PhD-FDEF-2019-10
The Faculty of Law, Economics and Finance

DISSERTATION

Defence held on 04/12/2019 in Luxembourg

to obtain the degree of

DOCTEUR DE L'UNIVERSITÉ DU LUXEMBOURG

EN SCIENCES FINANCIÈRES

by

Carmine GABRIELE

Born on 23 July 1984 in Napoli (Italy)

ESSAYS ON FINANCIAL STABILITY

Dissertation defence committee

Dr Thorsten Lehnert, dissertation supervisor

Professor, University of Luxembourg

Dr Theoharry Grammatikos, Chairman

Professor, University of Luxembourg

Dr Aitor Erce

European Investment Bank

Dr Vilem Valenta

European Central Bank

Essays on Financial Stability

CARMINE GABRIELE

Abstract

This dissertation consists of an introduction and three self-contained papers, each organised in a separate chapter.

In the first paper (chapter 2), we develop an early warning model of systemic banking crises that combines regression tree technology with a statistical algorithm (CRAGGING), with the goal of improving its accuracy and overcoming the drawbacks of previously used models. Our model has a large set of desirable features. It provides endogenously-determined critical thresholds for a set of useful indicators, presented in the intuitive form of a decision tree structure. Our framework takes into account the conditional relations between various indicators when setting early warning thresholds. This facilitates the production of accurate early warning signals as compared to the signals from a logit model and from a standard regression tree. Our model also suggests that high credit aggregates, both in terms of volume and as compared to a long-term trend, as well as low market risk perception, are amongst the most important indicators for predicting the build-up of vulnerabilities in the banking sector.

In the second paper (chapter 3), "Macro-financial determinants of NPLs over the credit cycle, and the role of uncertainty", we assess the macroeconomic and financial determinants of non-performing loans (NPLs) in the euro area. We investigate the role of these determinants in affecting NPLs over the credit cycle, and we analyse the role of uncertainty. We find that some of the determinants affect NPLs differently over the different stages of the credit cycle. In particular, when the credit cycle is positive, some of the standard relations weaken or change. We also find that economic and financial uncertainty may play a role, through the different degrees of risk-aversion of the banks at different levels of uncertainty. Low uncertainty, if associated to low levels of risk-aversion, could determine excessive risk taking by financial institutions, especially during credit expansions, which could frustrate the positive effects that economic growth has on NPLs during normal times.

In the third and last paper (chapter 4), written with my former colleague at the European Stability Mechanism, Dr. Aitor Erce, we focus on a frequent counterpart of banks when they experience (systemic) crises, i.e., the liability side of the general government's balance sheet. Traditional approaches to study sovereign debt sustainability, heavily dependent on debt level metrics and giving little role to debt flow indicators, are insufficient. We inform this debate by analysing the ability of gross financing needs, the debt flow metric that currently complements the traditional debt sustainability analysis, to provide information about a sovereign's likelihood of distress beyond that provided by debt stock metrics. We show that stock and flow metrics need to be assessed in combination. We also document an important role for gross financing needs when debt stocks are high. If debt is above 60%, reducing gross financing needs by one percent of GDP translates into 10 basis points lower sovereign spreads. In addition, we find that roll-over needs play a critical role in driving this effect. Our findings help understand how countries can sustain large debt stocks without suffering fiscal crises and, to the extent that official lending affects refinancing needs, they also inform the literature on crisis resolution.

Acknowledgements

Before starting my list of acknowledgements, I would like to say that this section is longer than I thought. This is due to the parsimony I used when I wrote it in previous occasions.

I would like to thank professor Lehnert for trusting me and making this possible. I am also thankful to the DSEF people for all the courses and seminars I attended where I have learnt a lot.

I am thankful to my employer, the European Stability Mechanism (ESM), for giving me this opportunity and for supporting me in this project. Especially my former manager, Agnes, for giving the first approval, and Rolf, my reference point at the ESM. I feel lucky for being a member of staff there.

A few former and current colleagues -all still friends- deserve a special mention. Antonello for encouraging me in this decision and for the technical help in the early work. Aitor, for challenging my work in every possible occasion, and for co-authoring one of the papers in this essay. Daragh, for the time spent reading part of this essay, and for listening to me, contributing to improve it. I am also thankful to him for constantly reminding me that, whatever I was working on, I should have been working on this dissertation instead. Mattia, for the help with the third chapter.

Many thanks also go to all the people that over the past years have attended my seminars.

Outside the academic and work environment, there is a group of people, which have contributed differently-but possibly more- to this achievement.

I would like to thank my family, i.e. Patrizia and Fulvio, my parents, and Alessandro and Maria Giovanna, my siblings. The last time that I have written that I was thankful to them was about 8 years ago, for my master dissertation acknowledgements. Since then, many things have changed, by all meanings. Just a few good ones: I built my own family, as my sister did, giving birth to my sweet nephew Claudio, and my brother has made many progresses in his career and in his life. However, something has never changed. The deep union, which has remained (a) constant over time, and which helps indirectly in many life events. So, thanks mum and dad for the love. Thanks Ale for being my first partner in the first adventure away from home and for the help you gave me when it was needed. Thanks Giovi for being my little sister and for sharing my same sensitiveness, we always need someone who understands us better within a family.

My thanks also go to my parents-in-law, Isa and Antonello, and to Nonna Giovi. They have accepted me immediately in their family, making me feel as one of them. They have been very helpful to me in countless circumstances. They supported me (us) in many important moments, like all the times they helped my family not to be affected by my absence when I was busy with work or PhD issues. I also want to give a special thanks to prof. Antonello for reviewing and improving the English of part of my final draft.

I would like to thank my second pair of parents, zia Anna and zio Enzo, for their parental love. A special thanks to zio Enzo for challenging me with math exercises when I was a little kid, just to prove me that I was good.

Finally, last but definitely not least, I would like to thank my family. The family I have built myself, together with my wife Ilaria. The family that we will keep strengthening, together with our little star, Maria Caterina.

I would like to start thanking Maria Caterina, for pulling out the best of me. Every single moment I feel like I have to be a better person just because she exists. I want to thank her also for the joy, love and happiness she gives to me every day. She arrived during the last year, and, contrarily to what most of my colleagues and friends told me about working, doing a part-time PhD and being a father, contemporaneously, she has been an unbelievable boost.

Finally, I would like to thank my wife. I am grateful to her for so many things. She has supported me since the beginning of this project in many different ways. In fact, I am thankful to her for supporting the decision to start the PhD, despite it could have affected our relationship, for the many hours spent reading my work and listening to me. I am also grateful for all the times she took care of literally everything else around me, when I was busy on this dissertation. She always motivated me in the right direction (sometimes even too much ☺). I never felt like the PhD was taking time away from her, or from our family, and this is thanks to Ilaria’s efforts.

In so many occasions, she took over our relationship and our family, to let me focus. She is the reason for which having a full-time job, doing a part-time PhD and building a family, have been compatible. This was costly to her, and I know. Thank you. Without all the people mentioned here, this thesis would have been worse. Without Ilaria, it would have not been possible.

As a final note, I would like to thank myself. I know why.

Table of Contents

Chapter 1

1.1 Introduction

Systemic banking system crises are generally associated with the most painful and long-lasting recessions, which most of the times end up involving the general government balance sheet, creating liquidity and solvency issues for the sovereigns.

Being able to understand whether banking systems are more or less vulnerable to sudden shocks has become of fundamental importance, especially after the global financial crisis (GFC). For this reason a broad and growing literature on early warning models for systemic banking crises is nowadays extensively available. Good early warning models might have helped preventing, or at least getting ready for the GFC.

Large financial crises usually start with (or end leaving as a legacy) deteriorated assets in financial institutions' balance sheets. This represents a substantial issue, as high Non-Performing Loan (NPL) stocks impair the supply and an efficient allocation of credit to the economy. In fact, most of the early warning models for banking crises use excessively-high NPLs as a possible banking crisis event, together with other definitions.

It is very important to understand from a systemic perspective what drives NPLs to inform appropriately early warning models, to allow policymakers to implement the right policies, as well as to do it with the right timing.

As observed during the past decade, a financial crisis of very large scale has spilled-over to the public sector in a relatively short time. Some European governments have run extremely large deficits and increased public debt in order to recapitalise banks or entire banking systems. This has raised concerns regarding both the sustainability of public debt, and the ways this latter is assessed. The traditional approaches to public debt sustainability analysis used to focus on the level of debt, whereas the flows generated by such debt stock result very important in assessing whether a sovereign is able to service its debt.

In Chapter 2, we combine alternative methodologies for early warning systems (EWS) with the classic techniques available in the literature to build an early warning system for systemic banking crises. We use an ensemble method developed by Savona and Vezzoli (2012) called CRAGGING, which aggregates the results of many "learners" (Regression Trees in this case), and we adapt it to issue early warning signals of systemic banking crises.

The global financial crisis has led researchers and policymakers around the world to put considerable effort into understanding and preventing systemic banking crises. In doing so, the empirical literature concerned has been focusing on developing early warning systems (EWS) which seek to predict the build-up of dangerous vulnerabilities within banking systems.

In fact, the goal of an EWS is to predict whether a system is vulnerable, to the extent that a sudden adverse shock may lead to a crisis. With regard to banking crises, EWSs turn out to be very useful in giving early signals to the authorities, allowing them to activate pre-emptive actions early enough (e.g., macroprudential measures). The role of policy institutions (national or supranational) is not to predict shocks, but to make sure that when they hit, the economic and financial systems are resilient enough to withstand them.

Binary regression trees overcome some of the well-known issues of the standard EWS methodologies (signalling approach and logistic regression) as they enable to identify rules-of-thumb with endogenously-determined thresholds in a (ex-post) multivariate framework for a set of interrelated indicators. On the one hand, this means that the ordering of the indicators to look at, and their critical thresholds, are determined by the algorithm and not arbitrarily. On the other hand, every *rule* subsequent to the first one comes from a sub-sample, and this causes the non-global optimality of the final prediction.

The easy interpretability of the results of a binary regression tree, coming from the decision tree structure, is one of the advantages of this methodology together with other desirable features that we will describe in Chapter 2.

However, decision trees also come with some drawbacks, amongst which the most important is the low out-of-sample accuracy. This is due to the tendency of decision trees to overfit the data and produce unstable results, as they change when we add new variables or new data points. For this reason we use an ensemble method to build an early warning model that keeps all these features, such that it allows us to achieve the right balance between accuracy and interpretability.

The combination of the two methodologies helps to improve the accuracy of the standard regression trees. The resulting early warning model has a set of desirable features for this class of models. It provides endogenously-determined critical thresholds for a set of indicators, related amongst each other, presenting them in the form of an intuitive decision tree structure. In fact, it takes into account the conditional relations between various indicators when setting early warning thresholds. In doing so, it produces accurate early warning signals as it results from a comparison with the signals of a logit model and of a standard regression tree.

In Chapter 3, we assess a series of macro-financial determinants of Non-performing loans (NPLs) as we believe that the economic and financial conditions of a country are fundamental determinants of NPL dynamics. Good economic performances should help curbing the build-up of problematic loans in the banking system balance sheet. In the same way, the banking system of a country that is experiencing a period of weak economic conditions is more likely to accumulate NPLs, at a faster pace.

However, other features play a role in the formation of NPLs. Such features affect the way the economic and financial variables drive NPLs, as well as they are more difficult to measure. As a result, the inverse relation between good economic performances in a country and NPLs in the banking system of the same country becomes more complex. On the one hand, during a credit boom the economy may be performing very well, though excessive risk taking from financial institutions may arise. This, in turn, may frustrate the beneficial effect of economic growth on NPLs by leading to an excessive lending to non-creditworthy borrowers. On the other hand, economic performances of a country can affect NPLs differently over different stages of the credit cycle, and this might depend upon banks' behaviour and their perception of risks.

Therefore, apart from assessing the macro-financial determinants of NPLs, we investigate whether they drive NPLs differently over different stages of the credit cycle. This represents the first main contribution of the paper to the existing literature. Moreover, we also investigate whether uncertainty, as a proxy of the degree of risk-taking by banks, influences the way some economic determinants affect NPLs, and how this happens over different stages of the credit cycle. This is the second main contribution of this paper to the existing literature.

We find that standard relations between economic and financial variables, and NPLs, change over the credit cycle. Most of the times, the standard relations hold in the lower part of the cycle, while they change in the upper part, becoming weaker, disappearing or even changing sign.

We also find that this is partly due to the way the banks lend, and how they do it over the credit cycle. By relating economic/financial uncertainty measures to the risk propensity of banks, we observe that when there is low uncertainty (which we relate to more risk taking by financial institutions) the beneficial effects of good economic performances reduce or vanish. While, in times of increased level of uncertainty, the standard negative relation between good economic performances and NPLs comes back.

Knowing how NPLs relate to the state of the economy may be useful not only to understand their dynamics. It might also be useful to have a better grasp on how, and at what speed, the risk reduction in the European banking system should proceed, in order to complete the banking union. Moreover, understanding the relation between NPLs and the economic developments in a country can better inform the set-up of scenarios for stress tests of banks/banking systems.

In Chapter 4, we quantify the extent to which differences in sovereign refinancing needs for a given level of debt matter for the perception markets have of sovereign risk. We document sizeable effects. When debt is above sixty percent of GDP, reducing refinancing needs by one percentage point of GDP can translate into ten basis points lower spreads.

In assessing sovereign risk, the key question is whether a government can secure the necessary funds to cover its financing needs. The stock of debt represents the amount of money *borrowed*, but not the flow of obligations required thereafter. Different amortisation schedules, different coupon structure and instruments (e.g., floating, fixed, linked, other currency issuances) can give different meanings to the same debt stock.

This issue came into focus when the International Monetary Fund (IMF) engaged in the euro area. By providing concessional and back-loaded loans, euro official creditors reduced both debt financing costs and the need to roll it over.

This triggered a change in official debt sustainability analyses (DSA). To limit rollover risk, official DSA now monitors whether gross financing needs (GFN) exceed a pre-determined threshold. Gross financing needs capture forthcoming financing needs by adding up interest payments, principal repayments, and primary deficits. Given that liquidity crises arise from mismatches between financing needs and sources, one would expect stress periods to be more likely when the GFN are larger. According to this rationale, while a too large debt stock could signal solvency problems, significant financing needs create liquidity risk. While linking solvency and liquidity to stock and flow debt metrics is logic, setting independent thresholds for level and flow of debt is not enough.

The GFN is very close to the concept of distress barrier as defined in the Contingent Claim Analysis (CCA) applied to the sovereigns. According to the CCA, a worsening of the primary balance, larger interest or principal payments, increases in debt stocks, or any combination of these factors, will bring a country closer to its default barrier, increasing the probability of a debt crisis.

We find that the effect of debt stocks on sovereign risk is dependent on the level of gross financing needs. Countries with large debt levels face more intense pressure if GFN increases. We decompose the effect of GFN into its subcomponent, and find that roll-over needs are the main drivers of this effect.

Our analysis complements the existing literature by showing that jointly considering flow and stock debt measures delivers a more accurate picture of impending risks to sustainability. These findings have two important implications. Firstly, they show that assessing sovereign solvency requires a simultaneous consideration of both flow and stock features of public debt. In fact, our results confirm that focusing on stock and flow metrics separately is more likely to lead to wrong conclusions. Secondly, our results on the role that large debt redemptions and debt stocks have in driving sovereign stress underline one channel through which official lending can enhance its effectiveness in the resolution of fiscal stress.

Chapter 2:

Learning from trees: A mixed approach to building early warning systems for systemic banking crises ♣

Banking crises can be extremely costly. The early detection of vulnerabilities can help prevent or mitigate those costs. We develop an early warning model of systemic banking crises that combines regression tree technology with a statistical algorithm (CRAGGING), with the objective to improve its accuracy and overcome the drawbacks of previously used models. Our model has a large set of desirable features. It provides endogenously-determined critical thresholds for a set of useful indicators, presented in the intuitive form of a decision tree structure. Our framework takes into account the conditional relations between various indicators when setting early warning thresholds. This facilitates the production of accurate early warning signals as compared to the signals from a logit model and from a standard regression tree. Our model also suggests that high credit aggregates, both in terms of volume and as compared to a long-term trend, as well as low market risk perception, are amongst the most important indicators for predicting the build-up of vulnerabilities in the banking sector.

Keywords: Early warning system, banking crises, regression tree, ensemble methods

JEL Codes: C40, G01, G21, E44, F37,

♣ I thank Daragh Clancy, Antonello D'Agostino, Aitor Erce, Ilaria Mameli, Lorenzo Ricci, Rolf Strauch, and seminar participants at the European Stability Mechanism. I also thank for the useful comments the participants of the seminar and poster session where this work was presented. The views herein are the author's and not those of the European Stability Mechanism.

2.1 Introduction

The global financial crisis has led researchers and policymakers around the world to put considerable effort into understanding and preventing systemic banking crises. In doing so, the empirical literature concerned has been focusing on developing early warning systems (EWS) which seek to predict the build-up of dangerous vulnerabilities within banking systems.

The aim of early warning models is not to forecast crisis events. Economic and financial crises are usually triggered by unpredictable shocks. The goal of an EWS is to predict whether a system is vulnerable, to the extent that a sudden adverse shock may lead to a crisis. With regard to banking crises, EWSs turn out to be very useful in giving early signals to the authorities, allowing them to activate pre-emptive actions early enough (e.g., macroprudential measures). The role of policy institutions (national or supranational) is not to predict shocks, but to make sure that when they hit, the economic and financial systems are resilient enough to withstand them.

Typically, early warning models rely on the univariate signalling approach, which allows identifying critical thresholds in a set of indicators. However, this methodology is too simplistic and all the indicators, with their critical thresholds, are unrelated and can deliver counterintuitive signals. Early warning models with a higher degree of complexity rely on the logistic regression technique. However, regression-based models are unable to capture important nonlinearities and complex interactions between macroeconomic and financial variables that may exist in the run-up to crises.[1]

Binary regression trees overcome some of the listed issues as they enable to identify rules-of-thumb with endogenously-determined thresholds in a (ex-post) multivariate framework for a set of interrelated indicators.[2] On the one hand, this means that the ordering of the indicators to look at, and their critical thresholds, are determined by the algorithm and not arbitrarily. On the other hand, every *rule* subsequent to the first one comes from a sub-sample, and this causes the non-global optimality of the final prediction.

The easy interpretability of the results of a binary regression tree, coming from the decision tree structure, is one of the advantages of this methodology, together with other desirable features, which we will describe in the next sections. However, decision trees also come with some drawbacks, amongst which the most important is the low out-of-sample accuracy. This is due to the tendency of decision trees to overfit the data and produce unstable results, as they change when we add new variables or new data points.

---

[1] Potentially, regression models can capture nonlinearities by using interaction terms. However, with a large number of indicators, the number of potential interaction terms becomes very high, unless one does not choose arbitrarily what interaction term to use. In this latter case, there would be a difference between arbitrarily chosen interaction terms for a regression-based model, and data-driven interactions for tree-based models.

[2] By ex-post multivariate framework, we mean that every split of the sample that attempts to separate the events from the non-events, is performed by looking at all the indicators, but eventually using just one of them. We call it ex-post multivariate, because the splitting algorithm is recursive and repeats the same procedure on the sub-sample generated by the previous spits, ending up with a prediction that is based on a series of rules-of-thumb computed in a univariate way, but with each of them, which is true conditional to the previous ones.

An early warning model usually includes a dependent variable, listing the series of events that we are aiming to predict, a set of early warning indicators, chosen by the researcher/analyst according to the literature and expert judgement, and a methodology that uses the indicators to predict the events being studied. Early warning models are not models that predict crises. We use them in order to understand whether imbalances are building up in the economy in such a way that the system becomes more vulnerable and therefore more prone to a crisis. A good early warning model is able to issue accurate warning signals, and, at the same time, it is able to show us where the vulnerabilities are likely to come from in an intuitive way.

Our objective in this paper is to build an early warning model that keeps all these features, so that it achieves the right balance between accuracy and interpretability.

In this paper, we combine some alternative methodologies with the classic techniques available in the literature to build an early warning system for systemic banking crises. We use an ensemble method developed by Savona and Vezzoli (2012) called CRAGGING, which aggregates the results of many "learners" (Regression Trees in this case), and we adapt it to issue early warning signals of systemic banking crises. The paper has the following outline. In section 2, we give an overview of the literature where this paper is placed. In section 3, we describe the data we use, while in section 4, we describe the methodology with a sub-section dedicated to the decision trees and another sub-section dedicated to a methodology that helps to overcome the regression trees drawback, i.e. the ensemble methods. Section 5, outlines the results and some implications, while in section 6, we will describe some robustness exercises we have carried out, before closing the paper in section 7.

2.2 Related literature

Most of the literature on early warning models focuses on two main approaches. The first is the univariate signalling approach (Kaminsky and Reinhart, 1999; Borio and Lowe, 2002; Borio and Drehmann, 2009), which essentially maps the historical time series of a single indicator into past crises and extracts a threshold value for each indicator (independently), above which crises are more likely to happen.

A "second generation" of early warning models estimates the probability of being in a pre-crisis period using a set of several potential early warning indicators jointly. This approach is multivariate and parametric (whereas the signalling approach is non-parametric). A leading example in the literature is given by the logit model (Demirgüç-Kunt and Detragiache, 1998; 2005, Bussiere and Fratzscher, 2006), but recently more formal procedures such as Bayesian model averaging have also been implemented (Babecký et al., 2012).

In order to overcome some of the drawbacks of the two main techniques used in the literature, some have started using Classification and Regression Tree (CaRT) technology to build early warning models. It is a methodology borrowed from computer sciences and only during the last decade, it started being applied to economic and financial studies. A number of papers use CaRT to explore the triggers of sovereign debt crises (Manasse et al., 2003; Manasse and Roubini, 2009; Savona and Vezzoli, 2012 and 2015), currency or balance of payment crises (Ghosh and Ghosh, 2002; Frankel and Wei, 2004) and banking crises (Dattagupta and Cashin, 2011; Davis and Karim (2008), Davis et al., 2011). Alessi and Detken (2014) apply the random forest (RF), a popular statistical method based on the aggregation of the results of a large amount of single decision trees, to the issue of identifying excessive credit and the associated build-up of systemic risk in the banking system. The RF is an extremely accurate predictor and a solid basis for the selection of the relevant indicators. One problem stemming from the exclusive use of the RF is that it does not provide a tree structure like the simple CaRT does. This means that the results come out of a complicated black box and lack of interpretability, especially for monitoring and policy purposes.

We develop an early warning system for systemic banking crises by combining CaRT technology with the CRAGGING algorithm. This ensemble method helps to overcome the drawbacks of the CaRT method (i.e., lack of robustness and poor out-of-sample performances).

Ensemble methods are statistical tools that allow the combination and the aggregation of results coming from large numbers of "single learners" (the generic name for the basic models aggregated using ensemble methods, single decision trees in our case) and that, therefore, help to overcome the weaknesses of single models by aggregating them. Some examples of ensemble methods are bootstrapping and aggregating (BAGGING), Boosting, RF, and the CRAGGING. Alessi and Detken (2014) apply the RF to select the most important variables in identifying systemic banking crises events and then use only those variables to run a single classification tree. An issue with this methodology is that the ordering coming from the RF does not always match the ordering coming from a single decision tree. For instance, the most important variable is not necessarily going to be at the first node of a decision tree run using the full sample.

Savona and Vezzoli (2012) introduced a methodology that makes the BRT results more robust. They use the V-fold cross validation to build a new dependant variable by averaging the predictions of large a number of BRTs estimated by rotating the sub-samples and all their possible permutations of these latter. With the CRAGGING approach, we employ an ensemble method to build a new dependent variable, which carries information from a large number of trees and which is "trained" to do out-of-sample prediction. Therefore, after building this new dependent variable, using it to run a regression tree should help us improve upon the out-of-sample performance of a standard regression tree.

2.3 Data

We use a quarterly (unbalanced) panel over the period that goes from 1980 Q1 to 2015 Q4 for a sample of 15 European Union countries.[3] This panel has the positive feature of including a homogeneous set of economies and banking systems.

2.3.1 The dependent variable

We study a binary dependent variable that captures systemic banking crises events. In the early warning models literature, there are mainly two ways of defining the dependant variable. Indeed, it is possible either to use a continuous stress indicator, like the ECB's Composite Indicator of Systemic Stress, or to use a discrete crises database. We use a binary dependent variable that ranges from 1970 Q1 to 2012 Q4[4] that defines two different states: crisis periods and tranquil periods. This dataset was developed by the European Systemic Risk Board in order to better understand how to implement the countercyclical capital buffer. This dataset is an updated and amended version of the banking crises dataset built by Babecky et al. (2012), where the authors identify the quarters in which European Union countries' banking sectors are in a crisis between 1970 Q1 and 2010 Q4. They do it by combining already existing databases with academic studies and in some cases by supplementing these data with the data coming from a comprehensive survey among country experts. According to the ESRB database, systemic banking crises are periods in which:

> 1) The banking system shows signs of financial distress (non-performing loans above 20% of GDP or bank closures amounting to at least 20% of banking system assets);
>
> 2) Public intervention takes place, in response to (or to prevent) losses in the banking system.

The ESRB amended the resulting dataset as follows:

> 1) Non-systemic crises were excluded;
>
> 2) Systemic banking crises that were not associated with a domestic credit/financial cycle were excluded;
>
> 3) Periods where domestic developments related to the credit/financial cycle could have caused a systemic banking crisis had it not been for policy action or an external event that dampened the financial cycle – henceforth "near misses" – were added.

Tables 1 and 2 in the annex show the dates and length of the crisis events we use in the analysis and provide some descriptive statistics of these events.

[Tables 1 and 2]

[3] There are crisis episodes earlier than 1980, but many indicators we use do not have such long time series. The countries we include in the sample are Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden and the United Kingdom. Our sample choice was dictated by data availability.

[4] We only start using it after 1980 because of the availability of most of the early warning indicators.

As shown in the tables 1 and 2, more than half of the crisis periods are concentrated in the years from 2008 onwards (i.e. the global financial crisis), while the remaining 46% happen between 1980 and 2007. All countries, except Luxembourg, have at least one banking crisis event. The frequency of the crisis events in the sample is of 15.4%, meaning that the distribution of the events is skewed towards tranquil periods (one quarter of crisis every 6.5 quarters). Since in building the EWSs we use the pre-crisis periods as events, and delete the crisis periods from the sample (together with three quarters before the crisis and 4 the four quarters after the crisis), we also compute the frequency of the pre-crisis periods, in order to have the exact distribution of zeros and ones in our dependent variable. Although it is slightly higher, the message remains the same, the frequency of the ones in the dependent variable is about 20% (one pre-crisis quarter every five quarters). This should already hint to the fact that, once we estimate a critical threshold in the probability of being in a pre-crisis period, this probability should be relatively lower than the *naïve* 0.5.

Using such a qualitative crisis events variable has a drawback in the arbitrariness of the choice of the events, but, on the other hand, it gives more clarity in the type of events under study, as opposed to continuous financial distress variables.

As we are building an early warning model, we are not interested in defining whether we are in a crisis, or to forecast when a crisis is coming. An early warning model is rather meant to issue a warning signal (or a series of signals) when, according to the past behaviour of economic, financial and other variables, imbalances are building up in the economy in such a way that a banking system becomes more likely to experience a stress/crisis event (i.e., it is more vulnerable). This means that a good early warning model has to be able to deliver both a relative and absolute measure of vulnerability of banking systems in different countries. In order to do so, we rearrange the binary dependent variable as follows:

- We assign a one to the periods that go from 20 quarters to 4 quarters before the beginning of the crisis;

- We delete the three quarters before a crisis, the crisis periods themselves and 4 quarters after a crisis, to avoid biased results due to the deteriorating (right ahead of a crisis) or deteriorated (crisis and post-crisis) dynamics of the relevant variables;[5]

- Finally, we assign a zero to every other non-crisis period.

2.3.2 Early warning indicators

In order to understand what happens in advance of a banking crisis, and in order to extract an accurate warning signal, consistent with past events, we select as explanatory variables a set of economic and financial variables. For each variable we have downloaded the longest possible series from different sources (Eurostat OECD, BIS, ECB and IMF) starting in 1980 Q1 to 2016 Q4.

[5] Removing the post-crisis periods is suggested also in Bussiere and Fratzscher (2006).

We downloaded the series by giving priority to their length (getting the longest possible series) and to their comparability amongst different countries (for the same variable we downloaded the data if that variable was available from the same source for all the countries). We have in total 30 explanatory variables, which we use as early warning indicators and that we group in six different areas for clarity. The areas are Credit, Real estate, Macro, Global, Financial and Contagion.

BIS data on credit to the private non-financial sector are employed to build nine credit variables. We include in the model credit from all sectors of the economy to the private non-financial sector (broad credit henceforth).[6] Moreover, we also include its breakdown in the form of broad credit to non-financial corporations and broad credit to households and non-profit institutions serving households (NPISH). We take these three measures both as a percentage of GDP and in growth rates year-over-year. Another two variables are built by using bank credit to the private non-financial sector, as a percentage of GDP and year-over-year growth rate. Finally, we also employ as an indicator also the so-called Basel gap, which is the distance of the broad measure of credit to the private non-financial economy from its long-term trend, estimated using a one-sided recursive Hodrick-Prescott filter.[7]

In general, we expect that all these credit variables become "higher than normal" in the pre-crisis periods, signalling that the exposures of the financial sector, as well as of other sectors, are becoming dangerous.

Real estate sector dynamics are very important in explaining banking crises (Aldasoro et al., 2018) and, for this purpose, we use four variables that should help us capture some of these dynamics. Using OECD data, we build the real house price growth rate (deflated by HICP inflation), the house price to income ratio and house price to rent ratio.

Excessive house price inflation creates distortions in the balance-sheet and in the behaviour of both banks and private sector agents, especially because of the collateral value effect. A sudden shock may reverse the house prices dynamics, creating the conditions for a banking crisis.

The macroeconomic variables we included in the model are the real GDP growth rate, the inflation rate, the unemployment rate, the current account as a percentage of GDP, the real effective exchange rate deflated by CPI, the growth in gross fixed capital formation. We also include here the financial sector employment growth, the debt to GDP ratio (in level and the variation) and the overall population. These are all Eurostat data.

---

[6] This data includes the credit to the private non-financial sector (essentially firms and households) from all sectors of the economy (i.e., banks, but also non-bank financial institutions, general government, trade credit, and so on).

[7] We acknowledge that gap variables estimated using the HP filter suffer from a lack of reliability of end-of-sample estimates of the series' trend due to a methodology drawback. Edge and Meisenzahl (2011), using U.S. data, find that ex post revisions to the credit-to-GDP ratio gap are sizable and as large as the gap itself. We think that these issues are important, but given that, once we have estimated the model, we use the last 16 quarters of data to issue signals, only the last ones could be affected. On the other hand, this variable turns out to be very important in splitting pre-crisis from normal times, as it is also used by the BIS, the ESRB, the ECB and the national authorities. For these reasons we prefer to keep it.

The performance of the economy where banking systems reside is important, both because they determine the profitability of banks and the creditworthiness of their customers, and because they contribute to determine the level of risk-taking by banks. We expect that, when an economy is performing well and moves towards overheating, there is a higher likelihood that the conditions for a future banking crisis could appear.

In order to track financial markets dynamics, we use OECD data for the 3-month interbank offer rate as the short-term interest rate, the 10-year government bond yield[8], and the difference between these two as an approximation of the slope of the yield curve[9]. Furthermore, we include the annual growth rate of equity prices (OECD) and the growth rate in the money aggregate M3 (national sources).

These variables, in principle, are important given that they can proxy for the banks' cost of funding, for the sovereign risk (which is always important in European banks given the home-bias in purchasing sovereign bonds), and for the health of the national corporates (the equity price growth).

We also consider the spillover effect coming from other countries' banking systems. Using BIS data on cross-boarder exposures, more specifically the foreign claims of national banks and foreign banks' claims on national banks, we build four variables to proxy for spillover. They are foreign exposures of national banks in percent of total assets of the banking system of the country as well as its annual growth rate, and the exposure of foreign banks to national banks as a percentage of total assets of the banking system as well as its annual growth rate.[10] Another spillover variable that we use is the trade openness, the sum of import and export in percent of GDP.

We expect that these variables signal the build-up of vulnerabilities involving excessive cross-border exposures, especially ahead of widespread banking crises, rather than around isolated episodes.

Finally, we also account for some global determinants by including in the set of early warning indicators the growth rate of global broad credit, the Baa-Aaa spread and the VIX (respectively from St. Louis FRED and CBOE).[11] In particular, we expect that the VIX indicator signals the presence of vulnerabilities when it is too low, given that, historically, a persistently low VIX has corresponded to excessive risk taking by financial institutions (and by economic agents in general). Table 3 in the annex shows the list of indicators used in the analysis.

[Table 3]

[8] We source these data from the ECB, Eurostat and IMF depending on which is the longest series, provided that they have the same definition.
[9] Long series of short-term interest rates on government bonds were not available for all the countries in the sample and we used the interbank rate, as we preferred to be consistent across the sample instead of using different measures to compute it.
[10] We use this measure on immediate counterpart basis instead of ultimate risk basis in order to have longer series.
[11] We also had an indicator of the global economy, i.e. the global GDP growth, but as it was never relevant we dropped it out.

2.4 Methodology

Early warning models are typically based on the analysis of signals from single variables, or on empirical binary dependent variable regression. Both of these methodologies have some desirable features which make them the most utilised in this field. The signalling approach is very simple to apply and to explain to a policy audience. It delivers a number of critical thresholds for a set of indicators, but in this case, all the thresholds are unrelated to each other and it is not possible to effectively judge whether (in our case) a banking system is becoming more vulnerable or not. In fact, it is possible that, using the signalling approach, all the indicators are below their critical thresholds, signalling that there is no concern, but it could be that jointly, these indicators would be able to give a different signal if they could interact.

The binary regression tree helps to overcome this issue, as every prediction is conditional on the interaction of a series of variables and their thresholds. Moreover, when applying the signalling approach to a set of indicator, it does not provide any ordering of importance of the various indicators, whereas the binary regression trees endogenously decides which is the first indicator to look at, then the second, and so on.[12]

Unlike the signalling approach, logit/probit regression allows a multivariate framework. Using this model it is possible to estimate the contribution of each indicator to the increase/decrease in the probability of being in a pre-crisis period. However, it does not (easily) allow the estimation of critical thresholds for the explanatory variables (which is a desirable feature for an early warning system) and their ordering. Furthermore, this framework potentially allows for interactions amongst explanatory variables, albeit there is a limit to the amount of interactions that can be introduced in the equation. However, the researcher would arbitrarily decide these interactions.

We apply the Binary Regression Tree (BRT) methodology to build an early warning system for systemic banking crises. It is a technique developed by Breiman (1984) and widely used in genetics, engineering, marketing, biology, chemometrics and many other scientific fields. We improve upon this methodology by integrating the binary regression tree within the CRAGGING algorithm, developed by Vezzoli and Stone (2007). Given that this methodology is not widely used in economics, we use a more standard logistic regression as a benchmark in order to compare its prediction performances, as it is easier to grasp given its wide use in this field.

2.4.1 Classification and regression Trees (CaRT)

Using Binary Regression Trees, we construct a prediction model that offers a non-parametric framework for uncovering non-linear and interactive structures in the data. It is a partitioning algorithm which recursively and endogenously identifies the variables and the respective thresholds, which are able to split the sample into homogeneous subsamples from the perspective of the dependent variable. Within

[12] However, with the decision tree, every time that the algorithm does a split and grows the tree further down, the new variable and relevant threshold are found over a sub-sample compared to the previous one.

each partition, a simple prediction model is fitted. As a result, the partitioning can be represented graphically as a decision tree. This boosts interpretation and provides greater insights to policymakers.

Regression trees allow for the possibility that different relationships may hold between predictors at different times and under different cross-sectional conditions without having to inflate the set of indicators with hundreds of interaction terms. The algorithm starts by finding the one binary split that delivers the best homogeneity of the dependent variable in the resulting two sub-samples. This split delivers our root node and two child nodes, one where the probability of a crisis increases and another where it decreases as compared to the parent node. The algorithm goes through every possible explanatory variable available in the dataset, and, for each of them, it assesses every value as a possible threshold value to split the dependent variable. The algorithm associates a value to each possible threshold. This value is the weighted average of the mean square errors of the dependent variable in the two subsamples stemming from the use of the assessed threshold. The indicator/threshold pair that minimises the mean square error will characterise the first split (root node), which separates all the observations in the sample in two child nodes. Once the first split is done, the algorithm proceeds recursively to further split the resulting subsamples using the same method, creating more child nodes, and it will continue until a stopping rule becomes binding or a full tree is grown (i.e., until each final leaf contains one observation).

Figure 2 in the annex shows a sample of regression tree, which should help with the terminology and with understanding it better. The tree starts with a root node, which is given by the first split done by the algorithm. The root note splits the sample in two according to the estimated critical threshold of one of the early warning indicators. After the first split, we are working on two separate sub-samples on which the algorithm repeats the same procedure, creating two child nodes. At this point, separate parts of the tree (for instance the root node and the right child note with its predictions are called branches). The final prediction is also called a leaf. If we keep splitting, the last child node at the end of each branch is called a terminal node. Once we have a full tree (it is not possible to proceed to further splits, as there is only one observation in the terminal nodes), we can prune (cut) nodes or entire branches according to an "optimal pruning" rule. Alternatively, we can stop growing the tree before it reaches its maximum degree of deepness (i.e., when all the terminal nodes include one observation).

[Figure 2]

Starting from the root node, each node of the tree has a question and a rule. The early warning indicator and its threshold represent the question. According to our answer to the question, we will have a rule. If we answer 'yes' to the question, we go right, otherwise we go left, to the next nodes (or to the predictions, if we are in the terminal node). For instance, in the sample tree, if we are in the first child node on the left, we are asked the question whether the credit gap is larger or smaller than 3.2. If it is larger, we go to the right, ending up in another node where we are asked about the VIX. Otherwise, if the credit gap is smaller than 3.2, we end up in the child node on the left, where another question will be asked.

In this case, starting from the credit gap, both the child nodes are also terminal nodes and the split inside them will end up with a prediction.

A tree should never reach its maximum deepness, and for this we have already mentioned above possible stopping rules in growing the trees, or an alternative methodology to decrease its size. Possible stopping rules could be a minimum number of observations in the final leaves, a minimum number of branch node observations, a maximum number of levels of splits. When growing a full tree, we can reduce its dimension by merging the final leaves when the sum of the errors (weighted by the relative frequency of 'zeroes' and 'ones' in the leaves) is larger or equal to the error in the parent node or by pruning the tree by following an optimal pruning sequence.

In this paper, we grow a full tree and then we prune it back by using an optimal pruning rule. The optimal pruning rule helps to balance the trade-off between a good in-sample fit of the model and avoiding the over-fitting of the data. It consists of minimising the following error-complexity measure:

$$EC(\alpha) = Err(T) + \alpha \times \#T \quad (1)\text{, where}$$

the Err(T) is the re-substitution error estimate of the tree (the larger the tree the smaller the error), #T is the number of leaves of the tree and α is the complexity parameter and defines the cost of having a larger or a smaller tree. If we would only consider minimising the in-sample error of the model, we would build a very large and complex tree that would not result credible when applied to new data.

