Performance of the Macroeconomic Imbalance Procedure in Light of Historical Experience in CEE Region

Draft, 31/10/2016

Maria SIRANOVA[1], Marek RADVANSKÝ[2]

Abstract

The creation of the MIP framework in the EU28 country group as an early warning indicator system meant a huge step forward an efficient and reliable system of crisis prediction. Yet, many authors have criticised its feature based on its backward looking orientation, choice and specification of early warning indicators along with their "ad-hoc" indicative threshold and lack of country group-specific features. Acknowledging this critique, this article applies set of 24 baseline and auxiliary indicators included into the MIP 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 thresholds proposed within the common MIP framework might either be too accommodative (public and private, total or external debt levels), overly conservative (current account balance, export market share, nominal unit labour costs), or not helpful at all (labour market characteristics) once used for 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, REER). There are only a few other indicators with relatively good prediction, such as change in the house price index or share of young people outside education and training.

Key words: early warning indicators, macroeconomic imbalance procedure, CEE region

JEL classification: G01, F47, F53

Acknowledgments

This research was financed under the Horizon 2020 program, project No. 649261–FIRSTRUN. Thanks are due seminar participants at the National Bank of Slovakia of several valuable comments, with special thanks going to Mr Gertler, Mr Suster and Mr Toth.

[1] Corresponding author.

Mária Širaňová, Institute of Economic Research SAS, Šancová 56, 811 05 Bratislava, Slovak Republic; e-mail: maria.siranova@savba.sk & Department of Banking and International Finance, Faculty of National Economy, University of Economics, Dolnozemska cesta 1, 852 01 Bratislava, Slovak Republic.

[2] Marek Radvanský, Institute of Economic Research SAS, Šancová 56, 811 05 Bratislava, Slovak Republic; e-mail: marek.radvansky@savba.sk

Introduction

Early warning indicators (EWIs) are an essential component of macro-prudential policies. The EWIs system is a system or procedure to warn of the potential or impending problem. The recent financial crisis highlighted the issue of monetary union heterogeneity in terms of different stage of economic development accompanied by unsuccessful convergence. In such conditions, it is crucial to prevent the occurrence of asymmetric shocks in the EMU by early identifying macroeconomic imbalances and by adopting necessary measures to reduce or eliminate the existing imbalances (Essl and Stiglbauer, 2012).

Only handful of studies (e.g. Csortos and Szalai, 2013; Knedlik, 2014, 2015; Domonkos et al., 2016b) have so far empirically assessed the efficiency of new Macroeconomic Imbalance Procedure (MIP) introduced recently at the EU28 level (European Commission, 2012a, 2012b). Yet, the need for efficient and reliable early warning indicators system has been widely recognised. This is true especially for the CEE[3] economies that are highly sensitive towards the adverse evolution in the core EU28 countries via their external trade and financial linkages. While the different forms of Early Warning Indicators (EWI) system has been tested on conditions of many developed and developing economies over the course of last two decades (started by seminar paper by Kaminsky et al., 1998), the CEE region as a separate country group has been widely neglected. At the same time, recent studies by Knedlik (2014) and Knedlik (2015) argue that grouping of countries based on their level of economic development and internal homogeneity might lead to a significant increase in the performance of EWI systems.

This article supplements broad literature on EWI systems by evaluating the performance of individual EW indicators included in the MIP framework, the only early warning indicator system officially adopted for a set of economically divergent EU28 countries, in the conditions of CEE countries.

As the key objective of the MIP procedure is to provide sound input for policy makers, their preferences are taken into account and modelled by a loss function. The most reliable and efficient EWIs therefore integrate both, a high predictive power and minimum costs related to their potential failure. This paper departs from the paper by Csortos and Szalai (2013) and Domonkos, Ostrihon and Sikulova (2016) by extending the dataset used for analysis and enhancing the methodology by implementing policy maker loss function as specified in Sarlin (2013), Csortos and Szalai (2014), and Knedlik (2014). We estimate optimum thresholds for all the indicators, as the common critique of the MIP procedure questions the specification of the trigger mechanism by "rule of thumbs" rules (Kamps et al., 2013).

In the empirical part of the paper, we apply signalling approach to assess the performance of different individual EWIs, both baseline and auxiliary, included into the MIP scoreboard. Our sample consist of 17 CEE economies covering yearly time series starting in 1991 to 2014.

Based on the previous, our analysis consists of three steps. Firstly, the efficiency of MIP benchmark EWIs along with various alternatives is tested individually by the signalling approach. Secondly, the performance of different EWIs combinations is evaluated by three different concepts (absolute utility, AUROC score and inverse aNtS ratio). Thirdly, optimal thresholds are compared with the official pre-determined ones in order to specify the extent of too accommodating or too restrictive policy-maker behaviour. The best performers among the benchmark and potential EWIs are selected given the outcomes from this three-step analytical procedure.

