EUROPEAN CENTRAL BANK

EUROSYSTEM

Working Paper Series

David Sondermann, Nico Zorell

A macroeconomic vulnerability model for the euro area

No 2306 / August 2019

Abstract

Macroeconomic imbalances increase the vulnerability of an economy to adverse shocks, which in turn can lead to crises with severe economic and social costs. We propose an early warning model that predicts such crises. We identify a set of macroeconomic indicators capturing domestic and external imbalances that jointly predict severe recessions (i.e. growth crises) in a multivariate discrete choice framework. The approach allows us to quantify an economy's macroeconomic vulnerabilities at any point in time. In particular, the model would have pointed early on to emerging vulnerabilities in all the euro area countries that registered severe recessions in the years after 2007. We also show that the model can be applied beyond the euro area crisis in that its main results remain robust to changes in assumptions and sample composition.

Keywords: macroeconomic imbalances, early warning models, growth crises
JEL codes: E37, E44, F47, O52

Non-technical summary

Macroeconomic imbalances increase the vulnerability of an economy to adverse shocks. They often build up over several years, for instance in the form of credit-fuelled housing booms, soaring indebtedness and excessive current account deficits. When an adverse shock eventually hits the economy, the imbalances suddenly need to be unwound. This can result in a crisis with substantial economic and social costs.

The severe recessions experienced by several euro area countries after 2007 are a prominent example. They were preceded by a prolonged period of tranquility in which broad-based macroeconomic imbalances were accumulated. Several countries registered exuberant growth in house prices, credit and wages. These domestic imbalances often went hand in hand with steady losses in cost and price competitiveness, deteriorating export performance and widening current account deficits. At the same time, policy makers failed to build sufficient fiscal buffers. Moreover, a lack of structural reforms left most of these economies with severe rigidities, preventing an efficient allocation of resources. When the US sub-prime crisis eventually sent its shock waves across the globe, euro area countries with broad-based macroeconomic imbalances witnessed particularly severe recessions. The already elevated private and public debt ratios surged further, impeding market access. Several countries were ultimately forced to request international financial assistance. Overall, the euro area experience demonstrates that macroeconomic imbalances can impair economic resilience and thereby lead to particularly severe recessions.

Existing early warning models did not anticipate the severe post-2007 crises in the euro area. Even with the benefit of hindsight, these models have had difficulties explaining the cross-country variation in the incidence and severity of the post-2007 crises. In part, this is because the models did not fully capture the risks stemming from broader macroeconomic vulnerabilities. The literature was more concerned with narrowly-defined vulnerabilities typically haunting developing countries, such as a lack of reserves prior to currency crises. The studies published after the inception of the global financial crisis have focused more on crisis aspects related to the financial sector and sovereign debt. Thus, some of the macroeconomic vulnerabilities that arguably played an important role in explaining the cross-country variation and severity of the post-2007 crises in the euro area have so far been largely ignored in the early warning literature.

In this paper, we develop a model that gauges the likelihood of crises related to macroeconomic vulnerabilities. To this end, we match past patterns of macroeconomic imbalances, spreading over a sample of 32 OECD economies and almost 40 years, with episodes of severe recessions (i.e. growth crises) that followed the accumulation of such imbalances. In this sense, our model complements the existing early warning literature.

We employ a multivariate discrete choice model that links the likelihood of growth crises to a set of explanatory macroeconomic indicators. The advantage of this approach, as compared with the more traditional signal extraction method, is that is allows us to look at patterns of indicators and their interactions. Most variables cannot be seen in isolation but often pose vulnerabilities only when they appear together with other vulnerabilities. A case in point is a surge in house prices, which is particularly dangerous when fuelled by excessive credit growth.

We identify ten core variables that are able to predict growth crises. External imbalances are covered by the current account balance and a country's export market share. We also include the CPI-deflated real effective exchange rate and compensation per employee as measures of price and cost competitiveness. In turn, domestic imbalances that can predict growth crises are captured by the stock of government, non-financial corporations (NFC) and household debt as well as house price and credit growth. We interact government debt with a proxy of financial market volatility to gauge a country's vulnerability to sudden changes in market sentiment.

The benchmark model performs well both in sample and out of sample. Moreover, it is very robust to changes in assumptions (e.g. related to the identification of crises) and the sample composition.

We show that our model would have been able to identify ex ante, on the basis of a set of macroeconomic indicators, the euro area countries that underwent particularly severe recessions in the years after 2007. For many of the countries that entered into EU-IMF financial assistance programmes in 2010 and subsequent years, the model would have already pointed to significant vulnerabilities several years in advance. At the same time, the model also performs well when we exclude those crises.

Our model can be seen as a quantitative complement to the scoreboard used in the context of the EU's Macroeconomic Imbalance Procedure (MIP). Most of the 14 variables included in the scoreboard are mirrored in our macroeconomic vulnerability model. While the scoreboard is only used for qualitative analysis, our early warning model allows us to attach a concrete quantitative risk score to an economy's macroeconomic vulnerability at any point in time.

1 Introduction

The global financial crisis and the subsequent euro area crisis exposed severe macroeconomic vulnerabilities in many euro area countries. The weaknesses were not only linked to over-indebtedness of the public sector, but encompassed a wide set of macroeconomic imbalances (see Pierluigi and Sondermann 2018 for an overview). Several countries registered exuberant growth in house prices, credit and wages. These domestic imbalances often went hand in hand with steady losses in cost and price competitiveness, deteriorating export performance and widening current account deficits. At the same time, policy makers failed to build sufficient fiscal buffers. Moreover, a lack of structural reforms left most of these economies with severe rigidities, preventing an efficient allocation of resources. When the US sub-prime crisis eventually sent its shock waves across the globe, euro area countries with broad-based macroeconomic imbalances witnessed particularly severe recessions. The already elevated private and public debt ratios surged further, impeding market access. Several countries were ultimately forced to request international financial assistance. The euro area experience vividly demonstrates that macroeconomic imbalances can impair economic resilience and thereby lead to particularly severe recessions.

In the mid-2000s, some academics (e.g. Lane 2006; Mongelli and Wyplosz 2008; Zemanek et al. 2009; Berger and Nitsch 2010) and institutions (e.g. Trichet 2005, 2006) started hinting towards macroeconomic heterogeneity across euro area countries and their implications for the smooth functioning of EMU. Although a few observers pointed to increasing risks, market participants and economists alike overall failed to anticipate the timing and severity of the post-2007 crises in the euro area.

It is not that early warning models did not exist at the time. Yet, the suite of early warning models used in the years preceding the euro area crisis arguably had a somewhat different focus. The first set of tools were applied to so-called currency crises, such as the Asian crisis in the late 1990s (e.g. Kaminsky and Reinhart 1999; Cuaresma and Slacik 2009; Gochoco-Bautista 2000). In the same vein, banking or more broadly financial crises were recognised and modelled, such as in Demirgüç-Kunt and Detragiache (1998), Bussière and Fratzscher (2006) or later by Laeven and Valencia (2013). However, the subsequent crises in euro area countries were characterised by a broader set of domestic imbalances and external sector weaknesses. Those weaknesses were partly linked to house price and credit growth, but also to exuberant wage growth in some countries as well as increasing vulnerabilities in the balance sheets of non-financial corporations and the public sector. Moreover, external imbalances were visible in the current account and export performance.

We contribute to the existing literature by developing a dedicated macroeconomic vulnerability model focusing on a broad set of macroeconomic domestic and external imbalances. We argue that such a model can be a complement to existing early warning models. This follows from the finding in the literature (e.g. Christofides et al. 2016; Basu et al. 2017) that early warning models work most effectively if separated by type of crisis. Against this backdrop, we adapt both sides of the early warning model. On the left hand side, we suggest to be more agnostic and broader on the crisis definition by focusing on pronounced drops in economic activity. We thus suggest identifying growth crises, somewhat along the lines of similar concepts employed by the IMF and the OECD (Basu et al. 2017; Prasad et al. 2019; Hermansen and Röhn 2017; Röhn et al. 2015). On the right-hand-side of the equation we overall limit ourselves to macroeconomic variables while mostly abstracting from country-specific financial variables. We show that the benchmark model performs very well compared to the standards set by other early warning models.

Our approach is also motivated by EU policy considerations. With the inception of the crisis in the euro area, the EU created the Macroeconomic Imbalance Procedure (MIP) in 2012 as part of the already existing economic governance framework with a view to improving the coordination of economic policies among EU Member States. The procedure draws on a set of indicators and corresponding alert thresholds summarised in a cross-country scoreboard. In our baseline model, we remain very close to the type of indicators selected in the scoreboard. However, an important difference between the MIP scoreboard and our model is that the data in the scoreboard are not synthesised but are standing next to each other. Hence, the scoreboard ignores interdependencies between indicators and does not measure the overall severity of imbalances in a Member State. Our macroeconomic vulnerability model, by contrast, features a multivariate framework attempting to complement the more qualitative analysis by a quantitative framework synthesising information into one macroeconomic vulnerability score.

While a few recent papers have studied the forecasting performance of the MIP indicators, they either suffer from a relatively short sample period or focus on a subset of EU Member States that does not include euro area countries. For instance, Siranova and Radvanský (2018) look at growth crises in central and eastern European EU Member States only. Domonkos et al. (2017) study growth crises in all EU Member States, but focus on a very short time period (2011-15). Our paper adds to these existing studies in that it broadens the sample across countries and time, while also employing a state-of-the-art discrete choice framework.[1]

The remainder of the paper is organised as follows. Section 2 introduces the methodology and data underlying our analysis while Section 3 presents the main results. Section 4 discusses the implications of our model-based results for individual euro area countries, while Section 5 concludes.

[1] Knedlik (2014) finds that the alert thresholds used in the context of the MIP scoreboard put a relatively heavy weight on crisis prevention.

2 Methodology and data used

2.1 Early warning methodology

The early warning literature usually aims to predict crisis events, captured by a binary variable, on the basis of relevant indicators. The most widely used methodologies can be split into three groups. The first approach, signal extraction, goes back to the seminal papers by Kaminsky et al. (1998) and Kaminsky and Reinhart (1999). These papers calculate a separate threshold value for each indicator such that observations on one side of the threshold are seen as a crisis signal while those on the other side point to tranquil times. This approach has the advantage of straightforward computations imposing hardly any constraints. In particular, the approach can easily handle differences in data availability across indicators. Moreover, the results for individual variables can be aggregated without the multi-collinearity faced in regression environments, albeit often at the cost of resorting to ad-hoc aggregation methods.

