**Macroeconomic Imbalances as Indicators
for Debt Crises in Europe**

*Tobias Knedlik,
Gregor von Schweinitz*

August 2011   No. 12

Macroeconomic Imbalances as Indicators for Debt Crises in Europe

Tobias Knedlik,
Gregor von Schweinitz

August 2011   No. 12

Authors: *Tobias Knedlik*
             Department of Macroeconomics
             E-mail: tobias.knedlik@iwh-halle.de
             Phone: +49 (0) 345 7753-740

             Gregor von Schweinitz
             Department of Macroeconomics
             E-mail: gregorvon.schweinitz@iwh-halle.de
             Phone: +49 (0) 345 7753-744

Comments and suggestions on the methods and results presented are welcome.

IWH Discussion Papers are indexed in RePEc-Econpapers and in ECONIS.

Macroeconomic Imbalances as Indicators
for Debt Crises in Europe

Abstract

European authorities and scholars published proposals on which indicators of macroeconomic imbalances might be used to uncover risks for the sustainability of public debt in the European Union. We test the ability of four proposed sets of indicators to send early-warnings of debt crises using a signals approach for the study of indicators and the construction of composite indicators. We find that a broad composite indicator has the highest predictive power. This fact still holds true if equal weights are used for the construction of the composite indicator in order to reflect the uncertainty about the origin of future crises.

Keywords: macroeconomic surveillance, macroeconomic imbalances, economic governance, signals approach, European Union (EU), European Monetary Union (EMU)

JEL Classification: C14, E61, E62, F40

Makroökonomische Ungleichgewichte als Indikatoren für Staatsschuldkrisen in Europa

Zusammenfassung

Europäische Institutionen und Wissenschaftler haben Vorschläge für die Verwendung bestimmter Indikatoren makroökonomischer Ungleichgewichte zur Aufdeckung von Risiken für die Tragbarkeit der Staatsverschuldung in der Europäischen Union unterbreitet. Wir testen, ob die vier vorgeschlagenen Indikatoren-Sets in der Lage sind, Frühwarnsignale für Schuldenkrisen zu senden. Wir verwenden einen Signalansatz zur Analyse der einzelnen Indikatoren und zur Konstruktion von Verbundindikatoren. Es zeigt sich, dass ein breiter Verbundindikator die höchste Vorhersagekraft hat. Dieses Ergebnis wird auch bestätigt, wenn die Indikatoren gleichgewichtet verwendet werden, um die Unsicherheit bezüglich künftiger Krisenursachen abzubilden.

Schlagwörter: makroökonomische Überwachung, makroökonomische Ungleichgewichte, wirtschaftspolitische Koordinierung, Europäische Union (EU), Europäische Währungsunion

JEL-Klassifikation: C14, E61, E62, F40

Macroeconomic Imbalances as Indicators for Debt Crises in Europe

1   Introduction

The debt crises in some countries of the European Monetary Union (EMU) following the recent global financial crisis made the failure of existing political instruments for the prevention of excessive public debts evident. In particular the focus on the ratio of public deficits to GDP in the Stability and Growth Pact (SGP) was insufficient for the analysis of debt sustainability. For example, the burst of a real estate bubble and accompanying decline of real economic activity in Spain led to a slump of public income that could not be compensated by cuts in public expenditure on such a short notice. In Ireland, the drawing of public guarantees by faltering banks put high pressure on public finances. As in Spain, also in Ireland neither the level of existing public debt nor newly accruing deficits could warn of the massive risks for public solvency.

Because of these insufficiencies, political and economic institutions on national and European levels discuss longer-term fundamental developments and risks underlying the sustainability of public debt under the keywords "Macroeconomic surveillance" and "Economic Governance". Proposals for institutional reforms were for example published by the German Federal Ministry for Economy and Technology (BMWi) (2010), by the European Commission (EC) (2010a, 2010b, 2010c), by the European Central Bank (2010), and by a taskforce of the European Council (2010). In these documents an evaluation of macroeconomic imbalances using an indicator approach is proposed. The surveillance of a set of indicators will allow the evaluation of the macroeconomic situations in countries belonging to the EMU. A debate, founded on these proposals, evolved on the question, which indicators of macroeconomic imbalances are capable to send early and reliable warning signals.

In this paper we analyze for the first time whether the proposed indicators are able to signal risks for excessive public debt in the corresponding countries of the EMU. The question is answered using the early-warning capability of the proposed indicators in the framework of a signals approach. This approach was presented first in Kaminsky, Lizondo and Reinhart (1998) for currency crises and in Kaminsky and Reinhart (1999) and Kaminsky (1999) for twin-crises (joint occurrence of banking and currency crises). Reinhart (2002) uses the signals approach to link currency crises to sovereign debt crises in emerging markets. Sovereign debt crises in emerging markets have been the focus of many studies, as for example Manasse, Roubini and Schimmelpfennig (2003), Manasse and Roubini (2009), Pescatori and Sy (2004). However, to our knowledge the signals approach is applied for the first time to signal debt crises in developed markets in

this paper. Another novelty of this paper is that, for the first time, we apply this debt crisis forecasting technique to a case of a currency union, where only four out of eleven analyzed countries experienced a crisis. We find that a combination of all proposed indicators could have sent strong signals of upcoming public debt crises. Particularly appealing in our results is the stability of the high predictive power to several modifications.

The rest of the paper is organized as follows: In the following section we discuss the term "macroeconomic imbalances" and how the existence of these imbalances might lead to public debt crises. In the third section we present indicators for an early-warning system out of four different proposals. The method of the signals approach, especially the condensation of different indicators to a composite indicator, is given in section four. In section five, we present and discuss the obtained results, followed by a proposal for a new composite indicator in section 6. The two final sections then contain robustness analyses and the conclusions.

2   Public Debt Crises and Macroeconomic Imbalances

Before the introduction of the Euro as the common currency for the European Monetary Union it had been argued by some scholars that a monetary union should ideally be accompanied by some form of fiscal policy coordination (e.g. von Hagen and Hammond, 1998). In the absence of such a coordination, excessive debt in some countries of the monetary union might cause costs for other members via increased interest rates or might even lead to higher inflation and risks for the external stability of the common currency if the central bank is not willing or able to counteract the expansionary fiscal policies (Beetsma and Uhlig, 1999). Thus, the minimum form of coordination should be limits to public debt and deficits in the EMU. Accordingly, member states agreed on what is called the Stability and Growth Pact (SGP). It lined out that public debt should not exceed a sustainable level (60 percent of GDP) and, therefore, public spending and public revenues should be balanced, allowing for deficits of up to 3 percent of GDP. For countries not fulfilling the requirements, measures of punishment have been implemented.

Whilst the Stability and Growth Pact (SGP) has already been weakened by avoiding the punishment of Germany and France, who failed to stick to the rules in 2003, it has proven to fail to cope with the sudden changes in the financial sector, its real economy effects, and its fiscal effects in the aftermath of the global financial crisis. Countries that had been able to fulfill all the criteria of the SGP were not safe from facing turbulent and unsustainable public finances. It became evident that reacting to present fiscal deficits that exceeded the limits of the SGP might be too late to avoid excessive debt, refinancing difficulties, and default pressures.

Regarding fiscal policy coordination, two proposals have been suggested by European institutions to address the shortcomings of the current institutions (Council of the European Union, 2011). The first proposal is to tighten the SGP, including further sanctions and the involvement of European institutions in the budget planning process (the European Semester). The second proposal is to discuss potential risks for the sustainability of public debt. If the sudden emergence of large public deficits is not seen as an early indicator for fiscal distress that might cause sovereign default, but already as an expression of unsustainable public debt, it is important to focus on potential determinants of the sudden changes in public finances. Stylized facts about the countries with refinancing difficulties and the ones without, lead to the hypothesis that evolving imbalances in macroeconomic and financial parameters, e.g. property price bubbles and foreign trade deficits, are at the root of sudden changes in public finances. The proposal of monitoring, discussion, and regulation of these imbalances is currently debated under the term of macroeconomic surveillance.

Public institutions and scholars proposed sets of indicators to be observed to detect imbalances that might result in fiscal stress. Accordingly, if imbalances emerge, European institutions might request deeper analysis of the potential vulnerabilities of economies and – if countries fail to react – to impose punishment on the countries. So far, the literature does not provide an analysis of which indicators are useful early indicators for unsustainable public debt and eventually public debt crisis. This paper provides an analysis of the ability of different proposed indicators and sets of indicators to forecast public debt crises.