The output is a decision tree structure displaying various sequences of conditions to hold in order to obtain different predictions. In our case, the conditions are sequences of thresholds in the relevant explanatory variables, which we use to understand what happens ahead of a banking crisis, whereas the predictions are given by averaging the dependent variable values within each final leaf and represent the probability of being in a pre-crisis period (or the degree of vulnerability of the banking system).

This methodology is particularly useful for building early warning systems for banking crises, as it recognises combinations of vulnerabilities that can trigger crises rather than identifying them in the deterioration of a single indicator. Moreover, the methodology allows us to recognise that economic indicators may have a nonlinear impact on the vulnerability of a country's banking system. Unlike other statistical methods, the binary regression trees method does not need any distributional assumptions on data and it does not assume any underlying functional form. This means that we are not assuming, for example, normality of errors, and we do not assume linearity, additivity or other functional forms. Furthermore, this methodology allows the use of a high number of explanatory variables (even if collinear), can deal with missing values and it is not affected by monotonic transformations of the variables. Last, but not least, the advantage that stems from using this method is the interpretability of the final output. In fact, the tree structure (with its simple rules of thumb) offers a visual outline that significantly simplifies the interpretation of results for policymakers and non-technical audience.

As previously shown, the tree structure is very easy to read. In fact, starting from the root node, what we have is a set of simple rules (i.e., if the indicator is above the threshold value go right, otherwise go left) that end up in a prediction. Hence, conditional to the path, the prediction can be different, and this allows for the fact that not all crises are alike. On the other hand, using this methodology also carries some disadvantages. Indeed, it is not possible to use it to find any causality or to establish any relationships that hold true through the entire dataset. In fact, every time we split the sample (starting from the root node) the relationships become more localised. Another disadvantage is that, being a non-parametric method with no distributional assumptions, there is no possibility to conduct statistical tests on the results or to compute marginal effects. However, it is possible to compute at each node the variation in the probability linked to breaching the threshold.

Another issue is related to the so-called masking problem, where one of two explanatory variables (both very good in explaining the dependent variable) does not show in the final tree because it is very much collinear with the other one. This is comparable to the standard regression analysis, when one of two collinear variables drops out. Regression tree technique tends to be "too" good in-sample as they tend to over-fit the data, at the expense of the goodness of the out-of-sample predictions. As already mentioned, a proper pruning helps to overcome this drawback.

Finally, the main two drawbacks of this methodology are that the model tends to be instable because of the way it partitions the data space and that the variable selected by the algorithm in the first split takes a disproportionate effect on the following choice of predictors and thresholds. In fact, even a relatively small change in the data set (new observations becoming available, or the addition of new indicators to the dataset) can lead to very different trees. As the model is not particularly robust when adding new predictors or observations, other models outperform its out-of-sample prediction ability.

Many contributions in this literature have attempted to overcome these drawbacks in the regression (and classification) tree models by using the "perturbation and combination" approach. The main idea underlying this approach is to create many artificial samples from the original dataset and then perturbing them, in order to estimate multiple models and to generate multiple pseudo out-of-sample predictions, which we then average over time. Breiman (1996) in this direction has introduced the Bagging (Bootstrap and AGGregatING) algorithm, which generates a number of new datasets by bootstrapping the original dataset.

In the second step, we estimate a model for each bootstrapped dataset and aggregate all the predictions by averaging. The aim of this procedure is to reduce potential instability of forecasts and to address the over-fitting problem. Another contribution to this literature comes from Breiman (2001) with the Random Forest algorithm, which is similar to the bagging with the difference that in the random forest algorithm, every bootstrapped dataset has a different dimension and for each of them, only a subset of explanatory variables is selected for prediction. Freund and Schapire (1996), proposed the "Boosting" approach. The idea is to generate multiple simple models for a random portion of the data and then to combine them. As noted, all these statistical approaches generate thousands of different models, and from these models, they generate predictions in the form of a probability. The predictions coming from these models are quite accurate, but they come at the cost of transparency and economic interpretability. Indeed, these approaches are sort of black boxes, which allow for no "economic intuition" and are impossible to interpret.

The CRAGGING algorithm exploits the idea of "perturbation and combination" for achieving predictive accuracy without sacrificing economic intuition. Furthermore, the CRAGGING algorithm can preserve the structure of the data by using in an efficient way the panel data structure of the data. The CRAGGING algorithm improves the stability and the out-of-sample performances of the regression tree model without sacrificing the easy interpretability deriving from the tree structure of the final output.

2.4.2 CRAGGING

In this section we describe the state of the art of the CRAGGING algorithm as in Savona and Vezzoli (2012), and we add the description of an improvement to the empirical strategy.

Let (Y, X) be an unbalanced panel data with N observations. Each unit[13] j of the panel, with $j = 1, \dots, J$, has a number of periods t, with $t = 1, \dots, T_j$ and $J\, x\, T_j = N$. Denote with $L = \{1, 2, \dots, J\}$ the set of units and with $x_{j,t-1} = (x_{1,t-1}, x_{2,t-1}, \dots, x_{r,t-1}, \dots, x_{R,t-1})$ the vector of predictors of unit j observed at time $t - 1$ where $j \in L$ and R the number of predictors for each country. As the name CRAGGING suggests, using the V-fold cross-validation, L is randomly partitioned into V subsets[14] denoted by L_v, with $v = 1, \dots, V$, each containing J_v units and N_v observations[15]. Denote with L_v^c the complementary set of L_v containing J_v^c units and $L_{v/l}^c$ the set where the *l*-th unit is removed by L_v^c $(l \in L_v^c \; and \; L_{v/l}^c \cup l = L_v^c)$.

The cost complexity parameter $\alpha \geq 0$, is the tuning parameter of the cross-validation. Hence, for a fixed α, for each L_v and for each $l \in L_v^c$ let

$$\hat{f}_{\alpha, L_{v/l}^c}(\cdot) \qquad (2)$$

be the prediction function of a single tree trained on data $\{y_{j,t}, x_{j,t-1}\}_{j \in L_{v/l}^c}, t = 1,2,\cdots,T_j$ and pruned with cost-complexity parameter α. The corresponding prediction in the test set is

$$\hat{y}_{jt,\alpha l} = \hat{f}_{\alpha, L_{v/l}^c}(x_{j,t-1}) \text{ with } j \in L_v \text{, and } t = 1,2,\cdots,T_j. \qquad (3)$$

Therefore, at each step, we exclude one unit (country) from the training set. Then, we use the training set without the excluded country, to grow a tree. If this perturbation causes significant changes in the obtained J_v^c trees, the accuracy of the predictors improves by running the following equation:

$$\hat{y}_{jt,\alpha} = \frac{1}{J_v^c} \sum_{l \in L_v^c} \hat{f}_{\alpha, L_{v/l}^c}(x_{j,t-1}) \text{ with } j \in L_v \text{, and } t = 1,2,\cdots,T_j \qquad (4)$$

[13] In our case, units are countries.
[14] In the partition, it is necessary that the number of subsets V is smaller than the number of units J.
[15] The dimension of each subset is of as nearly equal size as possible.

which is the average[16] of the functions (3) fitted over the units contained within the test set $\{y_{j,t}, x_{j,t-1}\}_{j \in \mathrm{L_v}}, t = 1,2,\cdots, T_j$. The objective of the CRAGGING is to improve accuracy, reduce the variance of out-of-sample prediction errors, and reduce the variance in the model selection process.

The first step of the CRAGGING algorithm, called leave-one-unit-out cross-validation, is used for perturbing the training set by removing one unit per time. Furthermore, we have to note that such cross-validation does not destroy the structure of the data, unlike the common cross-validation that partitions the observations randomly. Hence, the CRAGGING algorithm tries to solve the sampling of the observations in the case of dataset with time-varying predictors.

The second step of CRAGGING, called v-fold cross-validation, is implemented on the test sets with $v = 1, \ldots, V$, with the purpose to find the optimal tuning parameter, α^*, that minimizes the estimate of the prediction error on all the test sets. Formally,

$$\alpha^* = \mathrm{argmin}_{\alpha}\ \mathrm{LF}(\mathrm{y}, \hat{\mathrm{y}}) \text{ with } \mathrm{j} \in \mathrm{L}, \text{and } \mathrm{t} = 1,2,\cdots,\textstyle\sum_{\mathrm{j}=1}^{\mathrm{J}} \mathrm{T_j} \qquad (5)$$

where LF(.) is a generic loss function.

The entire procedure described above is run M times to minimize the generalization error, which is the prediction error over an independent test sample, then averaging the results in order to get the CRAGGING predictions to use in the second step. Using the Strong Law of Large Numbers, Breiman (2001a) has indeed shown that, as the number of trees grows larger ($\mathrm{M} \to \infty$), the generalization error has a limiting value and the algorithm does not over-fit the data. As a result, the CRAGGING predictions are given by:

$$\tilde{\mathrm{y}}_{\mathrm{jt}}^{\mathrm{crag}} = \mathrm{M}^{-1} \textstyle\sum_{\mathrm{m}=1}^{\mathrm{M}} \hat{\mathrm{y}}_{\mathrm{jt},\alpha^*} \text{ with } \mathrm{j} \in \mathrm{L}, \text{and } \mathrm{t} = 1,2,\cdots,\textstyle\sum_{\mathrm{j}=1}^{\mathrm{J}} \mathrm{T_j}. \qquad (6)$$

In the third step, a single tree, which we name as Final Tree, is fitted on ($\tilde{\mathrm{y}}^{\mathrm{crag}}$, $\boldsymbol{X}$) with cost complexity parameter $\alpha^{**} = \mathrm{M}^{-1} \sum_{\mathrm{m}=1}^{\mathrm{M}} \alpha^*$. Here, through the replacement of Y with CRAGGING predictions we do four things:

1) Mitigate the effects of noisy data on the estimation process that affect both the predictors and the dependent variable itself;
2) Give the tool a better grasp of how to predict out of sample;
3) Avoid the cliff effect due to the binary dependent variable Y;
4) Grow a final RT that encompasses the overall forecasting ability arising from multiple trees.

Using this process, we obtain a parsimonious model, with good predictions (accuracy), good interpretability and minimal instability. In other words, the second step of our procedure is conceived to deliver a single tree to better understand the complex CRAGGING predictions. This is in line with the idea of

[16] The base learners $\hat{\mathrm{f}}_{\alpha,L^c_{v/l}}(.)$ are linearly combined so that the $\hat{y}_{\mathrm{jt},\alpha}$ will act as a good predictor for future (y|x) in the test set.

assigning the simplest representations to the most accurate models suggested by Catlett (1991) and others.

In order to be able to issue a warning signal, we need to add a further step in the methodology. Indeed, while classification trees predicts whether a new observation classifies as pre-crisis or normal, regression trees return a probability of being in a pre-crisis period. However, a probability does not tell us enough in absolute terms, as for instance 0.4 could be either small or big. For this reason, we apply the signalling approach to estimate a critical threshold value for the dependent variable coming from the CRAGGING routine.

The signalling approach allows us to identify the threshold for predictions coming from the CRAGGING dependent variable as well as from the final tree. The thresholds that we identify are the ones that, in each of the two cases, best separate normal periods from pre-crisis periods in the initial binary dependent variable. The algorithm that implements this method recursively sets each of the predicted values (for both CRAGGING and final tree prediction) as possible critical thresholds, and then classifies the observations: above the threshold, it signals pre-crisis and below it does not.

For each threshold that the algorithm tries, there are four possible outcomes. The final predictions may issue a warning signal and it is correct (A), fail to issue a warning signal when it should have signalled it (i.e., missed pre-crisis, B), issue a warning signal but it is wrong (i.e., false alarm C), not issue a warning signal and it is right (D). These four outcomes fill the confusion matrix (Figure 1).

[Figure 1]

This means that the algorithm will classify the observations in a number of confusion matrices, which is the same as the number of possible thresholds it attempts to do the splits. Then, for each "filled-in" confusion matrix, it computes the value of a loss function. It will eventually use the threshold associated with the lowest value of the loss function. The loss function we use is called the policymaker loss function (PMLF), which is simply a weighted average of the two types of errors, and it is defined as follows:

$$PMLF = \theta P_1 \left(\frac{B}{A+B}\right) + (1-\theta) P_2 \left(\frac{C}{C+D}\right) \qquad (6)$$

where θ is the parameter that defines the policymaker's aversion for the two types of error. If $\theta > 0.5$, we give more weight to the missed alarm rates, with a policymaker that cares more about the potential loss from missing a stress event. If $\theta < 0.5$, we give more weight to the false alarm rate. In this case, the policymaker is more concerned about not issuing too many false alarms, as they are more averse to the potential loss of output given by taking too many pre-emptive measures following the false alarms.

P_1 and P_2represent the relative frequencies of having either of the two final outcomes as a percentage of the total number of observations. This way the weight is also distributed according to which of the two events (pre-crisis and normal times) is more frequent in the sample. This is the most generic version of the PMLF. However, we omit P_1 and P_2 as also in Alessi and Detken (2014). We want to avoid ending up overweighing the false alarms error rate, and consequently come up with thresholds that are too high. In fact, by using these weights in the case of banking crises, we would attribute a larger weight to the false alarm rate because the distribution of the events is skewed towards the tranquil periods. Finally, in the exercises that we carry out in this paper, we set the policymaker's preference $\theta = 0.5$ (the policymaker is equally conservative regarding the two types of errors).

2.5 Results

The outcome of the described routine is a tree structure which, through a series of rules of thumb, provides a prediction about whether we are in a pre-systemic banking crisis period or not, conditional on some of the mentioned rules. The prediction is a value between zero and one, which is interpretable as the probability of being in a pre-crisis period. It can also be interpreted as a systemic risk indicator. Figure 3 shows the resulting final tree.

[Figure 3]

Using the signal extraction method, we can then identify a threshold in this probability by relating it to the binary crisis variable, such that we will be able to say whether an outcome, in the form of a probability, is high or low, and therefore we would be able to issue a warning signal or not. [17]

The outcome could be used either in absolute terms, or in relative terms. In the first case, once we compute the relevant threshold, as said, we could issue a warning signal in case the model suggests it. In the second case, for the panel of countries that we use in the estimation, we could sort the banking systems from the least to the most risky. Apart from the results, the interesting thing is that the tree structure allows us to understand where in the economy the vulnerabilities are building up.

2.5.1 The Final Tree

The estimated final tree shows that, according to the past data, the most relevant variable in splitting normal periods from pre-crisis ones is the bank credit to GDP ratio, with a threshold of about 94%. After this step, we could go either right or left, and find another rule. For instance, if the bank credit to GDP ratio is larger than 94%, we look at the rule on the right, which has as splitting rule the VIX with a critical threshold of 16.7. We need to take the same steps, until we eventually reach a final leaf that shows a prediction in the form of a number between zero and one. We can interpret it in various ways, as, for instance, the probability of being in a pre-crisis period, the degree of vulnerability of a banking system, or as an indicator of systemic risk.

The final tree (Figure 3) shows that, using this methodology, and according to the data we are using, the most important variables to determine whether the banking

[17] The relevant threshold computed using the signalling approach is 0.3.

system is vulnerable are the credit, and the global variables. Regarding the former, the bank credit as a percentage of GDP appears in the root node of the tree, resulting in the most relevant variable when using the entire sample. Year-over-year growth rate of bank credit appears four times, but it is less relevant because for three times it delivers two predictions which are both below the relevant threshold estimated using the signalling approach, as described at the end of the methodology section.

The broad credit measure appears both in distance from its long-term trend (Basel gap) and as a percentage of GDP. The first one, consistently with the early warning literature, is quite relevant as it appears at the first child node, on the left of the root node, which includes a large part of the observations (almost 70%). Finally, amongst the credit measures, also household credit as a percentage of GDP appears once.

The global variables also seem to be important, as they appear three times (the VIX twice and the Baa-Aaa spread once) and, in the nodes where they are the relevant indicator, they include half of the sample.

Other indicators seem to be less relevant, unless they are considered in the broader context of the entire tree.

The final tree mainly shows three different states of the world. One in which the bank credit to the private non-financial sector is high. The other two states of the world have bank credit lower than 94% of GDP, but in one case we are likely to be in a credit boom (Basel gap > 3.2), while in the other case we are not.

In the first case (high level of credit from banks), we end up with four predictions, all of them showing a high probability of being in a pre-crisis period. This signals that banks may be accumulating too much risk in their balance-sheets. In case there is low global market uncertainty, then the model judges the banking system riskier, since the risk perception by economic agents is low. Otherwise, if the VIX is higher than the critical threshold, the model would still issue a warning signal, but there would be a lower probability associated to it.

When bank credit is below its relevant threshold, as mentioned above, we need to look at the Basel gap. If we are in a credit boom, then a low perception of risk (low VIX), and low unemployment (e.g., overheating economy), would trigger a warning signal from our model. However, if the VIX is higher than its critical threshold, the model issues a warning signal only if the year-over-year growth rate of bank credit is above 8.8%.

Finally, when the Basel gap is lower than its critical threshold, our model issues a warning signal only in two cases. One is when broad credit is higher than 153%, signalling that credit to the private non-financial sector is high, despite not coming from banks. In the second case the broad measure of credit is low. However, there is low house affordability (house price to income ratio higher than the critical threshold), combined with high household debt and low perception of risk (low Baa-Aaa spread, which signals that investors are more prone to risks and search for yield by purchasing lower rated corporate bonds).

In order to understand whether the CRAGGING methodology could be of some use, we compare the accuracy of its results with the results coming from other models.[18] Specifically, we compare the in-sample fit of the final tree estimated using the "CRAGGED" dependent variable, where the grouping for the cross-validation was "by country" (Final tree-country henceforth), to the fit of:

- The same model, where the grouping for the cross-validation was "by time" rather than "by country" (Final tree-time henceforth);
- The prediction coming from the CRAGGING (with both types of grouping; CRAGGED-time and CRAGGED-country henceforth);
- A standard regression tree;
- A stepwise Logit;
- A stepwise Logit augmented with interaction terms.[19]

More importantly, we compare the accuracy of the early warning signals, by carrying out an out-of-sample prediction evaluation exercise for the following models:

- The final tree estimated using the "CRAGGED" dependent variable (and where the grouping for the cross-validation where "by country");
- A standard regression tree;
- A stepwise Logit.

2.5.2 In-sample accuracy comparison

In Table 4, we show the results of the in-sample comparison. In the first column, the table shows the value of the PMLF, which is the average of the two error rates shown in the second and third columns (the lower the better). The accuracy rate in the fourth column is the number of times the model issues the right signal, independent of whether it is a warning or not, as a percentage of the total number of observations in the sample. In the fifth column of the Table, we report the Area Under the Receiving Operator Curve (AUROC), which is a performance measurement for classification problems at various thresholds settings.[20] Finally, in column 6 there is the threshold associated with the model used, which is important in the assessment, as it shows that different models, while using the same dataset, could set different thresholds.

[Table 4]

It is clear that there is a trade-off between the missed crisis and false alarm rate, and that, for instance, in order to reduce the missed crises error rate, we would need to give it more weight, and this would lower the critical threshold. However, by lowering the threshold, we automatically increase the false alarm rate, given that now we issue a warning signal more often.

The results show that the standard regression tree and the CRAGGED-time have the lowest value for the loss function and the highest accuracy rate. However, they both achieve this result particularly for the very low rate of false alarms. If we focus on both types of errors separately, we notice that the only model that

[18] The comparison between models will take into account that the tree-based models can deal with missing data, so the datasets used will not be exactly the same than when we use the Logit model.
[19] We also attempted to exploit the partitioning of the data space provided by regression tree methodology in order to choose the interaction terms for the Logit model, however the in-sample improvement was marginal.
[20] The AUROC has to be larger than 0.5 in order for the classifier to be of some use.

has a missed crisis error rate lower than false alarm error rate, is the Final tree-country, which has also the lowest accuracy rate and the second lowest AUROC. The AUROC and the accuracy rate do not distinguish between the two errors, and since the distribution of events is skewed towards the normal periods, the false alarm errors will weigh far more when looking at these two measures. For this reason, we also look at the two error rates separately and at the PMLF in order to compare the performance of different methodologies.

According to the PMLF, the regression tree, the final tree-time and the CRAGGED-time are the models with the best in-sample prediction. The final tree-country is the only model that achieves a missed crises rate lower than the false alarm rate (which, for some monitoring institutions, is desirable). This means that all other models are better (at least in-sample) in predicting when we are not in a pre-crisis period. The Logit has a similar overall result, compared to the final tree-country, but with the figures for the error rates inverted. Apart from the standard regression tree, all the AUROCs are comparable. When we use the tree-based models, we find lower thresholds than when we use the Logit model.

2.5.3 Out-of-sample accuracy comparison: full forecasting horizon

To compare the out-of-sample predictions, we narrow down the set of models for various reasons. We implemented the out-of-sample predictions also for the CRAGGED-country, but the results were similar to the ones from the final tree-country model (slightly worse), and we decided not to report them.[21] Finally, we do not implement the Augmented Logit as the in-sample results were very close to the Logit results without interaction terms. When performing the out-of-sample prediction exercise, having the interaction terms adds complexity to the model, which would have likely deteriorated the forecasting performances.

We implement the exercise by stopping the sample at some point in the past, and predicting whether the upcoming 16 quarters are pre-crisis or not, and repeating it recursively every time a new data point is added. We start the exercise by stopping the dataset in 2000 Q1, and repeating it up to 2008 Q1 each time that we add a new quarter of data. We stop the exercise in 2008 Q1 because with the beginning of the global financial crisis, it becomes more difficult to assess the prediction since most of the "actual" values would be a crisis, but we are predicting-pre-crisis. This means that we would end up with too few observations to compute the relevant statistics and, as we assessed, they become unstable and unreliable. Because of the way the dependent variable is structured, we can only update it up to four years ahead of the current date, because we assign the ones to the pre-crisis periods. This means that, for instance, we may assign a zero to the last three years now, but if a crisis happens within six months from now, then we had assigned zeros erroneously, given that now all the previous 16 quarters ahead of the mentioned crisis will have to be a one. For this reason, we implement the out-of-sample by estimating the models using the data up to when we have the dependent variable available, and use the successive 16 quarters of data to predict out-of-sample. We compute some measures to assess and compare the model for the 16 quarters ahead all together, and the we do the same looking only

[21] Results available upon request.

at the first 8 quarters ahead, and then at the next 8 quarters (from the 9^{th} to the 16^{th}).

Table 5, 6, and 7 show the results. Looking at the 16-quarters-ahead predictions assessment (Table 5), we can say that the final tree-country and the stepwise logit model perform overall similarly, given the similar values of the PMLF. The simple tree out-of-sample prediction performances become (as expected) the poorest of the three models. Looking deeper in the performances of the three models, a few points should be noted. The standard tree is not viable as it makes a missed crisis error rate larger than 0.5. It means that from that perspective, tossing a coin is more reliable.

[Table 5]

Comparing the final tree to the logit, we notice that they achieve a similar PMLF with opposite error rates. As compared to the logit, a gain of 12 p.p. in the missed crises rate (for the final tree) costs 16 p.p. more in the false alarm rate. This means that, overall they perform similarly, according to our PMLF. In absolute terms, if we do not distinguish between the two types of errors, the logit would be more accurate. If, instead, we only look at the missed crisis error (while still getting acceptable false alarm rates, at least below the naïve threshold of 0.5), then the final tree overperforms the logit.

Although the thresholds of the final tree and the logit are not comparable, it is worth mentioning that they have, on average, the same threshold over the various iterations of the out-of-sample predictions.[22] However, this threshold splits the events very differently as it effectively minimises different types of error rates in different cases.

Figure 4 shows the evolution of the missed crisis rate for the three models over the evaluation period. The standard tree is constantly higher than the others, while the final tree has the smaller error most of the time. The missed crisis error rate for all the models, starts decreasing when the global financial crisis approaches, as it becomes easier for the models to issue a right warning signal. The reasons are mostly that the models learn every time that a data point is added, and that there are, at each iteration, more pre-crisis events (ones) in the evaluation sample.

[Figure 4]

Figure 5 shows the evolution of the false alarm rate for the three models over the evaluation period. The logit and the standard tree have the best performance from this perspective over time, but this comes at the cost of an excessive missed crisis rate. The spike in error rates between 2003 and 2004 probably come from the fact that most countries were in pre-crisis period and the models, especially because of global and credit variables, started to issue warning signals for every country (also the ones that were not – yet – in a pre-crisis periods), creating a spike in the false alarm rate.

[22] As already mentioned, the tree-based models can deal with missing data while the logit ones cannot.

[Figure 5]

Figure 6 shows the evolution of the estimated thresholds over the evaluation period. Apart from the standard tree, which has unstable thresholds, the other two models show quite stable thresholds over time.

[Figure 6]

2.5.3.1 Out-of-sample accuracy comparison: shorter and longer horizons

Looking only at the first 8 quarters out of sample (Table 6), the results remain unchanged from the perspective of false alarms, while the missed crisis rates become smaller. Qualitatively the results remain similar, with the final tree-country outperforming the other models on the missed crises rates, and the logit outperforming the final tree-country on the false alarm rate. The standard tree would be even better than the logit on the false alarm rates, but its missed crises rate is 0.5, and it is not acceptable as an early warning indicator.

[Table 6]

Looking at the forecasting horizon that goes from 9 quarters to 16 quarters ahead (Table 7), the performances deteriorate on average for all models. The final tree-country is the only model that has both error rates below 0.5, but it has the highest false alarm rate.

[Table 7]

We observe that the false alarm rate is quite stable over the forecasting horizon, and even improves (marginally) for longer horizons. It means that these models are at least as good at not issuing false alarms in the shorter term, than in the long term when they are not needed. On the other hand, all the models, when passing from shorter horizons to longer ones, become less precise in issuing warning signals when they are actually needed.

Figures 7 through 10, show that the evolution of the error rates for all the models over the shorter and the longer term, is similar to the ones computed over the entire forecasting horizon (16 quarters ahead).

[Figures 7-10]

We believe that our method to build an early warning system represents a good complement to the mainstream early warning system tools. As introduced at the beginning of this paper, it has some desirable features while maintaining a good prediction performance, sometimes outperforming the classic models.

However, it is not a perfect methodology. Apart from the already mentioned drawbacks, we have to point out that the final predictions are not a global optimum. This means that, once the decision tree algorithm finds the first variable and threshold for the first split, it keeps going, until a binding rule stops it (or until all the final leaves include only one observation). However, it could be that another variable, or another threshold within the same variable, could be a worse splitter at that point, but could end up in a better final prediction. However, this is something that would be so "expensive" from the intensity of calculation that to the best of our knowledge, nobody has tried to overcome this issue.

2.6 Further work for implementation and future research

In order to make this methodology implementable within a monitoring institution, there would be additional work to do. One should be able to update independently the banking crises dataset, such that every quarter the model could be run again. The set of early warning indicators needs to be expanded, in order to capture more features (e.g., capital flows, bank specific variables, etc.). At the same time, increasing the countries in the panel, keeping into account possible data constraints, would be of help for practitioners.

We could introduce another layer to the methodology by using (as mentioned earlier in the paper) the random forest technique in order to choose only the most important variables from a larger amount, before going through the CRAGGING procedure.

Finally, this methodology could be transferred also to other types of crises, as it is not bound to work only for banking crises.

Future research would involve a comparison of the results of the CRAGGING where the grouping units are countries with the results of the CRAGGING where the groupings are done according to time and randomly. We believe that, given that we use mainly macro data, with strong interdependence amongst countries and time persistence, when grouping according to time, the results could improve.

2.7 Conclusions

In this paper, we build a multivariate early warning model for systemic banking crises combining a statistical algorithm (CRAGGING) with the regression tree technology. The combination of the two methodologies helps to improve the accuracy of the standard regression trees. The resulting early warning model has a set of desirable features for this class of models. It provides endogenously-determined critical thresholds for a set of indicators, related amongst each other, presenting them in the form of an intuitive decision tree structure. In fact, it takes into account the conditional relations between various indicators when setting early warning thresholds. In doing so, it produces accurate early warning signals as it results from a comparison with the signals of a logit model and of a standard regression tree.

Early warning models are not models that predict crises. We use them to understand whether imbalances are building up in the economy in such a way that the system becomes more vulnerable and therefore more prone to a crisis. A good early warning model is able to issue accurate warning signals, and, at the same time, it is able to show us where the vulnerability is likely to come from in an intuitive way.

An early warning system represents one of the tools that monitoring institutions should have in their toolkit, as a complement to other tools. It should issue warning signals and provide insights on the roots of the vulnerability, so that the practitioners can start a more judgemental analysis in an informed way.

Different institutions might have a different preference/aversion towards either of the two error types. Unfortunately, in classification problems, we cannot minimise both errors at the same time. Therefore, depending on the type of institution, one should give more weight to either of the two error types. The result is that in some cases, the same model applied to the same data could be preferable for one institution but not for another one.

2.8 References

Alessi, L., Detken, C., (2014). "Identifying excessive credit growth and leverage". *Working Paper Series* 1723, European Central Bank.

Aldasoro, I., Borio, C., and M. Drehmann, (2018). "Early warning indicators of banking crises: expanding the family". *BIS Quarterly Review,* 2018.

Babecký, J., Havránek, T., Matěju, J., Rusnák, M., Šmídková, K. and Vašíček, B., (2012). "Banking, Debt and Currency Crises: Early Warning Indicators for Developed Countries". *Working Paper Series 1485, European Central Bank.*

Borio, C. and Drehmann M. (2009): "Towards an Operational Framework for Financial Stability: 'Fuzzy' Measurement and Its Consequences.", *BIS Working Paper No. 284.*

Borio, C. and Lowe, P., (2002). "Assessing the Risk of Banking Crisis". *BIS Quarterly Review, December, 43–54.*

Breiman, L., Friedman, J., Olshen, R. and Stone,C., (1984). "Classification and Regression Trees". *Wadsworth and Brooks, Monterey, CA.*

Breiman, L., (2001). "Random Forests." *Machine Learning* 45 (1): 5–32.

Bussiere, M. and Fratzscher, M. (2006). "Towards a New Early Warning System of Financial Crises". *Journal of International Money and Finance, 25, 953–973.*

J. Catlett, Megainduction, (1991): "Machine learning on very large databases", *PhD Thesis, School of Computer Science, University of Technology, Sydney, Australia, 1991.*

Davis, E.P., Karim, D., (2008). "Comparing early warning systems for banking crises". *Journal of Financial Stability 4, 89–120.*

Davis, E. P., Karim D. and Liadze I. (2011): "Should Multivariate Early Warning Systems for Banking Crises Pool Across Regions?" *Review of World Economics 147, pp. 693–716.*

Demirgüç-Kunt, A. and Detragiache, E., (1998). "The Determinants of Banking Crises in Developing and Developed Countries". *IMF Staff Papers*, *45(1), 81–109.*

Demirgüç-Kunt, A., Detragiache, E., (2005). "Cross-country empirical studies of systemic bank distress: a survey". *National Institute of Economic Review 192, 68–83.*

Duttagupta, R., Cashin, P., (2011). "Anatomy of banking crises in developing and emerging market countries". *Journal of International Money and Finance 30, 354–376.*

Frankel, J. A. and Wei S.-J. (2004): "Managing Macroeconomic Crises." *NBER Working Paper No. 10907.*

Freund, Y., and Robert E. Schapire. "Experiments with a new boosting algorithm." *In Machine Learning: Proceedings of the Thirteenth International Conference, pages 148–156, 1996*

Ghosh, S. R. and Ghosh A. R. (2003): "Structural Vulnerabilities and Currency Crises." *IMF Staff Papers 50(3), pp. 481–506.*

Kaminsky, G.L., Reinhart, C.M., (1999). "The twin crises: the causes of banking and balance-of-payments problems". *American Economic Review 89, 473–500.*

Manasse, P. and Roubini,N., (2009). ""Rules of thumb" for sovereign debt crises". *Journal of International Economics, 78: 192-205.*

Manasse, P., Roubini N., and Schimmelpfennig A. (2003): "Predicting Sovereign Debt Crises." *IMF Working Paper No. 03/221.*

Savona, R. and Vezzoli M. (2012): "Multidimensional Distance-to-Collapse Point and Sovereign Default Prediction." *Intelligent Systems in Accounting, Finance and Management 19(4), pp. 205–228.*

Savona, R. and Vezzoli M. (2015): "Fitting and Forecasting Sovereign Defaults using Multiple Risk Signals." *Oxford Bulletin of Economics and Statistics 77(1), pp. 66–92.*

Vezzoli, M. and C. Stone (2007). "Cragging", manuscript, *Department of Statistics, University of California, Berkeley also published in Book of Short Papers CLADAG 2007, University of Macerata, September 12-14, 2007.*

2.9 Annex

Table 1: Dating of the crisis events

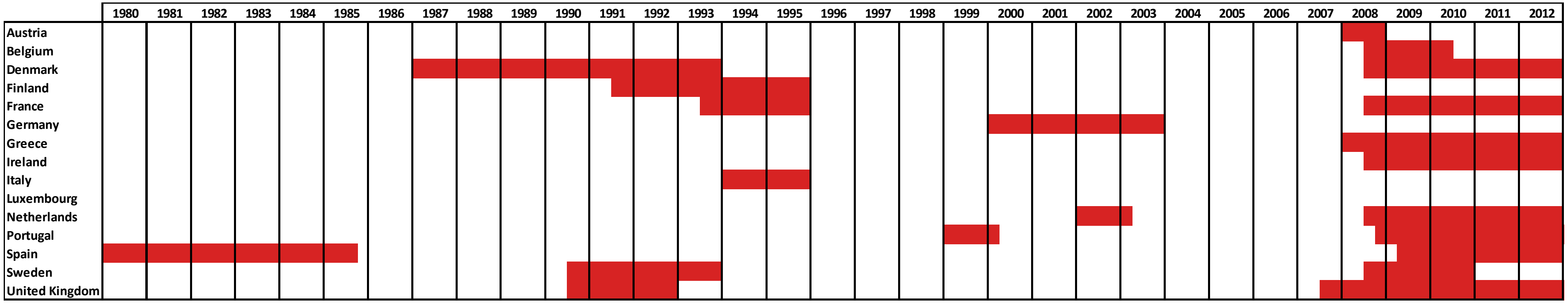


Table 2: Descriptive statistics of the dependent variable

Frequency of crisis events	15.4%
Frequency of pre-crisis periods	20.4%
Share of pre-crisis before 2008	45.6%
Share of pre-crisis from 2008 onwards	54.4%

Table 3: Early warning indicators

Credit	Macro	Financial	Real estate	Spillover	Global
Bank Credit/GDP	Inflation rate	Short term rate	House price/income	Foreign Claims of Banks Own by Nationals (% total Assets)	VIX
Bank credit growth	Unemployment rate	Equity price growth	House price growth	Tot Claim on National banks of Foreign Banks (% total Assets)	Baa-Aaa spread
HH credit/GDP	Real GDP growth	Long term government bond yield	House price/rent	% change in Foreign Claims of Banks Own by Nationals	
HH credit growth	General governent debt (%GDP)	M3 growth		% change in Tot Claim on National banks of Foreign Banks	
NFC credit/gdp	Change in general government debt	Sovereign Yield Curve Slope		Openness	
NFC credit growth	Real Effective exchange rate				
Broad credit/gdp	Employment growth in financial sector				
Broad credit growth	Investment (%GDP)				
Basel gap	Population				

Table 4: In-sample prediction comparison

	PMLF	Missed crisis rate	False alarm rate	Accuracy rate	AUROC	Threshold
Final tree (Time)	0.20	0.27	0.12	0.84	0.86	0.26
Final tree (Country)	0.24	0.16	0.33	0.71	0.79	0.27
Final tree_crag (Time)	0.16	0.22	0.11	0.87	0.91	0.30
Final tree_crag (Country)	0.28	0.29	0.27	0.73	0.75	0.26
Regression tree	0.15	0.18	0.11	0.87	0.95	0.23
Logit	0.25	0.35	0.15	0.80	0.83	0.43
Augmented Logit	0.24	0.35	0.14	0.81	0.83	0.43

PMLF is the simple average between missed crisis and false alarm rates (there could be some approximation error due to rounding). Missed crisis and false alarm rates are as a percentage of the number of observations in their respective actual events (pre-crisis or tranquil). Accuracy rate is the complement to 1 of the total error rate, irrespective of the type of error. The Area Under the Receiving Operator Curve (AUROC) is also an accuracy measure. The ROC curve is created by plotting the true positive rate (A / A+B from the confusion matrix) against the false positive rate (C/C+D from the confusion matrix, false alarm rate) at various threshold settings. The threshold column shows the splitting criteria for the probability of being in a pre-crisis period, computed using the signalling approach as described in the paper.

Table 5: Out-of-sample prediction comparison (16 quarters ahead)

	Missed crises	False alarms	Avg threshold	PMLF
Simple tree	0.61	0.23	0.30	0.42
Final tree (country)	0.34	0.40	0.20	0.37
Logit	0.46	0.24	0.22	0.35

PMLF is the simple average between missed crisis and false alarm rates (there could be some approximation error due to rounding). Missed crisis and false alarm rates are as a percentage of the number of observations in their respective actual events (pre-crisis or tranquil). The average threshold (over the evaluation period) column shows the splitting criteria for the probability of being in a pre-crisis period, computed using the signalling approach as described in the paper.

Table 6: Out-of-sample prediction comparison (1-8 quarters ahead)

	Missed crises	False alarms	PMLF
Simple tree	0.50	0.21	0.35
Final tree (country)	0.28	0.41	0.34
Logit	0.39	0.23	0.31

PMLF is the simple average between missed crisis and false alarm rates (there could be some approximation error due to rounding). Missed crisis and false alarm rates are as a percentage of the number of observations in their respective actual events (pre-crisis or tranquil).

Table 7: Out-of-sample prediction comparison (9-16 quarters ahead)

	Missed crises	False alarms	PMLF
Simple tree	0.64	0.26	0.45
Final tree (country)	0.42	0.39	0.40
Logit	0.52	0.17	0.34

PMLF is the simple average between missed crisis and false alarm rates (there could be some approximation error due to rounding). Missed crisis and false alarm rates are as a percentage of the number of observations in their respective actual events (pre-crisis or tranquil).

Figure 1: Confusion matrix

		Actual	
		Pre-crisis	Normal
Signal	Pre-crisis	A	C
	Normal	B	D

Type I error rate = $\frac{B}{A+B}$; Type II error rate = $\frac{C}{C+D}$

The EWS can have final predictions that:

1) Issues a warning signal and it is correct (A);
2) Fails to issue a warning signal when it should have signalled it (i.e., missed pre-crisis, B);
3) Issues a warning signal but it is wrong (i.e., false alarm C);
4) Does not issue a warning signal and it is right (D). These four outcomes fill the confusion matrix.

