vs. Random Effect model* ..... *37*
*3.2.6. Hausman test* ..... *39*
*3.2.7. Additional empirical tests* ..... *40*
*3.2.8. Components of the institutional quality index* ..... *41*
*3.2.9. Does the commercial trade determine exchange rate impact over NPLs?* ..... *42*
*3.2.10. Remarks of the final results using fixed effect estimation* ..... *43*
*3.2.11. Empirical literature overview* ..... *44*
*3.2.12. Dynamic GMM method* ..... *45*

**IV. ROBUSTNESS CHECKS** ..... **53**

**V. INTERPRETATION OF RESULTS** ..... **54**

**VI. CONCLUSION** ..... **59**

**BIBLIOGRAPHY** ..... **61**

**APPENDIX A: EMPIRICAL PART** ..... **72**

**APPENDIX B: ROBUSTNESS CHECKS** ..... **84**

**MASTER THESIS PROPOSAL** ..... **85**

# List of Tables

Table 2.1: Bank NPL ratio in the world (1998-2011) ..... 12
Table 3.1.1: Table 3.1.1: List of countries and division of the sample ..... 23
Table 3.1.2: List of economic and non-economic variables of the basic model ..... 27
Table 3.1.3: Stationary data ..... 30
Table 3.1.4: Unit root test results and transformations ..... 32
Table 3.2.1: Chow test results for structural break ..... 37
Table 3.2.2: Hausman test specifications ..... 39
Table 3.2.3: GMM estimation on full sample with dummy specifications ..... 50

Table A. 1: Descriptive statistics full sample ..... 72
Table A. 2: Descriptive statistics for groups of countries ..... 72
Table A. 3: Unit root presence for variables in level ..... 73
Table A. 4: Correlation matrix ..... 74
Table A. 5: VIF a), b), c), d) ..... 76
Table A. 6: Pooled OLS with robust standard errors (SE) ..... 77
Table A. 7: Preliminary results of Fixed Effect model ..... 78
Table A. 8: Additional empirical tests ..... 79
Table A. 9: Hausman tests ..... 79
Table A. 10: Export/import dominant countries ..... 80
Table A. 11: Estimation of the model including the extension of exchange rate hypothesis (Fixed Effect model) ..... 81
Table A. 12: Autocorrelation test and Sargan test for overidentification or valid instruments; Robust SE ..... 82
Table A. 13: System GMM results for inflation rate and FDI in middle and high income groups ..... 83
Table B. 1: Robustness checks results ..... 84

# List of Figures

Figure 2.2: NPL plot for the sampled countries ..... 14
Figure 2.3: Macroeconomics and banking sector ..... 18

Figure A. 1: Plots of unemployment rate and GDP growth ..... 74

# Acronyms

| | |
|---|---|
| **CAD** | Capital Adequacy Ratio |
| **CEE** | Central and Eastern European |
| **CIS** | Commonwealth of Independent States |
| **IMF** | International Monetary Fund |
| **IQ** | Institutional Quality |
| **IV** | Instrumental Variable |
| **EBRD** | European Bank for Reconstruction and Development |
| **EU** | European Union |
| **FDI** | Foreign Direct Investments |
| **FE** | Fixed Effects |
| **GCC** | Gulf Cooperation Council |
| **GDP** | Gross Domestic Product |
| **GIIPS** | Greece, Italy, Ireland, Portugal, and Spain |
| **GMM** | Generalized Method of Moments |
| **GNI** | Gross National Income |
| **HI** | High Income |
| **OLS** | Ordinary Least Square |
| **MI** | Middle Income |
| **NPLs** | Non-Performing Loans |
| **PLN** | Polish Zloty |
| **PRS ICRG** | Political Risk Services International Country Risk Guide |
| **RE** | Random Effects |
| **ROE** | Return on Equity |
| **UNDP** | United National Development Program |
| **USD** | United States Dollar |
| **WEO** | World Economic Outlook |
| **WB** | World Bank |

# I. Introduction

Since the banks' failures during the financial crisis, the predictions for credit default and prudential measures became a central topic and goal for a healthy financial environment. Many economists and regulators engage in developing more advanced and detailed macro stress tests which base their results also on the determinants or factors driving the quality of a bank's assets. Despite an immense attention paid to non-performing loans (NPLs) determinants and various studies that estimate their influence, their importance is not well-established yet. Most of the studies, based on the literature review in subsection 2.2, consider that pooling different countries from different regions of the world may explain in general the NPL ratio increase. This implies that similar financial tools can bring the same results among all of the selected economies. Others find empirical proofs that reactions of economic shocks to NPLs are different, only that the econometric techniques are employed separately for each country.

Depending on the purpose of a study and the targeted audience, criticisms arouse whether both the macroeconomic and microeconomic determinants are included in the model. In our work, the economic determinants at the national level for NPLs (at the aggregate level) only are considered, with an extension to non-economic explanatory factors. The determinants refer to those variables which control for the most important economic conditions (GDP, unemployment rate, inflation rate, FDI, exchange rate) and the non-economic factors (institutional quality, education level).

According to our literature research, a few empirical works concern most of the selected countries, especially the several selected emerging or developing markets caused by limited data availability. An interesting fact is to examine the determinants of NPLs on a comparative approach between the middle and high income countries within two regions from Europe (Western and Southeastern). This direction may provide an outlook over the most prominent determinants that can be influenced by the public policy makers especially in the middle income states. All the 14 states included represent the full sample, but we divide it into the subsample 1, or middle income (MI) group (Bosnia and Herzegovina, Bulgaria, Macedonia, Moldova,

Romania, Serbia and Ukraine), and the subsample 2, or high income (HI) group (Austria, Finland, France, Germany, Ireland, Portugal, United Kingdom).

With respect to previous works limited to this topic, the contributions of the thesis include the findings and construction of the educational index at the aggregate level to capture any impact of the human capital, followed by the results of the exchange rate hypothesis of the extended model, of the institutional quality index averaging its components, and the comparative approach of sampled countries with relatively extended time period (2002-2012).

Therefore, in section II, an introduction of the NPL definitions and the differences between types of loans as well as their treatment is developed. To build our own ideas and hypotheses, compared to previous works, a subsection for the literature background is provided to create an extension of the studies that have been done already. Considering that data of our selected countries must be well-behaved, information about each country's economic situation is investigated to provide the interconnections of the economic cycle and the NPL evolution. The last part of section II is the output of the mentioned subsections that is the hypotheses development and the motivations for choosing our variables. In section III, the thesis approaches the econometric techniques succeeded by data descriptions, variables and model issues and some stylized facts. These are followed by the main subsections of applying two methods (static and dynamic). Moreover, the robustness checks in section IV are employed to strengthen our final results. Then, we choose to interpret our results based on two-step system GMM of Arellano-Bond and Arellano-Bover/Blundell-Bond contrasting the fixed effect estimations in section V. In the last part, the interpretation of results and conclusion are emphasized.

# II. Implications of the topic

## 2.1. Definitions and remarks of NPLs

An indisputable fact in the banking system is that the risk of default and inability to meet the financial obligations aroused from the moment the banking system was created, along with an intensive economic progress. For the treatment of credit risk, new rules and regulations for banks (Basel Committee and the 3 Pillars) were launched eventually, after the financial or Asian currency crisis from 1997. For identification of factors that impact the credit risk, some researchers consider that NPL ratio is a suitable indicator for assessment of credit risk (Ahmad and Ariff, 2007). In the same line, Shingjergji (2013) uses the macroeconomic variables for assessment of the NPL ratio in Albania, while an evidence for Malaysian banks is based on the same indicator to measure the credit risk vulnerability (Janvisloo and Muhammad, 2013).

After an investigation for 100 states (Jose and Georgiou, 2008), the *NPL* variable seen by IMF as a financial soundness indicator for the asset quality, must bring together the signals for financial stability agreed by all countries' definitions of NPL.

Knowing the variety of the most prominent factors affecting NPLs, the analysis is very noticeable for macroeconomic stress testing scenarios, and several works use NPL to total loans for its estimation (Zeman and Jurca, 2008; I. Babouček and M. Jančar, 2005). When ECB (2013) issued the monthly bulletin, it was clarified that indicators for credit risk measurement may be designed by NPLs or loan loss reserves.

In literature and institutional documents, an international definition for NPLs, not difficult to apply for a comparison among groups of states, cannot be found because the interpretations vary from country to country including the European Union. EMF study (2010) on NPL in the European Union concluded that NPL elements distinguish stronger from each other across the available information within the investigated countries. The definitions vary globally in terms of a loan's due date settings for which is classified as doubtful or loss, but also in terms of the balance sheet items' corresponding. Recognition of this shortcoming was addressed by IMF

in 2005 where the *"Treatment of non-performing loans"* (Bolem and Freeman, 2005) defines a non-performing loan as: "when payments of interest and/or principal are past due by 90 days or more, or interest payments are equal to 90 days or more have been capitalized, refinanced or delayed by agreement or payments are less than 90 days overdue". Nevertheless, for simplicity in case of results interpretation, it is useful to follow the loan classification provided by Bloem and Gorter (2001):

- Standard – lending credit is safe, principal and interest payments are expected to occur regularly without difficulties under actual circumstances
- Watch – the loan repayment needs to be monitored more in case it remains uncorrected, otherwise a risk of non-fully repayment occurs
- Substandard – concern about the fully reimbursement is a cause of collateral mismatching; the value of the loan and/or principal and interest are not paid more than 90 days. In this case the risk of default or the risk to become an impaired loan is high
- Doubtful – bank's management determines the credits' full repayment because of its actual conditions and/or principal and interest are not repaid more than 180 days. These credits are not losses, but impaired assets
- Loss - loan cannot be repaid and/or principal and/or are not repaid more than 1 year.

In result, the last three types (substandard, doubtful and loss) are considered to raise concern and they can constitute elements of non-performing loans.

It is denoted by Barisitz (2013) in his studies to find more elements of adjusting the national definitions for a common comparison of selected Western European countries, besides the one which considers the due date of NPL for more than 90 days and the weaknesses of the debtor. Additionally, he refers to calculations of applicable items in case of a restructured loan, the performance of the total or partial value of the loan or the type of protection of the credit. Barisitz (2013) implies that from all the countries included in the sample from Western Europe, only 4 of them are consistent with the NPL definition, but some of them showing a downward bias meaning that the NPL ratio does not account for the total value of impaired loans, making the ratio lower than it is in reality. As a result, changes should be made at the country level towards the NPL definition for correcting the bias.

Furthermore, the uncertainty arises more from the standardized definitions implying that NPL definition is subject to a nation's own adjustments. The problem

can be justified due to a country's accounting standards, but at the same time it can bring contradictories as one state can reduce the period of 90 days for the reason to prevent the complete loss. Recently, Moody's Analytics provided the information for the asset quality (Balduini, 2013) which was reviewed by EBA (European Bank Authority) to introduce the definition of NPL as those loans "past due more than 90 days and/or unlikely to pay", meaning that all the on-balance sheet items (loans and debt securities) and some off-balance sheet items, except those related to trading, would be classified as a NPL. Along with the British Banker's Association opinion that does not consider the review is changing the definition much, two German banks addressed a document to the ECB, the German Bundesbank and Baffin, where they stated clearly the excessive burdens that may appear by changing the accounting calculations, according to Bloomberg (Groendahl, 2013).

Even if the disparities between NPL measurements are significant, the new trends for a common agreement to define NPL ratio is not concrete, therefore one should consider the dataset of countries based on the information available being consistent with the purpose of their studies.

## 2.2. Literature review

The dynamic situation in the banking and financial activities for a couple of years back generated highly concentrated works or research studies on explaining the non-performing loans - an essential preoccupation for Central Banks including regulators and supervisors. The literature that is based on exploring the determinants of NPL ratio is quite broad and beside the macroeconomic level determinants, it includes, as well, the factors specific to banks or factors at the microeconomic level. Many researches base their hypotheses and econometric models on the main macroeconomic factors: the real GDP growth, real interest rates and proxies for financial market development as stock market indices. Depending on which countries the authors intend to analyze the NPL ratio determinants, the results slightly or sometimes significantly differ from each other.

In most of the relevant economic estimations, it is broadly found a strong significance and indirect relationship between the GDP growth and the NPL ratio. Signals of a financial health during the financial turmoil are heterogeneous at the country level because many authors include in their analysis single country based

estimations (Khemraj, 2009; A. Belgrave, et. al, 2012) and less for a multi-country modelling. Shijaku and Ceca (2011) are estimating a fixed effect model for the credit risk in Albania concluding a substantial after effect of NPLs to shocks in GDP, given a 2-year financial world distress. Further, the same authors enlarge the model by examining new factors affecting the NPLs alike the change of regulation for credit risk in Albania, leading to findings of significant determinants of exchange rates and reference rates.

The recent investigation (Jakubík and Reininger, 2013), explaining factors that have a negative or positive contribution on NPLs in a sample of Central, Eastern and Southeastern European countries, take into consideration the macroeconomic variables for the economic activity as the stock index standing for the risk aversion of international investors towards the home country. Under the same study, the authors consider the aggregated credit (the private sectors' loans) to be included in the regression defined as credit aggregate to GDP ratio. Additionally, to seize for economic conditions, the authors introduce real exports and real domestic demand into their study as noticed by Beck et al. (2013). In case of the real exports, we transfer the idea that, if the domestic products are sold abroad then the home economy has a positive impact on investments, labor workforce and productivity. However, in case of the shrinkage of domestic demand, caused frequently by recession periods, the economy is said to point out to a decreasing trend leading to a higher probability of default for loan borrowers and, consequently lenders.

Jakubík and Reininger (2013) are estimating the correlation of the chosen dependent variables with the NPL ratio for Bulgaria, Croatia, Czech Republic, Hungary, Poland, Romania, Russia, Slovakia and Ukraine. The limited country sample is explained by their goal to bring to interested parties a benchmark for regions from developing Europe in case the analysis is done for other emerging states, and for the reason of available data. An interesting result suggested for a further research is established on the additional explanatory variable of exchange rate changes which is viewed to have a significant role for NPL ratio, in their case of currency depreciation in the total foreign denominated loans and risk of increasing interest rate for foreign currency loans. Also, it is attested the difference between credits borrowed in foreign currency and those in domestic currency, resulting in a higher magnitude on NPLs ratio, with the latter performing worse that can be explained by pegged or floating exchange rate regimes.

Heterogeneous effects are detected by using time series data analysis in Festic and Romih's work (2008), similar to many other studies as Fainstein and Novikov (2011), where they focus only on three countries: the Czech Republic, Slovakia and Slovenia. The results emphasize that individual features of an economy may not be followed up persistently by other states' progress. Festic and Romih (2008) found as the most acknowledged dependent variable – GDP growth causes decline of NPL ratio in case of the Czech Republic and Slovenia, while this hypothesis is rejected for Slovakia where GDP linkage to credit risk pursues the counter cyclicality. They include savings and inflation, as additional explanatory variables, providing empirical conclusions for the Czech Republic that the rise of unemployment rate slows down the NPL ratio and for Slovakia and Slovenia, the upturn of NPL growth is explained by acceleration of savings.

Linkage of the credit performance or the quality of the bank asset and the economic activity is therefore, by no reason a surprise and it is empirically proven. Loans are indispensable for financing individuals and businesses, which consists a reasonable argument to draw attention on the differences across a large sample of countries as Beck et al. (2013) did in their paper showing the main factors of NPLs during the bear financial markets. The differences among 80 countries are testified to show that part of the panel data have an unchanged NPL ratio, while for others the loans' performance recovered, given the recession of 2009. As a measure for underlining these differences, Beck et al. (2013) consider the results based on 4 criteria to represent the following situations jointly with the exampled countries: a) developed economies with floating exchange rate regime and bank financial system; b) developed economies with floating exchange rate regime and a capital financial system; c) developing economies with a fixed exchange rate regime, large foreign currency denominated loans with a stable exchange rate during 2008; d) developing economies with a large depreciation of domestic currency during crisis.

Setting the different exchange rate regimes during crisis across countries provides additional remarks that an economy (in this case Ukraine as the example given by Beck et al., 2013) has a much higher exposure to the NPL rise and credit default risk due to the mismatching characteristic that banks borrow in foreign currency but lend in local one. The mismatch appears when the depreciation cannot be avoided, but creditors are not able to repay back. In case of managing a fixed exchange rate during crisis the probability is also higher, experienced by previous financial crisis that the NPL will tend to boost up.

Having introduced the times of crisis for exploring the development of NPLs and their determinants, some papers reveal explanatory variables with specific crisis that mostly influence or is linked to the economy of a single state. For these variables some authors find that a proxy as the Greek crisis can assess the pressure on NPL ratio. Romania's banking system is witnessed to be affected by a transmission channel of the Greek crisis generating a negative impact on NPLs (Vogiazas, Nikoloidu, 2011). The motivation that is limited to the extent of Greek crisis, chosen as a specific variable, originates from the large presence of foreign-owned banks in this country, where 30.7% of them are Greek subsidiaries (Vogiazas, Nikoloidu, 2011). Consequently, the authors conclude that the spillover effect, transmitted through Greek debt, can consist a financial soundness indicator for a contagion effect.

Moreover, the research of Vogiazas and Nikoloidu (2011) is covering other explanatory evidence for NPLs assessment including into their model macroeconomic variables as the type of business (namely construction), investments, debt-to-GDP ratio and the monetary aggregate M2. The incentives are the loans provision survey of the National Bank from Romania that decides the risk of NPL upsurge is driven mainly by construction and real estate sectors. Overall, the results show significant influence of determinants on NPLs, beside the hypothesis that Romanian bank level factors, financial market situation, interest rates do not show any empirical prove that these variables have an explanatory power to the model.

Louzis et al. (2011) argues that NPLs must be analyzed separately on consumer loans, business loans and mortgages. The authors claim to include as additional variables - the bank-specific factors affecting NPLs, for the reason the NPLs change is insufficiently estimated as banks and financial institutions have own management abilities that can harm the asset quality (liquidity, CAD, ROE, etc.). In this way, the bank-specific factors are based on the hypotheses of bad management, monitoring loans, moral hazard, diversification (bank size), "Too Big to fail" and the quality of management. At the beginning, Louzis et al. (2011) highlighted that few studies combine determinants at microeconomic and macroeconomic level, but in the end their results count upon the explanatory power and significance for macroeconomic variables and management quality. Therefore, the hypotheses that "Too big to fail" doctrine increases the NPL ratio is accepted in general, but when considering the size of banks then the variable does not indicate any effect on NPLs. Another implication of their hypotheses resulted in a significant impact on NPLs shift, driven by the concentration of shareholders.

The research of Louzis et al. (2011) is highly appreciated due to the incorporations of a new approach for NPL diagnosis. Comparing to previous studies, the examined NPLs in their study, divided by their types, bring to light the intrinsic influence on the credit risk more than if one controls for the general non-divided NPL amount. In result, it is shown that for Greek banking system, there are differences among the variance of consumer loans, business loans and mortgages, where the consumer loans have the highest volatility. In the same manner, there is evidence for differences of macroeconomic variables towards the three types of loans. Nonetheless, the period used is rather short (2003-2009) and Greece may have specific features caused by the local crisis and can augment the estimators of the separated NPLs. Furthermore, on one side the results of bank-specific factors are consistent with the idea to generate more objective and correct estimators of explanatory variables after including them into the model, but on the other side Central Bankers would not have a very significant effect on how much the "Too Big to fail" doctrine can influence the quality of the bank assets. In the same manner the concentration of number of owners can be improved by regulators and in this way the vast power explaining the shift of NPLs is driven by the macroeconomic factors.

Indeed, in literature, the authors are concerned mostly either with macroeconomic variables explaining the development of the proxy for the probability of default – NPLs (Jakubík, Reininger, 2013), or with the firm level factors or bank characteristics and there are less studies engaged on both implications for the credit risk proxy.

Shijaku and Ceca (2011) are testing the forceful asymmetries in loan quality as a response to shocks in macroeconomic variables and bank characteristics inclusive, ascertaining that there are no different responses across idiosyncratic factors. Vogiazas and Nikoloidu (2011), supplemented by the Greek crisis collision, questioned why econometricians should ignore the other variables except macroeconomic factors. After they engage into hypothesis testing that growth of NPLs is shifted upward along with factors rooted in the bank level system, Vogiazas and Nikoloidu (2011) find that for Romania, the bank-specific factors do not explain the model well. In his research, Głogowski (2008) includes the debt burden variables (disposable income to GDP of households, loans to sales income, etc.) and finds to be insignificant to the model.

Concerning the impact on loan losses that also refers to the NPL definition, Głogowski (2008) is performing a panel data study for Polish bank's loan losses,

determined by their ties to the business cycle conditions. Despite the paper displays a significant impact of GDP, change in real interest rate, change in unemployment rate on NPL outcome, the author concludes controversial and unpredictable result. The influence of exchange rates is uncertain, even though the foreign currency loans assume a high percentage in total loans to households and is added to the model as the change of exchange rate weighted with the respective share. This research points out to the importance of providing the NPLs models for macro stress tests based on the macro scenarios which are employed by the author concluding that the loan losses are increased mainly by the oil prices scenarios.

Głogowski (2008) uses the explanatory variables lagged one quarter at least, to control for the gap between the classifications of loan losses, that is, the day of becoming non-performing and the day when the loan is considered a loss.