To analyze the explanatory power of potential determinants of changes in the sustainability of public debt it is necessary to define a dependent variable, i.e. what is meant by unsustainable public debt, fiscal stress, or public debt crises. From the above elaborations it is straight forward that a sovereign default is a too narrow definition, because even before a default, excessive debt might cause risks for the success of the common monetary policy. The same argument holds true for crises definitions relying on debt restructurings, IMF loans or large arrears (as, for example, in Reinhart and Rogoff (2010) or Manasse and Roubini (2009)). It is also evident, that the variable of fiscal deficits or debts in percent of GDP is too imprecise, because different countries might be able to cope with different levels of debt and different deficits. Pressure on the monetary union is caused by the potential inability of a member country to serve its debt, which would lead to bail outs or (partial) default. In line with the literature on sovereign debt crisis we choose the criterion of extraordinary high default risk premiums, which can be measured by the difference between the yield of a country's bond and a proxy for a safe investment. For example Pescatori and Sy (2004) in their analysis of debt crises in developing countries use the spreads of government bond yields against the yields of U.S. treasury bonds. A high spread of government bond yields is interpreted as a serious

doubt in the governments' capacity to service its debt in the future.[1] We refer to extraordinary high bond spreads as public debt crises in the rest of the paper.

Most of the European debt is denoted in Euro. Because we want to exclude exchange rate risks in our analysis, we do not choose the yields of U.S. treasury bonds as the reference in the spreads, but the average of the yields of AAA-rated countries of the EMU. Our crisis variable is thus

$$S_t^k = R_t^k - R_t^{AAA},$$

the yield R of a government bond of country k at time t with a maturity of ten years minus the average yield of government bonds of EMU-countries with AAA-rating with maturity of ten years at time t.[2] In this paper, we use monthly data on the yields of government bonds for all Euro-12 countries except Luxembourg (i.e.: Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, Portugal, Spain) since the introduction of the Euro (01.01.1999 or 01.01.2001 for Greece, respectively) until 01.04.2011. We calculate the mean μ_S and standard deviation σ_S of all spreads. As in Knedlik and Scheufele (2008), a crisis occurs in a country, when $S_t^k \geq \mu_S + 1.65\,\sigma_S \cong 166.4bp$.[3] Using tis definition we find four countries with debt crises in our sample: Greece, starting in January 2010, Ireland and Spain starting in June 2010, and Portugal starting in July 2010. Since both, countries and first manifestations of debt servicing difficulties are in line with common sense on crisis countries in the euro area, we use these crisis dates for testing the goodness of indicator sets to signal debt crises in the euro area.[4]

3   Proposed Indicators

We discuss four different sets of macroeconomic imbalances indicators that should warn against possible debt crises in a country up to 24 months in advance. The different sets were proposed by BMWi (2010), European Commission (EC, 2010b), European Central Bank (ECB, 2010) and Heise (2011).[5] While the indicators in Heise (2011) are

---

[1] One could argue that due to imperfect and irrational capital markets, difficulties of debt service are not fully reflected in government yields. For example, yields of European government bonds converged during the first years after the introduction of the Euro. A discrimination of the different countries of the EMU was only triggered by the advent of the financial crisis in 2007. On the other hand, the yields of government bonds might rise during speculative attacks. However, it is the emergence of the spreads that eventually indicates and leads to stress to public finance.

[2] The yields are provided by Datastream. At every point in time, they calculate a yield curve via interpolation of existing yield/maturity combinations by cubic splines. The yields of an average AAA-rated government bond of a country in the EMU are also provided by Datastream.

[3] For skewed data, the quantile of the normal distribution is not equivalent to the quantiles of the empirical distribution. Using this definition thus keeps the number of crisis periods endogenous.

[4] For similar results obtained with two other crisis definitions we refer to the robustness analysis in section 7.

[5] For the latter see also *Heise et al.* (2010).

analyzed and combined to a scorecard (the so-called Euro Monitor), the three other proposals do not contain any analysis on the performance of the proposed indicators. In fact, they are quite vague with respect to how they plan to establish a macroeconomic surveillance. The BMWi set contains five indicators: unit labor costs, current account balance, price level developments, unemployment rate, and private debt. The EC set contains seven indicators: government debt, current account balance, real effective exchange rates based on unit labor costs and GDP deflator, increases in real house prices, net foreign asset position, and ratio of private sector credit to GDP. The ECB set contains ten indicators, subdivided into main indicators (four) and qualitative control indicators (six). The former are: competitiveness indexes based on consumer prices, based on a GDP-deflator, and based on unit labor costs, as well as deviations from stability-oriented wage developments. The latter include: indebtedness of public sector, external imbalances, degree of convergence, asset prices, indebtedness of private sector, and credit booms. The Heise set contains 15 indicators: government debt, government deficit, government interest payments as share of government expenditure, required adjustments in primary balances due to demographic change, unit labor costs, current account balance, export shares in global trade, domestic demand, unemployment rate, employment ratio, labor productivity, consumption of energy, private debt-to-GDP ratios for households and non-financial corporations, and net international investment position.

To discuss the economic relevance of the indicators they can be grouped into five categories: fiscal indicators, competitiveness indicators, asset prices, labor market indicators, and private and foreign debt indicators. For the matching of our indicators with that of the four sets and a description of sources and transformations of the indicators compare the table in annex II.

The variables of the first group, fiscal indicators, are an obvious candidate to signal potential debt crises. They include measures of government debt, government deficits, and interest payments as share of government expenditure. It is proposed, that extraordinary high levels of debt, large deficits, and a relatively high share of interest payments might signal dangers for the sustainability of debt. All sets, except for the BMWi set, include fiscal indicators. One of the main reasons to develop the framework of macroeconomic surveillance was the assumption that it would not be sufficient to observe fiscal indicators only. Therefore other groups of indicators are proposed.

The second and largest group of indicators comprises competitiveness indicators. The idea behind choosing these indicators is, that in a currency union without the possibility of nominal devaluations the emergence of disadvantages in price competitiveness cannot be easily corrected and might therefore, by erosion of the tax base and ever-increasing spending on social security systems, lead to risks for debt crises. The group of indicators includes current account balances (in all sets) and trade shares (in Heise), whereas high deficits and declining shares would be interpreted as results of low or declining competitiveness. Heise includes longer term developments of domestic demand, arguing that relying on trade surpluses might not be a sufficient long term strategy,

since growth and the tax base also depend on local developments. The group also includes constructed measures of competitiveness, which are real effective exchange rates based on different prices, e.g. consumer prices, GDP-deflators, and unit labor costs. Decreasing competitiveness is expected to constitute a risk for debt sustainability. The constructed competitiveness measures are proposed by EC and ECB. Heise and BMWi in turn suggest looking directly at unit labor costs. Additionally, BMWi suggests considering consumer prices directly. Accordingly, high increases in unit labor costs or consumer prices are expected to reduce competitiveness.

The third group consists of asset prices as indicators of macroeconomic imbalances. The EC recommends looking at house prices, whilst the ECB more generally considers asset prices. Extraordinary increases in asset prices could be a valuable predictor of debt crises, because they might indicate the emergence of a bubble that might in turn, if bursting, not only cause other forms of financial crises, such as banking crises, but might also lead to debt crises, e.g. because of government bail outs of banks, the growth effects due to banking crises or reduced collateral. We are including increases of property and asset price indices.[6]

Labor market indicators build up the fourth group of indicators. They can be found in the sets of Heise and BMWi. Both of the sets use the unemployment rate as indicator, Heise includes additionally labor-force participation and labor productivity. Labor market indicators might picture the flexibility of an economy to react to shocks and their longer term growth prospects. Thus, extraordinary high increases in unemployment, decreases in labor-force participation and labor productivity might signal risks for government's budgets.

The final group of indicators is private and foreign debt. All sets include a private debt indicator, but Heise subdivides it into household debt and debt of non-financial institutions. We include all three. Extraordinary high levels of private debt might indicate a risk for public debt, because the government might be in charge to bail-out private debt in the case of crises, and because high levels of private debt might indicate a position in a financial cycle that is related to speculative investment (Minsky, 1972). Additionally to private debt, Heise and EC include net foreign assets. The indicator derives from accumulated current account deficits and shows the foreign indebtedness of a country. It is therefore not only an indicator of long-term competitiveness but also one of vulnerability to withdrawals and forced reversals in the balance of payments.

Some characteristics of the different sets include: The Heise set is the largest, but it does not cover asset prices and real exchange rates. The BMWi set is the smallest, but includes a relatively wide range of indicators, excluding only fiscal indicators. The EC set is small but covers all categories except for labor. The ECB set is similar to that of the

---

6   High decreases – a burst of a bubble – could also be a signal. But since countermeasures should be taken before the burst of a bubble, these signals are not very useful for policymakers.

