EUROPEAN CENTRAL BANK

EUROSYSTEM

WORKING PAPER SERIES

NO 1604 / NOVEMBER 2013

SETTING COUNTERCYCLICAL CAPITAL BUFFERS BASED ON EARLY WARNING MODELS

WOULD IT WORK?

Markus Behn, Carsten Detken,
Tuomas A. Peltonen and Willem Schudel

In 2013 all ECB publications feature a motif taken from the €5 banknote.

MACROPRUDENTIAL RESEARCH NETWORK

Macroprudential Research Network

This paper presents research conducted within the Macroprudential Research Network (MaRs). The network is composed of economists from the European System of Central Banks (ESCB), i.e. the national central banks of the 27 European Union (EU) Member States and the European Central Bank. The objective of MaRs is to develop core conceptual frameworks, models and/or tools supporting macro-prudential supervision in the EU.

The research is carried out in three work streams: 1) Macro-financial models linking financial stability and the performance of the economy; 2) Early warning systems and systemic risk indicators; 3) Assessing contagion risks.

MaRs is chaired by Philipp Hartmann (ECB). Paolo Angelini (Banca d'Italia), Laurent Clerc (Banque de France), Carsten Detken (ECB), Simone Manganelli (ECB) and Katerina Šmídková (Czech National Bank) are workstream coordinators. Javier Suarez (Center for Monetary and Financial Studies) and Hans Degryse (Katholieke Universiteit Leuven and Tilburg University) act as external consultants. Fiorella De Fiore (ECB) and Kalin Nikolov (ECB) share responsibility for the MaRs Secretariat.

The refereeing process of this paper has been coordinated by a team composed of Gerhard Rünstler, Kalin Nikolov and Bernd Schwaab (all ECB).

The paper is released in order to make the research of MaRs generally available, in preliminary form, to encourage comments and suggestions prior to final publication. The views expressed in the paper are the ones of the author(s) and do not necessarily reflect those of the ECB or of the ESCB.

Acknowledgements

We would like to thank Philipp Hartmann, members of the ESRB Expert Group on Countercyclical Capital Buffers as well as participants of the Banque de France Workshop on Countercyclical Capital Buffers on 7 June 2013 in Paris and the 12th International Conference on Credit Risk Evaluation on 26-27 September 2013 in Venice for valuable comments and useful discussions. This paper is related to the comprehensive analysis being conducted by the ESRB Expert Group on Countercyclical Capital Buffers. The views expressed do not necessarily reflect those of the European Central Bank, the Eurosystem, Deutsche Bundesbank, or the European Systemic Risk Board.

Markus Behn

Deutsche Bundesbank; e-mail: markus.behn@bundesbank.de

Carsten Detken

European Central Bank; e-mail: carsten.detken@ecb.europa.eu

Tuomas A. Peltonen

European Central Bank; e-mail: tuomas.peltonen@ecb.europa.eu

Willem Schudel

European Central Bank; e-mail: willem.schudel@ecb.europa.eu

Address	Kaiserstrasse 29, 60311 Frankfurt am Main, Germany
Postal address	Postfach 16 03 19, 60066 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Internet	http://www.ecb.europa.eu
Fax	+49 69 1344 6000

ISSN	1725-2806 (online)
EU Catalogue No	QB-AR-13-101-EN-N (online)

ABSTRACT

This paper assesses the usefulness of private credit variables and other macro-financial and banking sector indicators for the setting of Basel III / CRD IV countercyclical capital buffers (CCBs) in a multivariate early warning model framework, using data for 23 EU Members States from 1982Q2 to 2012Q3. We find that in addition to credit variables, other domestic and global financial factors such as equity and house prices as well as banking sector variables help to predict vulnerable states of the economy in EU Member States. We therefore suggest that policy makers take a broad approach in their analytical models supporting CCB policy measures.
Keywords: Basel III; CRD IV; countercyclical capital buffer; financial regulation; banking crises; early warning model
JEL Classification: G01, G21, G28

Non-technical summary

The countercyclical capital buffer (CCB), a policy instrument proposed by the Basel III and the EU Capital Requirements Directive (CRD IV) as a response to the recent financial crisis, aims at increasing the resilience of the banking system in times of financial crisis by ensuring that banks set aside capital when credit growth is strong and allowing them to use this capital in periods of financial stress. Motivated by the discussion in the literature on the appropriate methodology for setting the CCB, this paper assesses credit and various other macro-financial variables in predicting banking system vulnerabilities preceding systemic banking crises on a sample of 23 EU Member States across a period from 1982Q2 to 2012Q3.

This paper contributes to the literature in the following ways. First, to our knowledge, it is the first paper to apply the most recent early warning modelling techniques for the use of setting the CCB, calibrating models to predict a vulnerable state of the financial system which could lead to a systemic banking crisis. Second, the paper applies the latest (and most complete) database on credit to the private sector for EU countries and evaluates the importance of this variable, which plays a central role in the CCB legislation, in both uni- and multivariate settings predicting financial vulnerabilities.

The paper finds that the credit-to-GDP gap (or the deviation of the credit-to-GDP ratio from its long-term trend) is the best early warning indicator among domestic credit variables in terms of prediction ability. Interestingly, however, global credit variables seem to outperform domestic credit variables. In particular, global credit growth and the global credit-to-GDP gap provide strong early warning signals both as single variables in a non-parametric signalling approach and through consistent and significant effects in multivariate logit models. It is important to keep in mind, however, that the strong predictive abilities of the global variables are subject to a caveat related to the evaluation period that includes the global financial crisis, where there is a strong clustering of crisis episodes across countries. In addition to the salience of private credit variables, the paper finds that other macro-financial variables such as domestic house price growth and global equity price growth can

be positively associated with future banking crises. Moreover, the paper presents evidence that a high level of banking sector capitalisation decreases the probability of entering a state of financial vulnerability. Thus, our results suggest that even though private credit variables are very important in predicting future financial instability, there is good reason for policy makers and researchers to take a broad empirical approach by incorporating also other macro-financial variables in such analyses. In this respect, multivariate models, such as the one introduced in this study, are found to be more useful as they include the combined information of the joint behaviour of several indicators. Finally, as a validation of our analysis, we find a good out-of-sample performance of the models in predicting the vulnerable states preceding the financial crises in Finland and Sweden in the early 1990s as well as those in Italy and the U.K. in the mid-1990s.

1. Introduction

Being faced with the longest and most severe financial crisis in decades, policy makers around the globe have actively searched for policy tools which could help to prevent or at least reduce the intensity of future financial crises. A tool that is an integral part of the Basel III regulations and the EU Capital Requirements Directive (CRD IV) is the counter-cyclical capital buffer (CCB), which has been proposed by the Basel Committee on Banking Supervision (BCBS) at the Bank for International Settlements (BIS).

The CCB aims to increase the resilience of the banking system in case of a financial crisis by ensuring that banks set aside capital in times of *"aggregate growth in credit [...] associated with a build-up of systemic risk"*, which can be *"drawn down during stressed periods."* (EU [20]). In order to promote international consistency in setting CCB rates, the BCBS has developed a methodology based on the ratio of aggregate credit to GDP (Basel Committee on Banking Supervision [5]). The CRD IV, while acknowledging the importance of credit growth and the credit-to-GDP ratio, specifies that buffer rates should also account for *"other variables relevant to the risks to financial stability."* (EU [20]).[1] This provides the motivation for this paper: We assess the usefulness of credit and other macro-financial variables for the prediction of banking sector vulnerabilities in a multivariate framework, hence enabling a more informed decision on the setting of CCB rates.

The BCBS guidelines are based on an analysis that uses a sample of 26 countries from all over the world, for which the credit-to-GDP gap (defined as the deviation of the credit-to-GDP ratio from its long-term trend) performs as the best single indicator in terms of signalling a coming financial crisis. However, from the evidence presented by the BCBS it is not clear whether the credit-to-GDP gap provides a warning signal that is early enough to account for the 12 month implementation period for raising the capital buffers specified in

[1] In particular, the CRD IV specifies that the deviation of the credit-to-GDP ratio from its long-term trend should serve as *"a common starting point for decisions on buffer rates by the relevant national authorities, but should not give rise to an automatic buffer setting or bind the designated authority. The buffer shall reflect, in a meaningful way, the credit cycle and the risks due to excess credit growth in the Member State and shall duly take into account specificities of the national economy."* (EU [20]).

the CRD IV regulation.[2] In other words, the credit gap may be an early warning indicator that is not early enough for policy implementation purposes.[3] Moreover, the guidelines (or the work by Drehmann et al. [17]) do not directly compare the predictive power of the credit-to-GDP gap to that of other potentially relevant variables related to risks to financial stability (as stated in the CRD IV) in a *multivariate framework*. Acknowledging the potentially very large implications that this policy has for the international banking sector, our paper aims to address these non-trivial omissions.

The main findings of the paper are the following: First, we find that global variables and especially global credit variables are strong predictors of macro-financial vulnerability, providing good signals when used as single variables and demonstrating consistent and significant effects in multivariate logit models. Domestic credit-to-GDP also affects the probability of being in a vulnerable state, even though the effect is clearly smaller than that of global credit variables. However, despite the importance of credit variables, we also find evidence suggesting that other variables play a salient role in predicting vulnerable states of the economy.[4] For example, domestic house price growth and global equity growth are positively associated with macro-financial vulnerabilities. Moreover, we find that banking sector variables exert significant effects: Strong banking sector profitability may incur excessive risk-taking, leading to increased vulnerability, while a level of high banking sector capitalisation decreases the probability of entering a vulnerable state. This result is potentially important for policy makers involved in setting the CCB, as it reinforces the notion that

[2] According to Article 126(6) of the CRD IV, *"when a designated authority sets the countercyclical buffer rate above zero for the first time, or when thereafter a designated authority increases the prevailing countercyclical buffer rate setting, it shall also decide the date from which the institutions must apply that increased buffer for the purposes of calculating their institution specific countercyclical capital buffer. That date may be no later than 12 months after the date when the increased buffer setting is announced [in accordance with paragraph 8]. If the date is less than 12 months after the increased buffer setting is announced, that shorter deadline for application shall be justified by exceptional circumstances"*.

[3] Several other potential shortcomings of the credit-to-GDP gap have been discussed in the literature. For example, Edge and Meisenzahl [19] argue that gap measures are sensitive to the exact specification of the trending variable, in particular with regards to end-of-sample estimates of the credit-to-GDP ratio. For other critical views on the reliability or suitability of the credit-to-GDP gap in the context of the CCB, see for example Repullo and Saurina [37] and Seidler and Gersl [41].

[4] See also Drehmann and Juselius [18], who find that the debt service ratio performs well as a supplementary early warning indicator to credit variables for horizons up to two years prior to banking crises. The authors also find that the debt service ratio prior to economic slumps is related to the size of subsequent output losses. Moreover, Hahm et al. [22] find that growth of banks' non-core liabilities is an indicator for a lending boom and a source of vulnerability to a crisis.

higher CCB rates and bank capital ratios overall reduce financial vulnerability. As such, our findings suggest that even though credit variables are near-essential in early warning models, other macro-financial and banking sector variables are important covariates to control for and improve the predictive power of these models. In this respect, multivariate models, such as the one introduced in this study, are found to be more useful as they include the combined information of the joint behaviour of several indicators. Finally, as a validation of our analysis, we find a good out-of-sample performance (see Berg et al. [6]) of the models in predicting the vulnerable states preceding the financial crises in Finland and Sweden in the early 1990s as well as those in Italy and the U.K. in the mid-1990s.

This paper contributes to the literature in the following ways: First, we apply state-of-the-art modelling techniques from the early warning system (EWS) literature to see whether they could be useful for decisions on countercyclical capital buffers in EU countries. In line with the forthcoming legislation for the CCB, the models are calibrated so that they aim to predict a vulnerable state of the economy (or banking sector), i.e., a build-up of system-wide risk that, with a suitable trigger, could turn into a banking crisis. In practice, we analyse the out-of-sample predictive abilities of a variety of models for those states of the economy that have preceded earlier banking crises by twelve to seven quarters. This would, hopefully, allow a timely build-up of the CCB. Following the methodological approach of Frankel and Rose [21] and Demirgüç-Kunt and Detragiache [13], the analysis is conducted in a multivariate logit model framework using data for 23 EU Member States spanning over 1982Q2-2012Q3, where we complement the credit variables with several domestic macro-financial and banking sector variables, and following e.g. Frankel and Rose [21] and Lo Duca and Peltonen [32], also include global variables in our models in order to account for potential spillover effects. In a similar fashion as Alessi and Detken [2], Lo Duca and Peltonen [32] and Sarlin [38], the models are evaluated using a framework that takes into account a policy maker's preferences between type I (missing a crisis) and type II errors (false alarms of crises). Moreover, the paper focuses exclusively on EU countries, including the largest possible sample (limited by data availability) instead of focusing on a few large economies as is common in the literature.

