Routledge
Taylor & Francis Group

Journal of Economic Policy Reform

ISSN: 1748-7870 (Print) 1748-7889 (Online) Journal homepage: http://www.tandfonline.com/loi/gpre20

Performance of the Macroeconomic Imbalance Procedure in light of historical experience in the CEE region

Maria Siranova & Marek Radvanský

To cite this article: Maria Siranova & Marek Radvanský (2018) Performance of the Macroeconomic Imbalance Procedure in light of historical experience in the CEE region, Journal of Economic Policy Reform, 21:4, 335-352, DOI: 10.1080/17487870.2017.1364642

To link to this article: https://doi.org/10.1080/17487870.2017.1364642

Journal of Economic Policy Reform, 2018
Vol. 21, No. 4, 335–352, https://doi.org/10.1080/17487870.2017.1364642

Performance of the Macroeconomic Imbalance Procedure in light of historical experience in the CEE region

Maria Siranova[a,b]* and Marek Radvanský[a]

[a]*Institute of Economic Research, Slovak Academy of Sciences, Bratislava, Slovak Republic;*
[b]*Department of Banking and International Finance, Faculty of National Economy, University of Economics, Bratislava, Slovak Republic*

This article applies set of 24 baseline and auxiliary indicators included into the Macroeconomic Imbalance Procedure framework on the conditions of 17 CEE countries to assess their predictive power given the policy pre-determined and optimal thresholds for the period 1991–2014. Our results suggest that the optimal official thresholds might either be excessively too accommodative (public and private, total or external debt levels), overly conservative (current account balance, export market share and nominal unit labour costs), or with less informative value (labour market characteristics) for set of transition economies. Indicators with higher predictive power belong predominantly to the group of external imbalances indicators.

Keywords: early warning indicators; macroeconomic imbalance procedure; CEE region

JEL classifications: G01, F47, F53

1. Introduction

Early warning indicators (EWIs) represent an essential component of macro-prudential policies by being part of the EWIs system used to warn of potential or impending problems. The recent financial crisis highlighted the issue of monetary union heterogeneity in terms of different stages of economic development accompanied by lack of convergence. Adoption of necessary measures to reduce or eliminate existing imbalances might help to prevent the occurrence of asymmetric shocks in case of heterogeneous economic conditions within the EMU.

Only handful of studies (e.g. Csortos and Szalai 2013; Knedlik 2014, 2015; Domonkos et al. 2016, 2017) have so far empirically assessed the efficiency of the new Macroeconomic Imbalance Procedure (MIP) which was recently introduced at the EU28 level (European Commission 2012a, 2012b). Yet, the need for efficient and reliable EWIs system has been widely recognised. This is true particularly for the CEE[1] economies that are highly sensitive towards the adverse evolution in the core EU28 countries via their external trade and financial linkages. While different forms of EWI systems have been tested in the conditions of many developed and developing economies (started by seminar paper by Kaminsky, Lizondo, and Reinhart 1998), the CEE region, as a separate country group, has been widely neglected. At the same time, recent studies by Knedlik (2014, 2015) argue that grouping of countries based on their level of economic development and internal homogeneity might lead to a significant

*Corresponding author. Email: maria.siranova@savba.sk

increase in the performance of EWI systems. Additionally, as the key objective of the MIP procedure is to provide sound input for policy-makers, their preferences need to be accounted for. Hence, the most reliable and efficient EWIs integrate both, a high predictive power and minimum costs related to their potential failure.

This article supplements broad literature on EWI systems by the following means. First, we evaluate the overall performance of available individual EWIs included in the MIP framework, the unique EWI system officially adopted for a set of economically divergent countries, under the conditions of a specific group – the CEE transition economies. This aims is to address the recent acknowledgement by the Economic Commission that some extent of differentiation might be justified if heterogenous economic circumstances are presents (European Commission 2016). Additionally, we calculate optimum thresholds for all available indicators, both baseline and auxiliary, given the maximisation of policy-maker's utility function in order to address the common critique of the MIP procedure that questions the specification of the trigger mechanism by "rule of thumbs" rules (Kamps et al. 2013). Therefore, our results might represent an important input for policy-makers when evaluating potential threat of burbling macroeconomic imbalances either on an individual country level or as part of the MIP in-depth review assessment.

Methodologically, we depart from the paper by Csortos and Szalai (2013) and Domonkos et al. (2017) by extending the data-set used for analysis and by implementing a policy-maker loss function as specified in Sarlin (2013). Our sample consists of 17 CEE economies covering the 1991–2014 periods on a yearly basis. Based on the previous, our analysis consists of the following steps reflecting Drehmann and Juselius (2013) recommendation on efficient EWIs. First, the efficiency of MIP benchmark EWIs along with auxiliary indicators is tested individually using the signalling approach. Second, the performance of different EWIs is evaluated by three different concepts (absolute utility, AUROC score and inverse aNtS ratio). Third, optimal thresholds are compared with the official pre-determined thresholds to specify the extent of overly accommodating or too restrictive policy-maker behaviour represented by the European Commission. Lastly, the performance of all EWIs is commented on by examining the AUROC-based signalling properties over the course of the last two decades on a year-by-year basis. The best performers among the benchmark and potential EWIs are selected given the outcomes from this four-step analytical approach.

Our outcomes support a call for a more country group-specific (or country individual-specific) indicative thresholds differentiating by the level of socio-economic development and regional specifics. The optimal thresholds proposed within the common MIP framework for the EU28 group of countries might either be excessively accommodative (public and private, total or external debt levels), overly conservative (current account balance, export market share and nominal unit labour costs), or have less information value (most of labour market characteristics). Indicators that show higher predictive power belong predominantly to the group of external imbalances indicators (current and capital account balance, export market share and REER) and are accompanied by a change in the house price index and young people outside education and training indicators.

The paper is structured as follows. First, we briefly describe basic features of the newly adopted MIP procedure along with its critique. In the second chapter, the signalling approach with its variations is presented. The third chapter discusses outcomes and further possible advancements. The last chapter summarizes findings and reveals recommendations for further development of MIP related policies.

2. Macroeconomic imbalance procedure as the EWI system

The EWI procedure might serve as a powerful tool in order to assist policy-makers in forming such pre-emptive steps that would prohibit macroeconomic imbalances to deepen already existing heterogeneity potentially leading to the severe economic crisis. EWIs in this context must not only have sound statistical forecasting power, but also need to satisfy several additional requirements. According to Drehmann and Juselius (2013) there are three crucial criteria that should identify the most powerful EWIs: (i) the appropriate timing requires a signal that does not come at a very early stage due to the potentially high costs of its mitigation but that does not arrive too late prohibiting any action at all; (ii) the stability of the signal allows policy-makers to affect expectations more efficiently and reduce uncertainty regarding trends, thus allowing for a more decisive policy actions; (iii) the EWI signals should be easy to interpret, as any more complicated forecast might be ignored by policy-makers.