EC, but has a stronger focus on competitiveness indices. Only two variables are included in all set: current account balances and private debt measures.

Whilst we were trying to include as many variables as possible from the proposed sets into our sample, we had to exclude a few due to unavailability of data at high enough frequency (at least quarterly) or because the description has been too broad. These include two variables from the set of Heise (required adjustment of primary balances due to demographic change and energy consumption) and three variables from the ECB set (deviations from stability oriented wage developments, degree of convergence, and credit booms).

All remaining twenty variables are present either in monthly or quarterly frequency.[7] We drew the data from Eurostat, ECB and Datastream (including original data from the IMF, OECD, MSCI, ECB, and CPB). Most of the variables are transformed, e.g. expressed in terms of GDP or year over year changes are employed. For the details of the variable transformations see appendix A-1.

4   Empirical Method

The framework used to translate values of the different indicators I of macroeconomic imbalances into an early-warning system for public debt crises is the signals approach. We use – with small adaptations – the methodology as presented in Kaminsky, Lizondo and Reinhart (1998) and Kaminsky and Reinhart (1999). We employ the signals approach because of its simplicity and its ability to allow for detecting the single signaling indicators even if a composite indicator is used. The idea behind the signals approach is that an indicator sends an early-warning signal in time t and country k when the indicator value I_t^k exceeds a certain threshold Q_I.[8] The signals of the different indicators are then combined and condensed to one composite indicator that can be translated into pseudo-probabilities for the occurrence of a crisis. There are four questions that have to be answered before employing the signals approach:

The first question is: How long is the early-warning horizon? In the literature, the horizon before a crisis, in which signals should appear, ranges from 12 to 24 month. The suitability of the chosen horizon depends on the scope of the early-warning system. Our warning is aimed at the early identification of possible causes for difficulties with the servicing of government debt. Since we are concerned with longer-term developments and have to allow for policy reactions, we chose a comparably long early-warning horizon of 24 months, as for example in Kaminsky, Lizondo and Reinhart (1998).

---

[7]   Monthly data are obtained from quarterly using cubic interpolation. Where necessary, we calculated seasonally adjusted data using the berlin procedure.

[8]   The indicator I can be viewed as the multi-country time series, I_t^k is the realization in time t and country k.

The second question is: Which data could be used for the calibration of the system? As in Detragiache, E. and Spilimbergo, A. (2001), we exclude periods up to four years after the outbreak of a crisis from our analysis. The implicit assumption is that being in a crisis leads to a different behavior of the indicators compared to normal times or tranquil periods. For countries that did not experience a crisis in the observation period, we exclude the last 24 month of our sample (01.05.2009 – 01.04.2011) from the calibration because we do not know if the now crisis-free countries will perhaps experience a crisis in the near future or not (and thus, if the indicators should signal now or not).

The third question is: How is the threshold Q_I determined? The indicator sends a signal of a crisis within the next 24 months if its value exceeds a certain threshold. As we build an early-warning system for one currency union, we use uniform thresholds for all countries for a given indicator (for example current account over GDP). We express the threshold as a quantile of the indicators empirical distribution (using only the time range defined above and all eleven countries).

The determination of the threshold is crucial for the predictive power of the indicator. The prediction of an indicator is correct, if it is above the threshold before a crisis or if being below is not followed by a crisis in the next 24 months. Both, being above and below the threshold, can be seen as a signal. However, we refer to a signal if an indicator value is above its threshold. As in Kaminsky and Reinhart (1999), we have four possible outcomes (states) of the indicator for every point in time as defined in table 1.[9]

Table 1:

Indicator State Definition. Different combinations of signals and the occurrence of a crisis. We abstract for convenience from sub- and superscripts for indicator I, time t and country k ($\mathcal{S}$ instead of $\mathcal{S}_{I,t}^{k}$, $\mathcal{C}$ instead of $\mathcal{C}_{t}^{k}$, A instead of A_I).

	Crisis within the next 24 months $(\mathcal{C} = 1)$	No crisis within the next 24 months $(\mathcal{C} = 0)$
Signal issued $(\mathcal{S} = 1)$	A	B
No signal issued $(\mathcal{S} = 0)$	C	D

A is thus the number of months where a signal for a crisis was followed by a crisis in the next 24 months. B corresponds to the number of periods where a signal was not followed by a crisis period, while C contains all periods without signal despite a following crisis. D, last but not least, is the number of months where the absence of a signal is followed by 24 tranquil periods. States with correct signals are thus A and D, while those with erroneous (type II error) or missing (type I error) signals are B and C, respectively.

---

9 More technically, the table contains the number of periods that the indicator is in one of the four states depending on the level of the indicator and the occurrence of a crisis.

The share of type I errors to pre-crisis periods is $\frac{C}{A+C}$, the share of type II errors to tranquil periods is $\frac{B}{B+D}$. It is clear, that a low threshold will result in many signals (correct and wrong) and thus to possibly a high type II error probability, while a high threshold leads to few signals, but potentially missed crisis periods and thus a high type I error probability. In Kaminsky, Lizondo and Reinhart (1998) the optimal threshold is determined such that the Noise-to-Signal Ratio (NSR), the percentage of wrong signals over the percentage of correctly signalled crisis, is minimal:

$$NSR = \frac{\frac{B}{B+D}}{\frac{A}{A+C}} = \frac{\frac{B}{B+D}}{1 - \frac{C}{A+C}}.$$

An indicator is said to be acceptable, if $NSR < 1$. It can be seen, that this is equivalent to

$$\frac{C}{A+C} + \frac{B}{B+D} < 1.$$

This is a setup where type I and type II errors are equally weighted. But the question if a missed crisis or a wrongly signaled one is worse, depends highly on the preferences of the political decision makers that found their actions on an early warning system. For a discussion of this topic, see for example Bussière and Fratzscher (2008). Instead of using the NSR in order to determine the optimal threshold, we use the same utility function as in Alessi and Detken (2011). The utility function for every indicator I contains the weighted loss of type I and type II error probabilities:

$$U_I(\theta) = \min(\theta, 1-\theta) - \left(\theta \frac{C}{A+C} + (1-\theta)\frac{B}{B+D}\right).$$

As Alessi and Detken (2011) point out, an indicator is useful only if $U_I(\theta) > 0$.[10]

Because a crisis without a warning (type I error) will normally be more costly than preemptive action that is not followed by a crisis (type II error), one could argue, that θ should be bigger than 0.5. On the other hand, a decisionmaker may find type II errors more preferable because of the problematic justification of costly preemptive actions without a following crisis (a control mechanism is often only established after a crisis). Because of these opposing tendencies, we (as in Alessi and Detken (2011)) will work with a balanced risk aversion ($\theta = 0.5$) in this paper. For every indicator I, the optimal threshold is chosen as the quantile[11] which maximizes $U_I := U_I(0.5)$.

---

[10] $U_I(\theta) = 0$ is always achievable by regarding either every period as a signal (if type I errors are worse) or no period as a signal (if type II errors are worse).

[11] Between the 50% and 95% quantile of the empirical distribution of the indicator, steps in 1%.

The fourth question to be answered before employing the signals approach is: How can a composite indicator be constructed? For the assessment of the performance of the indicator set as a whole, we condense the information contained in every single indicator into a composite indicator. The indicator I sends a signal, if its value is larger than the threshold ($I_t^k > Q_I$). We introduce a binary variable ς for every indicator I, which is 1 if a signal is issued and 0 otherwise:

$$\varsigma_{I,t}^k = 1_{I_t^k > Q_I} = \begin{cases} 1, & if\ I_t^k > Q_I \\ 0, & otherwise \end{cases}$$

Different methods for the construction of a composite indicator have been proposed, for example in Kaminsky (1999) or Edison (2003). We will only use one of the different proposed methods and define the composite indicator CI as the weighted average of all binary indicators. As weights, we use the (normed) utilities U_I of the different indicators I (setting the usefulness of an indicator with negative utility to 0)[12]:

$$w_I = \frac{U_I}{\sum_{I\ Indicator} U_I}$$

$$CI_t^k = \sum_{I\ Indicator} w_I \varsigma_{I,t}^k.$$

In ECB (2010), we find the proposal that the early-warning system should work like a traffic light in three stages. This can be achieved using the composite indicator. Below a certain threshold, the composite indicator does not send a signal (green light). Above this threshold, a warning is issued (yellow light). The third level is reached, when a further thorough and detailed analysis reveals that there are serious macroeconomic imbalances that should be addressed in order to avoid a public debt crisis in the near future. Our analysis reported here is restricted to the differentiation between the green and the yellow phase.