Figure 2: Sample regression tree

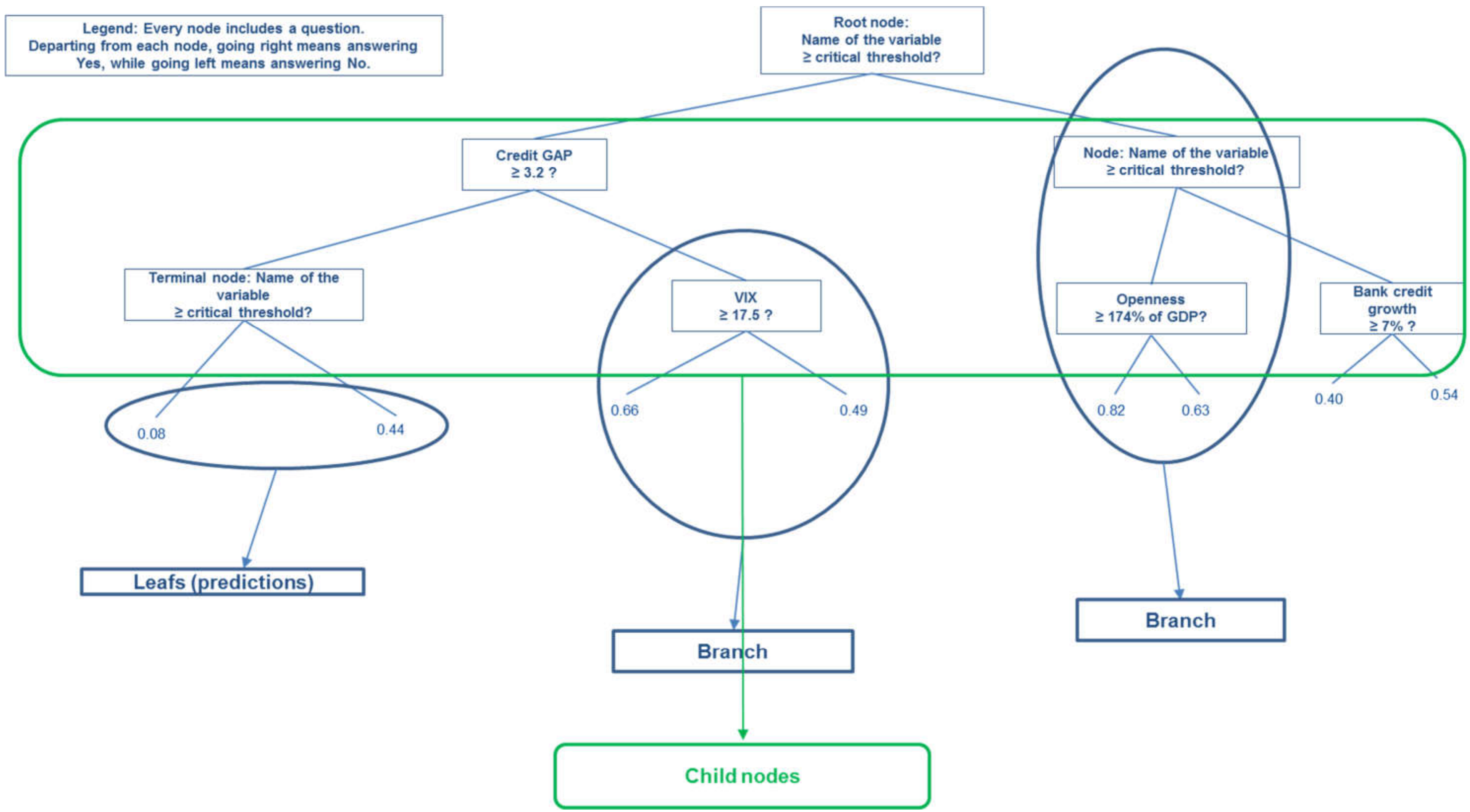


Figure 3: The final tree (predictions are probabilities)

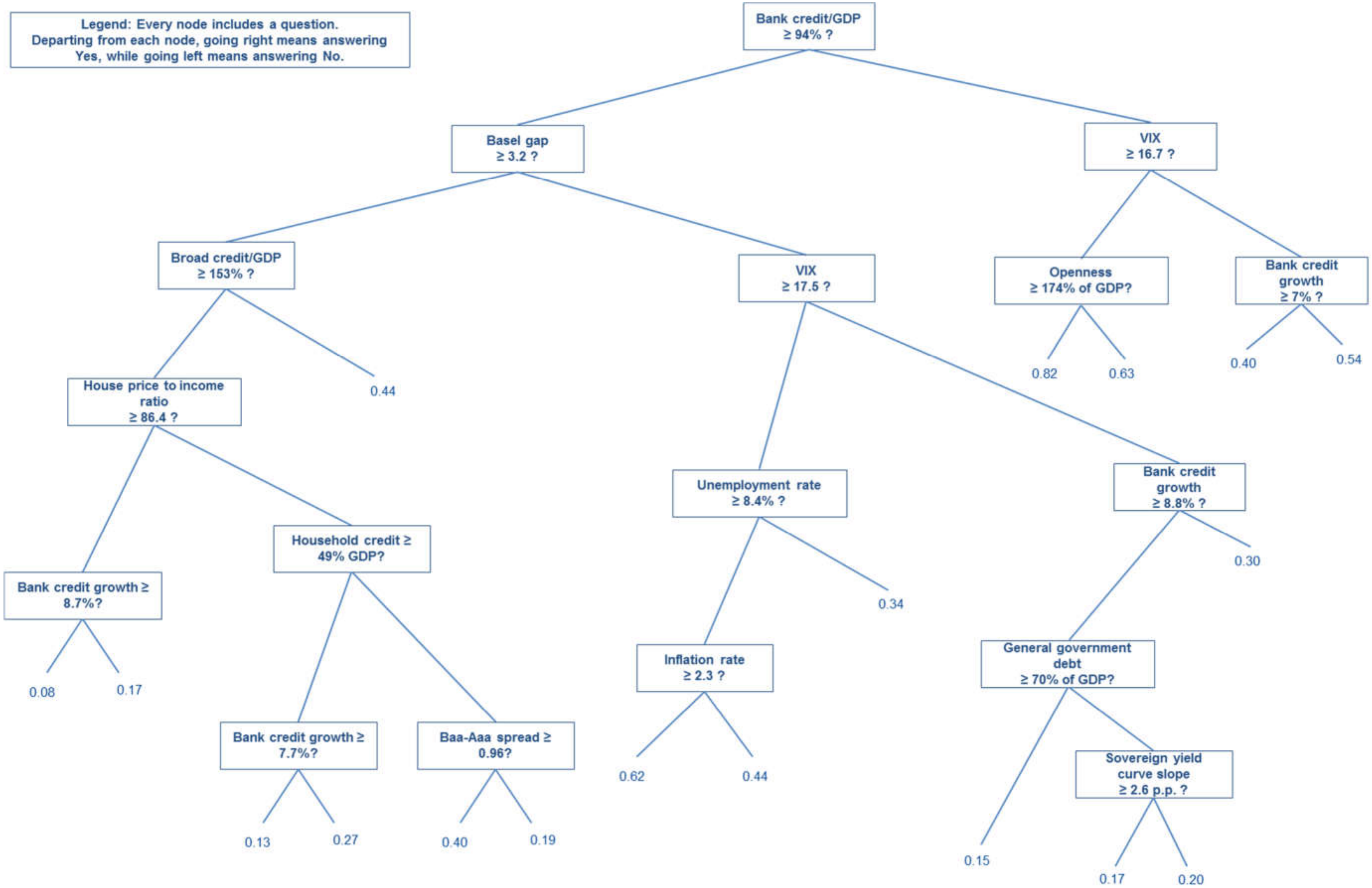
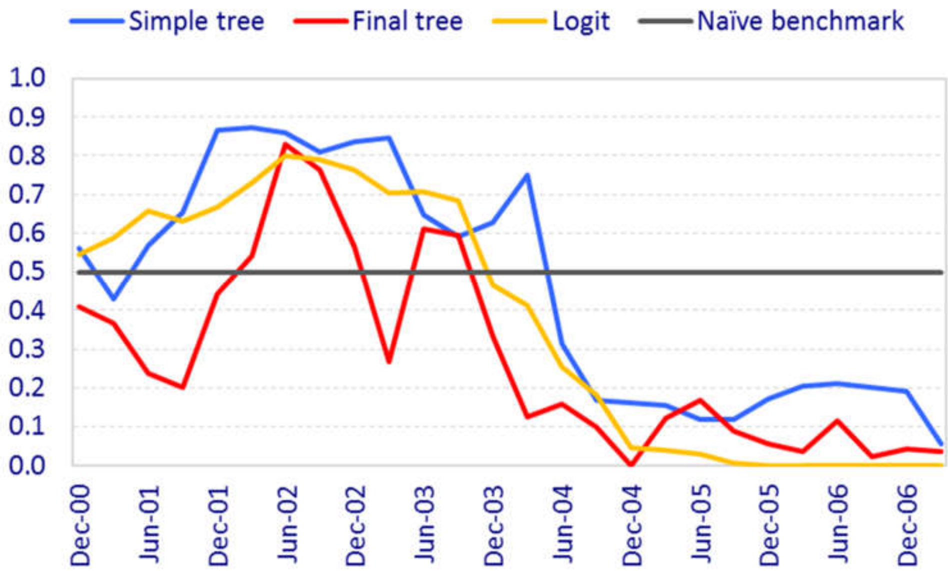
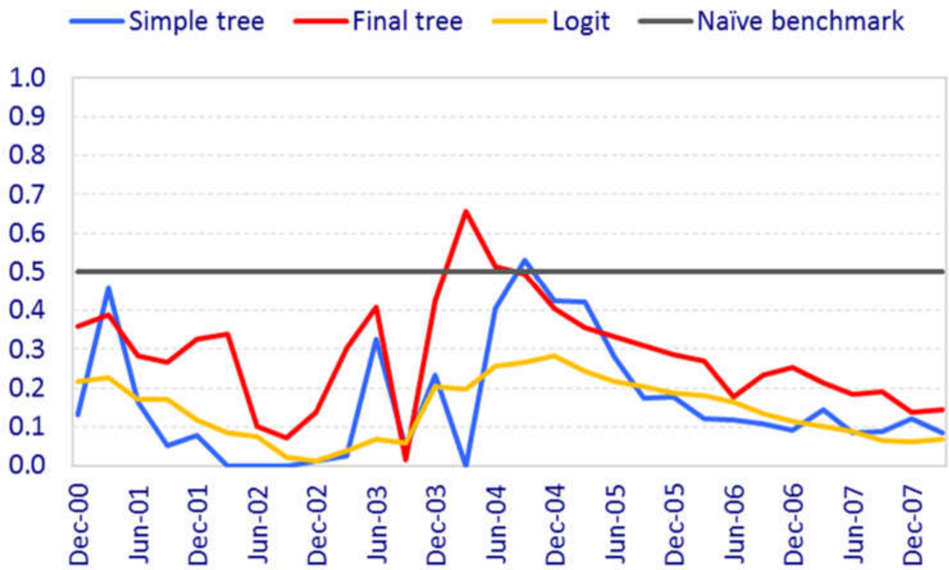


Figure 4: Out-of-sample missed crises rate



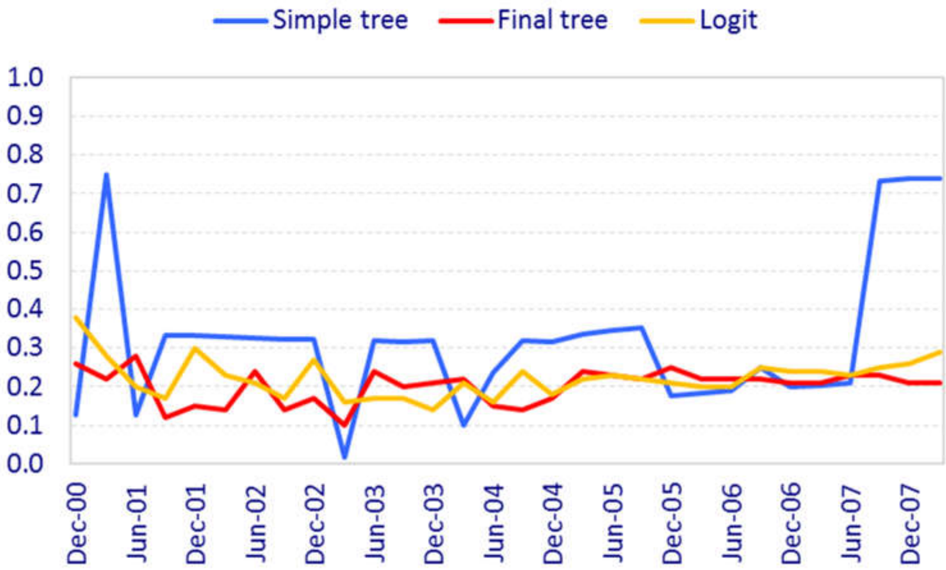
Vertical axis is in percent.

Figure 5: Out-of-sample false alarm rate



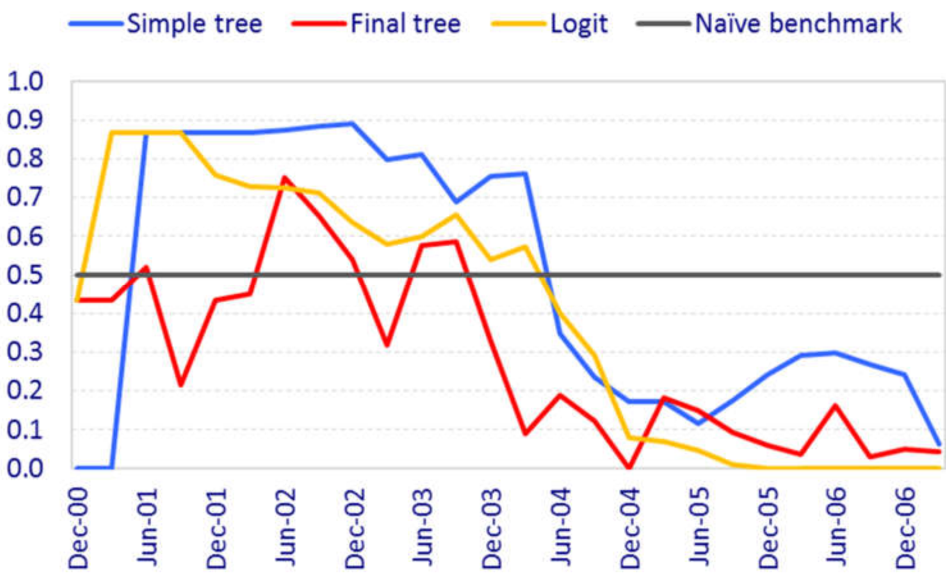
Vertical axis is in percent.

Figure 6: Out-of-sample critical thresholds



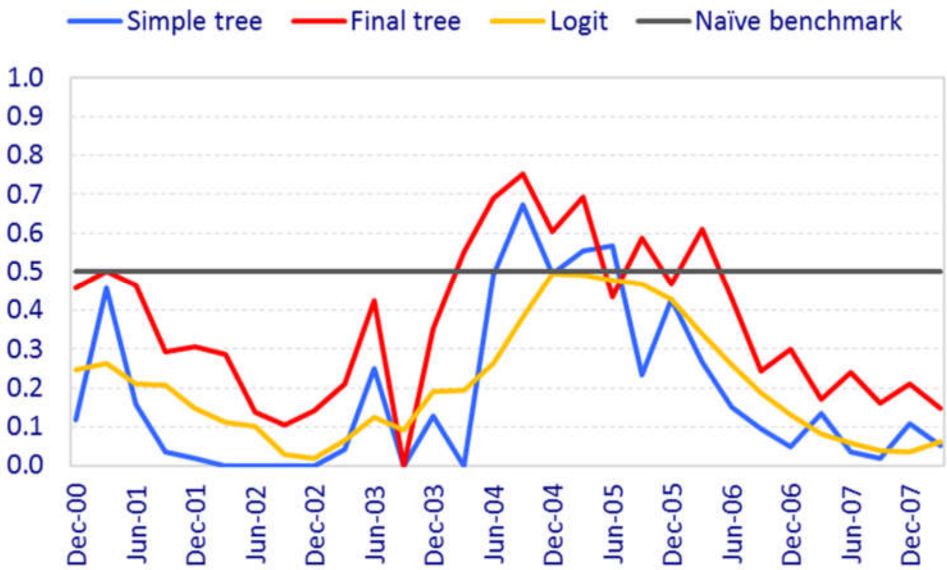
Vertical axis represents a probability.

Figure 7: Out-of-sample missed crises rate (1-8 quarters ahead)



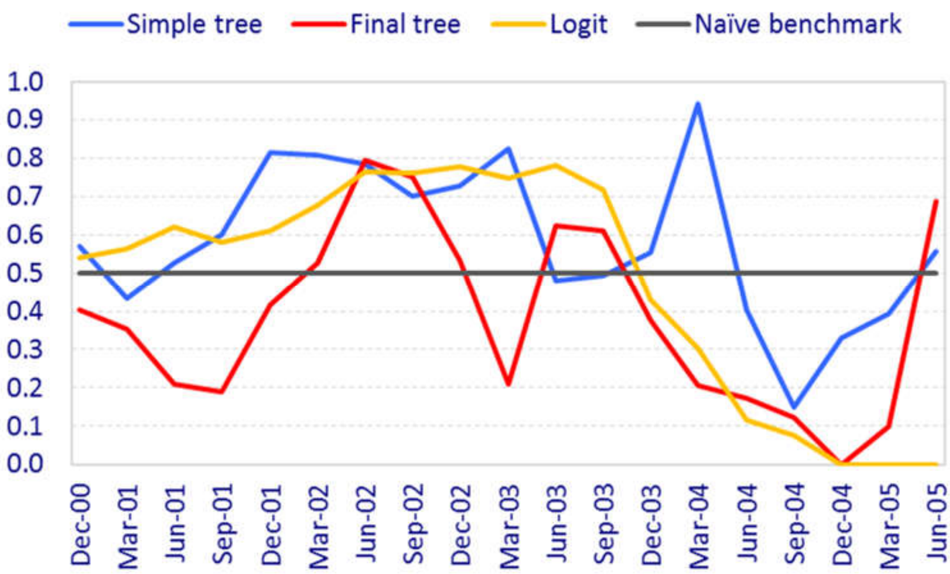
Vertical axis is in percent.

Figure 8: Out-of-sample false alarm rate (1-8 quarters ahead)



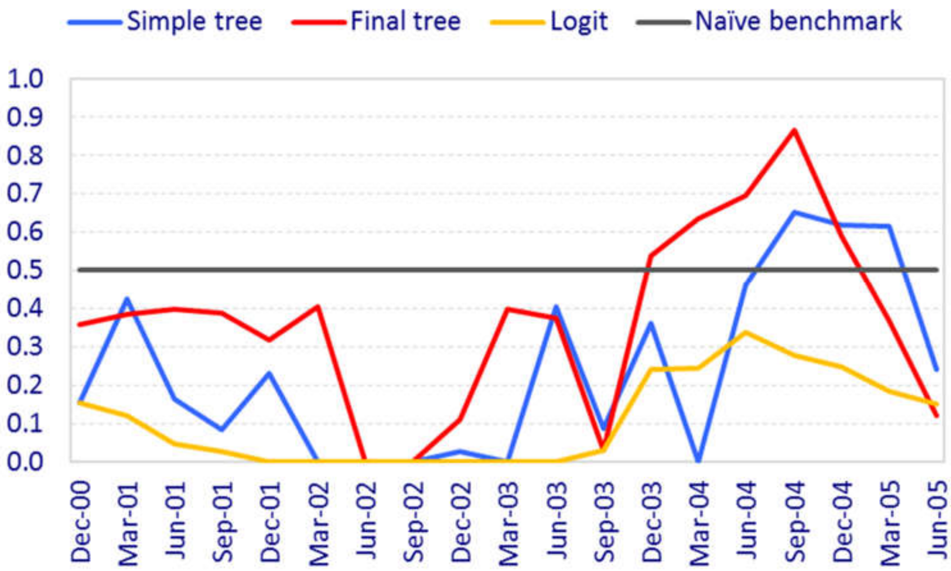
Vertical axis is in percent.

Figure 9: Out-of-sample missed crises rate (9-16 quarters ahead)



Vertical axis is in percent.

Figure 10: Out-of-sample false alarm rate (9-16 quarters ahead)



Vertical axis is in percent.

Chapter 3

Macro-financial determinants of NPLs over the credit cycle, and the role of uncertainty[*,♦]

In this paper, we assess the macroeconomic and financial determinants of Non-performing loans (NPLs) in the euro area. We investigate the role of these determinants in affecting NPLs over the credit cycle, and we analyse the role of uncertainty. We find that some of the determinants affect NPLs differently over the different stages of the credit cycle. In particular, when the credit cycle is positive, some of the standard relations weaken or change. We also find that economic and financial uncertainty may play a role, through the different degrees of risk-aversion of the banks at different levels of uncertainty. Low uncertainty, if associated to low levels of risk-aversion, could determine excessive risk taking by financial institutions, especially during credit expansions, which could frustrate the positive effects that economic growth has on NPLs during *normal times*.

Keywords: *Non-performing loans, credit cycle, uncertainty*

JEL Codes: *F3, G21, D81, H81, E44, E59*

[*] European Stability Mechanism, Circuit de la Foire Internationale 6a, L-1347, Luxembourg.

[♦] Corresponding author: C.Gabriele@esm.europa.eu

We thank all the seminar participants at the European Stability Mechanism. The views herein are the authors' and not those of The European Stability Mechanism or its Management Board.

3.1 Introduction

Large economic and financial crises usually leave as a legacy deteriorated assets in financial institutions' balance sheets. This represents a considerable issue, as high Non-Performing Loan (NPL) stocks impair the supply and an efficient allocation of credit to the economy. In fact, banking system crises are generally associated with the most painful and long-lasting recessions.[23]

Over the last decade, we observed a general increase in NPLs right after the global financial crisis (GFC). In some European countries, this increase continued for a prolonged period, mainly because of the second economic dip due to the sovereign debt crisis. After the double-dip, we observed a general, and in some cases slow, decrease, probably due to the economic recovery.

High stocks of NPLs are problematic for the performance of European banks. Firstly, they reduce banks' income and their profitability; secondly, their existence absorbs large amounts of banks' resources. As a result, the lending capacity of the banking system might be compromised creating in turn a negative effect on the real economy. For these reasons, the reduction in NPLs has often been discussed at the European level especially in view of finalizing the still incomplete Banking Union. According to the European Commission, completing the Banking Union remains one of the key policy objectives for deepening the Economic and Monetary Union.[24] Reaching a preliminary risk reduction objective and building a risk-sharing platform are considered fundamental steps in this direction and, in this regard, the European Institutions have frequently underlined the importance of a reduction in NPLs. In the last years, the debate on the topic has been quite active in the European *fora* and, in July 2017, it culminated with the European Commission and European Parliament launching an *Action Plan to Tackle NPLs in Europe*.

The aim of the plan was to harmonise non-performance management procedures and debt recovery frameworks, to develop a secondary market for NPLs[25], and to incentivise banks with NPL ratio larger than 5% to introduce a specific action plan for managing and reducing NPLs in their business strategy.[26]

Given the emphasis posed on these topics, it is quite important to study how NPLs relate to the state of the economy in order to understand their dynamics and to have a better grasp on how, and at what speed, the risk reduction in the European banking system should proceed, in order to complete the Banking Union.

When studying the evolution of NPLs it is worth looking at specific characteristics of banks. It is, nevertheless, equally important to understand what the macroeconomic and financial determinants of the deterioration in the financial system assets are. These two approaches are not alternative but complementary, and provide two different perspectives. One analyses the features of the banking business and of the banks' balance sheet, to understand the evolution of NPLs.

[23] See Boissay et al. (2016).
[24] See European Commission Communication of 11 October 2017.
[25] The proposal includes the development of a new platform with common rules and standards, which will eliminate the typical asymmetries existent in this market.
[26] See EBA Guidelines and European Central Bank (ECB)'s *Guidance* and *Addendum* for banks on NPLs.

The other looks at the issue from a systemic perspective, trying to reconcile economic and financial vulnerabilities with the build-up of NPLs. In this paper, we focus on the last approach, as we believe that the economic and financial conditions of a country are fundamental determinants of NPLs' dynamics. In fact, good economic performances should help curbing the build-up of problematic loans in the banking system balance sheet. In the same way, the banking system of a country that is experiencing a period of weak economic conditions is more likely to accumulate NPLs, at a faster pace. It is also important to mention the role of feedback loops, given that an impaired financial system balance sheet, by reducing the possibility for households and corporates to borrow, determines a lower economic output.

When we think about the effects of macroeconomic and financial dynamics on the formation of NPLs, we implicitly (in most cases) refer to a channel linking the economic and financial conditions in a country to the borrowers' capacity to repay their debt. However, other features play a role in the formation of NPLs. Such features affect the way the economic and financial variables drive NPLs, and they are more difficult to measure.

In some cases, borrowers unable to serve their debt may still be able to borrow money, regardless of their economic or financial conditions. In these cases, banks show either a low risk aversion or a poor screening activity of borrowers. Usually this happens during periods of credit expansions/booms. During these periods, we also observe rapid credit growth, which is positively associated to future loan losses.

Against this background, the inverse relation between good economic performances in a country and NPLs[27] in the banking system of the same country becomes more complex. During a credit boom, the economy may be performing very well. Nevertheless, the excessive risk taking from financial institutions may frustrate the beneficial effect of economic growth on NPLs, and it may lead to an excessive lending to non-creditworthy borrowers. On the other hand, economic performances of a country can affect NPLs differently over different stages of the credit cycle, and this could depend upon banks' behaviour and risk perception.

In this paper, we decided to focus on credit cycles because their peaks are usually followed with high probability by banking crises, currency crises or sudden stops episodes (Mendoza 2011, Eichengreen and Arteta, 2002).[28] In particular, we investigate whether economic and financial variables drive NPLs differently over different stages of the credit cycle. This represents the first main contribution of the paper to the existing literature. Moreover, we also investigate whether uncertainty, as a proxy of risk taking by banks, influences the way some economic determinants affect NPLs, and how this happens over different stages of the credit cycle. This is the second main contribution of this paper to the existing literature.

The paper has the following outline. In section 2, we provide readers with an overview of the literature in which we place this paper. Section 3 shows the data

[27] Erdiç and Abazi (2014); Us (2017);Jimenez and Saurina (2006)

[28] These events occur in 44% of the episodes studied in Mendoza (2011) and in 8 out of 12 episodes studied in Reinhart and Rogoff (2009).

and the variables used in the analysis, and in section 4we explain the methodology used. In section 5, we present and discuss the results, and in section 6we discuss more results obtained as robustness checks. Finally, in section 7, we conclude and indicate the way for future research on this topic.

3.2 Literature review

The existing literature shows that both macroeconomic and banking related metrics determine the NPLs dynamics. Indeed, we can link the build-up and/or the ability to resolve NPLs to both the external environment, i.e., the economic performance of a specific country or of the world economy, and to the operations of a specific bank.

Regarding the link between NPLs and the external environment, the literature shows us that this usually occurs with a lag that can vary according to the specific variable taken into account (Bofondi and Ropele, 2011).[29] A period of economic expansion is usually associated with an increase in consumers' income and companies' revenues. This implies an increase in the ability to pay back debts that in turn reduces the amount of NPLs. For this reason, a positive real GDP growth rate is negatively associated with NPLs.[30] The consensus is quite strong around these results that hold for groups of countries (Messai, 2013; Erdiç and Abazi, 2014; Nkusu, 2011; Beck, et al, 2013) and for single country case studies (Salas and Saurina, 2002; Jimenez and Saurina, 2006 for Spain; Bofondi and Ropele, 2011 for Italy; Us, 2017 for Turkey; Gosh, 2015 for the US; Louzis et al, 2012 for Greece). [31]

A similar logic applies to the positive link between unemployment rate and NPLs. A period of unfavourable labour market conditions increases unemployment. The lack of jobs reduces both current and future households' purchasing power. Debt servicing becomes more difficult potentially increasing NPLs even faster (Bofondi and Ropele, 2011; Gambera, 2000; Babouček and Jančar, 2005; Gosh, 2015; Louzis et al, 2010).

More controversial is the role of consumer prices. Inflation, indeed, might produce a double effect on NPLs. On the one hand, it makes debt repayment more difficult by envisaging higher nominal interest rates when loan rates are variable (Nkusu, 2011) or by reducing real income when wages are sticky (Michail, 2018; Nkusu, 2011). This should imply a positive relation between inflation and NPLs (Rinaldi and Sanchis-Arellano, 2006; Bofondi and Ropele, 2011; Erdiç and Abazi, 2014; Us, 2017; Gosh, 2015; Nkusu, 2011). On the other hand, inflation makes debt

[29] A positive real GDP growth for example reduces NPLs after 4-3 quarters. Unemployment rate instead seems to have a simultaneous effect.

[30] The literature also shows that prolonged periods of economic expansion might produce a loosening of credit conditions where credit is granted without considering the quality of the receivables (Messai, 2013). There is potential reverse causality between macroeconomic conditions and NPLs given that high NPLs are likely to affect negatively economic growth and to reduce the economic efficient allocation of resources. An increase in NPLs makes indeed banks more risk averse. This, combined with the physiologic increase in provisions, reduces their lending (Leon and Tracey, 2011) and increases the cost of borrowing (Accornero et al, 2017). See Nkusu (2011) for a study on the feedback generated between NPLs and their determinants.

[31] With a focus on Italy, Mohaddes et al. (2017) show that a sustained real GDP growth larger than 1.2% might half NPLs in five years.

servicing easier by reducing the real value of debt (Babouček and Jančar, 2005) justifying then a potential negative link with NPLs as well.

Strictly related is the role of interest rates, which we can consider a proxy for the cost of borrowing. The larger this cost, the higher the probability that debtors will not be able to service their debt (Cerulli et al., 2017; Jimenez and Saurina, 2006; Kalirai and Scheicher, 2002). Beck et al (2013) show how a rise in lending rate resulted in an increase in NPLs in 75 countries during the last decade.[32]

The literature also provides sporadic evidence for other macro variables such as: the level of judicial inefficiency in one country (Cerulli et al,2017), confidence indicators and stock market returns (Kalirai and Scheicher, 2002; Michail, 2018; Beck et al, 2013), the growth rate of property price, and the change in output gap (Hoggarth et al., 2005).

A recent branch of the literature started looking at the decomposition of NPLs between households and corporates. As these two groups are quite different, the determinants of NPLs might then differ as well. Bofondi and Ropele (2011) apply this decomposition to Italy in 1990-2010. They show that, for households, higher GDP growth rate and house prices are associated with lower NPLs whereas unemployment and short term nominal interest rate are associated with larger values. For firms instead, unemployment rate shows a positive relation with NPLs whereas a negative one is observed for the annual growth rate of durable goods consumption.[33] A further decomposition is realized in Louzis, Vouldis, and Metaxas (2010) whom divide between consumer, business and mortgage loans in order to study the determinants of NPLs in Greece. According to their results, real GDP growth is the most important determinant for consumer loans while unemployment is for business loans. Mortgages are instead less influenced by macro variables.

As mentioned above, the source of NPLs can also be internal to the banks. Amongst bank related metrics, it is established in the literature that an excessively loose credit policy might sow the seeds of future financial distress (Nkusu, 2011). An excessively loose credit policy can have multiple causes. It might be generated by herding behaviour (Kindleberger, 2000; Jimenez and Saurina, 2006), by an underestimation of risks (Boz and Mendoza, 2011; Borio et al., 2001). Further determinants could be a lowering of credit standards (Keeton, 2003; Dell'Ariccia and Marquez, 2006), and the presence of government guarantees (Corsetti et al., 1999). Such loosening is considered an important determinant of a reduction in the quality of loans (Keeton and Morris, 1987; Sinkey and Greenwalt, 1991; Keeton, 1999; Jimenez and Saurina, 2006; Erdiç and Abazi, 2014).[34] Similar is the effect that aggressive credit policies and a poor bank management can produce on NPLs (Berger and DeYoung, 1997; Kwan and Eisenbis, 1997; Podpiera and Weill, 2008; Salas and Saurina, 2002). From the positive side instead, profitability and net interest margin reduce NPLs (Gosh, 2015; Messai, 2013).

---

[32] Similar results were found by Erdic and Abazi (2013) for the emerging economies in Europe, by Louzis, et al, (2010) for Greece and by Caporale et al (2014) for Italy.

[33] A similar approach, which separates between households and firms, is used also in Michail (2018) for Cyprus.

[34] The causal effect of rapid/excessive credit growth on NPLs is often found with a lag (Jimenez and Saurina, 2006; Clair, 1992; Kwan and Eisenbeis, 1997; Keeton, 1999; Salas and Saurina, 2002).

Gilchrist et al (2014) show that uncertainty shocks negatively affect credit-supply.[35] Valencia (2013) shows that credit-supply decreases when uncertainty increases. This is relevant to our analysis, as we investigate the role of uncertainty (and its possible relation to the risk-aversion of banks) in affecting the way macroeconomic and financial variables affect NPLs formation. As suggested by Bloom (2009), a good proxy for uncertainty is provided by the financial markets' expectation of future volatility based on options on the S&P Index (VIX and VXO). Bekaert et al (2013) show that the VIX Index can be decomposed in an uncertainty component and a risk-aversion component.

3.3 Data and empirical methodology

In our analysis, we use a dataset that starts in the last quarter of 2007 and ends in the third quarter of 2019 for our dependent variable. The explanatory variables start earlier given that there is broader availability, and that we lag them in our equations.

3.3.1 Non-performing loans

Our dependent variable is NPLs, which are usually defined as loans that are either more than 90 days past-due, or that are unlikely to be paid in full.[36] Following the existing literature, we focus on the ratio of NPLs to total gross loans as this allows comparisons among countries by taking into account the specificity and size of each financial sector. We compute this ratio at country level from the consolidated balance sheet of a sample of banks included in the Single Supervisory Mechanism's (SSM) list of significant banks.[37] We download and mix consistently the data from Fitch Connect and SNL Financial databases. Our dataset for NPLs covers 19 euro area countries at a quarterly frequency from end 2007 to June 2018.

As shown in Figures 1 and 2 (Annex 2), the NPL ratio in the euro area has been increasing since the beginning of our sample at the end of 2007. It reached its peak in 2014 and since then it has been on a declining trend. Looking at NPL evolution at country level, we can notice a high degree of heterogeneity across countries. Although we can observe that the trend is homogeneously decreasing for most of the countries in the Euro area, NPLs in Greece and Portugal show an opposite pattern with peaks respectively of 47.41% and 12.25%.

[Figures 1 and 2]

Additionally, while NPL ratio averaged at 9.18% in the euro area, important differences can be noticed across countries. The country with the largest NPL ratio was Cyprus which experienced an average level of 35.95% whereas Finland was the country with the lowest level averaging only at 1.54%. Additionally, Cyprus,

[35] Alessandri and Bottero (2016) find that economic uncertainty it reduces banks likelihood to accept new credit applications

[36] Given that the precise definition of NPLs varies across jurisdictions, international comparison can be complicated (Beck, Jakubik, Piloiu, 2013) and results should be interpreted with caution.

[37] For Cyprus data we use aggregate data for all domestic and foreign credit institutions operating in Cyprus on a consolidated basis, as reported by the Central Bank of Cyprus according to the evolving definition perimeter of non-performing loans.

Ireland, Greece, and Slovenia all reached NPL ratios larger than 30% in the sample of our analysis whereas Belgium, France, Finland, Germany, Luxembourg, Malta, Netherlands, and Slovakia instead, always remained below 10%. Heterogeneity also appears with reference to the maximum values observed within countries and to their timing. While Cyprus experienced a maximum level of NPL ratio of 49.45% in the second quarter of 2016, Finland reached only a maximum of 2.02% in the first quarter of 2018.

3.3.2 Explanatory variables and expected results

Following the existing literature, we use macroeconomic and macro-financial indicators as explanatory variables.

As macroeconomic indicators, we use the real GDP growth rate, computed as year over year percentage change, and the unemployment rate (both are seasonally adjusted) to take into account how the economic performance of a specific country can negatively affect the soundness of its banking system. We expect that as the GDP increases, borrowers increase their ability to repay their debts. Conversely, a slow down or a negative economic growth might produce a reduction in companies' and households' cash flows, making it more difficult to repay bank loans. Same logic, but with reverse sign, applies to the link between unemployment rate and NPLs.

Country-specific institutional settings might play an important role when dealing with NPLs. A sound and efficient judicial system might be indeed fundamental to ease restructuring processes and to make faster NPLs' recovery. In such circumstances we consider as additional explanatory variables both regulatory quality and the degree of rule of law for each country.

To follow, we take into account the ratios of general government consolidated gross debt and of general government deficit/surplus to GDP in order to proxy for the government's fiscal space and its potential ability to intervene in case of problems in the domestic financial system.

The additional variables used in our analysis, which are still related to macroeconomic conditions, are the real effective exchange rate and the inflation rate. An exchange rate appreciation, by representing a deterioration in international competitiveness, might result in larger NPL ratio especially for export-oriented industries. A currency depreciation instead, increases debt servicing costs in domestic currency terms for borrowers with foreign currency-denominated loans. If such borrowers are not hedged against a depreciation, then defaults on loans denominated in foreign-currency will increase. Conversely, in countries with no currency mismatches, an exchange rate depreciation could be beneficial in term of NPLs since it might boost export and then improving the corporate sector's financial position. Similarly ambiguous is the link between inflation and NPLs. On one hand, higher inflation helps the borrower by reducing the real value of outstanding loan and, in turn, by making debt servicing easier. On the other hand, higher inflation reduces borrowers' real income when wages are sticky. We consider the growth rate of consumer price index, measured as year on year percentage change, to proxy for inflation.

As macro-financial indicators, we focus on credit growth rate, which is computed as the year on year change of credit provided to the private non-financial sector, from all sectors of the economy. This variable might give some broad information on the lending policies adopted by financial institutions. A fast growth rate in credit might for example signal potential easiness in quality standards, such as screening of borrowers and collateral requirements, and inadequate risk management (Jimenez and Saurina, 2006). Strictly related is the information coming from the credit cycle, which is usually measured in the literature taking the credit-to-GDP gap (Borio, 2012). They define this variable as the deviation of credit to private non-financial sector from all sectors of the economy from its long-term trend. To simplify the results, and to make the credit gap variable stationary, we transform it into a dummy variable that takes value 1 if the credit gap is positive, and zero if it is negative.

Given that part of NPLs is related to households' mortgages, we also consider the real house price growth rate, which is computed as the year on year percentage change (and deflate it using CPI inflation). Increases in house prices might improve loans quality, by increasing the value of the collateral, making debt repayment easier. Otherwise, in case of rising house prices, a troubled borrower might sell his house and extinguish the loan, without necessarily defaulting. However, excessive increases in house prices also relate to banking crises, as they could end up in asset prices bubbles.

Since we also study the role of uncertainty (and the possible link with the different degrees of risk taking by banks), we consider some indicators of expectations of economic conditions by economic agents coming from the European Commission, and the expectations of financial institutions about lending standards coming from the bank lending survey (European Central Bank). This survey should be particularly useful to our analysis since it provides an indication from banks on whether they intend to loosen or tighten the lending standards. Tightening the lending standards could dampen the economy and accelerate the formation of NPLs in the short-term, or could reduce the risks in the balance sheets of financial institutions (and private sector's indebtedness), slowing down the accumulation of NPLs.

Finally, we also take into account financial market variables. We use the EONIA to control for monetary policyand monetary conditions in general, and its effect on the real value of debt. We use the 10-year government bond yield to consider the influence of sovereign risk on the borrowing cost, and the year on year growth rate of the stock market since drop-in share prices might affect wealth and produce a decline in the value of collateral. Last, but not least, we take into account the VIX as a proxy for a global uncertainty[38]. The link between the VIX and NPLs is not clear ex-ante, and it might depend on the number of lags between the two variables. We expect the coefficient on the VIX to be positive if we observe the link between the two variables contemporaneously (or relatively close), as an increase in global uncertainty could correspond to a poor economic condition. However, many early warning models for banking crises show that historically low levels of VIX have usually preceded periods of distress in the banking system. In

[38] They also call it "the fear Index".

fact, prolonged periods of low uncertainty could create a bias (downward) in the perception of risks by market agents, and this could contribute to the build-up of dangerous imbalances in the balance sheet of banks.