Second, given the importance of the credit variables in the CRD IV regulatory framework, we use the same BIS database on credit as the BCBS and evaluate their salience in predicting vulnerable states of the economy, both in univariate and multivariate frameworks. Hence, we build on the work of Drehmann et al. [17], who use a univariate signal extraction methodology (see also Kaminsky et al. [25]) to find that the credit-to-GDP gap provides the best early warning signals for the build-up of capital buffers. Finally, we employ different definitions of banking crises (Babecky et al. [3]; Laeven and Valencia [28, 30]; Reinhart and Rogoff [35, 36]) as well as several other variations of the analysis to assess the robustness of the main results.

The remainder of the paper is organised as follows: We present our data set in Section 2 and introduce the methodology in Section 3. Estimation results and robustness analysis are presented in Section 4, while Section 5 is reserved for our concluding remarks.

2. Data

This section introduces the data used for our study. We begin with the identification of vulnerable states, i.e., the dependent variable in the study, based on banking crises in the European Union. We then proceed by introducing the independent variables used in the empirical analysis. Finally, we present some descriptive statistics on the development of key variables around banking sector crises in the sample countries.

2.1. Definition of vulnerable states

The paper develops an early warning model that attempts to predict vulnerable states of the economy from which—given a suitable trigger—banking crises could emerge. Thus, we are not trying to predict banking crises per se, even though we need to identify these crises in order to determine the vulnerable states. Specifically, we define a vulnerable state as the period twelve to seven quarters before the onset of a banking crisis. The time horizon accounts for the CCB announcement period of twelve months that is specified in the CRD

IV (EU [20], Art. 126(6)), and for a time lag required to impose such a policy. At the same time, extending the horizon too far into the past may weaken the link between observed variation in the independent variables and the onset of banking crises. To analyse this, we provide a number of alternative time horizons in the robustness section.

In order to identify banking crises, we use the dataset which has been compiled by Babecky et al. [3] as part of a data collection exercise by the European System of Central Banks (ESCB) Heads of Research Group (labeled as HoR database hereafter). This quarterly database contains information on banking crises in EU countries between 1970Q1 and 2012Q4.[5] The crisis index takes a value of 1 when a banking crisis occurred in a given quarter (and a value of 0 when no crisis occurred). The HoR database aggregates information on banking crises from "several influential papers", including (in alphabetical order): Caprio and Klingebiel [10]; Detragiache and Spilimbergo [15]; Kaminsky [26]; Kaminsky and Reinhart [27]; Laeven and Valencia [28, 29, 30]; Reinhart and Rogoff [35, 36]; and Yeyati and Panizza [42]. The crisis indices from these papers have subsequently been cross-checked with the ESCB Heads of Research before inclusion into the database. A list of the banking crisis dates for our sample countries based on this dataset is provided in Panel A of Table 1. In the robustness section, we test the robustness of the results by regressing the benchmark model on banking crisis data provided by Laeven and Valencia [30] and Reinhart and Rogoff [36].

We set the dependent variable to 1 between (and including) twelve to seven quarters prior to a banking crisis as identified by the ESCB HoR database and to 0 for all other quarters in the data. In order to overcome crisis and post-crisis bias (see e.g. Bussière and Fratzscher [8]), we omit all country quarters which either witnessed a banking crisis or which fall within six quarters after a banking crisis.

[5] Croatia, which joined the EU on 1 July 2013, has not yet been included in the database.

2.2. Macro-financial and banking sector variables

The panel dataset used in the analysis contains quarterly macro-financial and banking sector data spanning over 1982Q2-2012Q3 for 23 EU member states. The data is sourced through Haver Analytics and originally comes from the BIS, Eurostat, IMF, ECB, and OECD.[6] Panel A of Table 1 provides an overview of the data availability for our main variables, while Panel B summarises the variables included in our study.

Following Drehmann et al. [17], we first include variables measuring the supply of credit to the private sector. We use the "long series on total credit and domestic bank credit to the private non-financial sector" compiled by the BIS. This data includes "borrowing from non-financial corporations, households and non-profit institutions serving households. [...] In terms of lenders, the new total credit series aim to capture all sources independent of the country of origin or type of lender [...] [while] the coverage of financial instruments includes loans and debt securities such as bonds and securitised loans." (see Dembiermont et al. [12] for a description of the database). To our knowledge, the BIS credit series offers the broadest definition of credit provision to the private sector, while having been adjusted for data gaps and structural breaks.

We include four different measurements of credit in our models, accounting for credit growth and leverage, both at the domestic and at the global level. Credit growth is entered as a percentage (annual growth), while leverage is measured by the deviation of the credit-to-GDP ratio (using nominal GDP data) from its long-term backward-looking trend (using a backward-looking Hodrick-Prescott filter with a smoothing parameter λ of 400,000) as proposed by the Basel Committee on Banking Supervision [5] Consultative Document.[7]

[6] In particular, the individual series stem from the following original sources: Data on total credit to the private non-financial sector is obtained from the BIS and—for those countries where BIS data is not available—from Eurostat. Information on nominal GDP growth and inflation rates comes from the IMF's International Financial Statistics (IFS). Data on stock prices is obtained from the OECD, while data on house prices is provided by the BIS. Banking sector variables are obtained from two sources: The OECD provides relatively long series on banking sector capitalisation and profitability on an annual basis that we use in the empirical analysis. Additionally, for illustrative purposes, we use a shorter series of banking sector capitalisation in Figure 3 that is available on a quarterly basis and which is obtained from the ECB's Balance Sheet Items (BSI) statistics. Finally, quarterly data on the 10-year government bond yield and the 3-months interbank lending rate (money market rate) are obtained from the OECD.

[7] Recommendations in the BCBS Consultive Document are based on a paper by Borio et al. [7], who find

Global credit variables have been computed using a GDP-weighted average of the variable in question for several countries, including the United States, Japan, Canada, and all European countries which are in this study (see also Alessi and Detken [2]). In addition, we include four sets of interaction terms in the same fashion as Lo Duca and Peltonen [32], namely the product of the domestic variables, the product of the global variables and that between the domestic and the global credit variables. The results using different variations of the credit variables are discussed in Sections 4.1 and 4.2.[8]

In order to test the importance of credit variables in a comparative fashion as well as to analyse the potential importance of other factors, we include a number of additional variables in our study. These variables are available for fewer observations than the credit variables, which is why the number of observations in the full model differs from the number of observations in models that include only credit variables.[9] Variables are selected based on the existing literature and on data availability. In order to account for the macroeconomic environment and monetary stance, we include nominal GDP growth (domestic and global) and CPI inflation rates. Furthermore, following Reinhart and Rogoff [35], we include data on equity and residential house prices, both domestically and globally (using the same methodology to calculate the global variables as in the case of the credit variables), focusing on annual growth rates. Finally, to control for banking sector profitability and solvency, we include aggregate bank capitalisation (calculated by the ratio of equity over total assets) and aggregate banking sector profitability (defined as net income before tax as a percentage of total assets), which has e.g. been suggested by Barrell et al. [4].

As we are estimating binary choice models using panel data, non-stationarity of inde-

that trends calculated with a λ of 400,000 perform well in picking up the long-term development of private credit. In particular, a λ of 400,000 is consistent with the assumption of credit cycles being four times longer than business cycles if one follows a rule developed by Ravn and Uhlig [34], which states that the optimal λ of 1,600 for quarterly data should be adjusted by the fourth power of the observation frequency ratio (i.e., if credit cycles are four times longer than business cycles, λ should be equal to $4^4 \times 1,600 \approx 400,000$).

[8] In Section 4.1 we evaluate how individual credit variables perform in the prediction of banking sector vulnerabilities. In this section, we look at several other transformations of the credit variables, including the credit gap (defined as the deviation of private credit from its long-term trend), the credit-to-GDP ratio, several credit growth moving averages, and a variable defined as the difference between credit growth and nominal GDP growth. We evaluate all these variables on the domestic as well as on the global level.

[9] Estimating credit models on the reduced sample yields results that are very similar to the ones for the full sample.

pendent variables could be an issue (Park and Phillips [33]). We perform panel unit root tests suggested by Im et al. [23] as well as univariate unit root tests developed by Dickey and Fuller [16] in order to analyse the time series properties of the variables of interest. In the panel unit root test, the null hypothesis that all cross-sections contain unit roots can be rejected at least at the 10 percent level for all series except for the credit-to-GDP gap and global credit growth.[10] We complement the panel unit root analysis by using the Dickey and Fuller [16] test country-by-country, and can reject the null hypothesis of a unit root for the credit-to-GDP gap at least at the 10 percent level for all countries except for Estonia, Lithuania and Greece. Furthermore, in the country-by-country unit root tests, we can reject the null hypothesis for the global credit-to-GDP gap at least at the 10 percent level for all countries except for Estonia and Lithuania, while for global credit growth the null hypothesis can be rejected for all countries. This implies that sample periods for individual countries seem to affect unit root test results. Overall, the transformations done to the original variables, the results from the unit root tests and general economic theory make us confident that we have addressed potential non-stationarity concerns for the variables of interest.

2.3. Development of the key variables

Before entering the discussion of the main results, we shortly present some descriptive statistics, which provide the context of our main argument of moving beyond credit variables when predicting macro-financial vulnerabilities. Figure 1 presents the average development of the six main variables of interest over time before and after the onset of a banking crisis. For the purpose of predicting banking crises based on simple descriptive statistics, one would hope to find an indicator variable that displays a typical pattern in the run up to a crisis so that it can be used as a signal. In the current case of predicting vulnerable states of the economy, which precede future banking crises, one would be interested in variables that signal way ahead of the onset of a crisis (i.e. two to three years before the crisis), so that

[10] Alternatively, we perform a Fisher-type test by running a Dickey and Fuller [16] test by cross-section and then combining the p-values from these tests to produce an overall test statistic. The null hypothesis that all cross-sections contain unit roots can be rejected at least at the 10 percent level for all series except for the credit-to-GDP gap and the global credit-to-GDP gap.

policy makers can use this time to increase the resilience of banks.

In this context, we observe that among the six variables depicted here, the credit-to-GDP gap shows one of the least clear pictures in terms of signalling a coming crisis. On average, the credit gap increases slowly prior to a banking crisis and starts falling about one year into the crisis. Yet, this does not need to be a very surprising development, as this variable is a ratio and therefore requires the numerator to grow more slowly (or decrease faster) than the denominator in order for the variable to decrease in value. The BCBS itself concedes that the credit-to-GDP trend may not capture turning points well (Basel Committee on Banking Supervision [5]). Consequently, the ratio will not fall unless credit falls faster than GDP, something which is not at all certain during banking crises. Still, it shows that from a purely descriptive perspective, any signal to be derived from the credit gap is expected to come from the level of this variable breaching a threshold value, rather than from turning points in its development.

Unlike the credit gap, credit growth (as depicted in % year-on-year growth) does appear to hit a peak about two years before the onset of a banking crisis, even though its fall only becomes clear during the last pre-crisis year. A similar development can be observed in nominal GDP growth and equity price growth figures. These variables do (on average) peak already before the start of a crisis. Moreover, in our sample, the growth rate of residential house prices tends to peak about 3 years before a crisis happens on average, starting a clear descent (although prices are still rising) that lasts into the crisis where growth stalls. Based on this evidence, we would conclude that it is clear that it makes sense to gauge the developments of different macro-financial variables to predict or signal coming crises. Whether this result holds in a more rigorous comparative (multivariate) framework, will be discussed in the subsequent analysis.

3. Methodology

In this section we introduce the methodology used in the empirical analysis. We start by introducing the logistic regressions used in our multivariate framework. Thereafter, we explain how we evaluate individual indicators' and model predictions' usefulness for policy makers.

3.1. Multivariate models

In order to assess the predictive abilities of credit, macro-financial and banking sector variables in a multivariate framework, we estimate logistic regressions of the following form:

$$Prob(y_{it} = 1) = \frac{e^{\alpha_i + X_{it}'\beta}}{1 + e^{\alpha_i + X_{it}'\beta}} \qquad (1)$$

where $Prob(y_{it} = 1)$ denotes the probability that country i is in a vulnerable state, where a banking crisis could occur seven to twelve quarters ahead of quarter t. As described in Section 2.1, we set the dependent variable to 1 twelve to seven quarters before the onset of a banking crisis in the respective country and to 0 otherwise. As independent variables, the vector X_{it} includes credit and macro-financial variables on the domestic and on the global level as well as domestic banking sector variables (see Section 2.2 for a precise definition of the variables). The estimations also include a set of country dummy variables α_i in order to account for unobserved heterogeneity at the country level (country fixed effects).[11,12] Finally, we use robust standard errors clustered at the quarterly level in order to account for potential correlation in the error terms that might arise from the fact that global variables are

[11] There is an argument for omitting these dummies from the estimations as they automatically exclude all countries without a crisis from the estimation, hence introducing selection bias (see e.g. Demirgüç-Kunt and Detragiache [13] and Davis and Karim [11]). However, not including them also induces bias, namely omitted variable bias caused by unit effects. As it is unlikely that financial crises are homogeneously caused by identical factors (see also Candelon et al. [9]) and as a Hausman test indicates unit heterogeneity, we have decided to include unit dummy variables in our estimations. However, results for pooled models (without country dummies), where coefficients in these models are of course different, but they nevertheless carry the same sign in virtually all models that we have estimated, are available upon request.