Our outcomes support a call for a more country group (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 too accommodative (public and private, total or external debt levels), overly

[3] We are considering wider set of 17 CEE countries (Central and Eastern European Countres) in this paper. Detailed information about covered countries is in appendix.

conservative (current account balance, export market share, nominal unit labour costs), or not helpful at all (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, REER) and are accompanied by change in the house price index and young people outside education and training indicators.

The paper is structured as follows. Firstly, we shortly 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 concludes findings and reveals recommendations for further development of MIP related policies.

1. Macroeconomic Imbalance Procedure as the EWI system

The EWI procedure serves as a powerful tool that might help the policy makers to take pre-emptive steps that would prohibit macroeconomic imbalances to deepen already existing heterogeneity that would potentially lead to severe economic crisis, such as the recent debt crisis in 2009. EWIs in this context must not only have sound statistical forecasting power, but also need to satisfy several additional requirements. By Drehmann and Juselius (2013) there are three crucial criteria that should identify most powerful EWIs: i) the appropriate timing requires a signal that does not come at a very early stage due to potentially high costs of its mitigation but has 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 recognized. In 2010 when the European Commission adopted a legislative package consisting of six proposals, the so-called *six-pack* legislation, which aims to reinforce the monitoring and the surveillance of fiscal, macroeconomic and structural reform policies in the EU and the euro area compared to previously applied legislation. 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. It is a monitoring mechanism that aims to identify potential risks early on, correct the imbalances that are already in place and prevent them from re-emerging. The MIP presents an integral part of economic policy coordination within the EU and in particular within the EMU, where the need for such policy coordination is even stronger. In November 2011 the set of economic indicators as part of the MIP scoreboard used to detect macroeconomic imbalances was presented by the Ecofin/Eurogroup.

The MIP is built as a “two-step approach”. The first step is represented by an alert mechanism which works as a filter. As an early warning system, the objective of the alert mechanism is to focus attention on observed risks early on and identify the Member States for which, in the second step, more in-depth analysis appears warranted so as to assess their vulnerability and substantiate policy recommendations if appropriate (European Commission, 2012b). So it is in-depth analysis, and not the alert mechanism, which provides the basis for any recommendations to be addressed to the Member States under the preventive or corrective arm of the MIP.

The more detailed analysis of the efficiency of the MIP procedure is of high importance for both researchers and policy makers taking into account existing critique and possible shortcomings listed in the literature. Moschella (2014) points out that the MIP does not provide for mechanisms to prevent political considerations from interfering with the decision to activate sanctions and on 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 on the most efficient set of EWIs based on their individual or grouped performance reflecting the policy makers loss function. The Domonkos, Ostrihon and Sikulova (2016) study concludes that the highest predictive ability may be attributed to the indicators of the external (current account balance, net international investment position, export market shares), and internal imbalances measuring level of indebtedness (private sector debt, general government gross debt) or labour market inefficiencies (long-term unemployment rate, youth unemployment rate). 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). On top of that, 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).

2. Methodology

Our methodology assess predicting power of the different EWIs introduced as part of the MIP along with specifying optimal EWIs thresholds given the preferences of the decision maker in terms of committing various types of errors.

2.1. 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 (Sarlin, 2013; Csortos and Szalai, 2013; Csortos and Szalai, 2014; Domonkos, Ostrihon and Sikulova, 2016). Implementing EWIs system can be an useful tool for policymakers to tackle with forthcoming event. On the other hand, implementing EWIs system brings about the possibility of event prediction error. Prediction errors rely on a level of (arbitrary) individual preferences of policymakers, 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 way:

	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). 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 entire system (Kaminsky et al., 1998; Kaminsky and Reinhart, 1999; Alessi et al., 2015) by comparing false and true warning rate in the following way:

$$aNtS = \left({}^{B}/_{(B+D)}\right)/\left({}^{A}/_{(A+C)}\right) \quad [1]$$

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 gives rise 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 true positive rate ($A/(A+C)$) and false positive rate ($B/(B+D)$) for every possible threshold value. High thresholds generating small number of signals will be located close to 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 indicator having an informative value better than the random model.

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

The standard Alessi-Detken type of loss function is defined in the following way:

$$L(\mu) = \mu T_1 + (1-\mu)T_2 \qquad [2]$$

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 wrongly signalled crisis.

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

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

$$L(\mu) = \mu T_1 P + (1-\mu)T_2(1-P) \qquad [3]$$

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) at all. Positive values of absolute utility functions are treated as desirable outcomes signifying added value of following the EWI system recommendations over the random model. Hence the Sarlin-type absolute utility function is of the following form:

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

As apparent, the specification of the policy maker's risk aversion coefficient μ represents a crucial issue in the process of EWI assessment. In 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 ESRB (2014) recent publication 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. Sarlin-type of loss and utility function compensate 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 paper (2013).