The second approach builds on discrete choice models and goes back to pioneering papers from the late 1990s (e.g. Frankel and Rose 1996; Berg and Pattillo 1999a,b). It has been increasingly used in the literature over recent years (Lo Duca and Peltonen 2013; Catão and Milesi-Ferretti 2014; Jorda et al. 2015), as it offers several advantages compared to standard signal extraction tools. In particular, the explanatory variables can be assessed jointly, taking into account the correlation of the variables. The models also allow to assess the relative importance of individual indicators. Moreover, it is possible to test rigorously for the statistical significance of individual variables and the stability of coefficients across countries and time.

A third strand of the literature has lately started experimenting with machine learning techniques, such as random forests (e.g. Alessi and Detken 2018). While machine learning methods often deliver superior in-sample fit, they still tend to be outperformed by logit models in out-of-sample evaluations, as shown by Beutel et al. (2019) for the case of banking crises.

Against this backdrop, and in particular given the interconnections among the macroeconomic variables we are interested in, we choose a multivariate logit model for the empirical analysis. We follow Bussière and Fratzscher (2006) in the overall presentation of the model.

Suppose we look at N countries $i = \{1, 2, ..., N\}$ and T years $t = \{1, 2, ..., T\}$. For each country and year, the binary dependent variable captures the incidence of crises that we will define in the following Section 2.2:

$$\text{Crisis} = \begin{cases} 1 & \text{with probability } Pr(\text{Crisis} = 1) = P \\ 0 & \text{with probability } Pr(\text{Crisis} = 0) = 1 - P. \end{cases} \tag{1}$$

The paper aims to project the incidence of crises through a set of J explanatory variables X. Thus, X is a $JN \times T$ matrix of observations. In a logit model, the probability of a crisis is a non-linear function of the explanatory indicators and is given by

$$Pr(\text{Crisis} = 1) = \frac{e^{X\beta}}{1 + e^{X\beta}}. \tag{2}$$

Given the non-linearity of the logit, the marginal effect of a change in an explanatory variable on the probability outcome is not constant but depends on the precise state of X. To gain some intuition for this, it is useful to rewrite the logit model in terms of the odds ratio, i.e. the ratio of the crisis probability to its complement:

$$\Omega(\text{Crisis} = 1) = \frac{P}{1 - P} = e^{X\beta}. \tag{3}$$

This formulation shows that increasing the j-th regressor by one unit while holding all other variables constant will multiply the odds ratio by e^{β_j}. The preceding equation also demonstrates that the log of the odds ratio is a simple linear function of the regressors.

For any country-year observation, the model will produce a crisis probability ranging between zero and one. For evaluation and communication purposes, these continuous probabilities are usually translated into a binary alarm indicator. This variable equals 1 if the model-implied crisis probability exceeds a certain threshold $\tilde{P}$ and 0 otherwise. Thus, sufficiently high crisis probabilities are seen as a signal of an impending crisis. Ex post, the signal will turn out to be either correct or false. Thus, when comparing the incidence of actual crises with the binary alarm indicator, the results can be summarised by the following matrix:

	crisis	no crisis
signal issued	A	B
no signal issued	C	D

Here, A denotes the number of cases in which an indicator provides a correct crisis signal, while B is the number of false alarms. The second row captures the cases in which no signal is issued, either mistakenly (C) or rightly so (D). On the basis of this matrix, we can calculate the true positive rate as $TPR = A/(A+C)$. Hence, the share of missed crises (type I error) is given as $1 - TPR = C/(A + C)$. By contrast, the false positive rate is defined as $FPR = B/(B + D)$ and captures the share of false alarms (type II error).

The selection of the alert threshold involves a trade-off between maximising the number of good calls and minimising the number of false alarms. The literature often approaches this problem on the basis of a loss function modelling the preferences of policymakers that takes the following form:

$$L = \theta \frac{C}{A + C} + (1 - \theta) \frac{B}{B + D}. \tag{4}$$

Here, θ varies between zero and one, denoting the policy maker's aversion against type I errors as compared to type II errors. On the basis of such a loss function, Alessi and Detken (2011) postulate a so-called relative usefulness criterion, which flags whether the model is more useful for the policy maker than a naive benchmark:

$$RU = \frac{\min[\theta; 1-\theta] - L}{\min[\theta; 1-\theta]} \tag{5}$$

The naive benchmark assumes that the policymaker disregards the model and either always or never issues an alarm (depending on the value of θ), thereby realising a loss of $\min[\theta; 1-\theta]$. The relative usefulness indicator can take values between -1 and 1. All values below zero dismiss the model as being uninformative whereas a value of 1 suggests that the model correctly predicts all crises and never issues a false alarm. The threshold should be chosen such as to maximise the relative usefulness.

Sarlin (2013) developed an alternative version of the above loss function that explicitly takes into account that crises tend to be rare. On this basis, he showed that standard early warning systems are only useful to policymakers if they are much more concerned about missed crises than false alarms.[2]

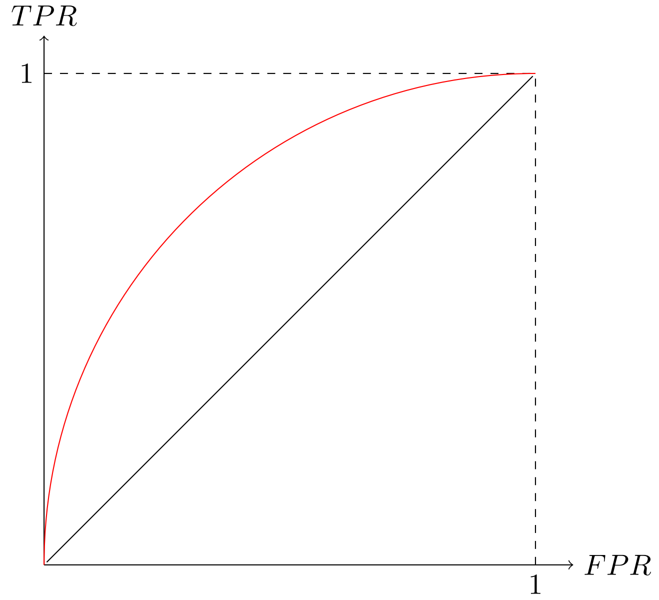
A drawback of all these loss functions is that the preferences of the policy maker are usually unobserved and therefore specified arbitrarily. Moreover, even relatively small variations in the preference parameters can result in significant changes in the optimal threshold. Therefore, we will consider different values of the preference parameter that are consistent with those used in the early warning literature.

A different way of looking at the trade-off between type I and type II errors is the so-called Receiver Operating Characteristic (ROC) curve, which plots the fraction of correctly predicted crises against the fraction of false alarms for various probability thresholds. A higher probability threshold will reduce the number of false alarms, yet at the expense of a larger number of missed crises. Finding an optimal alarm threshold amounts to choosing a point on the ROC curve, for instance based on the criteria outlined above.

Figure 1 plots the false positive rate (FPR) and the true positive rate (TPR) on the x- and y-axis respectively for every possible value of the threshold. The area under the ROC-curve (AUROC) ranges from zero to one, where a value larger than 0.5 (corresponding to a ROC curve that is to the left of the black line in Figure 1) indicates that the underlying model issues informative signals.

[2] Alessi and Detken (2014) argue however that the approach is not always robust to small perturbations of the key parameter.

Figure 1: The Receiver Operating Characteristics (ROC) curve



2.2 Data

The sample used in our analysis covers 32 OECD and EU countries, starting in 1980 and running up until 2017 (at annual frequency).[3] Annex A describes the country coverage. Depending on the availability of the various explanatory variables, some of our discrete choice models have a somewhat shorter sample.

2.2.1 Crisis definition

As for all early warning models, we first need to define the crisis we want to predict. Since we are interested in crises related to broad-based macroeconomic vulnerabilities, we choose economic activity as the most appropriate underlying measure. We argue that significant slumps in GDP, i.e. *growth crises*, are the most encompassing definition for the type of crisis we want to predict. Similar concepts are also used by the IMF and the OECD in their model-based country surveillance (Basu et al. 2017; Hermansen and Röhn 2017; Röhn et al. 2015). Naturally, such a broad, outcome-based definition includes the risk that crises other than those related to macroeconomic imbalances are subsumed in the set of identified crises. However, this just raises the bar for our early warning model. Assuming that by having a broader focus we involuntarily capture some crises that are unrelated to macroeconomic imbalances, each explanatory variable has to stand a more difficult test for being included in the model.

[3] Applying the framework on the basis of quarterly data faces severe limitations given the short history of many of our relevant explanatory variables. Accordingly, the estimation sample would be significantly shorter and thus cover only a small subset of crises. It is against this background that we prefer to use annual data.

Table 1: Growth crises - descriptive statistics

	No.	Peak-to-trough (%)	Length (quarters)
Full sample (1980-2017)	53	-4.9	5.0
1980s	9	-3.5	3.0
1990s	8	-3.2	4.5
2000s	28	-6.8	5.0
2010s	8	-5.0	7.0
Euro area countries	32	-5.2	5.0
Memo: all recessions	122	-2.1	4.0

Note: The peak-to-trough change in the level of real GDP and the crisis length refer to the median episode.

Technically, we define growth crises as recessions with a peak-to-trough decline in real GDP of at least 2.5% at quarterly frequency. Business cycle turning points are identified with the widely-used BBQ algorithm proposed by Harding and Pagan (2002). The BBQ algorithm identifies turning points as the local minima and maxima of a time series within a window of k quarters. For an observation $y(t)$ to be considered as a potential turning point, it needs to satisfy $y(t-k),...,y(t-1) < y(t) > y(t+1),...,y(t+k)$ (for a peak) or $y(t-k),...,y(t-1) > y(t) < y(t+1),...,y(t+k)$ (for a trough). In addition, the distance between peak and trough (i.e. the phase length) must be at least p quarters and the distance between successive peaks and troughs (i.e. the cycle length) at least c quarters. We follow the standard choice of parameters, with $k = 2$, $p = 2$ and $c = 5$. Once the turning points have been identified, all recessions with a peak-to-trough decline of less than 2.5% are discarded. This eliminates almost one half of all recessions in our sample. The remaining recessions constitute growth crises. Converted to annual frequency, the binary dependent variable will take the value 1 in the first year of a growth crisis and 0 in all other years.