From the above elaborations it is straightforward that the utility of the single indicators is an appropriate measure of the quality of an indicator for signaling proposes. In addition we also consider the empirical probability of type I $\left(\frac{C}{A+C}\right)$ and type II $\left(\frac{B}{B+D}\right)$ errors as components of the utility function.

However, not only these measures should be discussed, since the probability of type I and type II errors has a condition on the (non-)existence of a crisis. This can be seen only ex post. On the other hand, we might be interested in the ex ante probability of a crisis conditional on having a signal. Because of this, we will discuss in addition to the utility of the indicators the following two measures:

---

12   Results obtained with equal and inverse NSR weights are presented in section 7.

$\frac{A+D}{A+B+C+D}$ is the probability of a correctly signaled period (measured as the proportion of correct signals / no signals over all periods). $P(\mathcal{C} = 1|\mathcal{S} = 1) = P(\mathcal{C} = 1|I > \mathcal{Q}_I)$ (or, in short: $P(\mathcal{C}|\mathcal{S})$) is the probability of having a crisis now or within the next 24 months given the value of the indicator exceeds the threshold. For a good indicator, this measure should be monotonic increasing with $\mathcal{Q}_I$, because higher indicator values should be more likely to be followed by a crisis.

Still, the two measures of ex ante probability shown above might be somehow imprecise, because policy makers want to know, what an actually observed composite indicator value means. There will probably be a huge difference between a value of 0.4 and a value of 1 for the composite indicator (the latter meaning that all components of the indicator sent a signal of existing imbalances). Because of this, we also take a look at $P(\mathcal{C}|CI_t^k) := P(\mathcal{C}|lb \le CI_t^k < ub)$, the probability of having a crisis within the next 24 month given the value of the composite indicator is in a predefined interval, with $lb \in \{0, 0.1, \dots, 1\}$ and $ub = lb + 0.1$ (see Edison (2003), Kaminsky (1999), Brüggemann and Linne (2002)).[13] Like $P(\mathcal{C}|\mathcal{S})$ this measure should monotonically increase with lb.[14]

As a further quality measure for composite indicators, using the estimated probabilities and the actual realizations of a crisis, we can calculate the quadratic probability score (QPS) as described in Brüggemann and Linne (2002). The measure is essentially equivalent to the mean squared error. It sums the quadratic differences between the probability of a crisis within the next 24 months given the composite indicator value, and the actual outcome. Let T_k be the total number of periods that can be used for the calibration and quality assessment in country k. Then we have in analogy to Brier (1950) and Diebold and Rudebusch (1989) the normalized sum over the squared differences between the probability of a crisis given a certain composite indicator value and the actual existence of a crisis at time t in country k.

$$QPS = \frac{1}{\sum_k T_k} \sum_k \sum_{t=1}^{T_k} 2\big(P(\mathcal{C}|CI_t^k) - \mathcal{C}_t^k\big)^2.$$

A lower QPS means a lower mismatch between probabilities and actual outcomes.

With the above description of the signal approach and the introduction of different measures of quality for single indicators as well as for composite indicators we are able to present results for the four proposed indicator sets.

---

[13] This probability can also be calculated for individual indicators, given proper intervals. Since only the composite indicators are constructed in a way that ensures values between 0 and 1, we do not apply this measure to individual indicators.

[14] Just as $P(\mathcal{C}|\mathcal{S}) = P(\mathcal{C}|CI > x)$ can be interpreted as a complementary cumulative distribution function, $P(\mathcal{C}|lb \le CI_t < ub)$ is like a discrete density missing a scaling factor. The "interval" for $lb = 1$ is, of course, only a point value.

5   Results

We are presenting the results of our analysis in two parts. We first lay out the results of quality tests of single indicators, which is followed by the same exercise for the proposed composite indicators. In the next section we are then considering our own composite indicator.

The different indicator sets contain a total of 25 indicators, of which we use a total of 20. For each one, we calculate an optimal threshold, which results in the partition of the indicator in the four different states described above. The resulting utilities are then used to construct a composite indicator for every of the four indicator sets. We are now comparing the individual indicators. The analysis relies on the above described performance measures: utility, type I and type II error, probability of a correct crisis forecast, probability of correct forecast of crisis and non-crisis periods. . The results are summarized in table 2.

We can see, that higher optimal thresholds tend to produce more missed signals or a higher probability of type I errors, while lower thresholds go along with a higher share of type II errors. We see also, that indicators with a higher utility tend to be better at the other quality measures. A counterexample is domestic demand, that is an average indicator regarding $P(\mathcal{C}|\mathcal{S})$, while having a low utility. This can be explained by the fact, that domestic demand has a very high share of missed crisis, while sending nearly no false signals. The same holds true (to a lesser extent) for unit labor cost and foreign assets. At last, we see that the share of correctly signaled periods as share of all periods is mostly below the probability of not having a crisis (94%).[15]

We also see that some indicators perform very well, whilst others do not even yield a positive utility, and hence have no worth in forecasting a crisis in our sample. We are discussing the performance of the indicators by looking at the five groups defined in section 3.

The first group comprises the fiscal indicators, including government deficit, debt and interest payments. Government deficit (as share of GDP) is the best of the three indicators for coming crises and also the best indicator of the whole set of indicators considered. Government debt (as share of GDP) is lagging behind, as it is the accumulation of deficits. In both cases, signals are either given over most of the observation period (Greece for government debt and deficit; Italy and Belgium for government debt) or they start after the financial crisis. It is remarkable, that the calculated optimal thresholds are not too far from the limits given by the SGP (71% as a debt and 5.09% as deficit criterion). The interest payments as share of total government expenditure do not

---

15   For utility, the probability of type I and II errors have equal weight. A wrong signal (B) has a much lower share in the utility than a missed signal (C) due to sheer number of non-crisis periods. For the proportion of correctly signaled periods, the absolute number of wrong signals is relevant.

send any valid signals, because interest rates were declining (and converging) during most of the time since the introduction of the Euro leading to ever lower interest rates for newly issued debt and, until beginning 2008, declining shares of interest payments in almost all countries.

Table 2:
Different quality measures, sorted by utility.

Indicator	Utility	Optimal threshold	Type I error $\left(\frac{C}{A+C}\right)$	Type II error $\left(\frac{B}{B+D}\right)$	Prob. of correct crisis forecast $(P(c\mid s))$	Prob. of correct crisis/non-crisis period forecast $\left(\frac{A+D}{A+B+C+D}\right)$
Government deficit	0.4315	87%	0.0700	0.0669	0.5225	0.9328
Unemployment rate	0.3771	93%	0.2200	0.0258	0.6610	0.9625
Non-MFI debt	0.3367	88%	0.2500	0.0766	0.4032	0.9123
Household debt	0.3318	84%	0.2500	0.0864	0.5208	0.8954
Current account	0.3241	68%	0.0707	0.2812	0.1752	0.7315
Private debt	0.3153	84%	0.2500	0.1193	0.3024	0.8723
Unit labor cost	0.2506	86%	0.3900	0.1088	0.2711	0.8738
Foreign asset	0.1926	78%	0.4100	0.2048	0.1838	0.7803
Trade share	0.1861	85%	0.5000	0.1279	0.2008	0.8496
Labor-force participation	0.1858	64%	0.3000	0.3284	0.1639	0.6740
Labor productivity	0.1620	50%	0.2000	0.4759	0.1189	0.5446
Domestic demand	0.1506	90%	0.6200	0.0788	0.2714	0.8824
Government debt	0.0865	71%	0.5500	0.2769	0.1187	0.7021
ULC-competitiveness	0.0471	89%	0.8000	0.1058	0.0909	0.8593
Inflation	0.0319	91%	0.8500	0.0861	0.1007	0.8678
HICP-competitiveness	-0.0163	95%				
GDP-deflator-compet.	-0.0263	95%				
Interest payments	-0.0269	95%				
Asset prices	-0.0267	95%				
Property prices	-0.0273	95%				

Source: own calculations. Note: For the indicators with negative utility, the other measures are not shown, since there is no optimal threshold.