In the Sarlin-type utility function framework, the policy maker might be required to state his preferences in a very precise way 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 ten percent 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 underlying leading EWIS in the pre-crisis period (Kaminsky et al., 1998). To name few of the studies, Drehman and Juselius (2013) opt for one and a half years to five years interval, Alessi and Detken (2009) up to one and a half years window, Knedlik (2014) for one-to-two years lag. Paper by Babecky 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 way. Acknowledging that policy maker require 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 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.

2.2. 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 (Domonkos, Ostrihon and Sikulova, 2016; Domonkos et al., 2016a; Csortos and Szalai, 2014), significant negative deviation of real GDP from trend (Csortos and Szalai, 2013; Csortos and Szalai, 2014), list of cross-country crisis complied by other authors (Drehmann and Juselius, 2013). The study by Babecky et al. (2013) advocates continuous indicator of crisis real costs integrating output and employment loss with fiscal deficits.

Table 2 Crises windows in the CEE region

	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Albania						X					X							X		X		X		
Bosnia and Herceg.	…	…	…															X						
Bulgaria								X									X	X						
Croatia	…	…	…	…				X									X	X						
Czechia						X											X	X						
Estonia	…	…	…	…				X									X	X						
Hungary	X														X	X		X						
Kosovo	…	…	…	…	…	…	…	…	…														X	
Latvia	…	…	…	…	X												X	X						
Lithuania	…	…	…	…				X									X	X						X
Macedonia										X								X						
Montenegro	…	…	…	…	…	…		X										X						
Poland							X	X		X	X							X			X	X		
Romania						X								X				X						
Serbia	…	…	…	…				X										X					X	
Slovakia	…	X					X	X										X						
Slovenia	…	…	…	…			X			X							X	X						

Note: Crisis period indicated by "X" signifies period when the real GDP growth drops more than one standard deviation from its 5-year average. Year indicated by "…" has no data available.

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. 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 standardizes estimates of the potential GDP for CEE region up to 1991 are not available we instead focus on crisis periods captured by deviations of the real GDP growth from its 5-year average by more than one standard deviation, similar to Crostos and Szalai (2013) and Csortos and Szalai (2014) in use of real

GDP and to Knedlik (2014) in use of standard deviations from a long-term trend. In general, 5-year window corresponds to the average length of business cycle.

In the context of CEE region (Table 2) drops in real GDP growth bigger than one standard deviation split the sample into two main crisis periods clustering around 1997-1998 and 2007-2008 crisis years, along with handful of country-specific slumps in economic activity. This setup allows for investigation of MIP performance under the wider set of circumstances (similar to Csortos and Szalai, 2013; Csortos and Szalai, 2014) rather than focusing on the post 2000 period (Knedlik, 2012; Knedlik, 2014; Domonkos et al., 2016). On average, the frequency of crisis occurrence is set around 18 percent, however depending on the number of observations available per each indicator (Appendix I-IV).

2.3. Dataset and EWIs specification

Nowadays the scoreboard used for AMRs includes a list of fourteen indicators with their indicative thresholds (including an indicator on the financial sector and new employment indicators) as well as a supplementary list of other 28 auxiliary indicators without 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 into the scoreboard, the transformations used and the determination of threshold values is provided in the European Commission (2012b) or European Commission (2015a).

Due to the both data availability issues common among the CEE countries and non-existent reporting practices for selected indicators in general, the final list of indicators used in this exercise include 7 indicators for internal imbalances group mainly describing labour market conditions (Appendix I), additional 7 indicators for external imbalances and competitiveness group (Appendix II), next 7 domestic and foreign indebtedness-related indicators (Appendix III) and is complemented by last three auxiliary EWIs (Appendix IV).

In total, dataset consists of 17 countries from the CEE region covering the period from 1991 to 2014.[4] For some countries and variables, the data cover 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 V.

3. Results and discussion

In the Appendices 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, especially in estimating the optimal values for EWIs thresholds.

3.1. Labour market indicators

Among the labour market group EWIs there are four of them belonging to the baseline MIP scoreboard, hence 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 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; however the overall improvement remains only modest.

In the case of three baseline indicators, the officially pre-determined thresholds seem to be too accommodating to the needs of the CEE countries, only with the Activity rate being the exception. All implied policy-maker preferences are set below the optimal values of μ obtaining values from 0.73 to

[4] Albania, Bosnia, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Kosovo, Latvia, Lithuania, FYROM, Montenegro, Poland, Romania, Serbia, Slovakia, Slovenia.

0.81 revealing higher preference weights towards avoiding Type I error, given the frequency of crisis occurrence.

As a 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 UR, Long-term UR, UR, Employment) or deliver only very moderate outcomes practically indistinguishable from one (Activity rate, Participation rate).