Table 1 in Annex 1 contains a short description for all the variables included in the analysis and information regarding their sources, while Table 2 provides some summary statistics and a pairwise correlation matrix.

[Tables 1 and 2]

3.3.3 Econometric framework

Three separate steps compose our empirical approach. Firstly, we establish the relation between NPLs and the set of macroeconomic and financial variables in our dataset. Secondly, we continue by interacting the main macro-financial determinants of NPLs with the dummy variable that describes the credit cycle, in many different specifications. Finally, we investigate whether uncertainty plays a role, both in general and in the two parts of the credit cycle, by increasing the complexity of the model. To do so, we add another term to the previously mentioned double interaction. We augment the interaction term between the real GDP growth and the credit cycle dummy, first with the VIX, then with the EONIA rate, in a few different model specifications.

For each of the models mentioned above, we carry out the estimation both by adding the lag of the dependent variable, and without it. In both cases, we implement a panel model with country fixed effects[39]. All the explanatory variables are lagged by two periods (two quarters) in order to reduce the simultaneous determination problems (e.g., GDP growth determines NPLs, but increasing NPLs also affect contemporaneous GDP).

We also include year-dummies in our main specifications, given that they are useful to capture regulation changes. They may also capture the different timings in which each country has managed to overcome the financial crisis.

Before showing the mentioned models and before explaining better their rationale and what results we expect, we implement some preliminary work with our variables in order to make sure that we can use them correctly in our econometric models.

3.3.3.1 Preliminary work

Before starting the empirical analysis, we test the stationarity of the variables used in our paper. NPL ratio appears to be non-panel stationary and very persistent. Indeed, when running a panel unit root test[40], it shows that it is non-stationary. Therefore, by taking the log-differences (year-over-year) of the NPL ratio, we make it stationary, and we interpret the results as acceleration (or deceleration) of the ratio.

Despite the panel unit root test shows that the differentiated NPL ratio becomes stationary, it still retains some persistence. In fact, by running a set of regressions

[39] Within estimator.
[40] Fisher, as the panel is unbalanced. We conduct an augmented Dickey-Fuller test on each panel.

with different model specifications, we notice that country-wise the auto-correlogram of the residuals still shows some autocorrelation. Therefore, in our main model specification we include the first lag of the dependent variable in order to avoid issues in the error term. The presence of the lagged dependent variable on the right-hand-side of the equation should already be considered as a robustness test too, given that it captures a lot of the variability in the change of the NPLs percentage change. However, we also run all the regressions by excluding the lagged dependent variable, and (qualitatively) all the results hold.

We run the same panel unit root test for all the variables in our sample, and take the difference over the same quarters of the previous year for all the variables that showed a unit root.[41]

3.3.4 Macro-determinants of NPLs

At this step of our empirical strategy, we estimate the following equation by using standard errors robust to heteroskedastic and auto-correlated disturbances:

$$\Delta NPLs_{it} = \alpha + \beta_1 \Delta NPLs_{it-1} + \beta_2 X_{it} + \delta_t I(year_t) + \gamma_i + \varepsilon_{it} \qquad \textbf{(Equation 1)}$$

where *i* represents the country and *t* the time. X_{it} is a vector of macroeconomic explanatory variables, which are standard in this literature, and γ_i represent the country fixed effects. After estimating the model including the macroeconomic variables, we augment the model with groups of variables, representing some sectors which could be relevant in explaining the evolution of NPLs. We include first the fiscal variables, then the credit variables. Thus, we introduce the institutional variables, followed by the survey variables, and finally the VIX. We introduce each of these groups of variables individually, and we only keep the significant variables at each round.

To control for common changes to the banking systems of the euro area, we also include the year dummies, which in the equation are represented by $I(year_t)$.Finally, as mentioned, the dependent variable shows some persistence, and for this reason, we include its first lag on the right hand side of the equation to control for this persistence and avoid auto-correlated error term.

All the estimations are implemented by using a panel model with fixed effects. We acknowledge that it is not advised to estimate a model that includes the lagged dependent variable amongst the explanatory variables when using the fixed effects, as it might lead to a biased estimation of the coefficient on the lagged dependent variable. We believe that the stationarity of all the variables included in the model, and the fact that the time dimension is larger than the (small) cross-sectional dimension (T ≈ 40 and N=19), keep us safe from such bias.[42]

[41] The variables that we have made stationary by transforming them into a year-over-year differences are: 10 year yield on the sovereign bond, EONIA, government effectiveness, regulatory framework, rule of law, corruption, recovery rate, time insolvency, insolvency framework, enforcing contracts, real effective exchange rate, debt-to-GDP ratio, and the primary balance in percentage of GDP.

[42] Judson and Owen (1999) suggest that when T > 30 fixed effects models (and T > N), the bias on the coefficient on the lagged dependent variable is smaller when we use fixed effects models rather than GMM or other techniques. Fernandez-Val and Weidner (2017), show that a good approximation of the order of the bias is given by p/n, where *p* is the number of parameters to be estimated, and *n* is the sample size. In our case, it would be in the order of second decimal.

From this set of estimates, we mainly expect to confirm the already known standard relations between NPL change and macroeconomic, financial, and institutional variables.

3.3.5 NPL determinants over the credit cycle

In the second step of our empirical strategy, we want to investigate how our set of variables affect NPLs over the two different states of the credit cycle. To do this, we estimate the following equation:

$$\Delta NPLs_{it} = \alpha + \beta_1 \Delta NPLs_{it-1} + \beta_2 X_{it} + \beta_3 I(credit\ cycle_{it}) + \beta_4 \cdot X_{it} \cdot I(credit\ cycle_{it}) + \delta_{year} I(year_t) + \gamma_i + \varepsilon_{it} \qquad \textbf{(Equation 2)}$$

where all the variables are defined as in equation (1) and $I(credit\ cycle_{it})$ is the dummy variable that takes value 1, when the credit gap for a given country in a given year is positive, and zero otherwise.

We estimate the model under its baseline specification (i.e., with year dummies and lagged dependent variable) and under different specifications as described in the previous subsection, with the addition of an interaction term between the credit cycle dummy variable and (one-by-one) a group of determinants of NPLs already examined in the previous sub-section. We estimate these models by using the panel technique with fixed effects.

We expect that some of the determinants affect NPLs differently over the credit cycle. To find whether this claim is supported by the data, we look at the marginal effect of such variables on NPLs. These marginal effects depend on whether the credit cycle dummy is equal to zero or one. When the credit cycle dummy is equal to zero, we have the marginal effect of the determinant under study on the NPLs, during the lower part of the cycle. When the credit cycle dummy is equal to one, we get the marginal effect of the determinant on NPLs in the higher part of the cycle.

If the variable under study affects differently the NPLs over two different stages of the credit cycle, we expect the two numbers to be different.

3.3.6 Investigating the role of uncertainty

In the third step of our empirical exercise, we start from the results of the previous sub-section, i.e., the different reaction of NPLs to their determinants according to where the credit cycle is, and we take it onestep further. Here, we investigate whether we may attribute these results to the different behaviour of the banks at different degrees of economic and financial uncertainty. Then we check whether this holds across the two different stages of the credit cycle.

To do so, we start from the results of the previous two sub-sections, and modify the interaction term by substituting the credit cycle dummy with a variable that proxies for the degree of risk perception by the banks. Then we augment this interaction term (to become triple, i.e., composed of the multiplication between three variables) by including the credit cycle dummy.

At this step, as we have a triple interaction term and the analysis becomes more complex, we narrow down and only look for the marginal effect of GDP growth on NPLs, at different stages of the credit cycle and at different degrees of uncertainty, neglecting the marginal effect of other macro-financial determinants.

The triple interaction term allows us to condition the results on the two stages of the credit cycle, a variable that proxies for the degree of uncertainty, and the real GDP growth, for which we want to compute the marginal effect on NPLs.

In fact, with such two interaction terms, we are able to estimate the marginal effects of real GDP growth on NPLs, corresponding to various degrees of uncertainty, over the two states of the credit cycle.

We expect that the GDP growth has a more positive effect (i.e., accelerate, or slows the negative acceleration of NPLs) when the degree of uncertainty is low, as opposed to the case when it is high. We believe that, when the degree of uncertainty is low, the perception of risk is also low. If this happens in a period when the economy is strong, it is easier for banks to implement a weaker monitoring and to end up allocating inefficiently the financial resources, i.e., to non-creditworthy borrowers. Oppositely, when for some reason banks are more sensitive to risks, then they more carefully allocate the financial resources and more strictly monitor their borrowers.

Within our dataset, we identify two variables that could work as proxy of risk perception in an economy. The first one is the VIX, which allows us to proxy for risk perception by looking at uncertainty. As previously mentioned, Beaker et al (2013) have decomposed the VIX in an uncertainty component and a risk-aversion one. Despite the fact that it is a volatility index of the stock option on the US financial markets, it is often used as a global variable that proxies for economic and financial uncertainty. Gabriele (2019) finds that, in the context of an early warning model, low levels of VIX are often associated to a high probability of a systemic banking crisis.

The second variable is the euro area interbank rate (EONIA). The rationale behind the choice of this variable is that, when interest rates are very low, market participants start searching for yields, which gives incentive to take high risks for a yield considered low as compared to *normal* times. We assume that when the interbank overnight rate is high, banks are very sensitive to risks, as they do not even trust each other.

To understand whether our intuitions are consistent with the data at hand, we estimate the following equations:

$$\Delta NPLs_{it} = \alpha + \beta_1 \Delta NPLs_{it-1} + \beta_2 GDP_{it} + \beta_3 \cdot W_{it} + \beta_4 \cdot GDP_{it} \cdot W_{it} + \beta_5 X_{it} + \delta_{year} I(year_t) + \gamma_i + \varepsilon_{it} \quad \textbf{(Equation 3)}$$

$$\Delta NPLs_{it} = \alpha + \beta_1 \Delta NPLs_{it-1} + \beta_2 GDP_{it} + \beta_3 W_{it}) + \beta_3 I(credit\ cycle_{it}) + \beta_4 \cdot GDP_{it} \cdot W_{it} + \beta_5 \cdot GDP_{it} \cdot I(credit\ cycle_{it}) + \beta_6 \cdot W_{it} \cdot I(credit\ cycle_{it}) + \beta_7 \cdot GDP_{it} \cdot W_{it} \cdot I(credit\ cycle_{it}) + \beta_8 X_{it} + \delta_{year} I(year_t) + \gamma_i + \varepsilon_{it} \quad \textbf{(Equation 4)}$$

Where $X_{it}$is a vector of macroeconomic explanatory variables, which now does not include the real GDP growth, and W_{it} is a vector, which in turn represents the VIX and the EONIA rate. The other variables in the two equations are the same as explained in the previous sub-section

3.4 Empirical results

3.4.1 NPLs determinants

In Table 3 in the Annex 1, we report the estimated coefficients on the determinants of NPLs, and their level of significance, obtained by implementing the fixed effects estimation method on a number of different specifications both including and not including the year dummies.[43] We believe that the year dummies are very important especially because they capture the recovery of the economies after the double deep experienced in many countries of the euro area, which is a significant part of our sample, and the common changes in the banking regulation at European level.[44]

[Table 3]

As shown in most of the seven specifications, and as expected, there is a negative relationship between real GDP growth and NPLs. This means that, given the data available to us, good economic performances in a country should decelerate (or negatively accelerate) the formation of NPLs, as it should make it easier for the borrowers to service their debt.

The inflation rate has consistently a positive and very significant coefficient, which could imply that in our sample, increasing consumer price inflation have a negative effect on disposable income (and therefore on the ability to service the debt), which is stronger than the reduction in the real value of debt it causes.

The effect of the unemployment rate is not significantly different from zero in all the specifications including the lagged dependent variable, which could indicate that real GDP growth already captures all the effect of the economic activity on the NPLs.

The real effective exchange rate takes a positive and significant coefficient in most specifications, suggesting that external activity is important in euro area countries. In fact, an appreciation of the real effective exchange rate, in countries that cannot devaluate their currency, could be followed by a real devaluation to keep the competitiveness, or by a loss of competitiveness, and therefore a loss in exports.

[43] The other results on the research of the determinants of NPLs, when not using the main specification (with lagged dependent variable and year dummies) are in Annex 1a, Tables 10, 11, and 12.

[44] Only when we include the VIX in the list of co-variates, it becomes more difficult to choose whether looking at the result of the specification with or without year dummies. In fact, the VIX takes the same values for all the countries in the sample at every period, partly absorbing the effects absorbed by the year dummies.

The coefficients on the financial variables are all non-significantly different from zero, when we look at column 2. When we exclude the year dummies (Table 10, column 2) we observe that the percentage growth in stock prices of banks is negatively related to the change in NPLs, as it is a variable that gives a signal about the health status of the banks perceived by the markets. The EONIA rate has a positive coefficient and this indicates that tightening monetary policy could make it less sustainable for borrowers to service their debt.

The positive coefficient on the change in the government debt-to-GDP ratio (Table 3, column 5), might suggest that public sector borrowing could crowd out private sector borrowing, hampering the economy and making harder debt servicing.

The VIX shows a positive coefficient in column 6 of Table 3 and Table 10 (for the VIX we also mention the case without year dummies given that, as already explained, the VIX could capture some factors in common with the year dummies), suggesting us that global financial markets volatility relates positively to NPLs growth. This could happen through the generally low value of assets (and collateral) associated to high volatility in financial markets.

Finally, we only find significance in one institutional variable, and only in one specification. Generally, even if not significant, they always have a negative coefficient, signalling that good institutions are negatively related to the growth in NPLs. This could be due to the nature of these variables which are slow moving, and whose effect could be absorbed by country and year fixed effects. We also find no significance in the survey variables. We do not show the institutional variables and the survey variables in the results.

3.4.2 NPL determinants over the credit cycle

Table 4 shows the results of the previously estimated models[45] after adding an interaction term between the credit cycle dummy variable and the most relevant determinants of NPLs. Table 5 shows the marginal effects of the determinants of NPLs on this latter, when we are both in the negative and positive part of the credit cycle.

[Table 4 and 5]

From an overview of Table 4, we can see that after adding the mentioned interaction terms to all the specifications, the significance of the estimated coefficients that are in common with the ones in Table 3 remains broadly the same, and the same is true for the signs of the coefficients. Figure 3 (a trough h), provide a support to better visualise these results.

[Figure 3 (a-h)]

The results in Table 5, show us that some of the variables which we use in our specifications affect NPLs differently over the credit cycle. For instance, an increase in real GDP growth has a negative effect on NPLs when we are in the negative part of the credit cycle, while it has no effect (or an effect with the same sign, but with a smaller magnitude) when we are in the positive part of the credit

[45] We excluded the specification that included the institutional variables, as in the first part of the analysis, out of 13 variables available, none of them was significant in our model.

cycle. Regarding the negative effect of economic growth on NPLs, apart from the improvement in the capacity to service the debt, during the negative part of the credit cycle in a growing economy, we suppose that private sector economic agents might deleverage more easily.

The lower or no effect of real GDP growth on NPLs during the positive part of the credit cycle is harder to explain. One reason could be that a strong economic growth (perhaps overheating) during a credit expansion lowers the perception of risks of the banks, which might end up lending to non-creditworthy borrowers because of bad risk management practices.

The previously mentioned negative effect of increasing inflation on the disposable income during a credit expansion becomes smaller. The reasons might be manifold. High inflation and positive credit gap hint to an economy that is growing in that moment, making the disposable income higher and partly compensating the negative effects of inflation on it. Alternatively, when the private sector is highly indebted, as compared to a long-term trend, it is more likely that the economic agents will give more weight to the fact that increasing inflation can reduce the real value of debt.

An increasing unemployment, during the negative part of the credit cycle increases NPLs, as unemployed borrowers will not be able to service their debt. In the upper part of the credit cycle, an increase in unemployment reduces the growth rate of NPLs (in our main specifications), or it increases it, but by less than in the lower part of the cycle. The effect here is specular to the one of the GDP growth. An increase in the unemployment rate, during a credit expansion, may actually help in curbing the excessive allocation of credit.

When switching from negative to positive part of the credit cycle, the marginal effect of an appreciation of the real effective exchange rate on NPLs growth decreases, from positive to less positive or zero. The main rationale here stems from assuming that such an appreciation would deteriorate the trade balance of an economy. This is more problematic in the lower part of the credit cycle, rather than in the upper part. Issues with the cross-border lending activity would also support this result, but it would have been more appropriate to discuss about them if we had included the nominal exchange rate, rather than the real one.

A faster growth of equity prices determines a decrease in the growth rate of NPLs during the lower part of the credit cycle, whereas it has no effects on it during the upper part. This might be because better-capitalised corporates could leverage more on their capital and borrow more, and better-capitalised banks could concede loans more easily, increasing the probability of future NPLs. An increase in the growth rate of the credit to the private non-financial sector during the lower part of the credit cycle, is positively related to the increase in NPLs, by more than in the upper part of the cycle. Here we would have expected a higher effect of credit growth on NPLs during the upper part of the cycle, because, as documented by the literature (Mendoza and Terrones, 2012) most of the banking crises (which are also defined by excessively high NPLs) happen during the seven years around credit booms. Probably, to observe this type of effect, we should have had a longer lead-lag relation between the credit variables and the NPLs.

An increase in the debt-to-GDP ratio is positively related to the NPLs growth rate, with a stronger marginal effect during credit expansions. We can attribute the higher marginal effect during the upper part of the credit cycle, to the fact that, high public and private sector indebtedness, signal that the entire economy could be experiencing an episode of debt overhang.

Finally, low economic and financial uncertainty could decelerate NPLs during a credit contraction, while this effect disappears, or it becomes weaker, during a credit expansion, given that the low uncertainty determines faster growing lending. This extra-lending may be allocated to non-creditworthiness borrowers, since the low level of uncertainty could be related to low risk aversion by banks.

3.4.3 The role of uncertainty

In Tables 6 and 8, we tabulate the marginal effects of GDP on NPLs at different levels of uncertainty, and for two different types of uncertainty. In Tables 7 and 9, we tabulate the marginal effects of GDP on NPLs at different levels of uncertainty, over the credit cycle, for two different types of uncertainty.

Figures 4 through 7, provide a visual representation of these results, which is easier to interpret than the tables. From Figure 4 (a-d), we can see that GDP growth is negatively related to NPLs growth when the VIX is higher than some level between 15 and 20. This is consistent with the early warning models' literature, whereas a relevant threshold for the VIX to issue a warning of banking crisis is found around 17 (see Gabriele, 2019). This helps us in the interpretation of our result. When VIX is at normal levels, or high, the standard relation between GDP and NPLs applies, while "too low" VIX breaks such relation. This relates to the higher risks taken by banks when there is low uncertainty.

[Figure 4 (a-d) and Table 6]

[Figure 5 (a-d) and Table 7]

[Figure 6 (a-d) and Table 8]

[Figure 7 (a-d) and Table 9]

When we estimate the model that includes the triple interaction, we can observe these dynamics over the credit cycle. The results remain consistent with the ones obtained in the previous section, when the effect of GDP on NPLs was weaker or absent during the upper part of the credit cycle. As expected, during the lower part of the cycle, lower uncertainty weakens the beneficial effect of GDP growth on NPLs. These results hold when we add different types of controls, as we show in Table 6 and 7, and in figures 4 and 5.

Figure 6 (a-d), shows the marginal effect of GDP on NPLs, conditional on the change in the EONIA rate. When the EONIA is increasing or unchanged, the relation between GDP and NPLs is negative. When the EONIA is decreasing, the relation becomes statistically non-significant. Conditioning also on the credit cycle dummy (Figure 7a-d), we can see that during the negative part of the cycle, the relation between GDP and NPLs is negative, apart from when the negative change in the EONIA is large. During the upper part of the cycle, all the negative changes in EONIA are associated with a zero marginal effect of GDP on NPLs.

To follow, we suggest two ways of interpreting this. On the one hand, if we look at the EONIA as a mere interbank interest rate, we can assume that increasing EONIA could be associated with more uncertainty in the banking sector, given that banks apply increasing rates to lend money amongst each other. Low/decreasing EONIA rates, means that the uncertainty is low. On the other hand, the EONIA rate is driven directly by the monetary policy reference rates. This means that tightening the monetary policy stance when the economy is growing supports the slowdown of NPLs. This holds across both the stages of the credit cycle.

One last observation suggested by Figure 7, is that the effect of GDP on NPLs, when the EONIA is growing fast (above 2 percentage points), becomes stronger (with a minus sign) when the credit gap is positive, suggesting that tightening monetary policy during a credit expansion could help in slowing down NPLs, if supported by a growing economy.

3.5 Robustness checks

In order to check the robustness of our results, we implement alternative exercises, to test whether a change in the specification of the model, a correction of the standard errors or a change in the estimation method, can significantly change our results.

Tables 10-15 in Annex 1b, and columns 2-4 of Tables 5-9 of Annex 1 show some variations of the baseline model where in turn we remove the lagged dependent variable, the year dummies and both variables at the same time. The results are qualitatively similar, from the point of view of the signs and of the significance of the main parameters across the different specifications.

[Tables 10-15]

The quarterly frequency of the sample and its ten-years length, which includes a common double crisis, together with the likely dependence amongst the banking systems in the sample, required us further robustness checks.

We have used the Driscoll-Kraay correction of the standard errors in order to account for cross-sectional and time series dependence. The models for the research of the determinants keep broadly the significance. The models where we interact the main determinants with the credit cycle dummy lose some significance, but remain acceptably significant. The models where we investigate the role of uncertainty maintain the level of significance, and in some cases, they outperform the models without correction.[46]

[46] We have replicated the results of Tables 3 and 4, and the results underlying the marginal effects in Tables 5 through 9, using this correction of the standard error, as provided by the STATA command "xtscc". These results are not in the paper and are available upon request.

Finally, in order to check whether the use of the lagged dependent variable in a model estimated with fixed effects might present some issues, we implement a system GMM estimator, and the results hold from a qualitative perspective.

3.6 Conclusions and future research

In this paper, we identify a series of macroeconomic and financial determinants for NPLs, and we have investigated how they relate to a deterioration of the quality of loans over the credit cycle. We show that standard relations between economic and financial variables, and NPLs, change over the credit cycle. Most of the times, the standard relations hold in the lower part of the cycle, while they change in the upper part, becoming weaker, disappearing or even changing sign.

We further investigate whether this could be partly due to the way the banks lend, and how they do it over the credit cycle. By relating economic/financial uncertainty measures to the risk propensity of banks, we observe that when there is low uncertainty (which we relate to more risk taking by financial institutions) the beneficial effects of good economic performances reduce or vanish, whereas, when the level of uncertainty increases, the standard negative relation between good economic performances and NPLs comes back.

Knowing how NPLs relate to the state of the economy can be useful not only to understand their dynamics. It may also be useful to have a better grasp on how, and at what speed, the risk reduction in the European banking system should proceed, in order to complete the banking union.

Moreover, understanding the relation between NPLs and the economic developments in a country can better inform the set-up of scenarios for stress tests of banks/banking systems.

We believe that these results are important in understanding the different dynamics of NPLs across countries, and we credit that having more layers of complexity could help for this purpose. Moreover, these results could also inform those policy makers carrying out stress tests, in order to support them to better choose the macro-financial scenario to apply.

We stop this paper at a crossroad for two possible directions of future research. On the one hand, the same analysis could be repeated by using bank-level data. Apart from the possibility to mix macro-financial determinants with bank-business ones, it would help on the technical side, as the very large cross-sectional dimension of such panel would allow to implement a GMM estimation without experiencing efficiency issues.

On the other hand, this analysis could inform the use of panel VAR technique, in order to understand whether our findings could have a more structural interpretation and perhaps to better help the set-up of stress tests.

3.7 References

Accornero, M., Alessandri, P., Carpinelli, L. and A.M. Sorrentino (2017). "Non-performing loans and the supply of bank credit: evidence from Italy." *Bank of Italy Occasional Paper Series, 374.*

Alessandri, P. and M. Bottero (2017). "Bank lending in uncertain times." *Working Paper. Birkbeck, University of London, London, UK.*

Arcand, J.L., E. Berkes, and U. Panizza. (2015). "Too Much Finance?" *Journal of Economic Growth*, Vol. 20(2*), pp. 105-148.*

Babouček, I. and M. Jančar (2005). "A VAR analysis of the effects to macroeconomic shocks to the quality of the aggregate loan portfolio of the Czech banking sector." *Working paper series*, *No. 1, Czech National Bank.*

Basel Committee on Banking Supervision (2010). "Consultative Document: Countercyclical Capital Buffer Proposal"*, July.*

Beck, T., P. Jakubik and A. Piloiu (2013). "Non-Performing Loans: What Matters in Addition to the Economic Cycle?" *European Central Bank Working Paper 1515*.

Bekaert, Geert & Hoerova, Marie & Lo Duca, Marco, (2013). "Risk, uncertainty and monetary policy," *Journal of Monetary Economics, Elsevier, vol. 60(7), pages 771-788.*

Berger, A.N. and R. DeYoung (1997). "Problem loans and cost efficiency in commercial banks," *Journal of Banking and Finance, Vol. 21(6), 849-870.*

Bloom, Nicholas, (2009). "The Impact of Uncertainty Shocks," *Econometrica, 77(3): 623-685.*

Bofondi, M. and T. Ropele (2011). "Macroeconomic determinants of bad loans: evidence from Italian banks," *Bank of Italy Occasional Papers, 89*.

Frédéric Boissay & Fabrice Collard & Frank Smets, (2016). "Booms and Banking Crises," *Journal of Political Economy, University of Chicago Press, vol. 124(2), pages 489-538.*

Borio, C., Furfine, C. and P. Lowe (2001). "Procyclicality of Financial Systems and Financial Stability," *BIS Papers No.1*.

Borio, C. (2015). "The financial cycle and macroeconomics: What have we learnt?" *Journal of Banking & Finance, Vol. 45, pp. 182-198.*

Boz, E. and G. E. Mendoza (2014). "Financial Innovation, the Discovery of Risk and the U.S. Credit Crisis," *Journal of Monetary Economics*, *Vol. 62, pp. 1-22.*

Caporale, G.M., Di Colli, S. and J.S. Lopez (2014). "Bank lending procyclicality and credit quality during financial crises," *Economic Modelling, Vol. 43, pp. 142–157.*

Cerulli, G., D'Apice, V., Fiordelisi, F. and F. Masala (2017). "NPLs in Europe: The Role of Systematic and Idiosyncratic Factors," *SSRN Working paper*.

Clair, R. T. (1992). "Loan growth and loan quality: some preliminary evidence from Texas banks," *Economic Review, Federal Reserve Bank of Dallas*.

Corsetti, G., Presenti, P. and N. Roubini (1999). "What Caused the Asian Currency & Financial Crisis?" *Japan & the World Economy*, *Vol. 11, pp. 305-373.*

Dell'Ariccia. G. and R. Marquez (2006). "Lending Booms and Lending Standards," *The Journal of Finance, Vol. 51(5), pp. 2511-2546.*

Eichengreen, B. and C. Arteta (2002). "Banking Crises in Emerging Markets: Presumptions and Evidence," *in Mario Blejer and Marko Skreb (eds.) Financial Policies in Emerging Markets. MIT Press.*

Erdiç, D. and E. Abazi (2014). "The Determinants of NPLs in Emerging Europe, 2000-2011," *Journal of Economics and Political Economy*, *Vol. 1(2), pp. 112-125.*

Fernández-Val, Iván & Weidner, Martin, 2016. "Individual and time effects in nonlinear panel models with large N, T." *Journal of Econometrics, Elsevier, vol. 192(1), pages 291-312.*

Gabriele C. (2019). "Learning from trees: A mixed approach to building early warning systems for systemic banking crises." *European Stability Mechanism Working Paper No. 40.*

Gambera M. (2000). "Simple Forecasts of Bank Loan Quality in the Business Cycle," *Emerging Issues Series, Supervision and Regulation Department Federal Reserve Bank of Chicago.*

Gilchrist, S., Jae W. Sim, and Egon Zakrajšek (2014). "Uncertainty, Financial Frictions, and Investment Dynamics," *NBER Working Paper No. 20038.*

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

Hoggarth G., S. Sorensen and L. Zicchino (2005). "Stress Tests of UK Banks Using a VAR Approach," *Bank of England Working Paper No. 282*.

Jimenez, G. and J. Saurina (2006). "Credit cycles, credit risk, and prudential regulation." *International Journal of Central Banking, Vol. 2(2), 65-98.*

Jordá, O., Schularick, M. and A. M. Taylor (2011). "When credit bites back," *Journal of Money, Credit and Banking*, *Vol. 45(2), pp. 3-28.*

Judson, R. A., and A. L. Owen. 1999. "Estimating Dynamic Panel Data Models: A Guide for Macroeconomists." *Economics Letters 65: 9–15.10.1016/S0165-1765(99)00130-5*

Kalirai H. and M. Scheicher (2002). "Macroeconomic Stress Testing: Preliminary Evidence for Austria," *Austrian National Bank Financial Stability Report No. 3, May.*

Keeton, W. R. (1999). "Does Faster Loan Growth Lead to Higher Loan Losses?" *Economic Review, Vol. 84(2), pp. 57-75.*

Keeton, W. and C.S. Morris (1987). "Why do banks' loan losses differ?" *Economic Review*, *Vol. 72(3), pp. 3-21.*

Kindleberger, C. (2000). "Manias, Panics, and Crashes: A History of Financial Crises," *Palgrave Macmillan*, *New York.*

Kwan, S. and R. A. Eisenbeis (1997). "Bank risk, capitalization and operating efficiency," *Journal of Financial Services Research*, *Vol. 12(2/3), pp. 117-131.*

Leon, H., and M. Tracey (2011). "The Impact of Non-Performing Loans on Loan Growth," *IMF Working Paper*.

Louzis, D. P., Vouldis, A. T. and V. L. Metaxas (2012). "Macroeconomic and bank-specific determinants of nonperforming loans in Greece: a comparative study of mortgage, business and consumer loan portfolios," *Journal of Banking & Finance, Vol. 36(4), pp. 1012-1027.*

Mendoza E. G. and M. E. Terrones (2012). "An anatomy of credit booms and their demise," *NBER Working Paper No. 18379*.

Messai, A. S. and F.Jouini (2013). "Micro and Macro Determinants of Non-performing Loans." *International Journal of Economics and Financial Issues Vol. 3(4), pp.852-860.*

Michail, N. A. and C. S. Savva (2018). "Determinants of Non-Performing Loans in Cyprus: A Bayesian VAR approach." *University of Cyprus Economic Analysis Papers*.

Mohaddes, K., M. Raissi and A. Weber (2017). "Can Italy Grow Its Way Out of the NPL Overhang? A Panel Threshold Analysis," *Economic Letters*, *Vol. 159, pp. 195-189.*

Nkusu, M. (2011). "Nonperforming Loans and Macro financial Vulnerabilities in Advanced Economies," *IMF Working Paper No. 161*.

Podpiera, J., and L. Weill (2008). "Bad Luck or Bad Management? Emerging Banking Market Experience," *Journal of Financial Stability*, *Vol. 4(2), pp. 135-148.*

Reinhart, C., M. and K., S. Rogoff (2009). "The Modern History of Exchange Rate Arrangements: A Reinterpretation," *Quarterly Journal of Economics*, *Vol. 119(1), pp. 1-48.*

Rinaldi, L. and A. Sanchis-Arellano (2006). "Household Debt Sustainability: What explains Household Non-performing Loans? An Empirical Analysis." *ECB Working Paper*

Salas, V. and J. Saurina (2002). "Credit Risk in Two Institutional Regimes: Spanish Commercial and Savings Banks," *Journal of Financial Services Research*, *Vol. 22 (3), pp. 203-224.*

Sinkey, J.F. and M.Greenwalt (1991). "Loan-loss experience and risk-taking behaviour at large commercial banks," *Journal of Financial Services Research*, *Vol. 5(1), pp. 43-59.*

Us, V. (2017). "Dynamics of non-performing loans in the Turkish banking sector by an ownership breakdown: The impact of the global crisis," *Finance Research Letters*, *Vol. 20, pp.109-117.*

Valencia, Fabian, (2013). "Aggregate uncertainty and the supply of credit," *International Monetary Fund working paper no.241.*

3.8 Annex 1a: Tables

Table 1: Variables description

Variable	Definition	Source
NPLs growth rate	log. NPLs-to-total loans ratio y/y change	SNL/Fitch connect
Real GDP growth rate	Real Gross Domestic Product y/y % change	Eurostat
Consumer price index	Harmonized CPI: y/y % change	Eurostat
Unemployment rate	Percentage of unemployed people	Eurostat
Reer	Real Effective Exchange Rate	IMF/IFS
Stock market growth rate	DJ Euro STOXX price index: financials	ECB
10y gov. bond yield	Long Term Government Bond Yield: Average	ECB
EONIA	Overnight deposits rate	ECB
Real house price growth	Residential property price % growth	BIS
Credit growth rate	Percentage growth rate year on year of credit to private non-financial sector from all sectors of the economy	ECB
Credit gap dummy	Deviation of credit to private non-financial sector from all sectors, from long-term (HP filter) trend, transformed into binary	ECB with authors'computation
Budget balance to GDP	General governement balance as a % of GDP	ECB
Public debt to gdp	General Government: Consolidated Gross Debt as a % of GDP	Eurostat
Vix	CBOE volatility index	Wall Street Journal
Regulatory quality	Measure the government's ability to formulate and implement sound policies and regulations	World Bank
Rule of law	Measure how consumers have confidence in and abide by the rules of the society	World Bank

Table 2:a) Summary statistics

Variable	N. Obs.	Mean	Std. Dev.	Min.	Max.
NPLs growth rate	729	7.83	34.91	-86.89	176.15
Real GDP growth rate	729	1.07	4.35	-17.50	29.10
Consumer price index	729	1.39	1.62	-3.87	11.90
Unemployment rate	729	10.32	5.18	3.43	27.73
Reer	612	98.27	4.02	87.67	110.13
Stock market growth rate	729	-0.83	27.44	-57.77	42.63
10y gov. bond yield	689	3.41	3.11	-0.12	25.40
EONIA	729	0.00	0.01	0.00	0.04
Real house price growth	710	5.57	3.03	1.34	15.55
Credit growth rate	710	2.28	6.00	-15.39	50.92
Credit gap dummy	729	0.29	0.45	0.00	1.00
Deficit to GDP	729	-3.51	4.02	-32.06	4.18
Public debt to gdp	729	74.06	38.69	4.50	181.00
Vix	729	19.45	9.25	10.31	58.74
Regulatory quality	691	1.23	0.41	0.15	2.05
Rule of law	691	1.24	0.50	0.08	2.10

b) Correlations matrix

	NPL ratio growth	GDP growth	Inflation	Unemployment	REER	Equity prices growth	10 year bond yield	EONIA
NPL ratio growth	1							
GDP growth	-0.66*	1						
Inflation	0.02	0.14*	1					
Unemployment	0.26*	-0.19*	-0.16*	1				
REER	0.43*	-0.37*	-0.28*	-0.29*	1			
Equity prices growth	-0.22*	0.33*	-0.15*	0.02	-0.02	1		
10 year bond yield	0.62*	-0.31*	0.24*	0.40*	-0.04	-0.20*	1	
EONIA	0.38*	0.21*	0.49*	-0.24*	-0.06*	-0.16*	0.38*	1
Real house prices inflation	-0.04	-0.18*	-0.61*	0.20*	0.09*	0.00	-0.18*	-0.62*
Credit growth	-0.13*	0.47*	0.50*	-0.37*	-0.17*	0.03	0.00	0.54*
Credit Ggap	0.38*	-0.10*	0.14*	-0.10*	0.21*	-0.11*	0.31*	0.45*
Government budget balance to GDP	-0.49*	0.33*	0.08*	-0.49*	0.08*	0.06*	-0.41*	0.15*
Government debt to GDP	0.05	-0.29*	-0.36*	0.34*	-0.02	0.05*	0.11*	-0.31*
VIX	0.41*	-0.31*	0.19*	-0.05*	0.06*	-0.64*	0.32*	0.39*
Regulatory framework	-0.14*	0.10*	-0.06*	-0.55*	0.27*	-0.02	-0.35*	0.08*
Rule of law	-0.10*	-0.03	-0.14*	-0.52*	0.36*	0.00	-0.32*	0.01

	Real house prices inflation	Credit growth	Credit gap	Government budget balance to GDP	Government debt to GDP	VIX	Regulatory quality	Rule of law
NPL ratio growth								
GDP growth								
Inflation								
Unemployment								
REER								
Equity prices growth								
10 year bond yield								
EONIA								
Real house prices inflation	1							
Credit growth	-0.40*	1						
Credit Ggap	-0.22*	0.43*	1					
Government budget balance to GDP	-0.15*	0.31*	-0.20*	1				
Government debt to GDP	0.31*	-0.39*	-0.03	-0.39*	1			
VIX	-0.18*	0.05*	0.27*	-0.12*	-0.13*	1		
Regulatory framework	-0.03	0.10*	0.15*	0.38*	-0.39*	0.04	1	
Rule of law	0.06*	-0.02	0.12*	0.24*	-0.17*	0.02	0.84*	1

Table 3: NPLs determinants

Dependent variable: ΔNPL	1	2	3	4	5	6	7
Real GDP growth rate (t-2)	-0.25*	-0.36**	-0.36**	-0.27*	-0.07	-0.23	-0.26*
	(0.13)	(0.16)	(0.14)	(0.14)	(0.15)	(0.14)	(0.13)
Consumer price inflation (t-2)	2.27***	2.04*	2.07***	2.32***	2.35***	2.28***	2.31***
	(0.65)	(1.06)	(0.63)	(0.65)	(0.64)	(0.65)	(0.64)
Unemployment rate (t-2)	-0.06	0.13	0.00	-0.19	-0.10	-0.03	-0.08
	(0.57)	(0.53)	(0.59)	(0.61)	(0.56)	(0.57)	(0.56)
Reer (t-2)	0.31**	0.51**	0.30**	0.31**	0.34**	0.33**	0.29**
	(0.13)	(0.17)	(0.13)	(0.13)	(0.11)	(0.14)	(0.13)
Stock market growth rate (t-2)		-0.01					
		(0.03)					
10y gov. bond yield (t-2)		-0.24					
		(0.24)					
EONIA (t-2)		2.60					
		(1.49)					
Real house price growth (t-2)		0.23					
		(0.35)					
Credit growth rate (t-2)			0.15				
			(0.11)				
Credit gap dummy (t-2)			-2.62*				
			(1.27)				
Budget balance to GDP				-0.26			
				(0.31)			
Public debt to gdp (t-2)					0.22**		
					(0.10)		
Vix (t-2)						0.15*	
						(0.08)	
Regulatory quality (t-2)							-5.79
							(6.51)
Rule of law (t-2)							-0.57
							(6.50)
ΔNPL (t-1)	0.85***	0.85***	0.84***	0.85***	0.84***	0.85***	0.85***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Constant	3.57	6.09	4.44	3.12	3.10	-2.49	3.50
	(2.69)	(4.14)	(2.64)	(2.76)	(2.65)	(2.99)	(2.67)
Observations	596	594	596	596	596	596	596
R-squared	0.89	0.90	0.89	0.89	0.89	0.89	0.89
Number of countries	16	16	16	16	16	16	16
Country FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