A comparable analysis regarding the macroeconomic determinants of NPLs or the credit risk in this case is to notice the differences for the most vulnerable countries in the EU that were recently affected by the burdened governments and tough austerity measures. A dynamic panel dataset for GIIPS countries is estimated by Castro (2012) where the results are consistent with the previous studies on the NPLs determinants: NPLs increase when GDP growth and the share price are lower, while the unemployment rate increases. Additional explanatory variables are the credit growth, and the interest rate (long-term) to account for the probability when the client has the power to repay the debt.

Conclusion that higher credit growth causes higher NPLs (Castro, 2012) is in accordance with the study of Espinoza and Prasad (2010), where beside the examination of macroeconomic determinants of the NPL ratio, a return (feedback) investigation is used on how NPLs determine economic growth, employing VAR model. For further research, this paper contributes with an empirical statement that some variables may become redundant to the model, and the provided example is the unemployment rate in the GCC countries (or pegged exchange rate regime). The reason of excluding some variables is argued by the fact that the respective indicators may be relatively stable and low, in this way simplifying the future works for a better model specification.

A multi country comparison of credit risk determinants exists for developed countries: Australia, France, Japan, US on one hand, and for the emerging: India, Korea, Malaysia, Mexico and Thailand on the other hand, but the model is limited to NPL microeconomic determinants investigated at the level of two types of banking

systems (Ahmad and Ariff, 2007). Concerning the differences across the two types of countries the paper concludes that credit risk in emerging states is higher than in the developed ones. The author finds that the quality of management is significant in banks where loans are the predominant products, and across the banking systems the CAD is significant for banks with a diversified segment of products (Ahmad and Ariff, 2007).

The assessment of NPLs determinants in two regions of countries (CFA and non-CFA) from Sub-Saharan Africa in '90s showed significant disparities between them with respect to lending, predominantly for three sectors of economies: manufacturing, commercial activities and service (Fofack, 2005). A new variable explaining the causality of NPLs is the interbank loans (which is attested by the author to be classified at the microeconomic level) influence on credit risk development found to have a strong causality, measured by Granger-causality analysis. The paper concludes the importance of microeconomic factors (net interest margins and interbank loans).

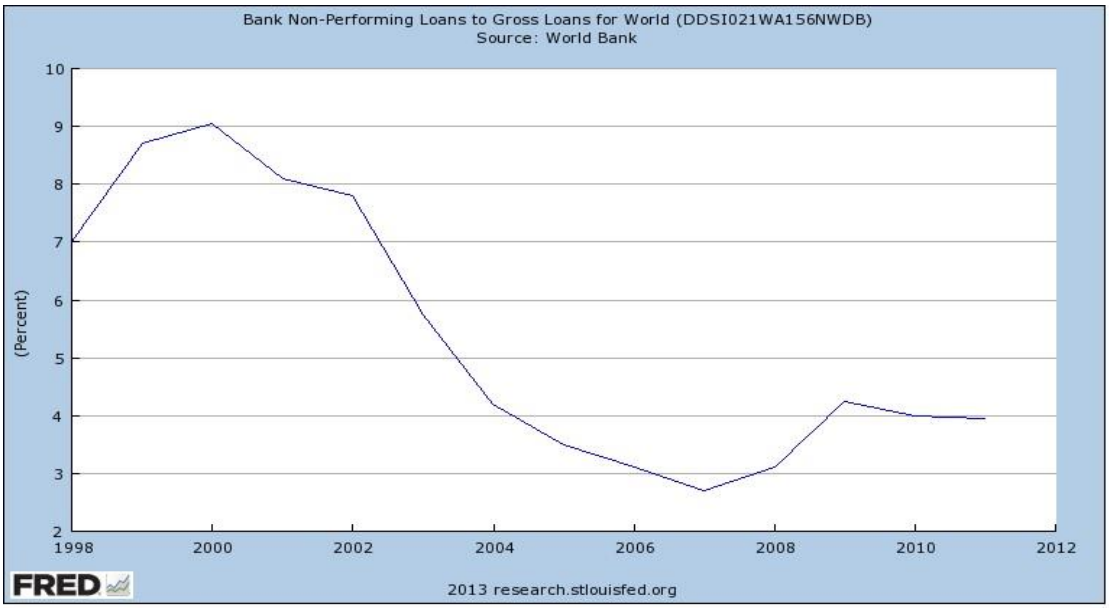
Fofack (2005) brought into attention the real exchange appreciation effect on the credit performance and assumes that this determinant of NPLs does not display a consistent estimator for one of the sub-sampled countries. Having concluded this, Fofack (2005) explains the inconsistency by observing the ambiguous sign and influence in the pre-crisis period and implies that is due to the monetary authority' regimes that have been anticipated. If Vogiazas and Nikoloidu (2011) have found a significant influence of M2 aggregate on NPLs ratio in Romania along with the Greek crisis, then Fofack (2005) ascertained an indirect relation - when the monetary indicator M2 is increasing, the NPL ratio has a decreasing trend.

## 2.3. NPL evolution and its interconnection with the economic conditions

The bank management faces the decision to allocate financial resources depending on clients' current condition which boils down to an acceptable risk if the prove of repayment is solid. However, predictions of a credit default or non-performing loan may not be accurate due to systematic risk which is influenced by the macroeconomic factors: unemployment rate, change of real GDP, stock prices, inflation rate, exchange rate, monetary policy, etc. (Castro, 2012). In times of crisis,

the banks' asset quality decreases, ascending the non-performing credit amount. The open markets, an indispensable world characteristic, assumes the economic development and systematic risk cause the countries, depended on foreign exports / imports, to respond in case of shocks. During the booming period, investors become more risk appetite increasing the demand for credits and the recession time implies more expensive loans, transforming gradually to NPLs. In Figure 2.1, we can notice the trend of NPL ratio in the world from 1998-2012. From 2002 the NPLs decreased due to the booming period until the US subprime mortgage crisis emerged and the amount of losses for banks increased. According to Beck et al. (2013) bank asset quality progressed in the emerging countries until 2008, while in 2009 the quality in these states did not depreciated (20% NPL ratio) as much as in the advanced economies, where NPL ratio reached even 60%.

**Table 2. 1: Bank NPL ratio in the world (1998-2011)**




*Source:* World Bank; Fred Economic Data

First of all we must analyze the overall recent evolution of the European banking system for the reason the work will be concentrated on European countries. After the US subprime sector crisis, domestic banks had to face diminished sources of foreign reserves for lending due to lessening of foreign flows to EBRD region in 2008-2009 (Haas and Knobloch, 2010). According to them, a balanced path of NPL was registered in Central European states and a growing trend of this ratio in Kazakhstan, Latvia, Mongolia, and Russia. Kazakhstan experienced a NPL level that rose significantly from 5% in 2008 to a level of almost 35% in 2009, September.

Even if the growth of the economy is around 9% per year (averaged), China is the country with the highest NPL ratio in the world.

Beginning with the increasing demand for credits, as a result until 2010, loans in foreign currency were high in Croatia (74.3%), Hungary (66.1%) and Romania (63%), especially. From 2010 to 2011, Spain, Ireland and Italy registered a high significant NPL ratio. Poland experienced an increasing NPL ratio in 2009 after a depreciation of PLN currency in 2008.

During 2011 the evolution of economic environment started to gradually slow down with EU unemployment rate slightly above 10 % and with a negative growth of GDP in some states. The exporting companies reduced their supply to Brazil, Russia, India, and China. Consequently, the EU governments had a 4.5% deficit.

With respect to previous year, bank assets had a 4.4% growth at the end of 2011 while loans and deposits increased, by 3.7% and 4.3% respectively. In the same period, Finland's financial assets grew by 33.7% and at the same time Ireland recorded a loss of bank assets of 14%. Other European countries had a negative decreasing trend in bank assets: Hungary (-8.7%), Greece (-74%), Estonia (-6.6%), Lithuania (-3.8%).

Referring to credit growth, it is observable that loans in the Euro area raised by 4% in 2011, in contrast to an increase of only 2% in the EU countries not included in the common currency area. In EBF report (Proskurovska, 2012) it is thought that the loan growth in the EU results from the interbank loans, amounting to 984 trillion EUR. The level of NPL ratio in 2011 was 6% in the EU and 5.6% in the Euro area, being a high but stable level, confirmed by IMF (EBF, 2012). Ireland and Lithuania had recorded 16.1% and 16.3%, to be the most increased NPL ratio. On the other side, the lowest NPL amount to total loans was registered by Luxemburg, Finland and Sweden with less than 1%.

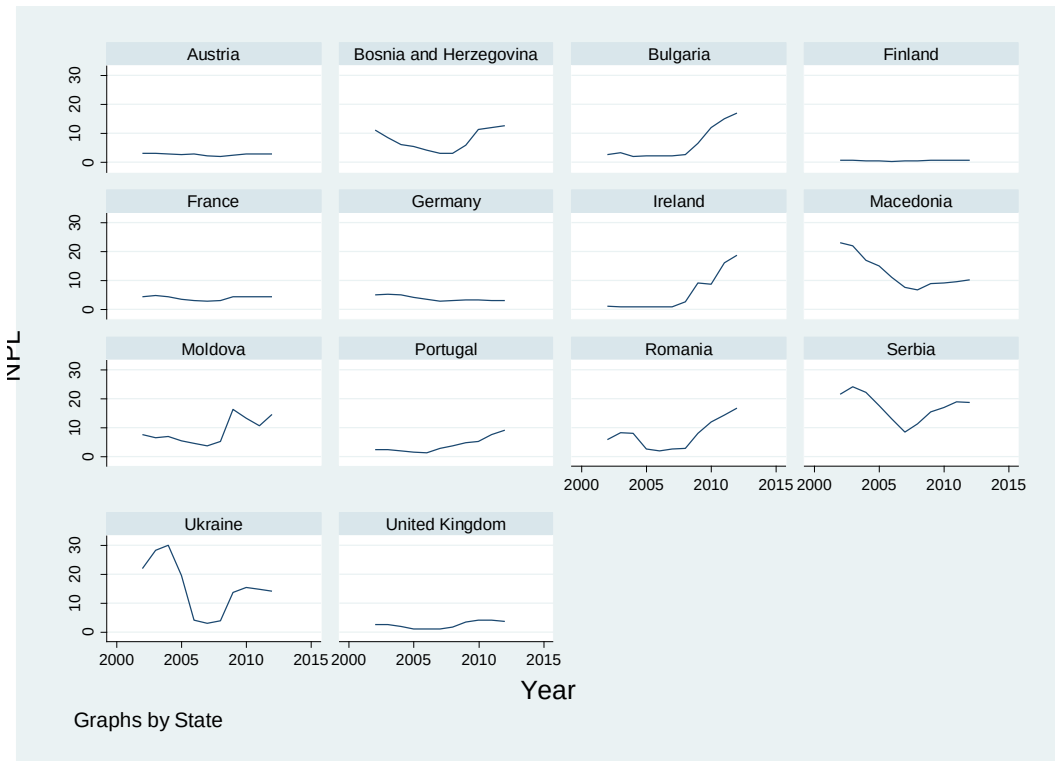
For 2012, the NPL ratio (averaged) recorded 10% in the CEE countries. In Hungary, as well as in Romania, NPL amount soared by 1.5%, but the Hungarian loan amount fell by 5.2%. In CIS region, the highest registered NPL amount belonged to Ukraine – 12 billion EUR.

From Köhler (2012) perspective, banks that are conservative to traditional banking system might indicate higher loan losses, as banks are affected by macroeconomic shocks, for the situation when the lending amount is increasing. In his study, the author indicates acute decline in their ROE.

As a response to crisis, depending on which level of economic development the countries are positioned, Aizenman and Noy (2012) find that middle income countries, experiencing before a banking crisis, will tend to be less sensitive to these crises in future. Concerning the measurement period, the different results for the middle and high income countries is emphasized when the authors include a larger time horizon, not only the recent banking crisis, finding that MI countries are more vulnerable.

Furthermore, the intended research is based on comparing randomly selected countries from two level of economic development, classified by the World Bank Database: HI OECD members (Austria, Finland, France, Germany, Ireland, Portugal, United Kingdom) and non-OECD MI countries (Bosnia and Herzegovina, Bulgaria, Macedonia, Moldova, Romania, Serbia, Ukraine). For the purpose of the present work, a brief revision of all these countries must be analyzed in part with a focus on recent situation. In the next figure (Figure 2.2), the plot of all the listed countries' NPL ratio, with logarithmic difference specification is provided, but more details and stylized facts will be provided later in the empirical part.

**Figure 2.2: NPL plot for the sampled countries**




*Source*: Author's elaboration in Stata

Among the countries Croatia, Bosnia and Herzegovina, Serbia and Slovenia, **Bosnia and Herzegovina** experienced the worst level for financial indicators and according to Dreca (2012), the weak financial situation is driven by an unestablished government, money outflows and large amount of loans from IMF. From 2001 until 2004, the proportion of total assets decreased at the same time with the declining trend of non-performing assets. In 2004 the share of non-performing assets decreased by 5.2% since 2001.

Even so, the effects of crisis on banks' financial statements consist mainly of the accumulation of credit risk (NPLs). At the end of 2008, the GDP growth registered a 5.4% increase only, but overall an increase of the share of loans in GDP is attested with a lowered level only in 2009 (Dreca, 2012). In 2012, the GDP growth started to shrink due to downturn of environment and natural climate conditions (European Commission, 2013). In the same year, real GDP growth decreased by 0.2% and the NPLs surged up to 13%.

After its economic exposure to the Greek crisis, given the ponderous importance of Greek banks (holding 22% of bank assets), **Bulgaria** endured a rise in bad loans up to 16% at the beginning of 2012 (NPL ratio) in contrast with 3.55% in 2009. The banking sector is treated lately, on the other side by the improvement of consumer credits. A major concern for the Bulgarian financial sector is its exposure to euro zone crisis involving the risk of recession over the economy, leading to unemployment and NPL increase.

**Macedonia** experienced in the past years a decreasing trend of bank profitability and at the end of 2005 the NPL ratio consisted of 18% out of total loans. Still, the economy is not as damaged by the global economy and financial crisis as its neighbors (Petrovska and Mihajlovska, 2013). From 2009 to 2010 the output has grown to 2.9% due to exports and global demand. By the end of 2011, the NPL ratio converged to less than 10%, as well as inflation rate which stepped up and then has moderated. The NPL ratio recorded monthly increasing trends and achieved a rate of 11.2% in 2012 (Mahmudi, 2013).

The NPL proportion in the economic sector in **Moldova**, rounded to 10, 6% at the end of 2011, a reduced share from 2010. In general, the annual growth of loans manifested a positive direction comparing the beginning of 2010 with the end of 2012, when the NPL rate reached 14.3%, according to World Finance Review (2013). The review is emphasizing the drop of interest rate on credits denominated in foreign currency during this period (from 8.8% to 8.3%). Nevertheless, by the end of 2012,

the World Bank explains that the growth of NPL ratio must be related to the legislative issues e.g. the approval of International Financial Reporting Standards (IFRS) and the worsening loan quality of state-owned banks.

Concerning the Southeastern Europe, NPL ratio is increasing in 2012 with an additional percentage of 2.5 in comparison with 2009. The lowest level of financial soundness indicator is registered in **Romania** in the first quarter of 2012 consisting of 20.1% (EMC, 2013). A similar situation as in Bulgaria is the foreign ownership of banks by Austrian and Greek banks which had more than half of shares of asset on the market during 2009-2011.

The NPLs were continually growing in **Serbia** and the analysis of Vukovic and Domazet (2013) concludes that, during financial crisis and afterwards, the systemic risk is conditioned mainly by the credit risk. According to availability of information, in 2011 the ratio of NPLs reached a record of 19% which is explained by the failure of credit sector due to decreasing number of corporate loans.

In **Ukraine**, the delicate issue on the uncertainty of NPL national definition is highlighted by Kirchner et al. (2011) as it is noticed that the Ukrainian NPL statistic numbers differ significantly from the international reporting of institutions and credit agencies in charge of analysis due to substantial loans that are not considered by the Ukrainian authorities. The National Bank of Ukraine shows the NPL ratio trend from 2008 to 2010 to have increased by 8.5% in contradiction to figures of IMF that NPL ratio increased by almost 42% during the same measurement period.

Within the same manifestation of economies during the credit boom, followed by the deterioration of the banking sector, the domestic credit demand increased fast from a proportion of 24% from GDP to 82%. In alignment with the global crisis factor and deterioration of banks' assets quality, Ukraine was affected mostly by the devaluation of local currency at the end of 2008, which explained the growth of NPLs (BSTDB, 2011).

Having noticed the facts until this moment, we may conclude that NPL in the selected countries heavily depend on the foreign global likelihood and the Government actions to move the NPL direction in a desired way. Therefore, we concentrate our work more on the MI countries overview concerning the NPL and economic situations.

As long as **Austria's** target countries are the regions of CEE and the forecasts of a positive GDP growth is higher than the average European GDP growth (PwC Austria, 2012), the country's banking sector registered a low level of NPL (2.7%) in

2012. The high development of Austrian banking system (EBF, 2012) is due to microeconomic level and the measures taken for advantageous exchange rate with respect to euro currency denominated loans noticed by EBF Report.

One of the most stable financial markets from the European countries is attributed to **Finland** and according to calculations of the Finnish Financial Supervision Authority, the country's banking sector has a good liquidity and solvency ratios alike the Nordic banks in general (Mattila 2011). In 2008 the NPL ratio was situated at a low level (0.4%), but which increased from 0.3%. The ratio was relatively stable in 2010 and 2011 recording a level of 0.6%, which is low, comparing to other states. One of the most important reasons is that Finnish banks have not been involved directly into the GIIPS financial markets by holding bonds or other assets into their banks (EBF, 2012).

One of the largest economies in the Euro area, **France**, deteriorated its economy, as well, in 2012 increasing unemployment rate and it is included in the six member states of the Euro zone (Germany, Ireland, Spain, Italy, UK and France) that reported NPLs in excess of 100 billion euro at the end of 2012.

**Germany** is considered to own the highest share of NPL proportion on the market in 2009 (PwC, 2012), but it reduced its NPL amount by almost 19% in a year (from 2008 to 2009).

For the period of 2000-2008, **Ireland** was improving its economy in view of expansion of construction sector and domestic demand. Nevertheless, the global financial downturn affected the GDP as in 2008 it declined to 3% and in 2009 to 7%. A specific characteristic set by the Irish budget was the establishment of the National Asset Management Agency in 2009 for the purpose of taking the NPLs off the balance sheet.

The market of NPL in the **United Kingdom** is viewed as stable in 2012 on the fact that the economy has the ability to eliminate any surplus of non-performing assets of 1 billion EUR (PwC, 2012). The impact on the level of NPL and loans in 2009 are considered to be the government interventions, shortage of funding and a low level of provisions held by the banks against distress (PwC Report, 2010).

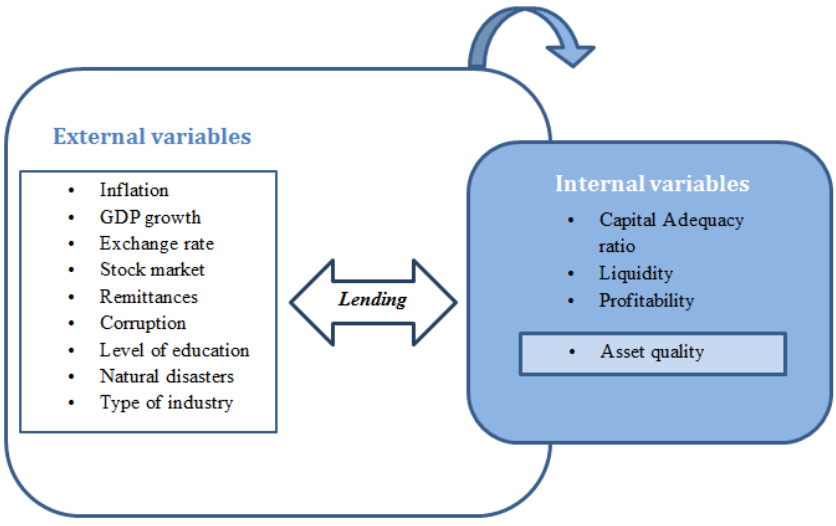
## 2.4. Hypotheses development

Keeping all economic variables constant, the demand and offer of credits have always been the drivers of a country's development from the social, political, economic point of view that create not only benefits for corporations, households and institutions, but several problems for short and long term for lenders (credit default, currency mismatching, etc.). Besides that NPL ratio is considered a financial soundness indicator (R. Babihuga, 2007) assessing a country's banking situation, this variable can be viewed as a preventive measure for bank asset quality to detect the default or credit risk before it occurs.

The work of this thesis is centered on the hypotheses discussed below that takes into consideration the determinants or factors derived from influential conditions that change the NPL ratio. The purpose of detecting the most important determinants of NPLs on the national or public policy level lies on the fact that Central Banks and political institutions are the major decision makers for regulation and monitoring the credit evolution even if the idiosyncratic factors (variables) play an important role for the volatility or stability of the bank financial development indicators. As economic factors affecting NPL ratio at the country level are of great interest for policymakers, the internal determinants cannot be directly influenced by them meaning that in our empirical research we will focus on determinants at the aggregate level.

In Figure 2.3, we can distinguish, in general, the link and interaction between the bank-specific variables and the macroeconomic factors:

**Figure 2.3: Macroeconomics and banking sector**




*Source:* Author's elaboration

For setting the variables that impact our dependent variable - NPL ratio, we must choose the variables of economic or other related conditions that impact the environment positively or negatively to affect the ratio. For the clarity of this work, a positive influence of a determinant on NPL ratio stands for an increase in Non-performing loans due to an increase in the respective variable and the other way round for a negative influence of a determinant.