On the European Union (EU) level the need for a conceptually robust EWIs system has been recently recognised. In 2010, the European Commission introduced a legislative package consisting of six proposals (the *six-pack* legislation), with the main objective to reinforce the monitoring and the surveillance of fiscal, macroeconomic and structural reform policies in the EU. In November 2011, the set of economic indicators that are part of the MIP scoreboard used to detect macroeconomic imbalances was presented by the Ecofin/Eurogroup. The structure of the official MIP scoreboard has undergone few changes during the five years of its existence. While the original scoreboard introduced in 2011 consisted of 10 main indicators, in 2012 the measure of change in total financial sector liabilities was introduced, and additional three labour market indicators were added in 2015. The current structure of the MIP scoreboard embodies outcomes from a continuous process of learning and interaction between European Commission, European Council and individual member states (European Commission 2016), hence reflecting the learning curve of stakeholders involved.

The introduction of a new procedure, MIP, within the EU's annual cycle of economic policy guidance and surveillance (the European Semester) serves as a tool for preventing and correcting macroeconomic imbalances in the euro area. As a robust monitoring mechanism, the MIP is used to identify the build-up of potential risks, correct existing imbalances and prevent them from re-emerging. Although that MIP scoreboard is not considered a pure EWS system, thus is not intended to predict the probability of the occurrence of crises (European Commission 2016), its scope and character are of having a similar nature and should help in the early detection and monitoring of emerging macroeconomic imbalances (EC Regulation 1176/2011).

The first stage of the MIP procedure (Alert Mechanism Report, AMR) involves screening member states by performance of MIP scoreboard indicators in order to select countries that might potentially face the risk of macroeconomic imbalances. The indicators contained in the MIP scoreboard were originally are intended to play the role of an EWI and are often directly referred to (European Commission 2011) and treated thus way (Csortos and Szalai 2013; Knedlik 2014, 2015, European Commission 2016; Domonkos et al. 2017).

The AMR report identifies those Member States for which more in-depth review (IDR) is consequently required to assess their inner vulnerabilities and, potentially, formulate policy recommendations if needed (European Commission 2012b). The preventive and corrective arms of the MIP based on IDRs thus provide the basis for any

recommendations to be addressed to the Member States. Thorough discussion on MIP procedure technicalities is further provided by Sikulova (2015).

The more detailed analysis of the efficiency of the MIP procedure is of high importance for both researchers and policy-makers considering the existing critique and the possible shortcomings listed in the literature. Moschella (2014) points out that the MIP does not deliver a mechanisms to prevent political considerations from interfering with the decision to activate sanctions as well as how to share the burden of adjustments. According to Kamps et al. (2013), several issues regarding the scoreboard indicators and thresholds established need to be targeted, either by amendments of the scoreboard itself or through a broader "economic reading" of the variables. Alcidi and Gros (2014) argue that the MIP should be based more on forward- (not backward-) looking variables, since the MIP system and its scoreboard are constructed as a preventive tool. Hallwirth (2014) points out several shortcomings of the procedure with respect to the surveillance of competitiveness divergences and current account imbalances.

Results from the recent studies have not yet provided a clear picture of the most efficient set of EWIs based on their individual or grouped performance reflecting the policy-makers loss function. The Domonkos et al. (2016) study concludes that the highest predictive ability may be attributed to the representatives of the external imbalances (current account balance and net international investment position) acting as very EWIs, and private sector debt and labour market inefficiencies (long-term unemployment rate, youth unemployment rate) serving as very short-term predictors. The weak performance of individual EWIs is highlighted in Csortos and Szalai (2013) with the exception of current account and unemployment rate. Financial system-based EWIs might represent an additional extension of the current MIP scoreboard as the credit-to-GDP ratio and capital flow indicators shows a promising performance in Csortos and Szalai (2014). Knedlik (2014) argues that policy-makers are less worried about the false crisis predictions (Type II error) than the missing crisis prediction (Type I error). In addition, the poor performance of individual indicators might be improved by taking into account estimation of optimal thresholds and by distinguishing the different level of economic development between most developed countries and countries from the CEE region, or by creating country-specific thresholds due to the high internal heterogeneity (Knedlik 2015).

A rather comprehensive coverage in terms of wide variety of EWIs and policy determined optimal thresholds allows us to investigate the performance of EWIs in the light of historical development in the CEE region since the transition period while acknowledging the critique by Knedlik (2015). This in turn results in the investigation of how the MIP could have fared over the past two decades if it would have been introduced and applied in 1993 on set of this highly specific group of countries. The calculation of group-specific optimal thresholds might provide an additional tool for policy-makers when assessing the extent of the threat presented by broiling macroeconomic imbalances in the second stage of the MIP procedure, the in-depth review evaluation.

3. Methodology

Our methodology assesses the predicting the power of the different EWIs introduced as part of the MIP and specifies the optimal EWIs thresholds given the preferences of the decision-maker in terms of committing various types of errors.

3.1. Signalling and the policy-maker utility function

The signalling approach that extends the early warning signal methodology is common in this type of the literature (Csortos and Szalai 2013, 2014; Sarlin 2013; Domonkos et al. 2016). Implementing an EWIs system can be an useful tool for policy-makers to address a forthcoming event. On the other hand, implementing an EWIs system leads to the possibility of event prediction error. Prediction errors rely on a level of (arbitrary) individual preferences of policy-makers, which directly affects the prediction threshold levels indicating an upcoming event. There are two types of prediction errors. If the threshold is too high, the event is not indicated (no warning). This is typically described as type I error – false negative. Type II error is related to low threshold level, when obtained signal is false positive (false warning). The confusion matrix is usually specified in the following manner:

	Crisis event	No crisis event
EWI prediction	A	B
No EWI prediction	C	D

For the evaluation of the predictive properties of the MIP indicators, three basic measures are normally employed (Csortos and Szalai 2014). The false negative rate (FNR) computes the ratio of missed events to all events, false positive rate (FPR) compares a number of false signals to all tranquil periods and measure of correctly predicted events as a percentage of all correctly predicted events to all events.

The adjusted noise-to-signal ratio (aNtS) serves as a tool for evaluating the performance of the individual EWI or the entire system (Kaminsky, Lizondo, and Reinhart 1998; Kaminsky and Reinhart 1999; Alessi et al. 2015) by comparing the false and true warning rate in the following manner:

$$\text{aNtS} = \left[\frac{\frac{B}{B+D}}{\frac{A}{A+C}}\right] \tag{1}$$

In general, the desirable outcome of [1] is below unity with a strategy to minimise the ratio given the set of plausible thresholds (Edison 2003). This approach also leads to the AUROC-based methodology using the inverse of the [1] in order to assess the reliability of the indicator benchmarking it with a random model.