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

In summary, while the outcomes in the Appendix I might still be used by policy makers to set up some sort of internal tentative thresholds (e.g. 60 percent for participation rate, 25 percent for NEET, -5.5 % change in employment rate) the signals emanated by the set of all labour market indicators should be taken with a high caution since they produce significant portion of noise.

Taking into historical experience of the CEE region, especially 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. Hence, signals issued by the labour market-oriented indicators will inherently embody a significant portion of false signals, especially during the relatively short-term time window. One might, therefore, expect that the performance of the labour market-focused indicators is to be improved by increasing the time window or by the inclusion of more advanced economies as a control group.

3.2. External imbalances indicators

As in the previous group of indicators, 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 to render 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 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. On the other hand, 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-maximizing thresholds justify a call for a more relaxed policy-maker attitudes, with the optimal value set on minus 18 percent change of export market share, minus 7.5 percent for current and minus 7 percent for current and capital account balance and minus 4 percent 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.

Current and capital account indicator being the supplementary MIP scoreboard 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 six percent of GDP does not deliver meaningful outcomes for 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. The optimal limit of one percent must be, therefore, viewed in light of this evidently different experience rather than an issue of lavish positive CA balances in Germany or Netherlands that, allegedly, stand behind the decision to even set an upper bound to 6 percent.

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 limit the healthy increase in domestic price level and nominal wages. The optimal bounds should therefore be specified with an utmost care, hence placing upper limit even below the official EC threshold at 9 percent and lower limit close to its pre-determined values at minus 10.5 percent. In both cases, the inverse aNtS ratio positively different from unity confirms added value of this early warning indicator 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, a finding further reinforced by negligible outcome for inverse aNtS ratio. The second indicator directly capturing the concept of competitiveness, drop in productivity rate, does not deliver desirable outcomes either with negative maximum absolute utility (-0.0035) and aNtS ratio smaller than one (0.701).

3.3. Indebtedness related indicators and others

Reliance of 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 rest of it supplement the MIP scoreboard as auxiliary ones. As in the case of implied preferences for positive current account balance, the data for gross external debt and private sector debt do not permit to reveal nor policy maker implicit preference nor derive maximum absolute utility.

The indicator of FDI inflows and private sector debt approximated by private credit-to-GDP might potentially generate useful signals for crisis prediction, according to the AUROC score and aNtS ratio outcomes. 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 too strict (optimal net IIP being twice the recommended net IIR balance) or too relaxed (public and private sector debt) for economic conditions of CEE countries. Not surprisingly, the optimal threshold levels mirror the underlying distribution of values of individual indicators with average private and public sector debt levels much lower than those of other European developed economies.

Among the group of others, two EWI stands out, namely a change in the house price index and level of gross fixed capital formation.

In the former case, the HPI indicator performers best among all 24 indicators in all three criteria considered: the AUROC score, inverse aNtS and absolute utility value. On top of that, the derived optimal threshold undoubtedly resembles the one imposed by the MIP procedure (6 percent). 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 let to an outburst of several domestic debt crises.

The problem of "bad" imbalances (Eichengreen, 2010) might be reflected in the second relatively successful indicator, the level of gross fixed capital formation. If economy indulges in the prolonged periods of over-consumption or spends borrowed capital on meaningless investment projects, this kind of behaviour might lead to increasing probability of a crisis. From economic point of view, too low value of optimal threshold (12 percent) needs to be understood in the broader context of CEE region historical experience with an average investments ratio fluctuating around 25 percent.

3.4. Discussion

The macroeconomic imbalance procedure implemented at the EU28 level represents systematic early warning indicators system bounding the heterogeneous economies by a set of predetermined common rules. As such it represents an important step forward by building an efficient crisis prevention system that could help to minimise potential costs of a normal business cycle downturn. Yet, it has its significant limitations especially 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 makers absolute utility differ significantly from those proposed for the set of advanced economies.

Assessment of the individual performance of specific indicators at MIP scoreboard should be an initial step to provide reliable EWS. Based on the results, only few of the indicators have strong prediction power. Moreover, some of the indicators require a relatively long time period for their construction (e.g. ten years average), which can limit the general usage of this procedure in many countries. On the other hand, one composite indicator with relatively strong prediction power could be constructed once combining indicators with higher predictive power and low mutual correlation.

Convergence process in the CEE economies might help them to sustain even higher levels of current account imbalances linked to expected increase in nominal unit labour costs or higher inflow of long-term investment capital. On the other hand, exposure towards speculative foreign capital and increase in public and private level of indebtedness thanks to the access to international capital markets must be investigated carefully as the indicative thresholds point towards much conservative policy maker stance than in the case of mature advanced economies (see Knedlik, 2014 and Knedlik, 2015 for comparison). In light of these considerations, any potential future analysis should make an endeavour towards more country-specific optimal thresholds for selected EWIs along with 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, especially in relatively young economies, not that far.

