ANALYSING THE RELEVANCE OF THE MIP SCOREBOARD'S INDICATORS

Tomáš Domonkos,* Filip Ostrihoň,** Ivana Šikulová** and Mária Širaňová***

The EU established an early warning system by introducing the Macroeconomic Imbalance Procedure (MIP) in the wake of the recent recession. Nevertheless, it has been found by some authors to be rather vague when launching the Excessive Imbalances Procedure. Performed analysis reflects on such views and treats the MIP indicators as a system while assessing the significance of all particular variables separately. This assessment was accomplished by applying a multivariate unbalanced logit model, utilising all 14 MIP headline indicators, using time horizons ranging from one to three years before crisis, which was represented by periods with output gap lower than negative 2 per cent. The approach was confronted with the estimates of a linear probability model to provide an idea about the robustness of the results. In the short term, activity rates, youth unemployment rates and private sector debt are the best performing indicators, complemented by current account balances in the long term.

Keywords: MIP scoreboard, Early warning systems, Binary response models, Logit

JEL Classifications: C25, C52, H12

1. Introduction

The economic and financial crisis followed by the debt crisis in 2009 revealed serious weaknesses in the governance framework of the European Monetary Union (EMU). In order to prevent possible future crises a complex governance reform was undertaken in the European Union (EU) in 2011. The reform included the introduction of a new procedure within the EU's annual cycle of economic policy guidance and surveillance (the European Semester) for preventing and correcting macroeconomic imbalances in the Euro Area – the so called Macroeconomic Imbalance Procedure (MIP). The aim of the new surveillance mechanism is to prevent the occurrence of asymmetric shocks by an early identification of potential risks, correcting existing imbalances that could lead to these shocks, and preventing them from re-emerging. In a monetary union with a single monetary policy, individual national fiscal policies constrained by the Stability and Growth Pact (SGP), and insufficiently flexible labour markets, the implementation of the enhanced surveillance mechanism seems to be inevitable if the Member States and the EU want to cope with potential asymmetric shocks (Essl and Stiglbauer, 2012).

The MIP legislation came into force in December 2011 as a part of the six-pack legislation which aims to reinforce the monitoring and surveillance of fiscal, macroeconomic and structural reform policies in the EU and the Euro Area compared to previously applied legislation. Thus, the MIP became an essential part of the European Semester in 2012. The MIP legislation consists of two regulations included in the six-pack: Regulation (EU) No 1176/2011 of the European Parliament and of the Council of 16 November 2011 on the prevention and correction of macroeconomic imbalances and Regulation (EU) No 1174/2011 of the European Parliament and of the Council of 16 November 2011 on enforcement measures to correct excessive macroeconomic imbalances in the Euro Area. While Regulation 1176/2011 covers all EU Member States and lays out the details of the surveillance procedure, Regulation 1174/2011 applies only on the Euro Area Member States and focuses on enforcement, including sanctions. Surveillance of macroeconomic imbalances is a part of broader efforts to move towards more integrated surveillance. It is supposed to go beyond the criteria on government debt and deficit of

*IER SAS and FSES CU in Bratislava; e–mail: tomas.domonkos@savba.sk. **IER SAS; e–mails: filip.ostrihon@savba.sk and ivana.sikulova@savba.sk. ***IER SAS and FNE EU in Bratislava; e–mail: maria.siranova@savba.sk. This research was financed under the Horizon 2020 programme, FIRSTRUN (grant 649261). We would like to thank the anonymous referees and the seminar participants at the National Bank of Slovakia for helpful suggestions.

the SGP. Therefore, the MIP is complementary to the SGP.

One of the criticisms of the MIP is its vague definition in terms of launching the Excessive Imbalances Procedure (EIP) resulting from the in-debt review (Kamps *et al*., 2013 and 2014).[1] They argue that this particular process of starting the corrective arm of the MIP is in fact subordinated to the judgement of the European Commission (henceforth referred to as 'the Commission'). In order to address this issue, we propose to enhance the decision-making process by taking into account additional information for each variable reflecting the entire MIP as one early warning system (EWS). This can be a useful complement to the current system based on individual signals and homogenous thresholds. A binary response model offers a way to assess the performance of individual indicators, while treating the MIP as one system. The magnitude and significance of average marginal effects might provide a fruitful insight, while evaluating the relevance of a signal issued by an individual indicator during the in-depth review. Moreover, the binary response models used are independent of any thresholds; thus, even if inappropriately set, they will not affect the results.

This paper estimates the average marginal contribution of individual MIP indicators to the overall probability of a recurrence of a crisis by means of a logit and linear probability model (LPM), which is meant as a robustness check for the logit results. In order to reduce the multicollinearity issue, factor analysis is utilised. We use data for EU28 countries with annual frequency. The composition of the MIP scoreboard is taken as given, and aims to evaluate the relevance and significance of all particular indicators as if they had been used to construct a common system. Additionally, three different time lags were tested to provide information about the dynamics of expected crises.

The key contribution of this research lies in complementing the discussion of how to expand the Alert Mechanism Report (AMR) and in-depth review stage before launching the EIP with an exact econometric model. Secondly, the use of factor analysis aims to address the multicollinearity issue by specifying which particular indicators tackle similar issues, so that policymakers may better understand the nature of individual signals. Thirdly, the results about the relevance of individual MIP indicators can be used while reassessing the composition of the scoreboard. Additionally, the results obtained may prove useful for the synthetisation of individual indicators into a common EWS, or one composite early warning indicator.

The paper is structured as follows. In the second section the MIP's background is elaborated. The third section discusses the data and methodology applied. Section four presents the results and discussion and finally in section five the conclusions are presented.

2. The Macroeconomic Imbalance Procedure

Similarly to the SGP, the MIP has two arms: a preventive and a corrective one. Within the preventive arm of the procedure, potential macroeconomic problems have to be identified and regularly analysed in order to detect emerging imbalances early on. The corrective arm provides the means effectively to enforce correction of imbalances and will come into effect if macroeconomic imbalances in a particular Member State prove to be excessive.[2] The preventive arm of the MIP consists of two steps: in the first step, an alert mechanism as an early warning system focuses attention on risks observed early on and identifies the Member States for which, in the second step, more in-depth analysis needs to be conducted to assess their vulnerability and substantiate policy recommendations, if appropriate (European Commission, 2012).

The MIP starts with the Alert Mechanism Report prepared by the Commission in November each year. The AMR covers all EU Member States not benefiting from financial assistance and is based on the so-called scoreboard, i.e., a set of fourteen (previously eleven, and initially ten) macroeconomic indicators of external imbalances, competitiveness, internal imbalances and the labour market with their threshold values. According to the Commission, the thresholds have been established using a statistical approach based on the distributions of the indicators' values, by identifying the thresholds as the lower and/or upper quartiles of the distributions, which are in line with the values found in the empirical literature (European Commission, 2012). Table 1 contains the indicators, their transformations and their indicative thresholds.

Using a larger set of indicators is basically in accordance with the general conclusion of Kaminsky, Lizando and Reinhart (1998) that an effective early warning system should consider a broad variety of indicators. The choice of the scoreboard indicators focuses on the most relevant dimensions of macroeconomic imbalances and competitiveness losses, with a particular emphasis on the smooth functioning of the EMU. Accordingly, the scoreboard includes both stock and flow indicators which can capture short-term deteriorations as well as long-term accumulation of imbalances.

As stated in the Occasional paper of the Commission on the scoreboard (2012), the choice of indicators is based

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

Table 1. Indicators of the MIP scoreboard, most recent version with fourteen indicators.

External imbalances and competitiveness		Indicative thresholds
Current account balance	% of GDP, 3 year backward moving average	+6% and –4%
Net international investment position	% of GDP	–35%
Real effective exchange rate	42 trading partners, HICP deflator, 3 years % change	±5% (EMU), ±11 (non–EMU)
Export market share	% of world exports, 5 years % change	–6%
Nominal unit labour cost	2010=100, 3 years % change	9% (EMU), 12% (non–EMU)
Internal imbalances		
House price index	deflated, 1 year % change	6%
Private sector debt	consolidated, % of GDP	133% (previously 160%)
Private sector credit flow	consolidated, % of GDP	14% (previously 15%)
General government gross debt	% of GDP	60%
Unemployment rate	3 year backward moving average	10%
Total financial sector liabilities	non–consolidated, 1 year % change	16.5%
New employment indicators		
Activity rate	% of total population aged 15–64, 3 years change	–0.2 p.p.
Long-term unemployment rate	% of active population aged 15–74, 3 years change	0.5 p.p.
Youth unemployment rate	% of active population aged 15–24, 3 years change	2.0 p.p.

Source: European Commission (2015c).

on the results found in the economic literature available at that time. It includes, in particular, the paper of Frankel and Saravelos (2010), who identified the causes and symptoms of financial crises that have been most consistent over time, country, and crisis. They found the real effective exchange rate (REER), the current account balance, credit growth and the level of external debt to be useful indicators for predicting the incidence of crises. Kaminsky, Lizando and Reinhart (1998), who examined the empirical evidence on currency crises and proposed a specific early warning system, had already identified the REER and domestic credit as indicators that proved to be particularly useful in anticipating crises. Later, Babecky *et al*. (2013) identified the most useful leading indicators by means of Bayesian model averaging and suggested that the current account balance to GDP ratio is robustly associated with the severity of crises (in line with Frankel and Saravelos, 2010), as well as private credit and the government debt-to-GDP ratio.[3]

In addition to the aforementioned REER, competitiveness development is monitored by another two headline indicators within the MIP. The first is the unit labour costs indicator, which (together with the REER indicator) allows a comprehensive assessment of cost/price competitiveness developments. The second is the export market share, which adds different aspects of competitiveness to the scoreboard that are not captured by price and cost competitiveness alone. Export performance was previously identified among the useful indicators in anticipating crises in Kaminsky, Lizando and Reinhart (1998).

The house price index has been chosen by the Commission as one of the headline indicators, since boom and bust in housing markets can be a source of macroeconomic imbalances. Later, Borgy, Clerc and Renne (2014) found evidence that house price booms are more likely to turn into a costly recession, or to trigger a banking crisis, than stock price booms. This finding is consistent with Barrell *et al*. (2010a), who showed that house price booms were a good leading indicator of banking crises, along with other indicators, in a logit early warning system framework.[4]

The reason for including the four labour market indicators (unemployment rate and the three new indicators) in the scoreboard is rather specific: monitoring these indicators helps us to understand better the social consequences of imbalances, including the correction phases of imbalances, and to fine tune policy recommendations under the MIP. It has been recently acknowledged that social and labour market indicators should not trigger the steps in the MIP procedure as the signals they potentially produce are not relevant for identifying macro financial risks (Council of the European Union 2016, §6).