Table 4: NPLs determinants conditional to the credit cycle

Dependent Variable: ΔNPL	1	2	3	4	5	6	7	8
Real GDP growth rate (t-2)	-0.61***	-0.26*	-0.26*	-0.25*	-0.38**	-0.34**	-0.03	-0.23
	(0.14)	(0.14)	(0.13)	(0.12)	(0.15)	(0.15)	(0.18)	(0.13)
Consumer price inflation (t-2)	2.05***	2.62***	1.92***	2.23***	1.82*	2.04***	2.35***	2.20***
	(0.59)	(0.58)	(0.59)	(0.62)	(0.93)	(0.60)	(0.62)	(0.57)
Unemployment rate (t-2)	-0.02	-0.10	0.70	0.04	0.19	-0.03	0.06	0.08
	(0.53)	(0.56)	(0.53)	(0.58)	(0.50)	(0.59)	(0.61)	(0.58)
Reer (t-2)	0.32**	0.30**	0.30**	0.30*	0.45**	0.31**	0.32**	0.31**
	(0.14)	(0.12)	(0.13)	(0.17)	(0.16)	(0.13)	(0.11)	(0.15)
Stock market growth rate (t-2)					-0.05			
					(0.03)			
Credit growth rate (t-2)						0.25**		
						(0.10)		
Public debt to gdp (t-2)							0.16*	
							(0.09)	
Vix (t-2)								0.22**
								(0.08)
Credit gap dummy (t-2)	-2.12*	-0.10	-1.10	-1.81	-1.49	-2.25*	-2.45	0.37
	(1.19)	(1.18)	(1.13)	(1.48)	(1.19)	(1.23)	(1.82)	(1.55)
Real GDP growth rate · Credit gap dummy (t-2)	0.54**							
	(0.23)							
Consumer price inflation · Credit gap dummy (t-2)		-1.00						
		(0.62)						
Unemployment rate · Credit gap dummy (t-2)			-1.50***					
			(0.26)					
Reer · Credit gap dummy (t-2)				-0.02				
`				(0.30)				
Stock market growth rate · Credit gap dummy (t-2)					0.10***			
					(0.03)			
Credit growth rate · Credit gap dummy (t-2)						-0.13		
						(0.10)		
Public debt to GDP · Credit gap dummy (t-2)							0.11	
							(0.16)	
Vix · Credit gap dummy (t-2)								-0.11
								(0.09)
ΔNPL (t-1)	0.84***	0.84***	0.85***	0.85***	0.84***	0.84***	0.84***	0.84***
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)
Other controls	NO	NO	NO	NO	YES	NO	NO	NO
Constant	5.96**	4.53*	5.93**	5.01*	8.92**	4.47	4.87*	-2.09
	(2.48)	(2.41)	(2.43)	(2.56)	(3.57)	(2.67)	(2.71)	(3.03)
Observations	596	596	596	596	594	596	596	596
R-squared	0.89	0.89	0.90	0.89	0.90	0.89	0.90	0.89
Number of countries	16	16	16	16	16	16	16	16
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

Table 5: Marginal effect of the determinants on NPLs conditional to the credit cycle

Marginal effect of:	on: When credit gap is:	ΔNPL 1	ΔNPL 2	ΔNPL 3	ΔNPL 4
Real GDP growth rate (t-2)	Negative	-0.614***	-0.799***	-2.253***	-3.390***
		(0.136)	(0.141)	(0.620)	(0.706)
	Positive	-0.0753	-0.269**	-1.161**	-2.287***
		(0.226)	(0.118)	(0.539)	(0.476)
Consumer price inflation (t-2)	Negative	2.618***	2.796***	1.990***	4.089***
		(0.575)	(0.522)	(0.748)	(0.744)
	Positive	1.620***	2.173***	0.180	2.712**
		(0.600)	(0.555)	(1.361)	(1.247)
Unemployment rate (t-2)	Negative	0.702	0.0991	6.124***	5.403***
		(0.534)	(0.571)	(1.376)	(1.603)
	Positive	-0.797*	-1.239***	4.944***	4.848**
		(0.466)	(0.411)	(1.414)	(2.001)
Reer (t-2)	Negative	0.304*	0.119	0.524	0.343
		(0.168)	(0.194)	(0.418)	(0.527)
	Positive	0.286	0.294	0.204	0.234
		(0.229)	(0.193)	(0.363)	(0.464)
Stock market growth rate (t-2)	Negative	-0.0511	-0.0900***	-0.131***	-0.203***
		(0.0332)	(0.0162)	(0.0429)	(0.0601)
	Positive	0.0476	-0.0261	0.0505	-0.104
		(0.0439)	(0.0216)	(0.104)	(0.0860)
Credit growth rate (t-2)	Negative	0.248**	0.462***	0.617	0.899**
		(0.103)	(0.135)	(0.383)	(0.429)
	Positive	0.113	0.338**	0.176	0.720**
		(0.122)	(0.142)	(0.184)	(0.281)
Public debt to gdp (t-2)	Negative	0.160*	-0.0307	0.586*	0.586*
		(0.0859)	(0.0844)	(0.303)	(0.327)
	Positive	0.270*	0.0670	0.785***	0.512**
		(0.141)	(0.172)	(0.262)	(0.249)
Vix (t-2)	Negative	0.217***	0.311***	0.528***	1.029***
		(0.0803)	(0.0822)	(0.121)	(0.186)
	Positive	0.111	0.228***	0.122	0.690***
		(0.0970)	(0.0859)	(0.202)	(0.251)

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Column 2: no year dummies. Column 3: no lagged dependent variables. Column 4: no year dummies and no lagged dependent variables. Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

Table 6: Marginal effect of real GDP growth on NPLs given the level of uncertainty (VIX)

Marginal effect of:	on: When VIX is:	ΔNPL 1	ΔNPL 2	ΔNPL 3	ΔNPL 4
Real GDP growth rate (t-2)	10	-0.0801	-0.0870	-0.346	0.0312
		(0.195)	(0.223)	(0.237)	(0.236)
	15	-0.166	-0.206	-0.389**	-0.0728
		(0.137)	(0.174)	(0.176)	(0.176)
	20	-0.252**	-0.324**	-0.432***	-0.177
		(0.116)	(0.151)	(0.141)	(0.143)
	25	-0.338**	-0.443***	-0.475***	-0.281*
		(0.147)	(0.163)	(0.152)	(0.158)
	30	-0.424**	-0.562***	-0.517**	-0.385*
		(0.210)	(0.205)	(0.201)	(0.210)

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Column 2: no year dummies. Column 3: no lagged dependent variables. Column 4: no year dummies and no lagged dependent variables. Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

Table 7: Marginal effect of real GDP growth on NPLs given the level of uncertainty (VIX) conditional to the credit cycle

Marginal effect of:	When VIX is:	on: and when credit gap is:	ΔNPL 1	ΔNPL 2	ΔNPL 3	ΔNPL 4
Real GDP growth rate (t-2)	10	Negative	-0.320	-0.350	-0.451**	-0.211
			(0.201)	(0.225)	(0.227)	(0.236)
		Positive	-0.00884	-0.0393	-0.521	0.109
			(0.210)	(0.224)	(0.317)	(0.227)
	15	Negative	-0.440***	-0.515***	-0.612***	-0.348*
			(0.158)	(0.198)	(0.206)	(0.186)
		Positive	-0.0783	-0.129	-0.527**	0.0211
			(0.170)	(0.194)	(0.263)	(0.181)
	20	Negative	-0.559***	-0.681***	-0.772***	-0.486***
			(0.136)	(0.180)	(0.194)	(0.154)
		Positive	-0.148	-0.219	-0.533**	-0.0672
			(0.163)	(0.190)	(0.228)	(0.166)
	25	Negative	-0.678***	-0.847***	-0.933***	-0.623***
			(0.146)	(0.176)	(0.193)	(0.152)
		Positive	-0.217	-0.309	-0.540**	-0.156
			(0.191)	(0.213)	(0.224)	(0.187)
	30	Negative	-0.798***	-1.012***	-1.093***	-0.760***
			(0.183)	(0.185)	(0.202)	(0.181)
		Positive	-0.287	-0.398	-0.546**	-0.244
			(0.242)	(0.257)	(0.250)	(0.237)

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Column 2: no year dummies. Column 3: no lagged dependent variables. Column 4: no year dummies and no lagged dependent variables. Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

Table 8: Marginal effect of real GDP growth on NPLs given the level of uncertainty (EONIA)

Marginal effect of:	on: When EONIA change is:	ΔNPL 1	ΔNPL 2	ΔNPL 3	ΔNPL 4
Real GDP growth rate (t-2)	-1.5	0.122	0.113	-0.117	0.219
		(0.241)	(0.238)	(0.245)	(0.241)
	-1	-0.0848	-0.0886	-0.317	0.0315
		(0.204)	(0.205)	(0.229)	(0.213)
	-0.5	-0.292*	-0.290	-0.517**	-0.156
		(0.176)	(0.186)	(0.226)	(0.199)
	0	-0.499***	-0.491***	-0.717***	-0.343*
		(0.161)	(0.185)	(0.237)	(0.200)
	0.5	-0.706***	-0.693***	-0.918***	-0.530**
		(0.163)	(0.202)	(0.259)	(0.216)
	1	-0.913***	-0.894***	-1.118***	-0.718***
		(0.182)	(0.233)	(0.291)	(0.244)
	1.5	-1.120***	-1.096***	-1.318***	-0.905***
		(0.213)	(0.273)	(0.330)	(0.281)
	2	-1.328***	-1.297***	-1.518***	-1.092***
		(0.252)	(0.319)	(0.373)	(0.324)
	2.5	-1.535***	-1.499***	-1.719***	-1.279***
		(0.296)	(0.369)	(0.419)	(0.370)
	3	-1.742***	-1.700***	-1.919***	-1.467***
		(0.342)	(0.422)	(0.468)	(0.419)

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Column 2: no year dummies. Column 3: no lagged dependent variables. Column 4: no year dummies and no lagged dependent variables. Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

Table 9: Marginal effect of real GDP growth on NPLs given the level of uncertainty (EONIA) conditional to the credit cycle

Marginal effect of:	on: When EONIA change is:	and when credit gap is:	ΔNPL 1	ΔNPL 2	ΔNPL 3	ΔNPL 4
Real GDP growth rate (t-2)	-1.5	Negative	-0.337	-0.538*	-0.455	-0.239
			(0.283)	(0.282)	(0.297)	(0.276)
		Positive	0.387	0.180	0.175	0.472*
			(0.249)	(0.219)	(0.299)	(0.263)
	-1	Negative	-0.512**	-0.669***	-0.607**	-0.398*
			(0.236)	(0.243)	(0.258)	(0.234)
		Positive	0.114	-0.0580	-0.104	0.220
			(0.216)	(0.187)	(0.287)	(0.237)
	-0.5	Negative	-0.687***	-0.801***	-0.758***	-0.556***
			(0.194)	(0.210)	(0.223)	(0.202)
		Positive	-0.158	-0.296*	-0.382	-0.0318
			(0.195)	(0.174)	(0.287)	(0.224)
	0	Negative	-0.862***	-0.932***	-0.910***	-0.714***
			(0.162)	(0.185)	(0.195)	(0.184)
		Positive	-0.431**	-0.534***	-0.660**	-0.284
			(0.191)	(0.183)	(0.299)	(0.225)
	0.5	Negative	-1.038***	-1.064***	-1.062***	-0.872***
			(0.146)	(0.171)	(0.178)	(0.185)
		Positive	-0.703***	-0.772***	-0.938***	-0.535**
			(0.204)	(0.213)	(0.322)	(0.239)
	1	Negative	-1.213***	-1.195***	-1.213***	-1.031***
			(0.150)	(0.172)	(0.175)	(0.203)
		Positive	-0.976***	-1.010***	-1.216***	-0.787***
			(0.231)	(0.255)	(0.354)	(0.266)
	1.5	Negative	-1.388***	-1.327***	-1.365***	-1.189***
			(0.174)	(0.186)	(0.187)	(0.236)
		Positive	-1.249***	-1.248***	-1.494***	-1.039***
			(0.268)	(0.306)	(0.392)	(0.302)
	2	Negative	-1.564***	-1.458***	-1.517***	-1.347***
			(0.211)	(0.212)	(0.210)	(0.278)
		Positive	-1.521***	-1.486***	-1.772***	-1.290***
			(0.312)	(0.360)	(0.435)	(0.343)
	2.5	Negative	-1.739***	-1.590***	-1.668***	-1.505***
			(0.256)	(0.246)	(0.242)	(0.325)
		Positive	-1.794***	-1.725***	-2.050***	-1.542***
			(0.360)	(0.418)	(0.482)	(0.388)
	3	Negative	-1.914***	-1.721***	-1.820***	-1.664***
			(0.304)	(0.284)	(0.280)	(0.376)
		Positive	-2.066***	-1.963***	-2.328***	-1.794***
			(0.410)	(0.477)	(0.531)	(0.437)

Standard errors in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Column 2: no year dummies. Column 3: no lagged dependent variables. Column 4: no year dummies and no lagged dependent variables. Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

3.9 Annex 1b: Other tables

Table 10: NPLs determinants (without year dummies)

Dependent Variable: ΔNPL	1	2	3	4	5	6	7
Real GDP growth rate (t-2)	-0.43***	-0.39**	-0.66***	-0.43***	-0.41**	-0.26**	-0.42***
	(0.13)	(0.14)	(0.10)	(0.12)	(0.15)	(0.12)	(0.14)
Consumer price inflation (t-2)	2.55***	1.92***	2.24***	2.63***	2.55***	2.27***	2.57***
	(0.47)	(0.57)	(0.41)	(0.48)	(0.47)	(0.46)	(0.46)
Unemployment rate (t-2)	-0.52	-0.10	-0.42	-0.63	-0.53	-0.32	-0.52
	(0.51)	(0.47)	(0.49)	(0.50)	(0.55)	(0.53)	(0.51)
Reer (t-2)	0.20	0.29*	0.19*	0.20	0.21	0.28**	0.20
	(0.13)	(0.15)	(0.10)	(0.13)	(0.13)	(0.12)	(0.13)
Stock market growth rate (t-2)		-0.06***					
		(0.01)					
10y gov. bond yield (t-2)		-0.25					
		(0.22)					
EONIA (t-2)		1.44*					
		(0.72)					
Real house price growth (t-2)		-0.16					
		(0.25)					
Credit growth rate (t-2)			0.37**				
			(0.14)				
Credit gap dummy (t-2)			-2.93*				
			(1.52)				
Budget balance to GDP (t-2)				-0.23			
				(0.31)			
Public debt to gdp (t-2)					0.02		
					(0.13)		
Vix (t-2)						0.23***	
						(0.07)	
Regulatory quality (t-2)							-3.61
							(6.27)
Rule of law (t-2)							-3.60
							(7.24)
ΔNPL (t-1)	0.87***	0.87***	0.86***	0.86***	0.87***	0.84***	0.87***
	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.04)	(0.04)
Constant	-2.72***	-0.82	-2.06***	-2.73***	-2.78***	-6.71***	-2.82***
	(0.62)	(1.55)	(0.43)	(0.62)	(0.58)	(1.35)	(0.62)
Observations	596	594	596	596	596	596	596
R-squared	0.88	0.89	0.89	0.88	0.88	0.89	0.88
Number of countries	16	16	16	16	16	16	16
Country FE	YES	YES	YES	YES	YES	YES	YES
Year FE	NO	NO	NO	NO	NO	NO	NO

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

Table 11: NPLs determinants (with year dummies and without lagged dependent variable)

Dependent Variable: ΔNPL	1	2	3	4	5	6	7
Real GDP growth rate (t-2)	-1.59***	-1.63***	-1.73***	-1.61***	-0.95*	-1.55***	-1.58***
	(0.41)	(0.45)	(0.49)	(0.41)	(0.50)	(0.41)	(0.43)
Consumer price inflation (t-2)	1.57	0.81	1.05	1.73*	1.79*	1.59	1.69
	(0.99)	(1.60)	(0.88)	(0.93)	(0.89)	(0.98)	(0.99)
Unemployment rate (t-2)	5.21***	4.92***	5.49***	4.68***	4.93***	5.25***	5.10***
	(1.30)	(1.21)	(1.24)	(1.21)	(1.21)	(1.28)	(1.31)
Reer (t-2)	0.45	0.51	0.39	0.45	0.54*	0.49*	0.40
	(0.27)	(0.31)	(0.28)	(0.28)	(0.27)	(0.27)	(0.28)
Stock market growth rate (t-2)		-0.07					
		(0.05)					
10y gov. bond yield (t-2)		1.17*					
		(0.61)					
EONIA (t-2)		3.46					
		(2.26)					
Real house price growth (t-2)		0.06					
		(0.42)					
Credit growth rate (t-2)			0.30*				
			(0.16)				
Credit gap dummy (t-2)			-10.89*				
			(5.87)				
Budget balance to GDP (t-2)				-0.96*			
				(0.52)			
Public debt to gdp (t-2)					0.69**		
					(0.26)		
Vix (t-2)						0.27***	
						(0.09)	
Regulatory quality (t-2)							-19.47*
							(9.67)
Rule of law (t-2)							-3.23
							(20.18)
Constant	23.44**	24.74**	29.52***	22.26**	21.77**	17.49*	23.05**
	(9.75)	(10.20)	(7.32)	(9.30)	(9.20)	(9.75)	(9.92)
Observations	612	610	612	612	612	612	612
R-squared	0.60	0.61	0.62	0.61	0.62	0.61	0.61
Number of countries	16	16	16	16	16	16	16
Country FE	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

Table 12: NPLs determinants (without year dummies and without lagged dependent variable)

Dependent Variable: ΔNPL	1	2	3	4	5	6	7
Real GDP growth rate (t-2)	-2.67***	-2.18***	-3.08***	-2.60***	-2.16***	-1.86***	-2.66***
	(0.52)	(0.43)	(0.59)	(0.51)	(0.62)	(0.35)	(0.53)
Consumer price inflation (t-2)	3.64***	2.10*	2.78***	4.00***	3.63***	2.70***	3.67***
	(0.70)	(1.02)	(0.76)	(0.77)	(0.71)	(0.73)	(0.71)
Unemployment rate (t-2)	5.23***	5.43***	5.14***	4.22***	4.77***	5.26***	5.07***
	(1.43)	(1.13)	(1.27)	(1.30)	(1.45)	(1.14)	(1.44)
Reer (t-2)	0.28	0.16	0.27	0.24	0.43	0.57*	0.27
	(0.42)	(0.48)	(0.36)	(0.40)	(0.39)	(0.32)	(0.43)
Stock market growth rate (t-2)		-0.16**					
		(0.06)					
10y gov. bond yield (t-2)		0.53					
		(0.43)					
EONIA (t-2)		-0.50					
		(1.73)					
Real house price growth (t-2)		-0.92					
		(0.59)					
Credit growth rate (t-2)			0.77***				
			(0.23)				
Credit gap dummy (t-2)			-3.66				
			(5.40)				
Budget balance to GDP (t-2)				-1.52***			
				(0.44)			
Public debt to gdp (t-2)					0.54**		
					(0.25)		
Vix (t-2)						0.77***	
						(0.24)	
Regulatory quality (t-2)							-22.24*
							(11.51)
Rule of law (t-2)							-0.40
							(22.10)
Constant	4.07**	8.65**	4.59**	3.76**	2.11	-10.42**	3.78**
	(1.43)	(3.58)	(2.11)	(1.42)	(1.72)	(4.52)	(1.44)
Observations	612	610	612	612	612	612	612
R-squared	0.53	0.56	0.55	0.55	0.54	0.58	0.54
Number of countries	16	16	16	16	16	16	16
Country FE	YES	YES	YES	YES	YES	YES	YES
Year FE	NO	NO	NO	NO	NO	NO	NO

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

Table 13: NPLs determinants conditional to the credit cycle (without year dummies)

Dependent Variable: ΔNPL	1	2	3	4	5	6	7	8
Real GDP growth rate (t-2)	-0.80***	-0.45***	-0.45***	-0.42***	-0.43***	-0.65***	-0.40**	-0.27**
	(0.14)	(0.12)	(0.10)	(0.13)	(0.13)	(0.10)	(0.18)	(0.12)
Consumer price inflation (t-2)	2.45***	2.80***	2.35***	2.57***	1.85***	2.22***	2.60***	2.24***
	(0.44)	(0.52)	(0.44)	(0.48)	(0.58)	(0.39)	(0.48)	(0.41)
Unemployment rate (t-2)	-0.55	-0.56	0.10	-0.44	-0.12	-0.46	-0.44	-0.13
	(0.48)	(0.52)	(0.57)	(0.53)	(0.48)	(0.49)	(0.58)	(0.58)
Reer (t-2)	0.20*	0.19	0.19*	0.12	0.30**	0.20*	0.19	0.25**
	(0.11)	(0.11)	(0.10)	(0.19)	(0.13)	(0.10)	(0.12)	(0.10)
Stock market growth rate (t-2)					-0.09***			
					(0.02)			
Credit growth rate (t-2)						0.46***		
						(0.14)		
Public debt to gdp (t-2)							-0.03	
							(0.08)	
Vix (t-2)								0.31***
								(0.08)
Credit gap dummy (t-2)	-1.08	0.23	-0.01	-0.64	-0.32	-2.56	-1.21	-0.91
	(1.34)	(1.56)	(1.24)	(1.55)	(1.44)	(1.47)	(1.70)	(1.49)
Real GDP growth rate · Credit gap dummy (t-2)	0.53**							
	(0.19)							
Consumer price inflation · Credit gap dummy (t-2)		-0.62						
		(0.66)						
Unemployment rate · Credit gap dummy (t-2)			-1.34***					
			(0.42)					
Reer · Credit gap dummy (t-2)				0.17				
`				(0.33)				
Stock market growth rate · Credit gap dummy (t-2)					0.06**			
					(0.03)			
Credit growth rate · Credit gap dummy (t-2)						-0.12		
						(0.08)		
Public debt to GDP · Credit gap dummy (t-2)							0.10	
							(0.15)	
Vix · Credit gap dummy (t-2)								-0.08
								(0.08)
ΔNPL (t-1)	0.87***	0.87***	0.87***	0.87***	0.87***	0.86***	0.87***	0.84***
	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.04)
Other controls	NO	NO	NO	NO	YES	NO	NO	NO
Constant	-1.72**	-2.78***	-2.05***	-2.53***	-0.24	-2.18***	-2.53***	-7.30***
	(0.61)	(0.52)	(0.52)	(0.61)	(1.10)	(0.42)	(0.62)	(1.41)
Observations	596	596	596	596	594	596	596	596
R-squared	0.89	0.88	0.89	0.88	0.89	0.89	0.88	0.89
Number of countries	16	16	16	16	16	16	16	16
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	NO	NO	NO	NO	NO	NO	NO	NO

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

Table 14: NPLs determinants conditional to the credit cycle (without lagged dependent variable)

Dependent Variable: ΔNPL	1	2	3	4	5	6	7	8
Real GDP growth rate (t-2)	-2.25***	-1.55***	-1.53***	-1.55***	-1.62***	-1.68***	-0.82	-1.50***
	(0.62)	(0.45)	(0.43)	(0.41)	(0.45)	(0.49)	(0.54)	(0.43)
Consumer price inflation (t-2)	1.07	1.99**	1.20	1.37	0.33	0.89	1.69**	1.31*
	(0.69)	(0.75)	(0.73)	(0.81)	(1.24)	(0.84)	(0.74)	(0.71)
Unemployment rate (t-2)	5.41***	5.32***	6.12***	5.53***	5.17***	5.37***	5.39***	5.60***
	(1.21)	(1.29)	(1.38)	(1.26)	(1.19)	(1.24)	(1.25)	(1.23)
Reer (t-2)	0.42	0.39	0.37	0.52	0.37	0.40	0.45	0.38
	(0.32)	(0.26)	(0.31)	(0.42)	(0.32)	(0.27)	(0.29)	(0.31)
Stock market growth rate (t-2)					-0.13***			
					(0.04)			
Credit growth rate (t-2)						0.62		
						(0.38)		
Public debt to gdp (t-2)							0.59*	
							(0.30)	
Vix (t-2)								0.53***
								(0.12)
Credit gap dummy (t-2)	-9.94*	-5.98	-8.75	-9.48	-8.15	-9.49*	-10.37	-0.80
	(5.55)	(6.29)	(5.82)	(6.12)	(5.43)	(5.25)	(6.56)	(6.27)
Real GDP growth rate · Credit gap dummy (t-2)	1.09**							
	(0.49)							
Consumer price inflation · Credit gap dummy (t-2)		-1.81						
		(1.58)						
Unemployment rate · Credit gap dummy (t-2)			-1.18					
			(1.27)					
Reer · Credit gap dummy (t-2)				-0.32				
`				(0.53)				
Stock market growth rate · Credit gap dummy (t-2)					0.18*			
					(0.10)			
Credit growth rate · Credit gap dummy (t-2)						-0.44		
						(0.44)		
Public debt to GDP · Credit gap dummy (t-2)							0.20	
							(0.19)	
Vix · Credit gap dummy (t-2)								-0.41
								(0.27)
Other controls	NO	NO	NO	NO	YES	NO	NO	NO
Constant	32.89***	30.54***	31.37***	31.06***	33.25***	29.07***	29.28***	20.06**
	(6.66)	(6.84)	(6.69)	(6.60)	(6.61)	(7.30)	(6.39)	(7.96)
Observations	612	612	612	612	610	612	612	612
R-squared	0.62	0.62	0.61	0.61	0.63	0.62	0.63	0.62
Number of countries	16	16	16	16	16	16	16	16
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

Table 15: NPLs determinants conditional to the credit cycle (without lagged dependent variable and without year dummies)

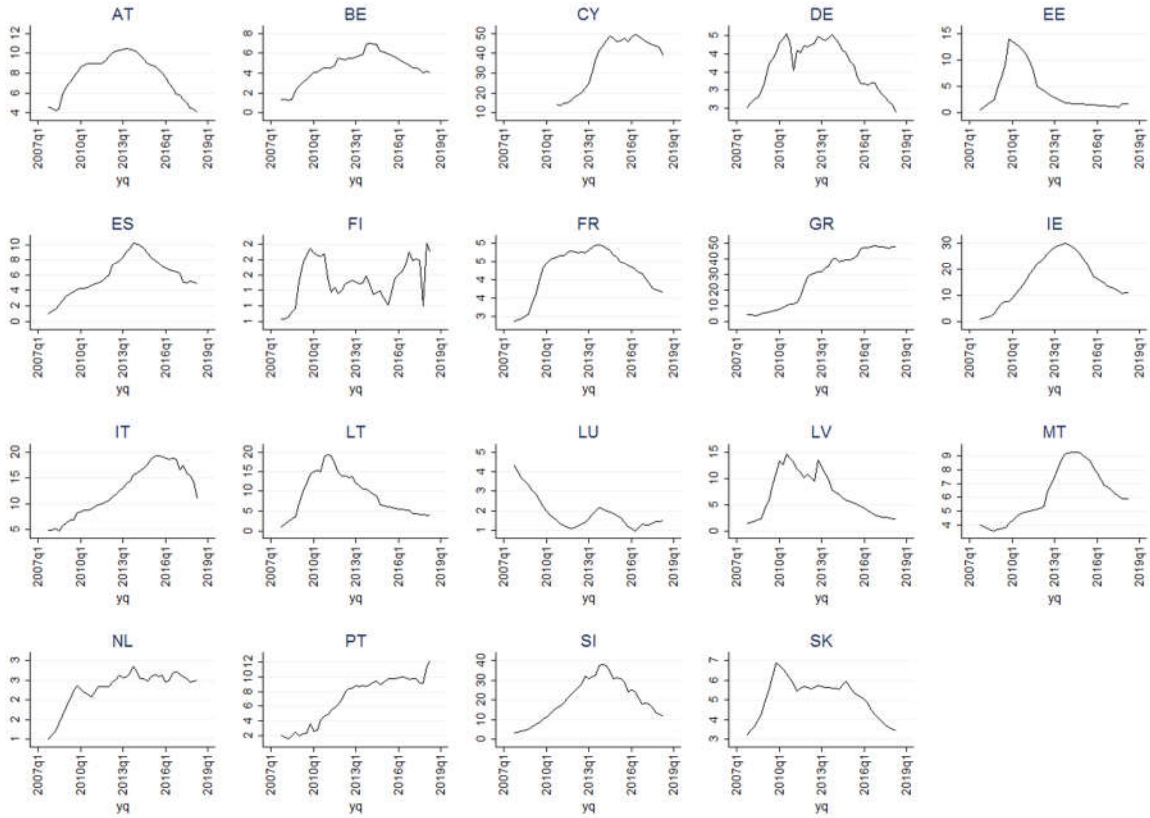
Dependent Variable: ΔNPL	1	2	3	4	5	6	7	8
Real GDP growth rate (t-2)	-3.39***	-2.68***	-2.66***	-2.67***	-2.24***	-3.07***	-2.17***	-1.86***
	(0.71)	(0.51)	(0.50)	(0.48)	(0.41)	(0.59)	(0.62)	(0.36)
Consumer price inflation (t-2)	3.40***	4.09***	3.54***	3.61***	1.99*	2.72***	3.60***	2.59***
	(0.64)	(0.74)	(0.70)	(0.70)	(0.99)	(0.81)	(0.68)	(0.65)
Unemployment rate (t-2)	4.93***	4.92***	5.40***	5.14***	5.40***	5.09***	4.73***	5.49***
	(1.48)	(1.58)	(1.60)	(1.51)	(1.21)	(1.28)	(1.55)	(1.21)
Reer (t-2)	0.34	0.31	0.30	0.34	0.17	0.28	0.43	0.54
	(0.40)	(0.38)	(0.40)	(0.53)	(0.44)	(0.37)	(0.39)	(0.35)
Stock market growth rate (t-2)					-0.20***			
					(0.06)			
Credit growth rate (t-2)						0.90*		
						(0.43)		
Public debt to gdp (t-2)							0.59*	
							(0.33)	
Vix (t-2)								1.03***
								(0.19)
Credit gap dummy (t-2)	0.35	3.44	1.36	0.99	-0.80	-3.07	0.35	2.29
	(6.12)	(5.91)	(6.03)	(6.40)	(5.48)	(4.81)	(6.57)	(5.81)
Real GDP growth rate · Credit gap dummy (t-2)	1.10*							
	(0.56)							
Consumer price inflation · Credit gap dummy (t-2)		-1.38						
		(1.54)						
Unemployment rate · Credit gap dummy (t-2)			-0.56					
			(1.89)					
Reer · Credit gap dummy (t-2)				-0.11				
`				(0.62)				
Stock market growth rate · Credit gap dummy (t-2)					0.10			
					(0.08)			
Credit growth rate · Credit gap dummy (t-2)						-0.18		
						(0.53)		
Public debt to GDP · Credit gap dummy (t-2)							-0.07	
							(0.23)	
Vix · Credit gap dummy (t-2)								-0.34
								(0.24)
Other controls	NO	NO	NO	NO	YES	NO	NO	NO
Constant	5.23**	3.04	3.90	3.76	9.65***	4.43**	2.08	-13.24***
	(2.44)	(2.20)	(2.59)	(2.49)	(2.40)	(1.91)	(2.50)	(4.44)
Observations	612	612	612	612	610	612	612	612
R-squared	0.54	0.53	0.53	0.53	0.56	0.55	0.54	0.59
Number of countries	16	16	16	16	16	16	16	16
Country FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Unemployment rate, real effective exchange rate, 10 year sovereign yield, EONIA, budget balance to GDP, debt to GDP, regulatory quality and rule of law are all take in year-on-year difference as they were non stationary.

3.10 Annex 2: Figures

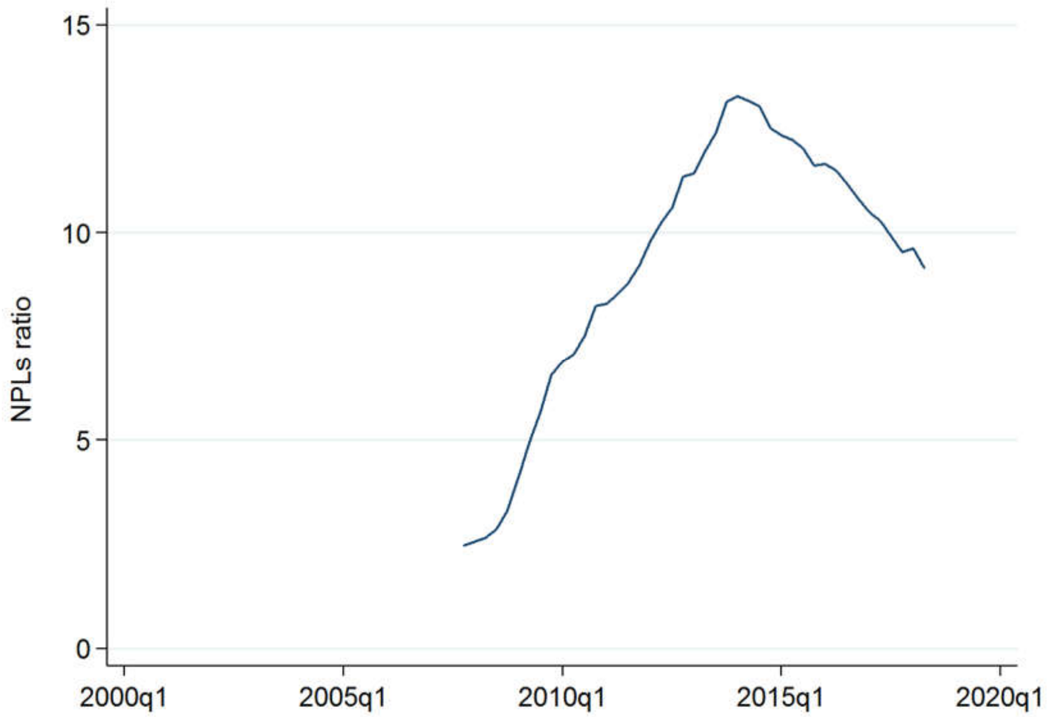
Figure 1: NPL-to-total gross loans ratios' evolution across countries (unit: %)



Source: author's computations, SNL Financials and national central banks.

Looking at NPL evolution at country level, we can notice a high degree of heterogeneity across countries. Although we can observe that the trend is homogeneously decreasing for most of the countries in the Euro area, NPLs in Greece and Portugal show an opposite pattern with peaks respectively of 47.41% and 12.25%.

Figure 2: NPL-to-total gross loans ratio in the Euro area (unit: %)



Source: author’s computation

The NPL ratio in the euro area has been increasing since the beginning of our sample at the end of 2007. It reached its peak in 2014 and since then it has been on a declining trend.