Overall, we identify 53 growth crises in the countries under observations for the period 1980-2017. As shown in Table 1, the median peak-to-trough decline in real GDP is 4.9% and the typical crisis length 5 quarters. The crises are relatively evenly spread across decades, with the notable exception of a sharp spike in the 2000s in the wake of the global financial crisis. Given the large share of this period in the overall number of crises identified, we will apply dedicated robustness checks to test the sensitivity of the results to the exclusion of the global financial crisis. Beyond the effect of the global financial crisis, we will more broadly check the robustness of our results to alternative definitions of growth crises.

The growth crises we have identified are distinct from other crises typically used in the literature, such as banking, currency and debt crises. This is corroborated by Table 2, which reports the correlations between different types of crisis. The correlations are very low, with a maximum of around 0.18 between growth and banking crises. Not surprisingly, most of the overlaps between these two crisis definitions occur during the global financial crisis.

Table 2: Correlation of different crises

	Growth crises	Banking crises	Currency crises	Debt crises
Growth crises	1.0000			
Banking crises	0.1767	1.0000		
Currency crises	0.0036	0.0081	1.0000	
Debt crises	0.0704	0.1867	0.0095	1.0000

Note: The table shows the phi coefficient, a correlation measure for binary variables. Growth crises follow our own definition. The other crises are taken from the dataset developed by Babecký et al. (2014), which stops in 2010.

How much in advance can the model predict a crisis? One would generally expect the model's predictive power to deteriorate with longer prediction horizons. To start with, it should be noted that two quasi-automatic lags are already included. First, there is the publication lag, since most annual data are released with a delay. Second, we will use three-year moving averages for many of the explanatory variables to avoid crisis signals on the basis of blips in the data. Against this background and in line with the majority of the early warning literature (e.g. Alessi and Detken 2011; Lo Duca and Peltonen 2013), we apply a lag, in our case one year, for the prediction of the identified crisis. Concretely, we will use data for 2015, 2016 and 2017 to project the probability of a crisis in 2018.

Another choice to make is how to tackle potential post-crisis bias (Bussière and Fratzscher 2006). Such a bias could arise given that our binary crisis variable takes the value 0 in tranquil periods but also in post-crisis periods, when macroeconomic variables often undergo an abrupt adjustment process. To address this issue, we follow the standard practice in the literature and remove the year directly following the start of a crisis from the sample (see e.g. Demirgüç-Kunt and Detragiache 1998; Fuertes and Kalotychou 2007; Catão and Milesi-Ferretti 2014; Savona and Vezzoli 2010 for comparison). When turning to robustness checks in Section 3.4, we will also analyse alternative ways of tackling the post-crisis bias.

2.2.2 Explanatory variables

As noted in Section 1 we aim to build a model that issues early warning signals for emerging macroeconomic imbalances that have the potential to lead to growth crises. The literature on macroeconomic imbalances is guiding our selection of possible explanatory variables. Another source of inspiration is the scoreboard of variables used in the context of the EU's Macroeconomic Imbalance Procedure (MIP). In fact, the most important indicators contained in the

MIP scoreboard are also included in our baseline model.

The literature usually distinguishes between external imbalances (including competitiveness issues) and internal imbalances. Of course, this does not preclude the possibility that both types of imbalance may interact with each other. In fact, the importance of such interaction effects is one of the reasons why we look at a multivariate logit model rather than a signal extraction tool with an indicator-by-indicator perspective.

The current account balance and the net international investment position (NIIP) have long been identified as good indicators of external macroeconomic imbalances (see e.g. Blanchard and Milesi-Ferretti 2009; Chen et al. 2012; Lane and Pels 2012; Zorell 2017). At the same time, current account deficits are not per se a symptom of imbalances. In particular in countries that are less advanced in terms of economic development, current account deficits can be the consequence of a healthy convergence process (Blanchard and Giavazzi 2002).

This again highlights the need to look at the current account not in isolation, but in a multivariate empirical framework linking current account deficits to persistent losses in cost and price competitiveness, as measured by the real effective exchange rate (REER), as also noted by Frankel and Saravelos (2012). Higher wage growth compared to trading partners, resulting in higher ULC growth and inflation rates, is not only a sign of reduced competitiveness, but often also a reflection of overheating tendencies. Lane (2006) argues that in many euro area countries where inflation rates were higher than in other parts of the currency area low real interest rates contributed to booming domestic demand and widening of current account deficits.

However, competitiveness extends beyond real effective exchange rates and cost or price competitiveness (Carlin et al. 2001). Weak non-cost competitiveness can also be detrimental for the trade performance. Against this background, it might be useful to look at export performance relative to peer countries, as captured by changes in export market shares.

In cases of eroding competitiveness and export market share losses, resources often flow into domestic activities (as also suggested by Chen et al. 2012). Therefore, our early warning model will also take into account indicators of exuberance in domestic demand. We look at widely-used indicators such as residential investment, house price growth (e.g. Adalid and Detken 2007) and private sector credit growth (e.g. Alessi and Detken 2011; Frankel and Saravelos 2012; Lane and Milesi-Ferretti 2011). There is ample evidence that domestic imbalances often go hand in hand with external imbalances and may amplify each other. In particular, buoyant domestic credit growth is often related to large net financial inflows from the rest of the world, in particular net debt inflows (Lane and McQuade 2014).

Most of the above-mentioned indicators are flow variables. However, a vast literature shows that some stock variables, in particular high levels of debt, can also increase a country's vulnerability to adverse shocks. High debt is found to hinder the ability of households and firms to smooth consumption and investment spending, and the ability of governments to cushion adverse shocks (e.g. Reinhart and Rogoff 2010). As regards the private sector, Berkmen et al. (1999) suggest that countries with a more leveraged financial system and higher credit growth were more affected by the global financial crisis. In general, negative feedback loops between high private and public sector debt and a weak financial sector seem to constrain investment decisions and economic growth (see the discussion in Sutherland and Hoeller 2012).

Evidence from past crises suggests that high debt levels often prevailed for some time before amplifying the severity of a crisis in times of increasing financial market volatility. Against this background, various proxies of financial market volatility, such as the VIX indicator related to equity prices or measures of bond spreads, are often applied (Bekaert et al. (2010)).

Labour market indicators are rarely found in the early warning literature as they usually respond with a lag to fluctuations in output. However, they might still contain information on an economy's vulnerability to adverse shocks going forward and are also included in the MIP scoreboard.

As noted in Section 2.2.1, we apply three-year changes for most of the flow variables to avoid false alarms driven by blips in the data. A notable exception is the current account balance, which is sufficiently persistent to be considered without this transformation. The same applies to all stock variables.[4] Only indicators of financial market volatility are used contemporaneously, since they are available at higher frequency and can therefore be updated in real time.

It should be noted that the model focuses on macroeconomic vulnerabilities that increase an economy's susceptibility to crises. By contrast, it is relatively silent on the concrete triggers of a crisis, which are often unpredictable.

3 Results

3.1 Non-parametric analysis

In Section 2.2 we have listed a range of potentially useful indicators suggested by the literature on macroeconomic imbalances. To obtain a first idea of their usefulness for the prediction of growth crises, we first apply some of the non-parametric methods described in Section 2.1 on an indicator-by-indicator basis. This will also allow us to arrive at a more parsimonious logit regression model.

For each of the variables put forward in the previous section, Table 3 reports descriptive

[4] For some of the variables, standard panel unit root tests cannot reject the presence of non-stationarity. However, the maximum likelihood estimator of a logit/probit model is also consistent in the presence of non-stationarity (Berg and Coke 2004; Park and Phillips 2000). Against this backdrop, the inclusion of non-stationary regressors is a relatively common practice in the early warning literature (e.g. Dawood et al. 2017; Lo Duca and Peltonen 2013; Schimmelpfennig et al. 2003). Of course, the presence of non-stationarity will be taken into account in the calculation of the standard errors.

Table 3: Non-parametric analysis of potential early warning indicators

	obs.	min.	max.	mean	s.d.	AUROC	RU^a	RU^b
External imbalances and competitiveness								
Current account (in % of GDP)	1041	-20.9	16.2	-0.4	5.1	0.62	0.25	0.24
Export market share (3-year growth)	963	-16.1	34.4	0.7	5.2	0.60	0.21	0.21
NIIP (in % of GDP)	751	-175.9	144.8	-16.4	46.2	0.56	0.16	0.08
REER (CPI-deflated, 3-year growth)	1020	-11.6	14.3	0.3	3.1	0.53	0.11	0.10
Domestic imbalances								
Household sector credit flow (% of GDP, 3-year change)	856	-14.2	11.0	1.3	2.5	0.68	0.30	0.23
House price growth (CPI-deflated, 3-year growth)	943	-24.1	39.0	2.1	7.1	0.60	0.19	0.15
Unemployment rate (3-year change)	972	-3.5	4.9	0.0	1.0	0.60	0.01	0.01
NFC debt (in % of GDP)	890	26.5	364.2	88.9	46.1	0.60	0.21	0.15
Compensation per employee (3-year growth)	985	-5.2	28.7	5.3	5.2	0.58	0.19	0.18
Residential investment (3-year growth)	687	-40.8	44.2	1.7	8.4	0.57	0.19	0.18
Government debt (in % of GDP)	981	3.7	236.4	59.7	36.3	0.55	0.03	0.00
Household debt (in % of GDP)	897	1.5	139.4	54.0	28.4	0.54	0.11	0.09
Budget balance (in % of GDP, 3-year change)	884	-10.8	8.6	0.1	1.2	0.53	0.07	0.06
Private consumption (3-year growth)	979	-7.5	15.3	2.5	2.4	0.53	0.09	0.08
Financial markets								
VIX	880	11.1	32.6	19.5	5.8	0.81	0.56	0.54
Government bond spread	888	-8.4	20.7	0.7	2.6	0.65	0.31	0.29
Government bond volatility	887	0.0	4.9	0.4	0.4	0.61	0.20	0.18

(a) Relative usefulness based on Alessi and Detken (2011) with preference parameter equal to 0.5.
(b) Relative usefulness based on Sarlin (2013) with preference parameter equal to 0.95.
Note: The three panels are sorted by the AUROC indicator. All variables are lagged by one year, except for the financial market indicators. All growth rates are annualised.

statistics along with standard measures of their usefulness in predicting growth crises. It turns out that all variables of interest exhibit an AUROC value above 0.5, suggesting they may help predict growth crises. The relative usefulness measures proposed by Alessi and Detken (2011) and Sarlin (2013) are also positive for all variables. The ranking of variables is broadly similar for the three different information criteria, although notable differences exist for a few variables. For example, household debt would rank rather low using the AUROC measure, but higher on the basis of the relative usefulness indicators.