From a methodological point of view, longer than annual averages or changes (over three or five years) are used in the calculations of several headline indicators, in order to capture their medium-term development and provide indications of the persistence of a potential imbalance. Several indicators, expressed as a share of GDP, are applied to allow for cross-country comparability, and differentiated thresholds are used for the EMU and non-EMU countries in cases of two indicators (REER

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

and nominal unit labour cost), given the different characteristics of the EU countries outside the monetary union.

The threshold values are not interpreted mechanically, but in conjunction with the accompanying qualitative analysis. The overall number of breaches of thresholds, the severity of individual breaches, as well as the combination of breaches potentially signalling broad based problems, is also taken into account (European Commission, 2011). The appropriateness of the scoreboard indicators is regularly reviewed by the Commission, in terms of the composition of indicators, the methodology used, and the indicative thresholds established. In accordance with the MIP legislation, it is possible to add new, or higher quality, indicators to the scoreboard, or, indeed, to replace some of the existing indicators.

Based on the reading of the scoreboard (headline indicators) in combination with relevant data beyond the scope of the scoreboard (auxiliary indicators for which no thresholds have been calculated), economic circumstances, and all relevant factors available specific to the country's situation, the Commission identifies the Member States that face risks of excessive imbalances. In these countries, closer analyses (so called in-depth reviews) are carried out. Following the in-debt reviews, the Commission determines whether imbalances exist in the Member States identified in the AMR and what is their nature.

In accordance with the Communication of the Commission 'On steps towards Completing Economic and Monetary Union' (European Commission, 2015b), the Commission has recently enhanced the transparency of the implementation of the MIP and stabilised the categorisation of macroeconomic imbalances by streamlining the number of imbalance categories from six to four: (1) No imbalance, (2) Imbalances, (3) Excessive imbalances, (4) Excessive imbalances with corrective action.

If the situation in a Member State is considered unproblematic, the Commission will not propose any further steps. If the Commission considers that macroeconomic imbalances exist, it issues policy recommendations on the correction of the imbalances to the Member State. However, if the Commission considers that there are excessive imbalances that may jeopardise the proper functioning of the EMU, it may recommend that the Council opens an EIP[5] as an enforcement mechanism, which falls under the corrective arm of the MIP.

After starting an EIP, the Member State concerned is obliged to submit a corrective action plan, based on a Council recommendation. The plan must contain adequate measures for the correction of the imbalances detected and specify deadlines for implementing the corrective action. In case of any contraventions, financial sanctions may be imposed for the EMU (but not for the non-EMU) Member States. The EIP will be terminated once the Council, based on a recommendation from the Commission, determines that the imbalances have been effectively eliminated.[6]

Several authors focus on weaknesses and possibilities of improving the MIP. Their suggestions relate to the choice of method used for computing the threshold values or indicators (Alcidi and Gros, 2014; Hallwirth, 2014), the need for some symmetry in the adjustment mechanism (De Grauwe, 2012), the single-country focus (Moschella, 2014), the relatively vague way of establishing excessive imbalances (Kamps *et al*., 2013) as well as the limited application of the RQMV (Moschella, 2014; Kamps *et al*., 2013). There is still ample room for further research that could contribute to the discussion of improving the MIP, in particular, by adjusting the overall design of the scoreboard of early warning indicators, which is allowed to evolve over time, while retaining its simplicity and clarity.

As the MIP can be considered an EWS there is another branch of the literature which focuses specifically on evaluation of EWS. It can, in general, be divided into univariate and multivariate according to the number of indicators simultaneously evaluated (Alessi *et al*., 2014). The two mainstream approaches currently applied are the signalling approach, proposed by Kaminsky and Reinhart (1999), or several versions of the binary response models which estimate the probability of crisis events, e.g. Antunes *et al*. (2014), Alessi *et al*. (2014), and Canova (1994). Alessi *et al*. (2014) point out that, regardless of the methods applied, the variable or model is evaluated in terms of the adjusted-noise-to-signal (aNtS) ratio, or a particular loss function of a policymaker. Other alternative criteria may be the number of correctly predicted crisis events. Kaminsky and Reinhart (1999), in their influential paper, discussed the signalling approach and used the aNtS ratio as a criterion. A similar approach is applied by Csortos and Szalai (2013) to the MIP scoreboard indicators with a simple loss function. The policymaker's loss function theory is discussed by Alessi and Detken (2011) and Sarlin (2013). Christensen and Li (2014) develop three composite indicators and, utilising the signal extraction approach, propose an EWS able to predict the probability

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

of financial stress incidence.[7] A model-free unified statistical framework aiming to assess crisis EWS is e.g. proposed by Dumitrescu, Hurlin and Candelon (2010).

3. Data and methodology

The quantitative part of this analysis relies on data published annually by Eurostat and included in the MIP scoreboard. The database used for the analysis consists of panel data featuring all 28 EU Member States, which for the assessment of predictive relevance of the MIP includes the headline indicators for three different time lags (one, two and three-year).

For the dependent variable, business cycle data, published by the Commission and covering the time period from 2001 to 2015, are used from the 2015 Autumn forecast (European Commission, 2015a).[8] As a crisis indicator we use the deviation of real GDP from potential GDP – i.e. output gap, estimated by the Commission according to production function method (European Commission, 2015a).

The definition of a crisis event was chosen based on the following assumptions: (1) currency crises are not relevant for the Euro Area Member States when focusing on proper functioning of the monetary union itself; (2) possible debt crises should be prevented primarily by the proper implementation of the SGP; and (3) financial/banking crises should be prevented, in the first place, by operation of the European System of Financial Supervision (ESFS), particularly the European Systemic Risk Board (ESRB) and the three European supervisory authorities (ESAs). Additionally, the frequency of crisis occurrences in our sample strongly resembles the frequency of crisis determined by the lower quartile approach often applied by the Commission when setting the threshold values for the MIP scoreboard indicators (European Commission, 2012).

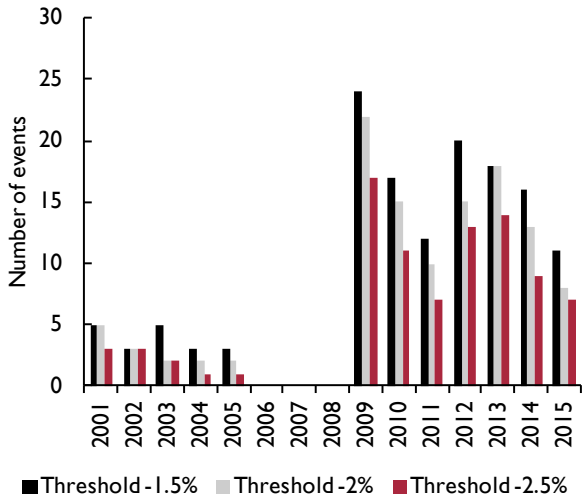
The lower quartile computed from the available output gap data for 28 EU Member States (except missing data for France in the years 2001–4 and for the UK in the years 2001–3) provided an indicative threshold value of –2.1 per cent. However, with slight changes in the sample years the threshold value tends to scale moderately upwards or downwards. In order to rely also on the empirical literature, the threshold value of –2 per cent was chosen for classification of crisis, which was also suggested by Csortos and Szalay (2013). Accordingly, the crisis event variable was subsequently defined to distinguish situations when the output gap is lower than –2.[9] Usage of dependent variable expressed as negative deviation from the average of the EU was also considered, but later dismissed as it would not capture crisis events which would affect the entire Union simultaneously.

As is apparent from figure 1, the distribution of crisis events in our sample is able to capture not only the most recent crisis period after the year 2008 but also idiosyncratic country-specific adverse disturbances in the pre-crisis period. Hence, defining the shock by output gap rather than the drop in GDP (figure 2) emanates more event signals potentially covering different types of crises. This further delivers more robustness by focusing on more than one type of negative economic shock.

For the analysis, we apply two different binary response models: the logit model and a standard linear probability model. The latter is used as a robustness check of the results estimated by the logit model. A lesser weight is, therefore, given to the LPM results during the evaluation because of its typical drawbacks concerning fitted probabilities greater than one, or lower than zero, and the constant partial effects of the explanatory variables (Wooldridge, 2013). The logit model is defined as $P(y=1|x)= G(\beta_0+x\beta)$ where G is the standard logistic cumulative distribution function (Wooldridge, 2013).

Given that the data for the 28 EU Member States were available over different time periods, the panel data estimation methods for imbalanced panels were employed. Assuming that cross-country heterogeneity is uncorrelated with the error term, as well as with explanatory variables, a pooled model was used both for the logit model and the LPM. The maximum likelihood estimation method was used for

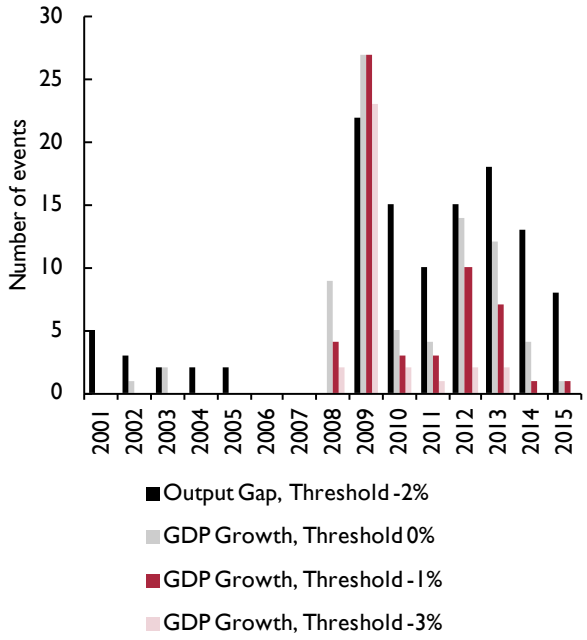
Figure 1. Crisis events based on output gap data



Source: Authors' calculations based on data from the Commission.