Conclusions

The creation of the MIP framework in the EU28 country group as an early warning indicator system meant a huge step forward an efficient and reliable system of crisis prediction. Yet, many authors have criticised its feature based on its backward looking orientation, choice and specification of early warning indicators along with their "ad-hoc" indicative threshold and lack of country group-specific features. Acknowledging this critique, this article applies set of baseline and auxiliary indicators included into the MIP framework on the conditions of 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 either be too accommodative (public and private, total or external debt levels), overly conservative (current account balance, export market share, nominal unit labour costs), or not helpful at all (labour market characteristics) once used for 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, REER) and are accompanied by change in the house price index and young people outside education and training indicators.

References

1. Alcidi, C. – Gros, D. (2014). Implications of EU Governance Reforms. Rationale and Practical Application. ETLA Reports No 25, 6 May 2014.
2. Alessi, L. – Detken, C. (2009*). 'Real time' early warning indicators for costly asset price boom/bust cycles: A role for global liquidity*. ECB Working Paper No 1039, March 2009.
3. Alessi, L. – Detken, C. (2014*).* On policymakers' loss functions and the evaluation of early warning systems: Comment. In *Economic Letters*, 124(3), pp. 338-340.
4. Alessi 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.
5. Babecky, J. et al. (2013). Leading indicators of crisis incidence: Evidence from developed countries. In Journal of International Money and Finance, 35, pp. 1-19.
6. Csortos, O., and Szalai, Z. (2013). *Assessment of Macroeconomic Imbalance Indicators*. MNB Bulletin, Magyar Nemzeti Bank, October 2013.
7. Csortos, O., and Szalai, Z. (2014). *Early Warning Indicators: Financial and Macroeconomic Imbalances in Central and Eastern European Countries*. MNB Working Papers 2, Magyar Nemzeti Bank.
8. Domonkos, T – Ostrihon, F. – Sikulova, I. (2016). *Ex-post Evaluation of the MIP Scoreboard Indicators Efficiency in Predicting Financial and Macroeconomic Imbalances in the EU*. Forthcomming.
9. Domonkos, T. – Ostrihon, F. – Sikulova, I. - Siranova, M. (2016). *Analyzing Macroeconomic Imbalances in the EU*. Ecomod Conference Proceedings, Lisbon, Portugal, July 6 – July 8, 2016. Downloadable at: http://ecomod.net/conferences/ecomod2016?tab=downloads
10. Drehmann, M. – Juselius, M. (2013). *Evaluating Early Warning Indicators of Banking Crisis: Satisfying Policy Requirement*s. BIS Working Papers No 421, Bank for International Settlements.
11. Edison, H. J. (2003). Do indicators of financial crises work? An evaluation of an early warning system. In *International Journal of Finance & Economics*, 8(1), pp. 11-53.
12. Eichengreen, B. (2010). Imbalances in the Euro Area. November 2010, University of California, Berkeley. Downloadable at: http://eml.berkeley.edu/~eichengr/Imbalances_Euro_Area_5-23-11.pdf
13. Essl, S. – Stiglbauer, A. (2012). Prevention and Correction of Macroeconomic Imbalances: The Excessive Imbalances Procedure. In *Monetary Policy & The Economy*, Q4/11, Oesterreichische National Bank, pp. 99-113.
14. European Commission. (2012a). Completing the Scoreboard for the MIP: Financial Sector Indicators. Commission Staff Working Document, 14.11.2012, SWD (2012) 389 Final.
15. European Commission. (2012b). Scoreboard for the Surveillance of Macroeconomic Imbalances. Occasional Papers 92, European Economy February 2012.
16. Hallwirth, V. (2014). Überwachung der Wettbewerbsfähigkeit in der Europäischen Wirtschafts- und Währungsunion. In *Wirtschaftsdienst*, 94(11), pp. 798-805.
17. Kaminsky, G. – Lizondo, S. – Reinhart, C. (1998). *Leading Indicators of Currency Crisis.* IMF Staff Papers, Vol. 45 No 1, International Monetary Fund.

18. Kaminsky, G. – Reinhart, C. (1999). The Twin Crisis: The Cause of Banking and Balance-of-Payments Problems. In *The American Economic Review*, 89(3), pp. 473-500.

19. Kamps, C. et al. (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.

20. 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 2012.

21. Knedlik, T. (2014). The Impact of Preferences on Early Warning System – The Case of the European Commission's Scoreboard. In *European Journal of Political Economy*, 34 (), pp. 157-166.

22. Knedlik, T. (2015). The Usefulness of the Scoreboard of the Macroeconomic Imbalances Procedure in the European Union: Potentials for Reform. In *European Policy Analysis*, June, Issue 2015:14, pp. 1:12.