Notwithstanding those variables for which indicators give different results, there are some variables that can be identified across the board as being particularly promising. In line with the early warning literature, the current account balance tends to be among the variables most helpful in predicting crises related to macroeconomic vulnerabilities. Moreover, competitiveness losses resulting from cost pressures, in this case strong rises in compensation per employee, as well as credit growth, house prices and some of the indebtedness variables fare rather well across the selected information criteria. The financial indicators, in particular the VIX, also perform very well. However, it should be noted already at this stage that the VIX indicator alone is not performing equally well for all sub-samples. By contrast, the predictive power of our baseline model that we introduce in the next section is very robust to perturbations of the estimation sample.

Less promising, by contrast, are variables such as private consumption and the real effective exchange rate. The fiscal variables, i.e. the budget balance and government debt, also appear to be surprisingly uninformative when it comes to predicting growth crises as defined in this paper. However, this might be related to the state contingency of public debt. High indebtedness of the public sector might be less problematic as long as financial markets are calm and investors have confidence in the respective government. We will look at the interaction of these factors in the following section.

3.2 Benchmark framework

The workhorse model for our empirical analysis is a discrete choice model, specifically a logit model, as explained in Section 2.1. The model relates a binary growth crisis variable (Section 2.2.1) to a range of explanatory factors (Section 2.2.2). Since the errors could be correlated, we use robust standard errors (clustered by country). All variables are lagged by one year with the exception of the financial market indicators which are included contemporaneously.

While we start from the variables identified as particularly promising in the univariate analysis in Section 3.1, the perspective changes when variables have to stand the test of their relevance in a multivariate framework. As can be expected, not all variables still add value to the crisis prediction if they are assessed jointly in a discrete choice framework. Most of the variables that were found to contain more information in the non-parametric analysis, however, also remain important in the multivariate context. This does not hold for the NIIP, the budget balance and the unemployment rate. In particular for the latter, however, this seems intuitive as the labour market developments are often seen as lagging output trends.

Table 4 depicts the evolution of the model in waterfall format. Our benchmark specification includes ten regressors (with two variables each being part of an interaction term) capturing specific macroeconomic imbalances. We find that a strong loss in cost competitiveness, as measured by the three-year cumulative wage growth, has a positive and significant impact on the probability of a growth crisis. Price competitiveness, as measured by the CPI-deflated real effective exchange rate, also has a positive and significant sign. Accordingly, an appreciation vis-à-vis a country's trading partners over three years increases the likelihood of a growth crisis. The finding of a significant REER is interesting, given that the univariate analysis in Section 3.1 did not highlight the variable's predictive power. This notwithstanding, it seems to add explanatory power as a complementary indicator.

In terms of external imbalances, we find the current account balance and export market share growth to be good predictors of future crises. Turning to domestic imbalances, we look at an interaction of credit growth and house prices, since strong house price increases only provide the nucleus for a severe crisis if they are associated with excessive credit growth. In cases where house prices are mounting without being fuelled by dynamic credit growth, a downturn and a subsequent deflation of house prices will not impact the economy to the same extent as households are not forced to deleverage quickly and banks are not confronted with rising non-performing loans. The interaction term is positive and significant.[5]

Table 4: Benchmark logit model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Compensation per employee (3-year change)	0.009***	0.012***	0.008***	0.014***	0.075***	0.077***	0.089***	0.088***
	(0.003)	(0.003)	(0.002)	(0.003)	(0.019)	(0.021)	(0.021)	(0.020)
Government debt (% of GDP) x VIX		0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Current account balance (% of GDP)			-0.133***	-0.128***	-0.069**	-0.072***	-0.081***	-0.080***
			(0.040)	(0.037)	(0.030)	(0.028)	(0.030)	(0.031)
Household debt (% of GDP)				0.014***	0.025***	0.019***	0.016**	0.016*
				(0.004)	(0.007)	(0.007)	(0.008)	(0.008)
Real house price growth x credit growth (3-year change)					0.005**	0.004*	0.005*	0.005*
					(0.003)	(0.002)	(0.003)	(0.003)
NFC debt (% of GDP)						0.008***	0.010***	0.010***
						(0.003)	(0.003)	(0.003)
Export market share (3-year change)							-0.036***	-0.040***
							(0.011)	(0.013)
Real effective exchange rate, HICP-deflated (3-year change)								0.069**
								(0.035)
Constant	-3.056***	-3.845***	-4.109***	-4.960***	-6.780***	-7.234***	-7.314***	-7.421***
	(0.107)	(0.263)	(0.407)	(0.503)	(1.077)	(1.121)	(1.131)	(1.156)
Pseudo R^2	0.011	0.040	0.085	0.099	0.149	0.159	0.177	0.182
Observations	1053	847	847	775	692	691	691	690

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; clustered standard errors in brackets.

Moreover, stock imbalances are very relevant to explain growth crises. Countries that are highly indebted, as measured here by the debt levels of non-financial corporations (NFCs), households and the government, are more vulnerable to adverse shocks and thus face a higher probability of a growth crisis. As outlined in Section 3.1, we include a measure of financial market volatility and interact it with government debt, thereby controlling for the fact that government debt can be state dependent. As long as financial market are tranquil, governments might find it easy to access markets. In periods of heightened uncertainty, however, high debt levels may increase an economy's susceptibility to a growth crisis. The interaction term is indeed positive and significant. We also looked at other measures of financial market uncertainty, such as government bond spreads and the volatility of sovereign bond yields, but did not find them to be significant. In Section 3.4 we will show that our main results survive a battery of robustness checks.

In order to avoid multi-collinearity issues, we also computed the cross-correlations between our explanatory variables. However, we only find very low correlations between the variables.

[5] A positive value of the interaction term could be the result of both variables jointly increasing or decreasing. The elevated likelihood of a crisis would mainly stem from the former. To ensure that the identification is correct, we also perform regressions with a dummy controlling for the two states. Given that the results remain highly robust, the baseline model is displayed without the additional dummy variable, but the results are available from the authors upon request.

Table 5: Mean value of explanatory variables during tranquil and crisis periods

	(1) Crisis=0	(2) Crisis=1	(3) t-test
Compensation per employee (3-year growth)	15.82	20.27	0.0001
Government debt (% of GDP)	61.05	55.12	0.8154
VIX	18.73	26.85	0.0000
Current account balance (% of GDP)	0.06	-2.83	0.0001
Household debt (% of GDP)	53.46	58.34	0.2023
Credit growth (3-year change)	3.55	8.89	0.0000
Real house price growth (3-year growth)	2.43	4.89	0.0006
NFC debt (% of GDP)	87.79	96.30	0.1083
Export market share (3-year growth)	2.72	-2.24	0.0878
Real effective exchange rate, HICP-deflated (3-year growth)	0.32	0.52	0.1619

Note: Column (1) reports the mean of the variables one year preceding a tranquil year, while column (2) refers to the mean one year before a crisis. The VIX is included contemporaneously. Column(3) reports the p-value of a one-sided t-test for equal means across the two outcomes. Only observations included in the baseline regression are taken into account.

Since the logit model is non-linear, the coefficients reported in Table 4 cannot be interpreted as marginal effects (see Section 2.1). The marginal effects can however be calculated at specific values of the independent variables. For illustrative purposes we compute the mean value of our right-hand-side variables for the years preceding tranquil and crisis times. Table 5 shows that one year before a crisis the variables point to greater vulnerabilities than prior to a tranquil year. Although these are purely descriptive, univariate statistics, they confirm the narrative of our early warning model.[6]

In a second step, we construct a hypothetical country for which all explanatory variables of the model equal the sample average in tranquil times as indicated in Table 5. We then compute the effect of one variable increasing to the average crisis value. Doing so for all ten independent variables individually, we find that the crisis probabilities increase up to two percentage points for some variables, such as wage growth and financial market uncertainty. For the remaining variables the increase is milder, with on average around one fourth of a percentage point. While in particular the latter value seems small, it is in fact reasonable when put into perspective. First, the estimated crisis probabilities in actual crisis periods tend to be relatively low for binary logit models estimated on samples with unequal frequencies of the two outcomes. We will elaborate more on the underlying technical reason for this when calculating appropriate alarm thresholds in Section 3.3.

A second reason for the small increase in probability of between 0.25% and around 2% is that a crisis is in fact rather unlikely if only one of the ten variables increases to the crisis level while the others remain at tranquil level. Such cases would rather reflect some idiosyncratic

[6] The only variable that is slightly lower in tranquil times than during crisis times is public debt in percent of GDP. This could reflect the cyclical impacts of exuberant economic activity on public finances. In the baseline logit, other regressors arguably control for such effects. For household debt and the real effective exchange rate, the differences in means are not statistically significant at standard confidence levels.

concerns of limited macroeconomic scope. Indeed, if all variables jump to their typical crisis level, the probability of a crisis increases to 13.2%.

3.3 Model evaluation

To show that our model is useful in predicting growth crises, we first evaluate its in-sample forecasting performance. The standard approach in the early warning literature is to translate the model-implied crisis probabilities into binary alarm signals and then compare these signals with the actual crisis periods. This mapping requires an alarm threshold. If the crisis probability exceeds the threshold, this is counted as a warning signal.

The selection of the alert threshold involves a trade-off between maximising the number of good calls and minimising the number of false alarms. To solve this optimisation problem, we use the loss-function-based approaches by Alessi and Detken (2011) and Sarlin (2013) outlined in Section 2.1.

Applied to our benchmark logit model, the Alessi/Detken approach delivers a probability threshold between 6% and 10%, assuming standard values of the preference parameter θ (Table 6). For the middle threshold of 8%, the model correctly predicts almost 70% of all growth crises in our sample while keeping the type II errors at a still acceptable level of 18%. Several of the false alarms are still followed by a crisis in one of the following years. Not surprisingly, the Sarlin approach requires a much higher aversion against type I errors in order to arrive at thresholds similar to those suggested by the Alessi/Detken method.