The ROC (receiver operating characteristic) curve plots combinations of the true positive rate $[A/(A+C)]$ and the false positive rate $[B/(B+D)]$ for every possible threshold value. High thresholds generating a small number of signals will be located close to the zero origin while strict thresholds indicating higher amount of signals will be plotted close to the $[1;1]$ origin. The integral of the area under the ROC curve (AUROC) larger than 0.5 value signifies the indicator has an informative value better than the random model.

Losses related to forecasting errors are generally defined by the loss function (Alessi and Detken 2009; Csortos and Szalai 2014). The loss function defines the cost of non-reaction to the crisis that does occur (cost of adaptation) and the cost of adaptation to a false warning.

The standard Alessi–Detken type of loss function is defined in the following manner:

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

where μ stands for the factor revealing policy-maker risk-aversion profile towards crisis avoidance, T_1 the type 1 error associated with share of missed crisis, and T_2 the type 2 error associated with share of incorrectly signalled crisis.

Usually, the decision range for policy-makers is to set preferences (threshold) from interval 0.25–0.75. Maximum costs related to an EWIs system are based on the probability of a prediction error and the cost of adoption by type of error.

The standard Alessi–Detken type of loss function as in [2] might be adjusted to account for the probability of crisis in the sample, as proposed in Sarlin (2013):

$$L(\mu) = \mu T_1 P + (1 - \mu) T_2 (1 - P) \tag{3}$$

where P represents the probability of crisis occurrence in the sample.

The absolute utility function calculates the difference of potential loss generated by the model and the cost of ignoring EWIs system (indicators) completely. Positive values of absolute utility functions are treated as desirable outcomes signifying the added value of following the EWI system recommendations over the random model. Hence, the Sarlin-type absolute utility function is of the following form:

$$U(\mu) = \min[\mu P; (1 - \mu)(1 - P)] - L(\mu) \tag{4}$$

As apparent, the specification of the policy-maker's risk aversion coefficient μ represents a crucial issue in the process of EWI assessment. In the light of the recent crisis it has become more evident that the policy-makers' preferences have been shifted to towards more prudent behaviour. In the European Systemic Risk Board recent publication (Detken et al. 2014) on capital buffer requirements the advisable sort of action places the μ coefficient within the $[0.5; 1]$ interval justifying the choice by considering the policy-makers at least as concerned about the missing the crisis as with pre-emptive actions based on false predictions. The Sarlin-type of loss and utility function compensates for the relatively low crisis occurrences by preferences being strongly biased towards the EWI efficiency in predicting the crisis, as pointed out by the Alessi and Detken (2014) response to Sarlin (2013) paper.

In the Sarlin-type utility function framework, the policy-maker might be required to state his preferences in a very precise manner, as even the relatively small change might render EWI useless. The sensitivity analysis of robustness in the estimated outcomes targets this issue by providing upper and lower bound of confidence intervals based on the up to 10% positive and negative variations in the values of underlying EWIs.

While various studies us different time windows for signalling the crisis, the expert-based choice is usually derived from the real-time behaviour of the underlying leading EWIS in the pre-crisis period (Kaminsky, Lizondo, and Reinhart 1998). To name few of the studies, Drehmann and Juselius (2013) opt for one and a half years to five years interval, Alessi and Detken (2009) for a maximum one and a half years window, Knedlik (2014) for one-to-two years lag. Paper by Babecký et al. (2013) explicitly tests for optimal time lag by distinguishing between three categories of indicators: the late EWI (1–3 quarters beforehand), early EWI (4–8 quarters beforehand), and ultra-early EWI (9 + quarters beforehand).

In our approach, we choose to make a compromise among all the aforementioned studies by specifying the time lag in the following manner. Acknowledging that the policy-maker requires a sensible time to react once a signal has been issued the lag between signal and crisis event is thus set for a one year. Additionally, the EWI issues a positive signal if there is at least one crisis occurrence within the following two-year period. Hence, we test the predictive power of selected EWIs treating them as early

and ultra-early indicators making it comparable with findings in Alessi and Detken (2009) and Knedlik (2014) studies. As part of the robustness check we also comment on the likely optimal time horizon of individual EWIs given the AUROC maximising criterion.[2] To look at the stability of EWIs performance across time we plot AUROC scores for individual countries on yearly basis.[3] As advocated by Candelon, Dumitrescu, and Hurlin (2012) the AUROC-based assessment of early warning systems should have a preference over other methods.

3.2. *Crisis specification in the context of CEE countries*

Different time horizons are tested to assess the predictive power of the various EWIs acknowledging the first requirement of an efficient and powerful EWI as specified in the Drehmann and Juselius (2013). In the literature there is a relatively broad variety of different indicators serving as the crisis signal ranging from years in which a new IMF lending arrangement is instituted (Knedlik 2012), spread on yields for government bonds over AAA-rated bonds exceeded the mean by one standard deviation (Knedlik 2014), negative deviation of real GDP from the potential GDP measured by an output gap (Csortos and Szalai 2014; Domonkos et al. 2016), significant negative deviation of real GDP from trend (Csortos and Szalai 2013, 2014), and list of cross-country crisis complied by other authors (Drehmann and Juselius 2013). The study by Babecký et al. (2013) advocates a continuous indicator of real crisis costs integrating output and employment loss with fiscal deficits.

As the MIP should be a general procedure that warns before the overall economic crisis caused by deepening of internal of external imbalances within the individual countries and EMU as a whole we opt for a more comprehensive definition of an economic crisis.[4] The study by Domonkos et al. (2016) argues that in the MIP context the deviations from the potential GDP should be used to capture the adverse impact of widening imbalances on the economic performance. Since the standardised estimates of the potential GDP for the CEE region until 1991 are not available we instead focus on crisis periods that are captured by deviations of the real GDP growth from its five-year average by more than one standard deviation, similar to Csortos and Szalai (2013, 2014) in the use of real GDP and to Knedlik (2014) in the use of standard deviations from a long-term trend. In general, a five-year window corresponds to the average length of the business cycle.

In the context of the CEE region (Appendix 3 in Supplemental data) drops in real GDP growth larger than one standard deviation split the sample into two main crisis periods clustering around the 1997–1998 and 2007–2008 crisis years, along with a handful of country-specific slumps in economic activity. This set-up allows for investigation of the MIP performance under the wider set of circumstances (similar to Csortos and Szalai 2013, 2014) rather than focusing on the post-2000 period (Knedlik 2012, 2014; Domonkos et al. 2016). On average, the frequency of crisis occurrence is approximately 18%, however, depending on the number of observations available per each indicator (Table 1).