Figure 3: Marginal effect of the determinants on NPLs conditional to the credit cycle (y-axis: percentage change in the NPL ratio; x-axis credit gap in percentage points)

Figure 3a: Marginal effect of real GDP growth on NPLs (95% confidence interval)

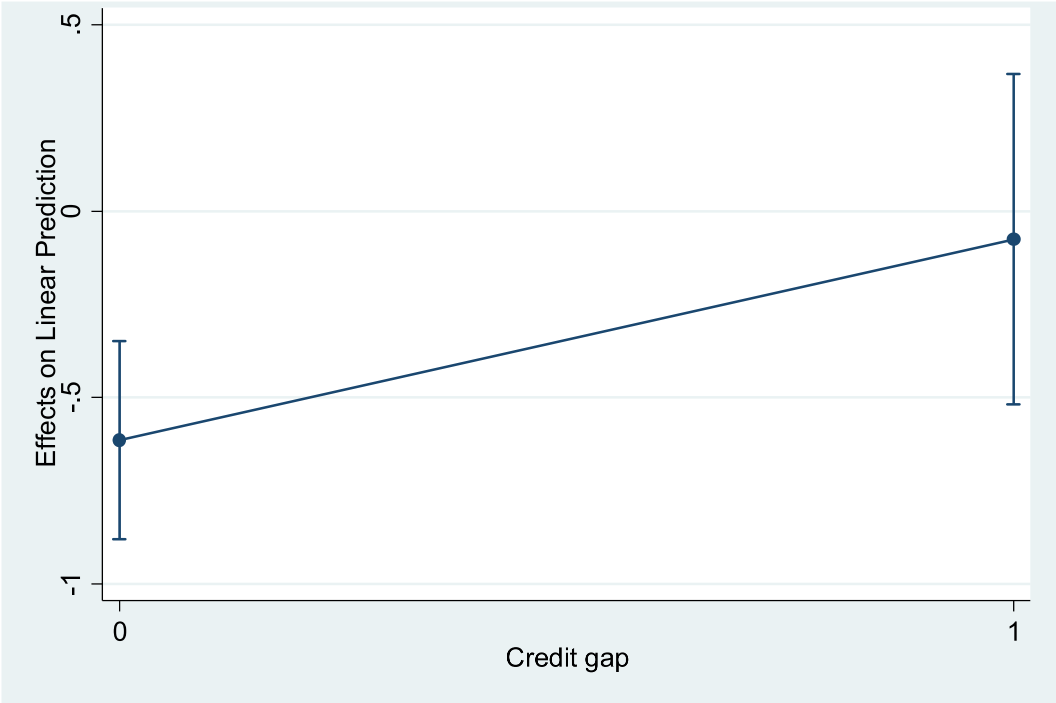


Figure 3b: Marginal effect of CPI inflation on NPLs (95% confidence interval)

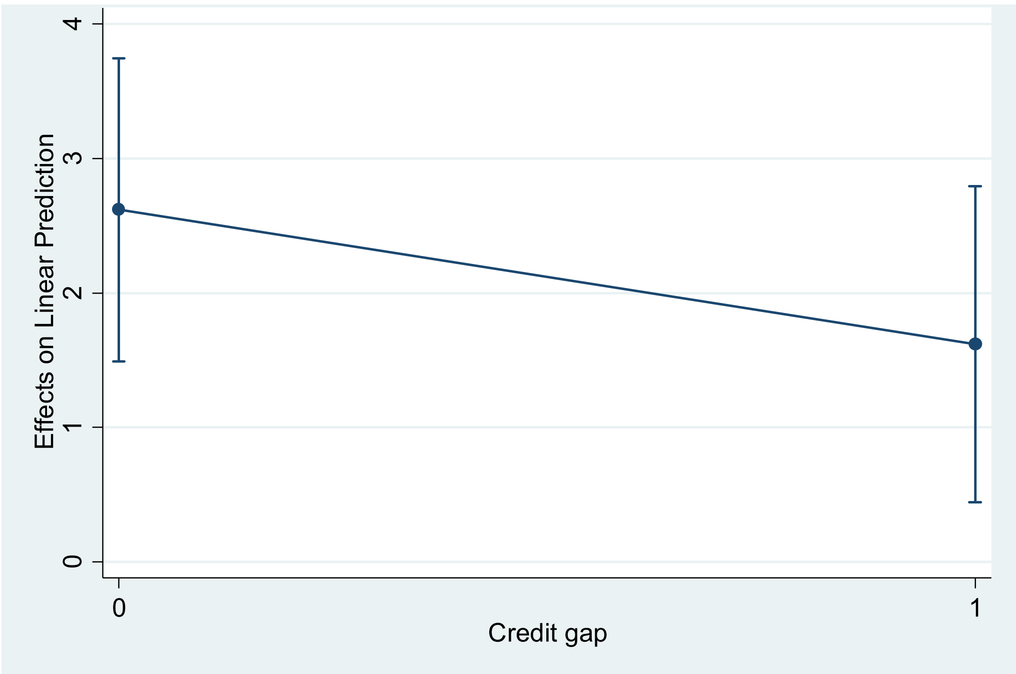


Figure 3c: Marginal effect of unemployment on NPLs (95% confidence interval)

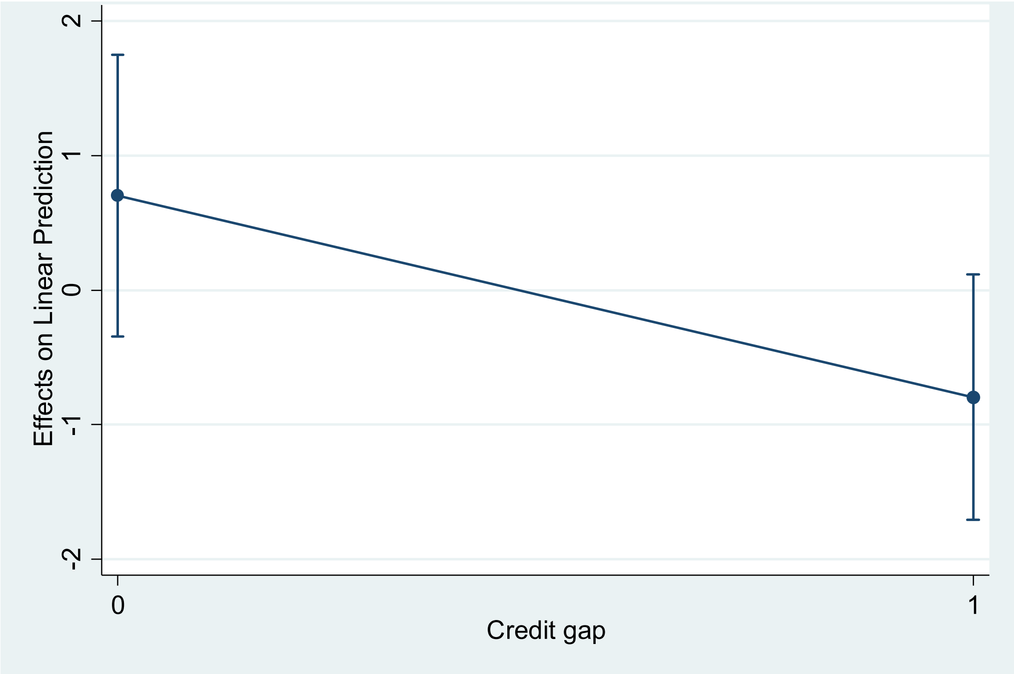


Figure 3d: Marginal effect of REER on NPLs (95% confidence interval)

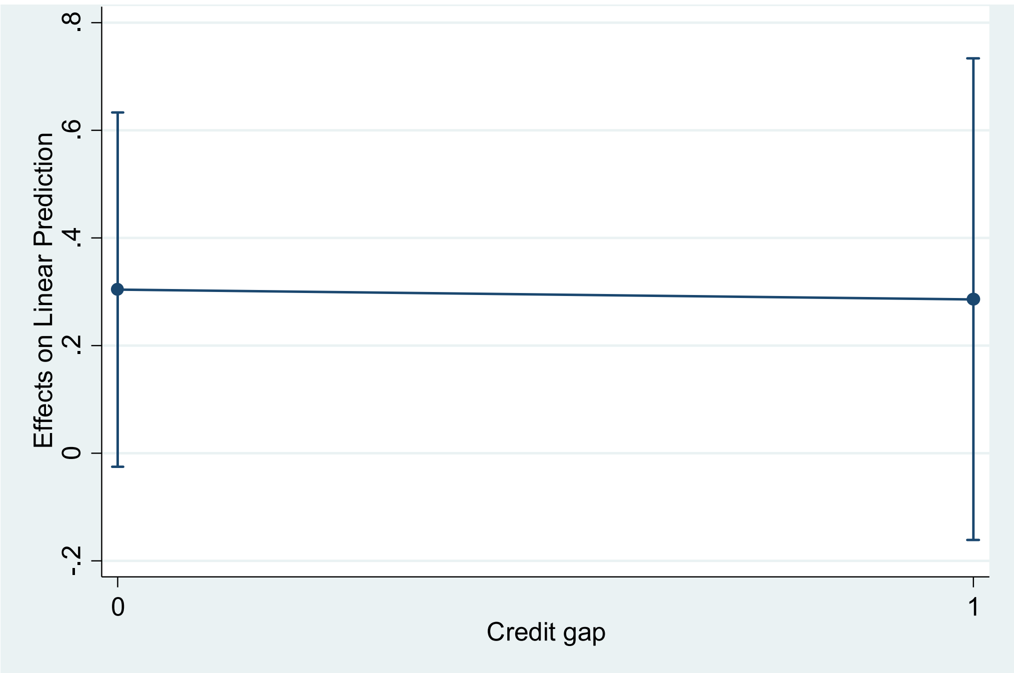


Figure 3e: Marginal effect of stock prices growth on NPLs (95% confidence interval)

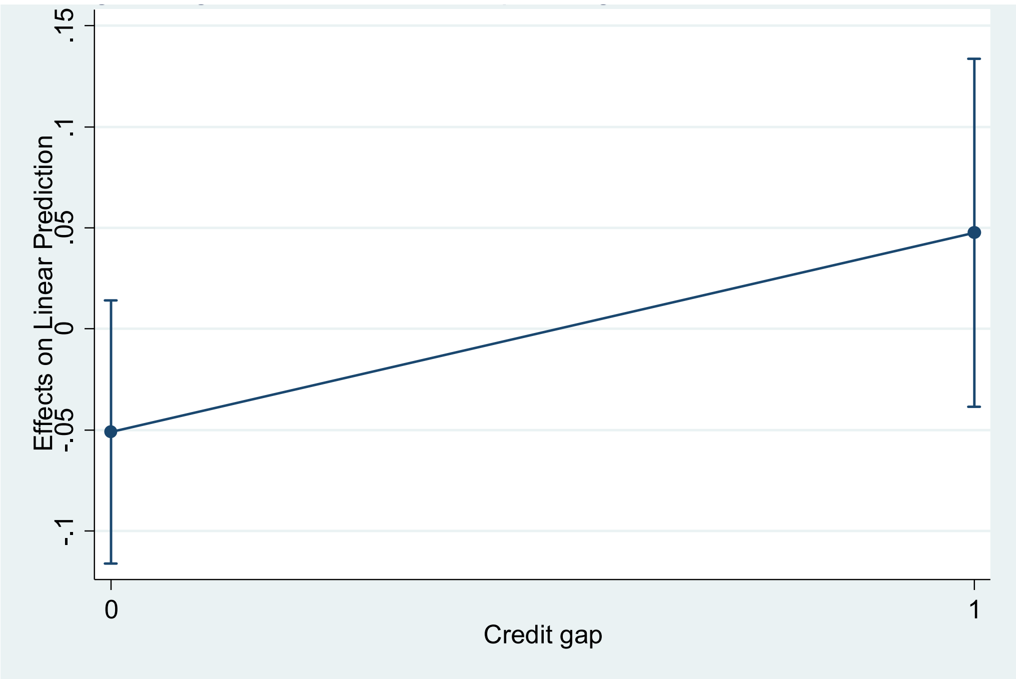


Figure 3f: Marginal effect of credit growth on NPLs (95% confidence interval)

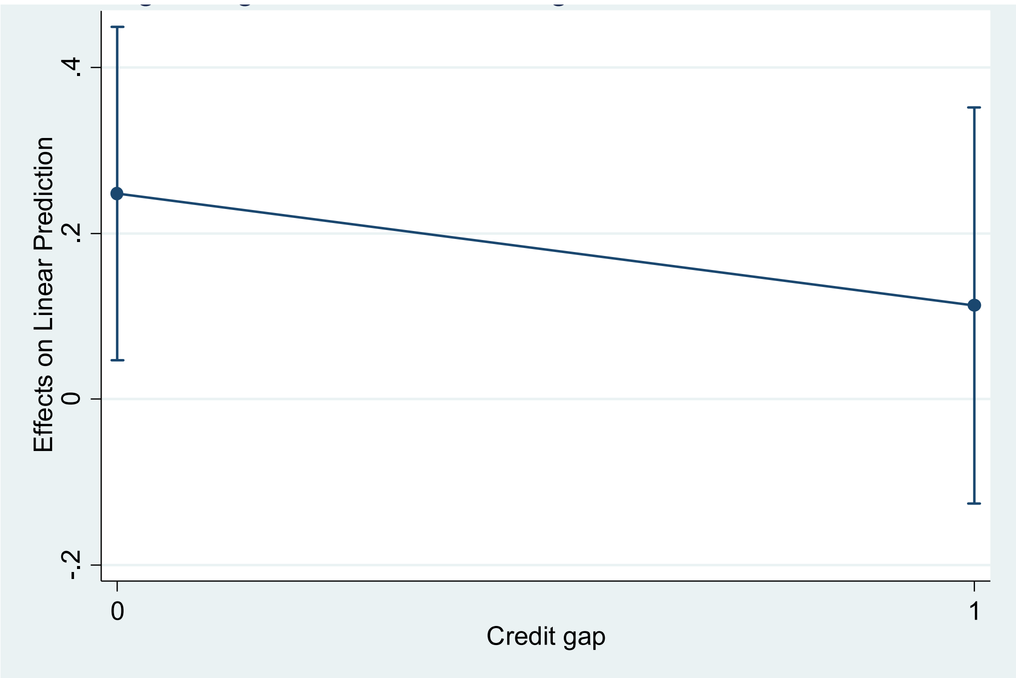


Figure 3g: Marginal effect of government debt-to-GDP on NPLs (95% confidence interval)

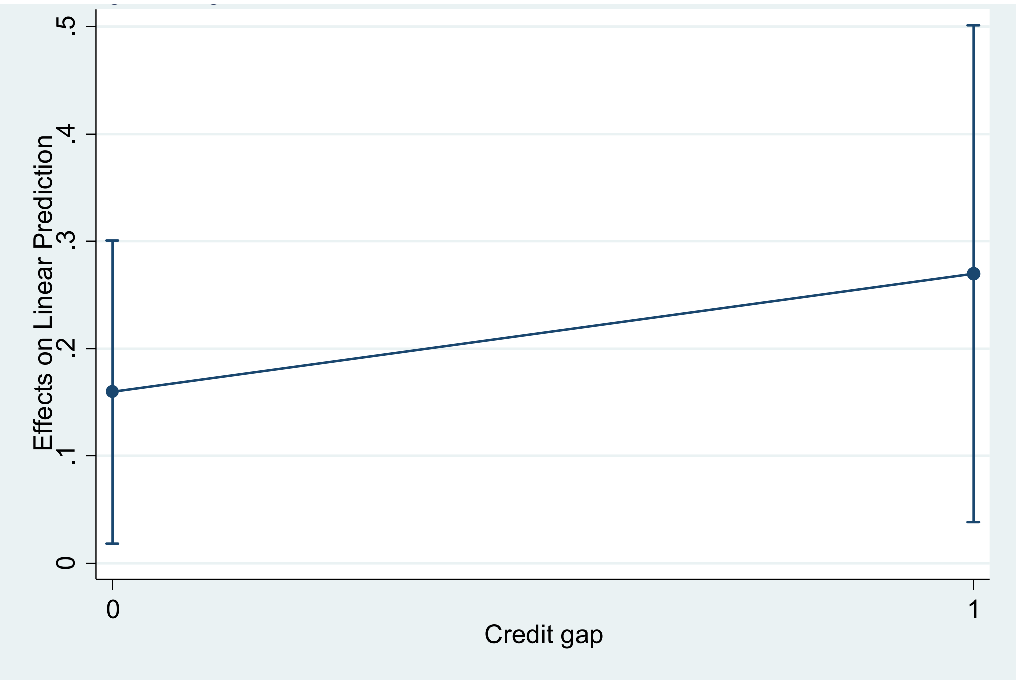


Figure 3h: Marginal effect of VIX on NPLs (95% confidence interval)

Figure 4: Marginal effect of real GDP growth on NPLs (95% confidence interval) given the level of uncertainty (VIX) (y-axis: percentage change in the NPL ratio; x-axis VIX in %)

Figure 4a: Specification 1 (Table 6 – Annex 1)

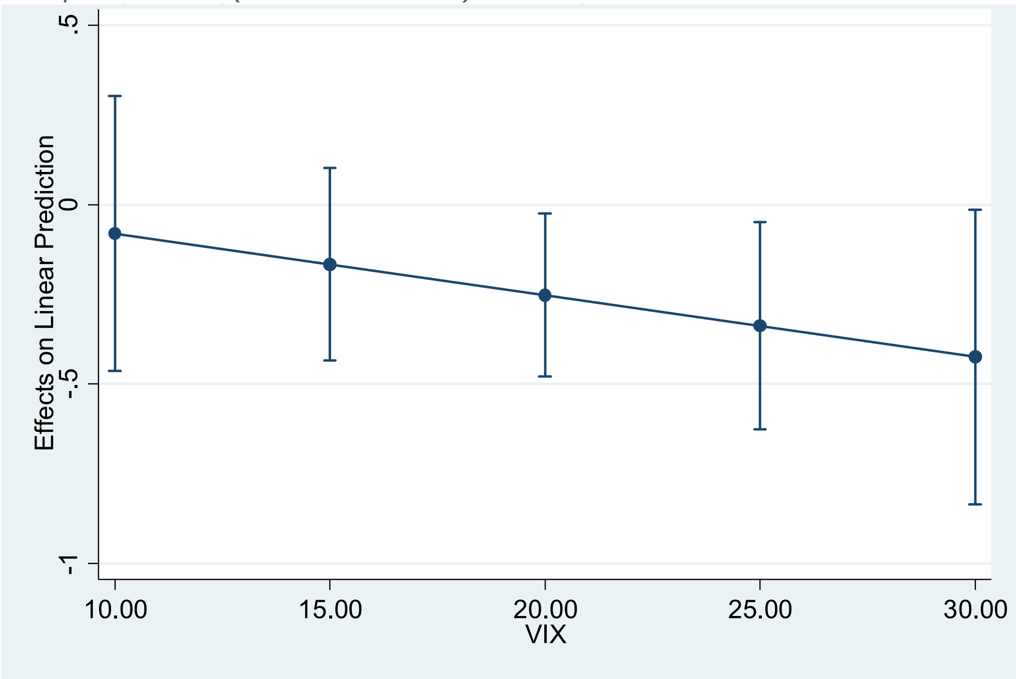


Figure 4b: Specification 2 (Table 6 – Annex 1)

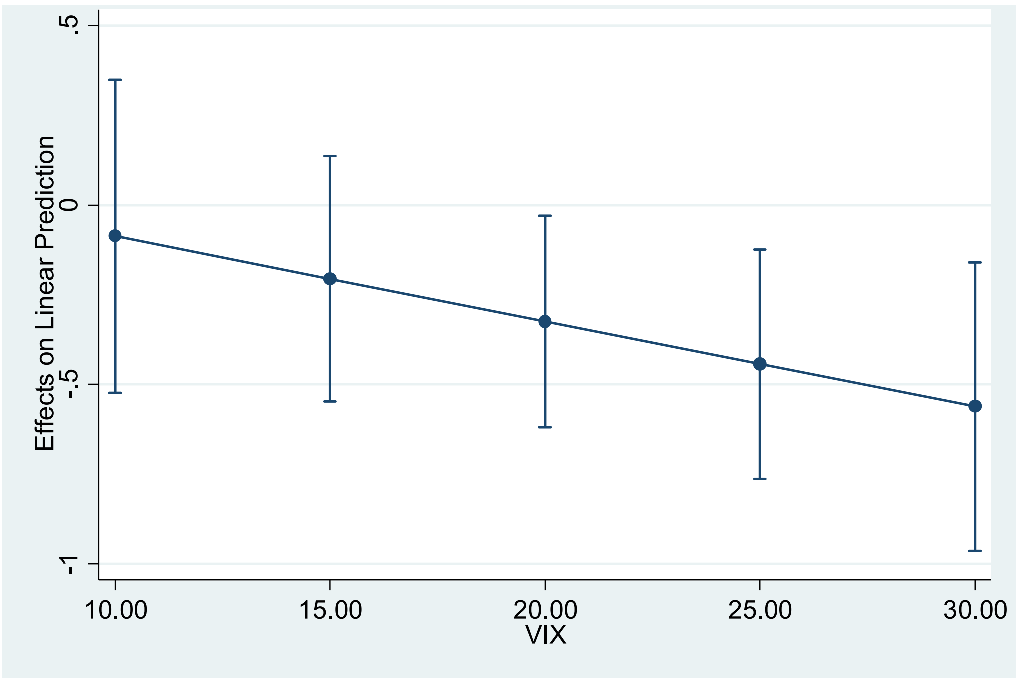


Figure 4c: Specification 3 (Table 6 – Annex 1)

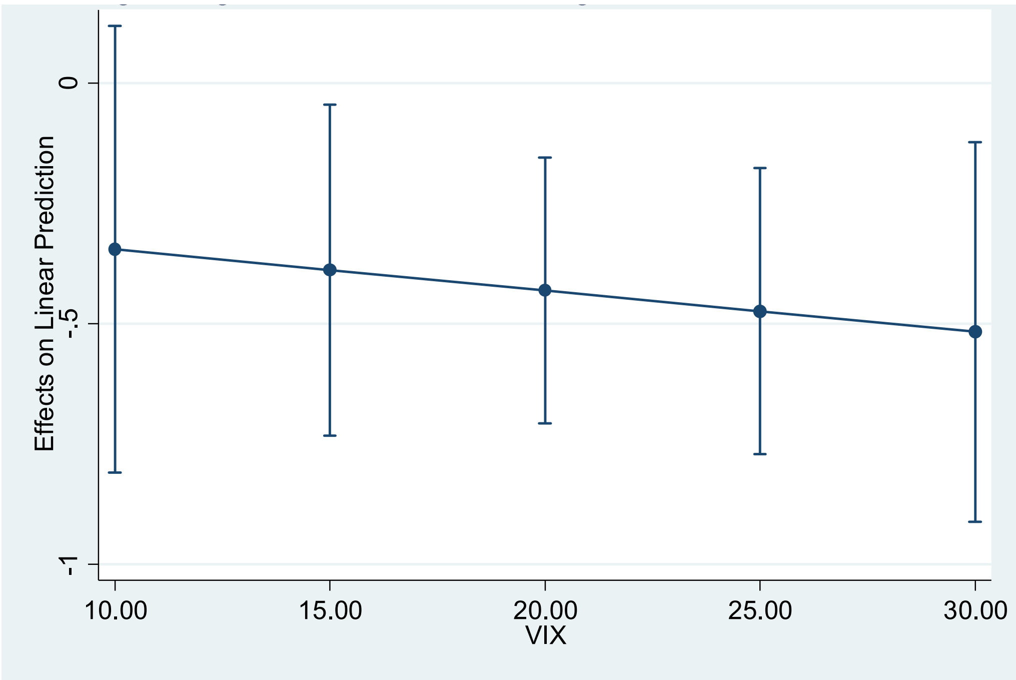


Figure 4d: Specification 4 (Table 6 – Annex 1)

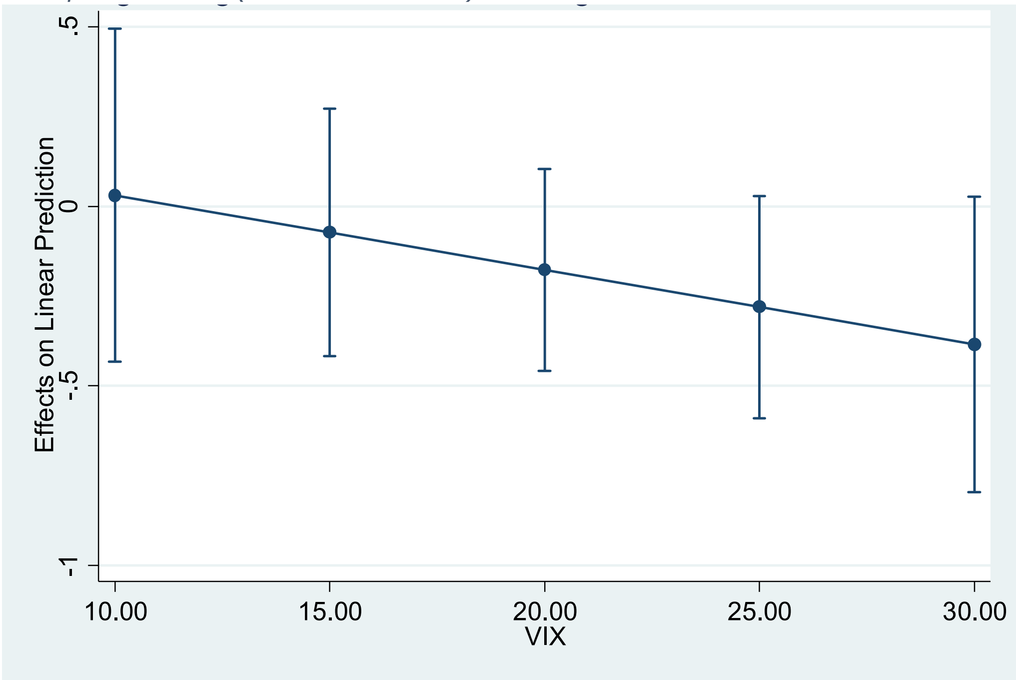


Figure 5: Marginal effect of real GDP growth on NPLs (95% confidence interval) given the level of uncertainty (VIX) conditional to the credit cycle (y-axis: percentage change in the NPL ratio; x-axis VIX in %)

Figure 5a: Specification 1 (Table 7 – Annex 1)

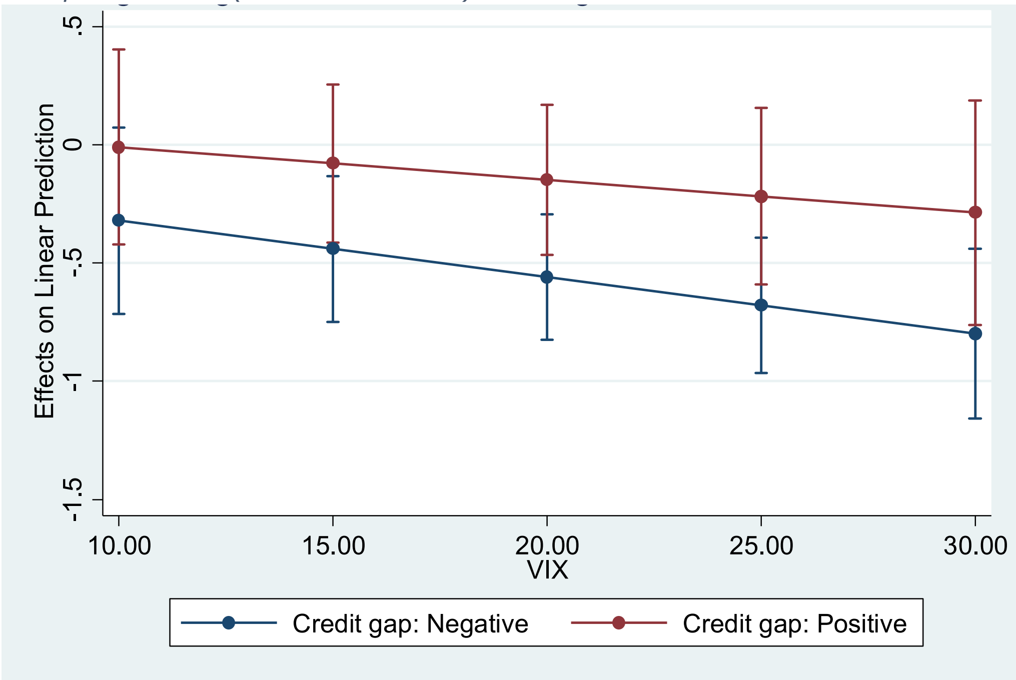


Figure 5b: Specification 2 (Table 7 – Annex 1)

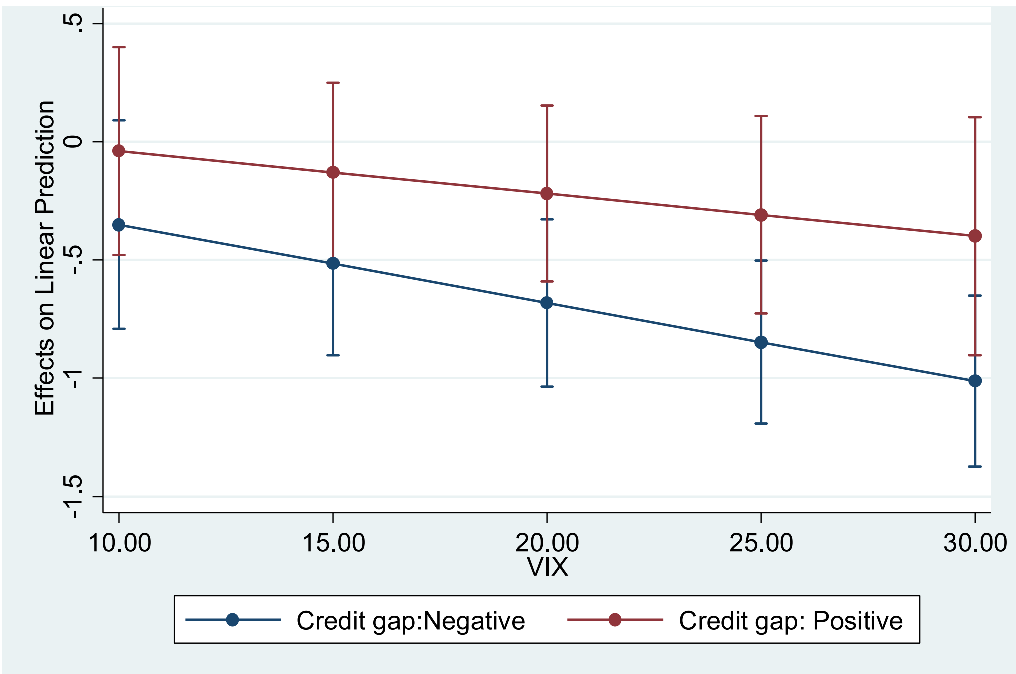


Figure 5c: Specification 3 (Table 7 – Annex 1)

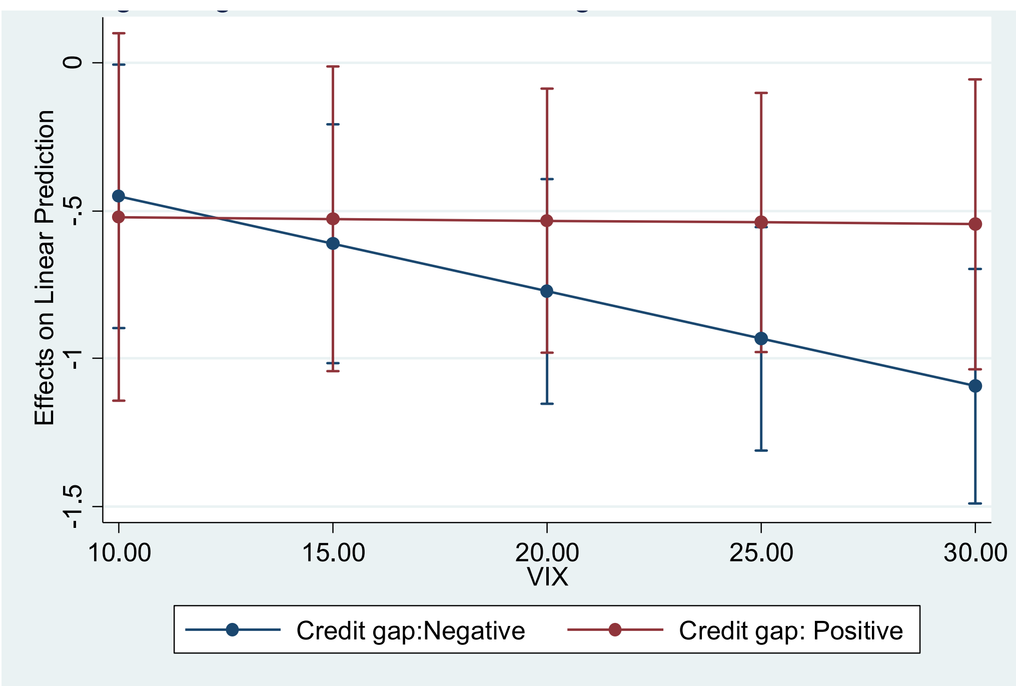


Figure 5d: Specification 4 (Table 7 – Annex 1)

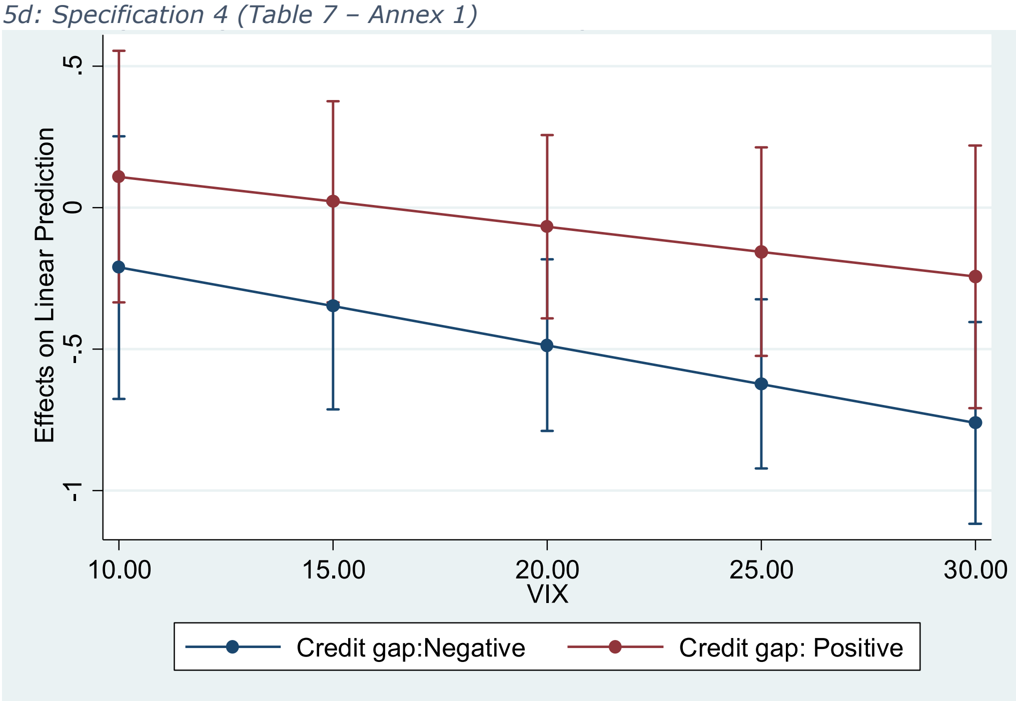


Figure 6: Marginal effect of real GDP growth on NPLs (95% confidence interval) given the level of uncertainty (EONIA) (y-axis: percentage change in the NPL ratio; x-axis change in EONIA rate in percentage points)

Figure 6a: Specification 1 (Table 8 – Annex 1)

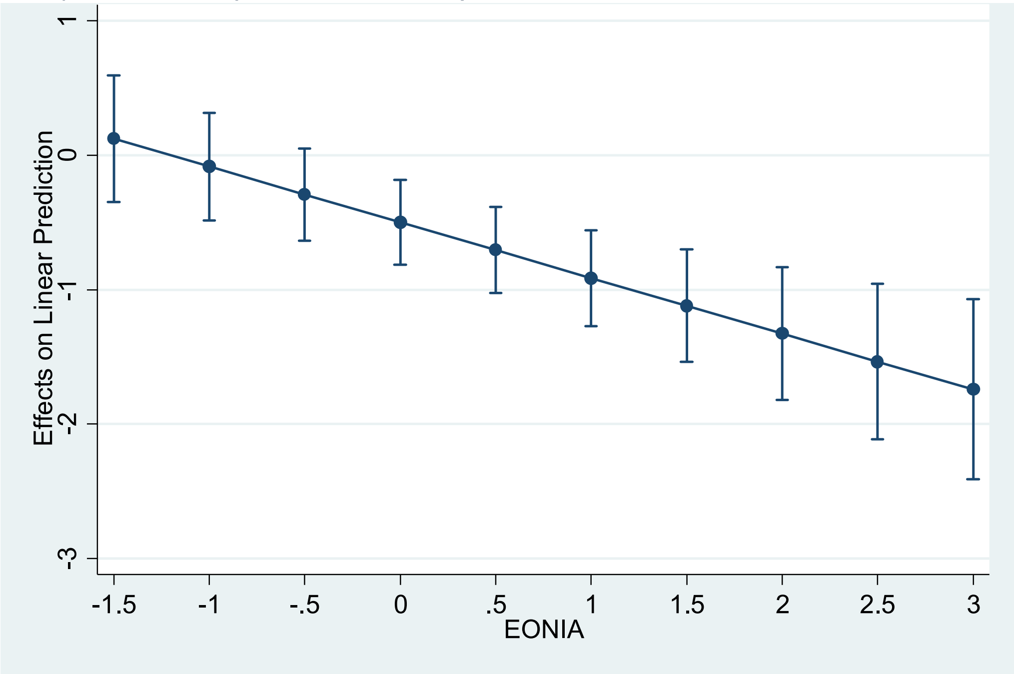


Figure 6b: Specification 2 (Table 8 – Annex 1)

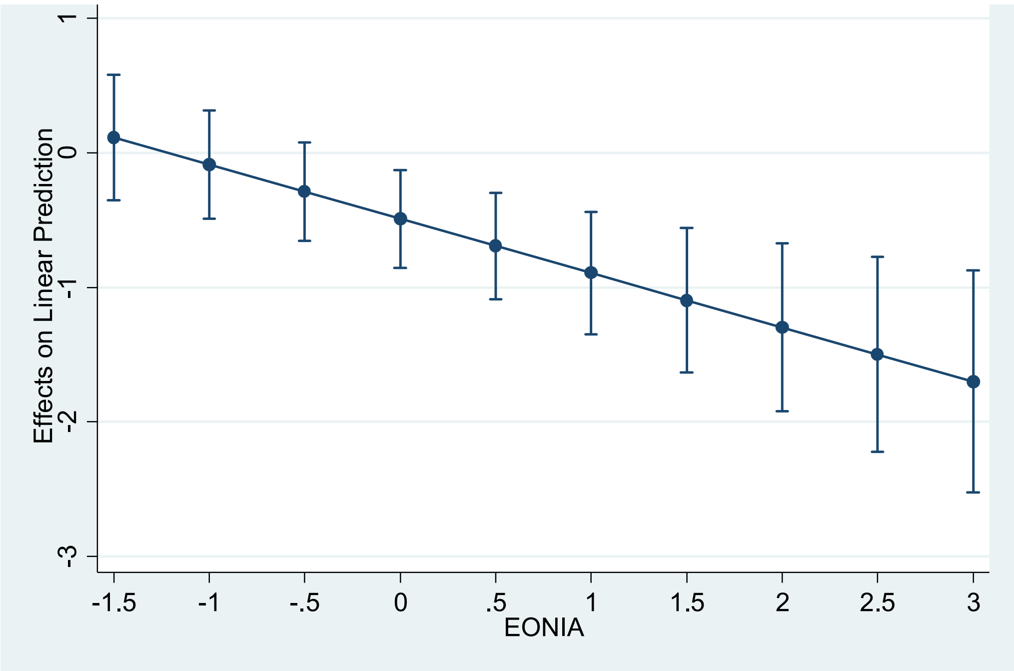


Figure 6c: Specification 3 (Table 8 – Annex 1)

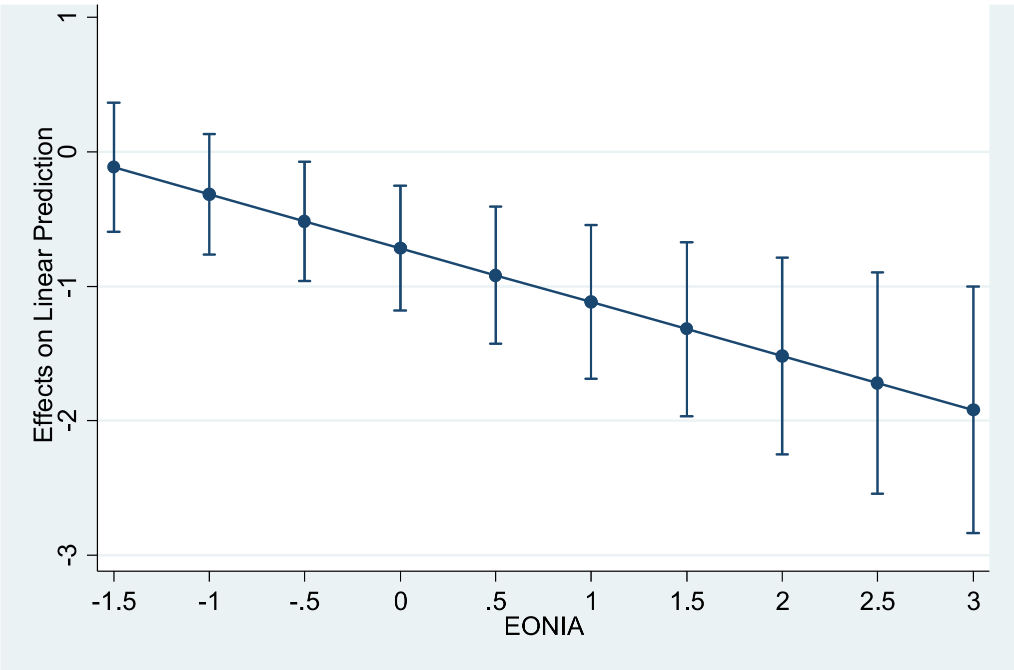


Figure 6d: Specification 4 (Table 8 – Annex 1)

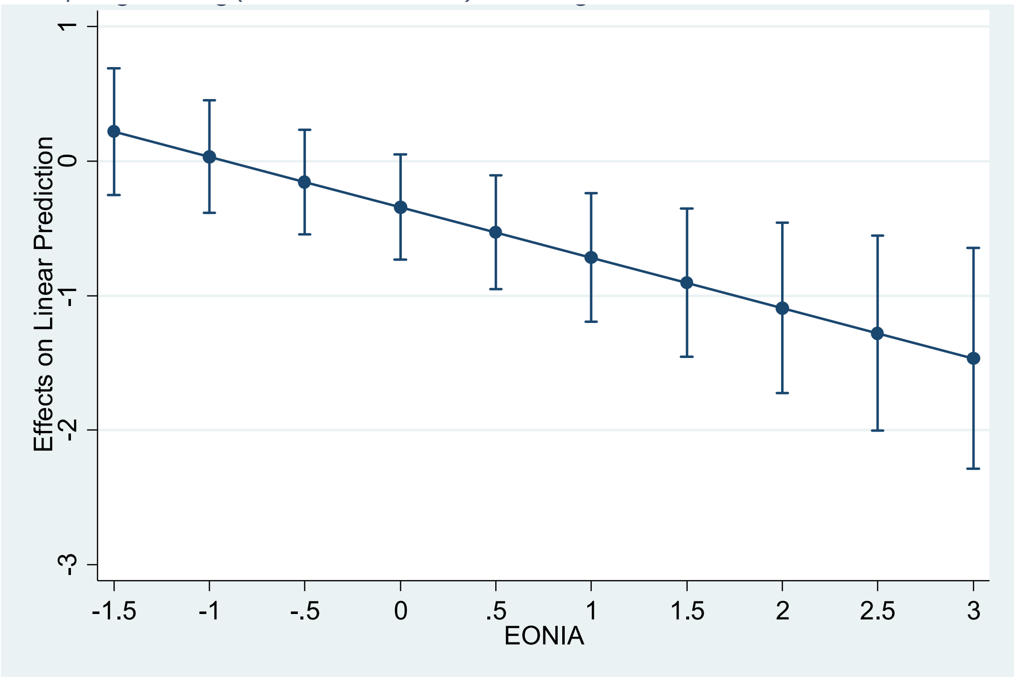


Figure 7: Marginal effect of real GDP growth on NPLs (95% confidence interval) given the level of uncertainty (EONIA) conditional to the credit cycle (y-axis: percentage change in the NPL ratio; x-axis change in EONIA rate in percentage points)