It should be noted that the relatively low alarm thresholds reflect a common feature of binary logit models estimated on samples with unequal frequencies of the two outcomes. In such a setting, the estimated crisis probabilities in actual crisis periods are on average lower than the predicted probabilities of no-crisis in tranquil periods (Cramer 1999). This translates into relatively low alarm thresholds.[7]

Table 6: Optimal alarm thresholds based on different criteria

	Alessi/Detken			Sarlin		
Preference parameter	*0.40*	*0.50*	*0.60*	*0.92*	*0.95*	*0.98*
Probability threshold	0.10	0.08	0.06	0.10	0.08	0.02
Type-I error (missed crises)	0.40	0.33	0.26	0.40	0.33	0.02
Type-II error (false alarms)	0.13	0.18	0.28	0.13	0.18	0.68

Note: The table reports the optimal alarm thresholds based on the approaches by Alessi and Detken (2011) and Sarlin (2013).

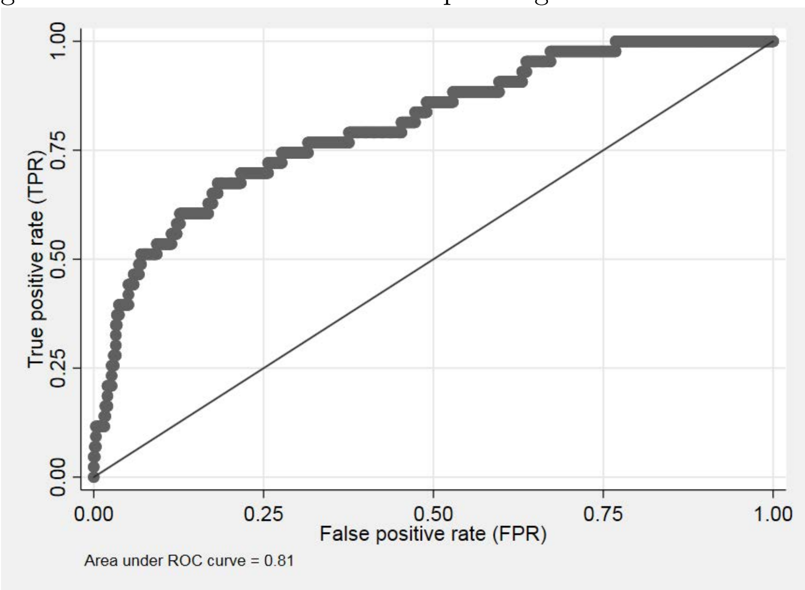
We conclude that standard selection criteria point to a range of plausible alarm thresholds

[7] It is also noteworthy that the average in-sample crisis probability predicted by a logit model is by construction equal to the actual crisis frequency.

for our benchmark logit model, although they cannot pin down a single optimal cutoff. In any case, policymakers may at times be interested in tentative signs of emerging vulnerabilities, even if they are not yet strong enough to warrant a fully-fledged alarm signal. In the subsequent analysis (Section 4), we will therefore consider a broader range of graduated thresholds that are consistent with Table 6. These thresholds partition the probability space into several risk categories, thereby also preserving a larger amount of the information contained in the model output.

Since the preceding model evaluation approach depends crucially on the alarm threshold chosen, we also look at an alternative performance metric, the AUROC measure introduced in Section 2.1. This indicator has the advantage of summarising the model's goodness-of-fit on the basis of *all* possible alarm thresholds (Candelon et al. 2012). To recall, the AUROC corresponds to the area under the Receiver Operating Characteristics Curve. It varies between 0 and 1, with higher values pointing to better model performance. As shown in Figure 2, our model delivers an AUROC of 0.81, pointing to very good in-sample performance.

Figure 2: Area under the Receiver Operating Characteristics Curve



Since early warning models tend to be vulnerable to in-sample overfitting, we now look into the model's out-of-sample properties. To this end, we employ the so-called k-fold validation method. More specifically, we split the sample into $k = 3$ random sub-samples, of which two are used to estimate the logit model and the third one is used to compute the out-of-sample AUROC. We repeat this exercise 1,000 times and calculate descriptive statistics for the AUROC across all these estimations. Table 7 shows that the average AUROC is 0.77 and that the model always beats an uninformative benchmark (which has an AUROC of 0.50). We take this as strong evidence that our baseline specification also performs well out of sample.

Table 7: Out-of-sample evaluation: K-fold validation

	AUROC
Mean	0.77
Minimum	0.62
Maximum	0.97
Memo: baseline	0.81

Note: The table reports descriptive statistics for the AUROC based on a k-fold validation. See main text for details.

3.4 Robustness checks

In the following, we undertake a range of robustness checks to document that our main results are not sensitive to changes in the assumptions taken (all tables and charts are included in Annex B).

To start with, we test the robustness of our main results with respect to alternative crisis definitions. Specifically, we relax the assumption that growth crises are associated with a peak-to-trough decline in output of at least 2.5%, considering less demanding thresholds of 2.0% and 1.5% instead. In addition, we deviate from the BBQ algorithm and try as an alternative approach the definition of a technical recession, which is commonly understood as negative GDP growth for at least two consecutive quarters. Table 9 in Annex B reports the incidence of crises with these alternative identification approaches. As shown by Table 10 in Annex B, our regression results overall remain robust to variations in the definition of the dependent variable. Only the household debt variable in two of the three specifications and the current account balance in one regression become (just) insignificant.

Another robustness check in relation to the crisis definition is to test for the specific nature of the global financial crisis. As already noted, a significant number of identified growth crises took place in the wake of the global financial crisis. While our motivation is to some extent related to the developments in euro area countries in those years, it is important to know whether the same macroeconomic imbalances also played a role in other crisis episodes. Table 11 depicts the model results without crises that occurred in 2007, 2008 and 2009. We find again that our results are overall robust to the exclusion of those years. Only household debt and, marginally so, the interacted house prices and credit growth variables become insignificant. Interestingly, the AUROC of 0.89 is even slightly higher than for the full sample.

A last remark on the identification of crises relates to the use of actual output. Some papers identify growth crises based on the output gap instead (e.g. Domonkos et al. 2017).

We do not follow this approach for two main reasons. First, our sample covers only advanced economies, in which recessions are usually also thought to be associated with contractions in actual output. Second, output gap measures based on standard filtering techniques (such as the HP filter) are known to suffer from substantial revisions over time and end-point problems.

We now scrutinise the sensitivity of the results to the sample of countries selected. To recall, a key motivation for this paper is to build a macroeconomic vulnerability model for euro area countries. However, the baseline regression includes a broader sample of 32 OECD and EU countries, since we wanted to feed the model with a more complete list of crisis experiences. This notwithstanding, to check the robustness of our results, we now run the same regression confined to euro area countries. As demonstrated by Table 12, the results are overall robust, although the current account balance becomes insignificant. The AUROC is very similar, with a value of 0.82.

As evident from our baseline model, not all variables listed in Table 3 made it into our model, since they turned out to be insignificant in a multivariate context. For completeness, however, we also integrate those other candidate variables in Table 13. As an alternative to the VIX, we experimented with country-specific measures of financial market volatility and uncertainty. However, neither the 10-year government bond spread to the German bund nor the volatility in 10-year government bond yields turned out to be significant. The same applies, not surprisingly, to the unemployment rate, as it usually lags output developments. The NIIP also turned out to be non-significant. Its debt component is in any case already covered by the indebtedness of the three sectors of the economy, i.e. households, NFCs and the sovereign. Fiscal policies, as captured by the change in the budget balance, were also not found to be significant. The same is true for private consumption and construction investment growth. In particular the latter, though, is partly covered by the house prices and credit growth variables.

Another set of robustness checks relates to the approach taken towards post-crisis years. As discussed in Section 2.2.1, there are various ways to deal with a possible post-crisis bias. We have followed among others Demirgüç-Kunt and Detragiache (1998) by dropping the first post-crisis year. To gauge the sensitivity of our results to this approach, we employ two robustness checks. The first relates to choosing more event-specific numbers of years that are dropped. The advantage of our crisis identification approach using the BBQ algorithm of Harding and Pagan (2002) is that next to the crisis start year we also obtain the end year in which output reaches a trough. This allows us to drop all years after the crisis year in which the BBQ algorithm has not yet identified a turning point. Interestingly, in the large majority of cases, our initial assumption holds, i.e. the business cycle turns again after the first year of a crisis. Table 14 shows in waterfall format that our results are very robust to this alternative approach, with the exception of the interacted credit and house price growth variable, that is borderline insignificant.

A second sensitivity analysis regarding the post-crisis bias is a multinomial logit, as proposed by Bussière and Fratzscher (2006). They argue that by dropping observations the forecaster ignores data that might offer valuable information. In this spirit, we run a multinomial logit. For the multinomial logit, we need to define two states of the world in addition to tranquil times (in which the variable takes the value 0). First, we assign the value 1 to the first crisis year, in line with Section 2.2.1. Second, for the multinomial logit we define the post-crisis phase as the year following the crisis (i.e. those years receive the value 2). Alternatively to just defining the year succeeding the start of the crisis we again use the BBQ-identified post-crisis years (as just described above and shown in Table 14). However, this does not qualitatively change the results.[8] Table 15 in the appendix summarises the regression results using a multinomial logit. The regression results are overall not substantially different in the two non-tranquil states of the world. We also find that the in-sample performance of the multinomial logit, as reflected in the type I and type II errors, is broadly comparable to that of the binomial baseline regression.

Firth (1993) and King and Zang (2001) have shown that logistic regressions might misspecify the probability of rare events. The maximum likelihood method used to estimate logistic regressions has desirable properties such as consistency, efficiency and normality. However, for those properties to hold, the sample size should be relatively large. In rare event studies, a small sample bias exists in logistic regressions. Both Firth (1993) and King and Zang (2001) have developed methods to account for this bias by applying penalised likelihood techniques. The results are displayed in Table 16. Again, the results from the benchmark model are overall confirmed, with the small exception of the current account variable that reaches a p-value of 11% and thus would be just insignificant.