3.3. *Data-set and EWIs specification*

Nowadays the scoreboard used for AMRs includes a list of 14 indicators with their indicative thresholds (including an indicator of the financial sector and new employment indicators) as well as a supplementary list of other 28 auxiliary indicators without

Table 1. (a) MIP labour market indicators – official and optimal threshold characteristics, (b) MIP external imbalances indicators – official and optimal threshold characteristics, (c) MIP indebtedness-related indicators – official and optimal threshold characteristics, and (d) MIP other indicators – official and optimal threshold characteristics.

					Official threshold				Optimal threshold				Time horizon	
		Auxiliary Indicator	#	Probability of event (%)	Value	Implied Preferences	Max Absolute Utility	AUROC	Value	Implied Preferences	Max Absolute Utility	Inverse aNtS	Lag/ Interval	AUROC
(a)														
Youth UR		NO	287	16.72	2.00	0.777	−0.052	0.377	0.62	0.833	−0.021	0.723	3/3	0.526
Sensitivity	*−10%*					0.816	−0.034	0.377	0.56	0.833	−0.021	0.723		
interval	*+10%*					0.777	−0.051	0.380	0.68	0.833	−0.021	0.723		
Long-term UR		NO	155	20.65	0.50	0.729	−0.046	0.442	0.41	0.794	−0.014	0.835	2/3	0.518
Sensitivity	*−10%*					0.642	−0.090	0.442	0.37	0.794	−0.014	0.835		
interval	*+10%*					0.729	−0.046	0.449	0.45	0.794	−0.014	0.835		
UR		NO	229	18.78	10.00	0.803	−0.048	0.365	6.40	0.813	−0.001	0.992	3/3	0.503
Sensitivity	*−10%*					0.809	−0.039	0.363	5.80	0.813	−0.001	0.992		
interval	*+10%*					0.803	−0.046	0.364	7.10	0.813	−0.001	0.992		
Activity rate		NO	295	16.61	−0.20	0.810	−0.009	0.503	−0.20	0.834	0.003	1.053	2/3	0.553
Sensitivity	*−10%*					0.810	−0.008	0.502	−0.18	0.834	0.003	1.053		
interval	*+10%*					0.810	−0.009	0.507	−0.22	0.834	0.003	1.053		
Employment		YES	302	16.56	NA	NA	NA	0.338	−5.34	0.834	−0.004	0.447	3/3	0.477
Sensitivity	*−10%*					NA	NA	0.339	−5.94	0.834	0.000	1.117		
interval	*+10%*					NA	NA	0.338	−5.88	0.834	−0.004	0.447		
Participation Rate		YES	302	16.56	NA	NA	NA	0.439	58.05	0.835	0.001	1.017	3/1	0.464
Sensitivity	*−10%*					NA	NA	0.441	52.20	0.835	0.001	1.017		
interval	*+10%*					NA	NA	0.440	71.10	0.835	0.002	1.013		
YNEET		YES	204	19.12	NA	NA	NA	0.392	25.65	0.788	0.012	3.362	3/3	0.588
Sensitivity	*−10%*					NA	NA	0.393	23.10	0.788	0.012	3.362		
interval	*+10%*					NA	NA	0.391	28.20	0.788	0.012	3.362		
(b)														
NULC		NO	196	18.88	12.00	0.864	−0.065	0.363	26.37	0.811	−0.017	0.337	3/3	0.533
Sensitivity	*−10%*					0.837	−0.045	0.360	24.39	0.811	−0.015	0.355		
interval	*+10%*					0.864	−0.067	0.360	20.79	0.811	−0.017	0.499		
Terms of trade		YES	112	23.21	NA	NA	NA	0.520	−3.90	0.768	0.017	1.470	1/1	0.548

Sensitivity	*−10%*						NA	NA	0.520	−3.50	0.750	0.013	1.470		
interval	*+10%*						NA	NA	0.519	−4.30	0.750	0.013	1.470		
REER		NO	+	161	18.01	11.00	0.821	0.022	0.600	8.85	0.820	0.027	1.454	1/1	0.638
Sensitivity	*−10%*						0.821	0.023	0.601	7.95	0.820	0.027	1.454		
interval	*+10%*						0.821	0.024	0.598	9.70	0.820	0.027	1.454		
			−	161	18.01	−11.00	0.846	−0.021	0.472	−10.45	0.819	0.006	2.985	3/3	0.556
Sensitivity	*−10%*						0.850	−0.025	0.472	−9.40	0.800	0.006	2.985		
interval	*+10%*						0.850	−0.026	0.472	−11.50	0.800	0.006	2.985		
EMS		NO		251	16.33	−6.00	0.929	0.002	0.697	−17.94	0.837	0.052	1.999	1/3	0.707
Sensitivity	*−10%*						0.950	−0.005	0.709	−17.34	0.850	0.048	2.163		
interval	*+10%*						0.950	−0.005	0.691	−17.64	0.850	0.044	1.916		
CA balance		NO	+	263	17.11	6.00	NA	NA	0.489	0.96	0.828	0.001	1.481	2/3	0.500
Sensitivity	*−10%*						NA	NA	0.489	0.84	0.828	0.001	1.481		
interval	*+10%*						NA	NA	0.489	1.08	0.828	0.001	1.481		
			−	263	17.11	−4.00	0.855	−0.005	0.562	−7.44	0.829	0.021	1.646	1/1	0.593
Sensitivity	*−10%*						0.850	−0.005	0.561	−8.68	0.800	0.015	2.105		
interval	*+10%*						0.900	−0.016	0.563	−10.64	0.800	0.015	2.105		
CA and CAP balance		YES		253	17.79	NA	NA	NA	0.643	−6.84	0.822	0.038	2.267	1/1	0.678
Sensitivity	*−10%*						NA	NA	0.642	−6.16	0.800	0.032	2.267		
interval	*+10%*						NA	NA	0.642	−7.52	0.800	0.032	2.267		
Productivity		YES		302	16.56	NA	NA	NA	0.433	−5.76	0.834	−0.004	0.701	1/1	0.492
Sensitivity	*−10%*						NA	NA	0.433	−5.46	0.834	−0.003	0.744		
interval	*+10%*						NA	NA	0.433	−6.00	0.834	−0.005	0.627		
(c)															
FDI flows		YES		301	16.28	NA	NA	NA	0.584	5.60	0.837	0.019	1.516	1/1	0.611
Sensitivity	*−10%*						NA	NA	0.584	5.05	0.837	0.019	1.516		
interval	*+10%*						NA	NA	0.583	6.15	0.837	0.019	1.487		
FDI stocks		YES		204	19.12	NA	NA	NA	0.444	29.20	0.809	0.002	1.023	3/3	0.506
Sensitivity	*−10%*						NA	NA	0.444	26.40	0.809	0.002	1.023		
interval	*+10%*						NA	NA	0.443	32.00	0.809	0.002	1.023		
Net IIR		NO		204	19.12	−35.00	0.824	−0.020	0.459	−70.70	0.809	0.007	1.260	1/1	0.589
Sensitivity	*−10%*						0.824	−0.020	0.460	−63.70	0.809	0.007	1.260		
interval	*+10%*						0.850	−0.022	0.460	−77.70	0.800	0.005	1.260		
Gross external debt		YES		98	11.22	NA	NA	NA	0.504	26.80	0.888	0.022	1.323	2/3	0.535

(*Continued*)

Table 1. (*Continued*).