23. Moschella, M. (2014). Monitoring Macroeconomic Imbalances: Is EU Surveillance more Effective than IMF Surveillance? In *Journal of Common Market Studies*, 52(6), pp. 1273-1289.

24. National Bureau of Economic Research. (2008). The NBER's Recession Dating Procedure. Business Cycle Dating Committee. Available online at:

    http://www.nber.org/cycles/jan08bcdc_memo.html

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

Appendix I MIP Labour market indicators - official and optimal threshold characteristics

		Auxiliary Indicator	#	Probability of event	Official threshold			AUROC	Optimal threshold			
					Value	Implied Preferences	Max Absolute Utility		Value	Implied Preferences	Max Absolute Utility	Inverse aNtS
Youth UR						0.777	-0.0518	**0.3771**	**0.62**	0.833	-0.0214	0.723
Sensitivity Interval	*-10 percent*	**NO**	287	16.72%	**2.00**	0.816	-0.0337	**0.3771**	**0.56**	0.833	-0.0214	0.723
	+10 percent					0.777	-0.0508	**0.3794**	**0.68**	0.833	-0.0214	0.723
Long-term UR						0.729	-0.0462	**0.4419**	**0.41**	0.794	-0.0142	0.835
Sensitivity Interval	*-10 percent*	**NO**	155	20.65%	**0.50**	0.642	-0.0895	**0.4419**	**0.37**	0.794	-0.0142	0.835
	+10 percent					0.729	-0.0462	**0.4491**	**0.45**	0.794	-0.0142	0.835
UR						0.803	-0.0482	**0.3635**	**6.40**	0.813	-0.0012	0.992
Sensitivity Interval	*-10 percent*	**NO**	229	18.78%	**10.00**	0.809	-0.0393	**0.3631**	**5.80**	0.813	-0.0012	0.992
	+10 percent					0.803	-0.0459	**0.3635**	**7.10**	0.813	-0.0012	0.992
Activity rate						0.810	-0.0088	**0.5028**	**-0.20**	0.834	0.0033	1.053
Sensitivity Interval	*-10 percent*	**NO**	295	16.61%	**-0.20**	0.810	-0.0076	**0.5022**	**-0.18**	0.834	0.0033	1.053
	+10 percent					0.810	-0.0091	**0.5065**	**-0.22**	0.834	0.0033	1.053
Employment						NA	NA	**0.3383**	**-5.34**	0.834	-0.0038	0.447
Sensitivity Interval	*-10 percent*	**YES**	302	16.56%	**NA**	NA	NA	**0.3388**	**-5.94**	0.834	0.0003	1.117
	+10 percent					NA	NA	**0.3384**	**-5.88**	0.834	-0.0038	0.447
Participation Rate						NA	NA	**0.4391**	**58.05**	0.835	0.0012	1.017
Sensitivity Interval	*-10 percent*	**YES**	302	16.56%	**NA**	NA	NA	**0.4413**	**52.20**	0.835	0.0012	1.017
	+10 percent					NA	NA	**0.4404**	**71.10**	0.835	0.0017	1.013
YNEET						NA	NA	**0.3923**	**25.65**	0.788	0.0121	3.362
Sensitivity Interval	*-10 percent*	**YES**	204	19.12%	**NA**	NA	NA	**0.3927**	**23.10**	0.788	0.0121	3.362
	+10 percent					NA	NA	**0.3907**	**28.20**	0.788	0.0121	3.362

Appendix II MIP External imbalances indicators - official and optimal threshold characteristics