The second group of indicators comprises competitiveness indicators including unit labor costs, current account, trade share, domestic demand, HIPC-, GDP deflator-, and ULC-competitiveness, as well as inflation. The unit labor costs indicator seems, on first sight, to be fairly useful. But, contrary to other indicators with comparable utility (private debt and foreign assets), the wrong signals are not concentrated in the crisis countries directly before our signaling period, but are rather equally distributed over all countries and times. Additionally, in comparison there are a lot of missed signals in pre-crisis periods. This arises from our choice of yearly changes of the unit labor cost index. A high increase in one period followed by a stable development or slow decrease in subsequent periods might not lead to problems, while the same high increase, followed by further increases, might. The current account indicator, on the other hand, is a very good indicator, that sends wrong signals (type II errors) mostly in the crisis countries, while countries that did not experience a crisis have positive or only slightly negative balances over the whole observation period. Overall, only Portugal, Greece and Spain have reported current account deficits for the whole observed period. Ireland, on the other hand, had a positive current account balance until mid 2004. Then it turns negative. Other countries with current account surpluses over the whole period are Germany, Austria, Finland, the Netherlands and Belgium (with a current account deficit during the financial crisis). For France and Italy, we observe slight surpluses in 1999 and a steady slow decline to small deficits in 2010. The current account indicator shows a higher utility but a lower probability of forecasting correct crisis and non-crisis periods as compared to the trade share indicator. The change in the share of world trade (exports) is a rather mean indicator in terms of utility. Basically, the problems with the trade shares are the same as with unit labor costs. Signals are sent nearly equally distributed over time and all countries. There are two main periods where all countries lost significant shares of world trade: after the introduction of the Euro until 2002 and during the financial crisis, where trade as a whole declined, but more so in the developed countries than in the rest of the world. This indicator is more dependent on the general competitiveness of the Euro area in the world and not on the country-specific imbalances compared to other countries in the Euro area. Another indicator of the group is domestic demand. The countries within the EMU having the strongest increase in domestic demand over most of the time have been Ireland, Spain and Greece (in that order). However, starting with the financial crisis, the development slowed down significantly, leading to the three lowest growth rates (declines, especially for Ireland) end of 2010. Because of this divided picture, neither low increases of domestic demand (signaling weak economic growth or even stagnation) nor high increases (signaling possible funding through high current account deficits) can be used as an indicator yielding results comparable to the best indicators. However, the utility of the domestic demand indicator is still positive. This does not hold true for all of the competitiveness indexes based on different price indices. The indicators are calculated from effective exchange rates that are deflated by either consumer prices (HICP), the GDP deflator or unit labor costs (ULC) and are pro-

vided by the ECB. We analyze the yearly development of the indices.[16] The results are not encouraging. First, the development of the indicators seems to depend strongly on the development of the Euro. This is especially true for the ones that are deflated by HICP and GDP. The nominal depreciation of the Euro after 1999, and afterwards an appreciation can be observed in the data. This dependence leads to a strong comovement of the indicator development across countries. The strongest increases of the three indicators are observed after the burst of the new economy bubble. Before the current fiscal crisis, only few signals were sent by the indicator deflated by ULC, no signals were sent by the two other indicators, leading to a negative utility for the indicators that are deflated by HICP and GDP. The final indicator in the group of competitiveness indicators is inflation. Most signals of high inflation were sent in and after 2001, when it spiked due a variety of price shocks, see Gregoriou, Kontonikas and Montagnoli (2011), and only some (mostly in Greece and Spain) during the financial crisis. Because of this, inflation doesn't seem to be a good indicator for the current public debt crises. At least, it has a slightly positive utility because of the few correct signals of high inflation in Greece before the outbreak of the current public debt crisis.

The third group of indicators are asset prices. We look at high asset and property price increases as they are signs of developing bubbles. These indicators are particularly unable to predict the current crises within a time horizon of 24 months because of two facts. Firstly, the highest increases were observed before and during the new economy bubble (as well as the highest declines of asset prices afterwards). Furthermore, the second period of high increases in asset and property prices ended with the financial crisis in the beginning of 2008 (residential property prices even earlier). This means, that it took more than two years from the burst of a bubble through different transmission mechanisms to a public debt crisis. Another problem might be that asset prices have a high colinearity because of the interdependency of financial markets. This means that it might be difficult to identify a problem for a single member country based on extremely high asset price increases.

The third group of indicators consists of three measures of the labor market, namely the unemployment rate, labor force participation, and labor productivity. We looked at the absolute change of the unemployment rate from one year's month to the next year's month. A strong increase in the unemployment rate leads to problems for public finances. This effect could especially be observed in Ireland and Spain, where declines in construction investment lead to recessions (Lane (2011), Suarez (2010)) and an increase in the unemployment rate from 4.5% (June 2007) to 13.6% (June 2010) in Ireland[17] and

---

16   As they are normalized in 1999, absolute values do not capture the initial differences of the countries at the introduction of the Euro. Because values are (qua definition) all close to the normalization value in the time around 1999, the detection of a crisis around 1999 would theoretically be possible, but very difficult in practice.

17   5.9% in June 2008 at the beginning of the warning period.

8.0% to 20.2% in Spain[18] for the same time period. Since June 2010, the unemployment rates in both countries have been fairly stable. The unemployment rate is the second best performing single early indicator for debt crisis in our sample. The labor-force participation rate and labor productivity are not performing that well but still yield positive utility. These indicators (yearly percentage change of labor-force participation rate and average change of labor productivity in the course of the last 5 years) have the two lowest thresholds.[19] This means, that a lot of signals will be sent – including a lot of false ones. Both indicators send these false signals over the whole time period and in all countries. Additionally, one cannot even deduce the more crisis-prone countries from the others by the level of the indicator, if a signal is sent. Belgium and Italy, for example, have both to fight relatively high losses in the labor-force participation rate. Italy has the strongest decline in labor productivity at most times. This does not necessarily mean that the indicator is useless, since Italy is frequently included in sets of crisis countries (PIIGS instead of PIGS), and Belgium could probably be one of the candidates, if the public debt crisis could not be contained and would spread over the rest of the EMU.[20]

The final group of indicators comprises private and foreign debt indicators, namely private debt, household and non-MFI debt, as well as foreign assets. For almost all countries in the EMU, the share of private debt to GDP has risen since the introduction of the Euro. This increase was particularly strong in Ireland, Spain, Portugal and the Netherlands, while Greek private debt only increased strongly after the emergence of the crisis (and the corresponding drop in output). Already in 2004, the levels of private debt became somewhat alarming, especially in Ireland and the Netherlands, meaning that warnings of possible problems generated by the transmission from private to public debt are issued much earlier in some countries. All three private debt indicators perform well. The performance of the foreign asset indicator on the contrary is rather mean. This indicator (as share of GDP) is the accumulation of the current account balances and shows the divergence in the EMU caused by consistently different current accounts. While the German net foreign asset position increased from 10% of GDP in 1999 to 70% in 2010, Italy stays close to 0% (between -1% and 5%). The strongest development has Portugal, where foreign assets valuing 10% of GDP in 1999 developed to foreign debt of almost 30% of GDP in the beginning of 2007. Most countries (Portugal is an exception) having current account surpluses, and thus accumulating foreign assets already started with a higher net foreign asset position in 1999. Because of that, the indicator does not offer much new insight compared to the current account balance.

As can already be noted from the elaborations above, there are four different causes for false signals in our data. The burst of the new-economy bubble lead to a period of ma-

---

18   11.0% in June 2008 at the beginning of the warning period.

19   Expressed as quantiles of the empirical distribution.

20   One of the robustness tests presented in section 6 includes different crises definitions, where we treat Italy as a crisis country.

croeconomic imbalances in some areas of the economy. The corresponding indicators (for example asset prices, trade shares) normally have two blocks of periods where most signals are sent. Another reason for false signals may be that some macroeconomic imbalances may need more than two years (and another external shock) to materialize into problems with public debt, if they do not become too large. A good example for this type of false signal is given by government deficits and current account balances. They could thus be used as "very early warning indicators", while other indicators, like inflation, sent their signals only a very short time directly before and during a crisis.[21] The third cause for false signals seems to be more or less random, affecting all countries at all times. Imbalances occur only for one or two periods and then disappear again – meaning, that they are much harder to explain than other causes of false signals. The last cause of wrong signals concerns the competitiveness indicators whose development shows a lot of comovement that depends on the exchange rate of the Euro. The signals of these indicators rather reflect the strength of the Euro than imbalances inside the EMU.

In ECB (2010), a strong focus is put on the point that ideally one indicator should be enough for the surveillance of macroeconomic imbalances. The goal of a single indicator can of course be achieved by the usage of a composite indicator, which can be created as described in section 4 for every one of the different indicator sets. The resulting four composite indicators differ because of the differences of the four sets as described in section 3. The results are presented below.

From the analysis of individual indicators it is intuitive that the success of a composite indicator depends on the number and diversity of individual indicators used in its construction. The same becomes even more evident when looking at quality measures of the composite indicators. As every other indicator, the composite indicator can be evaluated in terms of signals for an upcoming crisis or tranquil periods. As possible thresholds, we choose values in {0,0.1,0.2,...,0.9,1}, which is equivalent to the scoreboard as presented in Heise (2011). As the optimal threshold for all composite indicators, we choose $Q_{CI} = 0.4$[22]. We derive this value by looking at the utilities at different thresholds as shown in Figure 1.