					Official threshold				Optimal threshold				Time horizon	
		Auxiliary Indicator	#	Probability of event (%)	Value	Implied Preferences	Max Absolute Utility	AUROC	Value	Implied Preferences	Max Absolute Utility	Inverse aNtS	Lag/ Interval	AUROC
Sensitivity interval	*−10%*					NA	NA	0.504	24.00	0.888	0.022	1.323		
	+10%					NA	NA	0.505	29.60	0.888	0.022	1.323		
Public sector debt		NO	118	20.34	60.00	0.000	−0.153	0.466	19.80	0.797	0.018	1.183	3/3	0.650
Sensitivity interval	*−10%*					0.000	−0.109	0.468	17.40	0.797	0.018	1.183		
	+10%					0.769	−0.018	0.467	21.60	0.797	0.018	1.183		
Private sector debt		NO	297	15.82	133.00	NA	NA	0.536	33.25	0.842	0.014	1.198	1/1	0.549
Sensitivity interval	*−10%*					NA	NA	0.536	29.26	0.842	0.013	1.177		
	+10%					NA	NA	0.537	30.59	0.842	0.014	1.174		
HPI		NO	70	24.29	6.00	0.589	0.052	0.752	5.82	0.757	0.085	6.545	2/3	0.864
Sensitivity interval	*−10%*					0.589	0.052	0.753	5.28	0.757	0.085	6.545		
	+10%					0.589	0.046	0.752	6.42	0.757	0.085	6.545		
(d)														
Poverty		YES	204	19.12	NA	NA	NA	0.441	9.80	0.799	0.002	1.017	3/2	0.470
Sensitivity interval	*−10%*					NA	NA	0.439	6.20	0.799	0.002	1.013		
	+10%					NA	NA	0.441	10.80	0.799	0.002	1.017		
GFCF		YES	308	15.91	NA	NA	NA	0.342	11.60	0.840	0.002	2.341	3/3	0.545
Sensitivity interval	*−10%*					NA	NA	0.342	10.40	0.800	0.001	2.341		
	+10%					NA	NA	0.342	12.80	0.800	0.001	2.341		
RaD		YES	230	18.26	NA	NA	NA	0.481	1.43	0.818	0.009	1.065	2/3	0.504
Sensitivity interval	*−10%*					NA	NA	0.481	1.29	0.850	0.007	1.065		
	+10%					NA	NA	0.482	1.57	0.850	0.007	1.065		

indicative thresholds. The indicators include both stock and flow indicators which can capture shorter-term deteriorations as well as the longer-term accumulation of imbalances. The economic rationale behind the inclusion of individual indicators in the scoreboard, the transformations used and the determination of threshold values are provided in the European Commission (2012b) or European Commission (2015).

Due to the both data availability issues that are common among the CEE countries and non-existent reporting practices for selected indicators in general, the final list of indicators used in this study includes seven indicators for internal imbalances group mainly describing labour market conditions, additional seven indicators for the external imbalances and competitiveness group, next seven domestic and foreign indebtedness-related indicators and is complemented by last three auxiliary EWIs (Table 1). The construction of EWIs is in accordance with definitions used in the most updated version of the MIP scoreboard indicators as of today. As argued, the current version of the MIP scoreboard should encompass the experience of policy-makers since the MIP introduction in 2011.

In total, the data-set consists of 17 countries from the CEE region covering the period from 1991 to 2014 (Appendix 3 in Supplemental data). For certain countries and variables, the data cover a shorter time span given their availability. The list of countries and indicators along with a short description and data sources is given in the Appendix 2 in Supplemental data.

4. Results and discussion

In the Table 1, we report for each indicator its AUROC score, implied policy-maker preferences as specified in the [3], absolute utility from the [4], and inverse value of the aNtS ratio as defined in [1] along with a specification of optimal threshold values given the aforementioned characteristics. Sensitivity intervals for each indicator serve as a robustness check of estimated outcomes, particularly in estimating the optimal values for EWIs thresholds. On top of that, we add information on optimal time horizon given the maximisation of the AUROC score over the set of all time lag and prediction horizon combinations.

4.1. Labour market indicators

Among the labour market group EWIs there are four of them belonging to the baseline MIP scoreboard, hence they are given the pre-determined optimal threshold levels, and three auxiliary indicators without such recommended cut-off values. In practically all the cases, the officially adopted thresholds deliver a negative absolute utility function implying that the selected indicators do not provide more useful predictive power over the simple random model. The performance of individual indicators slightly improves once modelling the optimal threshold characteristics are modelled; however, the overall improvement remains only modest.

In the case of three baseline indicators, the officially pre-determined thresholds appear to be excessively accommodating to the needs of the CEE countries, only with the Activity rate as the exception. All implied policy-maker preferences are set below the optimal values of μ obtaining values from 0.73 to 0.81 revealing higher preference weights towards avoiding Type I error, given the frequency of crisis occurrence.

Consequence, the overall very poor AUROC scores (below 0.5) for almost all indicators do not recommend usage of any of these EWIs as they generate more nuisance than precise predictions. Even at the optimal thresholds the inverse aNtS ratios either

strongly underperform unity (Youth unemployment rate, Long-term unemployment rate, Unemployment rate, and Employment) or deliver only very moderate outcomes practically indistinguishable from one (Activity rate, and Participation rate).

In contrast, the behaviour of the youth NEET indicator (i.e. neither in employment nor in education and training) indicator illustrates how strikingly important is to precisely specify the utility-maximising optimal threshold in the assessment of the EWIs. Delivering mediocre outcomes of the overall AUROC score (0.39) calculated across all plausible thresholds it becomes one of the best performing given the aNtS inverse scores (3.36) once focusing solely on the optimal one. In this case, choosing the 25% of the youth NEET as the lower bound delivers both a positive absolute utility and the second best aNtS inverse score in the sample.