		Auxiliary Indicator		#	Probability of event	Official threshold			AUROC	Optimal threshold			
						Value	Implied Preferences	Max Absolute Utility		Value	Implied Preferences	Max Absolute Utility	Inverse aNtS
NULC							0.864	-0.0654	**0.3628**	**26.37**	0.811	-0.0165	0.337
Sensitivity Interval	*-10 percent*	**NO**		196	18.88%	**12.00**	0.837	-0.0449	**0.3602**	**24.39**	0.811	-0.0153	0.355
	+10 percent						0.864	-0.0672	**0.3601**	**20.79**	0.811	-0.0169	0.499
Terms of trade							NA	NA	**0.5198**	**-3.90**	0.768	0.0173	1.470
Sensitivity Interval	*-10 percent*	**YES**		112	23.21%	**NA**	NA	NA	**0.5203**	**-3.50**	0.750	0.0133	1.470
	+10 percent						NA	NA	**0.5193**	**-4.30**	0.750	0.0133	1.470
REER							0.821	0.0218	**0.6001**	**8.85**	0.820	0.0269	1.454
Sensitivity Interval	*-10 percent*		**upper**	161	18.01%	**11.00**	0.821	0.0226	**0.6009**	**7.95**	0.820	0.0269	1.454
	+10 percent	**NO**					0.821	0.0241	**0.5980**	**9.70**	0.820	0.0269	1.454
							0.846	-0.0212	**0.4717**	**-10.45**	0.819	0.0060	2.985
Sensitivity Interval	*-10 percent*		**lower**	161	18.01%	**-11.00**	0.850	-0.0249	**0.4716**	**-9.40**	0.800	0.0055	2.985
	+10 percent						0.850	-0.0261	**0.4716**	**-11.50**	0.800	0.0055	2.985
EMS							0.929	0.0019	**0.6972**	**-17.94**	0.837	0.0520	1.999
Sensitivity Interval	*-10 percent*	**NO**		251	16.33%	**-6.00**	0.950	-0.0050	**0.7089**	**-17.34**	0.850	0.0477	2.163
	+10 percent						0.950	-0.0053	**0.6907**	**-17.64**	0.850	0.0442	1.916
CA balance							NA	NA	**0.4889**	**0.96**	0.828	0.0009	1.481
Sensitivity Interval	*-10 percent*		**upper**	263	17.11%	**6.00**	NA	NA	**0.4888**	**0.84**	0.828	0.0009	1.481
	+10 percent	**NO**					NA	NA	**0.4889**	**1.08**	0.828	0.0009	1.481
							0.855	-0.0053	**0.5620**	**-7.44**	0.829	0.0209	1.646
Sensitivity Interval	*-10 percent*		**lower**	263	17.11%	**-4.00**	0.850	-0.0048	**0.5612**	**-8.68**	0.800	0.0148	2.105
	+10 percent						0.900	-0.0161	**0.5627**	**-10.64**	0.800	0.0148	2.105
CA and CAP balance							NA	NA	**0.6426**	**-6.84**	0.822	0.0377	2.267
Sensitivity Interval	*-10 percent*	**YES**		253	17.79%	**NA**	NA	NA	**0.6417**	**-6.16**	0.800	0.0322	2.267
	+10 percent						NA	NA	**0.6423**	**-7.52**	0.800	0.0322	2.267
Productivity							NA	NA	**0.4333**	**-5.76**	0.834	-0.0035	0.701
Sensitivity Interval	*-10 percent*	**YES**		302	16.56%	**NA**	NA	NA	**0.4334**	**-5.46**	0.834	-0.0028	0.744
	+10 percent						NA	NA	**0.4330**	**-6.00**	0.834	-0.0049	0.627

Appendix III MIP Indebtedness-related indicators - official and optimal threshold characteristics

		Auxiliary Indicator	#	Probability of event	Official threshold			AUROC	Optimal threshold			
					Value	Implied Preferences	Max Absolute Utility		Value	Implied Preferences	Max Absolute Utility	Inverse aNtS
FDI flows						NA	NA	**0.5840**	**5.60**	0.837	0.0193	1.516
Sensitivity Interval	*-10 percent*	**YES**	301	16.28%	**NA**	NA	NA	**0.5842**	**5.05**	0.837	0.0193	1.516
	+10 percent					NA	NA	**0.5834**	**6.15**	0.837	0.0186	1.487
FDI stocks						NA	NA	**0.4435**	**29.20**	0.809	0.0020	1.023
Sensitivity Interval	*-10 percent*	**YES**	204	19.12%	**NA**	NA	NA	**0.4438**	**26.40**	0.809	0.0020	1.023
	+10 percent					NA	NA	**0.4432**	**32.00**	0.809	0.0020	1.023
Net IIR						0.824	-0.0196	**0.4592**	**-70.70**	0.809	0.0067	1.260
Sensitivity Interval	*-10 percent*	**NO**	204	19.12%	**-35.00**	0.824	-0.0194	**0.4600**	**-63.70**	0.809	0.0067	1.260
	+10 percent					0.850	-0.0224	**0.4600**	**-77.70**	0.800	0.0052	1.260
Gross external debt						NA	NA	**0.5037**	**26.80**	0.888	0.0216	1.323
Sensitivity Interval	*-10 percent*	**YES**	98	11.22%	**NA**	NA	NA	**0.5043**	**24.00**	0.888	0.0216	1.323
	+10 percent					NA	NA	**0.5052**	**29.60**	0.888	0.0216	1.323
Public sector debt						0.000	-0.1528	**0.4660**	**19.80**	0.797	0.0181	1.183
Sensitivity Interval	*-10 percent*	**NO**	118	20.34%	**60.00**	0.000	-0.1091	**0.4675**	**17.40**	0.797	0.0181	1.183
	+10 percent					0.769	-0.0175	**0.4665**	**21.60**	0.797	0.0181	1.183
Private sector debt						NA	NA	**0.5359**	**33.25**	0.842	0.0136	1.198
Sensitivity Interval	*-10 percent*	**NO**	297	15.82%	**133.00**	NA	NA	**0.5363**	**29.26**	0.842	0.0125	1.177
	+10 percent					NA	NA	**0.5374**	**30.59**	0.842	0.0139	1.174
HPI						0.589	0.0521	**0.7524**	**5.82**	0.757	0.0849	6.545
Sensitivity Interval	*-10 percent*	**NO**	70	24.29%	**6.00**	0.589	0.0521	**0.7533**	**5.28**	0.757	0.0849	6.545
	+10 percent					0.589	0.0456	**0.7524**	**6.42**	0.757	0.0849	6.545