Finally, we have checked the sensitivity of our results to outliers and influential observations. In particular, the estimation results remain broadly unchanged if we exclude all observations with a leverage above 0.1.[9]

4 Country results

Our paper has two main objectives. First, we want to demonstrate that a macroeconomic vulnerability model is a useful complement to existing early warning models. Second, we are interested in the extent to which our model would have been able to predict the severe recessions that occurred in several euro area countries in the wake of the global financial crisis and the euro area crisis. In line with the second objective, we now look through the lens of our model at the macroeconomic vulnerabilities in euro area countries at that time.

[8] The results are therefore omitted but available from the authors on request.

[9] The leverage indicator proposed by Pregibon (1981) is able to detect observations with particularly high influence on the logit estimates. The adjusted estimation results are not reported here but available on request.

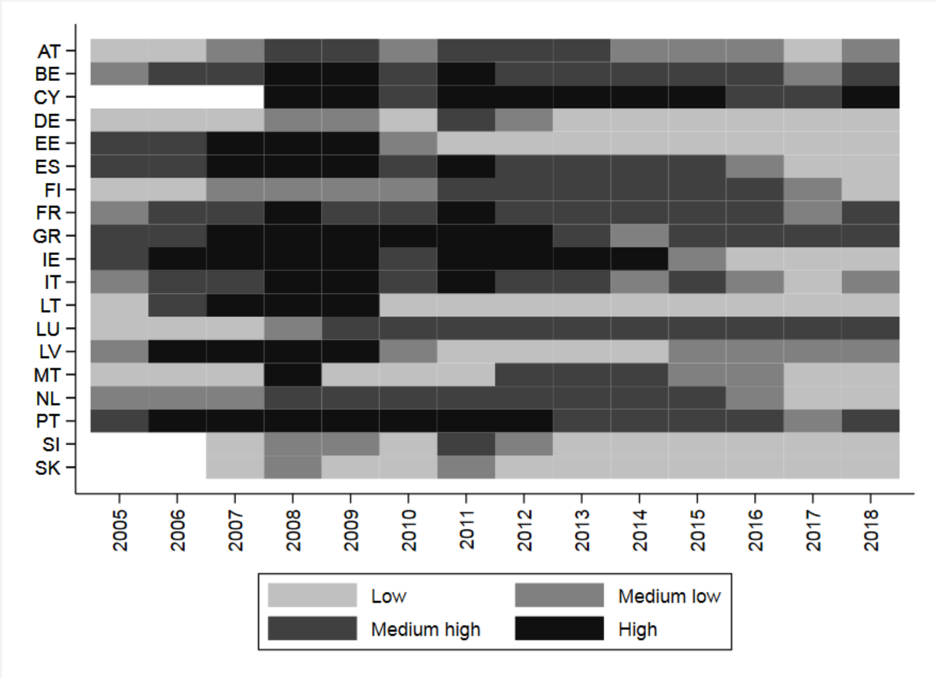
For illustration purposes, we create a heat map translating the model-implied crisis probabilities into colour-coded vulnerability scores, with darker colours corresponding to more severe macroeconomic vulnerabilities. We consider four vulnerability categories (low, medium low, medium high and high). The underlying probability thresholds of 1.5%, 3% and 10% are consistent with the range of thresholds identified by standard selection criteria (Table 6). As already noted in Section 3.3, the graduated thresholds reflect the idea that policymakers are often interested in tentative signs of emerging vulnerabilities, even if they are not yet strong enough to warrant a fully-fledged alarm signal. In any case, the thresholds should not be interpreted as sharp watersheds, but rather as indicative yardsticks in a continuous, albeit non-linear, mapping of macroeconomic imbalances into crisis probabilities.

The resulting heat map in Figure 3 allows us to compare macroeconomic vulnerabilities across countries and time. It shows that the model would have pointed early on to emerging macroeconomic vulnerabilities in many euro area countries, in particular those that later requested international financial assistance.[10] Table 8 confirms that the future programme countries (i.e. Cyprus, Spain, Greece, Ireland and Portugal) would have been assigned to the two highest risk categories already well before the severe recessions started (often in the form of a so-called double dip). The same applies to the Baltic countries, which experienced some of the most severe recessions. Also for most of the other euro area countries, perceptible albeit less forceful alarm signals would have been issued already a few years before the start of their crises. Hence, the model would have pointed to the build-up of significant macroeconomic vulnerabilities in several euro area countries early on. This, in turn, would have in principle allowed policy makers to take corrective action.

The heat map also gives a good impression of the degree of macroeconomic adjustment that took place after 2011 in most euro area countries. On the back of shrinking flow imbalances, the model-implied crisis probabilities gradually declined. While the model results would indeed point to an uptick in some crisis probabilities in 2018, it is important to put this into perspective. Figure 4 compares the model-implied crisis probabilities for 2018 with those seen a decade earlier. Some of the current euro area countries faced crisis probabilities well above 10% in 2008. By comparison, the crisis probabilities for 2018 (which are based on data for 2015-18) are significantly lower. Again, this reflects the macroeconomic rebalancing seen over the past decade, in particular the correction of flow imbalances. Notwithstanding this, there is significant heterogeneity across countries, partly due to persistent stock imbalances in the countries that underwent particularly severe crises.

[10]It should be borne in mind that the model-implied crisis probabilities for year t are only based on data up to year $t-1$. Thus, they would have been available at the beginning of year t. While the VIX enters the model contemporaneously, it is available at high frequency and can therefore be updated in real time.

Figure 3: Macroeconomic vulnerabilities in euro area countries over time



Note: The heat map shows the model-implied crisis probabilities, grouped into four categories. Darker colours signal more severe macroeconomic vulnerabilities. The underlying probability thresholds are 1.5%, 3% and 10%.

Figure 4: Crisis probabilities before the global financial crisis and more recently

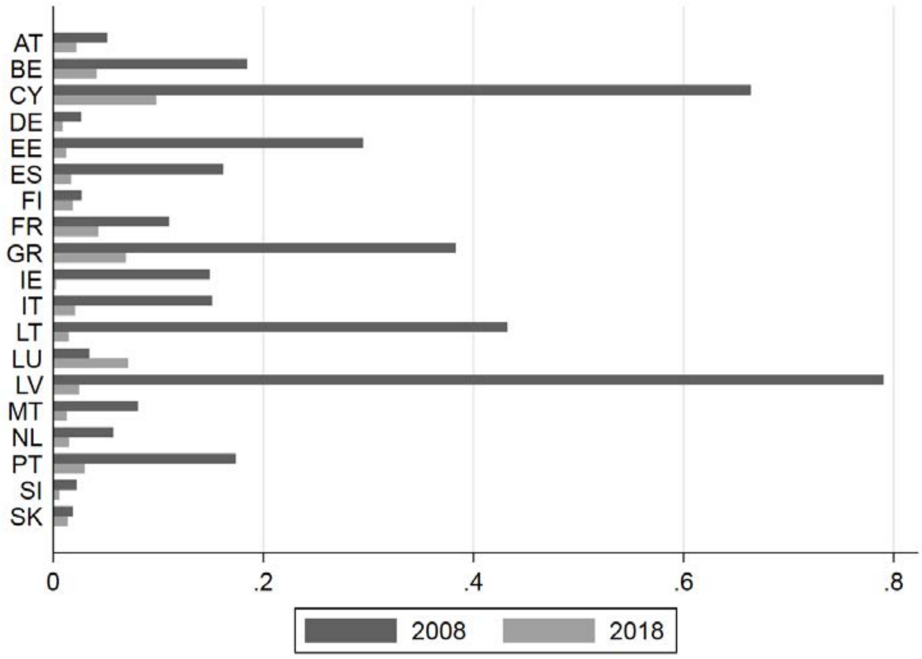


Table 8: Growth crises in euro area countries (2005-2017)

Country	Crisis start	Crisis end	Peak-to-trough change in GDP (%)	Vulnerability score in crisis year	Earliest signal
AT	2008	2009	-5.2	Medium high	2008
BE	2008	2009	-3.9	High	2006
CY	2008	2009	-2.7	High	*na*
CY	2011	2014	-11.6	High	2010
DE	2008	2009	-7.2	Medium low	-
EE	2008	2009	-23.4	High	2005
ES	2008	2009	-4.7	High	2005
ES	2010	2013	-5.9	Medium high	2009
FI	2008	2009	-10.5	Medium low	-
FI	2012	2013	-2.7	Medium high	2011
FR	2008	2009	-4.1	High	2006
GR	2007	2013	-32.1	High	2005
IE	2007	2009	-10.5	High	2005
IT	2008	2009	-8.3	High	2006
IT	2011	2013	-5.3	High	2009
LT	2008	2009	-18.4	High	2006
LU	2008	2009	-8.4	Medium low	-
LV	2007	2010	-25.8	High	2006
MT	2008	2009	-5.2	High	2008
NL	2008	2009	-4.6	Medium high	2008
PT	2008	2009	-4.4	High	2005
PT	2010	2012	-8.4	High	2009
SI	2008	2009	-10.0	Medium low	-
SI	2011	2013	-4.8	Medium high	2011

Note: The last column reports the first post-2005 year in which the vulnerability score reached either the level ”medium high” or ”high”. (-) indicates that this never happened prior to the crisis. (na) indicates missing observations. For countries with two crises and continuous alarm signals in-between, the end of the first crisis is reported. Recessions with a peak-to-rough decline below 2.5% are not reported, in line with our definition of growth crises.

5 Conclusions

Macroeconomic imbalances increase the vulnerability of an economy to adverse shocks. When such a shock eventually hits the economy, the economic and social costs are amplified given the previously accumulated imbalances.

We have developed a model that gauges the likelihood of such crises related to macroeconomic vulnerabilities. Specifically, we have matched past growth crises with patterns of macroeconomic imbalances spreading over a sample of 32 OECD and EU economies and almost 40 years.

Our model is not meant to replace existing early warning models focusing on different crises, such as banking or sovereign debt crises. By contrast, our model is meant to complement the existing suite of models. After all, early warning models are most successful when limited to a specific type of crisis.

We employ a multivariate discrete choice model that links the likelihood of growth crises to a set of explanatory indicators capturing domestic and external imbalances. Our approach allows us to look at patterns of indicators and their interactions. Most indicators cannot be seen in isolation but often pose risks only when they flash simultaneously.