Taking into account historical experience of the CEE region, particularly when considering transition period, these outcomes are to be partially expected since the CEE countries have had traditionally long-lasting internal imbalance problems due to the less efficiently functioning labour markets and costs or structural market makeover. Looking at the performance of indicators over time, the AUROC scores surpassing the indicative threshold of 0.5 tend to be concentrated around 1998–2003 and 2008–2010 periods (Appendix 1(a) in Supplemental data). With the forecasting horizon of two to three years, these outcomes might suggest that rather than warning of impending crisis, labour market-related indicators might perform more reliably in its aftermath in short-time horizons.

On the other hand, these indicators tend to capture long-term structural distortions in a real economy, as indicated by the relatively long optimal time horizon of four to six years (Table 1). The highest improvement in the AUROC score once optimising over time horizons is achieved in the case of youth unemployment rate, youth NEET and the total unemployment rate, but with only the YNEET remains distinctly far from the 0.5 benchmark.

In summary, while the outcomes in the Table 1 might still be used by policy-makers to establish a sort of internal tentative thresholds (e.g. 60% for participation rate, 25% for NEET, −5.5 % change in employment rate) the signals emanated by the set of all labour market indicators *in the short run* should be taken with a high caution since they produce significant amount of noise. Recently, even the Ecofin Council in its notes from the January 15 meeting acknowledged that since social and labour market indicators are not relevant for identifying macro-financial risks the developments in these indicators cannot trigger the steps in the MIP procedure (Council of the European Union 2016, §6). However, as persistent structural distortions, particularly longer-lasting adverse consequences of young people unemployment, might ultimately make any crisis more severe, the MIP procedure should take this into account when assessing potential risks in its in-depth review stage.

4.2. External imbalances indicators

As in the previous group of indicators, the possible extent of external imbalances is illustrated by the four baseline indicators (change in nominal unit labour costs, REER, export market share and current account balance) and three auxiliary indicators (change in terms of trade and productivity and current and capital account balance). Contrary to the internal imbalances class of indicators measures of external exposure succeed in rendering three very good performing and two relatively good performing indicators, a

finding that apparently reflects the strong orientation of the CEE economies on their external sector.

The high openness of the CEE region in terms of external trade and reliance on foreign financing might bring about long-lasting adverse consequences of currency and balance of payment crisis. From this perspective, it is of utmost importance for respective authorities to recognise relevant warning signals of external imbalances in ample time. Conversely, relatively strict official thresholds, as imposed by the MIP procedure, tend to deliver only sub-optimal outcomes potentially forcing policy makers in the CEE countries to adopt economically harmful counter-measures. Negative absolute utility associated with an excessive inclination towards conservative behaviour (avoiding Type I error) present in all cases, except a change in the REER indicator (maximum absolute utility, 0.022).

Values of utility-maximising thresholds justify a call for a more relaxed policy-maker attitude, with the optimal value set at a minus 18% change in export market share, minus 7.5% for current and minus 7% for current and capital account balance and minus 4% change for terms of trade indicators. In all four cases, the AUROC score indicates better than random model performance with inverse aNtS ratio achieving results significantly different from unity, thus signifying a positive information value of the underlying EWI.

As the supplementary MIP scoreboard indicator, the current and capital account indicator performs even more efficiently (aNtS 2.27) than the generally accepted current account balance taking into account possible inflow of long-term capital transfers into the CEE economies over the past two decades.

On the positive side, the upper cap on the current account balance set on 6% of GDP does not deliver meaningful outcomes for the estimation of indirect utility since, historically, all the countries in the CEE region have been operating in the environment of predominantly negative current account balances. Therefore, the optimal limit of 1% must be, therefore, viewed in the light of this evidently different experience rather than as an issue of lavishly positive CA balances in Germany or Netherlands that, allegedly, are behind the decision to even set the upper bound to 6%.

Appreciation of the REER is usually viewed as a signal warning from the possible deterioration in price competitiveness. Conversely, persistent and excessive REER depreciation might spill over into the domestic economy by driving up the imported inflation and/or limiting the healthy increase in domestic price level and nominal wages. Therefore, optimal bounds should therefore be specified with an utmost care, hence placing the upper limit even below the official EC threshold at 9% and lower limit close to its pre-determined values at minus 10.5%. In both cases, that the inverse aNtS ratio positively differs from unity confirms the added value of this EWI using the optimal threshold values.

The change in the nominal unit labour costs, as a measure of economic competitiveness, does perform poorly even with the upper optimal threshold double the size of the official one, which is a finding further reinforced by the negligible outcome for the inverse aNtS ratio. The second indicator that directly captures the concept of competitiveness, decrease in productivity rate, does not deliver desirable outcomes either with negative maximum absolute utility (−0.0035) and aNtS ratio smaller than one (0.701).

External imbalances indicators might be potentially grouped into a set of late indicators (ToT, positive change in REER, positive current account balance, current and capital account balance, and productivity) and very early EWIs (NULC, negative change in REER, EMS, positive current account balance) given the AUROC-maximising predictive

time horizon (Table 1). The existence of such a bisection line might point out to the prevalent differences between nominal price devaluation previously materialised in the balance of payments measures and deterioration in competitiveness nominal effects of which are yet to be realised. Contrary to the labour and social market indicators, the AUROC-based yearly performance of EWIs in this group does not tend to be clustered around one specific historical period. While current and capital account performance tends to improve while approaching the recent global crisis in 2008, the REER and EMS behaviour allows for the detection of impending crisis across an entire time horizon making use of these indicators potentially more robust to different types of crises.

4.3. Indebtedness related indicators and others

The reliance of the CEE economies on external sources of financing is reflected in the findings representing the third list of MIP indicators. Half of the indicators belong to the baseline scoreboard indicators, the remainder supplement the MIP scoreboard as auxiliary indicators. As in the case of implied preferences for a positive current account balance, the data for gross external debt and private sector debt do not allow the revealing of the policy-maker implicit preferences or the derivation of the maximum absolute utility.

The indicators of FDI inflows and private sector debt approximated by private credit-to-GDP might potentially generate useful signals for shorter term crisis prediction, according to the AUROC score and aNtS ratio outcomes. However, few of instruments, however, perform relatively poorly obtaining modest success if measured by the inverse aNtS ratio (net IIIR with 1.26, gross external debt with 1.32). Economically, the levels of threshold adopted by the MIP procedure are either excessively strict (optimal net IIP being twice the recommended net IIR balance) or overly relaxed (public and private sector debt) for the economic conditions of the CEE countries. Not surprisingly, the optimal threshold levels mirror the underlying distribution of the values of individual indicators with average private and public sector debt levels much lower than those of other European developed economies.