[12] In principle, we could have included time dummies in addition to country dummies in order to account for heterogeneity in crisis probabilities over time. However, we decided against the inclusion of these dummies for two reasons: First, only quarters where at least one country experiences a banking crisis could be used for identification in such a specification. As our sample includes many quarters where none of the countries experienced a crisis the inclusion of time dummies would significantly reduce our sample size. Second, the focus in our paper is on the prediction of future banking crises. Specifically, we aim to develop an early warning model that policy makers can use for the detection of vulnerabilities in the banking sector. While time dummies might improve the ex post fit of a model, they are of little use for out-of-sample forecasting since they are not known ex ante (see e.g. Schularick and Taylor [40]).

identical across countries in a given quarter.[13]

The analysis is conducted as much as possible in a real-time fashion, meaning that only information that is available at a particular point in time is used. As such, all de-trended variables have been calculated using backward trends, thereby only using information that was available up to that point. Furthermore, the explanatory variables have been lagged by one quarter, also to account for endogeneity bias through simultaneity. We are well aware that this simple procedure cannot crowd out all endogeneity-related bias, but we note that the dependent variable itself is an early warning variable. The time horizon for which this variable is equal to 1 has been chosen in the context of our exercise and has not been exogenously determined. Therefore, we consider endogeneity to be a somewhat smaller problem in this study. Nevertheless, we have tested our models for different specifications of the dependent variable, both in terms of the pre-crisis period chosen (12-1/20-13 quarters before the onset of a crisis) and the definition and data source of banking crises in the robustness section.

3.2. Model evaluation

Banking crises are (thankfully) rare events in the sense that most EU countries have encountered none or only one over the past two decades. Still, when they occur, banking crises tend to be very costly, both directly through bailouts and fiscal interventions and indirectly through the loss of economic output that oftentimes (particularly in systemic banking crises) tends to follow these crises. Thus, policy makers have a clear incentive to be able to detect early enough potential signs of vulnerabilities that might precede banking crises in order to take measures to prevent further building up of vulnerabilities or to strengthen the resilience of the banking sector. Yet, at the same time, policy makers may not want to be signalling crises when in fact they do not happen afterwards. Doing so may (a) reduce the credibility of their signals, weakening decision-making and damaging their reputation, and (b) needlessly incur costs on the banking sector, endangering credit supply. As a consequence, policy mak-

[13] Alternatively, we cluster standard errors at the country level, which results in smaller estimates in particular for the global variables.

ers also have an incentive to avoid false alarms, i.e., they do not want to issue warnings when a crisis is not imminent. As pointed out by Alessi and Detken [2], an evaluation framework for an early warning model needs to take into account policy makers' relative aversion with respect to type I errors (not issuing a signal when a crisis is imminent) and type II errors (issuing a signal when no crisis is imminent).

The evaluation approach in this paper is based on the so-called 'signalling approach' that was originally developed by Kaminsky et al. [25], and extended by Demirgüç-Kunt and Detragiache [14], Alessi and Detken [2], Lo Duca and Peltonen [32] and Sarlin [38]. In this framework, an indicator issues a warning signal whenever its value in a certain period exceeds a threshold τ, defined by a percentile of the indicator's country-specific distribution. Similarly, a multivariate probability model issues a warning signal whenever the predicted probability from this model exceeds a threshold $\tau \in [0,1]$, again defined as a percentile of the country-specific distribution of predicted probabilities. In this way, individual variables and model predictions for each observation j are transformed into binary predictions P_j that are equal to 1 if the respective thresholds are exceeded for this observation and 0 otherwise. Predictive abilities of the variables and the models can then be evaluated by comparing the signals issued by the respective variable or model to the actual outcome C_j for each observation.[14] Each observation can be allocated to one of the quadrants in the contingency matrix depicted in Table 2: A period with a signal by a specific indicator can either be followed by a banking crisis twelve to seven quarters ahead (TP) or not (FP). Similarly, a period without a signal can be followed by a banking crisis twelve to seven quarters ahead (FN) or not (TN). Importantly, the number of observations classified into each category depends on the threshold τ.

In order to obtain the optimal threshold τ one needs to take the policy maker's preferences vis-à-vis type I errors (missing a crisis, $T_1(\tau) = FN/(TP+FN) \in [0,1]$) and type II errors (issuing a false alarm, $T_2(\tau) = FP/(FP+TN) \in [0,1]$) into account. This can be done by defining a loss function that depends on the two types of errors as well as the policy maker's relative preference for either type. The optimal threshold is then the one that mini-

[14] C_j is equal to 1 if the country experiences a banking sector crisis twelve to seven quarters ahead of the respective period and 0 otherwise.

mizes the loss function. Taking into account the relative frequencies of crises $P_1 = P(C_j = 1)$ and tranquil periods $P_2 = P(C_j = 0)$, the loss function is defined as follows:[15]

$$L(\mu, \tau) = \mu P_1 T_1(\tau) + (1-\mu) P_2 T_2(\tau), \tag{2}$$

where $\mu \in [0,1]$ denotes the policy makers' relative preference between type I and type II errors. A μ larger than 0.5 indicates that the policy maker is more averse against missing a crisis than against issuing a false alarm, which—in particular following the recent financial crisis—is a realistic assumption in our view.

Using the loss function $L(\mu, \tau)$, the usefulness of a model can be defined in two ways. First, following the idea of Alessi and Detken [2] and as in Sarlin [38], the absolute usefulness is defined as:

$$U_a = min(\mu P_1, (1-\mu) P_2) - L(\mu, \tau). \tag{3}$$

Note that U_a computes the extent to which having the model is better than having no model. This is because a policy maker can always achieve a loss of $min(\mu P_1, (1-\mu) P_2)$ by either always issuing a signal (in which case $T_1(\tau) = 0$) or never issuing a signal (in which case $T_2(\tau) = 0$). The fact that P_1 is significantly smaller than P_2 in our sample (i.e., there are relatively few vulnerable states preceding banking crises) implies that, in order to achieve a high usefulness of the model, a policy maker needs to be more concerned about the detection of vulnerable states potentially preceding banking crises in comparison to the avoidance of false alarms.[16] Otherwise, with a suboptimal performing model, it would easily pay off for the policy maker to never issue a signal given the distribution of vulnerable states and tranquil periods (see Sarlin [38] for a detailed discussion of this issue).

A second measure, the relative usefulness U_r, is computed as follows (see Sarlin [38]):

$$U_r = \frac{U_a}{min(\mu P_1, (1-\mu) P_2)} \tag{4}$$

[15] As pointed out by Sarlin [38], policy makers should be concerned about the absolute number of misclassification rather than the share of misclassifications in relation to class size (i.e., unweighted type I and type II errors). Therefore, a failure to account for the relative frequency of crisis episodes and tranquil periods—as in previous studies—results in a bias on the weighting of type I and type II errors in the loss function.

[16] The share of observations that is followed by a banking crisis twelve to seven quarters ahead—P_1—is approximately equal to 10 % in our sample.

The relative usefulness U_r reports U_a as a percentage of the usefulness that a policy maker would gain from a perfectly performing model.[17] The relative usefulness is our preferred performance indicator as it allows the comparison of models for policy makers with different values for the preference parameter μ.

In addition to assessing the relative and absolute usefulness of a model, we also employ receiver operating characteristics (ROC) curves and the area under the ROC curve (AUROC) as these are also viable measures for comparing performance of early warning models. The ROC curve shows the trade-off between the benefits and costs of a certain threshold τ. When two models are compared, the better model has a higher benefit (TP rate (TPR) on the vertical axis) at the same cost (FP rate (FPR) on the horizontal axis).[18] Thus, as each FP rate is associated with a threshold, the measure shows performance over all thresholds.[19] In this paper, the size of the AUROC is computed using trapezoidal approximations. The AUROC measures the probability that a randomly chosen vulnerable state is ranked higher than a tranquil period. A perfect ranking has an AUROC equal to 1, whereas a coin toss has an expected AUROC of 0.5.

4. Empirical results

In this section we present the empirical results. We first explore the usefulness of credit variables for the identification of vulnerable states of the banking sector, and proceed by extending the framework to a multivariate model including other macro-financial and banking

[17] A perfectly performing indicator would achieve $T_1 = T_2 = 0$, implying $L = 0$. Consequently, U_a would reduce to $min(\mu P_1, (1-\mu)P_2)$.

[18] The TPR (also called sensitivity) gives the ratio of periods where the model correctly issues a warning to all periods where a warning should have been issued, formally $TPR = TP/(TP+FN)$. The FPR (also called specificity) gives the ratio of periods where the model wrongly issues a signal to all periods where no signal should have been issued, formally $FPR = FP/(FP+TN)$. An ideal model would achieve a TPR of one (no missed crises) and a FPR of zero (no false alarms).

[19] The measure can also be interpreted as showing the performance over all preference parameters μ of the policy maker: The lower the threshold τ, the more aggressive is the policy maker in making crisis calls as almost all signals are above the threshold. Hence, a low τ corresponds to a policy maker with a strong aversion against type I errors, i.e., a policy maker with a strong preference for correctly calling all crises. Equivalently, the larger the threshold τ the more conservative is the policy maker in making crisis calls. Therefore, a high τ corresponds to a policy maker with a strong aversion against type II errors, i.e., a policy with a strong preference for the avoidance of false alarms.

sector indicators. Thereafter, we evaluate the out-of-sample performance of the estimated models and—finally—present some robustness checks.

4.1. Estimation and evaluation

As the CRD IV regulations emphasise the role of credit variables for setting the countercyclical capital buffer rate—in particular the role of credit growth and the credit-to-GDP gap—we start by evaluating the usefulness of these variables for the identification of vulnerable states within the EU banking sector.

4.1.1. Individual indicators

First, we evaluate the usefulness of domestic credit variables by using a simple signalling approach. Using a preference parameter of μ equal to 0.9, Panel A of Table 3 reports the optimal threshold for several credit variable indicators.[20] Given the optimal threshold, the table also shows the number of observations in each quadrant of the matrix depicted in Table 2, the percentage of type 1 and type 2 errors, as well as several performance measures, such as the absolute and the relative usefulness, the adjusted noise-to-signal (aNtS) ratio[21], the percentage of vulnerable states correctly predicted by the indicator (% Predicted), the probability of a vulnerable state conditional on a signal being issued (Cond Prob) and the difference between the conditional and the unconditional probability of a vulnerable state (Diff Prob).

Among the domestic indicators, indeed, the credit-to-GDP gap performs best in the

[20] A preference parameter of μ equal to 0.9 indicates a strong preference for the detection of crises by the policy maker. In our view this is a reasonable assumption as the current crisis illustrated once more that financial crises often translate into large costs for the economy. As Sarlin [38] points out, using a μ equal to 0.9 and simultaneously taking into account the unconditional probability of a crisis (which is about 10 % in our sample) is equivalent to using a μ equal to 0.5 without adjusting for the unconditional probabilities (as in Alessi and Detken [2] or Lo Duca and Peltonen [32]). Results for different values of the preference parameter are available upon request.

[21] The aNtS ratio is the ratio of false signals measured as a proportion of quarters where false signals could have been issued to good signals as a proportion of quarters where good signals could have been issued, or $(FP/(FP+TN))/(TP/(TP+FN))$. A lower aNtS ratio indicates better predictive abilities of the model.

sense that it generates the highest relative usefulness.[22] This indicator issues a signal whenever the credit-to-GDP gap is above the 40^{th} percentile of its country-specific distribution and achieves 25.6 % of the usefulness a policy maker would gain from a perfectly performing model. The indicator correctly calls 81.3 % of the vulnerable states and displays an adjusted noise-to-signal ratio of 0.678. Conditional on a signal being issued, the probability of a vulnerable state is 16.8 %, which is 4.7 % higher than the unconditional probability of a vulnerable state in our sample. Other variables that perform relatively well are annual credit growth, the credit-to-GDP ratio and the credit gap (defined as the deviation of the stock of credit from its long term trend, see Section 2.2).