The basic econometric model for measuring the impact of macroeconomic determinants must take into consideration the following characteristics of market conditions: i) “health condition” of the economic activity; ii) price level of all sectors including real estate, construction, consumers’ goods products and services for purchasing power of corporations and households; iii) banks' vulnerability to monetary policy; iv) foreign impact; v) financial interaction between countries; vi) labor market. Despite that many researches have established their findings on GDP growth, interest rate, inflation, exchange rates, foreign direct investments, unemployment rate, etc., that is those basic variables mentioned in this thesis, this work considers a further research or extensions that can be done in this area and its contribution consists in adding to our basic model new variables and determinants that have low level of empirical findings. Nevertheless, the most important and novel work is done for a range of countries for which does not exist yet similar empirical findings regarding differences of the determinants influencing NPLs.

Proceeding, the additional variables that could describe the determinants for NPLs, in the same proportion of importance, are the aggregated education level and institutional factors.

The GDP growth is the fundamental variable because it describes the best the economic environment, and based on the literature, it is an important explanatory effect showing a significant impact on NPLs changes. For the relationship between them, we hypothesize that the GDP growth leads to a negative impact on NPL ratio, meaning that an increase in GDP causes a reduction on the amount of NPL and vice versa. Fostering the economic activity during good times determines the individuals and corporate sector to provide successful payments of the borrowed credit due to a healthy economic activity.

The inflation rate, as an explanatory variable for our model, assumes that a higher price level reduces the real value of offered loans, diminishing credit risk exposure for banks. In contrast, it weakens the real income of borrowers. Moreover, following the peak of the financial crisis, most central banks have lowered interest

rates' level to near zero bound. One of the many reasons behind this decision was to cope with the low inflationary pressure of that period. However none of the expected favorable changes has happened since the interest rates continued to decline along with the inflation rate, which was supposed to rise to at least the 2% target. Therefore we suppose that this determinant might have ambiguous effects on the evolution of NPL ratio.

Capital inflows create the opportunity for researches to control for the influence of its components as Foreign Direct Investments (FDI), other acquisitions of companies from abroad, remittances or other injections of international financial resources from investors. The reason of choosing the FDI as a representative variable that impacts NPL ratio derives from the fact that the source of capital flows matter in terms of country's economic development. Emerging economies are the major receivers of FDI from advanced states creating the NPLs amount to vary, but which imperatively depends on the contagion effect of the financial crisis. In this order, our hypothesis is based on the foreign influence in the counties of interest, motivating us to introduce it in the empirical model.

Financial interaction between countries is allowed to explain NPL ratio for the reason all the economies aim to increase their current account balance and net exports. The depreciation/appreciation of the local currency impacts the NPL ratio negatively/positively. In general, the exchange rate depreciation is assumed to significantly increase the NPLs. Aside from this general hypothesis testing, we extend our model by creating specifications for the type of commercial trade of a country. Depending on the type of the dominancy in the commercial trade, either export or import status, we introduce different hypotheses for the exchange rate specifications. In case of the dominant export countries, an appreciation of the domestic currency will lead to an increase of NPL ratio on the fact that agents exporting abroad have higher chances to gain during the depreciation of the exchange rate to meet their debt payments faster. In contrast, the import dominant countries are perceived to react differently when the depreciation of the domestic currency occurs, meaning a higher exchange rate or depreciation of domestic currency causes NPLs to go up.

Under the fact that monetary policy affects the banking financial development including NPLs, we control for the situation on the labor market, adding the unemployment rate to our econometric model.

Besides the macroeconomic determinants, other non-economic determinants are considered available as an instrument for monitoring and controlling the credit effect, explained further.

A major public policy, controlled by government and its institutions, is the level of education of a country. Whether education at the aggregate level is an explanatory factor causing NPLs to decrease/increase is based on the amount of loans borrowed by highly educated people, but which can bring the risk of higher default for banks. Therefore the intuition behind this new variable is that people who have attained a higher level of education are more likely to be granted higher volumes of loans, as their qualifications can be viewed as a guarantee of their ability to repay the loan as per contractual conditions. This will increase the volume of outstanding loans which in turn will expose banks to more credit risk. On other hand, the private sector clients are likely to ask for credits due to the reasons of unemployment status.

In contrast with the previous chosen variable, similar to explain the efficiency of institutions, education is partly conditioned by the demand of borrowers to acquire knowledge and/or skills. This leads to add indicators of institutions efficiency which would incorporate all factors at the same time into one index (not separately) assuming that the developing countries associate legal framework, reforms and politics with the problem of transparency, democracy and corruption. To see whether such an index will help other findings for setting a parsimonious model, we test the consistency of it, under several specifications.

Respectively, the developed *hypotheses* originate from the motivation of chosen determinants and are elaborated on a multi-country comparison between different economic development levels defined as:

- GDP growth has a negative consequence on NPLs;
- An increase in inflation will increase the level of asset quality, assessed by NPL ratio;
- Foreign Direct Investments will have a significant impact on NPLs in all groups of countries and a higher FDI from GDP means a lower NPL ratio;
- Impact of exchange rate on NPLs is significant for all groups of countries under import/export dominant specifications;
- Employment amelioration provides a better NPL ratio;
- Acquiring more education provides incentives for lenders to enlarge the amount offered which leads to higher NPL ratio;

- Impact of a single institutional index is significant and their increase have negative impact on the NPLs ratio for all groups of countries;

Having defined our hypotheses, the main research questions are:

*1) What is the influence of the educational index on the NPL ratio in both groups of countries?*
*2) Does the commercial trade of a country determine exchange rate impact over NPLs?*
*3) Does the institutional quality index, compiling all six indicators, provide an overall significant explanation to the NPL ratio and can be used in other econometric models to control for it?*

As it was mentioned above, the countries that were chosen for the study are based on their classification from the World Bank Group which divides them into middle income (including low and upper) and high income.

The first group of countries which belongs to the MI economies or developing states (Bosnia and Herzegovina, Bulgaria, Macedonia, Moldova, Romania, Serbia and Ukraine) is chosen based on their common geographic region (Southeastern Europe) and the OECD membership.

The second group of countries (Austria, Finland, France, Germany, Ireland, Portugal, United Kingdom) is selected from a sample of economies of Western Europe, based on the previous study of Barisitz (2013), that concludes the definition of NPL across these countries are comparable among them including their qualification as OECD member states.

# III. Empirical analysis

## 3.1. Data description

### 3.1.1. Variables and data sources

This study aims at explaining the determinants of NPLs at the aggregate level, their distinctive features for two different classes: developed and developing economies, with a focus on the latter group, which could have a faster improvement to follow up the advanced ones. Even if the presentation and motivation of the states were done in previous sections, this chapter will approach more the practical aspect, beginning with a description of sources and extra details about their choices.

The dataset used encompasses 14 countries, representing the whole sample, and spans from 2002 until 2012 on annual basis, meaning there could be 154 observations in total. The sample is divided into 2 subsamples: the middle income (MI) countries (Subsample 1) and the high income (HI) countries (Subsample 2). The list of countries and their division on how the split is done is showed in table 3.1, explaining how the data will be divided and analyzed.

**Table 3.1.1: List of countries and division of the sample**

| Sample | Subsample 1 | Subsample 2 |
|---|---|---|
| Bosnia and Herzegovina | Bosnia and Herzegovina | Austria |
| Bulgaria | Bulgaria | Finland |
| Macedonia | Macedonia | France |
| Moldova | Moldova | Germany |
| Romania | Romania | Ireland |
| Serbia | Serbia | Portugal |
| Ukraine | Ukraine | United Kingdom |
| Austria | | |
| Finland | | |
| France | | |
| Germany | | |
| Ireland | | |
| Portugal | | |
| United Kingdom | | |

*Source:* Author's elaboration

The main purpose is to analyze the NPL determinants in the first subsample that were randomly selected, along the availability of data with which we had to deal. The criteria to select the countries (both developing and developed) were motivated by the classification of the World Bank Group. From the viewpoint of economic development WB Group divides the MI economies into lower middle and upper middle income countries, based on GNI, ranging from 1,035 USD until 12,615 USD. For the HI countries, the classification considers countries recording a GNI per capita more than 12,615 USD. Besides the motivation of NPL definition and the choice of the countries, the geographical criteria must be taken into consideration for the purpose of comparability between Western and Southeastern countries. The classification is based on the United Nations Population division (UN, 2012). At the beginning of this study, some other countries were chosen for their data availability e.g. Belarus, but according to media public information and rating agencies a flawed reported data on NPL ratio (The Banker, 2013) is believed to persist. The argument arises from the fact that the publicized ratio is too low compared to the size of corporate lending that is increasing and represents a large amount of money. In order to avoid a biased estimation originating from wrong observations, we excluded Belarus from the sample and chose to include Serbia instead, even if the number of missing observations inclined.

For the selection of the dependent variable, NPL ratio, but as well as one of the independent variables – FDI, as a percentage from GDP, we used the World Bank database as it contains the most available observations even for the MI countries. All in all, the sample extended to relatively more time periods than the previous empirical works. Hence, the trial to build up an enlarged time series (NPL ratio) from the Bankscope Bureau van Dijk database, starting with 1998 or 2000 did not allow us to use it objectively due to the significantly reduced number of banks per country. Thus, we consider the period of 2002-2012 from World Bank database as more objective, with less unbiased and inconsistent errors. The annual frequency of the observations is one of the thesis' goals because, firstly, the aggregated data manifest the major changes at the national level and NPL ratio is viewed as more representative in our case. Secondly, the choice is explained by the inability to construct or add all the variables on annual basis to the regression model.

For the sources of the explanatory variables - the GDP growth rate, inflation rate and the unemployment rate, the WEO from IMF database was the most appropriate. In case of exchange rates, we used the official nominal exchange rates

according to various sources and National Banks of some countries. For some states, the foreign currency was chosen in accordance with the exchange rate's reference currency that is the foreign currency in which the majority of the transactions are denominated. In case of Ukraine, Moldova and the Eurozone countries, the exchange rates against US dollar are used, and against EURO for the rest. An equitable comparability and influence of this determinant of all mentioned countries motivated us to choose a benchmark currency for each of them. The reason arises from the fact that the appreciation or depreciation matters mostly, and not the currency on which the exchange rate is referred to, for capturing the movements and their importance on the international financial markets.

One of the contributions to this topic, for the interest of policymakers and regulators, is the explanatory variable of education incorporated into the educational index. Its calculation follows the algorithm of Human Development Index that covers the Educational Index[1], but due to its unavailability for the selected countries, we constructed it according to the formulas below, based on Human Development Report (2013):

$$Education\ Index = \frac{\sqrt{Mean\ years\ of\ schooling\ index * Expected\ years\ of\ schooling\ index} - 0}{Combined\ eduational\ index - 0}$$

The mean years of schooling and the expected years of schooling indices are computed according to the dimension index with the minimum and maximum values in order to transform the final index into observations ranging from 0 to 1:

$$Dimension\ index = \frac{x_0 - x_{min}}{x_{max} - x_{min}},$$

where:

$x_0$ - actual value

$x_{min}$ - minimum value

$x_{max}$ - maximum value

It should be denoted that the combined educational index must be computed as for each year separately and for all the countries in the world. Instead we make the assumption that combined educational index is constant through time and consider

[1] An educational index can be found at OECD Better Life Index on their official website, but which offers observations only for 2011, at the time of the thesis elaboration

the example for calculation in the UN report revealing that New Zealand in 2010 is the country with the highest value of 0.971.

The source used for the sub-indices of the ratio was the UNDP database, but which lacks the observations for Bosnia and Herzegovina, as well as Macedonia and that the time spans from 2005 until 2012, for these 2 countries.

The explanatory variable, also included as an index in the basic econometric model, is the institutional quality index set up as a simple average of the institutional indicators. Some authors consider measuring the institutional factors taking the average of the 6 indices of Kaufman or ICRG political risk index creating a single institutional index (Papaioannou, 2007; Bojnec and Ferto, 2009). For capturing the general effect of the government quality and result of their policies, we construct the institutional quality index from six indicators of Political Risk Services International Country Risk Guide (PRS) that encompasses other 12 sub-indicators:

- **Voice and Accountability:** ability of the population to express their choice to select the government and show an interest in their decision
- **Political Stability and Absence of Violence:** at what degree the terrorism, demonstrations and the potential risk of violent acts are kept under control
- **Government Effectiveness:** the facility of the government to create conditions for investments with effective bureaucracy and continuous policy and reform implementation
- **Regulatory Quality:** to what extent legislation is clear and transparent, including effectiveness of taxation system
- **Rule of Law:** evaluation of a state's ability to solve conflicts, act efficiently in case of crimes and offer security to citizens and residents
- **Control of Corruption:** the assessment of a country's ability to avoid corruption

The meaning of one single institutional quality index is: the higher the index, the higher the quality within a country, explained by the ranges of institutional quality and country risk below:

- Very High Risk: 0 - 49.9%
- High Risk: 50% - 59.9%
- Moderate Risk: 60% - 69.9 %
- Low Risk: 70% - 79.9 %

- Very Low Risk: 80% - 100 %

In case of Bosnia and Herzegovina, as well as Macedonia, the PRS ICRG does not offer any observations for all the institutional indicators. As a solution, another source was used (Global Insight Business Conditions and Risk Indicators) that gives similar indicators with ICRG for the rest of the countries.

Altogether, our data has missing observations for the exchange rates for Bosnia and Herzegovina along with Romania from 2002-2004. The educational index data is mentioned above. For the variables that we did not have data for the respective years in the sample, but we had found the observation from previous year, we used the interpolation method and constructed the missing observations, meaning for Serbia (2005 and 2006) and Romania (2002), in case of NPL ratio. For the reason to keep up with the time series' length, the NPL ratio for Germany in 2012 was found in the Ernst and Young Report (2013).

All the economic and non-economic variables of the basic model are summarized in Table 3.2, where the measurement type of each of them is mentioned.

**Table 3.1.2: List of economic and non-economic variables of the basic model**

| Variable and its abbreviation | Characteristics | Source |
|---|---|---|
| *Non-performing loans ratio (NPL)* | share of gross loans; % | WB |
| *Gross Domestic Product (gdpG)* | constant prices; percent change | WEO |
| *Inflation rate (Infl)* | average consumer prices; percent change | WEO |
| *Foreign Direct Investments (FDI_gdp)* | percent of GDP, % | WB |
| *Exchange Rate (ER)* | Local currency per USD/EUR | National banks, official exchange rates websites |
| *Unemployment rate (Unempl)* | percent of total labor force; % | WEO |
| *Educational Index (Edind)* | interval from 0 - 1 | for construction used Human Development Report and UNDP |
| *Institutional Quality Index (IQind)* | interval from 0 - 1 | PRS ICRG |

*Source:* Author's elaboration

### 3.1.2. Descriptive statistics

The summary statistics or descriptive statistics allows the data user to prevent errors in the further estimations and offers the possibility to create a general framework about the behavior of the dataset. In Table A.1, we can distinguish the descriptive statistics for all the countries (observations = N) and for each group of countries (Table A. 2). When the overall observations are examined, we elucidate a high variance across countries as highlighted by the minimum and maximum of most of the variables. NPL ratio varies from 0.2% until 30% or the unemployment rate with a minimum of 3.80% until 37.25%. On one hand, for all the variables, the standard deviation is relatively high, meaning that the presence of heterogeneity in the sample is also high, with an exception for the indices of education and institutional quality that do not manifest a very high deviation from their mean. This explains the near level of education and quality for institutions within the middle and high groups, comparing to the overall level of development including low income countries for the MI ones.

Concerning the differences between the two subsamples, we highlight some interesting points derived from the summary statistics, from Table A. 2. When the observations for the MI countries are analyzed (dum_coun = 0), we detect that the maximum level of NPL ratio (= 30%) originates from this group, the figure belonging to Ukraine in 2004, while the minimum of NPL ratio is lower in the developing countries (the observation from 2004 in Romania) than in the developed ones. When the figures for NPL ratio and GDP growth are examined, the mean and standard deviation are higher in the MI countries for both variables including the GDP growth. A higher GDP growth interacted with a higher NPL ratio in the MI countries can be justified by financial lending to the private and public sectors improving faster the economic conditions, but which should not viewed as a general fact due to side effects of accelerated lending during good times.

To check for the indices' stability included in the analysis, keeping in mind that a higher value means a better quality for both variables, we notice that the minimum of educational index variable is almost similar, with a difference of 0.026 units only. That corresponds to Macedonia for 2005-2006 and Portugal in 2002, which to our mind is viewed as reasonable. Regarding the institutional quality index, we can observe the maximum of 0.990, which corresponds to Finland according to the database, is much higher than in case of the MI group with a maximum of 0.680.

With respect to this variable, we conclude the institutional factors that display a better quality in the advanced countries than the developing states which confirms the indices are appropriate, for a correct interpretation in the next steps.

### 3.1.3. Stationarity and variable specifications

The analysis follows the procedures characterized by a panel data setting, known as longitudinal or cross sectional time series in which the econometric model takes the form:

$$y_{it} = \alpha + x_{it}\,\beta + \varepsilon_{it} \qquad (1)$$

where $i$ = 1, …, N representing the country and $t$ = 1, …, T standing for each year in our case, $\varepsilon_{it}$ is the error term, $\alpha$ the constant and $x_{it}\,\beta$ the coefficients of the variables explaining the dependent variable $y_{it}$. Our panel data for the sampled countries is considered to be unbalanced due to the missing data for exchange rates and educational index. The practical aspect is performed with the help of the statistical software Stata, and Gretl used occasionally.

Knowing that panel data consists of time series observations, the stationarity problem of the variables must be taken into consideration in order to avoid the spurious regression which can give us wrong inference for our coefficients. Moreover, importance of this sub-section is greater as the unit root is persistent in macroeconomic variables, especially. According to Greene (2012), if the unit root in macroeconomic variable data is integrated of order one I(1), then shocks which can occur may become remarkable and permanent. In this way, the unit root (non-stationarity) tests are used to detect the random walk which could characterize the data. In case the results of the tests give a positive response for presence of the unit root, we are required to transform the variables that cause the non-stationarity issue.

Considering the model from equation (1), the first-order autoregressive process can be introduced so as:

$$y_{it} = \alpha + \rho_i\, y_{i,t-1} + x_{it}\,\beta + \varepsilon_{it} \qquad (2)$$

where the tests for the unit root presence are used to test the null hypothesis (Stata): $H_0$: $\rho_i = 1$ or the data is non-satationary against the alternative hypothesis $H_A$: $\rho_i < 1$, or in other words the variable is stationary.

**Table 3.1.3: Stationary data**

| Sample Unit root tests | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| VARIABLES | npl | gdpg | er | unempl | fdi_gdp | edind | iqind | infl |
| Levin-Lin-Chu (LLC) | -5.5091 | -4.452 | - | -4.2636 | -5.7009 | - | -5.0934 | -5.2427 |
| | 0.0000 | 0.0000 | - | 0.0000 | 0.0000 | - | 0.0000 | 0.0000 |
| Harris-Tzavalis (HT) | 0.3255 | 0.2015 | - | 0.2822 | 0.0617 | - | -0.2475 | 0.3143 |
| | 0.0000 | 0.0000 | - | 0.0000 | 0.0000 | - | 0.0000 | 0.0000 |
| Breitung test | -5.1366 | -5.5436 | - | -2.9876 | -4.2237 | - | -4.2747 | -2.3973 |
| | 0.0000 | 0.0000 | - | 0.0014 | 0.0000 | - | 0.0000 | 0.0083 |
| Im-Pesaran-Shin (IPS) | -1.9076 | -2.2516 | - | -2.7139 | -2.4923 | - | -3.3676 | -2.7764 |
| | 0.0282 | 0.0026 | - | 0.0009 | 0.0018 | - | 0.0000 | 0.0001 |
| Fisher type (ADF) | -1.6571 | 51.5933 | 70.4271 | -5.1958 | 78.714 | 81.066 | -8.1232 | 90.3237 |
| | 0.0488 | 0.0043 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Subsample 1 Unit root tests | | | | | | | | |
| VARIABLES | npl | gdpg | er | unempl | fdi_gdp | edind | iqind | infl |
| Levin-Lin-Chu (LLC) | -7.6576 | - | - | -2.6702 | -1.2250 | - | -2.8363 | -3.2267 |
| | 0.0000 | - | - | 0.0038 | 0.1103 | - | 0.0023 | 0.0006 |
| Harris-Tzavalis (HT) | - | - | - | 0.2778 | 0.5214 | - | 0.3252 | 0.3155 |
| | - | - | - | 0.0000 | 0.0086 | - | 0.6275 | 0.0000 |
| Breitung test | - | - | - | -1.5424 | -2.2964 | - | 0.5039 | -0.9448 |
| | - | - | - | 0.0615 | 0.0108 | - | 0.0051 | 0.1724 |
| Im-Pesaran-Shin (IPS) | -3.1973 | - | - | -3.1777 | -1.6588 | - | -2.8642 | -2.9525 |
| | 0.0007 | - | - | 0.0050 | 0.3243 | - | 0.122 | 0.0018 |
| Fisher type (ADF) | 70.2598 | 93.1343 | 48.4314 | 75.6506 | 16.4564 | 56.4729 | 89.5443 | 51.7722 |
| | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.2863 | 0.0000 | 0.0000 | 0.0000 |
| Subsample 2 Unit root tests | | | | | | | | |
| VARIABLES | npl | gdpg | er | unempl | fdi_gdp | edind | iqind | infl |
| Levin-Lin-Chu (LLC) | - | -4.3387 | -2.1024 | -3.4146 | -6.2696 | -2.6485 | -2.7038 | -4.3527 |
| | - | 0.0000 | 0.0178 | 0.0003 | 0.0000 | 0.0040 | 0.0034 | 0.0000 |
| Harris-Tzavalis (HT) | - | 0.2505 | 0.5168 | 0.2989 | -0.2728 | 0.1964 | -0.0677 | 0.2991 |
| | - | 0.0000 | 0.0075 | 0.0000 | 0.0000 | 0.0000 | 0.0000 | 0.0000 |
| Breitung test | - | -4.052 | 0.2054 | -3.1114 | -2.9245 | -2.9792 | -4.3367 | -3.1084 |
| | - | 0.0000 | 0.5814 | 0.0009 | 0.0000 | 0.0014 | 0.0000 | 0.0009 |
| Im-Pesaran-Shin (IPS) | - | -2.3168 | -3.2106 | -2.2500 | -3.1664 | -2.1227 | -2.8356 | -2.6003 |
| | - | 0.0165 | 0.0006 | 0.0330 | 0.0011 | 0.0611 | 0.0096 | 0.0072 |
| Fisher type (ADF) | 239.891 | 26.6334 | 59.9282 | 26.6945 | 62.4279 | 24.7545 | 55.8510 | 38.5515 |
| | 0.0000 | 0.0215 | 0.0000 | 0.0211 | 0.0000 | 0.0371 | 0.0000 | 0.0004 |

*Source*: Author's elaboration; Note: adjusted t*, t-bar, chi-squared or inverse normal statistics reported; Results from Stata, and Gretl occasionally

A range of tests can be implemented for our data, including Levin-Lin-Chu (LLC), Harris-Tzavalis (HT), Breitung test, Im-Pesaran-Shin (IPS) and Fisher-type of Augmented Dickey Fuller (ADF), excluding Hadri LM test which has an opposite null hypothesis and is strongly recommended by its author to be used when the time spans over a very long period (Stata).