Level variables, such as FDI stocks, gross external debt or private sector debt achieve only one significant peak in AUROC score around year 2004 making them also less robust to an indication of different types of economic or financial crises (Appendix 1(c) in Supplemental data). Change in time horizons, as indicated by maximising the individual AUROC scores, does not lead to a significant improvement in the overall AUROC scores. In contrast, the FDI inflows indicator gains in predictive power in year 2002 and scores higher than 0.5 benchmark values in the majority of the subsequent years the reflecting both the historical importance of foreign capital in the CEE region as well as the flow dimension of the indicator. The optimal time horizon set at two years classifies this indicator as a late warning indicator among others EWIs. Public sector debt indicator achieves the best performance at longer time horizons with a significantly improved AUROC score of 0.650 making it an influential indicator to consider as part of the MIP in-depth assessment.

The HPI indicator performs best among all 24 indicators in all three criteria considered: the AUROC score, the inverse aNtS and the absolute utility value. Furthermore, the derived optimal threshold undoubtedly resembles that imposed by the MIP procedure (6%). The eminent success of this indicator might be partially attributed to the underlying data structure with a small number of observations (70) clustered within the post 2008 crisis period. As generally argued real asset price bubbles had accentuated

problems in domestic financial markets and led to an outburst of several domestic debt crises. Interestingly, while usually considered a late EWI, the AUROC-maximising criterion justifies its use even as a early warning indicator performing at optimum in a three to five year horizon.

The problem of "bad" imbalances (Eichengreen 2010) might be reflected in the second indicator, the level of gross fixed capital formation, once looking at its performance is examined over a longer time horizon (four to six years). If an economy indulges in prolonged periods of over-consumption or spends borrowed capital on meaningless investment projects, this kind of behaviour might lead to increasing the probability of a crisis. From an economic point of view, an excessively low value of the optimal threshold (12%) needs to be understood in the broader context of the CEE region's historical experience with an average investments ratio fluctuating around 25%.

4.4. Discussion

The Macroeconomic imbalance procedure implemented at the EU28 level represents systematic EWIs system bounding the heterogeneous economies by a set of predetermined common rules. As such this procedure represents an important step forward by building an efficient crisis prevention system that could not only help to minimise potential costs of a normal business cycle downturn but could warn of an impending severe crisis far beyond the normal economic cycle. The decision to include a relatively wide variety of indicators into the final MIP scoreboard might reflect the need to create a robust system that would be able to monitor the accumulation of macroeconomic imbalances that differ in their nature. Even if not ultimately materialised in a single crisis, their existence might still serve as an amplifier of future adverse shocks.

However, the MIP procedure encounters significant limitations particularly if not taking into account heterogeneous socio-economic conditions of individual member states. While many of the CEE countries might potentially use the MIP as a good starting point to build their own customised early warning indication systems, the optimal thresholds maximising policy-maker's absolute utility differ significantly from those proposed for the set of advanced economies.

The assessment of the individual performance of specific indicators on the MIP scoreboard should represent an initial step towards the creation of reliable EWS. Based on our results, only a few of the indicators have strong and robust prediction power (REER, EMS, current and capital account balance, and FDI inflows) for the CEE countries. Moreover, certain indicators require a relatively long time period for their construction (e.g. 10 years average), which can limit the general usage of this procedure in developing countries. Conversely, one composite indicator with relatively strong prediction power could be constructed once combining indicators with higher predictive power and low mutual correlation.

The convergence process in the CEE economies might help them to sustain even higher levels of current account imbalances linked to an expected increase in nominal unit labour costs or a higher inflow of long-term investment capital, as indicated by more relaxed optimal threshold values for many external imbalances-related EWIs. Conversely, exposure towards speculative foreign capital and an increase in the public and private level of indebtedness thanks to the access to international capital markets must be investigated carefully as the indicative thresholds point towards a much more conservative policy-maker stance than for mature advanced economies (for comparison, see Knedlik 2014, 2015). Long-lasting structural distortions in the labour market of the

CEE countries might be behind the relatively poorer performance of related EWIs. However, the imbalances in this market might continue to magnify the severity of the potential crisis, which is a possibility that should be addressed as part of the in-depth review process once other relevant signals in the first MIP stage are issued.

Given these considerations, any potential future analysis should make an endeavour towards more country-specific optimal thresholds for selected EWIs (e.g. Comunale 2015) in addition to determining a better specification of proposed indicators. Without more detailed theory-based specifications of long-term equilibrium values the empirical analysis might only proceed as far as historical memory allows which is, particularly in relatively young economies, not that far.

5. Conclusions

The creation of the MIP framework in the EU28 country group as an EWI system meant a huge step forward to an efficient and reliable system of crisis prediction. However, many authors have criticised the system's feature based on its backward looking orientation, choice and specification of EWIs in addition to their "ad hoc" indicative threshold and lack of country group-specific features. Acknowledging this critique, this article applies a set of baseline and auxiliary indicators incorporated into the MIP framework in the conditions of the CEE countries to assess their predictive power given the policy pre-determined and optimal thresholds.

Our results suggest that the optimal thresholds proposed within the common MIP framework might be either excessively accommodative (public and private, total or external debt levels), overly conservative (current account balance, export market share, and nominal unit labour costs), or with low informative value (labour market characteristics) once used for a set of less mature economies in transition. Indicators that show higher predictive power belong predominantly to the group of external imbalances indicators (current and capital account balance, export market share, and REER) and are accompanied by a change in the house price index and young people outside education and training indicators.

Acknowledgments

Thanks are due to seminar participants at the National Bank of Slovakia for several valuable comments, with special thanks going to Mr Gertler, Mr Suster and Mr Toth. We also extend our gratitude to Mr Domonkos, Mr Ostrihon and Ms Sikulova.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Horizon 2020 program [grant number 649261–FIRSTRUN].

Supplementary material

Supplementary material for this article can be accessed here. [https://doi.org/10.1080/17487870.2017.1364642]

Notes

1. We are considering wider set of 17 CEE countries (Central and Eastern European Countries) in this paper. Detailed information are available in the Appendix 2 in Supplemental data.
2. We calculate AUROC score for all possible combinations of time lag and time window during which an event might occur with cap set on three years for both lag and time window. While this allows for differentiating between late and ultra-early indicators by searching for optimum lag-window combination outcomes might be biased due to corresponding reduction of the data-set by excluding last year observations.
3. Since the calculation of the AUROC score is based on ratio of true and false warning rate, it might happen that for specific years the AUROC score cannot be computed. This might occur in a case when there does not exist a combination when EWI issued a signal that did not indicated a crisis (B). In our sample this describes a situation in 2005–2007 period when crisis was either indicated in all cases or every no-crisis event left was correctly predicted.
4. The most recent report by the European Commission (European Commission 2016, 31) does not provide any clear direction how to empirically capture the notion of macroeconomic imbalances but rather uses a very vague definition of "trends or states that could jeopardise macroeconomic stability if not corrected."

ORCID


Maria Siranova http://orcid.org/0000-0002-0267-0943

References

Alcidi, C., and D. Gros. 2014. *Implications of EU Governance Reforms. Rationale and Practical Application*. ETLA Reports No 25, 6 May.