---

[21] It can be argued, that the term "false" signal is misleading at that point since the signals are basically correct – they are just sent too early. This might give rise to further analysis of an early warning system with varying early-warning horizons for different indicators.

[22] This threshold is optimal for the Heise indicator and nearly so for the BMWi indicator (as the two fairly balanced and stable sets). Thus, it is an approximation of the real optimal value. This would theoretically be 0.5, if type I and II errors would be independently distributed over time. Since we expect especially for the pre-crisis periods increasing values of the composite indicator (and thus time-dependency), the optimal threshold lowers. The choice of 0.4 is founded on the results of the larger indicator sets having a more stable development of the composite indicator.

Figure 1:
Utility of the composite indicators of the different sets at different thresholds

Utility of the composite indicators of the different sets at different thresholds
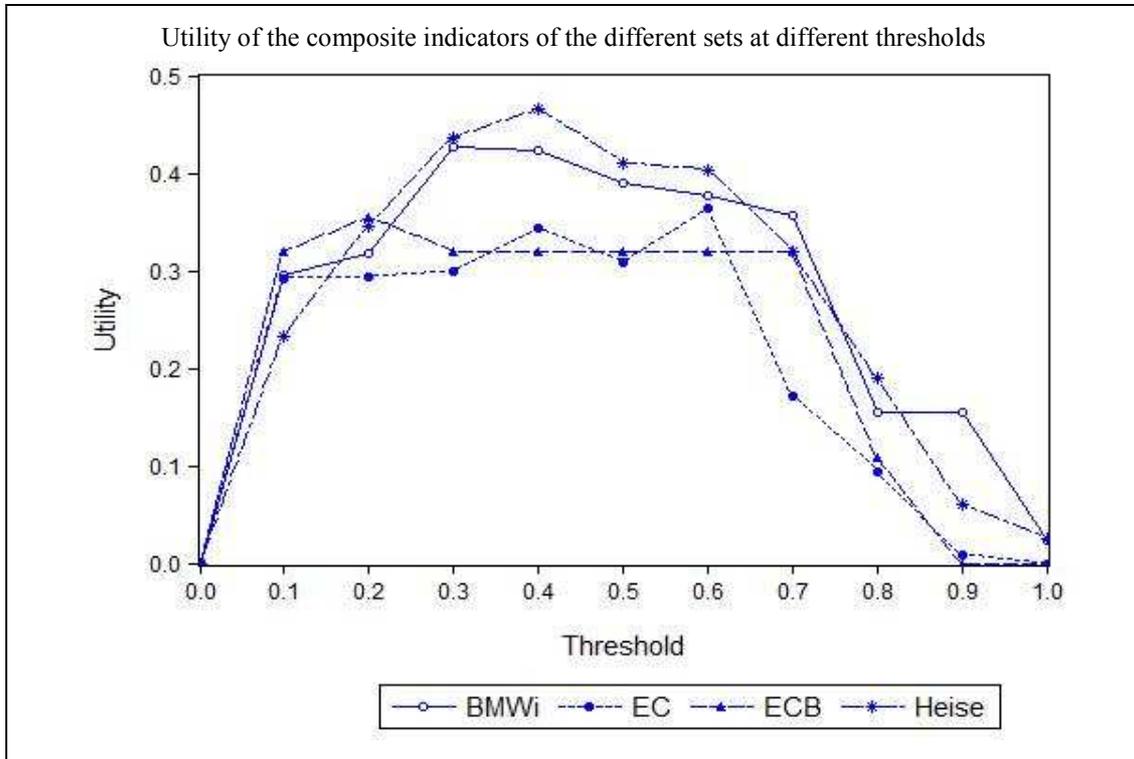
In Figure 2, we show the probability of correct crisis forecasts ($P(\mathcal{C}|\mathcal{S})$) for different thresholds. We can see the expected monotonic development of $P(\mathcal{C}|\mathcal{S})$ very good for the Heise indicator, but also (partly) for the BMWi indicator.[23] These two are also the only indicators reaching probability values above 90%, while the ECB indicator approaches the probability $P(\mathcal{C}|\mathcal{S})$ of private debt (30%).

Figure 2:
Probability of correct crisis forecasts ($P(\mathcal{C}|\mathcal{S})$) of the composite indicators of the different sets at different thresholds (a signal being issued for values of the composite indicator above the threshold)


---
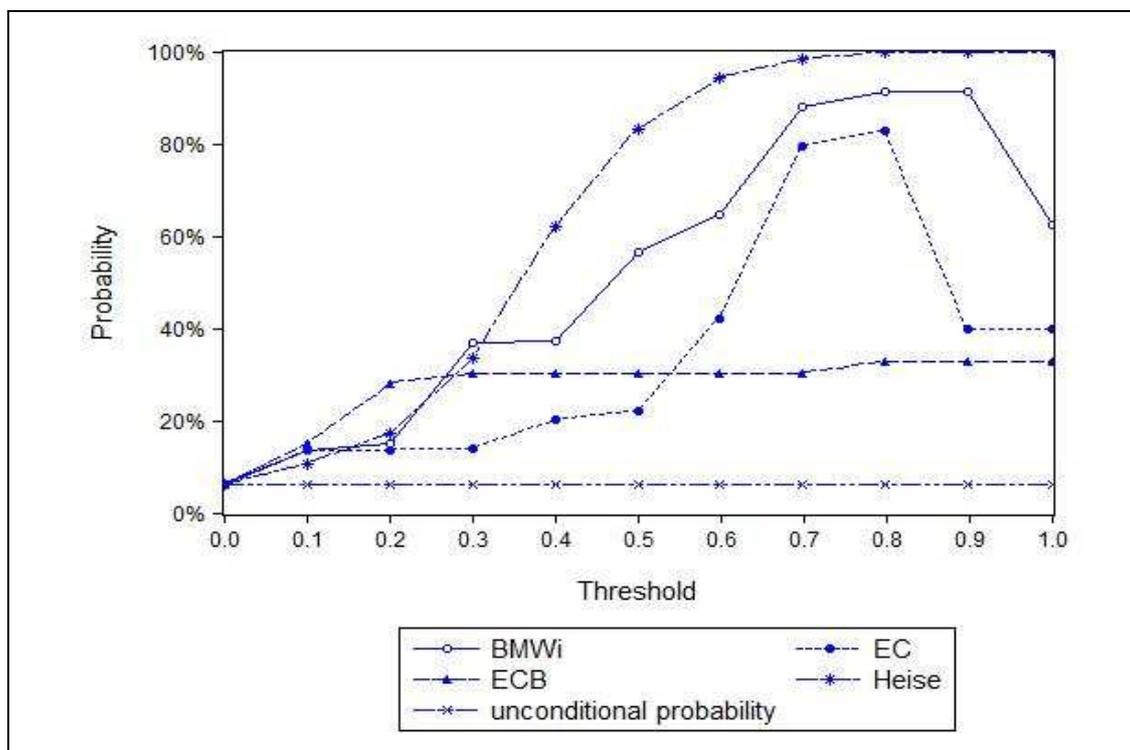
[23] Due to the small number of indicators for all sets but the Heise set, the indicator does not show the theoretically expected monotonic increase.

Figure 3 shows the estimated probabilities of a crisis given the values of the composite indicator are in a certain interval.[24] We can see that for the Heise and BMWi sets indicator values of 0.4 or above correspond with a probability for a debt crisis within the next 24 months above the unconditional probability. For values of 0.6 or higher this becomes true for the EC indicator as well.