Interestingly, global variables seem to outperform domestic variables in terms of usefulness. Panel B of Table 3 shows that these indicators usually exert a higher relative usefulness, a lower adjusted noise-to-signal ratio, and are able to predict a larger share of the vulnerable states in our sample. This suggests that focusing on the development of domestic credit variables might not be sufficient. In an increasingly integrated economy, vulnerabilities that develop at a global level potentially transmit to countries around the world. Therefore, policy makers should also take these developments into account when deciding on countercyclical buffer rates.[23]

The evaluation of the predictive abilities of global variables is subject to a caveat: As these variables do not vary across countries, and as most countries had a crisis starting in 2008, the good performance of these variables can in part be explained by a clustering of crisis episodes within the same year, i.e., indicators based on global credit variables correctly predicted the current crisis in several of our sample countries. To a certain extent this puts the higher usefulness of global as compared to domestic variables in a perspective. However, the current crisis is certainly one of the best examples for a non-domestic vulnerability that spread to banking systems around the world. Thus, if the aim of the CCB is to increase the resilience of the banking system, it appears to be beneficial to take into account both

[22] This is consistent with findings by Drehmann et al. [17] for a different set of countries and seems to support the approach taken in the CRD IV regulation. However, the main argument of our paper will be that performance of the individual indicators can be improved if they are combined in a multivariate approach. Moreover, also global variables are useful for the identification of vulnerabilities in the banking sector.

[23] In the CRD IV, the CCB rate is calculated based on a weighted average of banks' country exposures.

domestic and global developments.

4.1.2. Multivariate models

While the signalling approach is a simple and useful way to assess the predictive abilities of individual indicators, a multivariate framework has the advantage of being able to assess the joint performance of several indicators. We therefore estimate simple logit models including several of the individual credit variables and assess their performance and usefulness. In order to account for unobserved heterogeneity across countries that might otherwise bias our results, we include a set of country dummies.

Results for these models are presented in Table 4. Again, we start by considering only the domestic variables and focus on credit growth and the credit-to-GDP gap, as these variables performed well in the signalling approach and play a prominent role in the CRD IV regulations. Credit growth seems to dominate the credit-to-GDP gap, which is statistically not significant, in this simple model. Next, we gradually include the global credit variables, interactions between growth and leverage on the domestic and the global level as well as interactions between the domestic and the global variables.[24] The predictive power of the model improves with each step.[25]

In order to compare the models' predictive abilities with those of the individual indicators we once more apply the signalling approach by translating the predicted probabilities into country specific percentiles and determining the optimal threshold for the issuance of warnings as the one that maximizes the relative usefulness of the model (see Section 3.2).

[24] We orthogonalise interaction terms with first-order predictors in order to avoid problems of multicollinearity (see e.g. Little et al. [31]). In particular, when interacting two variables X and Y, we first form the simple product $X \times Y$ and then regress it on the original variables: $X \times Y = \alpha + \beta_1 \times X + \beta_2 \times Y + \varepsilon$. We then take the residual from this regression—ε, which is orthogonal to X and Y—to represent the interaction between the two original variables. Variance inflation factors (VIF) smaller than ten for all variables indicate that we are able to get rid of multicollinearity problems in this way.

[25] Note that the interpretation of interaction effects in logit models is cumbersome. As pointed out by Ai and Norton [1], the interaction effect is conditional on the independent variables (unlike interaction effects in linear models) and may have different signs for different values of the covariates. Moreover, the statistical significance of these effects cannot be evaluated with a simple t-test, but should be evaluated for each observation separately. Doing so allows us to conclude that for most observations only the Interaction(GC1×GC2) is significantly positive, while the other interactions are insignificant (although e.g. the Interaction(DC2×GC2) has a significantly negative sign in the regression itself).

Table 5 shows that the relative usefulness of the domestic model is 0.236, which is lower than the one of the best individual indicators. However, the stepwise inclusion of the remaining variables improves the usefulness, so that Model 3 surpasses the best domestic as well as the best global indicators in terms of relative usefulness. This indicates the benefits of a multivariate framework as compared to single indicators. We will elaborate more on these benefits by taking into account not only credit variables, but also other variables that might affect the stability of the banking sector.

Models 4-7 provide the estimation results for the extended models. The sample size is somewhat smaller than in the Models 1-3, as the data is not available for all variables across the whole period (see Table 1). In order to make results comparable, Model 4 re-estimates Model 3 on the reduced sample. The most striking difference between the two regressions is the coefficient for the domestic credit-to-GDP gap, which turns significant in the reduced sample. This indicates that any evaluation depends on the respective sample and should make policy makers cautious when generalizing findings from a particular sample of countries. However, the predictive abilities of the models are quite impressive. For example, Model 5, to which we refer as our *benchmark model*, achieves 60.3 % of the usefulness of a perfectly performing model and thus outperforms any individual indicator. The model issues a warning whenever the predicted probability is above its 63^{rd} percentile within the respective country. In this way, a warning is issued in 94.8 % of the quarters in our sample where a banking crisis occurs seven to twelve quarters ahead. The probability of a crisis conditional on a signal being issued is 28.5 %, which is 15.9 % higher than the unconditional probability of a crisis. Finally, the area under the ROC curve for this model is equal to 0.865, indicating a good predictive ability of the model for a wide range of policy maker's preference parameters (see Figure 2 for an illustration of the ROC curve for our benchmark model).[26]

We find that the credit variables are indeed among the most important predictors of vulnerable states of the economy. However, both model fit and model performance increase

[26] In contrast to the individual indicators and most of the credit models, the extended models perform well also for lower values of the preference parameter μ, which we see as another advantage of these models. These results are available upon request.

significantly when we include the other variables. For example, the consistently positive coefficient for house price growth indicates that asset price booms promote the build-up of vulnerabilities in the financial sector. This suggests that regulators should keep an eye on these developments instead of focusing exclusively on the development of credit variables (see also International Monetary Fund [24]). Moreover, Model 7 shows that banking sector variables exert a significant influence on the build-up of financial vulnerabilities.[27] We make the following observations: First, a country is more likely to be in a vulnerable state, when aggregate bank capitalisation within the country is relatively low. This is a particularly important finding in the context of countercyclical capital buffers as it indicates that indeed regulators could improve the resilience of the banking system by requiring banks to hold more capital when vulnerabilities build up. Second, we find that future banking crises are more likely when profits in the banking sector are relatively high. As Borio et al. [7] point out, periods of high bank profitability are typically associated with rapid credit growth, increased risk-taking and building up of vulnerabilities, which could explain the positive coefficient for the profitability variable preciding banking crises.

Figure 3 illustrates the relationship between predicted crisis probabilities from our benchmark model (Model 5) and actual banking sector capitalisation in the countries that had a banking crisis in 2007/2008. Most countries exerted declining or constantly low levels of bank capitalisation prior to the crisis, which is consistent with the evidence from Model 7.[28] At the same time, the benchmark model issues a warning already in late-2004/early-2005 in most cases.[29] Hence, if they had relied on this signal, regulators would have had enough time for the activation of the CCB prior to the crisis—even if we account for an announcement period of twelve months for the CCB.

[27] Again, the sample for Models 6 and 7 is reduced as banking sector variables are not available for all our sample countries. As before, we first re-estimate Model 5 on the reduced sample (see Model 6) in order to make results comparable.

[28] A notable exception is Austria (and to some extent Denmark), where aggregate banking sector capitalisation actually increased prior to the crisis.

[29] Again, the model issues a warning whenever the predicted probability is higher than the optimal threshold within the country (indicated by the dashed horizontal line in the figure).

4.2. Out-of-sample performance of the models

Given the objective of the early warning systems, any assessment should focus on the out-of-sample performance. Moreover, as shown by e.g. Berg et al. [6], successful in-sample predictions are much easier to achieve than successful out-of-sample predictions. In order to assess the out-of-sample usefulness of the models we proceed as follows: First, we consecutively exclude countries that had a banking crisis prior to 2007 from the estimation of the benchmark model. Then, we test whether the model based on the remaining countries is able to predict the crises in the excluded ones.[30]

The results of this exercise are presented in Figure 4. The benchmark model signals the banking crises in the Nordic countries well before their onset in the early 1990s.[31] In both Finland and Sweden, the indicator is consistently above the threshold from 1988Q2 onwards, which is 11 quarters ahead of the crisis for Finland and 9 quarters ahead for Sweden. In both cases, banks would have had enough time to build up capital before the crisis if the countercyclical capital buffer had been activated. Similarly, the model issues a warning signal for Italy from 1991Q2 onwards, 11 quarters ahead of the crisis in 1994. In the United Kingdom, the crisis is relatively close to the beginning of the sample period. Yet, in those quarters preceding the crisis of 1991, the benchmark model consistently issues a warning signal. Overall, the benchmark model exhibits strong out-of-sample properties. Information from the current crisis seems to be useful for the prediction of other systemic banking crises in the European Union.

[30]In principle we could have tried to fit a model to the observations prior to 2007 in order to see whether this model would be able to predict the current crisis. However, as most of the crisis episodes in our sample occur after 2007, and as we particularly want to learn something from these episodes, we prefer the approach described above, i.e., we use the information from the current crisis and check whether it would have been useful for the prediction of past crises.

[31]The model issues a warning whenever the predicted probability is higher than the optimal threshold within the country (indicated by the dashed horizontal line in the figure).

4.3. Robustness checks

In this section we modify the benchmark model (Model 5 of Table 4) in several ways in order to further assess the robustness of our results. The results from the robustness analysis are presented in Tables 6 and 7.

First, we check whether our results depend on the definition of the dependent variable. Apart from the ESCB Heads of Research database used in our analysis, the most common definitions of systemic banking crises are provided by Reinhart and Rogoff [36] and Laeven and Valencia [30]. Although the various databases are broadly consistent with each other, there are some deviations in the timing of crises as the definition of a systemic event in the banking sector requires a considerable amount of judgment. Columns 2 and 3 show that overall results are relatively similar for all three crisis definitions. Moreover, the area under the ROC curve is also greater than 0.8 for the other two models with the alternative crisis definitions, indicating good predictive abilities of the models.

Second, we include a dummy variable that is equal to one for each quarter in which the respective country is a member of the European Monetary Union (EMU). As expected, the coefficient for this dummy variable is positive and significant as most crises in our sample occur after the establishment of the EMU in 1999 (see column 4). However, the coefficients of the other variables remain largely unaffected by the inclusion of this dummy variable. Furthermore, the results are robust if we restrict the sample to include only countries from the EU-15 (column 5) or only countries that are part of the EMU (column 6).

Third, we augment the model with a money market rate (column 7). The estimated negative coefficient is potentially related to the 'great moderation', i.e., the general decline of inflation and money market rates over the sample period. The high R-squared and AUROC indicate that the fit of the model is superior compared to the other models. Despite this, we do not select this model as our benchmark model as its out-of-sample forecast abilities are inferior to the benchmark model, potentially due to an overfitting problem.

Fourth, following Lo Duca and Peltonen [32], we transform all variables into country-specific percentiles before using them in the regression. This method can be seen as an alternative way to account for heterogeneity across countries as differences in levels of indicators between countries vanish for the transformed variables. Columns 8 and 9 show that most of the estimated coefficients have the same sign as in the benchmark model if we use this alternative method.

Finally, we analyse model performance across different forecast horizons (see also Schudel [39]). Specifically, we check how the performance of the benchmark model and the indicator properties of variables change if the time window of the vulnerable state preceding a systemic banking crisis is altered from the twelve to seven quarters used in the standard specifications. Results in Table 7 show that although the benchmark model is broadly robust to an alteration of the forecast horizon, the relative importance and the estimated signs of the coefficients tend to vary a bit. Particularly important are the reversed signs for domestic and global credit growth in the model with the twenty to thirteen quarters ahead definition of a vulnerable state and the strong influence of global asset prices in this model (Model R3). As shown in Table 5, the benchmark model with a forecast horizon of twelve to seven quarters (Model 5) provides the highest absolute and relative usefulness measures, followed by the model with a forecast horizon of twelve to one quarter (Model R2). The performances of the models with the early (six to one quarter) and late (twenty to thirteen quarters) pre-crisis time horizons in terms of absolute and relative usefulness are broadly similar, but markedly lower than that of the benchmark model.

5. Conclusion

As a response to recent financial crises, the Basel III / CRD IV regulatory framework includes a countercyclical capital buffer (CCB) to increase the resilience of the banking sector and its ability to absorb shocks arising from financial and economic stress. In this context, this paper seeks to provide an early warning model, which can be used to guide the build-up and release of capital in the banking sector. Given the prominence of private credit variables in the upcoming regulations, the paper first examines the evolution of credit variables preceding banking crises in the EU Member States, and assesses their usefulness in guiding the setting of the CCB. Furthermore, the paper examines the potential benefits of complementing private credit variables with other macro-financial and banking sector indicators in a multivariate logit framework. The evaluation of the policy usefulness of the credit indicators and models follows the methodology applied in Alessi and Detken [2], Lo Duca and Peltonen [32] and Sarlin [38].