Alessi, L., and C. Detken. 2009. "'Real Time' Early Warning Indicators for Costly Asset Price Boom/Bust Cycles: A Role for Global Liquidity." ECB Working Paper No 1039, March.

Alessi, L., and C. Detken. 2014. "On Policymakers' Loss Functions and the Evaluation of Early Warning Systems: Comment." *Economic Letters* 124 (3): 338–340.

Alessi, L., A. Antunes, J. Babecky, S. Baltussen, M. Behn, D. Bonfim, O. Bush, et al. 2015. *Comparing Different Early Warning Systems: Results from a Horse Race Competition among Members of the Macro-prudential Research Network*. MPRA Paper 62194, University Library of Munich.

Babecký, J., T. Havránek, J. Matějů, M. Rusnák, K. Šmídková, B. Vašíček. 2013. "Leading Indicators of Crisis Incidence: Evidence from Developed Countries." *Journal of International Money and Finance* 35: 1–19.

Candelon, B., E. Dumitrescu, and C. Hurlin. 2012. "How to Evaluate an Early-warning System: Toward a Unified Statistical Framework for Assessing Financial Crises Forecasting Methods." *In IMF Economic Review* 60: 75–113.

Comunale, M. 2015. *Current Account and REER Misalignments in Central Eastern EU Countries: An Update Using the Macroeconomic Balance Approach*. BOFIT Discussion Papers 28/2015.

Council of the European Union. 2016. Outcome of Proceedings, Alert Mechanism Report 2106 – ECOFIN Council Conclusions. January 15. https://data.consilium.europa.eu/doc/document/ST-5318-2016-INIT/en/pdf.

Csortos, O., and Z. Szalai. 2013. *Assessment of Macroeconomic Imbalance Indicators*. MNB Bulletin, Magyar Nemzeti Bank, October.

Csortos, O., and Z. Szalai. 2014. *Early Warning Indicators: Financial and Macroeconomic Imbalances in Central and Eastern European Countries*. MNB Working Papers 2, Magyar Nemzeti Bank.

Detken, C., O. Weeken, L. Alessi, D. Bonfim, M. M. Boucinha, C. Castro, S. Frontczak, et al. 2014. *Operationalising the Countercyclical Capital Buffer: Indicator Selection, Threshold Identification and Calibration Options*. European Systemic Risk Board Occasional Paper Series No 5, June.

Domonkos, T., F. Ostrihon, I. Sikulova, and M. Siranova. 2016. "Analysing Macroeconomic Imbalances in the EU." Ecomod Conference Proceedings, Lisbon, Portugal, July 8. https://ecomod.net/conferences/ecomod2016?tab-downloads.

Domonkos, T., F. Ostrihon, I. Sikulova, and M. Siranova. 2017. "Analysing the Relevance of the MIP Scoreboard's Indicators." *National Institute Economic Review* 239 (1): R32–R52.

Drehmann, M., and M. Juselius. 2013. *Evaluating Early Warning Indicators of Banking Crisis: Satisfying Policy Requirements*. BIS Working Papers No 421, Bank for International Settlements.

Edison, H. J. 2003. "Do Indicators of Financial Crises Work? An Evaluation of an Early Warning System." *International Journal of Finance & Economics* 8 (1): 11–53.

Eichengreen, B. 2010. *Imbalances in the Euro Area*. Berkeley, CA: University of California. November. https://eml.berkeley.edu/~eichengr/Imbalances_Euro_Area_5-23-11.pdf

European Commission. 2011. *Scoreboard for the Surveillance of Macroeconomic Imbalances: Envisaged Initial Design*. Commission Staff Working Paper, 8.11.2011, SEC(2011) 1361 Final.

European Commission. 2012a. *Completing the Scoreboard for the MIP: Financial Sector Indicators.* Commission Staff Working Document, 14.11.2012, SWD (2012) 389 Final.

European Commission. 2012b. *Scoreboard for the Surveillance of Macroeconomic Imbalances*. Occasional Papers 92, European Economy, February 2012.

European Commission. 2015. *Adding Employment Indicators to the Scoreboard of the Macroeconomic Imbalance Procedure to Better Capture Employment and Social Developments*. Ref. Ares(2015)5426195 - 27/11/20154.9.2015, September 4 2015. https://ec.europa.eu/info/sites/info/files/employment_indicators_mip_en.pdf

European Commission. 2016. *The Macroeconomic Imbalance Procedure. Rationale, Process, Application: A Compendium*. Institutional Paper 039, November 2016.

Hallwirth, V. 2014. "Überwachung Der Wettbewerbsfähigkeit in Der Europäischen Wirtschafts- Und Währungsunion." [The Surveillance of Competitiveness in the Economic and Monetary Union.] *Wirtschaftsdienst* 94 (11): 798–805.

Kaminsky, G., and C. Reinhart. 1999. "The Twin Crises: The Causes of Banking and Balance-of-Payments Problems." *American Economic Review* 89 (3): 473–500.

Kaminsky, G., S. Lizondo, and C. Reinhart. 1998. "Leading Indicators of Currency Crisis." *IMF Staff Papers* 45 (1): 1–48, International Monetary Fund.

Kamps, C., R. De Stefani, N. Leiner-Kilinger, R. Ruffer, and D. Sondermann. 2013. "Would the Strengthened EU Fiscal and Economic Governance Framework Have Helped Signaling Sovereign Debt Crises?" Paper presented at Banca d'Italia 15th Public Finance Workshop.

Knedlik, T. 2012. *The European Commission's Scoreboard of Macroeconomic Imbalances – The Impact of Preferences on an Early Warning System*. IWH Discussion Papers No 10, Halle Institute for Economic Research, November.

Knedlik, T. 2014. "The Impact of Preferences on Early Warning System – The Case of the European Commission's Scoreboard." *European Journal of Political Economy* 34: 157–166.

Knedlik, T. 2015. "The Usefulness of the Scoreboard of the Macroeconomic Imbalances Procedure in the European Union: Potentials for Reform." *European Policy Analysis*, 14 (Jun.): 1–12.

Moschella, M. 2014. "Monitoring Macroeconomic Imbalances: Is EU Surveillance More Effective than IMF Surveillance?" *Journal of Common Market Studies* 52 (6): 1273–1289.

Sarlin, P. 2013. *On Policymakers' Loss Functions and the Evaluation of Early Warning Systems.* ECB Working Paper Series No 1509, European Central Bank.

Sikulova, I. 2015. "Macroeconomic Imbalance Procedure in the Euro Area – Ex Post Analysis." European Scientific Journal, August/Special/edition: 161–175. http://eujournal.org/index.php/esj/article/view/6147