It is worth mentioning that some tests are performed and preferred more as the Fisher ADF test and IPS because they allow for unbalanced data presence. Nevertheless, the other tests are limited only for our variables that have missing observations. The unit root tests are performed for all of the variables and show the results of the variables in level in Table A.3, separately on the sample, subsample 1 (MI group) and subsample 2 (HI group).

According to the results for the full sample, where the highlighted values are the p-values of the t-statistic, 3 variables are found to be non-stationary: the NPL ratio, unemployment rate and the institutional quality index. The rest or more exactly the regressors left, all are stationary and we fail to accept the null hypothesis of a unit root at the 5% confidence level. Notwithstanding the non-stationary variables, most of the relevant tests that were carried out (IPS and ADF tests) accept the null hypothesis as the p-values are insignificant and the t-statistic is less than the critical value. When we take the logarithmic differences of NPL and the institutional quality index and set the order of integration denoted as I(1), when first differencing in case of the unemployment rate, we achieve our goal of correct specification. The results of no unit root of the variables in level are displayed in Table 3.3 from above.

In the same order, we proceed with the subsamples, testing for the presence of unit root and in case of a positive response, we transform our variables. Hence for the first subsample we find that instead of logarithmic differences of NPL, the stationarity is supported only when 3 lags are added to this variable. As for the GDP growth, the results are confusing when using Stata and Gretl and in order to avoid the problem of misspecification, the test from Gretl is used (ADF) to conclude that *gdpg* must be first differenced[2]. The unemployment rate takes the same form as in the sample specification (first differenced), like the exchange rate does. Regarding the FDI, as a percentage from GDP, when taking the logarithm of it, we assume that it

[2] We try to avoid the first difference as much as we can because it removes many observations and we lose information. Even though, the non-stationarity problem is considered to be more important in our study, then we choose to differentiate if there is no other possibility.

does not have a unit root according to Breitung or Harris Tzavalis, though the result of Fisher test accepted the unit root.

For the high income group of countries, it was required to add 3 lags of the logarithmic differenced NPL variable and to first differentiate the unemployment rate, institutional quality index and the educational index. Therefore, our data is qualified as correctly specified because we could test and accept the absence of a unit root for our panel data.

To sum up this subsection, a general table of how all variables are abbreviated and specified per each subsample and the full sample (based on stationarity results) is provided:

**Table 3.1.4: Unit root test results and transformations**

| Variable | Full sample | | Subsample 1 | | Subsample 2 | |
|---|---|---|---|---|---|---|
| *npl* | non-stationary | logarithmic difference[3]; | non-stationary | lag length = 3 | non-stationary | logarithmic difference (3 lags) |
| *gdp* | stationary | no transformation | non-stationary | first difference | stationary | no transformation |
| *er* | stationary | no transformation | non-stationary | first difference | stationary | no transformation |
| *unempl* | non-stationary | first difference | non-stationary | first difference | non-stationary | first difference |
| *fdi_gdp* | stationary | no transformation | non-stationary | logarithm[4] | stationary | no transformation |
| *edind* | stationary | no transformation | stationary | no transformation | non-stationary | first difference |
| *iqind* | non-stationary | logarithmic difference | stationary | no transformation | non-stationary | first difference |
| *infl* | stationary | no transformation | stationary | no transformation | stationary | no transformation |

*Source*: Author's elaboration in Stata, Gretl occasionally

An attention should be paid to the relevant researches to our topic as the stationarity issue is approached differently so that testing the specified variables varies from author to author. Some consider an important start for an empirical analysis (Jakubik and Reininger, 2013; Babouček and Jančar, 2005; De Bock and Demyanets, 2012; Vogiazas and Nikolaidou, 2011; Moinescu and Codirlaşu, 2011; Castro, 2012; Nkusu, 2011; Klein, 2013; Espinoza and Prasad, 2010, etc.), whereas

[3] Instead of logarithmic difference we also found that NPL is stationary under 3 lag length and first differenced, but we choose to take logarithm form, not lose more observations

[4] This variable (for the middle income economies) proved to be stationary only under the logarithmic specification and the first difference, which is avoid, for keeping maximum the number of observations

others do not put an imperative weight on it (Boudriga et al., 2010; Boudriga et al., 2009; Jimenez and Saurina, 2006; Ahmad and Ariff, 2007, etc.). To our mind the stationarity and specification of variables is important.

For our panel data, split into subsamples, the unit root testing proved to detect different results of non-stationarity of the variables in all subsamples and lead to different specifications for them. Overall, the explanations may derive from the fact that the setting of the panel is different. Besides this, even if the short time series do not offer very reliable results, the main stylized facts over NPL ratio, that unemployment rate and/or GDP growth with their correct signs and significance, should indicate when the stationarity transformations of variables is well-behaved. However, the extent to which a study is pursued is mainly driven by the subjectivity for the setup of the variables, but which must give sensible results in case the unit root tests are considered or not.

### 3.1.4. Stylized facts

The NPL ratio (measured as logarithmic difference) in our sample is characterized by very volatile dynamics for some countries and a rather smoothed trend for others, meaning the macroeconomic or non-macroeconomic development captures the favorable periods as well as the period of financial distress. In Figure 2.2, we can detect that for the HI countries (Austria, France, Germany, and less for Portugal, Finland) the NPL plot does not detect severe and alarming features. In contrast, countries as Ukraine, Romania, Bosnia and Herzegovina or Moldova reflect a worse situation with high volatility in the data. Nevertheless, both groups are characterized by a rise of NPL ratio after the burst of the financial crisis and European sovereign debt crisis around 2009, where the Southeastern countries reacted much stronger with a sudden acceleration of the ratio.

Although we can see the periods of fluctuations in the NPL plot of sampled states, for a rigorous analysis, we could follow and compare the preliminary facts based on 2 key macroeconomic variables: the unemployment rate and the GDP growth rate, depicted in the plot from Figure A. 1. Accordingly, we can confirm that there is an evident opposite trend between NPL ratio and the GDP growth rate and a similar trend between NPL ratio and the unemployment rate. All in all, we can shed more light on the determinants of NPLs through econometric models beginning with the static ones and concluding with a dynamic panel model.

## 3.2. Econometric models

We have introduced the panel data structure in the previous section and it is worth mentioning that this type of data offers a curious insight on the behavior of the cross sections (countries) in time meaning that our motivation for panel data arises from our interest into the heterogeneity across countries and understanding their differences. One of the great advantages using panel data models is the precise estimation because of many observations, combining time series and cross sections (Cameron and Trivedi, 2005). Moreover, panel data framework estimation allows to control for biases derived from the heterogeneity effect, solving the omitted variable bias.

In general, the basic framework for panel data regression is (Greene, 2012):

$$y_{it} = \boldsymbol{X'_{it}}\boldsymbol{\beta} + \boldsymbol{Z'_{i}}\boldsymbol{\alpha} + \varepsilon_{it} \qquad (3)$$
$$y_{it} = \boldsymbol{X'_{it}}\boldsymbol{\beta} + c_i + \varepsilon_{it}$$

where $\boldsymbol{Z'_{i}}\boldsymbol{\alpha}$ or $c_i$ represents the heterogeneity or individual-specific effects that may be observed or unobserved. Their explanations and role for our work are described later in subsection 3.2.5, but before developing these concepts, we are interested in data processing to take advantage of its true information as much as we can. Consecutively, an analysis starting with pooled OLS is imperative, in order to understand our strong motivations for choosing a correct estimation method.

### 3.2.1. Pooled OLS

For the beginning, the estimation framework applied to our panel data will consider the simplest model, but also the most restrictive panel data model that is not used much in the literature. However, Boudriga et al. (2010) implements the pooled regression as it has the advantages to increase the degrees of freedom, decrease collinearity among regressors, when number of years is reduced.

Pooled OLS is implemented, and in this case, it is the most consistent and efficient among all estimators only if the individual effect from equation (3) is observed for all countries and meets the assumptions (Greene, 2012):

1. Linearity in the model
2. Exogeneity of explanatory variables: $E[\varepsilon_i \mid x_{j1}, x_{j2}, \dots, x_{jk}] = 0, i, j = 1, \dots, n$

3. Full rank of matrix $\boldsymbol{X}$
4. Homoskedasticity (constant error term) and non-autocorrelation: $Var[\varepsilon_{it} \mid x_{i1}, x_{i2}, \dots, x_{iT}] = \sigma_{\varepsilon}^2$ and $Cov[\varepsilon_{it}, \varepsilon_{js} \mid x_{i1}, x_{i2}, \dots, x_{iT}] = 0 \; if \; i \neq j \; or \; t \neq s$
5. Asymptotically normal distribution of error terms

In case one of the assumptions is not met, then OLS becomes unbiased and inconsistent. Our basic model of the study will take the form based on the variables in levels (not displaying the transformation of variables as it changes to both subsamples and it would be difficult to follow, but regressions of the full sample, subsample 1 and subsample 2 are based on transformations from Table 3.4):

$$\begin{aligned} NPL\,ratio = \; & \beta_1 * gdpG_{it} + \beta_2 * ER_{it} + \beta_3 * Unempl_{it} + \beta_4 * FDI_{gdp_{it}} \\ & + \beta_5 * Edind_{it} + \beta_6 * IQind_{it} + \beta_7 * Infl_{it} + c_i \\ & + \varepsilon_{it} \end{aligned} \tag{4}$$

### 3.2.2. Multicollinearity

First of all, the multicollinearity problem is diagnosed because its presence can cause perfect linear dependencies (high correlation) between the regressors and wrong magnitude of estimators. To detect the perfect collinearity in order to decide whether to exclude a regressor or not, beside the correlation matrix (Table A. 4), we use the Variance Inflation Factor (VIF) on which our decisions are based. For the full sample, the correlation matrix reveals us the correct signs for the key variables with respect to NPL ratio. The highest correlation is detected between the NPL and GDP growth (-0.60), and the unemployment rate and GDP growth (-0.48). The insignificant result of the correlation between some variables is questioned e.g. the inflation rate and the NPL ratio or the educational index and the institutional quality index. Even so, the low correlation may not imply low collinearity (Chennamaneni et al., 2008). As such, VIF measures the inflated variances in case the explanatory variables are highly correlated. Some authors precisely Neter et al. (1990) or Ahmad and Ariff (2007) recommend that a rule of thumb would be to observe the values of VIF higher than 10, and those variables corresponding to values higher than 10 should be omitted from the regression. From Table A. 5 a), b), c), d), the regression on the full sample under the VIF method, does not imply high collinearity, meaning

we do not exclude any variable. In the subsample 1 or the MI countries regression, the collinearity problem is detected as VIF yielded a value of 12.69 for the second lag of NPL. After its omission from the model, the VIF analysis gives the results presented in Table A. 5 c).

### 3.2.3. Robust standard errors

To correctly analyze the information from our data, the clustered robust standard errors must be applied in order to produce valid estimates and significance tests. Across the analyzed countries, the observations must be independent on the other observations and not correlated between them, but the opposite may happen often in a dataset. For this reason, we show our results using robust standard errors or Huber–White sandwich estimator of the variance (Baum, 2006), clustered by country, in Table A. 6 along with the estimation with the simple standard errors. The results from pooled regression with clustered robust standard errors offer lower standard errors than using simple standard errors (SE). The significance tests show that robust standard errors increase the variables to be significant for the model in case of all the countries and the MI group, with exception of the HI group, where the simple OLS suggests that only unemployment rate influences NPL ratio, and moreover at 10% level.

### 3.2.4. Structural break

Allowing for sudden economic events to persist in a model refers to the presence of a structural break into time series leading to less precise estimators and spurious regression. The events (Liao, 2008) can emerge from financial crises, economic liberalization, changes on monetary policies or exchange rate regimes. To test the structural change we use the Chow test by assuming there was a breaking point in 2009 due to the financial debt crisis. Furthermore, the full time series take a dummy variable that divides the series into 2 periods: i) from 2002 – 2009 and ii) from 2010 – 2012. The underlying assumption of the Chow test relies on the fact that the coefficients in the two regressions of the mentioned periods are the same, meaning they are constant and do not change through the entire time series. The null hypothesis of the Chow test is that the restricted model is not different from the

unrestricted model ($\beta_1 = \beta_2$, $\alpha_1 = \alpha_2$), or in other words, there is no structural break. The results of the test are shown below in Table 3. 4, suggesting the test did not detect any structural changes. Even though, the reduced time period or the uncertainty of break point reduces the reliability to use the test.

**Table 3.2.1: Chow test results for structural break**

| Full sample | Subsample 1 | Subsample 2 |
|---|---|---|
| F( 8, 118) = 0.38 | F( 8, 46) = 0.83 | F( 8, 54) = 1.68 |
| Prob > F = 0.9302 | Prob > F = 0.5845 | Prob > F = 0.1244 |

*Source:* Author's elaboration; Results from Stata

### 3.2.5. Fixed Effect model vs. Random Effect model

It is mandatory for a correct estimation of the model to assess the regression from equation (3) because of the presence of the unobserved variables that do not change over time. In contrast to pooled OLS, the Fixed Effect (FE) and Random Effect (RE) models can produce unbiased and consistent estimators if the individual effects are detected in our sample and subsamples. On the other hand, if there are no fixed or random effects, the least square estimator is the most efficient and robust coefficient. Even though the panel dataset containing the heterogeneity among the countries can be modeled by first differences, Least Square Dummy Variable (LSDV), within (FE) and between-groups estimators, the thesis will be focused on the FE (within) and RE estimators as the literature and theoretical background place these two models at a higher level of precision.

In case the regression model (3) does include the individual effects $c_i$ or the coefficient of $\boldsymbol{Z'_i}$ and they are unobserved, constant across time and are correlated with at least one of the explanatory variables $\boldsymbol{X'_{it}}$ ($Cov(\boldsymbol{X'_{it}}, c_i) \neq 0$), then fixed effects are present in our panel. As a result the omitted variable problem occurs in the presented situations. Consequently, econometricians developed the FE model, viewed as a tool (Wooldridge, 2006) to transform the unobserved fixed effects by removing them from the estimation, which equals the pooled OLS after the time-demeaned data or the within transformation taking the form:

$$y_{it} - \bar{y}_{i.} = (x_{it} - \bar{x}_{i.})'\beta + (\varepsilon_{it} - \bar{\varepsilon}_{i.}) \qquad (5)$$

where the $\bar{y}_{i.}$, $\bar{x}_{i.}$and $\bar{\varepsilon}_{i.}$ are the mean of each of group $i$.

In this way, the unobserved effects are eliminated, but together with the time-invariant variables included in the regression like the education of individuals. In our case the disadvantage of this method does not threaten the study, as the education index with respect to countries varies slowly and is not constant for long periods.

The unobserved variables $\boldsymbol{Z'_i}$ may be uncorrelated with the explanatory ones, and for this situation, the RE model is useful for a robust estimation. The RE model assumes an error component model where the error term comprises the random effects ($u_i$) (distributed randomly across countries) and the error term ($\varepsilon_{it}$):

$$y_{it} = \boldsymbol{X'_{it}}\boldsymbol{\beta} + (\alpha + u_i) + \varepsilon_{it} \quad (6)$$

The assumptions of non-correlation of the error term and all other components of the model, constant variance of the error term impose stricter restrictions than the FE model. While the FE model assumes the individual effects causes different intercepts across countries, the RE model makes the difference in the specific errors but the slopes and the intercepts are similar for all sampled states.

Even if the FE model is more effective for panel data derived from weaker assumptions than the pooled OLS or RE model, the assumptions come at a cost of losing empirical findings for the time-invariant regressors. However, the RE model provides inconsistent coefficients due to correlation of the unobserved with the explanatory variables, concluding that FE model in this case leads to consistent estimate. For a better understanding, the assumptions of FE model can be detected in our study. In this case, there are explanatory variables which cannot be observed (measured). Intuitively, some examples may be provided as the attitude of investors towards a country influenced by international rankings, or the geographical area of the subsampled states due to historical and political regimes correlated with institutional factors, etc. For most of the times, it is not clear if one should use random or fixed effect model and according to Lee et al. (2006), without sufficient proves of assumptions regarding distribution to use RE model instead of FE, the preference is given to fixed effects. Therefore, we hypothesize that our data contain rather fixed effects than random effects and the tests performed will allow confirming or rejecting this. The fact is also supported by Wooldridge (2006) debating that in aggregated data, the random effects cannot be implied as the observations come from

large units (states, regions) and FE model is much more convincing when choosing between RE and FE models for the interest of policy analysis.

### 3.2.6. Hausman test

For identification of which model fits better our panel data - fixed or random effects, the Hausman test is performed. The mentioned distinction before - the correlation between the unobserved effects and the regressors, is the criteria to decide whether to use the fixed or random specification method. In this way, the following hypotheses are developed:

$$H_0: Cov(\boldsymbol{X'_{it}}, c_i) = 0$$
$$H_A: Cov(\boldsymbol{X'_{it}}, c_i) \neq 0$$

Under the null hypothesis ($H_0$), all priority is given to the random effect model, meaning that in case the result of the test is insignificant (p-value > 0.05) then $H_0$ is accepted. The opposite interpretation is specific to the alternative hypothesis that suggests using the FE model if the null is rejected. The motivation for choosing the right model is determined by the features depicted in table 3.5:

**Table 3.2.2: Hausman test specifications**

| Hausman test | $H_0$ | $H_A$ |
|---|---|---|
| FE model | consistent + inefficient | consistent + efficient |
| RE model | consistent + efficient | inconsistent |

*Source:* Cameron, 2012

Hence, it is denoted the fact that under null hypothesis, we must choose the RE model motivated by the consistency and efficiency of the estimator against the FE estimator that is only consistent, but not efficient. The FE model is appropriate under the alternative hypothesis because it gains both the consistency and efficiency while RE model does not assume any one of them.

All things considered, the Hausman test performed in the statistical software give the results from Table A. 9. All three settings of our models (full sample and subsamples) incorporate different behavior at the country level across time. The test result shows that all p-values are significant at less than 5% level, meaning that $H_0$ is

rejected, and RE model is inconsistent, allowing us to choose the FE model which is the most consistent and efficient one among them. In this way we confirmed our expectation that there are individual effects, and the parameter estimates for the countries involved have different intercepts. As for the preliminary results for our main hypotheses, we regress the basic model using the sample and the subsamples, including the robust standard errors clustered by country. In the Table A. 7, it is evident that there are large differences in what matters the significance of the NPLs determinants. In case of MI countries, all variables are significant with the exception of the inflation and unemployment rate. By contrast, in the HI groups of countries only unemployment rate showed a significant result at 5% level. For our hypotheses concerning the indices, we notice that IQ index is highly important for the MI states, and not significant for the other group of economic development. In case of the educational index, we conclude that this explanatory variable is not affecting the NPL ratio in a country, based on the preliminary results of the FE model.

According to the literature, a dynamic model which includes lagged dependent variable may lead to biased and inappropriate estimators due to endogeneity and autocorrelation problem. Therefore, we apply several tests to detect these problems including the test for heteroskedasticity.

### 3.2.7. Additional empirical tests

Looking in Table A. 8, several tests were executed for checking the precision of the FE estimation. Before all else, if the heteroskedasticity and serial correlation problems are found in our data, the robust standard errors option is used to allow their presence.

The Wald statistic test is applied for checking the heteroskedasticity in error terms. According to the statistical result, we reject the null hypothesis of homoskedasticity, meaning the variance of residuals is not constant in time.