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

Figure 2. Crisis events based on output gap data and GDP growth data



Source: Authors' calculations based on data from the Commission.

the estimation of the logit model and ordinary least squares for the LPM. Potential serial correlation within clusters was accounted for by using cluster-robust standard errors, as was suggested by Cameron and Trivedi (2010). However, the possible attenuation bias for the logit model might be substantial in case of unobserved heterogeneity. Therefore the analysis was focused on average marginal effects or average partial effects, as designated by Wooldridge (2010), which quantify the change of response probability with the change of particular explanatory variable averaged across the observed population. According to Wooldridge (2010), these are consistently estimated even in such cases. The cluster-robust standard errors of the logit model's average marginal effects were computed by using the delta method. Descriptive statistics such as number of observations, correct prediction rate, and R-squared in case of LPM's or pseudo R-squared in case of logit models are presented with the results.

Both models applied were estimated using all fourteen indicators of the MIP scoreboard. The latent-variable specification for logit model is as follows:

$$y_{it}^{*} = \beta_0 + \beta_1 AR_{it} + \beta_2 LTUR_{it} + \beta_3 YUR_{it} + \beta_4 EMS_{it} + \beta_5 PSD_{it} + \beta_6 UR_{it} + \beta_7 GGD_{it} + \beta_8 NIIP_{it} + \beta_9 REER_{it} + \beta_{10} CA_{it} + \beta_{11} NULC_{it} + \beta_{12} PSCF_{it} + \beta_{13} TFSL_{it} + \beta_{14} HPI_{it} + e_{it} \quad (1)$$

The estimated linear probability model had the following specification:

$$y_{it} = \beta_0 + \beta_1 AR_{it} + \beta_2 LTUR_{it} + \beta_3 YUR_{it} + \beta_4 EMS_{it} + \beta_5 PSD_{it} + \beta_6 UR_{it} + \beta_7 GGD_{it} + \beta_8 NIIP_{it} + \beta_9 REER_{it} + \beta_{10} CA_{it} + \beta_{11} NULC_{it} + \beta_{12} PSCF_{it} + \beta_{13} TFSL_{it} + \beta_{14} HPI_{it} + e_{it} \quad (2)$$

The meanings of the abbreviations used in both specifications (1) and (2) are listed below:

AR	activity rate,
LTUR	long-term unemployment rate,
YUR	youth unemployment rate,
EMS	export market share,
PSD	private sector debt,
UR	unemployment rate,
GGD	general government debt,
NIIP	net international investment position,
REER	real effective exchange rate,
CA	current account balance,
NULC	nominal unit labour cost,
PSCF	private sector credit flow,
TFSL	total financial sector liabilities,
HPI	house price index.

Given that the estimated coefficients β are the same across all EU Member States and all observed periods, the estimated effects are describing general relationships of the explanatory variables with crisis probability across all EU countries over the entire examined period rather than country specific ones. Nevertheless, country specific performance of the models is presented in the Appendix, table A4, where the performance of the models is evaluated from the perspective of the PIIGS group of countries and Hungary.[10]

Due to the inclusion of many complementary indicators into the MIP system, the issue of potential multicollinearity ought to be tackled. The results (table 2) suggest that the identification of the marginal effects of some of the indicators examined might be affected by multicollinearity of high degree. To obtain an idea about the stability of the marginal effects – unaffected by the presence of multicollinearity – both models were estimated repeatedly focusing on one explanatory variable, while omitting other explanatory variables for which the pair-wise coefficient of correlation with the explanatory variable of interest was in absolute value higher than $\rho = 0.5$.

Average marginal effects obtained this way are listed under the estimates for the adjusted logit and LPM

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

Table 2. Correlation matrix of the MIP scoreboard indicators

	YUR	LTUR	UR	NIIP	EMS	NULC	TFSL	HPI	CA	REER	GGD	PSD	PSCF	AR
YUR	1.00													
LTUR	0.82	1.00												
UR	0.39	0.63	1.00											
NIIP	–0.33	–0.37	–0.59	1.00										
EMS	–0.24	–0.30	–0.09	–0.14	1.00									
NULC	–0.29	–0.52	–0.39	0.03	0.50	1.00								
TFSL	–0.39	–0.40	–0.31	0.15	0.37	0.33	1.00							
HPI	–0.52	–0.35	–0.18	0.27	0.15	0.02	0.51	1.00						
CA	–0.08	–0.02	–0.20	0.61	–0.38	–0.41	–0.09	0.16	1.00					
REER	–0.03	–0.26	–0.16	–0.04	0.38	0.37	0.02	–0.03	–0.20	1.00				
GGD	0.34	0.43	0.39	–0.16	–0.55	–0.42	–0.40	–0.23	–0.11	–0.26	1.00			
PSD	0.16	0.18	–0.15	0.14	–0.32	–0.11	–0.03	0.00	0.35	–0.15	0.07	1.00		
PSCF	–0.24	–0.20	–0.22	0.11	0.17	0.13	0.32	0.29	0.03	0.05	–0.22	0.21	1.00	
AR	–0.31	–0.29	–0.07	0.07	0.27	0.29	0.18	0.17	–0.08	0.07	–0.22	–0.21	0.03	1.00

Source: Authors' calculations based on Eurostat data. Note: Pair-wise coefficients of correlation with value higher than 0.5 are in red.

model. The adjusted model of latent variable for e.g., the long-term unemployment rate explanatory variable had the following specification:

$$y^{*}_{it} = \beta_0 + \beta_1 AR_{it} + \beta_2 LTUR_{it} + \beta_3 EMS_{it} + \beta_4 PSD_{it} + \beta_5 GGD_{it} + \beta_6 NIIP_{it} + \beta_7 REER_{it} + \beta_8 CA_{it} + \beta_9 PSCF_{it} + \beta_{10} TFSL_{it} + \beta_{11} HPI_{it} + e_{it} \quad (3)$$

In this specification, the possible collinear explanatory variables with the long-term unemployment rate (youth unemployment rate, unemployment rate, and nominal unit labour cost) were omitted. The adjusted models for all other explanatory variables were estimated following the same approach. If no other explanatory variable was sufficiently correlated, then only the estimates of the original model using all fourteen indicators were reported.

Additionally, factor analysis can be used to evaluate the nature of relationships among the indicators of interest, and to define groups of significantly correlated variables, but with not significantly correlated groups at the same time. Thanks to this approach, the large number of variables may be reduced to a smaller set of unobserved factors independent of one another with minimum loss of information (Stankovičová and Vojtková, 2007). These factors were subsequently applied as explanatory variables in the LPM and the Logit models. All of the mentioned models for one, two, and three-year lags are presented in the Appendix as table A1, table A2, and table A3, respectively. The grouping of variables into factors shows which variables address similar imbalances, and what is each particular indicator's contribution to a specific factor.

According to the unconditional correlation coefficients (table 2), the strongest link is among labour market related indicators, namely the long-term unemployment rate and the youth unemployment rate, as well as the unemployment rate and long-term unemployment rate. Medium significant correlation was found in the group of other variables including net international investment position and current account balance, nominal unit labour cost, house price index, gross government debt and export market share.

In order to mitigate the potential multicollinearity issues, we performed factor analysis on the correlation matrix while following the procedure described by Stankovičová and Vojtková (2007). We selected three factors based on the Screeplot (available upon request), the Kaiser-Meyer-Olkin measure, and the variance explained by eigenvectors. To facilitate the interpretation of the results, we rotated the factors by means of the orthogonal quartimax method. Relevant factors are extracted after the aforementioned rotation (table 3).

The first factor, labelled as 'Labour–Capital Nexus', combines labour market characteristics (youth, long-term and total unemployment rate) with a measure of external vulnerability captured by the net international investment position indicator. This factor integrates the development in the labour market with the behaviour of the net international investment position. In terms of the factor loadings, labour market variables have positive signs while the net international investment position enters negatively. This finding is consistent with the underlying data structure showing that countries with an increasing net international investment position through the export of capital are usually associated with

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

Table 3. Rotated factor loadings and unique variances

	Factor1	Factor2	Factor3	Uniqueness
Youth unemployment rate	0.54	–0.24	–0.59	0.31
Long-term unemployment rate	0.72	–0.43	–0.34	0.18
Unemployment rate	0.87	–0.18	–0.01	0.20
Net international investment position	–0.80	–0.31	0.14	0.24
Nominal unit labour cost	–0.33	0.83	–0.07	0.19
Export market shares	0.10	0.82	0.22	0.27
House price index	–0.16	–0.06	0.91	0.15
Total financial sector liabilities	–0.15	0.39	0.68	0.37

Source: Authors' calculations based on Eurostat data. Note: Factor loadings with value higher than 0.5 are marked in red.

lower unemployment rates (Germany or Netherlands). Conversely, countries with a deteriorating net international investment position are likely to record rises in unemployment rates (Cyprus or Ireland). Naturally, exceptions like Slovakia occur where a remarkably high negative net international investment position resulting from accumulated FDI inflows translates into an improvement in domestic unemployment rates (Hošková, 2001). The relationship of the unemployment and net international investment position might be connected to the hypothesis of demography-induced capital flows, as presented in Marchiori, Pierrard and Sneessens (2011). Ageing countries (e.g. Germany, Netherlands and Austria) tend to accumulate capital faster, which leads to higher labour productivity and decreases in unemployment and, at the same time, export capital due to a higher demand for capital in a higher working-age population ratio, which has a positive impact on their net international investment position. Analogously, importers of capital driven by their lower stage of ageing worsen their international investment position and also have simultaneously higher unemployment.

Two of the indicators belonging to the external imbalances and competitiveness group (nominal unit labour costs and export market share) are explained by the second factor. Positive signs associated with both factor loadings (table 3) confirm the existence of a positive linear relationship between the underlying indicators already noticed in the correlation analysis (table 2). We denote this factor as the 'Competitiveness and Catch-up Effect'. Since the early 1990s, the developing countries of the EU have made significant progress in economic convergence in the direction of the economically developed EU Member States. A recent report by the ECB (2015) even argues that the real convergence in the EU as a whole since 1999 might be attributed to the catching-up of Central and Eastern European economies, while a lack of convergence haunts the twelve countries that adopted the euro in 1999 and 2001. The Central and Eastern European Countries have not only been successful in increasing their world export market share, but also managed to both shift towards more sophisticated products outperforming the EU benchmark (Cazacu Bancu, 2015) and to profoundly change their export market structure to resemble their EU counterparts (Crespo and Fontoura, 2007). This shift in industrial structure is also associated with their convergence in income levels (Barrios, Barry and Strobl, 2002) and this is ultimately transmitted into growth rates in unit labour costs which are higher than those of their EU trade partners. Hence, both variables are positively correlated and obtain positive factor loadings.