The paper finds that, in addition to credit variables, other domestic and global financial factors such as equity and house prices and banking sector variables help to predict macro-financial vulnerabilities in EU Member States. Consequently, the main policy implication of this study is that in the context of setting up CCB measures, policy makers could benefit from considering a wide range of macro-financial and banking sector indicators. In this respect, multivariate models, such as the one introduced in this study, are found to be more useful as they include the combined information of the joint behaviour of several indicators. The models demonstrate good out-of-sample predictive power, signalling the Swedish and Finnish banking crises of the early 1990s at least 6 quarters in advance.

References

[1] Ai, C. and Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, 80(1):123–129.

[2] Alessi, L. and Detken, C. (2011). Quasi real time early warning indicators for costly asset price boom/bust cycles: A role for global liquidity. *European Journal of Political Economy*, 27(3):520–533.

[3] Babecky, J., Havranek, T., Mateju, J., Rusnák, M., Smidkova, K., and Vasicek, B. (2012). Banking, debt and currency crisis: Early warning indicators for developed countries. *ECB Working Paper No. 1485*.

[4] Barrell, R., Davis, E. P., Karim, D., and Liadze, I. (2010). Bank regulation, property prices and early warning systems for banking crises in OECD countries. *Journal of Banking & Finance*, 34(9):2255–2264.

[5] Basel Committee on Banking Supervision (2010). Guidance for national authorities operating the countercyclical capital buffer. *Bank for International Settlements*.

[6] Berg, A., Borensztein, E., and Patillo, C. (2005). Assesing early warning systems: How have they worked in practice? *IMF Staff Papers*, 52(3):462–502.

[7] Borio, C., Drehmann, M., Gambacorta, L., Jimenez, G., and Trucharte, C. (2010). Countercyclical capital buffers: Exploring options. *BIS Working Paper No. 317*.

[8] Bussière, M. and Fratzscher, M. (2006). Towards a new early warning system of financial crises. *Journal of International Money and Finance*, 25(6):953–973.

[9] Candelon, B., Urbain, J.-P., and Van den Berg, J. (2008). A cautious note on the use of panel models to predict financial crises. *Economics Letters*, 101(1):80–83.

[10] Caprio, G. and Klingebiel, D. (2003). Episodes of systemic and borderline financial crises. *World Bank Research Dataset*.

[11] Davis, E. P. and Karim, D. (2008). Could early warning systems have helped to predict the sub-prime crisis? *National Institute Economic Review*, 206(1):35–47.

[12] Dembiermont, C., Drehmann, M., and Muksakunratana, S. (2013). How much does the private sector really borrow? A new database for total credit to the private non-financial sector. *Bank for International Settlements Quarterly Review, March 2013*.

[13] Demirgüç-Kunt, A. and Detragiache, E. (1998). The determinants of banking crises in developing and developed countries. *IMF Staff Papers*, 45(1):81–109.

[14] Demirgüç-Kunt, A. and Detragiache, E. (2000). Monitoring banking sector fragility. A multivariate logit approach. *World Bank Economic Review*, 14(2):287–307.

[15] Detragiache, E. and Spilimbergo, A. (2001). Crises and liquidity – Evidence and interpretation. *International Monetary Fund Working Paper No. WP/01/2*.

[16] Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74:427–431.

[17] Drehmann, M., Borio, C., and Tsatsaronis, K. (2011). Anchoring countercyclical capital buffers: The role of credit aggregates. *International Journal of Central Banking*, 7(4):189–240.

[18] Drehmann, M. and Juselius, M. (2013). Evaluating early warning indicators of banking crises: Satisfying policy requirements. *BIS Working Papers No. 421*.

[19] Edge, R. M. and Meisenzahl, R. R. (2011). The unreliability of credit-to-gdp ratio gaps in real-time: Implications for countercyclical capital buffers. *International Journal of Central Banking*, 7(4):261–298.

[20] EU (2013). Capital Requirements Directive IV. *Available online:* `http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32013L0036:EN:NOT`.

[21] Frankel, J. A. and Rose, A. K. (1996). Currency crashes in emerging markets: An empirical treatment. *Journal of International Economics*, 41(3):351–366.

[22] Hahm, J.-H., Shin, H. S., and Shin, K. (2013). Noncore bank liabilities and financial vulnerability. *Journal of Money, Credit and Banking*, 45(s1):3–36.

[23] Im, K. S., Pesaran, M. H., and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics*, 115(1):53–74.

[24] International Monetary Fund (2013). Key aspects of macroprudential policy.

[25] Kaminsky, G., Lizondo, S., and Reinhart, C. M. (1998). Leading indicators of currency crises. *IMF Staff Papers*, 45(1):1–48.

[26] Kaminsky, G. L. (2006). Currency crises: Are they all the same? *Journal of International Money and Finance*, 25(3):503–527.

[27] Kaminsky, G. L. and Reinhart, C. M. (1999). The twin crises: The causes of banking and balance-of-payments problems. *American Economic Review*, 89(3):473–500.

[28] Laeven, L. and Valencia, F. (2008). Systemic banking crises: A new database. *IMF Working Papers No. WP/08/224*.

[29] Laeven, L. and Valencia, F. (2010). Resolution of banking crises: The good, the bad, and the ugly. *IMF Working Papers No. 10/146*.

[30] Laeven, L. and Valencia, F. (2012). Systemic banking crises database: An update. *IMF Working Papers No. 12/163*.

[31] Little, T. D., Bovaird, J. A., and Widaman, K. F. (2006). On the merits of orthogonalizing powered and product terms: Implications for modeling interactions among latent variables. *Structural Equation Modeling*, 13(4):497–519.

[32] Lo Duca, M. and Peltonen, T. (2013). Assessing systemic risks and predicting systemic events. *Journal of Banking & Finance*, 37(7):2183–2195.

[33] Park, J. Y. and Phillips, P. C. (2000). Nonstationary binary choice. *Econometrica*, 68(5):1249–1280.

[34] Ravn, M. O. and Uhlig, H. (2002). On adjusting the Hodrick-Prescott filter for the frequency of observations. *Review of Economics and Statistics*, 84(2):371–376.

[35] Reinhart, C. M. and Rogoff, K. S. (2008). Banking crises: An equal opportunity menace. *National Bureau of Economic Research Working Paper No. 14587*.

[36] Reinhart, C. M. and Rogoff, K. S. (2011). From financial crash to debt crisis. *American Economic Review*, 101(5):1676–1706.

[37] Repullo, R. and Saurina, J. (2011). The countercyclical capital buffer of Basel III: A critical assessment. *CEPR Discussion Paper No. DP8304*.

[38] Sarlin, P. (2013). On policymakers loss functions and the evaluation of early warning systems. *Economics Letters*, 119(1):1–7.

[39] Schudel, W. (2013). Shifting horizons: Assessing macro trends before, during, and following systemic banking crises. *Mimeo*.

[40] Schularick, M. and Taylor, A. M. (2012). Credit booms gone bust. *American Economic Review*, 102(2):1029–1069.

[41] Seidler, J. and Gersl, A. (2011). Credit growth and capital buffers: Empirical evidence from Central and Eastern European countries. *Czech National Bank, Research and Policy Note No. 2/2011*.

[42] Yeyati, E. L. and Panizza, U. (2011). The elusive costs of sovereign defaults. *Journal of Development Economics*, 94(1):95–105.

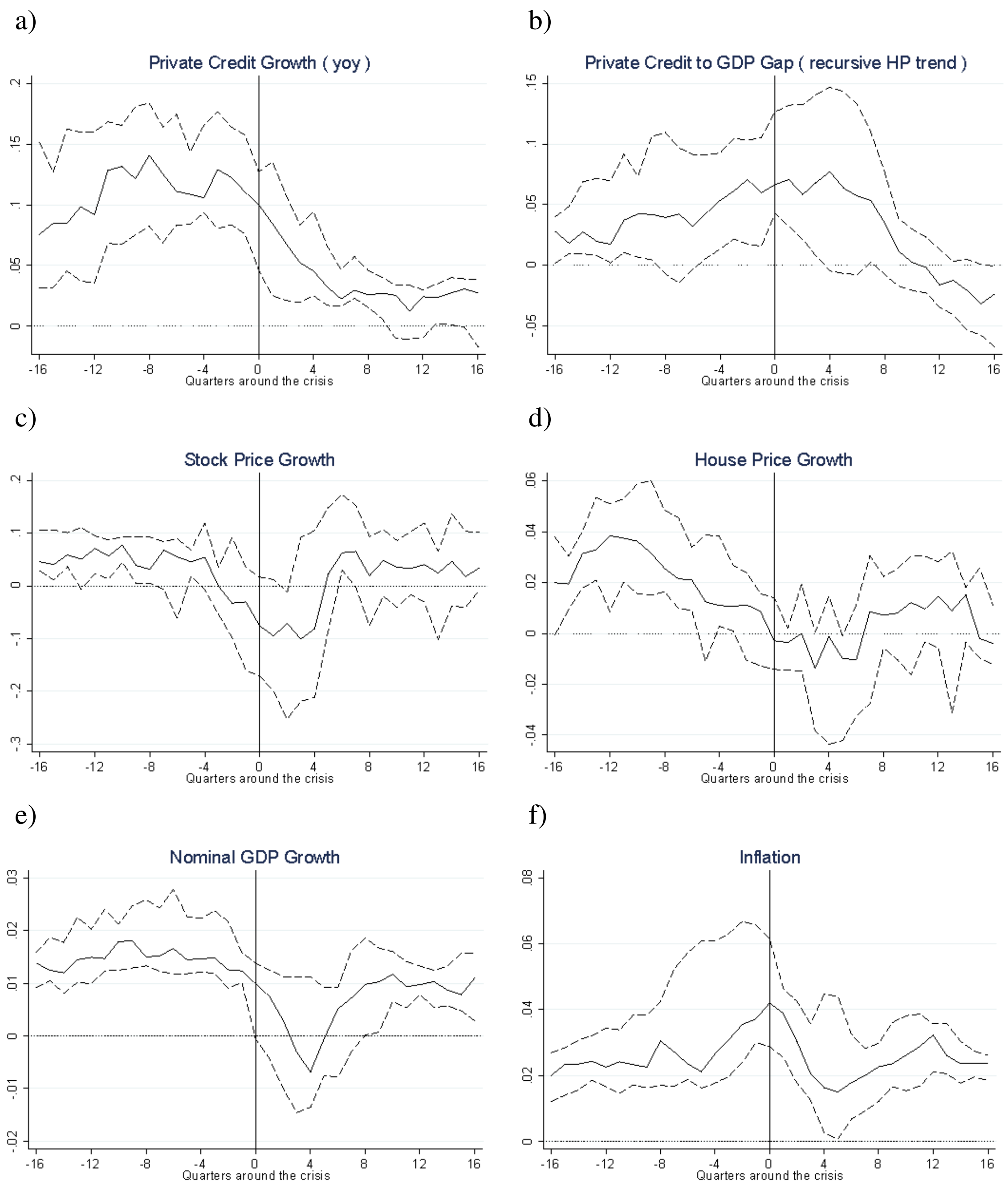

Figure 1: Development of key variables around banking crises

The figure depicts the development of selected key variables around banking crises within the sample countries. The start date of a banking crisis is indicated by the vertical line, while the solid line shows the development in the median country and the dashed lines represent the countries at the 25th and the 75th percentile, respectively.

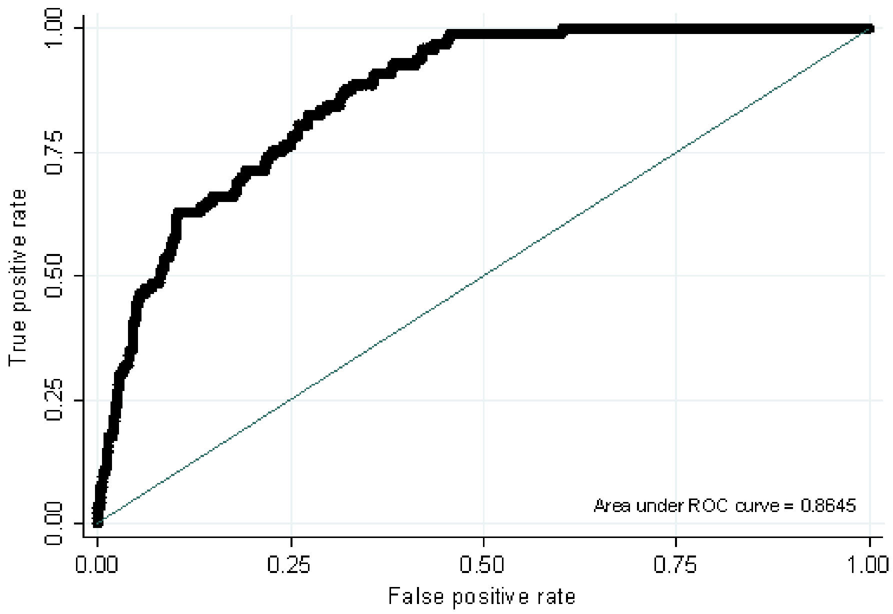

Figure 2: ROC Curve for benchmark model (Model 5)

The figure shows the Receiver Operating Characteristic (ROC) curve for our benchmark model. The area under the ROC curve (AUROC) is equal to 0.8645.