Figure 7a: Specification 1 (Table 9 – Annex 1)

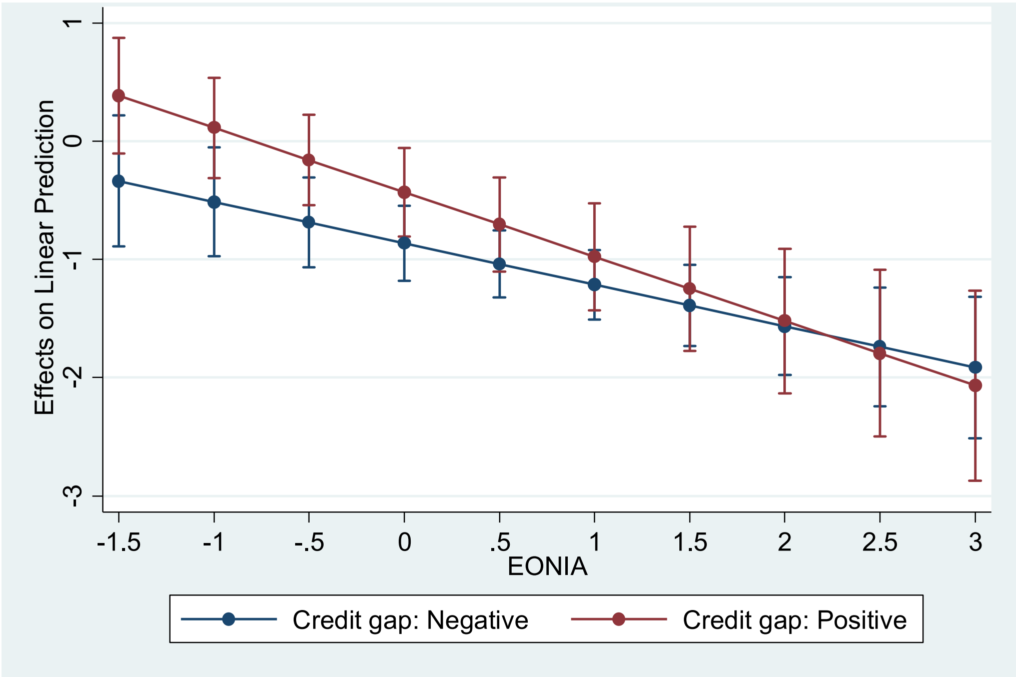


Figure 7b: Specification 2 (Table 9 – Annex 1)

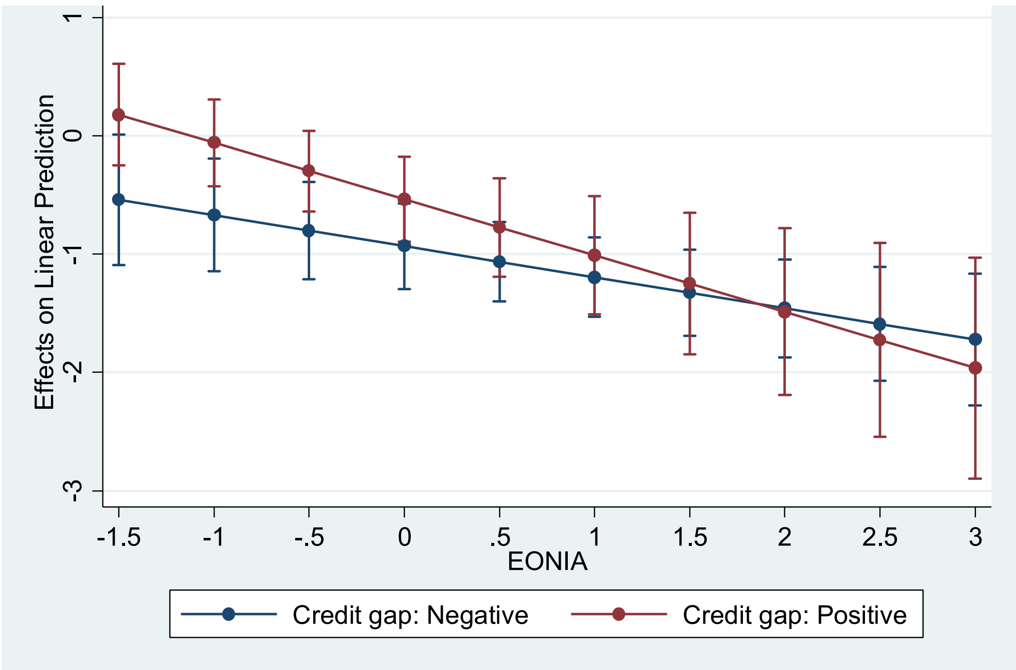


Figure 7c: Specification 3 (Table 9 – Annex 1)

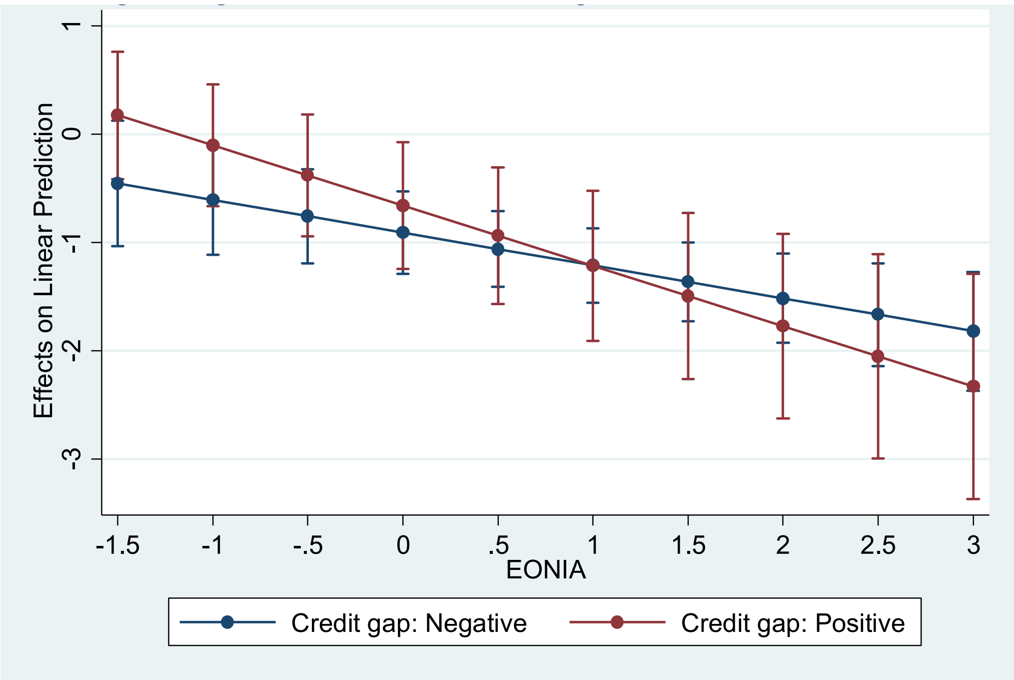
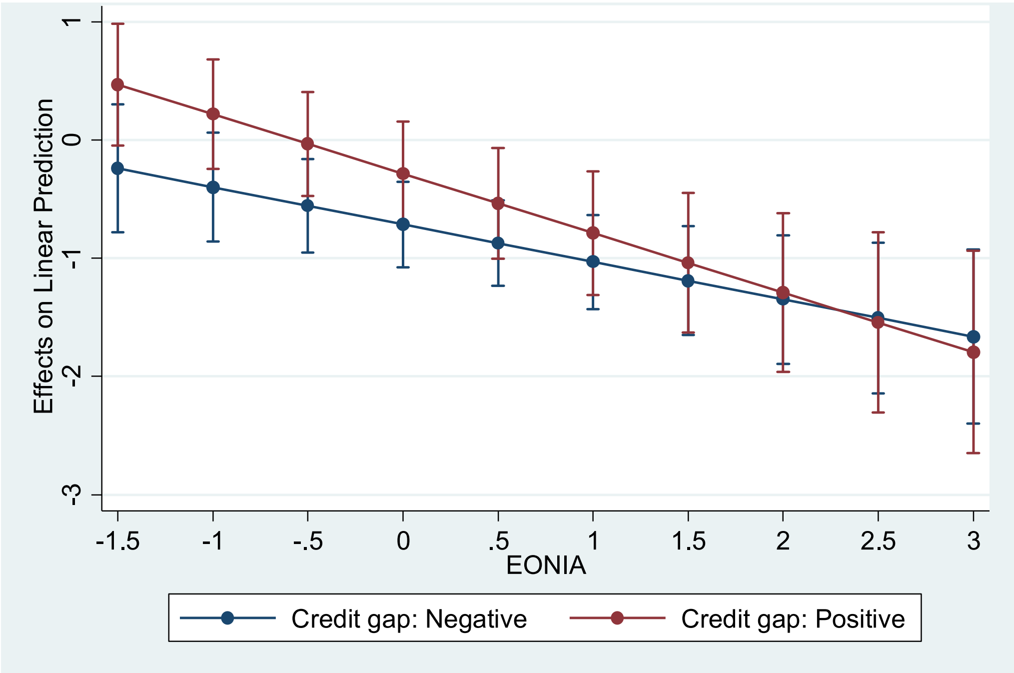


Figure 7d: Specification 4 (Table 9 – Annex 1)



Chapter 4

Beyond Debt Stocks: A Flow View of Sovereign Debt Sustainability[♣,*,*,♦]

This paper is a joint work with Dr. Aitor Erce.

Traditional approaches to study sovereign debt sustainability, heavily dependent on debt stock metrics and giving little role to debt flow indicators, are increasingly seen as insufficient. We inform this debate by analysing the ability of gross financing needs, the debt flow metric that currently complements the traditional debt sustainability analysis, to provide information about a sovereign's likelihood of distress beyond that provided by debt stock metrics. We show that stock and flow metrics need to be assessed in combination, and document an important role for gross financing needs when debt stocks are high. If debt is above 60%, reducing gross financing needs by one percent of GDP translates into 10 basis points lower sovereign spreads. We find that roll-over needs play a critical role in driving this effect. Our findings help understand how countries can sustain large debt stocks without suffering fiscal crises and, to the extent that official lending affects refinancing needs, also inform the literature on crisis resolution.

Keywords: Sovereign sustainability, debt stocks, financing needs, debt maturity.

JEL Codes: H62, H63, F34

[♣] We thank Harry Huizinga, Dennis Bams, Daragh Clancy, Andreja Lenarcic, Lorenzo Ricci, Rolf Strauch, Javier Suarez, Vilem Valenta and seminar participants at the European Stability Mechanism, European Commission, International PhD colloquium and 42nd Meeting of the Spanish Economic Association for their insightful comments. The views herein are the authors' and not those of The European Stability Mechanism or its Management Board.

[*] European Stability Mechanism, Circuit de la Foire Internationale 6a, L-1347, Luxembourg.

[♦] Corresponding author: C.Gabriele@esm.europa.eu.

4.1 Introduction

An integral element of any official support program is the analysis of public debt sustainability (IMF, 2014). As no single metric can provide a reliable risk assessment of debt sustainability on a cross-country basis, this is often done by combining debt sustainability analysis (DSA) with risk thresholds (IMF, 2013b).[47] This approach, although also touching upon economic growth and roll-over risks, pays most attention to the level of debt.[48] In practice, the focus on debt stocks has led to policy recommendations along the lines of "*for debt to be sustainable, the debt stock should decline to level X by year Y"*. [49]

In this paper we show that, in assessing sovereign risk, the key question is whether the country can secure the necessary funds to cover its financing needs. The stock of debt represents the amount of money *borrowed*, but not the flow of obligations required thereafter. Different amortisation schedules, different coupon structure and instruments (e.g., floating, fixed, linked, other currency issuances) can give different meanings to the same debt stock (Dias et al., 2014). In this paper we quantify the extent to which differences in refinancing needs for a given level of debt matter for markets' perception of sovereign risk. We document sizeable effects. When debt is above sixty percent of GDP, reducing refinancing needs by one percentage point of GDP can translate into ten basis points lower spreads.

This issue came into focus when the International Monetary Fund (IMF) engaged in the euro area. By providing concessional and back-loaded loans, euro official creditors reduced both debt financing costs and the need to roll it over. According to Corsetti et al. (2017, 2018) or Gourinchas et al. (2017), this helped reduce the risk of a disordered default for a given debt level. Emerging voices of discomfort argued that, in such circumstances, debt sustainability should also be linked to flow-related debt metrics. This triggered a change in official debt sustainability analyses (DSA).[50] To limit rollover risk, official DSA now monitors whether *gross financing needs* exceeds a pre-determined threshold (Hagan et al., 2017). As described in Zettelmeyer et al. (2017), *gross financing needs (*GFN) captures forthcoming financing needs by adding up interest payments, principal repayments, and primary deficits. Given that liquidity crises arise from mismatches between financing needs and sources, one would expect stress periods to be more likely when GFN are larger.

---

[47] The IMF guidelines dictate a "risk-based" approach. The approach categorizes countries as "lower-scrutiny" or "higher-scrutiny", based on at least one of the following benchmarks being meet: (i) current or projected public debt above 60 (50) percent of GDP in advanced (emerging) economies, (ii) current or projected gross financing needs exceed 20 (15) percent of GDP, and (iii) seeking or having exceptional access to IMF resources.

[48] The IMF defines a forward-looking view of sustainability (IMF, 2013a): "*In general terms, public debt can be regarded as sustainable when the primary balance needed to at least stabilize debt under both the baseline and realistic shock scenarios is economically and politically feasible, such that the level of debt is consistent with an acceptably low rollover risk and with preserving potential growth at a satisfactory level. Conversely, if no realistic adjustment in the primary balance – i.e., one that is both economically and politically feasible – can bring debt to below such a level, public debt would be considered unsustainable*."

[49] As an example, according to the 2012 Euro group framework reaching a debt-to-GDP ratio of 124% in 2020 and remaining substantially lower than 110 percent of GDP in 2022 would ensure Greece's debt sustainability.

[50] See Schumacher and Weder di Mauro (2016) or Zettelmeyer et al. (2017).

According to this rationale, while a too large debt stock could signal solvency problems, significant financing needs create liquidity risk. In fact, one of the main lessons from the theoretical literature on this field (see Cole and Kehoe 2000, Aguiar and Amador 2014, or Aguiar et al. 2016) is that solvency and illiquidity are intertwined, as the possibility of suffering a liquidity crisis (a non-fundamental run in the bond market) increases at relatively high debt levels (when solvency may not be guaranteed).[51] Our translation of this crucial link between insolvency and illiquidity into our empirical analysis is that while linking solvency and liquidity to stock and flow debt metrics is logic setting separate thresholds for stock and flow debt is not enough (see also Corsetti et al. 2018). A vast literature focuses on the linkage between debt stocks and sovereign risk.[52] Similarly, numerous contributions relate debt flows and sovereign spreads.[53] Instead, few papers combine stock and flow debt features. One exception is Dias et al. (2014), who study net present value measures of debt.[54] In this paper we show that studying the interaction between debt stocks and debt flows provides a more accurate way to study sovereign risk than setting separate thresholds for stocks and flows.[55]

Our analysis also relates to an insightful finance literature that studies sovereign risk using the contingent claim analysis (Gray et al. 2007).[56] According to the contingent claim analysis (CCA), a decline in the value of assets below the level of promised debt payments helps determine the default likelihood. Gray et al. (2007) assess sovereign asset values by assuming that sovereign liabilities are a contingent claim on the asset side of the balance sheet.[57] Within the CCA framework, the debtor’s default likelihood is a non-linear function of the volatility and average of liabilities, the risk-free rate and the distress barrier.[58] The distress barrier, which collects the present value of the promised payments on debt, is the critical concept linking CCA and our approach. Gray et al. (2007) estimates it as the sum of interest payments, short-term debt and one half of the stock of long-term debt, which is remarkably close to GFN.[59] According to the CCA, a worsening of the primary balance, larger interest or principal payments, increases

[51] While in those models, the extent to which a liquidity crisis breaks out depends on a sun-spot, the same feature is shared by models where the liquidity run does not require sun-spots (see Morris and Shin 2006 or Corsetti et al. 2006).

[52] See, for instance, Afonso et al. (2015) or Bernoth & Erdogan (2010).

[53] The component of GFN most often found as an explanatory factor is the primary balance. Attinasi et al. (2009) and Ardagna et al. (2004), who study the determinants of yields, account for the debt service.

[54] Also related, Bassanetti et al. (2016) show that whether debt is increasing or decreasing affects spread.

[55] We focus on how this is true on past data, but we do not aim to use this relation to perform forecast.

[56] Born as a generalization of option pricing theory (Black–Scholes 1973 or Merton 1973), the contingent claims approach has been applied to a wide variety of claims. When applied to credit risk, it is known as the “Merton Model”.

[57] Thus, using observed prices and volatilities of market-traded securities one can estimate implied asset values and volatilities. Gray et al. (2007) use this framework to calculate the implied value and volatility of assets.

[58] The CCA defines the default likelihood is a non-linear function of the volatility (σ_{debt}) and mean (μ_{debt}) of liabilities, the risk-free rate (r^f) and the distress barrier (DB_t): $\Pr(Default_t) = f(\sigma_{debt}, \mu_{debt}, r^f, \mathrm{DB}_t, , \varepsilon_t)$, where ε_t represents a set of other factors and shocks driving default risk.

[59] The parallel between the CCA and this paper implies $\mathrm{DB}_t \cong \mathrm{GFN}_t - PB_t$, where PB_t stands for the primary balance.

in debt stocks, or any combination of these factors, will bring a country closer to its default barrier, increasing the probability of a debt crisis.

We apply panel regression techniques to study the joint behaviour of gross financing needs, debt stocks and various measures of a country's solvency, using a newly built annual dataset that spans 1995 to 2015 and includes 23 European Union countries. First, we build a binary indicator of fiscal stress using sovereign spreads, and study its relation with debt stocks, gross financing needs and the interaction of the two.[60] Then, we repeat the exercise using a continuous indicator of sovereign stress, the ten-year sovereign spread.

We find that the effect of debt stocks on sovereign risk is dependent on the level of gross financing needs. Countries with large debt levels face more intense pressure if GFN increases. According to our estimates, when public debt is around 110% of GDP, a one percentage point increase in GFN increases the probability of a fiscal crisis by one percentage point. When focusing on sovereign spreads, we find that if debt is above 60% of GDP, a one percentage point increase in GFN increases spreads by five basis points. We further decompose the effect of GFN into its subcomponent, and find that roll-over needs are the main drivers of this effect.

Our analysis complements the existing literature by showing that jointly considering flow and stock debt measures delivers a more accurate picture of impending risks to sustainability. These findings have two important implications. First, they show that assessing sovereign solvency requires a simultaneous consideration of both flow and stock features of public debt. In fact, our results confirm that focusing on stock and flow metrics separately is more likely to lead to wrong conclusions. Second, our results on the role that large debt redemptions and debt stocks have in driving sovereign stress underline one channel through which official lending can enhance its effectiveness in the resolution of fiscal stress.[61]

The rest of this paper is organised as follows. Section 2 gives an overview of the data and some stylised facts. Our econometric approach and results are described in sections 3 and 4 respectively. Finally, section 5 concludes.

4.2 Data

We build a panel of general government financing needs and debt stock series for twenty-three European Union countries for the period 1995 to 2015, using the European Central Bank Statistical Data Warehouse.[62] Our gross financing needs (GFN) indicator adds up upcoming interests (on accrual basis) and principal debt

[60] The explanation on how it is built is at the beginning of Section 3.

[61] See Sandri (2015), Muller et al. (2016), Corsetti et al. (2017) or Abraham et al. (2017).

[62] The countries are: Austria, Belgium, Finland, France, Greece, Italy, Portugal, Spain, Denmark, Poland, Romania, Slovenia, Slovakia, Sweden, Czech Republic, Croatia, Hungary, Malta, Cyprus, Bulgaria, Latvia, Lithuania and Netherlands.

payments (outstanding Maastricht debt maturing in less than one year ahead), and primary deficits.[63]

$$\mathrm{GFN}_t = \mathrm{PB}_t + \mathrm{IP}_t + PR_{t-1} = Deficit_t + PR_{t-1}$$

where PB_t stands for the primary balance at time t, and is defined such that a positive value refers to a deficit. IP_t stands for the interest payments in the corresponding year, and PR_{t-1} represents the amount of principal repayment due in one year at the end of year t-1. By construction the variable measures sovereign gross financing needs in the year ahead.[64] Figure 1 shows the evolution of the breakdown of the GFN ratio for a selection of our sample countries.[65] Following the global crisis, gross financing needs increased significantly for most countries. This increase was driven by increasing deficits and roll-over needs that the increasing debt stocks were generating. Since 2012, gross financing needs have decreased almost everywhere. Interestingly, in Romania, Portugal, and Cyprus, although debt stocks kept increasing, these dynamics decoupled from the dynamics of gross financing needs (see Figure 1).

[Figure 1]

To quantify combinations of the flow and stock properties of debt, we construct a Stock-Flow Pressure (SFP) Index. We do this by multiplying, for each country/year pair, debt stock (as % of GDP) and gross financing needs (as % of GDP). Figure 2 plots the SFP index, re-scaled between 0 and 1, against the debt stock.

[Figure 2]

In line with the literature, we proxy sovereign risk using the 10-year sovereign bond spreads, which we compute as spreads against the German Bund.[66] We use a set of variables to control for both country and global features. Table A1 in the Appendix enumerates all the variables we use, table A2 shows their sources, and table A3 shows the pairwise correlation matrix of our variables. Figure 3 summarizes, using all countries in the sample, simple OLS-based correlations between sovereign spreads and our three debt metrics: debt to GDP, GFN to GDP, and the SFP index. Panel A shows that while both lagged debt and lagged GFN contribute positively to the increase in spreads, the interaction of debt and flow measures (SFP Index) correlate yet more strongly with spreads. Panel B focuses on the period 2006-2015, when spreads became more reactive to fiscal events, and shows even stronger relations.[67]

[Figure 3]

[63] More specifically, we collected the following variables: primary balance, interest payable (accrual basis), and outstanding Maastricht debt with any original maturities, maturing in the coming year.

[64] PR_{t-1} includes all the outstanding debt at the end of the previous year (securities, loans and any debt under Maastricht rules) with all types of original maturities, which have short-term residual maturity.

[65] We note that both the time span and frequency of the dataset were limited by the availability of debt redemption data.

[66] Intuitively, the private market uses information on a country's gross financing needs, among other things, when it decides at what price to trade the sovereign's debt securities in secondary markets.

[67] For the three debt metrics, R-squared doubles in the sub-sample.

Finally, to evaluate whether combining flow and stock metrics can help predict the occurrence of crisis, we need a fiscal crisis indicator. As there is no single comprehensive dataset that consistently reports fiscal crises in our sample countries, we construct our own indicator. Our fiscal crisis variable is a binary indicator taking value one when a country is seen as suffering a large pressure in its secondary sovereign bond market. We define a country as being under fiscal stress when any of the following two conditions is met: (i) the annual average sovereign spread for a country/year pair is larger than 350 basis points and it went above 500 basis points at least for a month during the year, (ii) the country is under an official programme from the International Monetary Fund or one of the euro area bailout funds. We assign a one when the country is under fiscal stress according to our definitions, and a zero otherwise. Figure 4 details all the country/year observations meeting these requirements.[68] Most of the distress events in our binary dependent variable are concentrated around the sovereign debt crisis in Europe, some around the great recession and a few more at the beginning of the 2000s in some central and Eastern Europe countries.

[Figure 4]

4.3 Flows versus Stocks as an Early Warning

In this section we apply various econometric methods to evaluate the ability of stock and flow metrics to detect the occurrence of fiscal crises. In line with the literature on the measurement of sovereign risk in a cross-section of countries (see Lee et al. 2013), we use panel data techniques. We estimate the relationship between GFN (as percentage of GDP), debt (as percentage of GDP) and the fiscal crisis indicator using a simple specification in which stock and flow measures are entered independently. Furthermore, to assess whether combining flow and stock metrics allows us to capture the non-linear features described in the contingent claim approach, we augment the model with the interaction of GFN and debt, the SFP index.

We use two types of econometric model for estimating the probability of facing a fiscal crisis. First, we use a lineal probability model (LPM):

$$P(y_{it} = 1 | \, x_{it}) = \alpha + \beta \cdot \frac{Debt}{GDP}_{it-1} + \gamma \cdot \frac{GFN}{GDP_{it-1}} + \delta \cdot SFP_{it-1} + \theta \cdot Controls_{it-1} + \mu_i + \varepsilon_{it},$$

where y_{it} stand for the fiscal crisis indicator and x_{it} is our set of explanatory variables including real GDP growth, change in debt, and three global factors (VIX, world growth, and the US 10-year yield).[69] In addition, we estimate the model using a logit specification:

[68] Our fiscal stress episodes include more events than an IMF dataset on fiscal crises (Gerling et al., 2017). In particular, the main difference is that our variable assigns a stress event also when an EFSF/ESM financial assistance programme was requested by the government of a country (e.g., Cyprus, Greece, Spain, Portugal), or also an EU balance of payments assistance programme was granted (e.g., Romania and Hungary).

[69] The choice of the set of control variables is based on the existing literature on sovereign defaults and risk, and includes the real GDP growth, which is one of the main drivers of sovereign spreads. We use world GDP growth to control for economic developments not related to the countries. The change in debt enters the model in order to control for the increases in debt with maturity beyond one year. We include also VIX and

$P(y_{it} = 1 | x_{it}) = \frac{e^W}{1+e^W}$, with $W = \alpha + \beta \cdot \frac{Debt}{GDP}_{it-1} + \gamma \cdot \frac{GFN}{GDP}_{it-1} + \delta \cdot SFP_{it-1} + \theta \cdot Controls_{it-1} + \mu_i + \varepsilon_{it}$

Within these regressions, the key coefficients are β, γ and δ. According to our argument, for a given debt stock, countries with lower gross financing needs should suffer less fiscal risks ($\delta > 0$). While we find evidence that both flow and stock metrics are relevant, our results indicate that their interaction is the crucial element. Changes in the SFP Index lead to both higher probability of default and a significant and economically sizeable movement on sovereign spreads.[70] The results are detailed in Tables 1 through 4.

We start by establishing whether a relation between fiscal crises, debt, GFN and their interaction is reflected in the data. We show the results in Table 1. Column 1 shows the results when we regress our binary dependent variable against debt, GFN and the interaction among the two. In column 2 we control for a set of macroeconomic and financial controls. We define this as our *favourite* specification. In both specifications Gross financing needs have a negative and significant coefficient, stronger when we control for additional factors. The coefficient on the interaction term is positive and offsets the negative effect of gross financing needs once the debt-to-GDP ratio becomes larger than 60%. Column 3 shows very similar results when we correct for the presence of cross-sectional dependence.

In line with the theoretical literature, our results show that there is a level of public debt, above which, if gross financing needs increase, the probability of a fiscal stress episode follows suit. In fact, the way GFN affects the probability of a sovereign distress event now depends on the level of debt. Conversely, to understand what happens to the probability of a distress event when debt levels change we need to know the GFN level. One implication of this finding is that both GFN and debt stocks become increasingly relevant in determining sovereign risk as the other grows.

[Table 1]

We also quantify the contribution of gross financing needs to the probability of a fiscal crisis using panel logit models. Columns 4 and 5 in Table 1 show the results. Results in column 4 are qualitatively similar to those in Column 1. In column 5 we include additional controls. Our results show that the marginal effect of GFN on the probability of entering a period of fiscal stress becomes positive when debt is higher than 90% of GDP. For instance, as shown in Figure 6, a country with public debt of around 110% of GDP whose GFN increases by one percentage point of GDP, will increase its probability of going into a period of fiscal stress by one percentage point.

Next, we ask ourselves: does the use of an interaction between debt stocks and gross financing needs help in anticipating fiscal crises? In order to answer this

US 10-year sovereign bond yield to control for global market uncertainty and developments. US and global variables act similarly to time fixed effects. We included also the current account balance, the inflation rate and the real effective exchange rates, but given that they were not significant in any specification, we dropped them.

[70] In order to attenuate simultaneous determination issues, we lag all our independent variables by one year.

question, we compare two logit models. One is our favourite specification. The other is a model that excludes the interaction term.

Figure 7 plots the Receiver Operating Characteristic (ROC) curves for these two models. The ROC curve illustrates the diagnostic ability of a binary classifier system by plotting the true positive rate against the false positive rate at various thresholds. The larger the area, the better is the model in classifying events. It is clear from Figure 7 that the model including the interaction between debt and GFN performs better. In fact, Table 2 computes the areas under the ROC for both models and shows that they are statistically different.

[Figure 7 and Table 2]

This result has an important operational implication, as it points to the fact that indicators that combine information about gross financing needs and debt stocks have the potential to add to existing early warning indicators for sovereign distress.

4.4 The stock and flow of the link between sovereign spreads and debt

To study more granularly the relation between debt stock, debt flows and sovereign risk, we substitute our dummy dependent variable with secondary market spreads for 10-year benchmark sovereign bonds. Again, we begin with a simple specification in which stock and flow measures are entered independently. Then we augment the model with the SFP Index:

$$Spread_{it} = \alpha + \beta \cdot \frac{Debt}{GDP}_{it-1} + \gamma \cdot \frac{GFN}{GDP}_{it-1} + \delta \cdot SFP_{it-1} + \theta \cdot Controls_{it-1} + \mu_i + \varepsilon_{it}$$

We estimate these models by using standard panel OLS techniques. Again, our main interest is in coefficient δ. We assess the robustness of our results in various ways. First, we extend our model with additional controls. Second, we use different specifications for the error terms. We use both a standard correction for serial correlation and heteroscedasticity, and the Driscoll-Kraay (1998) estimator that also corrects for the presence of cross-sectional dependence. Next, we follow Ongena et al. (2017) and apply instrumental variables to our estimation. Finally, as in Bassanetti et al. (2016), we use a state-dependent approach and measure the effect of GFN in sub-samples with high and low debt.[71]

Table 3 shows the results. We start again by establishing that the relation between sovereign spreads, debt, GFN and their interaction is reflected in the data. As shown in the first column of Table 3, which presents the results of a panel regression with fixed effects, we find that the interaction term is positive and significant. In column 2 we add the same set of macroeconomic, financial and global controls as earlier. The results look even stronger.[72]

[71] This robustness exercise is presented in the appendix (see Table 5 and Figure 9).
[72] Again, GFN explains a part of the variation in spreads that the change in debt stocks cannot disentangle.

[Table 3]

Again, the interaction between debt flows and stocks appears as a fundamental driver of sovereign spreads within this framework. Columns 3 through 5 of Table 3 test robustness of these estimates: by adding year dummies (columns 3 and 5), and correcting the standard errors for cross sectional dependence (columns 4 and 5).[73]

In Figure 5, we use the estimated parameters for the debt-to-GDP ratio, GFN, and their interaction to create a visual representation of the implied non-linear relation between spreads, debt stocks, and GFN. The steepness of the curve highlights the importance of combined large debt stocks and flow measures in generating stress in sovereign bond markets. In turn, Figure 8 quantifies the extent to which GFN affects sovereign spreads depends on the level of debt (and vice versa) graphically. Panel A in Figure 8 plots the effect of increasing GFN by 1% of GDP (marginal effect) on spreads at different debt levels. Analogously, in Panel B we compute the marginal effect on spreads of increasing the debt stock by 1% of GDP. For low debt stocks, increases in GFN have no significant effect on spreads (and are thus not a concern). When debt ratios go above 60%, however, increases in GFN lead to larger sovereign spreads.[74] As shown in Panel A of Figure 8, if a country has a debt-to-GDP ratio of 100%, an increase of 1% in GFN translates, all else being equal, into a 10-basis point spread increase. Analogously, the same increase in GFN when the debt stock is at 80% would translate into a five basis point increase in spreads. On the other hand, Panel B shows that when GFN is low, debt stock increases do not affect spreads. But if GFN is above 20%, each percentage point of increase in debt stocks leads to a four-bps spread widening.[75]

Although our explanatory variables enter the estimations in lagged form, reverse causality might still be an issue. This is so because GFN is affected by sovereign spreads (mainly, but not only) through the interest payments component. We address this issue by instrumenting GFN and the interaction term using the previous year maturing debt and its interaction with contemporaneous debt. Our choice of instrument is motivated by the fact that maturing debt is exogenously pre-determined by governments' past issuance choices (see also Ongena et al., 2016). Column 6 of Table 3 shows the results of this panel instrumental variable regression.[76] Although with a smaller magnitude, the results hold.[77]

4.4.1 Decomposing Gross Financing Needs

Finally, we explore the role played by the different components of GFN. Given the important role roll-over needs plays in the theoretical literature (see Cole and

[73] A Pesaran test indicated the existence of cross-sectional correlation.

[74] We note that this endogenous turning point for the debt stock, obtained from our analysis, coincides with threshold for high scrutiny used by the IMF or that provided by the Maastricht rules.

[75] Note again that the threshold for GFN coincides with those currently used by the IMF and the ESM.

[76] The F-statistics of the first stage regressions for GFN and the interaction term are, respectively, 58 and 118.

[77] Using these estimates the marginal effect of GFN-to-GDP ratio on spreads is positive when the debt is above 111% of GDP, while the marginal effect of the debt-to-GDP ratio on spreads becomes positive when GFN is above 16.8% of GDP.

Kehoe 2002 or Hatchondo et al. 2016), we are especially interested on their effects. Specifically, we estimate the following model:

$$Spread_{it} = \alpha + \beta \cdot \frac{Debt}{GDP}_{it-1} + \sum_{X=\{P,I,A\}} \gamma^X \cdot X_{t-1} + \sum_{X=\{P,I,A\}} \delta^X \cdot \frac{Debt}{GDP}_{t-1} \cdot X_{t-1} + \theta \cdot Controls_{it-1} + \mu_i + \varepsilon_{it}$$

where P is the primary balance, I are the interest payments and A are the amortisations. The results from this model allow us to disentangle the channels through which changes in GFN affect sovereign risk. The coefficient δ^A, for example, informs us about the extent to which, at a given debt level, increases in debt amortisation (roll-over needs) affect sovereign spreads. In line with the theoretical insights from Cole and Kehoe (2002), if $\delta^A > 0$, the larger the underlying debt stock the more the increases in roll-over needs will affect sovereign spreads.

Table 4 shows all the results. In the first column, we show the results of the estimation of our *favourite* specification (with the macroeconomic and financial controls) with the only difference that now the subcomponents of GFN enter independently the model (and the same happens for the interactions with debt). As we can see from column 1, for the debt amortisation component of GFN, its interaction with debt, and marginally the interaction between deficit and debt, the correlations are significantly different from zero and the signs of the coefficients for debt amortisation are similar to the ones for GFN in Table 3. Columns 2 through 4 in Table 4 show the results of the same regression including year dummies (column 2) and repeating the estimations in columns 1 and 2, but correcting the standard errors using an estimator that is robust to cross sectional dependence. Roll-over needs are important in explaining sovereign spreads both independently but also through their interaction with the debt stock.

[Table 4]

4.5 Conclusions

In this paper we use twenty years of data for twenty-three European Union countries to study the relationship between stock (debt to GDP) and flow (gross financing needs to GDP) metrics of public debt, fiscal crises, and sovereign borrowing costs. We show that jointly considering these flow and stock measures delivers a more accurate picture of impending risks and helps in better understanding what drives sovereign risk. We find that the effects of stock and flow measures reinforce each other above a critical threshold, and that sovereign roll-over needs are a critical element explaining this effect.

These findings have two important implications. First, they reinforce the idea that any evaluation of a country's debt sustainability needs to simultaneously consider both flow and stock features of the underlying public debt. Focusing only on stock metrics is likely to lead to the wrong conclusions. Second, given the role of official lending in smoothening refinancing needs (see Corsetti et al. 2018), our findings also inform a flourishing literature on the role of official financing in the resolution of sovereign stress. The reinforcing negative effect of debt redemptions and debt stocks on sovereign spreads underlines one channel through which official lending can be beneficial for countries' market access. High-debt countries can manage their redemption profile to reduce sovereign stress and access financial markets at better conditions while working on debt reduction.