We show that our model would have been able to identify, on the basis of macroeconomic indicators, the euro area countries that experienced particularly severe recessions post-2007. In most cases, and in particular for the future EU-IMF programme countries, early warning signals would have been issued several years in advance.

Looking at more recent data, the model suggests that most euro area countries have achieved significant macroeconomic adjustment over recent years, which has strongly reduced the likelihood of a growth crisis going forward. However, the crisis probabilities remain heterogeneous across euro area countries, partly due to persistent stock imbalances in some countries.

Our model can be seen as a quantitative complement to the scoreboard used in the context of the EU's Macroeconomic Imbalance Procedure. Most of the 14 variables included in the scoreboard are mirrored in our macroeconomic vulnerability model. While the scoreboard only offers qualitative information, our early warning model allows us to calculate a quantitative vulnerability score for each country at any point in time, taking into account interdependencies between the relevant indicators.

References

Adalid, R. and Detken, C. (2007). Liquidity shocks and asset price boom/bust cycles. *ECB Working Paper Series*, No. 732.

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*, 27:520–533.

Alessi, L. and Detken, C. (2014). On policymakers loss functions and the evaluation of early warning systems: Comment. *Economics Letters*, 124(3):338–340.

Alessi, L. and Detken, C. (2018). Identifying excessive credit growth and leverage. *Journal of Financial Stability*, 35:215 – 225.

Babecký, J., Havránek, T., Matěju, J., Rusnák, M., Šmídková, K., and Vašíček, B. (2014). Banking, debt, and currency crises in developed countries: Stylized facts and early warning indicators. *Journal of Financial Stability*, 15(C):1–17.

Basu, S., Chamon, M., and Crowe, C. (2017). A model to assess the probabilities of growth, fiscal, and financial crises. *IMF Working Paper*, No. 282.

Bekaert, G., Hoerova, M., and Lo Duca, M. (2010). Risk, uncertainty and monetary policy. *NBER Working Paper*, No. 16397.

Berg, A. and Coke, R. N. (2004). Autocorrelation-corrected standard errors in panel probits: An application to currency crisis prediction. IMF Working Paper 04/39, International Monetary Fund.

Berg, A. and Pattillo, C. (1999a). Are currency crises predictable? A test. *IMF Staff Papers*, 46(2):108–138.

Berg, A. and Pattillo, C. (1999b). Predicting currency crises: The indicators approach and an alternative. *Journal of International Money and Finance*, 18(4):561–586.

Berger, H. and Nitsch, V. (2010). The euros effect on trade imbalances. *IMF Working Paper*, No. 10/226.

Berkmen, P., Gelos, G., Rennhack, R., and Walsh, J. (1999). The Global Financial Crisis: Explaining cross-country differences in the output impact. *IMF Working Paper*, No. 09/280.

Beutel, J., List, S., and von Schweinitz, G. (2019). An evaluation of early warning models for systemic banking crises: Does machine learning improve predictions? IWH Discussion Papers 2/2019, Halle Institute for Economic Research (IWH).

Blanchard, O. and Giavazzi, F. (2002). Current account deficits in the euro area: The end of the Feldestein-Horika puzzle? *Brookings Papers on Economic Activity*, 33(2):147–186.

Blanchard, O. and Milesi-Ferretti, G. (2009). Global imbalances: In midstream? *IMF Staff Position Notes*, 09(29).

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.

Candelon, B., Dumitrescu, E. I., and Hurlin, C. (2012). How to evaluate an early-warning system: Toward a unified statistical framework for assessing financial crises forecasting methods. *IMF Economic Review*, 60(1):75–113.

Carlin, W., Glyn, A., and van Reenen, J. (2001). Export performance of OECD countries: An empirical examination of the role of cost competitiveness. *The Economic Journal*, 111:128–162.

Catão, L. and Milesi-Ferretti, G. (2014). External liabilities and crises. *Journal of International Economics*, 94:18–32.

Chen, R., Milesi-Ferretti, G. M., and Tressel, T. (2012). External imbalances in the euro area. *IMF Working Paper*, No. 12/236.

Christofides, C., Eicher, T., and Papageorgiou, C. (2016). Did established early warning signals predict the 2008 crises? *European Economic Review*, 81:103–114.

Cramer, J. S. (1999). Predictive performance of the binary logit model in unbalanced samples. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 48(1):85–94.

Cuaresma, J. C. and Slacik, T. (2009). On the determinants of currency crises: The role of model uncertainty. *Journal of Macroeconomics*, 31(4):621 – 632.

Dawood, M., Horsewood, N., and Strobel, F. (2017). Predicting sovereign debt crises: An early warning system approach. *Journal of Financial Stability*, 28:16 – 28.

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

Domonkos, T., Ostrihoň, F., Šikulová, I., and Širaňová, M. (2017). Analysing the relevance of the MIP scoreboard's indicators. *National Institute Economic Review*, 239(1):32–52.

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1):2738.

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

Frankel, J. and Saravelos, G. (2012). Can leading indicators assess country vulnerability? Evidence from the 2008-09 global financial crisis. *Journal of International Economics*, 87:216–231.

Fuertes, A.-M. and Kalotychou, E. (2007). Optimal design of early warning systems for sovereign debt crises. *International Journal of Forecasting*, 23(1):85–100.

Gochoco-Bautista, M. S. (2000). Periods of currency pressure: Stylized facts and leading indicators. *Journal of Macroeconomics*, 22(1):125 – 158.

Harding, D. and Pagan, A. (2002). Dissecting the cycle: A methodological investigation. *Journal of Monetary Economics*, 49(2):365 – 381.

Hermansen, M. and Röhn, O. (2017). Economic resilience: The usefulness of early warning indicators in OECD countries. *OECD Journal: Economic Studies*, 2016(1):9–35.

Jorda, O., Schularick, M., and Taylor, A. (2015). Leveraged bubbles. *Journal of Monetary Economics*, 76:S1–S20.

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

Kaminsky, G. and Reinhart, C. (1999). The Twin Crises: The Causes of banking and balance-of-payments problems. *American Economic Review*, 89(3):473–500.

King, G. and Zang, L. (2001). Logistic Regression in Rare Events Data. *Political Analysis*, No. 9:137–163.

Knedlik, T. (2014). The impact of preferences on early warning systems: The case of the european commission's scoreboard. *European Journal of Political Economy*, 34(C):157–166.

Laeven, L. and Valencia, F. (2013). Systemic banking crises database. *IMF Economic Review*, 61(2):2225–270.

Lane, P. R. (2006). The real effects of European Monetary Union. *Journal of Economic Perspectives*, 20(4):47–66.

Lane, P. R. and McQuade, P. (2014). Domestic credit growth and international capital flows. *Scandinavian Journal of Economics*, 116(1):218–252.

Lane, P. R. and Milesi-Ferretti, G. (2011). The cross-country incidence of the global crisis. *IMF Economic Review*, 59(1):77–110.

Lane, P. R. and Pels, B. (2012). Current account imbalances in Europe. CEPR Discussion Papers 8958, C.E.P.R. Discussion Papers.

Lo Duca, M. and Peltonen, T. (2013). Assessing systemic risks and predicting systemic events. *Journal of banking and finance*, 37:2183–2195.

Mongelli, F. and Wyplosz, C. (2008). The euro at ten: Unfulfilled threat and unexpected challenges. In Mackowiak, B., Mongelli, F., Noblet, G., and Smets, F., editors, *The euro at 10: lessons and challenges*, pages 24–57. ECB.

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

Pierluigi, B. and Sondermann, D. (2018). Macroeconomic imbalances in the euro area: Where do we stand? *ECB Occasional Paper Series*, No. 210.

Prasad, A., Elekdag, S., Jeasakul, P., Lafarguette, R., Alter, A., Feng, A. X., and Wang, C. (2019). Growth at Risk: Concept and Application in IMF Country Surveillance. *IMF Working Paper*, No. 19/36.

Pregibon, D. (1981). Logistic regression diagnostics. *The Annals of Statistics*, 9(4):705–724.

Reinhart, C. and Rogoff, K. (2010). Growth in a time of debt. *American Economic Review: Papers & Proceedings*, No. 100:573–578.

Röhn, O., Caldera Sanchez, A., Hermansen, M., and Rasmussen, M. (2015). Economic resilience: A new set of vulnerability indicators for OECD countries. OECD Economics Department Working Papers No. 1249, OECD Publishing.

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

Savona, R. and Vezzoli, M. (2010). Fitting and forecasting sovereign defaults using multiple risk signals. *Oxford Bulletin of Economics and Statistics*, 77(1):66–92.

Schimmelpfennig, A., Roubini, N., and Manasse, P. (2003). Predicting sovereign debt crises. IMF Working Paper 03/221, International Monetary Fund.

Siranova, M. and Radvanský, M. (2018). Performance of the Macroeconomic Imbalance Procedure in light of historical experience in the CEE region. *Journal of Economic Policy Reform*, 21(4):335–352.

Sutherland, D. and Hoeller, P. (2012). Debt and macroeconomic stability: An overview of the literature and some empirics. *OECD Economics Department Working Paper*, No. 1006:573–578.

Trichet, J.-C. (2005). Testimony before the Committee on Economic and Monetary Affairs of the European Parliament. *Speech given on 21 November 2005.*

Trichet, J.-C. (2006). EMU after seven years: Successes and challenges. *Speech given on 5 May 2006.*

Zemanek, H., Belke, A., and Schnabl, G. (2009). Current account imbalances and structural adjustment in the euro area: How to rebalance competitiveness. *IZA Policy Paper*, No. 7.

Zorell, N. (2017). Large net foreign liabilities of euro area countries: Risks, prospects and policy options. *ECB Occasional Paper Series*, No. 198.

Appendices

A Description of the dataset

Country list: Australia, Austria, Belgium, Canada, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Japan, Latvia, Lithuania, Luxembourg, Malta, the Netherlands, New Zealand, Poland, Portugal, Slovakia, Slovenia, South Korea, Spain, Sweden, Switzerland, United Kingdom, United States.

The countries in central and eastern Europe (Czech Republic, Estonia, Hungary, Latvia, Lithuania, Slovakia, Slovenia and Poland) are excluded from the sample prior to 2000, i.e. during the transition period.