The last of the three factors, entitled 'Real Estate Bubble', is composed of the house price index, and total financial sector liabilities, while sharing the young unemployment rate indicator with the first factor. This factor embodies the harmful developments in real estate markets and, at the same time, captures the vulnerability of young people to unemployment when a real estate market bubble burst may occur. Considering the argument of Hartmann (2015), boom–bust cycles in real estate markets are important determinants of financial crises. Such crises tend to aggravate unemployment severely among young people, as pointed out by Verick (2011).

It is important to mention that the application of factor analysis procedure was purely mechanical and no other information, assumptions, or researcher preferences were used for including particular indicators into the factors. The choice of whether a particular indicator was included in a potential factor was solely an outcome of the procedure described above.

4. Results and discussion

Firstly, we discuss the performance of individual indicators according to the logit models while treating the results of LPM models as a robustness check, and then confront the outcomes from the estimations using the results from

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

the factor analysis. Outcomes of the estimations for all time lags are reported in the Appendix (tables A1, A2 and A3). It is evident from the Appendix that the number of available observations drops substantially with the increase in lags of variables, as shown by the summary statistics provided at the bottom of the tables. The quality of the fit of models (as described by R-squared in case of LPM and pseudo R-squared in case of logit) seems also to vary substantially. Regardless of the model used, the approach explains less than half of the variation of crisis binary dependant variable. However, the in-sample correct prediction rate (computed by using a positive outcome threshold set to 0.5) is in all cases much higher than in a scenario without a model (i.e. a random model with expected correct prediction rate close to 0.5).

When evaluating the current account balance indicator, the size of the average marginal effect depends on the number of years preceding the crisis event. While completely insignificant one year before the occurrence of the crisis, lagging the variable by two-year periods turns the current account balance into the second-best performing indicator, once controlling for the presence of multicollinearity. A strong performance is also evident for three-year lags. The negative sign associated with the estimated coefficient confirms the general knowledge that deterioration in the external balance brings about an increasing risk of a crisis event, with there being an average marginal effect in the case of a two-year time lag at –0.022 for logit model when controlling for multicollinearity and having a tendency to grow in magnitude as the time horizon increases. Conversely, accumulation of surpluses in the external sector lowers the overall probability of negative events. The negative sign assigned to this parameter supports the question mark over the positive 6 percentage point threshold. This positive threshold may play an important role, if the stability of the Eurozone or the EU28 as a whole is evaluated but, according to the findings of this research, in the case of one particular member state it seems to be irrelevant.[11]

The net international investment position accompanies the current account balance in the group of external imbalances and competitiveness indicators and, by definition, reflects its behaviour. Thus, while the indicator performs badly in one-year lag estimations, increases in the time horizon yield an improvement in the statistical significance of coefficients achieving better results with an indicator preceding the crisis event by two or three years. On the other hand, the overall size of the average marginal effect (–0.002 to –0.003) places the indicator near the bottom of the virtual ranking of indicators, thus questioning its further usefulness in policymaking decision processes. Nevertheless, the effect is in line with the intuition that a higher investment position decreases the probability of a crisis event.

As in the case of its peers, the real effective exchange rate indicator almost copies the current account balance behaviour, scoring relatively high in terms of the size of the average marginal effect (–0.013 vs –0.018) depending on the model or length of time horizon (two to three years) used. Hence the favourable change in price and cost competitiveness ultimately materialises in a decrease in the probability of a crisis. As in the previous studies (e.g. Kaminsky, Lizando and Reinhart, 1998; Kaminsky and Reinhart, 1999), the REER is likely to serve as a reliable early warning indicator within the MIP system. This is not the case for the fourth external imbalance and competitiveness variable, the export market share; this is neither statistically significant nor economically important, no matter what the number of periods preceding the crisis event is.

The group of external imbalances and competitiveness indicators is complemented by the nominal unit labour cost indicator. Increases in the number of years preceding the predicted event – up to three years – delivers a straightforward improvement in its performance, turning an initially statistically insignificant result into a strong positive association between the labour costs measure and the probability of a crisis event (about 0.01). The deterioration of cost competitiveness not accompanied by a rise in labour productivity is likely to be associated with an increasing probability of an economic crisis, yet the transmission of this adverse development is likely to take a few years to materialise (three years).

The second group of MIP indicators evaluates economic conditions from the internal sector perspective and can be further broken down into labour-market and financial system-oriented indicators. Compared to the previous group, these variables, in general, tend to perform better in closer proximity, timewise, to the predicted event, with some notable exceptions.

The general government gross debt is among the most often discussed variable, potentially triggering the recent Eurozone debt crisis. A vast number of studies examined the complicated role of government debt in influencing economic growth, confirming the existence of a nonlinear relationship between the debt-to-GDP ratio and economic growth in general (e.g. Checherita and Rother, 2010). Additionally, literature dealing with the early warning indicators often includes this indicator in its list of determinants e.g. Alessi *et al*. (2014). Yet,

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

the outcomes of our study do not confirm either of those, as we robustly reject the existence of short-term and medium-term links between government debt and changes in the probability of economic crises, given the presence of other MIP indicators.[12] Despite the general acceptance of this variable by policymakers, in terms of our results its role as an early warning indicator may seem to be over-exaggerated. However, this may reflect the statement mentioned in section 2 (above), that the role associated with this indicator is, rather, to indicate the country's current position and vulnerability to economic shocks, instead of monitoring its financial stability. Regardless of these findings, the indicator has its merit and should be kept in mind when policy conclusions and austerity measures are drawn.[13]

On the opposite side, the private sector debt indicator delivers an expected contribution to the probability of a crisis event. A statistically significant and positive relationship in the estimates seems to hold robustly for all three examined time horizons. Higher private sector debt levels tend to be associated with a rise in vulnerability to economic shocks, thus driving the potential incidence of crisis. Thus, changes in the level of private indebtedness, rather than the so widely debated public sector debt, is what matters for an increase in crisis probability, according to evidence drawn from many decades of experience. Highlighted by Reinhart and Rogoff (2009) and confirmed by Taylor (2012) or Schularick and Taylor (2012), too much private debt can simply jeopardise an economy.

The house price index aims to capture the phenomenon when real property prices are increasing extremely rapidly over a relatively short period of time due to a potential occurrence of destabilising bubbles. According to our estimations, the link between the risk of an occurrence of a crisis and the behaviour of this indicator yields powerful results, especially for the one-year lag period (–0.033 for the logit model), while the effect slowly dissipates over the course of the second year (drop in size to –0.011), and disappears in the third year, when controlling for multicollinearity.[14] Yet, the negative sign associated with the coefficient surprisingly suggests that rising property prices results in lowering the probability of a crisis event, an outcome strongly contradicting recent historical experience and economic intuition. One possible explanation may be that the models implicitly check for increases in prices connected with the expansion of the financial system and, hence, only price increases that are driven by other factors are considered. Thus, an increase in housing prices may reflect an increase in specific market demand, which may be caused by, for example, migration into the country, indicating positive expectations for the future. As discussed in Babecky *et al.* (2013), a fall in house prices and share prices could, therefore, be considered a late early warning indicator. On top of that, a similar experience is shared by two other indicators to some extent. The total financial sector liabilities indicator, a variable mirroring possible adverse developments in the financial sector, does not stand out from this group. With a statistically significant yet negative sign for estimates preceding the event in one-year lag (–0.007), the link breaks down into no effects for two- and three-year time lags between the crisis and this explanatory variable. Private sector credit flow on the other hand appears to be negative (–0.004) and statistically significant only for the LPM model at one-year lag. Therefore, there is only slight evidence of any effect on crisis probability. Furthermore, the initially negative response to increases in private sector credit flow converges to zero very rapidly in the longer time horizon (2-year lag), where the effect remains with additional increases to the time horizon (3-year lag). A possible explanation for this result can be that an expansion of the domestic financial system, with an upsurge in the provision of credit is likely to stimulate economic activity in the short run, with no long-term effects.

Labour market-oriented MIP indicators include an aggregate measure of the unemployment rate, as well as three additional variables, each measuring one specific dimension of labour market imbalances. Once controlling for potential multicollinearity, the unemployment rate performs poorly in all specifications, across all time lags. On the other hand, the inclusion of new labour-market indicators introduces a more disaggregated view in the labour market and delivers more promising outcomes.

The activity rate consistently outperforms all other indicators, in terms of the size of the coefficient over all time lags. Regardless of the time lag, the marginal effect is negative, thus, a positive increase in the activity rate is associated with a drop in the probability of a crisis event. The stability of sign of the marginal effects together with magnitude of the effect almost doubling with two-year lag marks the activity rate as the most promising indicator among all MIP indicators.

As in the previous case, the economic rationale linking deterioration in the youth unemployment rate to the higher probability of a crisis is confirmed by a positive and statistically significant coefficient in all specifications, with the exception of three-year horizon (maximum for adjusted logit models at 0.011), once the

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

multicollinearity issue has been addressed. The overall size of the marginal effect places this indicator behind the activity rate and current account indicators.

Changes in the long-term unemployment rate ought to gauge future developments in the labour market, since longer duration of unemployment decreases considerably the prospects of re-employment. However, the results obtained for three-year lags seem to contradict this assumption, as the logit results suggest that increases in the long-term unemployment rate may decrease the probability of crisis, when controlling for multicollinearity, although they are not supported by LPM estimates. This aside, the empirical findings advocate against the capabilities of this indicator to contribute to a change in the probability of a crisis once controlling for multicollinearity. The results are to some extent in line with the general thought about the social and labour market indicators in the MIP mentioned earlier, that they are not relevant predictors of macro-financial risks.

In general, factor analysis aims to resolve potential multicollinearity issues among variables by identifying latent existing structures in the list of MIP indicators. Eight of the indicators proposed by the Commission tend, to a greater or lesser extent, to cluster into three distinct factors.

Firstly, marginal changes in the Labour–Capital Nexus factor positively influence the probability of a crisis event over a very short-term horizon (one year). Deterioration in labour market conditions, especially among the most sensitive group of young job seekers, accompanied by negative foreign financial exposure, tends to be associated with a higher frequency of crises (0.073 for the logit model). Conversely, a decrease in the probability of an adverse crisis event is likely to happen in countries with a net creditor position and competitively superior labour market conditions. As we move to longer time horizons, the factor quickly loses its statistical significance, even if some of its elements do not (i.e., Net international investment position, Youth unemployment rate, and Long-term unemployment rate which also constitute the third factor). Apparently, part of the explanatory power of individual indicators is still left untapped, rather than extracted through the factor analysis.