Another problem to be detected is the possibility of serial correlation in the error terms. Allowing its presence in the regression, creates the likelihood of lowered standard errors, while the R-squared value is inflated upward. For testing the probability of serial correlation, we apply the Wooldridge test, which assumes that under null hypothesis there is no serial correlation. According to the p-value $< 0.05$ of the F test probability, only for the sampled MI countries the serial correlation is detected and for the rest of the regressions the null hypothesis is rejected.

The last test we run for the FE model is the Pesaran CD test for cross-sectional dependence, for which the null hypothesis is the following (without a robust estimation as it is not allowed by the program): the residuals across states are not correlated. The test indicates that we must choose the alternative hypothesis for the full sample because the p-value = 0.0000, but the null hypothesis is confirmed for the sub-samples.

Having applied the static models procedures up to this point, concentrates our attention to some extensions of our two hypotheses concerning the exchange rate and institutional quality index determinants.

### 3.2.8. Components of the institutional quality index

Despite the correlation matrix (Table A. 7) does not distinguish a high correlation between the variables and the IQ index using the VIF results, we denote the highest values (after the lagged NPL variable) are attested for it in case of both subsamples, although not including the full sample. For this reason, but as well for a deeper understanding of institutional factors that have an impact on NPL ratio, an attention is paid to the components of the single index. The components are included separately in the basic model due to the highly correlated effect between them, if all the 6 factors are added. As it was turned up, the IQ index was significant in the MI countries from the two groups that we differentiate in this thesis. Thus, the results depicted in Table A. 7 show that, among all the IQ index components, only the Regulatory Quality (RQ) factor for the first subsample was found to be an important determinant for NPLs. Even though the second subsample regression did not show any significant impact of the IQ index, a software run of the separated regression, indicates that only Political Stability and Absence of Violence (PSAV) variable is actually significant at the 5% level. This effect can be caused by the differenced level of the variable IQ index. For the comparison, we see that the significant sign was kept when using the index, or adding only the Voice and Accountability (VA) factor. For subsample 1 and full sample, the first difference of the IQ index is not applied. As a remark, we state that there are asymmetric factors in the studied countries that drive up the IQ index in order to decrease the NPLs. In this way, finding their strong point (in case of MI countries – the regulatory framework and for HI countries – Political Stability and Absence of Violence) could improve faster the asset quality.

### 3.2.9. Does the commercial trade determine exchange rate impact over NPLs?

In addition to our main hypotheses presented in subsection 2.4, one of the extensions we introduce into our basic model is how the commercial regime (trade) status of the selected countries differs in terms of its influence over NPL ratio. In this subsection, the idea underlying the exchange rate impact is presented and developed, specifying our assumptions when is needed and the reconstruction of the basic model.

The idea behind the hypothesis originates from the division of the commercial trade by its importance to the country. We divide the countries based on commercial trade criteria, explained below, into:

- Export dominant countries (when exports, in total exports plus imports, represent more than the threshold of 41%)
- Import dominant countries (when imports, in total exports plus imports, represent less than the threshold of 41%)

In order to materialize our specification in the model, the data should be collected, but to our knowledge a precise division of export/import dominant countries could not be found. Despite this drawback, a feasible and attractive method, and also at hand, was to set up a threshold for those states exporting more than importing to the world. The data for annual imports/exports from International Financial Statistics (IFS) IMF were used for the threshold construction.

The main criterion for setting the model was based on the simple computation of the share of exports into total exports plus total imports to the world. The observations are complete with the exception of Serbia, having missing observations from 2002 – 2004, which should not affect much our estimations. The calculations cover an interval from 20% as a minimum value, until 65% as for the maximum value. In the next step, the threshold criterion is set up, based on the difference between minimum and maximum values, with a deviation of 3 pp. due to our assumption that, for example, France or the United Kingdom registering a share of exports equal to 41-46% are not believed to be import dominant countries, thus imposing us to downward the threshold level with 3 pp. Therefore, a critical point of 41% and above is a cut off for a country considered to be export dominant and below 41% - an import dominant player on the market[5]. The reconstructed regression model will take the following form:

[5] The export/import dominant countries are included in Table A. 10

$$NPL\ ratio = \beta_1 * gdpG_{it} + \beta_2 * ER_{it} * \boldsymbol{Dum_{Exp}} + \beta_2 * ER_{it} * \boldsymbol{Dum_{Imp}} + \beta_3 * Unempl_{it} + \beta_4 * FDI_{gdp_{it}} + \beta_5 * Edind_{it} + \beta_6 * IQind_{it} + \beta_7 * Infl_{it} + c_i + \varepsilon_{it} \quad (7)$$

where the highlighted dummies are multiplied by the exchange rate variable along with its coefficients, with $Dum_{Exp}$ standing for dummy in case the country is an export dominant taking the value of 1, otherwise equals 0. The same approach defines the dummy for import dominant ($Dum_{Imp}$), the countries take the value of 1 if it is import dominant, and 0 if the country is export dominant.

To distinguish our results applying the FE model with corrected standard errors, the interpretation is deducted from Table A. 11.

### 3.2.10. Remarks of the final results using fixed effect estimation

Using a fixed effects panel model and at the same time controlling for cluster effects on countries, the results are in line with the stylized facts. This confirms that for the full sample, the GDP growth has a negative and significant impact on NPLs, while the examination on the two levels of economic development (subsamples) is rather slightly mixed. The estimated coefficient of the economic growth appears to be insignificant in the developed states, even if the sign is negative as well. The coefficient in case of developing countries is also higher by 10 units than the one for the developed ones. Another negative determinant affecting NPLs is exhibited by the estimator of institutional quality index, significant for the MI countries, compared to HI countries. A similar impact is provided by the FDI (measured as percentage from GDP) contributing to a reduction of the bad quality of bank assets. With a significant level of 1% in the MI countries, this coefficient provides strong incentives to facilitate the business environment and attract investors. An important fact, explained by the model concerns the exchange rate influence on the dependent variable, based on the dominance type in commercial trade. The results show that the impact is different from sample to sample. In case of the full sample, the impact of depreciation of domestic currency has a negative sign in the countries believed to be import dominants, confirming our hypothesis. For the subsamples, a depreciation of domestic currency in the export dominant countries among the developing economies (Ukraine, Romania or sometimes Macedonia) increases the NPL ratio. Regarding the

determinants of NPLs, the fixed effect estimation does not indicate a significant impact of the educational index and inflation rate on NPLs in any of the subsamples or full sample.

### 3.2.11. Empirical literature overview

In the previous subsections, the estimation frameworks took an approach to find the specific effects, either fixed or random, and deal with the last findings of heteroskedastic and serial correlated error terms. As it was mentioned in 3.2.5, the time-invariant variables are not present into our model, meaning that there is one more argument that justifies the choice of the FE model. Moreover, the implementation of the FE estimation enabled us to control for the unobserved heterogeneity, correlated with the explanatory variables. Even though the omitted variable bias and endogeneity problem are controlled in the FE model, the specification, in our case, of the dependent variable NPL ratio, yields non-reliable fixed effect estimator. In this way, a dynamic relationship present in the model, due to inclusion of the lagged dependent variable (and/or lagged independent variables), produces biased and inconsistent estimates, according to Athanasoglou (2005).

Many authors consider the FE model as a starting point in their analysis, allowing to control for the country-specific effects, being a simple and affordable approach (Klein, 2013; Castro, 2012; De Bock and Demyanets, 2012, etc.). Concerning the ratio of NPLs, Jakubik and Reininger (2013) as well as other literature sources (Beck et al., 2013, De Bock and Demyanets, 2012, Louzis et al., 2011), many authors addressed the pattern of persistent NPLs in time, explaining the lagged specification of this financial indicator and the autocorrelation problem. In line with the literature, the dynamic specification into an econometric relationship implies that a variable in a period is correlated with the same variable in the previous periods. The idea is rationalized by the long-run persistence of loans into balance sheet, as they are not excluded from it immediately.

Following the previous studies, to solve the correlation between error terms, to induce the exogeneity of the variables and let the error term behave regardless of the normally distributed assumption, we apply the dynamic generalized method of moments (GMM). Analogous approach is used as a technique for NPLs determinants by various authors because it is an appropriate method to gain efficiency (Beck, 2013; Jakubik and Reininger, 2013) in case OLS, RE or FE estimators for which the

principles of valid estimators do not hold anymore, or the functional form of the maximum likelihood for MLE (maximum likelihood estimator) is not known or is difficult to find.

After estimation of their models by the OLS and FE, De Bock and Demyanets (2012) complement their results with the difference GMM method of Arellano and Bond which corresponds to take the first difference of all variables and the lagged forms, in order to use them as instruments for the exogenous (predetermined) and endogenous variables in the regression. As an extension of the method, for higher precision and lower bias, the system GMM of Arellano and Bover and Blundell and Bond is employed as well.

The same methods are used by Castro (2012), employing the GMM of Arellano and Bond due to biasness in the static models and the efficiency for the relatively small T, introducing in their model lags of order (-1) of the dependent and independent variables.

Espinoza (2010), on the other hand, is emphasizing the efficiency of system GMM compared to difference GMM, as the latter is less important when the autoregressive term is close to 1.

Nevertheless, several drawbacks can occur during the GMM estimation as it is the case of too many specified instruments over the number of the individuals (countries) leading to overidentification problem (Roodman, 2009). Moreover, the weak instruments found for the endogenous variables may mislead the interpretation of results. To avoid the problem, Roodman (2009) advices to keep the number of instruments lower than the number of groups.

### 3.2.12. Dynamic GMM method

In the current subsection of the chapter, another model based on Roodman (2009) algorithm, is introduced and implemented for the fixed effects present in our panel data. In fact, the algorithm is derived from the origins of the method, named the dynamic panel estimator generalized method of moments (GMM) of Arellano-Bond and Arellano-Bover/Blundell-Bond. The motivation of choosing the dynamic GMM is based on (Roodman, 2009): i) short time series and large number of countries; ii) variables are not all strictly exogenous; iii) the presence of fixed effects; iv) the problem of autocorrelation and heteroskedasticity for the selected data; v) the dependent variable takes a dynamic form which depends on its past information; v)

the linear functional form of variables. Hence, our motivation is much stronger when the transformed dependent variable in the FE model took the form of 3 lags, due to unit root presence.

Following Greene (2012), we briefly introduce the main structure and properties of GMM, and later of the dynamic panel. As such, the method of moments theory, developed in econometrics, defends an important assumption and supported by the law of large numbers theorem, that a sample "moments" or sample statistics (mean, variance) of a finite sample will converge under probability limit to one constant:

$$plim\ \bar{m}'_k = \frac{1}{n}\sum_{i=1}^{n} y_i^k = \mu'_k = E\left[y_i^k\right] \qquad (8)$$

The derivation of parameters estimation $\theta_1, \dots, \theta_K$ originates from the moment estimators, being consistent by the Slutsky theorem. The K number of parameters will equal the number of equations and parameters K are the unknown ones to be estimated as a function of $\hat{\theta}_k = \hat{\theta}_k\ [\bar{m}'_1, \dots, \bar{m}'_k]$. Nevertheless, as it is specified with the examples by the author, the method of moments is efficient for exponential distributions, but as moving further, we see the moments are essential for minimizing the criterion function of the minimum distance estimator:

$$q = [\bar{m}_n - \boldsymbol{g}(\boldsymbol{\theta})]'\, \boldsymbol{W}\, [\bar{m}_n - \boldsymbol{g}(\boldsymbol{\theta})]' \qquad (9)$$

where $\boldsymbol{\theta}$ is the vector of K ≤ L, L the number of equations and $\boldsymbol{W}$ is any positive definite weighting matrix. For any choice of *W*, an optimal matrix must be found to for the covariance matrix of the minimum distance estimator. As an extension of the introduced estimator, the GMM provides another important condition, that of orthogonality. It would be much simpler to employ the instrumental variable method only, but the case of instrumental variable estimator does not consider the situation when the moment equations are more than parameters to be estimated. In the just identified case when the K = L, the weighting matrix is useless. In general, the GMM estimator is obtained in two steps:

*Step 1:* obtain the covariance matrix: $\widehat{\boldsymbol{V}} = \frac{1}{n}\sum_{i=1}^{n} e_i^2 \boldsymbol{z_i} \boldsymbol{z'_i}$ ; where $\boldsymbol{z_i}$ vector of instrumental variables

*Step 2:* calculation of the inverse of the matrix from step 2, used for GMM estimator.

Therefore the GMM coefficient is found by minimizing the criterion function with respect to parameter vector $\theta$ (Greene, 2012):

$$\hat{\theta}_{GMM} = \left[\left(\sum_{i=1}^{n} W'_i H Z_i\right)\left(\sum_{i=1}^{n} V'_i H' \hat{\varphi} H Z_i\right)^{-1}\left(\sum_{i=1}^{n} Z'_i H' W\right)\right]^{-1} * \\ * \left(\sum_{i=1}^{n} W'_i H Z_i\right)\left(\sum_{i=1}^{n} Z'_i H' \hat{\varphi} H Z_i\right)^{-1}\left(\sum_{i=1}^{n} Z'_i H' W\right) \qquad (10)$$

where $\hat{\varphi} = \hat{e}_i \hat{e}'_i$; and $e_i$ - the disturbance term; the matrix $Z_i$ is the matrix of instrumental variables, which in our case for the lagged dependent variable NPL ratio, takes the form:

$$\boldsymbol{Z_i} = \\ = \begin{bmatrix} NPL_{04-02} & d_{06} & 0' & 0 & 0' & 0 & \dots & 0' & 0 \\ 0' & 0 & NPL_{05-02} & d_{07} & 0' & 0 & 0' & 0' & 0 \\ 0' & 0 & 0' & 0 & NPL_{06-02} & d_{08} & \dots & 0' & 0 \\ 0' & 0 & 0' & 0 & 0' & 0 & 0' & 0' & 0 \\ 0' & 0 & 0' & \vdots & \vdots & 0 & \ddots & 0' & 0 \\ 0' & 0 & 0' & 0 & \dots & 0 & 0' & NPL_{10-02} & d_{12} \end{bmatrix}$$

where the $NPL_{t1-t0}$ stands for the NPL including the respective years (by 04 we denote e.g. year 2004 and the $NPL_{04-02}$ includes 2004, 2003, 2002 years); while $d_t$ is the time dummy. Due to the orthogonality condition - ($E[\varepsilon \mid z] = 0$) for all of the variables (dependent and independent), the instrumental variable estimator must satisfy the overidentification condition. For our data, considering the lag = 3 for the dependent variable only, it implies in total 49 instrumental variables. This is explained by the usable time series left, spanning from 2006 - 2012 as t = lag + 2 + … + 11, and the number of years then equals seven.

Concerning the lagged independent variables, we believe their presence is not highly required as the effect of the NPL ratio is much more drastic and specific, reflected in banks' balance sheets, in line with the literature investigations from section II. Moreover, the process of stationarity result brought itself the information about how long the recent information is linked to its past. We also keep tightly with the econometric settings to avoid overidentification problem and to obtain valid instruments. Even though the GDP, institutional factors or the unemployment rate would lose their exogeneity due to non-included lagged forms, our results would not

be reliable, and we would lose the properties of the GMM estimator. However, we agree that e.g. unemployment rate, GDP growth or institutional index would weaken their influence, but for our annual frequency, this should not threaten our results very much in contrast to quarterly or monthly frequency with the seasonality effect. We interpret the exogeneity of the unemployment rate from the viewpoint of an immediate change of the NPLs due to unemployment rate improvement (or deterioration). This means the past information of unemployment rate has already incorporated its effect into lags of NPL ratio, and its presence as an IV is redundant. As it was mentioned previously, the motivation of having introduced lags to the dependent variable is the cause of a slow adjustment in banks' balance sheet due to NPLs improvement in time or the long-term lending. Therefore, the slow improvement (deterioration) of the labor market is already reflected in the slow adjustment of NPLs, controlling for its effect (upward or downward), the reason for which we use the current form for the rest of the regressors.

Having introduced the GMM process, we can control for the entire information captured in our lagged variable, recognized by the main equation:

$$y_{it} = \boldsymbol{X'_{it}}\boldsymbol{\beta} + \delta y_{i,t-1} + c_i + \varepsilon_{it} \qquad (11)$$

where the right-hand side variables include all the regressors $\boldsymbol{X}c_i{}'_{\boldsymbol{it}}$ and the lagged dependent variable $y_{i,t-1}$, the country specific fixed effects $c_i$ and the error term $\varepsilon_{it}$.

According to the literature (Roodman, 2009; Greene, 2012), using the within transformation or FE model for the dynamic panel causes a dynamic panel bias. Under the within procedure the lagged dependent variable $y_{i,t-1}$ ($NPL_{i,t-1}$) and the error term $\varepsilon_{i,t-1}$are negatively correlated, only if T (number of periods) is not so large, in which case the term that correlates them would become insignificant and the estimator consistent. In our case, the number of years is not considered large enough to accept the results from FE model only, pointing to another argument for choosing GMM.

To estimate the model (11) with a higher precision, the difference GMM method proposed by Arellano-Bond is developed, an extension of Anderson and Hsiao method, to first-differentiate the equation and use the lags of the dependent variable as instruments for the endogeneity problem. In this way, the first-difference equation takes the form, where the fixed effects are removed (Greene, 2012):

$$y_{it} - y_{i,t-1} + = \left(\boldsymbol{x_{it}} - \boldsymbol{x_{i,t-1}}\right)' \boldsymbol{\beta} + \delta(y_{it-1} - y_{i,t-2}) + \left(\varepsilon_{it} - \varepsilon_{i,t-1}\right) \qquad (12)$$

One disadvantage of the difference GMM is that employing it for short time series (small T), the estimation is not very precise and for this reason, another method, system GMM developed by Arellano-Bover/Blundell-Bond is more appropriate. The system GMM uses the regression of differenced variables and of the variables in level, meaning that the strictly exogenous variables that are first differenced, are instrumented in the same form and the dependent, as well as the endogenous variables, use their lags as an instrumental variable for themselves (Klein, 2013). Nevertheless, a consistent GMM estimator should be performed in line with the assumptions of no second order of serial correlation and strong and valid instruments (Blanco and Gimeno, 2012). Referring to this, we apply a dynamic model using GMM for both sub-samples and the full sample, paying attention to stationarity issues considered in the first part of this section.

For the beginning, the general dynamic model for our panel data is expressed as:

$$NPL_{it} = \boldsymbol{X'_{it}}\,\boldsymbol{\beta} + \sum_{s=1}^{s} \delta NPL_{i,t-s} + c_i + \varepsilon_{it} \qquad (13)$$

where the variables treated to be strictly exogenous are the GDP growth, exchange rate, unemployment rate, institutional quality index and the educational quality index, while inflation rate and FDI are considered to be endogenous, for which good instruments must be found. The inflation rate is conditioned by the Central Bank, viewed as a result of a government's actions to establish the economic equilibrium providing the control of monetary policy. The growth rate of M2 or the aggregated money supply is used as an instrumental variable for the inflation rate, seen as a proxy for the issued money in the economy. The data is taken from World Bank, measured as annual growth rate, in percent. Another endogenous explanatory variable concerns the FDI, as a percentage from GDP, which depends on a variety of factors, but due to unavailable data, the trade openness of a country was chosen being highly

**Table 3.2.3: GMM estimation on full sample with dummy specifications**

| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| L.npl | 0.631*** | 0.584*** | 0.779*** | 0.596*** | 0.528*** | 0.594*** | 0.581*** | 0.526*** |
| | (0.0677) | (0.133) | (0.181) | (0.0784) | (0.125) | (0.0764) | (0.0953) | (0.132) |
| L3.npl | -0.182** | -0.122* | -0.150 | -0.0778 | -0.0393 | -0.0975 | -0.0548 | -0.0460 |
| | (0.0802) | (0.0678) | (0.100) | (0.0888) | (0.0650) | (0.0877) | (0.113) | (0.0606) |
| gdpg | -0.0383*** | -0.0338** | -0.0572*** | -0.0362*** | -0.0310*** | -0.0384*** | -0.0324*** | -0.0310*** |
| | (0.00972) | (0.0128) | (0.0169) | (0.00817) | (0.00860) | (0.00903) | (0.00966) | (0.00891) |
| er | 0.00497** | 0.00439 | -0.00962 | | 0.00242 | | | 0.00256 |
| | (0.00230) | (0.00272) | (0.0144) | | (0.00164) | | | (0.00180) |
| unempl | 0.0337 | 0.0548** | -0.00844 | 0.0530*** | 0.0569*** | 0.0498** | 0.0524*** | 0.0566*** |
| | (0.0224) | (0.0208) | (0.0350) | (0.0136) | (0.00903) | (0.0182) | (0.0169) | (0.00923) |
| fdi_gdp | -0.0129 | -0.0161 | -0.00474 | -0.0348* | -0.0321* | -0.0302* | -0.0362 | -0.0296 |
| | (0.0148) | (0.0115) | (0.0171) | (0.0192) | (0.0172) | (0.0157) | (0.0227) | (0.0177) |
| infl | -0.00899 | -0.0114 | 0.0140 | 0.00410 | 0.00376 | 0.00404 | 0.00353 | 0.00306 |
| | (0.0150) | (0.0293) | (0.0203) | (0.0242) | (0.0231) | (0.0250) | (0.0257) | (0.0233) |
| edind | 4.310*** | 4.521 | 8.591** | 5.300** | | 5.657** | 4.454** | 5.544 |
| | (0.998) | (2.675) | (3.788) | (1.955) | | (2.229) | (2.007) | (3.196) |
| iqind | -3.303*** | -3.433 | -1.418 | -4.098** | -4.429 | -4.529* | -3.407* | |
| | (0.993) | (2.456) | (1.350) | (1.794) | (2.872) | (2.202) | (1.887) | |
| erDE | | | | -0.0847** | | | | |
| | | | | (0.0388) | | | | |
| erDI | | | | -0.000265 | | | | |
| | | | | (0.00292) | | | | |
| edind_HI | | | | | 5.594* | | | |
| | | | | | (3.170) | | | |
| edind_MI | | | | | 5.305* | | | |
| | | | | | (2.646) | | | |
| erDE_HI | | | | | | -0.0200 | | |
| | | | | | | (0.108) | | |
| erDE_MI | | | | | | -0.0862** | | |
| | | | | | | (0.0309) | | |
| erDI_HI | | | | | | | 0.193 | |
| | | | | | | | (0.172) | |
| erDI_MI | | | | | | | 0.00228 | |
| | | | | | | | (0.00231) | |
| iqind_HI | | | | | | | | -4.425 |
| | | | | | | | | (2.942) |
| iqind_MI | | | | | | | | -4.735 |
| | | | | | | | | (3.559) |
| Observations | 112 | 112 | 98 | 112 | 112 | 112 | 112 | 112 |
| N (country) | 14 | 14 | 14 | 14 | 14 | 14 | 14 | 14 |

*Source:* Author's elaboration; Stata xtabond2 command; Note: model (1) by one-step difference GMM; model (2) by one-step system GMM; next models by two-step system GMM; erDE/erDI stands for exchange rate * dummy for export/import dominant countries; edind_HI/edind_MI is the education index in HI/MI groups (analogous for iqind_HI/MI); N - number of countries; variables are used with their transformations according to Table 3.4

correlated with the analyzed variable. Trade openness stands for the share of commercial balance (exports plus imports) in total GDP, measured in US dollars.