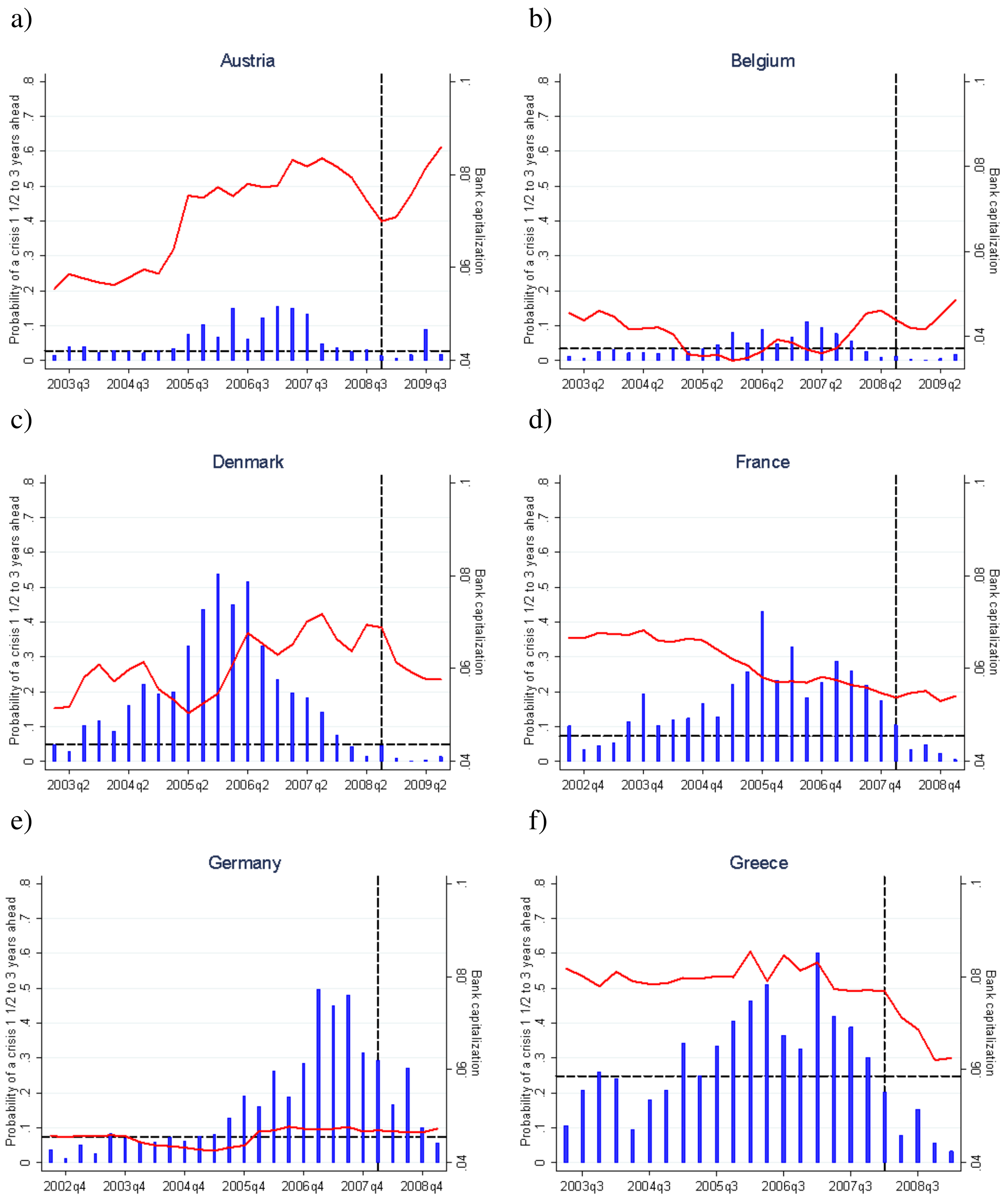

Figure 3: Predicted crisis probabilities and banking sector capitalisation

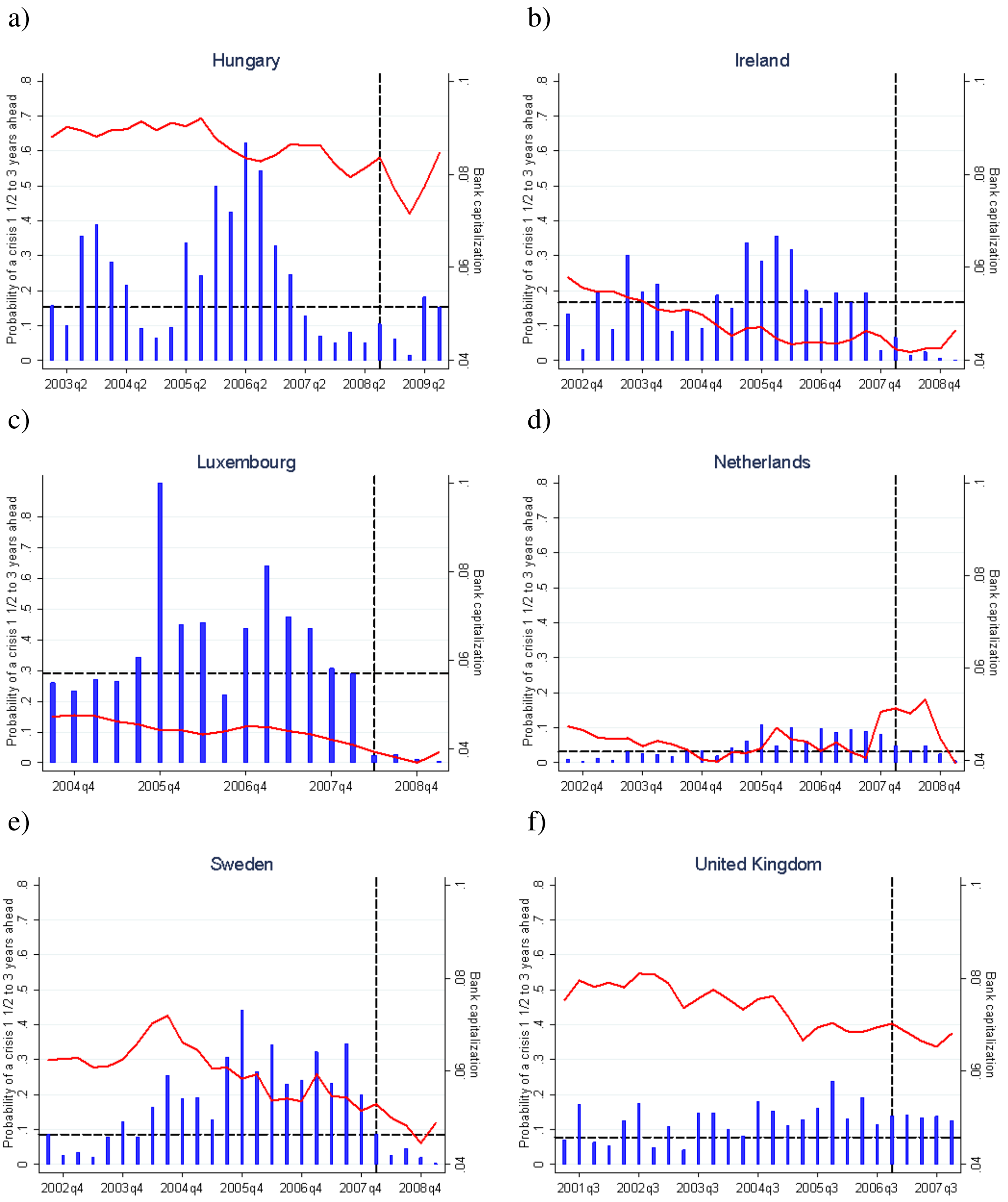

Figure 3 continued...

The figure plots the predicted probabilities (blue bars) from our benchmark model (Model 5 in Table 4) around the crises of 2008 in our sample countries (depicted by the dashed vertical lines). The optimal threshold for each country is depicted by the dashed horizontal line. The model issues a warning whenever the predicted probability is above this threshold. The red line shows the development of aggregate capitalisation in the banking sector defined as total banking sector equity over total banking sector assets.

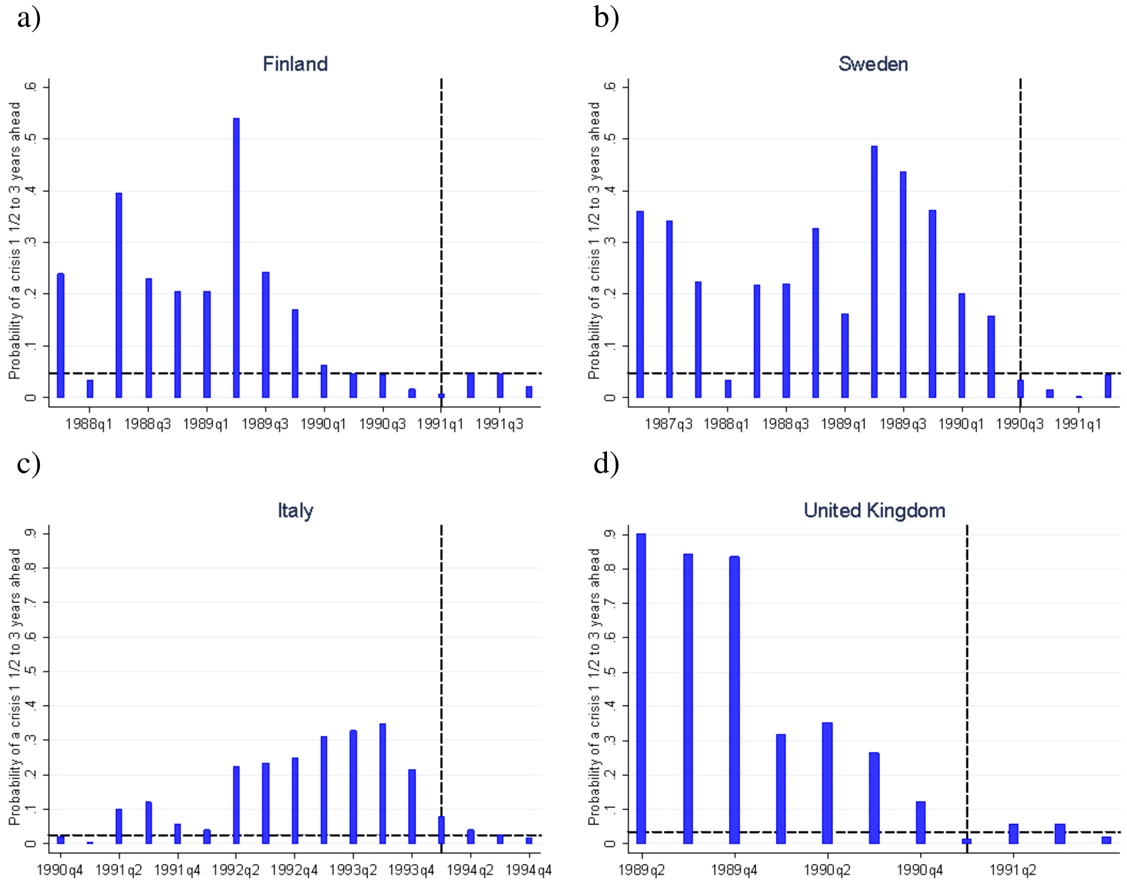

Figure 4: Out-of-sample performance of the model

The figure shows results for an out-of-sample evaluation of our benchmark model (Model 5 in Table 4). We exclude the respective country from the estimation and depict the predicted probabilities from a model based on the remaining countries around the crisis in the excluded country (dashed vertical line). The blue bars denote the predicted probabilities, while the horizontal dashed line presents the threshold.