A very similar story might be told about the second factor, the Competitiveness and Catch-Up Effect. By its nature, it takes longer for an adverse drop in external competitiveness negatively to affect the probability of a crisis. While a rise in nominal unit labour costs firstly indicates an increase in the probability of a crisis at statistically significant levels three years before the occurrence of the event, the estimated coefficient associated with this factor remains statistically insignificant for all three specifications (Appendix tables A1, A2 and A3). Hence, even in this case, information related to the crisis event and stored in the variability of underlying elements does not fully transmit into the final common factor.

The behaviour of the Real Estate Bubble Burst factor resembles the performance of the underlying individual indicators. Substantial variations in real estate market may go hand-in-hand with deterioration of the financial sector. This can strongly affect the most vulnerable employees group, young job seekers, and morphs into an increasing probability of a crisis event a year later (average marginal effects of –0.222 for one-year horizon and –0.105 for two-year horizon). A relatively early indicator by its nature, the factor loses its statistical significance three years before the occurrence of a crisis event.

The results from the logit, LPM and factor analysis indicate that the activity rate, house price index, youth unemployment rate and private sector debt are the best performing individual indicators from the MIP scoreboard in the short term. The current account balance provides correct warnings rather more so in the long-term perspective. On the other hand, indicators of export market share, and the unemployment rate do not deliver the expected performance in terms of the models applied. Their poor contribution to the warnings, regardless of the time lag, raises the question of their future usefulness in the MIP surveillance process. The results gained from the factor analysis further suggest that several indicators may be jointly affected by common factors, such as demography, structure of the industry, and real estate bubbles. After including these factors into model specifications, some of their average marginal effects overshadowed those of individual indicators, especially in the one-year horizon. This certainly complicates the correct evaluation of the MIP scoreboard's warnings.

5. Conclusions

The evidence from recent development shows that economic crises are very costly and painful events, from both economic and social perspectives. Thus, possessing a tool able to predict such developments is highly appreciated by the EU and all particular Member States. The Commission's early warning system found

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

its place in the preventive arm of the MIP, relying on a scoreboard of indicators covering a wide range of potential imbalances. This early warning system attracted researchers and practitioners and received criticism, as well as suggestions, for improvement. Regardless of this, one question remains still open – if new crises occur, will it be able to warn effectively, and on time? Another important aspect is the ability of the Commission to enforce the implementation of the most appropriate reform measures by the Member States on imbalances.

This research has contributed to the EWS literature by applying logit, LPM, and factor analysis to extend the evaluation of signals from MIP's AMR and the in-depth review where the Commission decides about launching the EIP. This is a critically important moment, because false alarms and wrong decisions can lead to costly results. Specifically, unbalanced pooled logit models with cluster-robust standard errors were estimated via a maximum likelihood method, together with unbalanced pooled linear probability models also with cluster-robust standard errors, which were estimated by the ordinary least square method and served as a robustness check. This was done by confronting the results of the linear probability model with the corresponding average marginal effects, with standard errors computed using the delta method. One of the advantages of this econometric exercise is its independence from homogenous thresholds for all Member States, which may be, in some cases, inappropriate due to national specificities. Moreover, we showed that different indicators give warnings with different time lags. Policymakers can reflect this information about the dynamics of potential crises while defining the measures necessary to address their adverse development. The results gained from the estimations and the procedure applied may be a useful source of information for policymakers in their decision-making in particular during the preventive part of the MIP.

Another contribution of the analysis presented is the identification of a high degree of multicollinearity among some of the MIP Scoreboard's indicators. This multicollinearity issue was further developed via factor analysis, which uncovered links among specific indicators, which can be explained as unobservable factors. Three resulting factors were described based on their relation to the MIP indicators: Labour–Capital Nexus, Competitiveness and Catch-Up Effect, Real Estate Bubble Burst. The Labour–Capital Nexus in EU countries seems to be driven by demographic-induced capital flows. The Catch-Up, or Convergence, effects seem to resemble the shift of industrial structure associated with convergences in income levels and increases in export market share. The last factor of the Real Estate Bubble Burst resembled – by its effects – the outbreak of a financial crisis.

NOTES

1 By the in–debt review the European Commission aims to evaluate the nature and significance of macroeconomic imbalances in the Member States. It is usually launched when required by outcomes of the Alert Mechanism Report (AMR) or due to the occurrence of sudden negative economic development (European Commission, 2016).

2 According to the Regulation 1176/2011, a macroeconomic imbalance means "any trend giving rise to macroeconomic developments which are adversely affecting, or have the potential adversely to affect, the proper functioning of the economy of a Member State or of the economic and monetary union, or of the Union as a whole", while the excessive imbalances are defined as "severe imbalances, including imbalances that jeopardise, or risk jeopardising, the proper functioning of the economic and monetary union".

3 The general government debt is included into the scoreboard not to monitor risks of unsustainable public finances (covered by the SGP), but to offer a broader picture of a country's indebtedness (together with private sector debt), which can increase the vulnerability of a Member State to economic shocks. Similarly, the MIP looks at the financial sector from the perspective of macroeconomic imbalances, while the European Systemic Risk Board (ESRB) monitors financial stability risks.

4 Barrell *et al.* (2010a) found in their work that along with property prices bank capital adequacy and bank liquidity also have influence on banking crisis probabilities. Moreover Barrell *et al.* (2010b) showed, when analysing the predictability of subprime crisis, that along with property prices, bank capital adequacy and bank liquidity the current account also has influence on crisis probabilities.

5 So far the Council has never launched any EIP.

6 An essential innovation of the MIP procedure is the use of relative qualified majority vote (RQMV), under which a Council decision on a Commission recommendation regarding the activation of sanctions against Euro Area Member States is deemed to be adopted unless the Council decides by qualified majority to reject the recommendation within ten days. This semi-automatic decision-making procedure enhances the likelihood that the surveillance and enforcement process will not be blocked by political considerations.

7 Alternative composite early warning indicator was e.g. proposed by Berti, Salto and Laquien (2012).

8 The original 2015 Autumn forecast covers the period 2001–15, but, given that the intersect for all other indicators is starting in later periods, observations before 2005 are not used in the econometric estimations.

9 Csortos and Szalay (2013) applied the same method for a different dataset and achieved –2 per cent for the threshold of the dependent variable. They further point out in their analysis, however, that the HP filter endpoint error is usually very likely; it is not the exact size of the output gap which matters, but the sudden drop in GDP growth.

10 The country-specific results present the number of available observations, number of crisis events, correct prediction rate, correct prediction rate of crisis events, and adjusted noise to signal ratio, which is an indicator addressing the distribution of

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

Type I and Type II prediction errors – their further description can be found in Csortos and Szalai (2013). Similarly, evaluation of uniform positive outcome threshold set to 0.5, which served for classification of the predictions of particular models, is provided in the Appendix, tables A5 and A6. The threshold was evaluated in regard to every EU Member State, whether it provides better predictions of crisis periods or tranquil periods.

11 In the two-country example, it holds true that one country's current account surplus must be reflected in the second country's current account deficit. Yet, as in the multi-country case, one relatively sizeable current account surplus might be distributed among other countries' current account deficits in a relatively uniform way, the usefulness of recently imposed limit on current account surpluses might be called into question. One might argue that the question of concentration and persistence of current account deficits, rather than issue of surpluses, is of crucial importance.

12 Babecky *et al*. (2013) even report a positive link between decreasing government debt and the occurrence of banking crises.

13 Ultimately, individual fiscal policy is equally as responsible as monetary policy for mitigating the possible consequences of a crisis, both through automatic stabilisers and discrete anti-crisis measures, especially in the EA region. If, due to any limitations imposed by government stakeholders, the proper functioning of counter-cyclical fiscal policy is reduced, adverse crisis effects might become even more pronounced. Thus, the search for fiscal austerity, even without empirical evidence in favour of existing links between changes in crisis probability and government debt, might be justified if this serves to ensure that a sufficiently large fiscal space is available if needed.

14 Barrell *et al*. (2010a) found the house price index a leading indicator with a three-year time lag. They argue that this indicator captures a lending issue as a typical aftermath of house price bubbles, thus, it should be considered with a longer time lag. We did not find the house price index using our data sample a leading indicator with three years time lag. However, following the approach of Barrell *et al*. (2010a) and allowing longer time lags for this particular indicator might improve the model.

REFERENCES

Alcidi, C. and Gros, D. (2014), 'Implications of EU governance reforms. Rationale and practical application', Etla Reports, no. 25.

Alessi, L., Antunes, A., Babaecky, J., Baltussen, S., Behn, M., Bonfim, D., Bush, O., Datken, C., Frost, J., Guimaraes, R., Havránek, T., Joy, M., Kauko, K., Matěju, J., Monteiro, N., Neudorfer, B., Peltonen, T., Rodrigues, P.M.M., Rusnák, M., Schudel, W., Sigmund, M., Stremmel, H., Šmídková, K., Van Tilburg, R., Vašíček, B. and Žigarová, D. (2014), 'Comparing different early warning systems: results from a horse race competition among members of the macro-prudential research network', preliminary and incomplete draft. Available at: < http://www.ecb.europa.eu/events/pdf/conferences/140623/Vasicek-et-al_Comparing-Different-Early-Warning-Systems.pdf?F96bbb525a26071ecf97f9154fb3cc73>.

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, pp. 520–33.

Antunes, A., Bonfim, D., Monteiro, N. and Rodrigues, P.M.M. (2014), 'Early warning indicators of banking crises: exploring new data and tools', Banco de Portugal, *Economic Bulletin*.

Babecky, J., Havránek, T., Matějů, J., Rusnák, M., Šmídková, K. and Vašíček, B. (2013), 'Leading indicators of crisis incidence: evidence from developed countries', *Journal of International Money and Finance*, 35, pp. 1–19.

Barrell, R., Davis, E.P., Karim, D. and Liadze, I. (2010a), 'Bank regulation, property prices and early warning systems for banking crises in OECD countries', *Journal of Banking and Finance*, 34, pp. 2255–64.

—(2010b), 'Was the subprime crisis unique? An analysis of the factors that help predict banking crises in OECD countries', NIESR Discussion Paper No. 363.

Barrios, S., Barry, F. and Strobl, E. (2002), 'FDI and structural convergence in the EU periphery', Centre for European Policy Research, Workshop and Conferences, no. 2319.

Berti, K., Salto, M. and Laquien M. (2012), 'An early-detection index of fiscal stress for EU countries', *European Commission Economic Papers*, no. 475.

Borgy, V., Clerc, L. and Renne, J. P. (2014), 'Measuring aggregate risk: can we robustly identify asset-price boom-bust cycles?', *Journal of Banking and Finance*, 46, pp. 132–50.