An initial estimation of the basic model, regressed for the full sample, is done with the difference GMM (model 1, Table 3.6), one-step approach, that implies the error terms are homoscedastic for the countries. For a consistent interpretation of results, the autocorrelation test and the Sargan test for overidentification are provided automatically after the regression.

One of the problem that appears in our sampled countries, divided into subsamples of 7 states per each, is the lower number of countries than the number of years (N < T) for which system GMM is not a valid estimator. Later, the full sample only is considered, due to the problem of omitted variables. In this case, the system GMM is applied on subsamples. For the purpose of our new hypotheses, we include in our model some dummies to interact with the variables from the main research questions.

For a consistent estimator, the two-step methods and robust standard errors are specified, to relax the assumptions of independence and homoscedastic error terms (Janvisloo and Muhammad, 2013). Therefore, the two-step system GMM is employed, releasing the results from model 3, Table 3.6 above. For an accurate estimation, the results of autocorrelation and valid instruments tests are reported in Table A. 12.

According to Beck et al. (2013), as it is required by the autocorrelation test, the serial correlation of AR(1) is allowed to persist in the model, but this is not characteristic for the second order autocorrelation AR(2). The first order serial correlation of error terms imply that the lagged values are endogenous when taking the first differences of the equation. As such, because the Arellano-Bond test uses the differenced error terms $\varepsilon_{it} - \varepsilon_{i,t-1}$ to detect the correlation between the difference and the $\varepsilon_{i,t-1}$, it is enough the case of accepting the null hypothesis of no second serial correlation (Roodman, 2009). Fitting the number of lags for the instrumented variables must be in concordance with the overidentification problem - of including too many instruments.

The number of lags of NPL ratio, included as the predetermined variables, are chosen to be the first ($NPL_{i,t-1}$) and the third ($NPL_{i,t-3}$) due to multicollinearity specified previously. While checking the p-value for Arellano-Bond test of AR(1), we conclude that for the final result of system GMM, the null hypothesis of no serial

correlation is rejected and for AR(2) test the null is accepted, meaning the estimators are valid. Concerning the Sargan test, the p-value = 0.165 indicates us to accept the hypothesis that the instruments are valid and that there is no overidentified specification. This also holds for the extended hypothesis, regarding the exchange rate, with the results from model 4.

Providing the small number of countries and relatively large number of time series, we interact the main hypotheses with dummy variables for the 2 groups of countries - (HI) and (MI). The new generated variables are included in the model 5 for educational index, in the $6^{th}$ model for the exchange rate, in case the country is an export dominant and if the country is import dominant (model 7). The last model incorporates the variables generated for the institutional quality index.

# IV. Robustness checks

Before switching to interpretation of our results and providing an analysis towards the empirical findings, some robustness checks will be performed. In order to make the results more reliable, in case of some hypotheses to be tested, different variables or IVs are added.

For the beginning, a new variable is used instead of the educational index – percentage of the total population enrolled in tertiary education, or school enrollment (% gross), which is the highest level of education. The data is taken from the World Bank, but we must exclude Germany from our sample due to missing observations for the whole time period. Using the extended model 4 (Table 3.6), the variable *sen* (school enrollment) represents the proxy for educational index and the results are shown in Table B. 1, model 1. The exogeneity of this variable is viewed differently from the educational index, as the percentage of enrolled population at the moment in the tertiary level depends on the birth rate in a country.

Given the availability of the data, live births (by mother's age at last birthday and legal marital status) are used, taken from Eurostat database with missing observations for year 2002. Therefore, we instrument the school enrollment variable with the live births and the estimated coefficients are shown in Table B. 1, model 2.

Another choice for building up our final results is associated with the increasing NPL ratio, in case of an increase in one unit of educational index, because as it was mentioned in our hypothesis, the education level may be linked to the amount of credit borrowed by highly educated clients. To check if the strict exogeneity holds, the system GMM model uses, as an IV for the educational index, the proxy for credit growth - domestic credit to private sector (% of GDP). The available data from World Bank helps us employ the method, including the credit growth as the instrumental variable for the educational index, but which may be also correlated with the NPL ratio.

To better understand the empirical results and their reasoning, a separate chapter for analysis of the results is developed, in the last part of the present work.

# V. Interpretation of results

At the beginning, for choosing an appropriate method to estimate our basic model, the unit root presence is tested, leading to a transformation of the variables in order to avoid the spurious regression. This is decisive for our estimators as it can change the results drastically, but the preliminary results of the applied technique are in line with the stylized facts of the two NPL determinants (GDP growth and the unemployment rate), meaning the problem of flawed estimation is highly diminished. In the next part of the econometric work, the simple panel models are employed (pooled OLS, fixed effect model, random effect model) to notice if our countries manifest a significant difference between their observations, if there are country-specific effects or random effects. The basic model shows the FE model fits best for the analyzed data for both subsamples, and the results (Table A. 7) produce substantial differences between them. While the determinants for MI countries (subsample 1) are significant for NPL ratio, with the exception of the educational index and inflation rate, the determinants for HI countries (subsample 2) are not significant at any level with the exception of the unemployment rate, being significant at 5% level. For the extended model, which assumes the export/import dominancy can determine the influence of exchange rate over NPLs, the full sample estimation expresses a high significance in case of import dominant countries which displays an unexpected sign of decreasing the NPL ratio if the domestic currency depreciates (or exchange rate depreciates). Even so, when splitting the sample, only for the export dominant countries among MI countries, the coefficient of exchange rate depreciation is significant, bearing a negative sign as well, being in opposite to our hypothesis.

Despite of biased and inconsistent estimates implementing the FE model, which is a simple OLS regression with control for the heterogeneity effect, we migrate to a more advanced tool in econometrics, due to problems of autocorrelation and endogeneity which can be relaxed and solved by GMM methods. Consistent with our literature review, the lagged NPL ratio is an important part of the regression model to attest for

persistence of NPL in time, based on the arguments of long-term lending, payment schedule and persistence of the item in the banks' balance sheet.

Employing System GMM[6], we analyze the results mainly from the full sample (Table 3.6, model 4) and for the main research questions of this work, additional specifications are added to the model in order to attest the differences between the 2 groups only. In this way we perform separately the system GMM including dummy variables for the MI/HI countries in case of the educational index, exchange rate for import/export dominant countries and the institutional quality index. For the rest of the variables, the model is employed with the specified dummies, and the difference from the full sample is evidenced only for the FDI determinant not detecting any significance for both groups (Table A. 13), but for which the instruments were not valid or weak, based on Sargan test. The coefficients of GDP growth, unemployment rate and the exchange rate did not show any major deviation from the full sample. In line with the literature done for other panel studies, our results are similar for the GDP growth (Louzis, 2011; Nkusu, 2011; Beck et al. 2013; Jakubík and Reininger, 2013), unemployment rate (Louzis, 2011; Nkusu, 2011) and the exchange rate (De Bock and Demyanets, 2012; Klein, 2013; etc.).

Under all specifications (Table 3.6, model 4-8), the *NPL ratio lagged* 1 year back, is statistically significant at 0.1% confidence level, meaning that the set of new information from the past, influences our dependent variable. The coefficient of the lagged NPL loses its influence, if the lags more than 1 or 2 years are included in the equation, due to a faster convergence of the asset quality indicator to its satisfactory level of good quality. Concerning the coefficient of convergence, $\delta$, a high value of more than 0.50 proves the persistence of NPL at an upper medium extent. It can be caused by the spreading effect of inability of the borrower to pay the loan in case of an economic shock in the economy, but which at the aggregate level, the $\delta$ estimator may display a higher value. The fact of the negative sign of NPL lagged 3 years back derives from the reclassified non-performing loans as performing loans motivating lenders to become optimistic in future regarding the payment of loans, thus a good historical payment of

[6] It should be noted that system GMM is estimated only for the full sample because it does not allow us to estimate at once the subsampled countries due to a lower number of countries (7) than the number of years (11), omitting our main variables from the model.

credits leads to an increase of credit amount but which increases, actually, the actual NPL ratio.

The *GDP growth* variable, measuring the economic environment conditions, is negatively significant for NPL ratio under all specifications. At 10% level with low standard errors, an increase of one unit in GDP growth determines the NPL ratio to decrease by 3.62% for the full sample, and the coefficient is stable for the other regressions. During good times, strong evidence is found that agents are likely to repay their loans and therefore the NPL ratio is reduced. It is interesting to notice that by employing the FE model, a much lower influence is found to affect the NPL ratio in the HI countries than in the middle ones, and the coefficient is not even significant at any level. Even so, the system GMM does provide a significant estimator of this determinant.

Regarding our hypothesis of the *exchange rate* influence on NPLs and whether a country is an export dominant or import dominant, we can conclude that the commercial trade matters only for the MI countries which are considered as export dominants. The FE model offers the same result with GMM method, in case of subsamples having provided the dummies for countries interacted with the exchange rates. The expected sign is confirmed given that a depreciation of the domestic currency decreases the NPL ratio. As such, in case of the full sample and when comparing the 2 subsamples, we find the same significance and sign of the determinant is spotted among the export dominant middle income countries. Even if for the import dominant countries, the depreciation of the exchange rate estimator is not significant, the sign of it is in line with our hypothesis. It means that a depreciation of the exchange rate (depreciation of domestic currency) for the import dominant states should have led to an increase of the NPL ratio due to a loss of value when buying foreign currency and less valued domestic currency. Hence, a successful macroeconomic goal to reduce the financial stress can arise in the MI countries that are more export-oriented based on their struggling efforts to grow faster. For this reason, the HI economies may concentrate on other determinants or fields, but need to keep up with a stable and smoothed trend of commercial trade. The result may not show a real significant impact as the selected states (European Union members) are mainly opened and motivated by free tariff and import quotas to trade between them and

because some share a common currency (EUR/USD) which may display a biased estimation for this subsample.

*Unemployment rate* is confirmed to have a positively significant impact on NPLs consistent with the GMM method and FE estimation. At a 1% level, an increase in one unit of unemployment rate, the NPL ratio increases by around 5%.

The determinant of NPL, the *foreign direct investments*, instrumented by trade openness has a negative significant impact, but at 10% level. Comparing to GMM, which did not show any difference between the subsamples in terms of significance, the FE estimation concluded that FDI is a significant determinant for NPLs in the MI countries, bearing a negative sign. This fact can be supported by increasing availability of investors in the emerging markets where labor force is relatively not as expensive as in the developed economies. However, we base our results on the GMM method and conclude that in general (full sample), a rising investment flow by one unit can reduce the NPL ratio by 3%.

Comparing to FE model, including the extended specification, the *inflation rate* is statistically insignificant even under different specifications, including the separated subsamples. Our expectation of ambiguous effect of this determinant is denoted, emerging from the construction of this variable as it is an average of consumer prices based on goods and services indexed each year causing an uncertain impact over NPL ratio. Moreover, beside its strong insignificance to the model (Table A. 13), meaning the coefficient is not different from zero, it took a negative form for both countries, that could be interpreted as the ability of a faster repayment due to increasing value of the collateral, or the anticipated inflation of banks' management.

A slight surprising result of the *educational index* after FE estimation arouse under the GMM method which released a statistically significant estimator, confirming our hypothesis of a positive sign of educational level over the NPL ratio. The reason for NPLs upsurge, explained by the educational level, refers to overestimation of the bank managers to lend to highly educated or intelligent clients more than to lower educated clients motivated either by better abilities to diversify their financial resources, as having a highly paid salary, or solid plans of this class to borrow more credits. Nevertheless, we are referring to the NPL term, not to a credit default probability or the last classifications

(doubtful and loss). We suppose, in case of a research to this extent, where educational level can explain the credit default, the level of education would change its sign for a negative impact on the bank's loan losses. Having concluded this, an important incentive for the financial system, would be to provide reforms for monitoring how banks manage their highly educated clients as their request for credits is also higher, or insert this factor in the financial stress tests. If we divide the sample into subsamples and interact the dummies for countries with this determinant, we do not find any major difference between the two subsamples, but for which the coefficient becomes significant at 10% level, meaning education level can be one of the determinants increasing the NPLs.

Another determinant for NPL ratio is the *institutional quality index* which is found to be negatively significant for the full sample, but insignificant for the both subsamples due to small subsample size. Even so, when regressing separately on each institutional component, almost all coefficients of them are significant and in line with the expected negative sign. As the purpose of the thesis is not focused primarily on institutional factors, we provide only the belief that using an averaged institutional index can give a better result only if the sample may be expanded, otherwise a single index may not bring additional value to the empirical researches.

The results are sustained by the robustness checks estimated for the full sample only (Table A. 1) concerning the proxy for the educational index – the school enrollment, that supports our result of positive influence over NPL ratio in case of increasing education participation in a country. Under the reasoning that a higher participation into education is motivated by higher birth rate, after instrumenting the school enrollment variable with the live births, the results are consistent with the previous mentioned. Another IV used – the credit growth of the private sector, explains there is a correlation with the educational index due to a lowered significance, and smaller coefficient of the index, in the model where the credit growth is used as an IV.

# VI. Conclusion

This thesis demonstrates the importance of determinants of NPLs at the aggregate level, including economic and non-economic variables. We focused on the educational index, exchange rate of import/export dominant countries and a single institutional index, averaging all six indicators of ICRG. In order to objectively respond to the main research questions, several macroeconomic variables are controlled for in the empirical model: GDP growth, exchange rate, unemployment rate, FDI and inflation rate.

Firstly, the static models for the panel dataset were applied, where the fixed effects were detected. In the next part, we decided to base our results on the system Generalized Method of Moments, due to dynamic features and persistence of NPL ratio in time, heteroskedasticity and endogeneity problems, as well as the small sample bias. Even so, this methodology did not afford us to implement it on separated subsamples. To overcome the problem, the dummy variables for Middle and High income groups were interacted with the variables of our interest. Nevertheless, this detail remained for a future work in case the data is found and the sample size is increased in both regions.

Unlike many other studies highlighting the main determinants of NPLs, we used a slightly extended time series with different countries providing a comparative approach for them. We also built the educational index when the data was not available, for which a previous empirical investigation, related to our dependent variable, does not exist.

Regarding the final results, on one side, various other studies find the same impact over NPLs in case of GDP growth (Louzis, 2011; Nkusu, 2011; Beck et al. 2013; Jakubík and Reininger, 2013), unemployment rate (Louzis, 2011; Nkusu, 2011) and the unspecified exchange rate (De Bock and Demyanets, 2012; Klein, 2013; etc.). On the other side, authors as Basu (2002), Coviello (2003), Bojnec and Ferto (2009) and Papaioannou (2007) construct a composite of the institutional indicators, either based on simple average, equally weighted average or principal components (ICRG, Kaufmman indicators), that had a significant impact onto their analyzed hypotheses. Despite different tools to comprise this indicator in the mentioned works, this thesis concludes

that the significance of institutional quality index is not consistent across all specified models, but for which a negative influence over the dependent variable is noticeable.

Another point, in contrast to several researches (Baselga-Pascual et al., 2013; Moinescu and Codirlaşu, 2011; Nkusu, 2011; Klein, 2013), is that we do not evidence a significant impact of the inflation rate on NPLs, neither when using the fixed effects model, nor the GMM methods. However, Castro (2012) observed, as well, the insignificant impact of this determinant on risk management practice in the GIIPS banks.

In case of FDI, other few studies associated the investments to significantly decrease the NPL ratio (De Bock and Demaynets, 2012; Festic and Beko, 2008; Kavkler and Festic, 2010), but in our work, when the groups are divided, this determinant is not found to be important compared to the full sample result where the economic or regional classification is not taken into consideration. The effect of FDI is rather unstable, implying the policy makers must establish long-term monetary, regulatory, social reforms in middle and high income states to attract investments which are more constructive for the economy.

A curious research question was related to the education, assumed to determine the acceleration of NPLs because, to the best of our knowledge, a previous estimation of this variable is not evidenced in any study conducted before. It is interesting enough to find that it is a significant determinant at the 10% level for both groups of classified countries, having a positive impact on the NPL ratio. The results are supported by the robustness checks, employing a proxy variable for it - school enrollment of tertiary education. Our interpretation concluded that in case of a research to this extent, the level of education would change its sign for a negative impact on the bank's loan losses if the credit default or losses are used as dependent variables.

Finally, it is worth mentioning that the contributions of this thesis, including the mentioned findings above, could be of a policy interest for the banking's credit risk analysis based on the significance of the education level, for the macroeconomic stress tests estimated by regulators and supervisors who may include the commercial trade dominancy status into their scenario analysis, and for the other research works trying to compress the institutional indicators into a single index.

# Bibliography

Ahmad N. H. and Ariff M. (2007): Multi-country study of banks credit risk determinants, Volume 5 Issue 1, Article 6; International Journal of Banking and Finance

Aizenman J. and Noy I. (2012): Macroeconomic adjustment and the history of crisis in open economies, 2012, http://www.voxeu.org (Research Based Policy Analysis and commentary from leading economists), 21 November 2012

Athanasoglou P. P., Brissimis S.N. and Matthaios D. D. (2005): Bank-specific, industry-specific and macroeconomic determinants of bank profitability; Working paper No. 25; Bank of Greece

Babihuga R. (2007) Macroeconomic and Financial Soundness Indicators: An Empirical Investigation; IMF Working Paper ; WP/07/115

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

Balduini A. (2013): Asset Quality Review: Setting the foundation for a standard stress testing framework – Moody's Analytics

Barisitz S. (2013): Nonperforming Loans in Western Europe – A Selective Comparison of Countries and National Definitions, Focus On European Economic Integration Q1/2013

Baselga-Pascual L., Trujillo-Ponce A., Cardone-Riportella C.(2013): Factors Influencing Bank Risk in Europe: Evidence from the Financial Crisis; Nº 722, ISSN: 1988-8767

Basu S.R. (2002): A new way to link development to institutions, policies and geography; United Nations New York and Geneva; Policy issues in international trade and commodities study series no. 37

Baum Ch. F. (2006): An introduction to modern econometrics using Stata; Stata press

Beck R., Jakubík P. and Piloiu A. (2013): Empirical Determinants of Non-performing loans ECB Financial Stability Review, Special Feature, No. 1515 February 2013

Belgrave A., Guy K. and Jackman M. (2012) Industry specific shocks and non-performing loans in Barbados; The Review of Finance and Banking Volume 04, Issue 2; Pages 123—133 S print ISSN 2067-2713, online ISSN 2067-3825

Blanco R. and Gimeno R. (2012): Determinants of default ratios in the segment of loans to households in Spain; No. 1210

Bloem A. M. and Gorter C.N. (2001): The Treatment of Nonperforming Loans in Macroeconomic Statistics

Bloem A. M. and Freeman R. (2005): The Treatment of Nonperforming Loans – SNA/M1.05/21, IMF

Bock R. De and Demyanets A. (2012): Bank Asset Quality in Emerging Markets: Determinants and Spillovers; IMF Working Paper Monetary and Capital Markets WP/12/71

Bojnec Š. and Ferto I. (2009): The institutional determinants of bilateral Agricultural and food trade; Applied Studies in Agribusiness and Commerce; Agroinform Publishing House, Budapest

Boudriga A., Taktak N. B. and Jellouli S. (2009): Banking supervision and nonperforming loans: a cross-country analysis; Journal of Financial Economic Policy Vol. 1 No. 4, 2009 pp. 286-318

Boudriga A., Taktak N. B. and Jellouli S. (2010): Bank specific, business and institutional environment determinants of banks nonperforming loans: evidence from mena countries; Working Paper 547, The Economic Research Forum (ERF)

British Bankers' Association response to EBA Draft Implementing Technical Standards on supervisory reporting on forbearance and non-performing exposures (CP/2013/06) under article 95 of the draft Capital Requirements Regulation (2013)

BSTDB (Black Sea Trade and Development Bank) (2011): Ukraine Country strategy 2011-2014, BSTDB, Policy Paper Series [PP/02/2011], March 2011

Cameron A. C. and Trivedi P. K. (2005): Microeconometrics - Methods and applications, Cambridge University Press

Castro V. (2012): Macroeconomic determinants of the credit risk in the banking system: The case of the GIPSI, NIPE WP 11/ 2012, University of Coimbra and NIPE, Portugal

Chennamaneni P., Echambadi R., Hess J. D., Syam N. (2008): How Do You Properly Diagnose Harmful Collinearity in Moderated Regressions?