4.6 References

Abraham, A., Carceles-Poveda, E., Liu, Y., and R. Marimon (2017), "On The Optimal Design of a Financial Stability Fund," *ADEMU Working Paper.*

Afonso, A., Aerghyrou, M., and A. Kontonoikas (2015), "The Determinants of Sovereign Bond Yields in the EMU." *ECB Working Paper 1781.*

Aguiar, M., and M. Amador (2014), "Sovereign Debt," *Handbook of International Economics Vol 4.*

Aguiar, M., Amador, M., Hopenhayn, H., and I. Werning (2016), "Take the Short Route: Equilibrium Default and Debt Maturity," *NBER Working Paper No. 22847*

Ardagna. S., Caselli, F., and T. Lane (2004), "Fiscal Discipline and the Cost of Public Debt Service – Some Estimates for OECD Countries," *ECB Working Paper 411.*

Attinasi, M., Checerita, C., and C. Nikel (2009), "What Explains the Surge in Euro-Area Sovereign Spreads During the Financial Crises of 2007-09?," *ECB Working Paper 1131.*

Bassanetti, A., Cottarelli, C., and A. Presbitero (2016), "Lost and Found: Market Access and Public Debt Dynamics," *IMF Working Paper 16/253.*

Bernoth, K., and Erdogan, B. (2010), "Sovereign Bond Yield Spreads: A Time-Varying Coefficient Approach," *DIW BERLIN, Discussion Paper 1078.*

Black, F. and Scholes M. (1973), "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy, 81(May–June): 637–654.*

Cole, H., and T. Kehoe (2000), "Self-Fulfilling Debt Crises," *The Review of Economic Studies, 67(1), 91-116.*

Corsetti, G., Guimaraes, B., and N. Roubini (2006), "International lending of last resort and moral hazard: A model of IMF's catalytic finance," *Journal of Monetary Economics, Elsevier, vol. 53(3), pages 441-471.*

Corsetti, G., Erce, A., and T. Uy (2018), "Debt Sustainability and the Terms of Official Support," *ADEMU Working Paper.*

Dias D., Richmond, C., and M. Wright (2014), "The Stock of External Sovereign Debt: Can We Take the Data at Face Value?," *Journal of International Economics, Volume 94, Issue 1.*

Gerling, K., Medas, P., Poghosyan, T., Farah-Yacoub, J. and Y. Xu (2017), "Fiscal Crises," *IMF Working Paper no. 17/86.*

Gourinchas, P.O., Martin, P., and T. Messer (2017), "The Economics of Sovereign Debt, Bailouts and the Eurozone Crisis," *Unpublished manuscript*

Gray, D.F., Merton, R.C., Z. Bodie (2007), "Contingent Claims Approach to Measuring and Managing Sovereign Credit Risk" - *Journal of Investment Management, 2007*

Hagan, S., Obstfeld, M., and P. Thomsen (2017), "Dealing with Sovereign Debt" *– The IMF Perspective, IMF Blog post.*

IMF (2013a), "Staff Guidance Note for public debt sustainability analysis in market- access countries," *International Monetary Fund.*

IMF (2013b), "Greece: Ex Post Evaluation of Exceptional Access under the 2010 Stand-By Arrangement," *International Monetary Fund.*

IMF (2014), "The Fund's Lending Framework and Sovereign Debt-Preliminary Considerations," *International Monetary Fund.*

IMF (2016), "IMF Survey: IMF Reforms Policy for Exceptional Access Lending," *International Monetary Fund.*

Lee, C., Chiu, Y., and C. Chang (2013), "Insurance Demand and Country Risks: A Nonlinear Panel Data Analysis," *Journal of International Money and Finance, Volume 36, Pages 68-85.*

Driscoll, J. and A. Kraay (1998), "Consistent Covariance Matrix Estimation With Spatially Dependent Panel Data," *The Review of Economics and Statistics, 80, issue 4, p. 549-560.*

Merton, R. C. (1973), "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science 4, 141–83.*

Morris, S. and H.S. Shin (2006), "Catalytic finance: When does it work?," *Journal of International Economics, 70, issue 1, p. 161-177.*

Müller, A., Storesletten, K., and F. Zilibotti (2016), "The Political Color of Fiscal Responsibility." *Journal of the European Economic Association, Volume 14, Issue.*

Ongena, S., Popov, A. and Van Horen, N. (2016), "The invisible hand of the government: "Moral suasion" during the European sovereign debt crisis," *Centre for Economic Policy Research.*

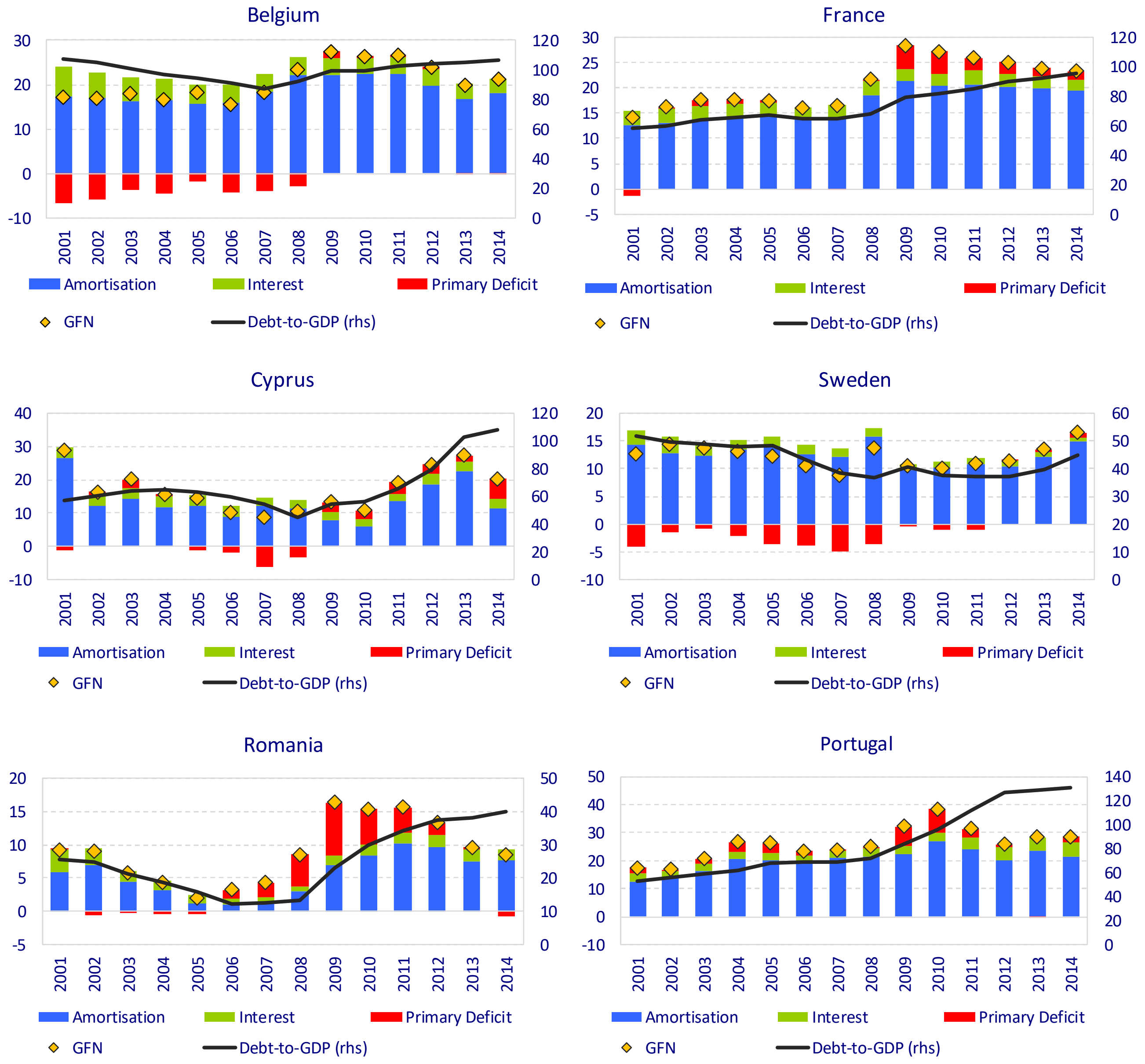
Schumacher, J., and B. Weder di Mauro (2016), "Greek Debt Sustainability and Official Crisis Lending," *Brookings Papers on Economic Activity, Fall 2016*

D. Sandri (2015), "Dealing With Systemic Sovereign Debt Crises: Fiscal Consolidation, Bail-Ins or Official Transfers?," *IMF Working Paper no. 15/233.*

J. Zettelmeyer, Kreplin, E., and U. Panizza (2017), "Does Greece Need More Official Debt Relief? If So, How Much?," *Peterson Institute for International Economics, Working Paper 17-6.*

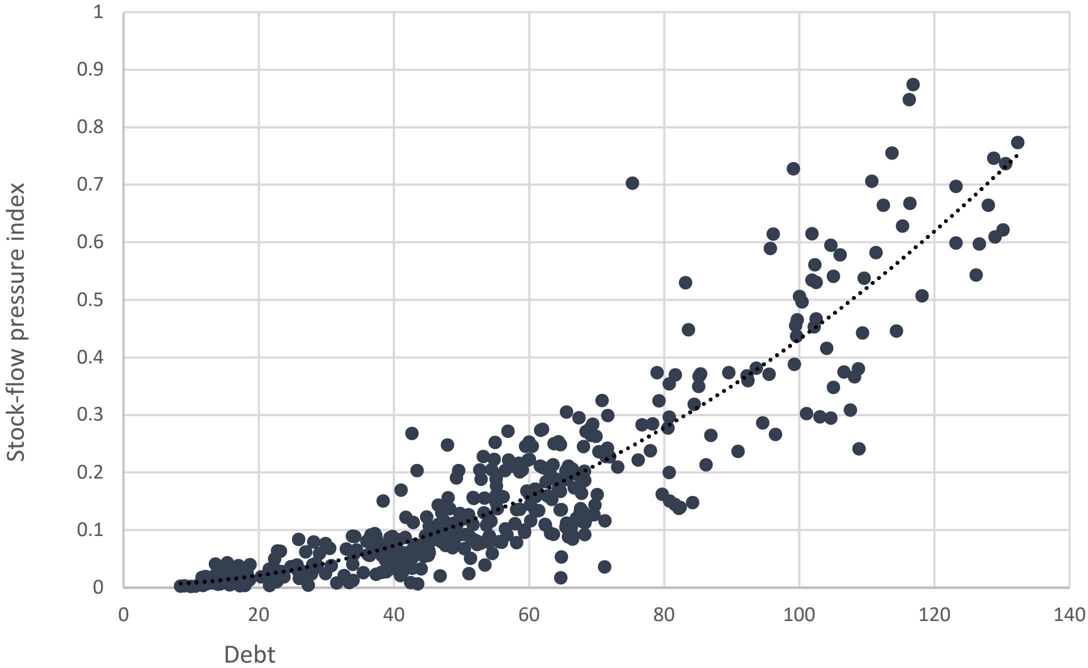
4.7 Annex 1. Data

Figure 1: Breakdown of GFN in a selection of EU countries



Source: Authors' calculations based on ECB data. GFN, its sub-components and debt are in percent of GDP. GFN and its sub-components are measured on the left-hand axis, while the debt-to-GDP ratio is measured on the right-hand axis. Both axis units are in % of GDP.

Figure 2: SFP Index (between 0 and 1) and debt stocks (% GDP)



Source: Authors' calculations based on ECB data. Debt is in percent of GDP.

Table A1: Coverage

Country	Start	End
Austria	1995	2015
Belgium	1995	2015
Denmark	1995	2015
France	1995	2015
Italy	1995	2015
Netherlands	1995	2015
Sweden	1995	2015
Finland	1995	2015
Greece	1995	2015
Malta	1998	2015
Portugal	1995	2015
Spain	1995	2015
Cyprus	1998	2015
Bulgaria	2000	2015
Czechia	1999	2015
Slovakia	1998	2015
Latvia	1998	2015
Hungary	1998	2015
Lithuania	1998	2015
Croatia	2003	2015
Slovenia	2000	2015
Poland	1998	2015
Romania	2002	2015

Table A2: Data sources

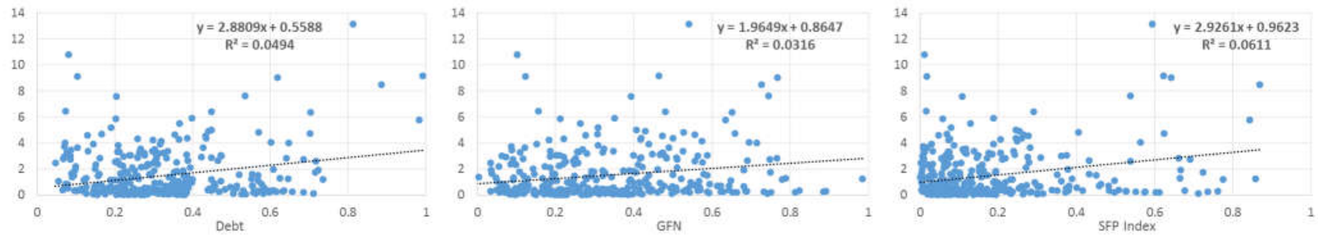
Variable	Source
Sovereign yields (10 year)	Eurostat
Amortisation	ECB
Primary deficit	Eurostat
Interest expenditure	Eurostat
Debt to GDP	ECB
Real GDP growth	Eurostat
World real GDP growth	World Bank
US 10 year yields	Federal Reserve Board
VIX	Wall Street Journal
EFSF/ESM programmes	EFSF/ESM
EU Balance of Payments programmes	European Commission
IMF programmes	IMF

Table A3: Correlations matrix

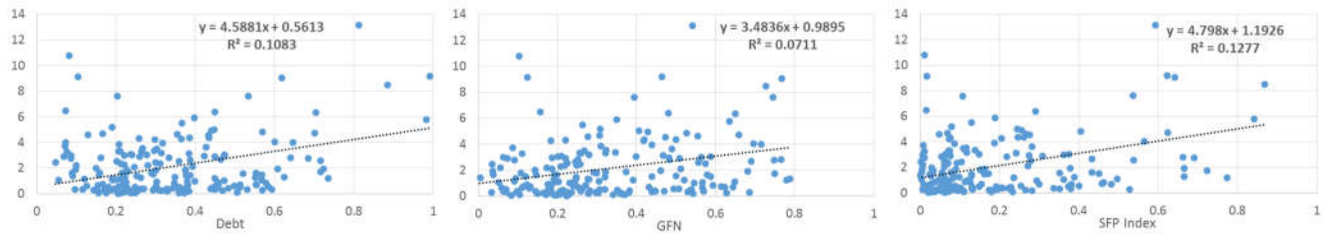
	Sovereign spreads	Debt · GFN	GFN to GDP	Debt to GDP	Real GDP growth	World GDP growth	US 10 year yield	VIX	HICP inflation	Current account to GDP	Primary balance to GDP	Interest payment to GDP	Debt amortisation to GDP	Change in debt to GDP	Change in REER
Sovereign spreads	1														
Debt · GFN	0.30*	1													
GFN	0.23*	0.91*	1												
Debt	0.21*	0.91*	0.78*	1											
Real GDP growth	-0.34*	-0.41*	-0.44*	-0.41*	1										
World GDP growth	-0.12*	-0.08	-0.14*	-0.06	0.61*	1									
US 10 year yield	-0.28*	-0.06	-0.03	-0.09*	0.44*	0.31*	1								
VIX	0.14*	-0.03	0.06	-0.07	-0.30*	-0.59*	0.02	1							
HICP inflation	0.33*	-0.19*	-0.17*	-0.31*	0.21*	0.1*	0.14*	0.18*	1						
Current account to GDP	-0.30*	0.05	0.06	0.18*	-0.27*	-0.09*	-0.1*	-0.07	-0.49*	1					
Primary balance to GDP	-0.41*	-0.05	-0.21*	0.05	0.33*	0.27*	0.41*	-0.10*	-0.05	0.30*	1				
Interest payment to GDP	0.12*	0.76*	0.70*	0.78*	-0.13*	0.04	0.42*	-0.02	-0.04	0.10*	0.28*	1			
Debt amortisation to GDP	0.06	0.89*	0.94*	0.78*	-0.33*	-0.05	0.06	0.02	-0.20*	0.18*	0.12*	0.74*	1		
Change in debt to GDP	0.28*	0.28*	0.42*	0.17*	-0.33*	-0.20*	-0.22*	0.13*	0.06	-0.12*	-0.54*	0.11*	0.2308*	1	
Change in REER	0.05	-0.13*	-0.12*	-0.19*	0.07	-0.07	0.06	0.07	0.36*	-0.22*	-0.06	-0.11*	-0.1359*	0.11*	1

Figure 3: Sub-samples correlations between sovereign spreads and debt, GFN and SFP Index

Panel A: Sample 1996-2015



Panel B: Sample 2006-2016



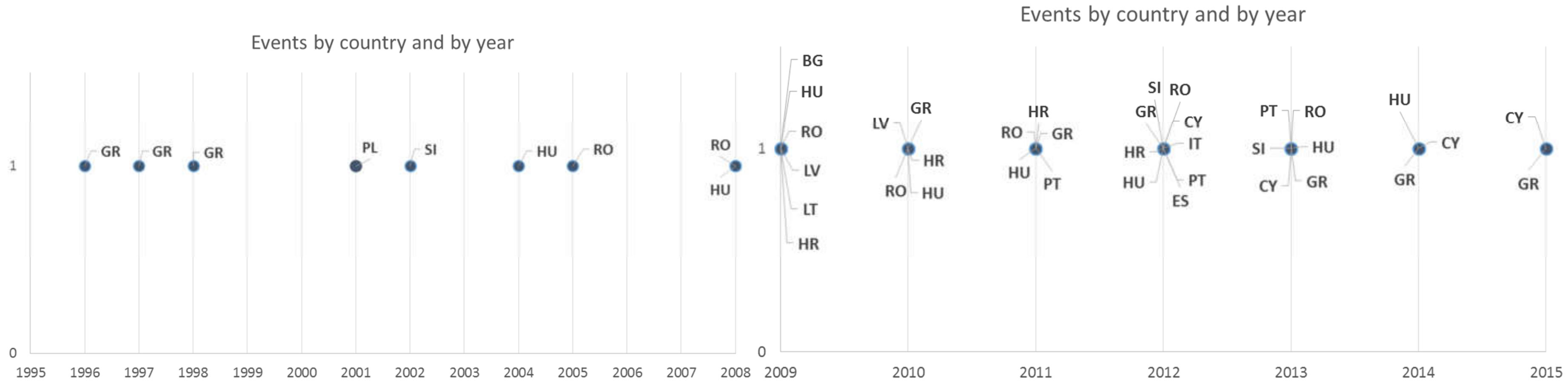
Source: Authors' calculations based on ECB data. The vertical axis measure spreads (in percent). Debt stands for the Government's debt-to-GDP ratio, and GFN for the gross financing needs, as percentage of GDP).

4.8 Annex 2: Results

Figure 4: Fiscal stress episodes

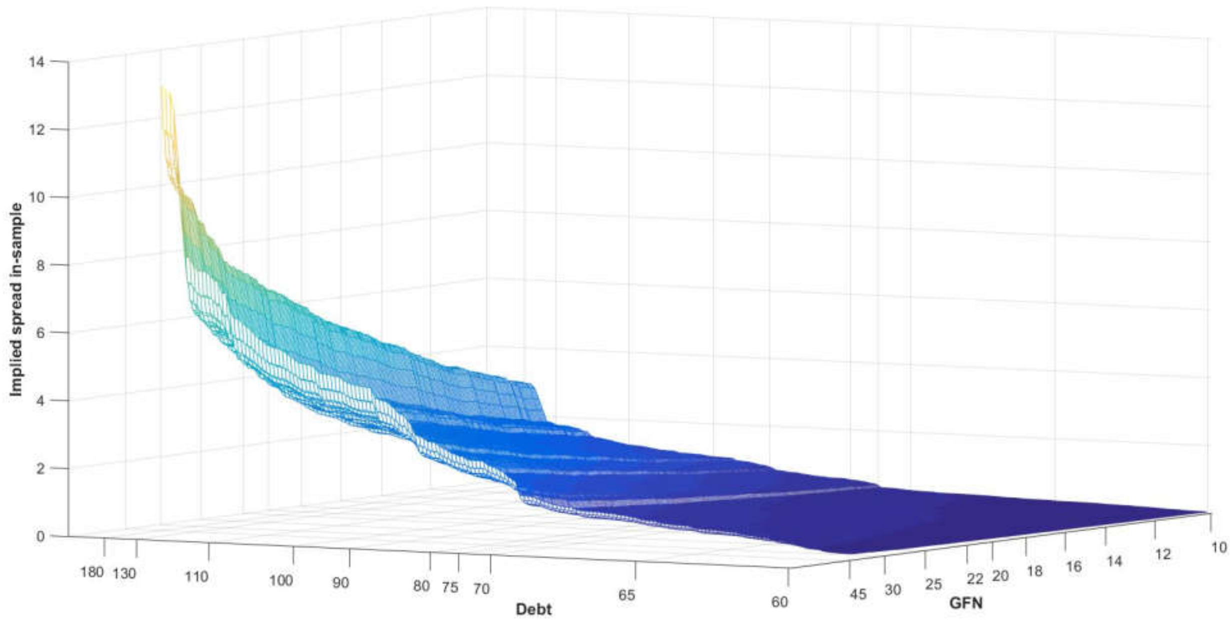
Panel A: Pre-crisis

Panel B: Post-crisis



Source: Authors' calculations based on Eurostat data. We define a country as being under fiscal stress when any of the following two conditions is met: (i) the annual average sovereign spread for a country/year pair is larger than 350 basis points and it went above 500 basis points at least for a month during the year, (ii) the country is under an official programme from the International Monetary Fund or one of the euro area bailout funds. We assign a one when the country is under fiscal stress according to our definitions, and a zero otherwise.

Figure 5: Spreads implied by debt, GFN, and the SFP Index



Source: Authors' calculations based on ECB data

Legend:

Debt and GFN are in % of GDP. Implied spreads are in percent. This 3D plot is produced in Matlab by using in-sample data and the coefficients estimated in Table 1, column 6. Each point in the chart is composed of three coordinates: GFN, Debt, and the outcome implied spread.

$$Implied\ spread_t = -0.142 * GFN_{t-1} + -0.012 * Debt_{t-1} + 0.00245 * (GFN * Debt)_{t-1}$$

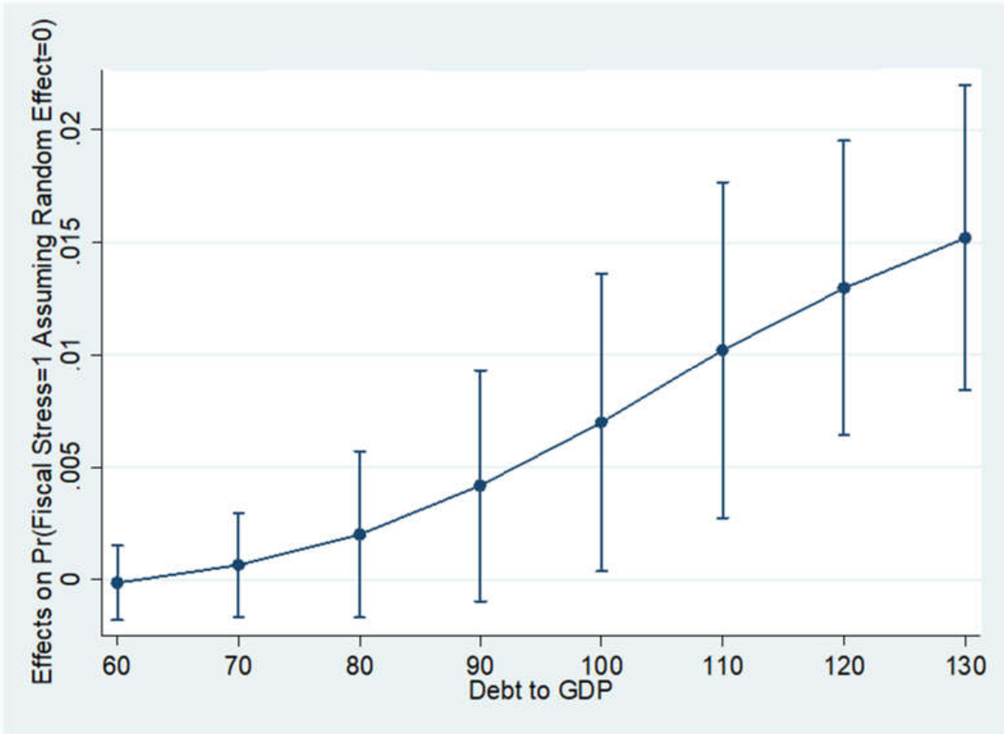
Table 1: Binary Dependent Variable

Dependent variable: Fiscal Stress Events	1	2	3	4	5
GFN	-0.004	-0.017***	-0.017***	-0.024	-0.349**
	-0.005	(0.005)	(0.005)	(0.067)	(0.148)
Debt	0.001	0.0003	0.0003	-0.012	-0.067
	(0.002)	(0.002)	(0.001)	(0.023)	(0.043)
Debt · GFN	0.00013*	0.0002***	0.0002***	0.00157*	0.00562***
	(0.0001)	(0.0001)	(0.0001)	(0.0009)	(0.002)
Change in Debt		-0.001	0.001		0.008
		(0.004)	(0.002)		(0.089)
Real GDP Growth		-0.031***	-0.031***		-0.348***
		(0.006)	(0.008)		(0.120)
World GDP growth		0.064***	0.064***		1.040***
		(0.014)	(0.016)		(0.303)
US 10 year yield		-0.021	-0.021*		-1.238***
		(0.013)	(0.012)		(0.453)
VIX		0.015***	0.015***		0.350***
		(0.003)	(0.002)		(0.095)
Constant	-0.028	-0.239*	-0.239**	-3.885***	-7.272*
	(0.089)	(0.137)	(0.085)	(1.309)	(3.884)
Observations	373	357	357	373	357
R-squared	0.062	0.221	0.221	-	-
Number of countries	23	23	23	23	23
Random Effects	NO	NO	NO	YES	YES
Country FE	YES	YES	YES	NO	NO

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

GFN and Debt are in percent of GDP. Change in Debt is the difference between Debt in t and Debt in (t-1), divided by GDP in (t-1). Columns 1 and 2 are estimated with standard panel regression techniques. Column 3 is estimated using Driscoll-Kraay estimator. Columns 1 through 3 are Linear Probability models. Columns 4 and 5 are estimated using Logit panel regression technique with random effects.

Figure 6: Marginal effectd of GFN on the probability of a fiscal stress event (90% confidence intervals)



Source: Authors' calculations based on ECB data. On the y-axis the units are increases in the probability of a fiscal stress event. On the x-axis, debt is in percent of GDP. The graph presents 10% error bands.

Legend: Figure 6 shows the marginal effect of gross financing needs on the increase in the probability of being in a fiscal stress event (as defined in the paper). Every point in this figure shows how the probability of a fiscal stress event would change given a change in gross financing needs of 1% at different levels of the debt-to-GDP ratio (all else equal). In order to draw this figure, we used estimates from Table 1, column 5.

Figure 7: Area under the receiver operating curve

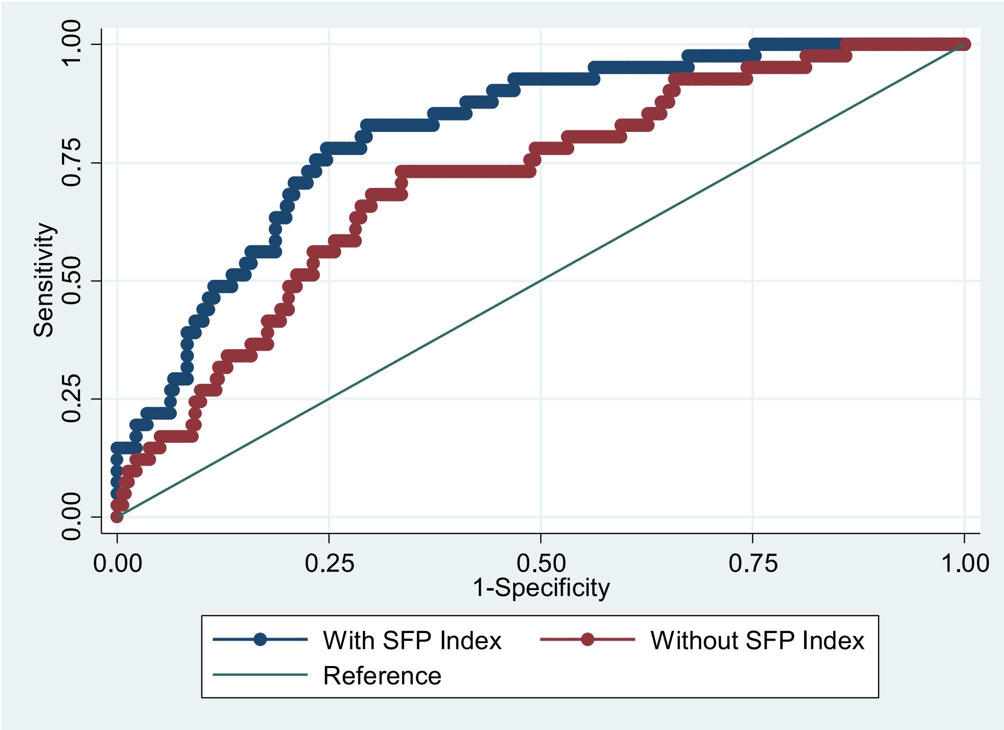


Table 2

Model	Observations	ROC Area	Standard Error	Asymptotic Normal [95% Confidence Interval]	
With SFP Index	357	0.814	0.032	0.752	0.877
Without SFP Index	357	0.711	0.041	0.631	0.792

Null Hypothesis: Area (With SFP Index) > Area (Without SFP Index)

$$\chi^2(1) = 16.23 \qquad Prob > \chi^2 = 0.0001$$

Table 3: Stock-Flow Combinations and Sovereign Spreads

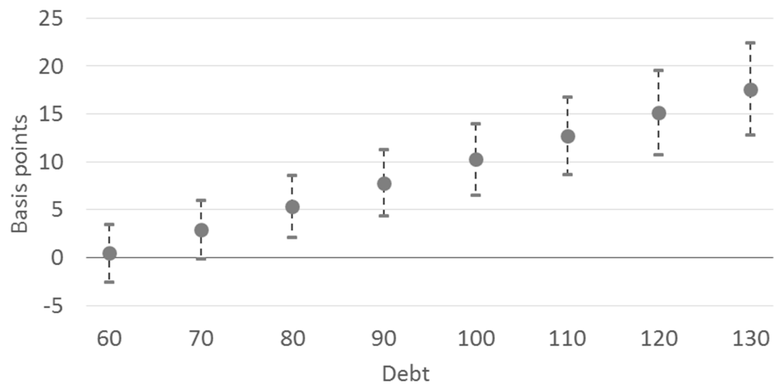
Dependent variable: 10y Spread	1	2	3	4	5	6
GFN	-0.027	-0.142**	-0.142***	-0.142**	-0.142**	-0.139***
	(0.027)	(0.052)	(0.0274)	(0.059)	(0.060)	(0.047)
Debt	-0.008	-0.012	-0.013	-0.012	-0.013	0.021
	(0.010)	(0.009)	(0.010)	(0.014)	(0.016)	(0.014)
Debt · GFN	0.0017***	0.00245***	0.00239***	0.00245**	0.00239**	0.00125**
	(0.0004)	(0.0008)	(0.0003)	(0.0011)	(0.0011)	(0.0006)
Change in Debt		0.021	0.024	0.021	0.024	0.056**
		(0.02)	(0.019)	(0.015)	(0.021)	(0.023)
Real GDP Growth		-0.240***	-0.255***	-0.240***	-0.255***	-0.237***
		(0.065)	(0.032)	(0.028)	(0.039)	(0.031)
World GDP growth		0.593***	1.518	0.593***	1.518***	0.585***
		(0.111)	(1.680)	(0.079)	(0.700)	(0.067)
US 10 year yield		-0.157	-0.178	-0.157***	-0.178*	-0.156**
		(0.095)	(0.343)	(0.054)	(0.095)	(0.068)
VIX		0.127***	0.129	0.127***	0.129***	0.128***
		(0.022)	(0.089)	(0.015)	(0.023)	(0.014)
Constant	0.500	-1.665***	-3.850	-1.665***	-	-
	(0.482)	(0.486)	(3.116)	(0.435)	-	-
Observations	373	357	357	357	357	346
R-squared	0.216	0.498	0.536	0.498	0.536	0.446
Number of countries	23	23	23	23	23	23
Country FE	YES	YES	YES	YES	YES	YES
Year FE	NO	NO	YES	NO	YES	YES

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

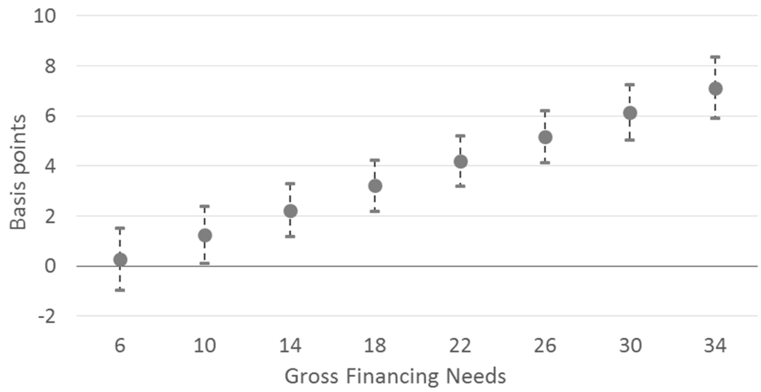
GFN and Debt are in percent of GDP. Change in Debt is the difference between Debt in t and Debt in (t-1), divided by GDP in (t-1). Columns 1 through 3 are estimated with standard panel regression techniques. Columns 4 and 5 are estimated using Driscoll-Kraay estimator. Column 6 reports the results of a panel instrumental variable estimation.

Figure 8: Debt stocks meet gross financing needs: the role of non-linearities

Panel A: Impact of gross financing needs on spreads, by debt levels



Panel B: Impact of debt on spreads, by gross financing need levels



Source: ECB and authors computations. GFN and debt are in percent of GDP. The graph presents 10% error bands.

Legend: Panel A shows the estimated marginal effect of gross financing needs on spreads, using estimates from Table 3, column 2. The figure reports how spreads react to 1% increases in gross financing needs by debt-to-GDP levels. Panel B shows the estimated marginal effect of the debt-to-GDP ratio on spreads, using estimates from Table 3, column 2. The figure reports how spreads react to 1% increases in debt-to-GDP gross financing needs by gross financing need levels.

Table 4: Stock-Flow Combinations and Sovereign Spreads - Breakdown of GFN

Dependent variable: 10y Spread	1	2	3	4
Amortisation	-0.126**	-0.128**	-0.126***	-0.128**
	(0.056)	(0.050)	(0.043)	(0.049)
Interest	-0.561	-0.700**	-0.561	-0.700
	(0.593)	(0.292)	(0.516)	(0.606)
Deficit	-0.055	-0.050	-0.055	-0.050
	(0.043)	(0.062)	(0.135)	(0.124)
Amortisation · Debt	0.00189**	0.00181***	0.00189**	0.00181*
	(0.0008)	(0.0007)	(0.0008)	(0.0009)
Deficit · Debt	0.00153*	0.0015*	0.00153	0.0015
	(0.0009)	(0.0008)	(0.0024)	(0.0024)
Interest · Debt	0.006	0.006**	0.006	0.006
	(0.007)	(0.003)	(0.006)	(0.006)
Debt	-0.004	0.001	-0.004	0.001
	(0.014)	(0.013)	(0.022)	(0.025)
Change in Debt	0.018	0.021	0.018	0.021
	(0.021)	(0.020)	(0.021)	(0.027)
Real GDP Growth	-0.221***	-0.241***	-0.221***	-0.241***
	(0.057)	(0.033)	(0.026)	(0.034)
World GDP Growth	0.574***	0.670	0.574***	0.397***
	(0.093)	(1.708)	(0.073)	(0.078)
US 10 year yield	-0.069	0.162	-0.069	0.175
	(0.103)	(0.382)	(0.164)	(0.213)
VIX	0.127***	0.210**	0.127***	0.032
	(0.019)	(0.098)	(0.013)	(0.034)
Constant	-1.775***	-3.752	-1.775*	-
	(0.588)	(3.152)	(0.923)	-
Observations	357	357	357	357
R-squared	0.507	0.548	0.507	0.548
Number of countries	23	23	23	23
Controls	YES	YES	YES	YES
Country FE	YES	YES	YES	YES
Year FE	NO	YES	NO	YES

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Amortisation stands for debt amortisation. Interest stands for interest payments. Deficit stands for primary deficit. Amortisation, Interest, Deficit and Debt are all in percent of GDP. Change in Debt is the difference betw een Debt in t and Debt in (t-1), divided by GDP in (t-1). Columns 1 and 2 are estimated w ith standard panel regression techniques. Columns 3 and 4 are estimated using Driscoll-Kraay estimator.

4.8.1 Other results: Threshold effects

Using the previous estimates, we can identify a threshold for debt above which increases of GFN add to a country's perceived solvency risk.[78] We find that the threshold is around 60% of GDP (using the results from Table 3, column 2). In this section, we use this threshold to perform another experiment. Specifically, we study the effect of GFN on spreads in two sub-samples: when debt is above 60% of GDP and when it is below. The results of this alternative model are presented in Table 5. Additionally, Figure 9 gives a visual sense of the economic significance of these results.

In Table 5 we show the results obtained by using our *favourite* specification in column 1. We observe that if the debt stock is above 60% of GDP, sovereign spreads may increase by six basis points for every 1% increase in GFN. If debt is below that threshold, the results indicate a change in spread that ranges from between zero to 4 basis points for every 1% increase in GFN. The negative marginal effect is, however, not significant. These results also hold when we add time-fixed effects (column 2), whereas, when we correct for the presence of cross-sectional correlation (column 3), despite maintaining the same magnitudes and signs, the coefficients on GFN lose significance.

[78] Using the signalling approach to identify the debt threshold which best splits stress events and normal periods (with stress events defined in the same way we define them in the paper) we get a threshold for debt of 69.4% of GDP. For robustness, we have implemented the "Threshold model" using the 69.4% threshold and the results were qualitatively similar, and in line with the results from our continuous model, with a higher marginal effect of GFN on spread when debt is above the threshold and with the same marginal effect when debt is below the threshold being non-significant.

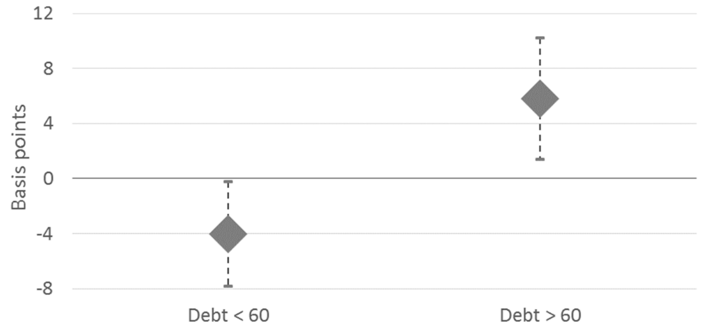
Table 5: Threshold Model – Spreads and Flows when debt is above or below 60%

Dependent variable: 10y Spread	1	2	3
GFN · High Debt Dummy	0.058**	0.052*	0.052
	(0.027)	(0.028)	(0.046)
GFN · Low Debt Dummy	-0.040*	-0.036	-0.036
	(0.023)	(0.025)	(0.029)
High Debt Dummy	-1.805***	-1.791***	-1.791
	(0.564)	(0.584)	(1.305)
Debt	0.037***	0.039***	0.039***
	(0.007)	(0.008)	(0.013)
Change in Debt	0.024	0.026	0.026
	(0.020)	(0.021)	(0.026)
Real GDP Growth	-0.221***	-0.232***	-0.232***
	(0.031)	(0.034)	(0.046)
World GDP growth	0.576***	0.649	0.324***
	(0.070)	(1.792)	(0.049)
US 10 year yield	-0.128*	0.019	-0.140
	(0.071)	(0.365)	(0.087)
VIX	0.127***	0.160*	-0.008
	(0.015)	(0.095)	(0.022)
Constant	-3.668***	-4.871	-
	(0.639)	(3.327)	-
Observations	357	357	357
R-squared	0.430	0.475	-
Number of countries	23	23	23
Country FE	YES	YES	YES
Year FE	NO	YES	YES

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1
GFN and Debt are in percent of GDP. Change in Debt is the difference between Debt in t and Debt in (t-1), divided by GDP in (t-1). Columns 1 and 2 are estimated with standard panel regression techniques. Column 3 is estimated using Driscoll-Kraay estimator. High Debt Dummy equals 1 when the Debt to GDP ratio is larger than 60%. Low Debt Dummy equals 1 when the Debt to GDP ratio is smaller than 60%.

Figure 9, shows the marginal effect of GFN on the sovereign spreads, when the debt-to-GDP ratio is below and above 60%, computed from the specification in column 2 of Table 3. The graph presents 10% error bands.

Figure 9: Marginal effects: gross financing needs and spreads when debt is high/low



Source: ECB and authors computations. Debt is in percent of GDP. The graph presents 10% error bands.

Legend: Figure 9, shows the computed marginal effect of gross financing needs on spreads using estimates from Table 5, column 1. The two points in the figure report how the spreads would change if gross financing needs change by 1% of GDP, when the debt-to-GDP ratio is below and above 60%.