Cameron, A.C. and Trivedi, P.K. (2010), *Microeconometrics Using Stata*, Revised Edition, Stata Press.

Canova, F. (1994), 'Were financial crises predictable?', *Journal of Money, Credit and Banking*, 26, 1, pp. 102–24.

Checherita, C. and Rother, P. (2010), 'The impact of high and growing government debt on economic growth: an empirical investigation for the Euro Area', ECB Working Paper, no. 1237.

Christensen, I. and Li, F. (2014), 'Predicting financial stress events: a signal extraction approach', *Journal of Financial Stability*, 14, pp. 54–65.

Cazacu (Bancu), A.-M. (2015), 'Export performance of central and eastern European countries: macro and micro fundamentals', *Procedia – Social and Behavioral Sciences*, 195, pp. 514–23.

Council of the European Union (2016), 'Outcome of proceedings, alert mechanism report 2106 – ECOFIN Council conclusions', 15 January, http://data.consilium.europa.eu/doc/document/ST–5318–2016–INIT/en/pdf.

Crespo, N. and Fontoura, M.R. (2007), 'Integration of CEECs into EU market: structural change and convergence', *Journal of Common Market Studies*, 45, 3, pp. 611–32.

Csortos, O. and Szalai, Z. (2013), 'Assessment of macroeconomic imbalance indicators', *MNB Bulletin*.

De Grauwe, P. (2012), 'In search of symmetry in the Eurozone', CEPS Policy Brief, no. 268.

Dumitrescu, E.I., Hurlin, CH., Candelon, B. (2010), 'How to evaluate an early warning system: towards a unified statistical framework for assessing financial crises forecasting methods', available at < https://hal.inria.fr/file/index/docid/450050/filename/EWS_candumhur.pdf>.

Essl, S. and Stiglbauer, A. (2012), 'Prevention and correction of macroeconomic imbalances: the excessive imbalance procedure', *Monetary Policy and The Economy Q4/11*, pp. 99–113.

European Central Bank (2015), 'Real convergence in the euro area: evidence, theory and policy implications', *ECB Economic Bulletin*, Issue 5, pp. 30–45.

European Commission (2011), 'Commission staff working document. Scoreboard for the surveillance of macroeconomic imbalances: envisaged initial design', Brussels, 8.11.2011, SEC (2011) 1361 final.

—(2012), 'Scoreboard for the surveillance of macroeconomic imbalances', *European Economy Occasional Papers*, no. 92.

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

—(2015a), 'Cyclical Adjustment of Budget Balances', Autumn 2015.
—(2015b), 'Communication from the Commission to the European Parliament, the Council and the European Central Bank. On steps towards completing economic and monetary union', Brussels, 21.10.2015, COM (2015) 600 final.
—(2015c), 'Report from the Commission to the European Parliament, the Council, the European Central Bank and the European Economic and Social Committee. Alert Mechanism Report 2016', Brussels, 26.11.2015, COM (2015) 691 final.
—(2016), 'The macroeconomic imbalance procedure. Rationale, process, application: a compendium', *European Economy Institutional Papers*, 039, November.
Frankel, J. and Saravelos, G. (2010), 'Can leading indicators assess country vulnerability? Evidence from the 2008–9 Global Financial Crisis', Harvard Kennedy School, mimeo.
Hallwirth, V. (2014), 'Überwachung der Wettbewerbsfähigkeit in der Europäischen Wirtschafts– und Währungsunion', *Wirtschaftsdienst*, 94, 11, pp. 798–805.
Hartmann, P. (2015), 'Real estate markets and macroprudential policy in Europe', ECB Working Paper Series, no. 1796.
Hošková, A. (2001), 'Impact of foreign direct investment on the economy of Slovakia', National Bank Slovakia.
Kaminsky, G., Lizando, S. and Reinhart, C. (1998), 'Leading indicators of currency crisis', *IMF Staff Papers*, 45, 1.
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, pp. 473–500.
Kamps, Ch., De Stefani, R., Leiner-Killinger, N., Rüffer, R. and Sondermann, D. (2013), 'Would the strengthened EU fiscal and economic governance framework have helped signalling sovereign debt crises?', Paper presented at Banca d'Italia 15th Public Finance Workshop.
—(2014), 'The identification of fiscal and macroeconomic imbalances. Unexploited synergies under the strengthened EU governance framework', ECB Occasional Paper Series, no 157.
Marchiori, L., Pierrard, O. and Sneessens, H. R. (2011), 'Demography, capital flows and unemployment', IZA DP, no. 6094.
Moschella, M. (2014), Monitoring macroeconomic imbalances: is EU surveillance more effective than IMF surveillance?', *Journal of Common Market Studies*, 52, 6, pp. 1273–89.
Official Journal of the European Union (2011a), ' Regulation (EU) No 1174/2011 of the European Parliament and of the Council of 16 November 2011 on enforcement measures to correct excessive macroeconomic imbalances in the euro area'.
—(2011b), 'Regulation (EU) No 1176/2011 of the European Parliament and of the Council of 16 November 2011 on the prevention and correction of macroeconomic imbalances'.
—(2012), 'Consolidated version of the Treaty on the Functioning of the European Union'.
Reinhart, C.M. and Rogoff, K.S. (2009), *This Time is Different: Eight Centuries of Financial Folly*, Princeton, N. J.: Princeton University Press.
Sarlin, P. (2013), 'On policymakers' loss functions and the evaluation of early warning systems', ECB Working Paper Series, no. 1509.
Schularick, M. and Taylor, A.M. (2012), 'Credit booms gone bust: monetary policy, leverage cycles, and financial crises, 1870–2008', *American Economic Review*, 102, 2, pp. 1029–61.
Stankovičová, I. and Vojtková, M. (2007), Viacrozmerné štatistické metódy s aplikáciami, Bratislava, Iura edition.
Taylor, A.M. (2012), 'The great leveraging', NBER Working Paper, no. 18290.
Verick, Sh. (2011), 'Giving up job search during a recession: the impact of the global financial crisis on the South African labour market', IZA DP, no. 6116.
Wooldridge, J.M. (2010), *Econometric Analysis of Cross Section and Panel Data, Second edition*, MIT Press.
—(2013), *Introductory Econometrics: A Modern Approach, Fifth international edition*, USA, Cengage Learning.

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

Appendix

Table A1. Estimations of the Linear Probability Model (LPM) and the logit models with one-year time lag

Indicator	Logit	Logit Adjusted	Logit FA	LPM	LPM Adjusted	LMP FA
Youth unemployment rate	–0.017*	0.011*		–0.011	0.012**	
	(0.089)	(0.051)		(0.323)	(0.041)	
Long-term unemployment rate	0.081**	0.022		0.068*	0.023	
	(0.037)	(0.126)		(0.057)	(0.118)	
Unemployment rate	–0.015	0.003		–0.014	0.007	
	(0.287)	(0.723)		(0.409)	(0.488)	
Net international investment position	–0.001	0.000		–0.001	–0.001	
	(0.508)	(0.661)		(0.414)	(0.252)	
Factor 1: Labor–Capital Nexus			*0.073***			*0.078**
			(0.019)			*(0.090)*
Export market shares	–0.002	–0.003		–0.001	–0.001	
	(0.272)	(0.202)		(0.718)	(0.731)	
Nominal unit labour cost	0.000	–0.005		0.005	0.000	
	(0.929)	(0.334)		(0.336)	(0.962)	
Factor 2: Catch-up Effect			*0.001*			*0.034*
			(0.979)			*(0.524)*
Total financial sector liabilities	0.001	–0.007*		0.002	–0.007**	
	(0.615)	(0.096)		(0.479)	(0.039)	
House price index	–0.035***	–0.033***		–0.021***	–0.019***	
	(0.000)	(0.000)		(0.000)	(0.000)	
Youth unemployment rate	–0.017*	0.011*		–0.011	0.012**	
	(0.089)	(0.051)		(0.323)	(0.041)	
Factor 3: Real estate bubble			*–0.222****			*–0.178****
			(0.000)			*(0.000)*
Current account balance	–0.001	–0.006	–0.005	–0.005	–0.013	–0.007
	(0.912)	(0.420)	(0.528)	(0.638)	(0.120)	(0.432)
Real effective exchange rate	0.005		–0.004	0.001		–0.005
	(0.413)		(0.446)	(0.937)		(0.390)
General government gross debt	0.000	0.001	0.000	0.000	0.001	0.001
	(0.880)	(0.529)	(0.782)	(0.781)	(0.683)	(0.766)
Private sector domestic credit	0.002***		0.002***	0.002***		0.002***
	(0.000)		(0.000)	(0.004)		(0.000)
Private sector credit flow	–0.004***		–0.004	–0.004***		–0.003***
	(0.094)		(0.211)	(0.000)		(0.000)
Activity rate	–0.044		–0.046*	–0.056*		–0.059**
	(0.107)		(0.098)	(0.062)		(0.028)
Number of observations	213		213	213		213
R-squared / Pseudo R-squared	0.418		0.334	0.389		0.355
Correct prediction rate	0.808		0.808	0.798		0.793

Source: Authors' calculations based on data from Eurostat and the Commission.

Note: Significance level at * 10 per cent, ** 5 per cent, *** 1 per cent. FA are estimates with variables reduced to factors by factor analysis. LPM denotes linear probability model.