Table 1: Data availability and descriptive statistics

Panel A: Data availability	Credit Variables	Other Variables	HoR Banking Crises
Austria	1982Q1-2012Q3	1986Q4-2012Q3	2008Q4
Belgium	1982Q2-2012Q3	1982Q1-2012Q3	2008Q3-2008Q4
Czech Republic	1994q2-2012Q2	—	1998Q1-2002Q2
Denmark	1982Q2-2012Q3	1992Q2-2012Q3	1987Q1-1993Q4, 2008Q3-ongoing
Estonia	2005Q1-2012Q2	2005Q2-2012Q2	–
Finland	1982Q2-2012Q3	1987Q2-2012Q3	1991Q1-1995Q4
France	1982Q2-2012Q3	1992Q2-2012Q3	1994Q1-1995Q4, 2008Q1-2009Q4
Germany	1982Q2-2012Q2	1991Q2-2011Q4	2008Q1-2008Q4
Greece	2003Q1-2012Q2	2003Q1-2012Q2	2008Q1-ongoing
Hungary	1997Q1-2012Q3	2002Q1-2012Q2	2008Q3-2009Q2
Ireland	1999Q1-2012Q3	1999Q1-2010Q4	2008Q1-ongoing
Italy	1982Q2-2012Q3	1990Q3-2012Q2	1994Q1-1995Q4
Lithuania	2005Q1-2012Q2	2005Q1-2012Q2	2009Q1-2009Q4
Luxembourg	2004Q2-2012Q3	2004Q2-2010Q4	2008Q2-ongoing
Malta	2006Q2-2012Q2	—	–
Netherlands	1982Q2-2012Q2	1982Q1-2011Q4	2008Q1-2008Q4
Poland	1997Q1-2012Q3	2003Q1-2012Q3	–
Portugal	1982Q2-2011Q4	1998Q2-2011Q4	–
Slovakia	2005Q2-2012Q2	—	–
Slovenia	2005Q3-2012Q2	—	–
Spain	1982Q2-2012Q3	1995Q2-2012Q3	1982Q2-1985Q3
Sweden	1982Q2-2012Q3	1986Q2-2012Q3	1990Q3-1993Q4, 2008Q3-2008Q4
United Kingdom	1982Q2-2012Q3	1988Q2-2012Q2	1991Q1-1995Q2, 2007Q1-2007Q4

Panel B: Descriptive statistics	Obs	Mean	Std. Dev.	Min	Max
Dom. Credit Growth (qoq)	1220	0.0228	0.0196	-0.0318	0.0989
Dom. Credit Growth (yoy)	1220	0.0926	0.0662	-0.0690	0.3579
Dom. Credit Gap	1220	0.1149	0.1186	-0.1570	0.4550
Dom. Credit Growth (4q MA)	1220	0.0232	0.0166	-0.0173	0.0897
Dom. Credit Growth (6q MA)	1220	0.0232	0.0154	-0.0122	0.0813
Dom. Credit Growth (8q MA)	1220	0.0233	0.0150	-0.0099	0.0805
Dom. Credit to GDP Ratio	1220	1.2756	0.4259	0.4426	2.4829
Dom. Credit to GDP Gap	1220	0.0346	0.0796	-0.1788	0.3249
Dom. Credit Growth - GDP Growth	1220	0.0081	0.0171	-0.0508	0.0715
Glo. Credit Growth (qoq)	1220	0.0152	0.0086	-0.0048	0.0335
Glo. Credit Growth (yoy)	1220	0.0614	0.0289	-0.0113	0.1095
Glo. Credit Gap	1220	0.0597	0.0431	-0.0101	0.1593
Glo. Credit Growth (4q MA)	1220	0.0154	0.0071	-0.0028	0.0274
Glo. Credit Growth (6q MA)	1220	0.0156	0.0069	-0.0021	0.0280
Glo. Credit Growth (8q MA)	1220	0.0158	0.0065	0.0005	0.0274
Glo. Credit to GDP Ratio	1220	0.7557	0.1193	0.5778	0.9933
Glo. Credit to GDP Gap	1220	0.0158	0.0285	-0.0420	0.0676
Glo. Credit Growth - Glo. GDP Growth	1220	0.0022	0.0235	-0.0492	0.0671
GDP Growth	919	0.0123	0.0088	-0.0232	0.0437
Inflation	919	0.0242	0.0166	-0.0108	0.1078
Equity Price Growth	919	0.0240	0.1199	-0.3759	0.3051
House Price Growth	919	0.0172	0.0289	-0.0735	0.1204
Banking Sector Capitalization	756	0.0507	0.0161	0.0238	0.1088
Banking Sector Profitability	756	0.0066	0.0040	-0.0142	0.0292
Gov. Bond Yield	862	0.0575	0.0237	0.0220	0.1385
Money Market Rate	862	0.0460	0.0292	0.0010	0.1643
Global GDP Growth	919	0.0117	0.0229	-0.0585	0.0616
Global Equity Price Growth	919	0.0135	0.0675	-0.3344	0.1122
Global House Price Growth	919	0.0066	0.0162	-0.0389	0.0539

Panel A shows the availability of credit and other variables as well as the crisis dates for the 23 countries in our sample. Credit variables are obtained from the BIS database for total credit to the private non-financial sector (see Dembiermont et al. [12]) and from Eurostat for those countries where the BIS data is not available. Other macro-financial and banking sector variables are obtained from various sources, including the BIS, IMF, and OECD. The crisis definitions are from the ESCB Heads of Research database described in Babecky et al. [3]. Panel B shows descriptive statistics for the credit as well as the other variables. Credit variables are available for a longer period of time in most countries, which is why the number of observations is larger for them.

Table 2: Contingency matrix

		Actual class C_j	
		1	0
Predicted class P_j	1	*True positive (TP)*	*False positive (FP)*
	0	*False negative (FN)*	*True negative (TN)*

The table shows the relationship between model prediction and actual outcomes. Observations are classified into those where the indicator issues a warning that is indeed followed by a banking crises twelve to seven quarters ahead (TP), those where the indicator issues a warning that is not followed by a crisis (FP), those where the indicator issues no warning and there is no crises seven to twelve quarters ahead (TN), and those where the indicator issues no warning although there is a crisis coming (FN).

Table 3: Evaluation of individual indicators

	μ	Threshold	TP	FP	TN	FN	T_1	T_2	Absolute Usefulness	Relative Usefulness	aNtS Ratio	% Predicted	Cond Prob	Diff Prob
Panel A: Domestic Variables														
Dom. Credit to GDP Gap	0.9	40	100	497	404	23	0.187	0.552	0.023	0.256	0.678	0.813	0.168	0.047
Dom. Credit Growth (yoy)	0.9	58	85	399	502	38	0.309	0.443	0.022	0.240	0.641	0.691	0.176	0.056
Dom. Credit to GDP Ratio	0.9	69	51	169	732	72	0.585	0.188	0.019	0.211	0.452	0.415	0.232	0.112
Dom. Credit Gap	0.9	37	104	577	324	19	0.154	0.640	0.018	0.201	0.757	0.846	0.153	0.033
Dom. Credit Growth (4q MA)	0.9	48	93	500	401	30	0.244	0.555	0.017	0.194	0.734	0.756	0.157	0.037
Dom. Credit Growth (6q MA)	0.9	61	72	364	537	51	0.415	0.404	0.015	0.170	0.690	0.585	0.165	0.045
Dom. Credit Growth (qoq)	0.9	46	92	530	371	31	0.252	0.588	0.014	0.153	0.786	0.748	0.148	0.028
Dom. Credit Growth - GDP Growth	0.9	54	70	409	492	53	0.431	0.454	0.009	0.103	0.798	0.569	0.146	0.026
Dom. Credit Growth (8q MA)	0.9	66	57	314	587	66	0.537	0.349	0.009	0.100	0.752	0.463	0.154	0.034
Panel B: Global Variables														
Glo. Credit Gap	0.9	45	113	427	474	10	0.081	0.474	0.040	0.443	0.516	0.919	0.209	0.089
Glo. Credit Growth (qoq)	0.9	60	100	357	544	23	0.187	0.396	0.037	0.412	0.487	0.813	0.219	0.099
Glo. Credit Growth (yoy)	0.9	57	101	365	536	22	0.179	0.405	0.037	0.411	0.493	0.821	0.217	0.097
Glo. Credit Growth (4q MA)	0.9	49	109	448	453	14	0.114	0.497	0.035	0.386	0.561	0.886	0.196	0.076
Glo. Credit Growth (6q MA)	0.9	46	110	467	434	13	0.106	0.518	0.033	0.373	0.580	0.894	0.191	0.071
Glo. Credit Growth (8q MA)	0.9	41	109	509	392	14	0.114	0.565	0.029	0.318	0.637	0.886	0.176	0.056
Glo. Credit to GDP Ratio	0.9	75	44	100	801	79	0.642	0.111	0.021	0.229	0.310	0.358	0.306	0.185
Glo. Credit to GDP Gap	0.9	37	105	571	330	18	0.146	0.634	0.019	0.216	0.742	0.854	0.155	0.035
Glo. Credit Growth - Glo. GDP Growth	0.9	83	46	161	740	77	0.626	0.179	0.016	0.178	0.478	0.374	0.222	0.102

The table shows results for the evaluation of individual indicator variables using the signalling approach (see Section 3.2 for a detailed description). The preference parameter of $\mu = 0.9$ indicates that policy makers have a strong preference for the detection of crises (i.e., avoiding type I errors) as compared to the avoidance of false alarms (i.e., type II errors). The optimal threshold is calculated as the one that maximises the relative usefulness and gives the percentile of the country-specific distribution at which the respective indicator issues a warning. The columns of the table report the number of observations: where the indicator issues a warning that is indeed followed by a banking crises seven to twelve quarters ahead (TP); where the indicator issues a warning that is not followed by a crisis (FP); where the indicator issues no warning and there is no crises seven to twelve quarters ahead (TN); and where the indicator issues no warning although there is a crisis coming (FN). Furthermore, the table reports the fraction of type I errors $T_1 = FN/(TP+FN)$, the fraction of type II errors $T_2 = FP/(FP+TN)$, the absolute and the relative usefulness (see Section 3.2 for details), the adjusted noise-to-signal ratio (i.e., the ratio of false signals measured as a proportion of months where false signals could have been issued to good signals as a proportion of months where good signals could have been issued, or $(FP/(FP+TN))/(TP/(TP+FN)))$, the probability of a crisis conditional on a signal being issued (Cond Prob) and the difference between the conditional and the unconditional probability of a crisis (Diff Prob). The domestic and the global variables are ranked in terms of relative usefulness, respectively.

Table 4: Multivariate models

	(1) Model 1	(2) Model 2	(3) Model 3	(4) Model 4	(5) Model 5	(6) Model 6	(7) Model 7
Dom. Credit Growth (DC1)	6.38**	2.66	1.61	3.93	1.54	-0.16	0.70
	(2.59)	(2.95)	(2.75)	(3.27)	(3.73)	(4.78)	(3.60)
Dom. Credit to GDP Gap (DC2)	3.01	2.71	3.70	8.55***	12.98***	13.24***	20.04***
	(1.93)	(2.80)	(2.66)	(2.27)	(2.27)	(3.19)	(2.68)
Interaction(DC1 $\times$ DC2)			26.42	55.68**	55.12**	83.05**	53.94
			(21.83)	(22.77)	(22.35)	(38.97)	(35.03)
Glo. Credit Growth (GC1)		16.71***	16.07***	29.01***	25.99***	19.72*	6.72
		(4.26)	(4.80)	(5.61)	(8.88)	(11.12)	(12.82)
Glo. Credit to GDP Gap (GC2)		1.96	-2.74	6.74	12.19	26.84**	41.15***
		(7.67)	(6.57)	(6.67)	(8.89)	(11.72)	(15.31)
Interaction(GC1 $\times$ GC2)			391.54**	-486.53*	-324.72	-472.64	-973.05***
			(188.05)	(258.07)	(305.59)	(312.51)	(317.56)
Interaction(DC1 $\times$ GC1)			45.98	-56.67	-28.48	-129.64	34.59
			(75.98)	(56.65)	(68.99)	(124.41)	(128.36)
Interaction(DC2 $\times$ GC2)			-239.65***	-417.35***	-472.20***	-410.73***	-582.20***
			(49.73)	(67.99)	(91.07)	(100.92)	(109.21)
GDP Growth					19.64	41.84	30.80
					(18.97)	(26.08)	(27.05)
Inflation					-29.04**	-10.04	14.19
					(11.73)	(12.23)	(12.18)
Equity Price Growth					-1.01	-0.38	-0.15
					(1.10)	(1.14)	(1.35)
House Price Growth					16.73***	19.80***	18.05***
					(5.40)	(5.56)	(5.35)
Global GDP Growth					-10.24	-10.58	-9.88
					(12.62)	(13.68)	(13.79)
Global Equity Price Growth					7.39	7.61	7.78
					(4.80)	(5.42)	(6.12)
Global House Price Growth					16.29	14.97	30.09
					(18.34)	(20.67)	(22.55)
Banking Sector Capitalization							-136.85***
							(39.63)
Banking Sector Profitability							324.89***
							(76.45)
Country dummies	YES	YES	YES	YES	YES	YES	YES
Observations	1,220	1,220	1,220	919	919	756	756
Pseudo R-Squared	0.0894	0.108	0.133	0.210	0.278	0.272	0.336
AUROC	0.710	0.733	0.780	0.824	0.865	0.846	0.892
Standard error	0.0266	0.0232	0.0185	0.0195	0.0160	0.0165	0.0157

The table shows estimation results for multivariate logit models, where the dependent variable is set to 1, twelve to seven quarters preceding a banking crisis in a respective country. Observations for banking crises and six quarters following banking crises are omitted, while other dependent variable observations are set to 0. All regressions include country-specific dummy variables to account for unobserved heterogeneity across countries. Robust standard errors adjusted for clustering at the quarterly level are reported in parentheses. * indicates statistical significance at the 10 %-level, ** at the 5 %-level, and *** at the 1 %-level.