Appendix IV MIP Other indicators - official and optimal threshold characteristics

		Auxiliary Indicator	#	Probability of event	Official threshold			AUROC	Optimal threshold			
					Value	Implied Preferences	Max Absolute Utility		Value	Implied Preferences	Max Absolute Utility	Inverse aNtS
Poverty						NA	NA	**0.4405**	**9.80**	0.799	0.0023	1.017
Sensitivity Interval	*-10 percent*	**YES**	204	19.12%	**NA**	NA	NA	**0.4391**	**6.20**	0.799	0.0021	1.013
	+10 percent					NA	NA	**0.4408**	**10.80**	0.799	0.0023	1.017
GFCF						NA	NA	**0.3422**	**11.60**	0.840	0.0019	2.341
Sensitivity Interval	*-10 percent*	**YES**	308	15.91%	**NA**	NA	NA	**0.3420**	**10.40**	0.800	0.0014	2.341
	+10 percent					NA	NA	**0.3419**	**12.80**	0.800	0.0014	2.341
RaD						NA	NA	**0.4813**	**1.43**	0.818	0.0089	1.065
Sensitivity Interval	*-10 percent*	**YES**	230	18.26%	**NA**	NA	NA	**0.4813**	**1.29**	0.850	0.0066	1.065
	+10 percent					NA	NA	**0.4817**	**1.57**	0.850	0.0066	1.065

Appendix V MIP baseline and auxiliary indicators and their approximations

Indicator description	Calculation	Source	Approximated by
Baseline Indicators			
Current account balance as % of GDP	3 year average	WB	
Net international investment position	% of GDP	WB	
Real effective exchange rate (42 trading partners)	3 year % change	WB	Real effective exchange rate (2010=100)
Export market share (% of world export)	5 year % change	WB	
Nominal unit labour cost index (2010=100)	3 year % change	WB	
House price index (2010=100, deflated)	1 year % change	BIS	Real BIS selected property prices
Private sector credit flow - consolidated	% of GDP	not available	
Private sector debt - consolidated	% of GDP	WB	Domestic credit to private sector (% of GDP)
General government gross debt	% of GDP	WB	Central government gross debt (% of GDP)
Unemployment rate	3 year average	WB	
Total financial sector liabilities	1 year % change	not available	
Activity rate - % of total population aged 15-64	3 year change in p.p.	WB	
Long term unemployment rate - % of total population aged 15-74	3 year change in p.p.	WB	
Youth unemployment rate - % of total population aged 15-24	3 year change in p.p.	WB	
Auxiliary Indicators			
Gross fixed capital formation	1 year % change	WB	
Gross domestic expenditures on research and development	% of GDP	WB	
Current plus capital account (net lending-borrowing)	% of GDP	WB	
Net external debt	% of GDP	WB	External debt stocks (% of GDP)
Foreign direct investment in the reporting economy (flows)	% of GDP	WB	
Foreign direct investment in the reporting economy (stocks)	% of GDP	WB	
Net trade balance of energy products	% of GDP	WB	
Real effective exchange rate (Euro area trading partners)	3 year % change	not available	
Export performance against advanced economies	5 year % change	not available	
Terms of trade (goods and services)	5 year % change	WB	Net barter terms of trade index (2000=100)
Export market share, volumes	1 year % change	not available	
Labour productivity	1 year % change	WB	GDP per person employed (constant 2011 PPP $)
Nominal unit labour cost index (2010=100)	10 year % change	not available	
Unit labour cost performance relative to Euro Area	10 year % change	not available	
House price index (2010=100) nominal	3 year % change	not available	
Residential construction	% of GDP	not available	
Private sector debt - non-consolidated	% of GDP	not available	
Financial sector leverage - non-consolidated	% debt-to-equity	not available	

Employment rate	1 year % change	WB	Employment to population ratio, 15+, total (%)
Young people neither in employment nor in education and training - % of total population aged 15-24	%	WB	
People at risk of poverty or social exclusion - % of total population	%	WB	Poverty headcount ratio at national poverty lines (% of population)
People at risk of poverty after social transfers - % of total population	%	not available	
Severely materially deprived people - % of total population	%	not available	
People living in households with very low work intensity - % of total population aged 0-59	%	not available	