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

Table A2. Estimations of the Linear Probability Model (LPM) and the logit models with two-year time lag

Indicator	Logit	Logit Adjusted	Logit FA	LPM	LPM Adjusted	LMP FA
Youth unemployment rate	0.000	0.011*		0.001	0.011*	
	(0.986)	(0.095)		(0.953)	(0.085)	
Long–term unemployment rate	0.035	0.004		0.034	0.005	
	(0.355)	(0.795)		(0.384)	(0.745)	
Unemployment rate	–0.024	–0.007		–0.024	–0.007	
	(0.226)	(0.588)		(0.267)	(0.658)	
Net international investment position	–0.001	–0.002*		–0.002	–0.002*	
	(0.248)	(0.077)		(0.350)	(0.095)	
Factor 1: Labor–Capital Nexus			*0.016*			*0.019*
			(0.711)			*(0.728)*
Export market shares	0.000	0.000		–0.001	0.000	
	(0.835)	(0.876)		(0.787)	(0.974)	
Nominal unit labour cost	0.007	0.002		0.008	0.003	
	(0.144)	(0.696)		(0.160)	(0.599)	
Factor 2: Catch-up Effect			*0.074*			*0.081*
			(0.246)			*(0.297)*
Total financial sector liabilities	0.003	–0.002		0.003	–0.002	
	(0.232)	(0.499)		(0.245)	(0.538)	
House price index	–0.011***	–0.011**		–0.011**	–0.011**	
	(0.007)	(0.025)		(0.011)	(0.018)	
Youth unemployment rate	0.000	0.011*		0.001	0.011*	
	(0.986)	(0.095)		(0.953)	(0.085)	
Factor 3: Real estate bubble			*–0.105***			*–0.107***
			(0.020)			*(0.023)*
Current account balance	–0.012	–0.022**	–0.019*	–0.013	–0.025***	–0.019
	(0.289)	(0.011)	(0.089)	(0.324)	(0.008)	(0.116)
Real effective exchange rate	–0.015**		–0.018**	–0.014**		–0.017***
	(0.037)		(0.010)	(0.036)		(0.008)
General government gross debt	0.000	0.000	0.000	0.000	0.000	0.000
	(0.944)	(0.965)	(0.921)	(0.929)	(0.999)	(1.000)
Private sector domestic credit	0.001***		0.002***	0.002**		0.002***
	(0.007)		(0.000)	(0.042)		(0.002)
Private sector credit flow	0.000		0.000	0.000		0.000
	(0.870)		(0.772)	(0.773)		(0.755)
Activity rate	–0.088***		–0.102***	–0.095**		–0.106***
	(0.003)		(0.000)	(0.010)		(0.003)
Number of observations	186		186	186		186
R-squared / Pseudo R-squared	0.284		0.261	0.329		0.304
Correct prediction rate	0.780		0.763	0.774		0.769

Source: Authors' calculations based on data from Eurostat and the Commission.
Note: Significance level at * 10 per cent, ** 5 per cent, *** 1 per cent. FA are estimates with variables reduced to factors by factor analysis. LPM denotes linear probability model.

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

Table A3. Estimations of the Linear Probability Model (LPM) and the logit models with three-year time lag

Indicator	Logit	Logit Adjusted	Logit FA	LPM	LPM Adjusted	LMP FA
Youth unemployment rate	0.026**	0.009		0.022**	0.009	
	(0.027)	(0.127)		(0.048)	(0.155)	
Long-term unemployment rate	–0.057	–0.033*		–0.042	–0.032	
	(0.173)	(0.091)		(0.328)	(0.130)	
Unemployment rate	–0.032	–0.022		–0.029	–0.020	
	(0.230)	(0.294)		(0.213)	(0.312)	
Net international investment position	–0.003*	–0.003***		–0.003	–0.003***	
	(0.066)	(0.002)		(0.151)	(0.013)	
Factor 1: Labor–Capital Nexus			*–0.042*			*–0.043*
			(0.531)			*(0.539)*
Export market shares	–0.001	0.000		–0.001	0.000	
	(0.705)	(0.908)		(0.706)	(0.961)	
Nominal unit labour cost	0.011*	0.009**		0.012*	0.010*	
	(0.056)	(0.043)		(0.085)	(0.052)	
Factor 2: Catch-up Effect			*0.080*			*0.084*
			(0.244)			*(0.271)*
Total financial sector liabilities	–0.002	0.001		–0.002	0.001	
	(0.522)	(0.602)		(0.523)	(0.649)	
House price index	0.011**	0.005		0.010**	0.005	
	(0.022)	(0.335)		(0.036)	(0.357)	
Youth unemployment rate	0.026**	0.009		0.022**	0.009	
	(0.027)	(0.127)		(0.048)	(0.155)	
Factor 3: Real estate bubble			*–0.001*			*0.000*
			(0.980)			*(0.995)*
Current account balance	–0.011	–0.027***	–0.023**	–0.011	–0.029***	–0.024**
	(0.326)	(0.000)	(0.040)	(0.442)	(0.001)	(0.049)
Real effective exchange rate	–0.018**		–0.013*	–0.018**		–0.013**
	(0.020)		(0.050)	(0.022)		(0.043)
General government gross debt	0.001	0.001	0.000	0.001	0.000	0.000
	(0.718)	(0.721)	(0.990)	(0.792)	(0.810)	(0.960)
Private sector domestic credit	0.002***		0.002***	0.002***		0.002***
	(0.008)		(0.000)	(0.022)		(0.000)
Private sector credit flow	0.000		–0.001	0.000		–0.001
	(0.886)		(0.371)	(0.886)		(0.426)
Activity rate	–0.073**		–0.096***	–0.075*		–0.094**
	(0.026)		(0.009)	(0.062)		(0.023)
Number of observations	159		159	159		159
R-squared / Pseudo R-squared	0.253		0.168	0.294		0.210
Correct prediction rate	0.755		0.723	0.736		0.711

Source: Authors' calculations based on data from Eurostat and the Commission.

Note: Significance level at * 10 per cent, ** 5 per cent, *** 1 per cent. FA are estimates with variables reduced to factors by factor analysis. LPM denotes linear probability model.

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

Table A4. Country-specific prediction rates for the members of PIIGS and Hungary

		Base model					Factor model				
		Obs.	N.Cri.	Pr.Cri.	aNtS	CPR	Obs.	N.Cri.	Pr.Cri.	aNtS	CPR
Lag 1											
EL	LPM	6	6	1.000	N/A	1.000	6	6	1.000	N/A	1.000
	Logit	6	6	1.000	N/A	1.000	6	6	0.833	N/A	0.833
ES	LPM	3	3	1.000	N/A	1.000	3	3	1.000	N/A	1.000
	Logit	3	3	1.000	N/A	1.000	3	3	1.000	N/A	1.000
HU	LPM	7	4	1.000	0.333	0.857	7	4	0.500	0.667	0.571
	Logit	7	4	1.000	0.333	0.857	7	4	0.750	0.444	0.714
IE	LPM	8	5	1.000	0.333	0.875	8	5	1.000	0.333	0.875
	Logit	8	5	1.000	0.333	0.875	8	5	1.000	0.333	0.875
IT	LPM	11	6	0.333	0.600	0.545	11	6	0.500	0.400	0.636
	Logit	11	6	0.500	0.400	0.636	11	6	0.667	0.300	0.727
PT	LPM	11	6	1.000	1.000	0.545	11	6	1.000	0.800	0.636
	Logit	11	6	0.667	1.500	0.364	11	6	1.000	0.600	0.727
Lag 2											
EL	LPM	5	5	1.000	N/A	1.000	5	5	1.000	N/A	1.000
	Logit	5	5	1.000	N/A	1.000	5	5	1.000	N/A	1.000
ES	LPM	2	2	0.500	N/A	0.500	2	2	1.000	N/A	1.000
	Logit	2	2	0.500	N/A	0.500	2	2	1.000	N/A	1.000
HU	LPM	6	3	0.667	0.500	0.667	6	3	0.667	0.500	0.667
	Logit	6	3	1.000	0.333	0.833	6	3	0.667	0.500	0.667
IE	LPM	7	5	0.800	1.250	0.571	7	5	0.800	1.250	0.571
	Logit	7	5	0.800	1.250	0.571	7	5	0.800	1.250	0.571
IT	LPM	10	6	0.333	0.000	0.600	10	6	0.333	0.000	0.600
	Logit	10	6	0.333	0.000	0.600	10	6	0.333	0.000	0.600
PT	LPM	10	6	1.000	1.000	0.600	10	6	1.000	1.000	0.600
	Logit	10	6	1.000	1.000	0.600	10	6	1.000	1.000	0.600
Lag 3											
EL	LPM	4	4	1.000	N/A	1.000	4	4	1.000	N/A	1.000
	Logit	4	4	1.000	N/A	1.000	4	4	1.000	N/A	1.000
ES	LPM	1	1	0.000	N/A	0.000	1	1	0.000	N/A	0.000
	Logit	1	1	0.000	N/A	0.000	1	1	0.000	N/A	1.000
HU	LPM	5	2	1.000	0.333	0.800	5	2	0.500	0.667	0.600
	Logit	5	2	1.000	0.333	0.800	5	2	0.500	0.667	0.600
IE	LPM	6	4	1.000	1.000	0.667	6	4	0.750	1.333	0.500
	Logit	6	4	1.000	1.000	0.667	6	4	0.750	1.333	0.500
IT	LPM	9	6	0.500	0.000	0.667	9	6	0.167	0.000	0.444
	Logit	9	6	0.667	0.000	0.778	9	6	0.333	0.000	0.556
PT	LPM	9	6	1.000	1.000	0.667	9	6	1.000	1.000	0.667
	Logit	9	6	1.000	1.000	0.667	9	6	1.000	1.000	0.667

Source: Authors' calculations based on data from Eurostat and the Commission.

Note: Abbreviations EL, ES, HU, IE, IT, PT stand for Spain, Greece, Hungary, Ireland, Italy, and Portugal, respectively and LPM stands for linear probability model. Furthermore, the available observations (Obs.), number of crisis events (N.Cri.), correct prediction rate of crisis event (Pr.Cri.), adjusted noise to signal ratio (aNtS), and correct prediction rate (CPR) are also presented in the table. The statistic for base model presents results for unadjusted Logit and LPM models while factor model results are obtained for the estimates with variables reduced to factors by factor analysis.

Commentary:

For evaluation of the estimated models at national level, additional summary statics were computed for members of PIIGS (Portugal, Italy, Ireland, Greece, and Spain) and Hungary, presented in table A4. These countries were chosen because of being hit substantially by recent financial and economic crises. Therefore they may provide a good real life example for studying the countries which were at the time of the crisis outbreak mostly susceptible to external shocks.

In general, the number of observations for individual countries is rather low, especially with increasing numbers of lags of used indicators, ranging from eleven observations to one observation. With a small number of observations, the variation in dependant variable is also low, leading to some rather extreme situations where all observed periods are either crisis periods or tranquil periods. Such situations arose for Greece and Spain, for which all of

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

the observed periods available are crisis periods. It is mostly for these reasons and for the setting of the positive outcome threshold at 0.5 that models do provide rather outstanding results of 100% correct prediction rates in the cases of the described countries. With increased lags the performance for Spain becomes less stellar, reaching 0% correct prediction rate in case of the three-year horizon, where only one observation is available. This crisis period is incorrectly predicted by all reported models except for the logit model, which included previously described factors. It is, however, short-sighted to generalise these results to some specific properties of the model or any particular features of the position of these two countries in regard to recent recession.