Table 5: Model evaluation

	μ	Threshold	TP	FP	TN	FN	T_1	T_2	Absolute Usefulness	Relative Usefulness	aNtS Ratio	% Predicted	Cond Prob	Diff Prob
Model 1	0.9	48	97	489	427	28	0.224	0.534	0.021	0.236	0.688	0.776	0.166	0.045
Model 2	0.9	43	114	525	391	11	0.088	0.573	0.030	0.336	0.628	0.912	0.178	0.058
Model 3	0.9	56	108	333	583	17	0.136	0.364	0.045	0.497	0.421	0.864	0.245	0.125
Model 4	0.9	67	71	174	501	26	0.268	0.258	0.041	0.456	0.352	0.732	0.290	0.164
Model 5	0.9	63	92	231	444	5	0.052	0.342	0.054	0.603	0.361	0.948	0.285	0.159
Model 6	0.9	63	66	179	408	6	0.083	0.305	0.051	0.595	0.333	0.917	0.269	0.160
Model 7	0.9	67	64	163	424	8	0.111	0.278	0.051	0.596	0.312	0.889	0.282	0.173
Model R1	0.9	42	91	417	366	6	0.062	0.533	0.036	0.401	0.568	0.938	0.179	0.069
Model R2	0.9	65	81	251	532	16	0.165	0.321	0.045	0.503	0.384	0.835	0.244	0.134
Model R3	0.9	69	34	224	451	14	0.292	0.332	0.032	0.357	0.468	0.708	0.132	0.065

The table shows results for the evaluation of the multivariate models presented in Tables 4 and 7. As for the individual indicators, we apply the signalling approach by transforming predicted probabilities into country-specific percentiles. The preference parameter of $\mu = 0.9$ indicates that a policy maker has a strong preference for the detection of crises (i.e., avoiding type I errors) as compared to the avoidance of false alarms (i.e., type II errors). The optimal threshold is calculated as the one that maximizes the relative usefulness and gives the percentile of the country-specific distribution at which the respective indicator issues a warning. The columns of the table report the number of observations: where the indicator issues a warning that is indeed followed by a banking crises seven to twelve quarters ahead (TP); where the indicator issues a warning that is not followed by a crisis (FP); where the indicator issues no warning and there is no crises seven to twelve quarters ahead (TN); and where the indicator issues no warning although there is a crisis coming (FN). Furthermore, the table reports the fraction of type I errors $T_1 = FN/(TP+FN)$, the fraction of type II errors $T_2 = FP/(FP+TN)$, the absolute and the relative usefulness (see Section 3.2 for details), the adjusted noise-to-signal ratio (i.e., the ratio of false signals measured as a proportion of months where false signals could have been issued to good signals as a proportion of months where good signals could have been issued, or $(FP/(FP+TN))/(TP/(TP+FN))$), the probability of a crisis conditional on a signal being issued (Cond Prob) and the difference between the conditional and the unconditional probability of a crisis (Diff Prob).

Table 6: Robustness checks

	(1) Benchmark	(2) RR	(3) LV	(4) EMU	(5) EU-15	(6) Euro	(7) Interest Rates	(8) Percentiles	(9) Percentiles
Dom. Credit Growth (DC1)	1.54	-10.01**	-0.58	-1.08	1.01	-4.23	2.83	-0.009	-0.000
	(3.73)	(4.71)	(3.28)	(4.29)	(4.57)	(6.77)	(5.86)	(0.008)	(0.010)
Dom. Credit to GDP Gap (DC2)	12.98***	23.91***	4.09	15.68***	12.46***	14.70***	37.86***	0.039***	0.050***
	(2.27)	(4.32)	(2.60)	(2.51)	(2.64)	(4.73)	(4.67)	(0.007)	(0.009)
Interaction(DC1 × DC2)	55.12**	35.24	26.75	57.40**	55.10*	138.12*	83.42*	0.018***	0.026***
	(22.35)	(36.30)	(26.96)	(26.67)	(31.69)	(71.79)	(46.34)	(0.005)	(0.005)
Glo. Credit Growth (GC1)	25.99***	32.12***	39.40***	49.67***	18.57**	13.00	108.66***	0.024**	0.018
	(8.88)	(10.77)	(15.26)	(10.14)	(9.42)	(10.23)	(16.80)	(0.010)	(0.011)
Glo. Credit to GDP Gap (GC2)	12.19	-8.68	42.49***	-4.67	20.22*	18.52*	-37.62**	-0.033***	0.004
	(8.89)	(10.83)	(15.62)	(12.87)	(11.24)	(10.52)	(17.82)	(0.008)	(0.013)
Interaction(GC1 × GC2)	-324.72	-255.97	1,718.19***	-806.86**	-390.66	-121.35	-1,941.20***	0.011	-0.005
	(305.59)	(320.49)	(587.11)	(345.60)	(296.43)	(301.21)	(527.98)	(0.009)	(0.010)
Interaction(DC1 × GC1)	-28.48	77.07	12.25	-17.42	-240.77**	-134.60	208.01	-0.001	-0.004
	(68.99)	(88.37)	(104.67)	(89.52)	(102.03)	(101.30)	(139.13)	(0.005)	(0.006)
Interaction(DC2 × GC2)	-472.20***	-754.19***	-339.87***	-617.47***	-436.37***	-276.74**	-1,466.07***	-0.042***	-0.062***
	(91.07)	(102.92)	(66.07)	(133.84)	(96.37)	(109.19)	(182.09)	(0.006)	(0.010)
GDP Growth	19.64	2.67	11.35	12.95	18.29	14.00	35.04	0.003	0.001
	(18.97)	(20.61)	(25.41)	(20.57)	(20.65)	(21.40)	(24.89)	(0.006)	(0.006)
Inflation	-29.04**	-3.98	-25.31*	-29.70**	-3.65	15.29	68.72***	-0.001	-0.004
	(11.73)	(11.60)	(15.21)	(12.35)	(10.79)	(10.41)	(20.10)	(0.005)	(0.006)
Equity Price Growth	-1.01	-0.20	-1.46	-1.57	-0.43	-0.14	-1.77	-0.008	-0.009
	(1.10)	(1.49)	(2.11)	(1.16)	(1.26)	(1.51)	(1.91)	(0.006)	(0.006)
House Price Growth	16.73***	14.03***	6.73	21.02***	23.45***	20.71***	20.73**	0.019***	0.015**
	(5.40)	(4.81)	(8.52)	(6.20)	(5.33)	(5.50)	(9.37)	(0.005)	(0.006)
Gov. Bond Yield							-28.69		
							(45.08)		
Money Market Rate							-125.08***		
							(38.55)		
Global GDP Growth	-10.24	-4.87	-40.28*	-8.78	-7.63	-10.67	0.08	-0.004	-0.004
	(12.62)	(12.25)	(20.80)	(12.60)	(12.95)	(15.21)	(11.66)	(0.009)	(0.009)
Global Equity Price Growth	7.39	7.11	21.31***	6.86	8.58*	10.72*	6.47	0.016	0.018*
	(4.80)	(4.56)	(7.62)	(4.99)	(5.07)	(6.14)	(4.61)	(0.010)	(0.010)
Global House Price Growth	16.29	8.29	51.73**	11.56	20.92	27.87	27.49	0.046***	0.054***
	(18.34)	(17.54)	(23.01)	(20.00)	(18.53)	(21.34)	(19.84)	(0.014)	(0.013)
D(EMU)				2.83***					
				(0.62)					
Country dummies	YES	YES	YES	YES	YES	YES	YES	NO	YES
Observations	919	835	893	919	869	664	862	919	919
Pseudo R-Squared	0.278	0.286	0.442	0.314	0.269	0.270	0.477	0.324	0.372
AUROC	0.865	0.827	0.807	0.887	0.847	0.836	0.942	0.892	0.909
Standard error	0.0160	0.0207	0.0223	0.0136	0.0172	0.0184	0.0091	0.0126	0.0117

The table shows several robustness checks for our benchmark model (Model 5 in Table 4). In column 2 we use the Reinhart and Rogoff [35] crisis definition instead of the Babecky et al. [3] definition, while in column 3, we use the definition by Laeven and Valencia [28]. Column 4 includes a dummy that is equal to one for each quarter in which the respective country is a member of the European Monetary Union (EMU) and column 5 restricts the sample to include only EU-15 countries for which data availability is better than for the rest of our sample countries. Column 6 restricts the sample to include only the countries that are a part of the EMU. In column 7 we include two interest rate variables, the 10-year government bond yield and the 3-months interbank lending rate. Finally, in columns 8 and 9 we transform all variables into country-specific percentiles before using them in the regression. All specifications (except column 8) include country-specific dummy variables to account for unobserved heterogeneity across countries. Robust standard erros adjusted for clustering at the quarterly level are reported in parantheses. * indicates statistical significance at the 10 %-level, ** at the 5 %-level, and *** at the 1 %-level.

Table 7: Robustness—Forecast horizon

	(1) Benchmark 7-12 quarters	(2) Model R1 1-6 quarters	(3) Model R2 1-12 quarters	(4) Model R3 13-20 quarters
Dom. Credit Growth (DC1)	1.54	-11.63	-3.06	-11.77***
	(3.73)	(8.06)	(4.28)	(4.07)
Dom. Credit to GDP Gap (DC2)	12.98***	61.69***	33.92***	16.03***
	(2.27)	(20.72)	(3.91)	(4.08)
Interaction(DC1 × DC2)	55.12**	-18.61	139.78***	-1.36
	(22.35)	(82.99)	(36.95)	(32.76)
Glo. Credit Growth (GC1)	25.99***	113.85***	93.09***	-50.73***
	(8.88)	(18.75)	(12.10)	(14.10)
Glo. Credit to GDP Gap (GC2)	12.19	17.90	2.08	28.57**
	(8.89)	(21.61)	(10.95)	(13.06)
Interaction(GC1 × GC2)	-324.72	6,895.95***	2,763.43***	-52.28
	(305.59)	(1,067.10)	(502.18)	(449.06)
Interaction(DC1 × GC1)	-28.48	453.38	223.97**	-606.47***
	(68.99)	(322.98)	(99.32)	(111.78)
Interaction(DC2 × GC2)	-472.20***	-1,947.93***	-1,193.54***	-458.57***
	(91.07)	(565.00)	(151.01)	(79.88)
GDP Growth	19.64	-28.92	-14.58	35.03
	(18.97)	(48.51)	(20.64)	(40.03)
Inflation	-29.04**	77.15***	22.54*	24.42**
	(11.73)	(24.02)	(13.25)	(10.59)
Equity Price Growth	-1.01	2.14	-0.45	-1.14
	(1.10)	(2.34)	(1.40)	(1.98)
House Price Growth	16.73***	-22.60**	7.29	8.62
	(5.40)	(9.98)	(6.19)	(5.91)
Global GDP Growth	-10.24	-0.39	-2.07	12.86
	(12.62)	(12.63)	(13.10)	(9.72)
Global Equity Price Growth	7.39	14.15***	15.06***	12.53***
	(4.80)	(5.08)	(4.91)	(4.76)
Global House Price Growth	16.29	-60.67**	-32.74	116.13***
	(18.34)	(30.62)	(23.07)	(21.94)
Observations	919	919	919	919
Pseudo R-Squared	0.278	0.781	0.617	0.340
AUROC	0.865	0.726	0.960	0.895
Standard error	0.0160	0.0179	0.0063	0.0134

The table shows the results of robustness analysis with respect to the forecast horizon. Column 1 re-estimates Model 5 from Table 4, which is referred as the benchmark model. The dependent variable in this regression is set to one twelve to seven quarters preceding a banking crisis in a respective country. In column 2, we replace the dependent variable with a dummy that is equal to one, six to one quarter before a banking crisis. Similarly, the dependent variable in column 3 equals one twelve to one quarter before a banking crisis in the respective country. Finally, the dependent variable in column 4 is equal to one twenty to thirteen quarters before the onset of a banking crisis in a respective country. Robust standard errors adjusted for clustering at the quarterly level are reported in parentheses. * indicates statistical significance at the 10 %-level, ** at the 5 %-level, and *** at the 1 %-level.