B Tables and charts containing robustness checks

Table 9: Crisis incidence - alternative crisis definitions

	BBQ	BBQ v2	BBQ v3	technical recession
threshold	*-2.5*	*-2.0*	*-1.5*	*-2.5*
Full sample (1980-2017)	53	66	80	53
1980s	9	12	17	9
1990s	8	12	14	7
2000s	28	31	33	29
2010s	8	11	16	8
Euro area countries	32	40	47	32

Note: The peak-to-trough change in the level of real GDP and the crisis length refer to the median episode.

Table 10: Alternative crisis definitions

	(1) BBQ v2	(2) BBQ v3	(3) technical recession
Compensation per employee (3-year growth)	0.055***	0.049***	0.067***
	(0.018)	(0.018)	(0.021)
Government debt (% of GDP) x VIX	0.001***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)
Current account balance (% of GDP)	-0.041	-0.048*	-0.062**
	(0.026)	(0.026)	(0.028)
Household debt (% of GDP)	0.009	0.003	0.014*
	(0.006)	(0.006)	(0.008)
Real house price growth x credit growth (3-year change)	0.003***	0.003***	0.003***
	(0.001)	(0.001)	(0.001)
NFC debt (% of GDP)	0.011***	0.013***	0.010***
	(0.002)	(0.002)	(0.003)
Export market share (3-year growth)	-0.043***	-0.044***	-0.058***
	(0.013)	(0.012)	(0.018)
Real effective exchange rate, HICP-deflated (3-year growth)	0.071*	0.060**	0.093**
	(0.038)	(0.029)	(0.037)
Constant	-6.296***	-5.926***	-7.140***
	(0.854)	(0.872)	(1.085)
Pseudo R^2	0.146	0.148	0.186
AUROC	0.8	0.8	0.8
Observations	733	733	733

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; clustered standard errors in brackets.
BBQ v2 (BBQ v3) refers to growth crises with a threshold of 2.0% (1.5%).

Table 11: Baseline regression excluding global financial crisis

Compensation per employee (3-year change)	0.100***
	(0.019)
Government debt (% of GDP) x VIX	0.001**
	(0.000)
Current account balance (% of GDP)	-0.182***
	(0.057)
Household debt (% of GDP)	0.002
	(0.011)
Real house price growth x credit growth (3-year change)	-0.002
	(0.001)
NFC debt (% of GDP)	0.017***
	(0.005)
Export market share (3-year growth)	-0.066***
	(0.024)
Real effective exchange rate, HICP-deflated (3-year growth)	0.157***
	(0.049)
Constant	-8.696***
	(1.521)
Pseudo R^2	0.228
AUROC	0.89
Observations	623

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; clustered standard errors in brackets.

Table 12: Baseline logit with euro area country sample

Compensation per employee (3-year growth)	0.066***
	(0.019)
Government debt (% of GDP) x VIX	0.001**
	(0.000)
Current account balance (% of GDP)	-0.043
	(0.032)
Household debt (% of GDP)	0.020**
	(0.009)
Real house price growth x credit growth (3-year change)	0.003**
	(0.001)
NFC debt (% of GDP)	0.008**
	(0.003)
Export market share (3-year growth)	-0.045***
	(0.015)
Real effective exchange rate, HICP-deflated (3-year growth)	0.187**
	(0.089)
Constant	-7.137***
	(1.386)
Pseudo R^2	0.207
AUROC	0.82
Observations	391

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; clustered standard errors in brackets.

Table 13: Other possible explanatory variables

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Compensation per employee (3-year growth)	0.040***	0.040***	0.067***	0.091**	0.081***	0.082***	0.084***
	(0.014)	(0.014)	(0.022)	(0.042)	(0.024)	(0.031)	(0.023)
Government debt (% of GDP) x VIX			0.001***	0.001***	0.001***	0.001***	0.001***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Current account balance (% of GDP)	-0.071**	-0.076**	-0.084**	-0.044	-0.059*	-0.052	-0.053*
	(0.034)	(0.033)	(0.033)	(0.042)	(0.032)	(0.038)	(0.031)
Household debt (% of GDP)	0.005	0.005	0.011	0.015	0.016*	0.020*	0.015*
	(0.006)	(0.006)	(0.010)	(0.010)	(0.008)	(0.011)	(0.008)
Real house price growth x credit growth (3-year change)	0.003***	0.002***	0.002**	0.002	0.003**	0.002	0.003**
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)
NFC debt (% of GDP)	0.008***	0.008***	0.010***	0.011**	0.010***	0.009	0.010***
	(0.002)	(0.002)	(0.003)	(0.005)	(0.003)	(0.009)	(0.003)
Export market share (3-year growth)	-0.046***	-0.047***	-0.038***	-0.051***	-0.042***	-0.061***	-0.048***
	(0.011)	(0.011)	(0.013)	(0.016)	(0.014)	(0.022)	(0.015)
Real effective exchange rate, HICP-deflated (3-year growth)	0.057	0.055	0.063*	0.116***	0.083**	0.085**	0.093***
	(0.038)	(0.037)	(0.034)	(0.038)	(0.034)	(0.035)	(0.036)
Government debt (% of GDP) x government bond spread	0.000						
	(0.000)						
Government debt (% of GDP) x government bond volatility		-0.001					
		(0.002)					
Unemployment rate (3-year change)			-0.113				
			(0.092)				
NIIP (in % of GDP)				-0.002			
				(0.006)			
Private consumption growth (3-year change)					-0.003		
					(0.034)		
Construction investment (3-year growth)						0.014	
						(0.034)	
Budget balance (3-year change)							0.005
							(0.031)
Constant	-4.628***	-4.576***	-6.167***	-7.734***	-7.522***	-7.857***	-7.562***
	(0.550)	(0.565)	(1.611)	(1.552)	(1.173)	(1.581)	(1.190)
Pseudo R^2	0.112	0.112	0.202	0.196	0.191	0.194	0.198
Observations	655	654	689	579	690	540	676

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; clustered standard errors in brackets.

Table 14: Benchmark logit with post-crisis years identified by BBQ-algorithm

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Compensation per employee (3-year growth)	0.017***	0.036***	0.021*	0.068***	0.085***	0.090***	0.103***	0.102***
	(0.006)	(0.012)	(0.012)	(0.014)	(0.018)	(0.018)	(0.021)	(0.021)
Government debt (% of GDP) x VIX		0.001***	0.001***	0.001***	0.001***	0.001***	0.001***	0.001***
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Current account balance (% of GDP)			-0.146***	-0.115***	-0.090**	-0.089***	-0.108***	-0.109***
			(0.043)	(0.033)	(0.037)	(0.034)	(0.039)	(0.040)
Household debt (% of GDP)				0.028***	0.031***	0.024***	0.020**	0.021**
				(0.007)	(0.009)	(0.008)	(0.009)	(0.009)
Real house price growth x credit growth (3-year change)					0.001	0.001	0.001	0.001
					(0.001)	(0.001)	(0.001)	(0.001)
NFC debt (% of GDP)						0.010***	0.012***	0.013***
						(0.003)	(0.003)	(0.004)
Export market share (3-year growth)							-0.044***	-0.050***
							(0.014)	(0.015)
Real effective exchange rate, HICP-deflated (3-year growth)								0.096***
								(0.037)
Constant	-3.181***	-4.340***	-4.488***	-7.095***	-7.589***	-8.164***	-8.307***	-8.543***
	(0.162)	(0.385)	(0.488)	(0.995)	(1.262)	(1.221)	(1.303)	(1.370)
Pseudo R^2	0.014	0.060	0.103	0.155	0.179	0.193	0.217	0.226
Observations	1009	803	803	733	673	672	672	672

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; clustered standard errors in brackets.

Table 15: multinomial logit

	Crisis	Post-crisis
Compensation per employee (3-year growth)	0.083***	0.079***
	(0.022)	(0.020)
Government debt (% of GDP) x VIX	0.001***	0.001***
	(0.000)	(0.000)
Current account balance (% of GDP)	-0.055*	-0.093***
	(0.033)	(0.029)
Household debt (% of GDP)	0.015*	0.014*
	(0.008)	(0.008)
Real house price growth x credit growth (3-year change)	0.002**	-0.002
	(0.001)	(0.002)
NFC debt (% of GDP)	0.011***	0.011***
	(0.003)	(0.003)
Export market share (3-year growth)	-0.043***	-0.022
	(0.014)	(0.020)
Real effective exchange rate, HICP-deflated (3-year growth)	0.080**	0.065*
	(0.039)	(0.036)
Constant	-7.435***	
	(1.193)	
Pseudo R^2	0.171	
Observations	733	

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; clustered standard errors in brackets.

Table 16: Penalised logistic regressions

	Firth (1993)	King/Zeng (2001)
Compensation per employee (3-year growth	0.077***	0.077***
	(0.017)	(0.020)
Government debt (% of GDP) x VIX	0.001***	0.001***
	(0.000)	(0.000)
Current account balance (% of GDP)	-0.056	-0.056
	(0.037)	(0.043)
Household debt (% of GDP)	0.015*	0.015
	(0.008)	(0.009)
Real house price growth x credit growth (3-year change)	0.002**	0.002**
	(0.001)	(0.001)
NFC debt (% of GDP)	0.011***	0.011**
	(0.004)	(0.004)
Export market share (3-year growth)	-0.040**	-0.040**
	(0.018)	(0.016)
Real effective exchange rate, HICP-deflated (3-year growth)	0.081	0.082**
	(0.053)	(0.037)
Constant	-7.311***	-7.306***
	(0.927)	(1.009)
Observations	690	690

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$; clustered standard errors in brackets.

Acknowledgements

We thank Jonne Lehtimäki for excellent research assistance. The paper has also benefited from comments of participants in two ECB and European Commission seminars. In particular, we thank Lucia Alessi, Jose González Mínguez, Klaus Masuch, Wolfgang Modery, Ralph Setzer, Alessandro Turrini, Andreas Westphal, Stefan Zeugner and an anonymous referee of the working paper series for helpful comments.

David Sondermann
European Central Bank, Frankfurt am Main, Germany; email: david.sondermann@ecb.europa.eu

Nico Zorell
European Central Bank, Frankfurt am Main, Germany; email: nico.zorell@ecb.europa.eu

Postal address 60640 Frankfurt am Main, Germany
Telephone +49 69 1344 0
Website www.ecb.europa.eu

PDF ISBN 978-92-899-3568-5 ISSN 1725-2806 doi:10.2866/967743 QB-AR-19-087-EN-N