Figure 3:
Probability of a crisis, given the values of the composite indicator in a certain interval ($P(\mathcal{C}|CI_t^k)$) of the composite indicators of the different sets
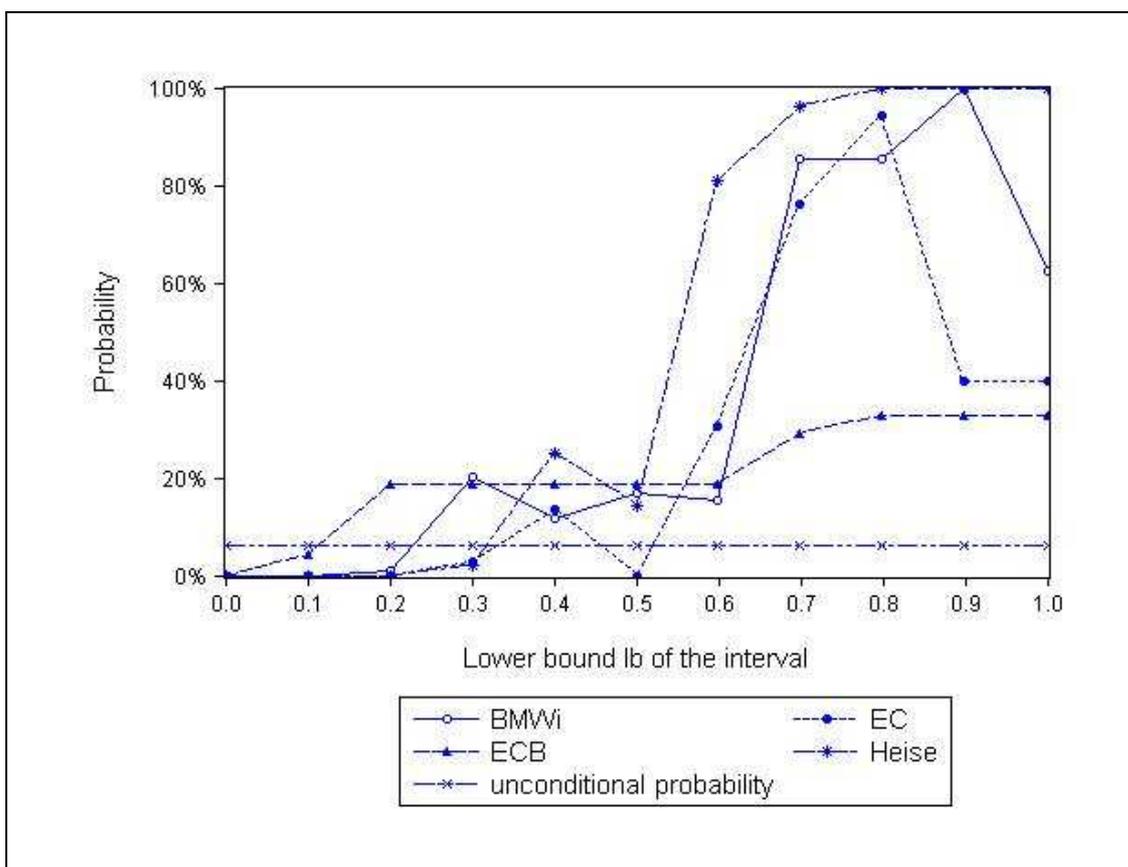
From the results presented in figure 3, we can calculate the quadratic probability score (QPS) as a summary measure of the quality of the composite indicator given indicator intervals. The quadratic probability score (where a lower value is more favorable) is presented in table 3 together with utility, probability of correct crisis forecasts and probability of a correctly signaled crisis and non-crisis periods.

---

24   The calculated probabilities are independent of country k and time t, they only depend on the actual value of the composite indicator.

Table 3:
Quality measures with $Q_{CI} = 0.4$ for composite indicators of the four different sets, calculated using utility weights.

Composite indicator	BMWi	ECB	EC	Heise
Utility	0.4233	0.3194	0.3440	0.4660
Probability of correct crisis forecast $\left(P(\mathcal{C}|\mathcal{S})\right)$	37.11%	30.24%	20.31%	62.18%
Probability of correct crisis/non-crisis period forecast $\left(\frac{A+D}{A+B+C+D}\right)$	89.98%	88.04%	77.72%	96.26%
Quadratic probability score ($\boldsymbol{QPS}$)	0.0393	0.0903	0.0668	0.0234

Table 3 shows that the broad indicator from Heise (2011) is extremely good at all different quality measures. Especially the measure $P(\mathcal{C}|\mathcal{S})$ will be of interest for the policy-maker: in our sample and given the composite indicator exceeds the value of 0.4, a crisis within the next 24 months occurred with a probability of only 20% in the EC set, but with a probability of 62% for the Heise composite indicator.

A closer look at the composition and performance of the different composite indicators provides the following picture. The composite indicator of BMWi is fairly balanced among the different groups of indicators. We can observe low, but not critical values for the composite indicator in many countries. These values only rise before and during the financial crisis. However, one main feature stands out: There is a long history of higher values (0.4 – 0.7) for Greece, Portugal and Ireland. The composite indicator of ECB has been constructed from all three competitiveness indicators, public and private debt and asset prices. Three of these indicators have a weight of 0 due to negative utility. This leaves private debt with a share of 70% in the composite indicator. [25] The composite indicator is thus nearly a bivariate signal of private debt. Because of the low number of valid indicators (3) and the concentration on one of them, the proposed signals approach is possibly not the best way to achieve signals from this set. As private debt levels in Greece are low until beginning of 2007 and only slowly rising until the beginning of 2010, a detection of the crisis in Greece is not possible with this indicator. On the other hand, signals are issued over a long time in the Netherlands, since private debt levels move around the threshold since 2001. The composite indicator of EC is actually built on five indicators, because the competitiveness indicator, based on HICP, and property prices have negative utility. Private debt, current account balance and net foreign asset position contribute a total of 86% to the indicator. Thus, the indicator is fairly imbalanced, though by far not as badly as the one proposed by the ECB. We can observe high indicator values in Spain and Portugal over most of the observation period. In Ireland, the indicator starts to increase in 2004 and stays among the highest values from

---

[25] Government Debt has a share of approximately 20%, the remaining 10% are covered by the competitiveness index, based on ULC.

2005 onwards. Other strong signals are sent for Italy, while Greece (due to its low private and foreign debts) stays at around 0.4 for most of the time until mid 2007, when they start to increase as well. The composite indicator of Heise has the largest and most diversified indicator set which tries to find a balance between four different areas of the economy. This leads to a broad range of possible values and, in general, nearly no zero values of the indicator. Comparable to the indicator of the BMWi, we can observe low indicator values until mid 2001 (with persistent deviations in Greece, Portugal, and possibly Italy and Spain). After a turbulent period following the burst of the new economy bubble (with most countries staying under a value of 0.3), we observe the long tranquil period until 2007. In this period, Ireland starts to build up imbalances, while for example Germany reduces them. During and after the financial crisis, all countries have increasing indicator values, while the mark of 0.4 is only crossed persistently by Ireland, Spain, Greece, Portugal, Italy (in that order).

6 A Proposal for a New Composite Indicator

In the previous section, we described the performance of individual indicators and of the composite indicators obtained from the four different sets (using utility weights). It became clear from table 3, that composite indicators calculated from larger and more diverse sets have better quality measures. This was probably due to the fact that imbalances showed up in nearly all different areas of the economy (captured by the five different categories of individual indicators). But the origin of future imbalances laying a foundation to possible crises is unknown. Because of this, a composite indicator built on a broad set of individual indicators will most likely (on average) be better than any composite indicator based on a small set. Using a large set means that fewer imbalances escape the eyes of the observer.

Therefore, we propose a composite indicator CI^{all} based on all individual indicators.[26] We will calculate this indicator both using utility and equal weights of $\frac{1}{N}$, N being the number of indicators Kaminsky (1999). The intuitive argument for an equally weighted composite indicator is again the ignorance as to the origins of future crises. The utility weighted composite indicator might just not give enough weight to the area where the next imbalances build up.[27] We show the indicators in figure 4 and 5 (the former shows the indicator based on utility, the latter the one based on equal weights).

---

[26] We only exclude the competitiveness indicator based on the GDP deflator because of its similarity to the one based on CPI. This way, we can keep a certain balance between the five categories. The whole set thus contains 19 indicators: the set of Heise (as we analyze it here) and additionally: competitiveness indicators based on CPI and ULC, private debt, inflation, asset and property prices.

[27] This solves at least partly the problem of overfitting: until now, we deducted our probabilities in a way that was optimal for the past. Using equal weights relaxes this restriction at least for the composite indicator. However, we do not change the (utility-optimal) threshold of the individual indicators. Indicators with a negative utility are counted as well and are evaluated at the 94%-quantile of the empirical distribution, corresponding to the unconditional probability of a crisis in the past, which was 6.04%.