Coviello D. (2003): Instrumental variables regressions in growth, geography and institutions: reconsidering some results

Dreca N. (2012) Evaluation of Financial Performance of Banking Sector: Evidence from Bosnia and Herzegovina, Croatia, Serbia and Slovenia; Journal of Economic and Social Studies

ECB (2013): A macro stress-testing framework for bank solvency analysis, Monthly Bulletin August 2013

EMC (2013) Non-Performing Loans in South Eastern Europe; Ensight Management Consulting

EMF (European Mortgage Federation) (2010): 2010 EMF Methodological note and survey on Non-Performing Loans in the EU

Ernst & Young Report (2013): EY Eurozone Forecast, Outlook for financial services

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

Fainstein G. and Novikov I. (2011) The Comparative Analysis of Credit Risk Determinants In the Banking Sector of the Baltic States; Review of Economics & Finance; Article ID: 1923-7529-2011-03-20-26

Festic M. and Beko J. (2008): The Banking Sector and Macroeconomic Indicators: Some Evidence for Hungary and Poland; NG , št. 5–6/2008 Strokovni članki – Professional papers

Festic M. and Romih D. (2008): Cyclicality of the banking sector performance and macroeconomic environment in the Czech Republic, Slovakia and Slovenia, Prague Economic Papers

Fofack M. (2005): Nonperforming loans in Sub-Saharan Africa: causal analysis and macroeconomic implications, World Bank Policy Research Working Paper 3769

Głogowski A. (2008): Macroeconomic determinants of Polish banks' loan losses – results of a panel data study, National Bank of Poland, Working Paper Nr. 53

Greene W. H. (2012): Econometric analysis, New York University; Prentice Hall

Groendahl B. and Suess O. (2013): Bloomberg "German Banks Warn ECB Against Stricter Rules in Asset Review"

Haas R. de and Knobloch S. (2010): In the wake of the crisis: dealing with distressed debt across the transition region, European Bank for Reconstruction and Development, EBRD, Working Paper No. 112

Human Development Report (2013): The rise of the South Human Progress in a Diverse World; Technical Notes. Calculating the human development indices

Janvisloo M.A. and Muhammad J. (2013): Non-Performing Loans Sensitivity to Macro Variables: Panel Evidence from Malaysian Commercial Banks; American Journal of Economics 2013, 3(5C): 16-21, DOI: 10.5923/c.economics.201301.04

Jakubík P. and Reininger Th. (2013): Determinants of Nonperforming Loans in Central, Eastern and Southeastern Europe, Focus On European Economic Integration Q3/2013

Jimenez G. and Saurina J. (2006): Credit Cycles, Credit Risk, and Prudential Regulation; Banco de Espana

Jose A. S. and Georgiou A. (2008): Financial soundness indicators (FSIs) framework and implementation; IFC Bulletin No 31, Bank for International Settlements

Kavkler A. and Festic M. (2010): The banking sector in the Baltics; Banks and Bank Systems, Volume 5, Issue 3

Khemraj T. (2009) The determinants of non-performing loans: an econometric case study of Guyana; Munich Personal RePEc Archive Paper No. 53128

Kirchner R., Giucci K., Rubia C. de la and Kravchuk V. (2011): The Banking Sector in Ukraine: Past Developments and Future Challenges, Institute for Economic Research and Policy Consulting

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

Köhler M. (2012): Which banks are more risky? The impact of loan growth and business model on bank risk-taking, Discussion Paper Deutsche Bundesbank No 33/2012

Lee Y., Nedler J. A., Pawitan Y. (2006): Generalized Linear Models with random effects: Unified Analysis via H-likelihood; Chapman & Hall/CRC Taylor & Francis Group

Liao W. (2008): Structural Breaks in Panel Data Models: A New Approach

Louzis D. P., Vouldis A. T. and Metaxas V. L. (2011): Macroeconomic and bank-specific determinants of non-performing loans in Greece: A comparative study of mortgage, business and consumer loan portfolios, Journal of Banking & Finance, doi:10.1016/j.jbankfin.2011.10.012

Mahmudi S. (2013) Correlation between Growth Rate and Non Performing Retail Loans in the Republic of Macedonia; International Journal of Academic Research in Accounting, Finance and Management Sciences; Vol. 3, no. 3, July 2013, pp. 133-138; ISSN: 2225-8329

Mattila V. (2011) FFI Federation of Finnish Financial Services; "Healthy financial sector bolsters Finnish economy"

Moinescu B.-G. and Codirlaşu A. (2011): Lending, economic growth and nonperforming loans: empirical evidences from the new EU member states

Neter J., Wasserman W. and Kunter M. (1990): Applied Linear Statistical Models, 3rd ed., Irwin

Nkusu M. (2011): Nonperforming Loans and Macrofinancial Vulnerabilities in Advanced Economies; IMF Working Paper WP/11/161

Papaioannou E. (2007): What Drives International Financial Flows? Politics, Institutions and Other Determinants

Petrovska M. and Mihajlovska E.M. (2013) Measures of Financial Stability in Macedonia; Journal of Central Banking Theory and Pratice; UDC: 336.13.051(497.7)

Proskurovska V. (2012): European banking sector (EBF): Facts and figures 2012

PwC (PricewaterhouseCoopers) Report (2010): NPL Europe, Issue 2, June 2010

PwC (PricewaterhouseCoopers) Austria (2012): Banking in times of financial distress Austria & CEE nonperforming assets overview 2012, PwC Austria

PwC (PricewaterhouseCoopers) Report (2012): Stable volume masks changing dynamics, The European NPL Barometer, PwC

PwC (PricewaterhouseCoopers) (2012): A growing non-core asset market; European outlook for non-core and non-performing loan portfolios, PwC

Roodman D. (2009): How to do xtabond2: An introduction to difference and system GMM in Stata; The Stata Journal 9, Number 1, pp. 86–136 Center for Global Development Washington, DC

Shijaku H. and Ceca K. (2011): A model for the credit risk in Albania using banks' panel data, Working Paper 06(9)011, Bank of Albania

Shingjergji A. (2013): The Impact of Macroeconomic Variables on the Non-Performing Loans in the Albanian Banking System During 2005-2012; Academic Journal of Interdisciplinary Studies, MCSER Publishing-Rome,Italy, E-ISSN 2281-4612, ISSN 2281-3993,Vol. 2, No. 9

The Banker (2013): “Belarus banks hit more turbulence”, author: Jan Cienski

UN (United Nations) (2012): World Population Prospects: The 2012 Revision, Classification of countries by major area and region of the world; UN Population Division, Department of Economic and Social Affairs

Vogiazas S. D. and Nikolaidou E. (2011): Investigating the determinants of NPL in the Romanian banking system: an empirical study with reference to the Greek crisis, Hindawi Publishing Corporation, Economics Research International, Volume 2011, Article ID 214689, doi:10.1155/2011/214689

Vukovic V. and Domazet I. (2013): Non-performing loans and systemic risk: comparative analysis of Serbia and countries in transition CESEE; Industrija, Vol.41, No.4

Wooldridge J. M. (2006): Introductory econometrics, Thomson South-Western; Thomson Learning Academic Center

World Finance Review (2013): Moldova: Recent Economic Developments and Outlook – World Bank, December 2013

Zeman J. and Jurča P. (2008): Macro stress testing of the Slovak banking sector , Working Paper 1 National Bank of Slovakia, ISSN: 1337-5830

Stata: Stata manual, xtunitroot - Panel-data unit-root tests

World Bank Governance

Global Insight Business Conditions and Risk Indicators

International Financial Statistics, International Monetary Fund

World Bank Database

World Economic Outlook

# Appendix A: Empirical part

**Table A. 1: Descriptive statistics full sample**

| VARIABLES | (1) N | (2) mean | (3) sd | (4) min | (5) max |
|---|---|---|---|---|---|
| npl | 154 | 6.940 | 6.491 | 0.200 | 30 |
| gdpg | 154 | 2.411 | 3.719 | -14.80 | 12.10 |
| infl | 154 | 4.403 | 4.439 | -1.710 | 25.20 |
| er | 148 | 13.36 | 26.10 | 0.628 | 113.1 |
| unempl | 154 | 11.96 | 8.934 | 3.800 | 37.25 |
| edind | 148 | 0.760 | 0.0748 | 0.624 | 0.913 |
| iqind | 154 | 0.720 | 0.177 | 0.420 | 0.990 |
| fdi_gdp | 154 | 5.593 | 6.064 | -6.744 | 32.95 |

*Source:* Author's elaboration in Stata

**Table A. 2: Descriptive statistics for groups of countries**

| | (1) dum_coun 0 | (2) | (3) | (4) | (5) | (6) dum_coun 1 | (7) | (8) | (9) | (10) |
|---|---|---|---|---|---|---|---|---|---|---|
| VARIABLES | N | mean | sd | min | max | N | mean | sd | min | max |
| npl | 77 | 10.69 | 6.854 | 1.800 | 30 | 77 | 3.186 | 3.045 | 0.200 | 18.70 |
| gdpg | 77 | 3.547 | 4.208 | -14.80 | 12.10 | 77 | 1.276 | 2.744 | -8.530 | 6.080 |
| infl | 77 | 6.744 | 5.215 | -0.810 | 25.20 | 77 | 2.062 | 1.171 | -1.710 | 4.730 |
| er | 71 | 26.55 | 33.03 | 1.949 | 113.1 | 77 | 1.203 | 0.228 | 0.628 | 1.471 |
| unempl | 77 | 16.31 | 10.71 | 4 | 37.25 | 77 | 7.601 | 2.727 | 3.800 | 15.65 |
| edind | 71 | 0.716 | 0.0524 | 0.624 | 0.811 | 77 | 0.800 | 0.0699 | 0.650 | 0.913 |
| iqind | 77 | 0.554 | 0.0588 | 0.420 | 0.680 | 77 | 0.887 | 0.0569 | 0.730 | 0.990 |
| fdi_gdp | 77 | 6.184 | 5.134 | 0.811 | 32.95 | 77 | 5.003 | 6.852 | -6.744 | 26.65 |

*Source:* Author's elaboration in Stata

**Table A. 3: Unit root presence for variables in level**

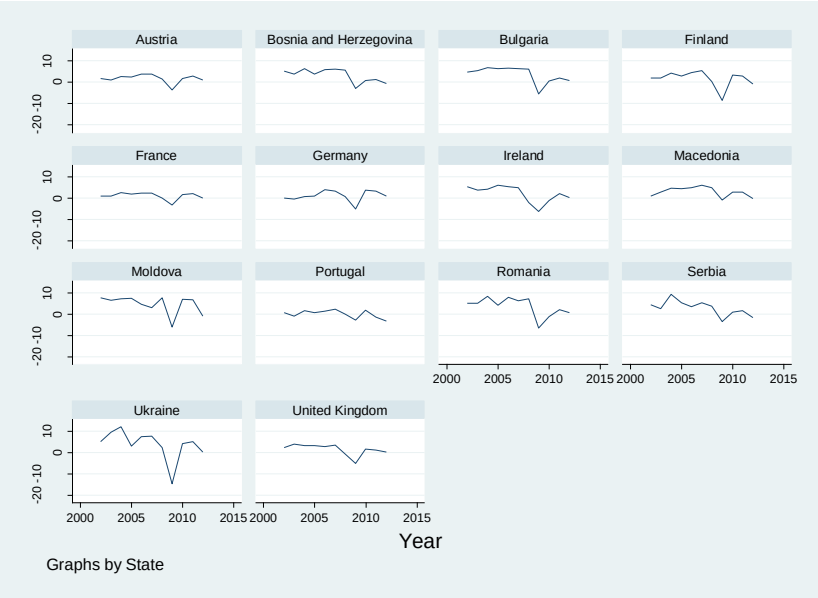
| Sample Unit root tests | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| VARIABLES | npl | gdpg | er | unempl | fdi_gdp | edind | iqind | infl |
| Levin-Lin-Chu (LLC) | -5.0903 | -4.4520 | - | -0.7879 | -5.7009 | - | -2.2209 | -5.2427 |
| | 0.0000 | 0.0000 | - | 0.2154 | 0.0000 | - | 0.0132 | 0.0000 |
| Harris-Tzavalis (HT) | 0.8316 | 0.2015 | - | 0.7948 | 0.0617 | - | 0.2279 | 0.3143 |
| | 0.8856 | 0.0000 | - | 0.7456 | 0.0000 | - | 0.5901 | 0.0000 |
| Breitung | 0.4813 | -5.5436 | - | 0.1143 | -4.2237 | - | -1.4686 | -2.3973 |
| | 0.6849 | 0.0000 | - | 0.5455 | 0.0000 | - | 0.9913 | 0.0083 |
| Im-Pesaran-Shin (IPS) | -0.6116 | -2.2516 | - | -1.1287 | -2.4923 | - | 0.6694 | -2.7764 |
| | 0.9999 | 0.0026 | - | 0.9436 | 0.0018 | - | 0.1172 | 0.0001 |
| Fisher type (ADF) | 3.5575 | 51.5933 | 70.4271 | 1.3489 | 78.714 | 81.066 | 0.9417 | 90.3237 |
| | 0.9998 | 0.0043 | 0.0000 | 0.9113 | 0.0000 | 0.0000 | 0.8268 | 0.0000 |

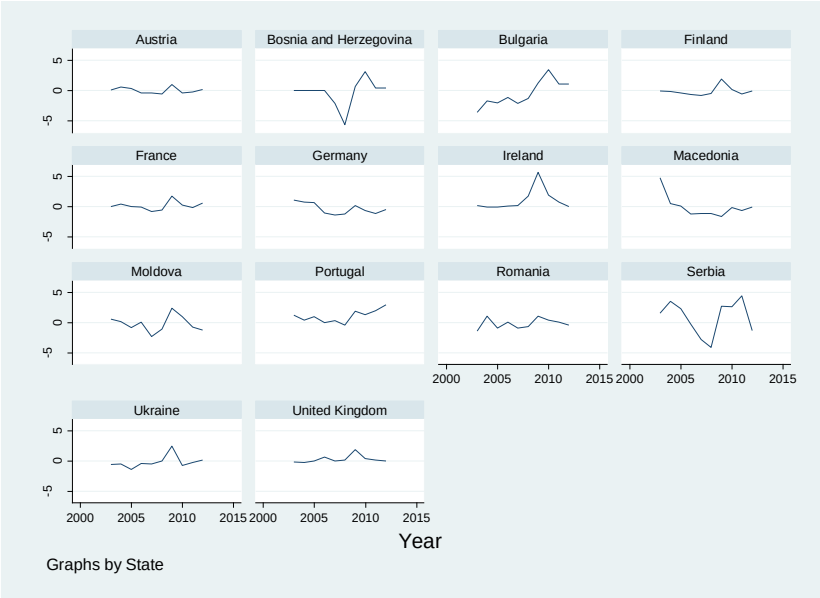
| Sub-sample 1 Unit root tests | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| VARIABLES | npl | gdpg | er | unempl | fdi_gdp | edind | iqind | infl |
| Levin-Lin-Chu (LLC) | 0.8268 | - | - | -1.9823 | -2.1706 | - | -2.8363 | -3.2267 |
| | 0.0000 | - | - | 0.0237 | 0.0150 | - | 0.0023 | 0.0006 |
| Harris-Tzavalis (HT) | 0.7800 | - | - | 0.6216 | 0.5739 | - | 0.3252 | 0.3155 |
| | 0.6229 | - | - | 0.0903 | 0.0331 | - | 0.6275 | 0.0000 |
| Breitung | 0.2024 | - | - | -1.1571 | -3.0514 | - | 0.5039 | -0.9448 |
| | 0.5802 | - | - | 0.1236 | 0.0011 | - | 0.0051 | 0.1724 |
| Im-Pesaran-Shin (IPS) | -0.7690 | - | - | -1.8422 | -1.8182 | - | -2.8642 | -2.9525 |
| | 0.9840 | - | - | 0.1415 | 0.1463 | - | 0.122 | 0.0018 |
| Fisher type (ADF) | 2.1634 | 21.8036 | 0.8722 | 17.2391 | 16.2861 | 56.4729 | 89.5443 | 51.7722 |
| | 0.9847 | 0.0827 | 0.8084 | 0.2437 | 0.2962 | 0.0000 | 0.0000 | 0.0000 |

| Sub-sample 2 Unit root tests | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| VARIABLES | npl | gdpg | er | unempl | fdi_gdp | edind | iqind | infl |
| Levin-Lin-Chu (LLC) | -2.2636 | -4.3387 | -2.1024 | 1.0394 | -6.2696 | - | 0.3216 | -4.3527 |
| | 0.0118 | 0.0000 | 0.0178 | 0.8507 | 0.0000 | - | 0.6261 | 0.0000 |
| Harris-Tzavalis (HT) | 1.1619 | 0.2505 | 0.5168 | 1.0393 | -0.2728 | - | 2.8191 | 0.2991 |
| | 1.0000 | 0.0000 | 0.0075 | 0.9987 | 0.0000 | - | 0.9994 | 0.0000 |
| Breitung | 0.5063 | -4.0520 | 0.2054 | 1.4764 | -2.9245 | - | -0.0730 | -3.1084 |
| | 0.6937 | 0.0000 | 0.5814 | 0.9301 | 0.0000 | - | 1.0000 | 0.0009 |
| Im-Pesaran-Shin (IPS) | -0.4542 | -2.3168 | -3.2106 | -0.4152 | -3.1664 | - | 0.9585 | -2.6003 |
| | 0.9988 | 0.0165 | 0.0006 | 0.9995 | 0.0011 | - | 0.9852 | 0.0072 |
| Fisher type (ADF) | 239.891 | 26.6334 | 59.9282 | 6.6067 | 62.4279 | -0.2669 | 1.3002 | 38.5515 |
| | 0.0000 | 0.0215 | 0.0000 | 0.9488 | 0.0000 | 0.3948 | 0.9032 | 0.0004 |

*Source*: Author's elaboration; Note: adjusted t*, t-bar, chi-squared or inverse normal statistics reported; Results from Stata (and Gretl occasionally)

**Figure A. 1: Plots of unemployment rate and GDP growth**




*Source:* Author's elaboration in Stata

**Table A. 4: Correlation matrix**

| | dlognpl | gdpg | er | diff_unempl | Fdi_gdp | edind | dlog_iqind | infl |
|---|---|---|---|---|---|---|---|---|
| dlognpl | 1 | | | | | | | |
| gdpg | - 0.603*** | 1 | | | | | | |
| er | - 0.0786 | 0.0645 | 1 | | | | | |
| diff_unempl | 0.407*** | -0.489*** | 0.0552 | 1 | | | | |
| Fdi_gdp | - 0.0254 | 0.237** | 0.00661 | -0.0880 | 1 | | | |
| edind | 0.0873 | -0.169 | -0.354*** | 0.107 | 0.0612 | 1 | | |
| dlog_iqind | - 0.220* | 0.137 | 0.120 | -0.176* | 0.0733 | -0.0210 | 1 | |
| infl | - 0.128 | 0.310*** | 0.300*** | -0.142 | 0.168 | -0.231** | 0.0525 | 1 |

*Source:* Author's elaboration in Stata

**Table A. 5: VIF a), b), c), d)**

a) sample

| VARIABLE | VIF | 1/VIF |
|---|---|---|
| gdpg | 1.50 | 0.667835 |
| infl | 1.36 | 0.733256 |
| er | 1.26 | 0.794202 |
| unempl | 1.24 | 0.804507 |
| edind | 1.21 | 0.827752 |
| iqind | 1.09 | 0.915922 |
| fdi_gdp | 1.06 | 0.944877 |

b) subsample 1

| VARIABLE | VIF | 1/VIF |
|---|---|---|
| l2npl | 12.69 | 0.078774 |
| l1npl | 6.54 | 0.152993 |
| l3npl | 5.75 | 0.173963 |
| iqind | 2.66 | 0.375486 |
| edind | 2.57 | 0.388744 |
| log_fdi_gdp | 1.89 | 0.527843 |
| infl | 1.81 | 0.553021 |
| diff_er | 1.45 | 0.690629 |
| diff_unempl | 1.35 | 0.740048 |
| diff_gdpg | 1.24 | 0.804844 |

c) subsample 1 check

| VARIABLE | VIF | 1/VIF |
|---|---|---|
| iqind | 2.61 | 0.383406 |
| edind | 2.56 | 0.390028 |
| l1npl | 1.93 | 0.517710 |
| log_fdi_gdp | 1.86 | 0.537321 |
| l3npl | 1.77 | 0.563434 |
| infl | 1.76 | 0.566943 |
| diff_er | 1.37 | 0.731262 |
| diff_unempl | 1.34 | 0.744456 |
| diff_gdpg | 1.16 | 0.859078 |

d) subsample 2

| VARIABLE | VIF | 1/VIF |
|---|---|---|
| diff_unempl | 4.19 | 0.238562 |
| gdpg | 2.94 | 0.339963 |
| L2dlognpl | 2.05 | 0.488857 |
| L1dlognpl | 1.93 | 0.518197 |
| Fdi_gdp | 1.71 | 0.585892 |
| infl | 1.64 | 0.611045 |
| diff_iqind | 1.50 | 0.665233 |
| L3dlognpl | 1.39 | 0.721934 |
| diff_edind | 1.38 | 0.724573 |
| er | 1.19 | 0.842521 |