Regarding the rest of the members of PIIGS and Hungary, more variation in statistics is observed for these countries as more observations are available. It is therefore possible to compare the country specific correction rate with one which was obtained using all available observations for the EU28 (tables A1–A3). In case of a one-year lag, the country specific rate for Hungary and Ireland surpasses the average correct prediction rate for the EU28 for the logit model and the LPM without factors and in the case of Ireland also with the inclusion of the derived factors. For the rest of the observed cases the one-year horizon is however lower than the correct prediction rate for the entire EU. This possibly indicates that the members of PIIGS and Hungary are on average more reliant on other unobserved phenomena than other EU Member States. This seems to be even more evident with more lags, the only exception being Hungary which performs rather well even on three-year horizon for models without factors.

Table A5. Country-specific assessment of positive outcome threshold

	Logit						LPM					
	L1 BA	L1 FA	L2 BA	L2 FA	L3 BA	L3 FA	L1 BA	L1 FA	L2 BA	L2 FA	L3 BA	L3 FA
AT	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
BE	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
BG	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
CY	APOT	APOT	APOT	APOT	N/A	N/A	APOT	APOT	APOT	APOT	N/A	N/A
CZ	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT
DE	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT
DK	APOT	APOT	APOT	APOT	APOT	APOT	APOT	APOT	APOT	APOT	APOT	APOT
EE	APOT	LPOT	HPOT	HPOT	LPOT	HPOT	LPOT	LPOT	HPOT	HPOT	LPOT	HPOT
EL	APOT	APOT	APOT	APOT	APOT	APOT	APOT	APOT	APOT	APOT	APOT	APOT
ES	APOT	APOT	APOT	APOT	N/A	N/A	APOT	APOT	APOT	APOT	N/A	N/A
FI	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT
FR	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
HR	LPOT	LPOT	LPOT	LPOT	APOT	APOT	LPOT	LPOT	LPOT	LPOT	APOT	APOT
HU	LPOT	LPOT	LPOT	APOT	LPOT	HPOT	LPOT	HPOT	APOT	APOT	LPOT	HPOT
IE	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT
IT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT
LT	APOT	APOT	HPOT	APOT	N/A	N/A	HPOT	APOT	HPOT	APOT	N/A	N/A
LU	HPOT	LPOT	APOT	LPOT	LPOT	LPOT	LPOT	LPOT	APOT	LPOT	LPOT	LPOT
LV	HPOT	APOT	HPOT	HPOT	APOT	APOT	HPOT	APOT	HPOT	HPOT	APOT	APOT
MT	HPOT	HPOT	N/A	N/A	N/A	N/A	HPOT	HPOT	N/A	N/A	N/A	N/A
NL	LPOT	LPOT	APOT	HPOT	HPOT	HPOT	LPOT	LPOT	APOT	HPOT	HPOT	HPOT
PL	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
PT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT	LPOT
RO	HPOT	HPOT	APOT	APOT	HPOT	HPOT	HPOT	HPOT	APOT	APOT	HPOT	HPOT
SE	APOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT
SI	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT
SK	LPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT	HPOT
UK	HPOT	HPOT	LPOT	APOT	LPOT	LPOT	HPOT	HPOT	LPOT	HPOT	LPOT	LPOT

Source: Authors' calculations based on data from Eurostat and the Commission.

Note: Abbreviations AT, BE, BG, CY, CZ, DE, DK, EE, EL, ES, FI, FR, HR, HU, IE, IT, LT, LU, LV, MT, NL, PL, PT, RO, SE, SI, SK, UK stand for Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Germany, Estonia, Greece, Spain, Croatia, Hungary, Ireland, Italy, Lithuania, Luxemburg, Latvia, Malta, Nederland, Poland, Portugal, Romania, Sweden, Slovenia, Slovakia, and United Kingdom, respectively and LPM stands for linear probability model. The table provides evaluation of the positive outcome thresholds for each country and model used. Whether it was a high positive outcome threshold (HPOT), low positive outcome threshold (LPOT), appropriate positive outcome thresholds (APOT) for a particular country, or was not possible to compute (N/A). The values for base model (BA) present results for unadjusted Logit and LPM models while factor model (FA) results are obtained for the estimates with variables reduced to factors by factor analysis. The set positive outcome threshold was 0.5.

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

Table A6. Country-specific assessment of positive outcome threshold (values)

	Logit						LPM					
	L1 BA	L1 FA	L2 BA	L2 FA	L3 BA	L3 FA	L1 BA	L1 FA	L2 BA	L2 FA	L3 BA	L3 FA
AT	.	.	.	.	.	.	.	.	.	.	.	.
BE	.	.	.	.	.	.	.	.	.	.	.	.
BG	.	.	.	.	.	.	.	.	.	.	.	.
CY	1.000	1.000	1.000	1.000	.	.	1.000	1.000	1.000	1.000	.	.
CZ	0.000	0.000	0.000	0.000	0.429	0.000	0.000	0.000	0.000	0.000	0.000	0.000
DE	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
DK	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
EE	1.000	1.125	0.800	0.800	1.167	0.875	1.125	1.125	0.800	0.800	1.167	0.875
EL	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
ES	1.000	1.000	1.000	1.000	.	.	1.000	1.000	1.000	1.000	.	.
FI	0.550	0.314	0.400	0.000	0.360	0.360	0.000	0.000	0.400	0.000	0.360	0.360
FR	.	.	.	.	.	.	.	.	.	.	.	.
HR	1.071	1.071	1.125	1.125	1.000	1.000	1.071	1.071	1.125	1.125	1.000	1.000
HU	1.167	1.050	1.200	1.000	1.250	0.833	1.167	0.875	1.000	1.000	1.250	0.833
IE	1.143	1.143	1.400	1.400	1.500	1.500	1.143	1.143	1.400	1.400	1.500	1.500
IT	0.786	0.917	0.556	0.556	0.857	0.600	0.611	0.786	0.556	0.556	0.750	0.375
LT	1.000	1.000	0.000	1.000	.	.	0.600	1.000	0.000	1.000	.	.
LU	0.917	1.100	1.000	1.042	1.250	1.071	1.100	1.100	1.000	1.042	1.250	1.071
LV	0.625	1.000	0.964	0.964	1.000	1.000	0.833	1.000	0.964	0.964	1.000	1.000
MT	0.000	0.000	.	.	.	.	0.000	0.000	.	.	.	.
NL	1.500	1.500	1.000	0.833	0.889	0.889	1.500	1.125	1.000	0.833	0.889	0.889
PL	.	.	.	.	.	.	.	.	.	.	.	.
PT	1.833	1.375	1.667	1.667	1.500	1.500	1.833	1.571	1.667	1.667	1.500	1.500
RO	0.900	0.900	1.000	1.000	0.889	0.889	0.960	0.900	1.000	1.000	0.889	0.889
SE	1.000	0.000	0.000	0.000	0.643	0.000	0.000	0.000	0.000	0.000	0.643	0.000
SI	0.825	0.978	0.857	0.750	0.600	0.600	0.978	0.825	0.857	0.750	0.720	0.360
SK	1.200	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
UK	0.880	0.978	1.250	1.000	1.500	1.500	0.880	0.733	1.250	0.889	1.500	1.500

Source: Authors' calculations based on data from Eurostat and the Commission.

Note: Abbreviations AT, BE, BG, CY, CZ, DE, DK, EE, EL, ES, FI, FR, HR, HU, IE, IT, LT, LU, LV, MT, NL, PL, PT, RO, SE, SI, SK, UK stand for Austria, Belgium, Bulgaria, Cyprus, Czech Republic, Germany, Estonia, Greece, Spain, Croatia, Hungary, Ireland, Italy, Lithuania, Luxemburg, Latvia, Malta, Nederland, Poland, Portugal, Romania, Sweden, Slovenia, Slovakia, and United Kingdom, respectively and LPM stands for linear probability model. The table provides evaluation of the positive outcome threshold for each country and model used, by comparing the correct prediction rate of crisis event with overall correct prediction rate. A number higher than one indicates that the model is better in predicting a crisis event than tranquil periods, while a lower number indicates that the model predicts tranquil periods rather than crisis events. The values for base model (BA) present results for unadjusted Logit and LPM models while factor model (FA) results are obtained for the estimates with variables reduced to factors by factor analysis. The set positive outcome threshold was 0.5.

Commentary:

Tables A5 and A6 provide an overview of the country-specific heterogeneity in the performance of the applied models in regard to their ability to predict crisis occurrence. Based on the available observations for each country, the correct prediction rate of crisis event was compared to the overall correct prediction rate. The computed ratio is higher than one in the examined country if the model performs better in predicting crisis periods than in predicting tranquil periods. Analogously, the ratio is lower than one when the model predicts tranquil periods more reliably than crisis periods in the examined countries. Given the small number of observations for each country, they should be regarded only as hints of potential signs of country-specific heterogeneity and only in regard to the models and positive outcome threshold.

Based on the results obtained, it is possible to see that the performance of models in particular countries changes with the time horizon considered for the prediction of crisis. There are, however, some countries examined for which the performance is consistently better for predicting tranquil periods than crisis periods, such as the Czech Republic, Germany, Finland, Italy, and Slovenia. For each of these countries, the ratio of performance is in favour of the prediction of tranquil periods regardless of the number of lags or model used. One can recommend decreasing the set positive outcome threshold value 0.5 of the estimated models for the listed countries in an effort to improve the overall performance of the models.

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press

On the other hand, all models perform more robustly in predicting the crisis periods than predicting the tranquil periods in Ireland, Portugal, and to a greater extent also in Croatia. Results for each of these countries consistently show better performance for predicting tranquil periods regardless of model or number of lags used, with the exception of Croatia, which shows balanced performance in case of the three-lag horizon. The models can be optimised by increasing the positive outcome threshold from the default 0.5 in the case of these countries to improve overall performance.

However, it was not possible to evaluate the performance of any of the models in the cases of Austria, Belgium, Bulgaria, France, Poland, and most of Malta specifications, with the exception of France, for which there was no intersection of observation for all indicators. This phenomenon was caused by low variation in the dependent variable, which obtained only value 0 in cases of the aforementioned countries, which meant that the ratio of correct prediction rate and correct prediction rate of crisis was not possible to compute. Nevertheless, the rest of the countries examined, which leaves almost half of the sample, cannot be unambiguously classified in any of the previous categories. Therefore, in most of the cases models were appropriately set or their performance in predicting crisis and tranquil periods was shifting with the number of lags. It is, however, important to bear in mind that the observed periods for computing these country-specific indicators are fairly short, and should not be used for policy recommendation alone.

https://doi.org/10.1177/002795011723900112 Published online by Cambridge University Press