*Source:* Author's elaboration in Stata

**Table A. 6: Pooled OLS with robust standard errors (SE)**

| VARIABLES | (Sample; SE) npl | (Sample; Robust SE) npl | (Subsample 1; SE) npl | (Subsample 1; Robust SE) npl | (Subsample 2; SE) npl | (Subsample 2; Robust SE) npl |
|---|---|---|---|---|---|---|
| npl (-1) | - | - | 0.658*** | 0.658*** | -0.00824 | -0.00824 |
| | - | - | (0.0742) | (0.0773) | (0.138) | (0.169) |
| npl (-2) | - | - | - | - | -0.0185 | -0.0185 |
| | - | - | - | - | (0.143) | (0.107) |
| npl (-3) | - | - | -0.179*** | -0.179* | -0.119 | -0.119 |
| | | | (0.0582) | (0.0868) | (0.118) | (0.155) |
| gdpg | -0.0574*** | -0.0574*** | -0.134** | -0.134** | -0.0215 | -0.0215 |
| | (0.00833) | (0.00596) | (0.0603) | (0.0514) | (0.0202) | (0.0136) |
| er | -0.00104 | -0.00104* | 0.182 | 0.182 | 0.138 | 0.138 |
| | (0.00113) | (0.000555) | (0.132) | (0.108) | (0.190) | (0.237) |
| unempl | 0.0348* | 0.0348** | 0.525** | 0.525*** | 0.147** | 0.147* |
| | (0.0207) | (0.0141) | (0.216) | (0.0762) | (0.0602) | (0.0616) |
| edind | -0.222 | -0.222 | 14.02 | 14.02** | 10.29 | 10.29 |
| | (0.379) | (0.393) | (9.506) | (4.807) | (7.280) | (7.831) |
| iqind | -1.273* | -1.273** | -17.74 | -17.74 | -1.300 | -1.300 |
| | (0.726) | (0.536) | (11.72) | (9.763) | (2.445) | (1.481) |
| infl | 0.00600 | 0.00600** | -0.0435 | -0.0435 | 0.0572 | 0.0572 |
| | (0.00693) | (0.00269) | (0.0892) | (0.0950) | (0.0364) | (0.0406) |
| fdi_gdp | 0.00762* | 0.00762* | -1.193* | -1.193*** | 0.00725 | 0.00725 |
| | (0.00439) | (0.00392) | (0.599) | (0.240) | (0.00747) | (0.00797) |
| Constant | 0.281 | 0.281 | 7.124 | 7.124 | -0.264 | -0.264 |
| | (0.299) | (0.312) | (4.976) | (6.661) | (0.274) | (0.340) |
| Observations | 134 | 134 | 54 | 54 | 49 | 49 |
| R-squared | 0.417 | 0.417 | 0.833 | 0.833 | 0.546 | 0.546 |

*Source:* Author's elaboration in Stata; Note: variables are used with their transformations according to Table 3.4 ;Standard errors (SE) and robust standard errors (Robust SE) in parentheses *** p<0.01, ** p<0.05, * p<0.1

**Table A. 7: Preliminary results of Fixed Effect model**

| VARIABLES | (full sample) npl | (full sample and VA) npl | (Subsample 1) npl | (Subsample 1 and RQ) npl | (Subsample 2) npl | (Subsample 2 and PSAV) npl |
|---|---|---|---|---|---|---|
| npl (-1) | | | 0.541*** | 0.596*** | -0.0765 | -0.223 |
| | | | (0.0588) | (0.0538) | (0.110) | (0.129) |
| npl (-2) | | | | | -0.247** | -0.253* |
| | | | | | (0.0895) | (0.121) |
| npl (-3) | | | -0.247** | -0.218*** | -0.0711 | -0.157 |
| | | | (0.0707) | (0.0585) | (0.209) | (0.173) |
| gdpg | -0.0589*** | -0.0623*** | -0.131** | -0.159*** | 0.000182 | 0.000963 |
| | (0.00543) | (0.00696) | (0.0384) | (0.0405) | (0.0162) | (0.0109) |
| er | -0.0109*** | -0.00289 | -0.0931** | -0.0724** | 1.744 | 2.137* |
| | (0.00181) | (0.00381) | (0.0329) | (0.0291) | (1.063) | (1.038) |
| unempl | 0.0237* | 0.0171 | 0.344*** | 0.428*** | 0.162** | 0.167*** |
| | (0.0125) | (0.0151) | (0.0649) | (0.0603) | (0.0470) | (0.0416) |
| edind | 2.126 | 4.993** | -33.13 | -34.32 | 7.065 | 11.06 |
| | (2.234) | (1.890) | (19.67) | (26.20) | (7.557) | (6.559) |
| iqind | -1.109* | | -56.80*** | | -1.954 | |
| | (0.543) | | (8.251) | | (1.711) | |
| infl | -0.00143 | 0.00942 | -0.0202 | -0.0645 | 0.00181 | 0.0175 |
| | (0.00552) | (0.0117) | (0.0869) | (0.107) | (0.0294) | (0.0321) |
| fdi_gdp | -0.00177 | -0.00242 | -2.145*** | -1.803*** | -0.0131 | -0.0153 |
| | (0.00390) | (0.00368) | (0.409) | (0.381) | (0.00756) | (0.00795) |
| va | | -2.401* | | | | |
| | | (1.260) | | | | |
| rq | | | | -17.87*** | | |
| | | | | (4.392) | | |
| psav | | | | | | -2.449** |
| | | | | | | (0.917) |
| Constant | -1.292 | -1.525 | 66.07** | 44.79* | -2.097 | -0.727 |
| | (1.687) | (1.274) | (17.86) | (21.75) | (1.288) | (1.068) |
| Observations | 134 | 134 | 54 | 54 | 49 | 49 |
| R-squared | 0.460 | 0.506 | 0.869 | 0.864 | 0.635 | 0.678 |
| Number of country | 14 | 14 | 7 | 7 | 7 | 7 |

*Source:* Author's elaboration in Stata; Note: variables are used with their transformations according to Table 3.4; Standard errors (SE) and robust standard errors (Robust SE) in parentheses *** p<0.01, ** p<0.05, * p<0.1

**Table A. 8: Additional empirical tests**

| TEST | Full sample | Subsample 1 | Subsample 2 |
|---|---|---|---|
| **Wald statistic test for heteroskedasticity after robust SE** | chi2(1) = 3.10 | chi2 (7) = 94.81 | chi2 (7) = 337.44 |
| | Prob>chi2 = 0.0000 | Prob>chi2 = 0.0000 | Prob>chi2 = 0.0000 |

| TEST | Full sample | Subsample 1 | Subsample 2 |
|---|---|---|---|
| **Wooldridge test of serial correlation** | F( 1, 13) = 15.965 | F( 1, 6) = 0.555 | F( 1, 6) = 13.245 |
| | Prob > F = 0.0015 | Prob > F = 0.4845 | Prob > F = 0.0108 |

| TEST | Full sample | Subsample 1 | Subsample 2 |
|---|---|---|---|
| **Cross sectional dependence Pesaran test** | 6.231 | -1.278 | -1.025 |
| | Pr = 0.0000 | Pr = 0.2012 | Pr = 0.3055 |

*Source:* Author's elaboration in Stata

**Table A. 9: Hausman tests**

| Full sample | Subsample 1 | Subsample 2 |
|---|---|---|
| chi2(7) = (b-B)'[(V_b-V_B)^(-1)](b-B) = 17.00 | chi2(9) = (b-B)'[(V_b-V_B)^(-1)](b-B) = 30.81 | chi2(10) = (b-B)'[(V_b-V_B)^(-1)](b-B) = 50.07 |
| Prob>chi2 = 0.0174 | Prob>chi2 = 0.0003 | Prob>chi2 = 0.0000 |

*Source:* Author's elaboration in Stata

**Table A. 10: Export/import dominant countries**

| | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 | 2012 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Middle Income Group* | | | | | | | | | | | |
| **Bosnia and Herzegovina** | | | | | | | | | | | |
| **Bulgaria** | | | | | | | | | | | |
| **Macedonia** | | | | | | | | | | | |
| **Moldova** | | | | | | | | | | | |
| **Romania** | | | | | | | | | | | |
| **Serbia** | | | | | | | | | | | |
| **Ukraine** | | | | | | | | | | | |
| *High Income Group* | | | | | | | | | | | |
| **Austria** | | | | | | | | | | | |
| **Finland** | | | | | | | | | | | |
| **France** | | | | | | | | | | | |
| **Germany** | | | | | | | | | | | |
| **Ireland** | | | | | | | | | | | |
| **Portugal** | | | | | | | | | | | |
| **United Kingdom** | | | | | | | | | | | |

*Source*: Author's elaboration based on cutoff point of 41%; Note: the highlighted cells are the export dominant countries, the rest import dominant countries

**Table A. 11: Estimation of the model including the extension of exchange rate hypothesis (Fixed Effect model)**

| VARIABLES | (full sample) npl | (subsample 1) npl | (subsample 2) npl |
|---|---|---|---|
| npl (-1) | | 0.518*** | -0.107 |
| | | (0.0639) | (0.102) |
| npl (-2) | | | -0.224* |
| | | | (0.0990) |
| npl (-3) | | -0.188** | -0.0792 |
| | | (0.0672) | (0.216) |
| gdpg | -0.0580*** | -0.110* | -0.00211 |
| | (0.00690) | (0.0511) | (0.0180) |
| unempl | 0.0238* | 0.406*** | 0.178** |
| | (0.0132) | (0.0621) | (0.0498) |
| edind | 1.902 | -18.46 | 6.688 |
| | (2.306) | (25.85) | (7.644) |
| iqind | -1.062* | -34.71** | -2.432 |
| | (0.556) | (12.91) | (1.584) |
| infl | -0.000772 | -0.0370 | 0.0192 |
| | (0.00608) | (0.113) | (0.0320) |
| fdi_gdp | -0.00167 | -1.966*** | -0.0103 |
| | (0.00390) | (0.216) | (0.00820) |
| erDE | 0.00756 | 0.750*** | 1.593 |
| | (0.0313) | (0.187) | (1.062) |
| erDI | -0.0107*** | -0.0373 | 1.726 |
| | (0.00206) | (0.0235) | (1.070) |
| Constant | -1.151 | 42.35 | -1.987 |
| | (1.728) | (22.40) | (1.293) |
| Observations | 134 | 56 | 49 |
| R-squared | 0.461 | 0.870 | 0.643 |
| Number of country | 14 | 7 | 7 |

*Source:* Author's elaboration in Stata; Note: variables are used with their transformations according to Table 3.4; Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.

**Table A. 12: Autocorrelation test and Sargan test for overidentification or valid instruments; Robust SE**

| TEST | one-step difference GMM | two-step system GMM | two-step system GMM (*erDE* and *erDI*) | two-step system GMM (dummies for *edind*) | two-step system GMM (dummies for *er_DE*) | two-step system GMM (dummies for *er_DI*) | two-step system GMM (dummies for *IQind* |
|---|---|---|---|---|---|---|---|
| **Autocorrelation test AR (1)** | z = -1.98 | z = -1.86 | z = -2.11 | z = -1.99 | z = -2.04 | z = -2.20 | z = -1.91 |
| | Pr > z = 0.048 | Pr > z = 0.063 | Pr > z = 0.035 | Pr > z = 0.046 | Pr > z = 0.041 | Pr > z = 0.028 | Pr > z = 0.056 |
| **Autocorrelation test AR (2)** | z = -1.49 | z = -0.93 | z = 0.78 | z = 0.84 | z = 0.68 | z = 0.60 | z = 0.75 |
| | Pr > z = 0.136 | Pr > z = 0.353 | Pr > z = 0.435 | Pr > z = 0.399 | Pr > z = 0.496 | Pr > z = 0.545 | Pr > z = 0.452 |
| **Sargan test** | chi2(13) = 16.21 | chi2(13) = 16.21 | chi2(41) = 49.72 | chi2(41) = 55.86 | chi2(41) = 42.70 | chi2(41) = 55.59 | chi2(41) = 55.29 |
| | Prob > chi2 = 0.238 | Prob > chi2 = 0.238 | Prob > chi2 = 0.165 | Prob > chi2 = 0.061 | Prob > chi2 = 0.398 | Prob > chi2 = 0.064 | Prob > chi2 = 0.067 |

*Source:* Author's elaboration in Stata

**Table A.13: System GMM results for inflation rate and FDI in middle and high income groups**

| VARIABLES | npl |
|---|---|
| gdp_dum | -0.0356** |
| | (0.0148) |
| gdp_dum2 | -0.0369*** |
| | (0.00794) |
| erDE_dum | -0.0200 |
| | (0.108) |
| erDE_dum2 | -0.0862** |
| | (0.0309) |
| erDI_dum | 0.193 |
| | (0.172) |
| erDI_dum2 | 0.00228 |
| | (0.00231) |
| unempl_dum | 0.125* |
| | (0.0600) |
| unempl_dum2 | 0.0320** |
| | (0.0111) |
| educ_dum | 5.594* |
| | (3.170) |
| educ_dum2 | 5.305* |
| | (2.646) |
| infl_dum | -0.00189 |
| | (0.153) |
| infl_dum2 | -0.00682 |
| | (0.0282) |
| iqind_dum | -4.425 |
| | (2.942) |
| iqind_dum2 | -4.735 |
| | (3.559) |
| fdi_gdp_dum | -0.00141 |
| | (0.0142) |
| fdi_gdp_dum2 | -0.0278 |
| | (0.0192) |
| Observations | 112 |
| Number of countries | 14 |

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

*Source:* Author's elaboration in Stata; Note: the coefficients are showed only, from all regressions interacted with the dummy variables; gdp_dum stands for the gdp from high income countries and gdp_dum2 for gdp from middle income countries, the same applies for all other determinants (dum_coun – high income and dum_coun2 – middle income); Each determinant is interacted with the dummy variables in the basic model; variables are used with their transformations according to Table 3.4, here only from full model specifications

# Appendix B: Robustness checks

**Table B. 1: Robustness checks results**

| VARIABLES | (1)<br>npl | (2)<br>npl | (3)<br>npl |
|---|---|---|---|
| L.npl | 0.845*** | 0.836*** | 0.578*** |
| | (0.0899) | (0.112) | (0.107) |
| L3.npl | -0.129 | -0.00844 | -0.0110 |
| | (0.199) | (0.244) | (0.0902) |
| gdpg | -0.0676*** | -0.0622*** | -0.0352*** |
| | (0.00992) | (0.0143) | (0.00889) |
| erDE | -0.208** | -0.200** | -0.0812* |
| | (0.0837) | (0.0746) | (0.0402) |
| erDI | 0.00106 | -0.00151 | -0.000882 |
| | (0.00382) | (0.00441) | (0.00334) |
| unempl | 0.0229 | 0.0438 | 0.0611 |
| | (0.0227) | (0.0404) | . |
| fdi_gdp | 0.0117 | -0.0103 | -0.0395* |
| | (0.0130) | (0.0202) | (0.0197) |
| iqind | -3.306* | -2.851 | -3.707 |
| | (1.836) | (1.809) | (2.405) |
| sen | 0.0536* | 0.0480* | |
| | (0.0280) | (0.0271) | |
| infl | 0.0200 | 0.0229 | 0.00742 |
| | (0.0342) | (0.0341) | (0.0278) |
| edind | | | 4.873* |
| | | | (2.668) |
| Observations | 92 | 92 | 112 |
| Number of country | 13 | 13 | 14 |

Source: Author's elaboration in Stata; Note: variables are used with their transformations according to Table 3.4; Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.

# Master Thesis Proposal

| | |
|---|---|
| **Author:** | Violeta Sandrovschi |
| **Supervisor:** | PhDr. Ing. Petr Jakubík Ph.D. |
| **Defense Planned:** | September 2014 |

**Proposed Topic:**

Determinants of NPLs at the aggregate level: A comparative approach for middle and high income countries

**Topic Characteristics:**

Past experience of risky assets facing the financial and currency crisis from Asian countries in 1997 motivated the banking and macroeconomic system to induce a higher pressure on the market through prudential rules and regulations to prevent the default of banks' loans. Credit risk measurement may be associated with estimating the non-performing loans(NPLs) by regulators and banks.

In my thesis, I will focus on macroeconomic determinants or factors that influence the change of NPLs, by estimating and using empirical study. For assessing the financial stability for a particular country or in regions, this research topic is of major relevance as many banks offer an excessive amount of loans in the boom period contributing to build up external imbalances. On the contrary, in recessions banks might dramatically reduce lending due to high volume of non-performing loans and negative feedback effect on their balance sheets could be emerged. Hence, regulators and bank supervisors try to use different disposable management tools though limited, for assessment of NPLs determinants as GDP, interest rates, unemployment, and foreign exchange, necessary for stress tests scenarios.

In case of a literature analysis, we see that many attempts are done for specific countries and less for group of countries within particular geographical region offering some results for NPL determinants at the aggregate level. This thesis will approach first of all an extension of previous studies on NPLs based on the idea to compare the determinants of the indicator for two groups of selected countries from Europe: middle and high income economies. Based on the World Bank classification, the selected 12 countries are divided into middle income countries: Bosnia and Herzegovina, Bulgaria, Macedonia, Moldova, Romania, Serbia, Ukraine, and high income countries: Austria, Finland, France, Germany, Ireland, Portugal, and United Kingdom.

Key innovation for this thesis is the work of comparing the 2 groups of economies under new additional variables and specifications. I am going to use panel data set with annual frequency covering randomly selected countries, and the material available from the World Bank's database, Central Banks of selected countries, IMF. The period of analysis would comprise the years 2002-2012.

**Hypotheses:**

1. GDP growth has a negative consequence on NPL ratio
2. An increase in inflation will increase the level of asset quality, assessed by NPLs ratio
3. Foreign Direct Investments will have a significant impact on NPLs in all groups of countries and a higher FDI from GDP a lower NPL ratio
4. Impact of exchange rate on NPLs is significant for all groups of countries under import/export dominant specifications
5. Employment amelioration provides a better NPLs ratio
6. Acquiring more education provides incentives for lenders to enlarge the amount offered which leads to higher NPL ratio
7. Indices for efficiency of a single institutional index are significant and their increase have negative impact on the NPLs ratio for all groups of countries

**Methodology:**

For an empirical testing of the hypotheses mentioned above, I plan to use the static panel data models at the beginning and continue with a dynamic panel model (Generalized Method of Moments). The independent variables are planned to be the GDP growth, unemployment rate, inflation rate, exchange rate, Foreign Direct Investments as % from GDP, educational index and insitutional index. The choice of independend variables must be correctly specified in order to keep a low number of regressors and relevant variables. As for the dependent variable (NPL ratio), the comparisons imply separate estimation for the developed and developing countries..

For the fourth hypothesis it must be generate a model which includes dummy variables for the case the countries are export dominant or import dominants. Thereafter the result should hihglight if there is a difference among the 2 groups of countries for an increase/ decrease of NPL. For the last 2 hypotheses, we construct the indices at the national level and interpret the results.

**Outline:**

1. Introduction
2. Implications of the topic
3. Empirical analysis
4. Robustness checks
5. Interpretation of results
6. Conclusions

**Core Bibliography:**

1. 1. Roland Beck, Petr Jakubik and Anamaria Piloiu (2013) „Non-performing loans: What matters in addition to the economic cycle? “ Macroprudential Research Network ECB, Working Paper Series NO 1515
2. Helmut Elsinger, Alfred Lehar, Martin Summer (2002) „Risk Assessment for Banking Systems“
3. Dimitrios P. Louzis, Angelos T. Vouldis, Vasilios L. Metaxas (2011) „Macroeconomic and bank-specific determinants of non-performing loans in Greece: A comparative study of mortgage, business and consumer loan portfolios“, Journal of Banking & Finance
4. Adam Głogowski (2008) „Macroeconomic determinants of Polish banks' loan losses – results of a panel data study“, National Bank of Poland Working Paper

NO 53
5. Raphael Espinoza and Ananthakrishnan Prasad (2010) „Nonperforming Loans in the GCC Banking System and their Macroeconomic Effects“, IMF Working Paper Middle East and Central Asia Department
6. Asghar Ali, Kevin Daly (2010) „What Explain Credit Defaults? A Comparative Study“, SBP Research Bulletin Volume 6, NO 1
7. Yuqi Li „Determinants of Banks Profitability and its Implication on Risk Management Practices: Panel Evidence from the UK in the Period 1999-2006“, The Univesity of Nottingham
8. Ricardo Schechtman and Wagner Piazza Gaglianone (2011) „Macro stress testing of credit risk focused on the tails“, Banco Central do Brasil
9. Nor Hayati Ahmad, Mohamed Ariff (2007) „Multi-country study of bank credit risk determinants“, International Journal of Banking and Finance, Vol. 5 Issue 1

**Author** **Supervisor**