Figure 4:
Values of our composite indicator CI^{all}, calculated using utility weights, for every country in the analysis. Crisis periods (indicator above a threshold of 0.4) shaded.
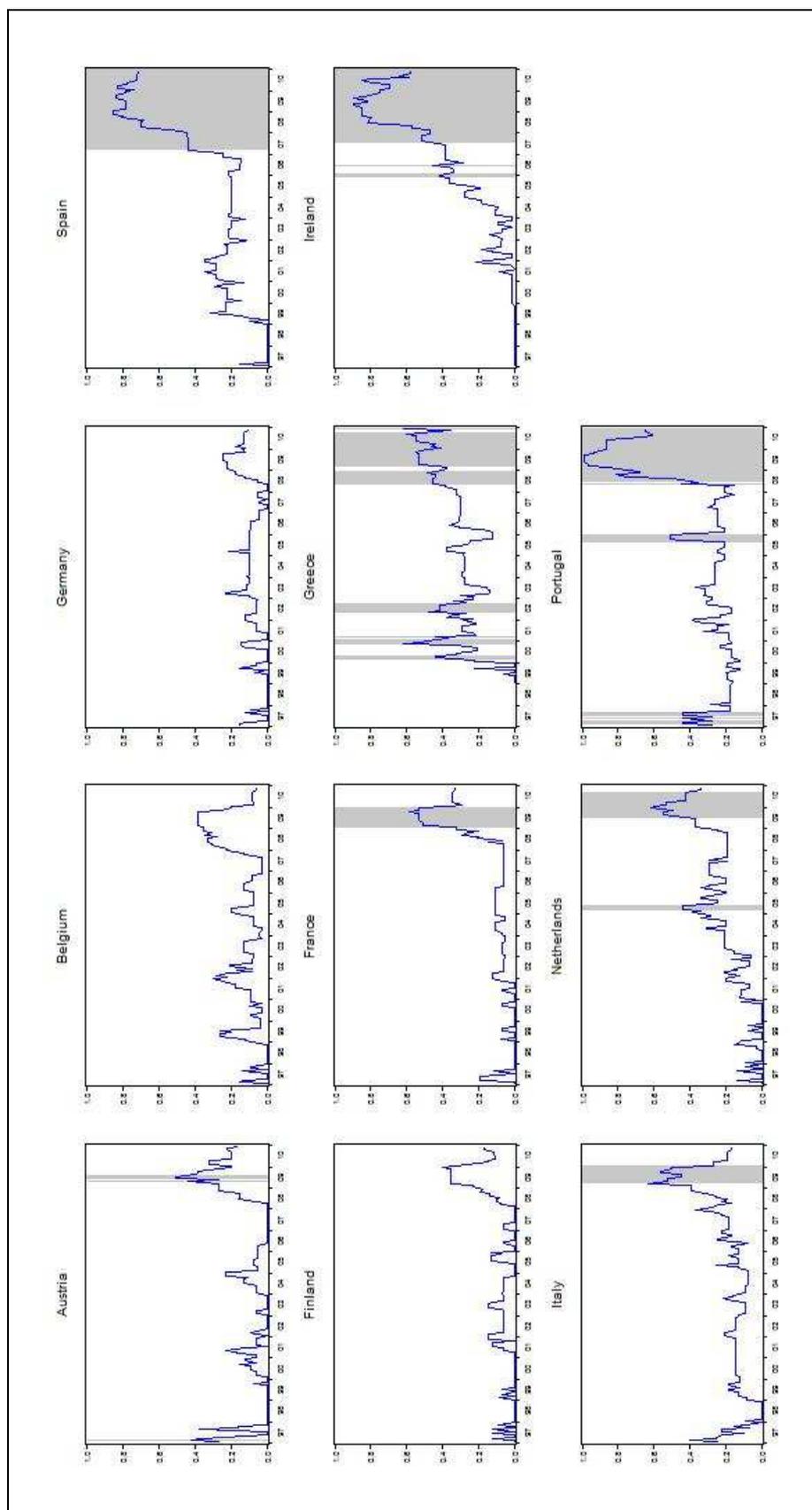
Figure 5:
Values of our composite indicator CI^{all}, calculated using equal weights, for every country in the analysis. Crisis periods (indicator above a threshold of 0.3) shaded. In general, the values are much lower due to higher weights of indicators that had negative utility in the past.
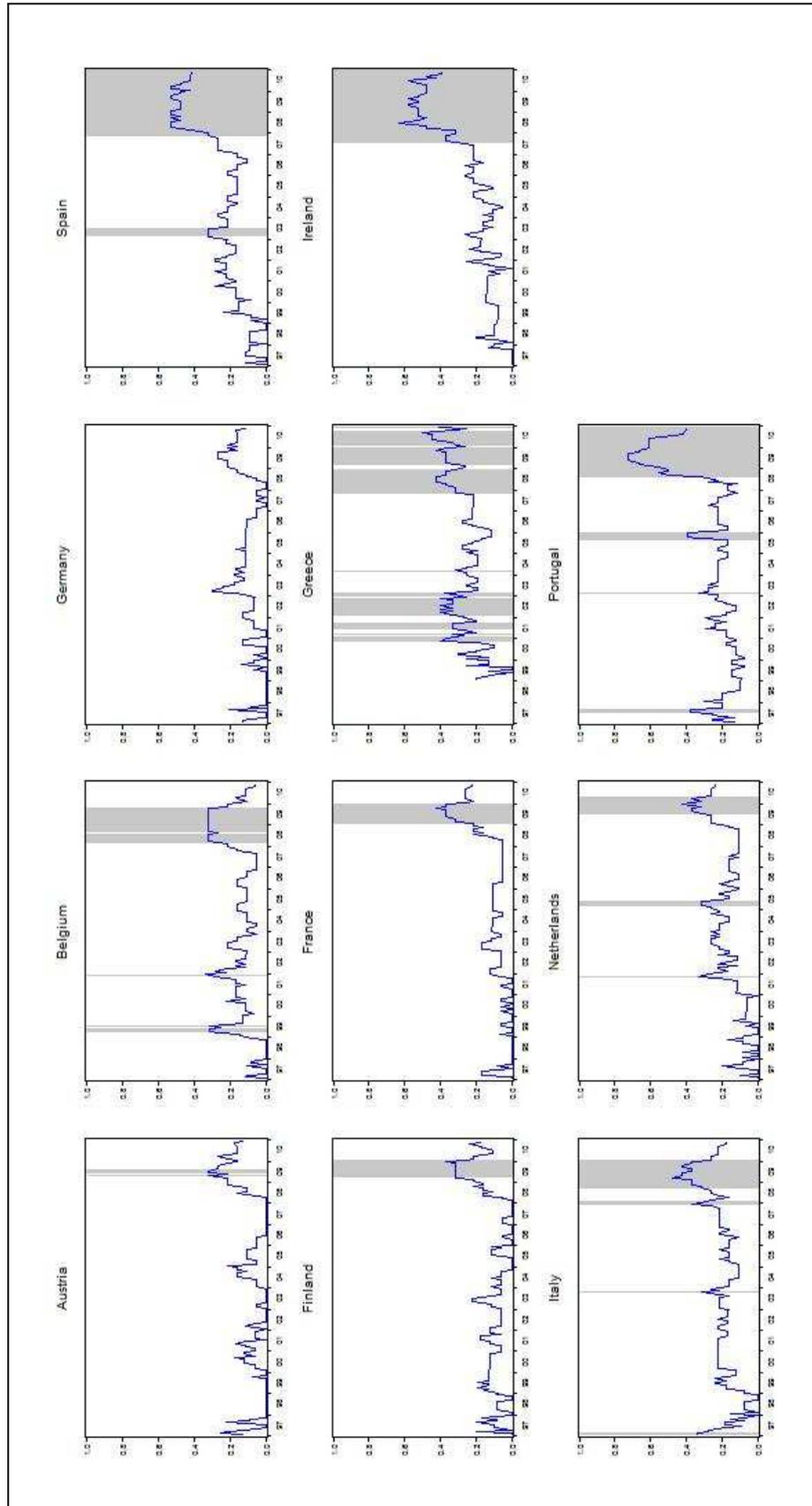
As could be expected, the highest values for the indicator based on equal weights do not exceed 0.75.[28] The crisis countries send warnings of a possible crisis (indicated by shaded bars) continuously since the financial crisis, Greece also in 2001 and 2002, directly after the introduction of the Euro. Other countries getting warnings during the height of the financial crisis are Italy, France, the Netherlands and Belgium (the latter only for the equally weighted indicator).

We can compare the performance of our composite indicator to the best performing composite indicator proposed by other authors, namely the Heise indicator. Due to the larger set of individual indicators, we would expect the following from a composite indicator: First, the utility, but also the other quality measures should be better (or at least comparable because of statistical uncertainty) to the ones of the 4 different composite indicators. Second, the utility function should have only one local (and global) maximum. Third, both the functions $P(\mathcal{C}|\mathcal{S}) = P(\mathcal{C} = 1|CI_t^k > Q_{CI})$ and $P(\mathcal{C}|CI_t^k)$ should be monotonic increasing with Q_{CI} and lb, respectively.

Figure 6:
Utility of the composite indicators CI^{all} (utility and equal weights) and from the Heise set.

---
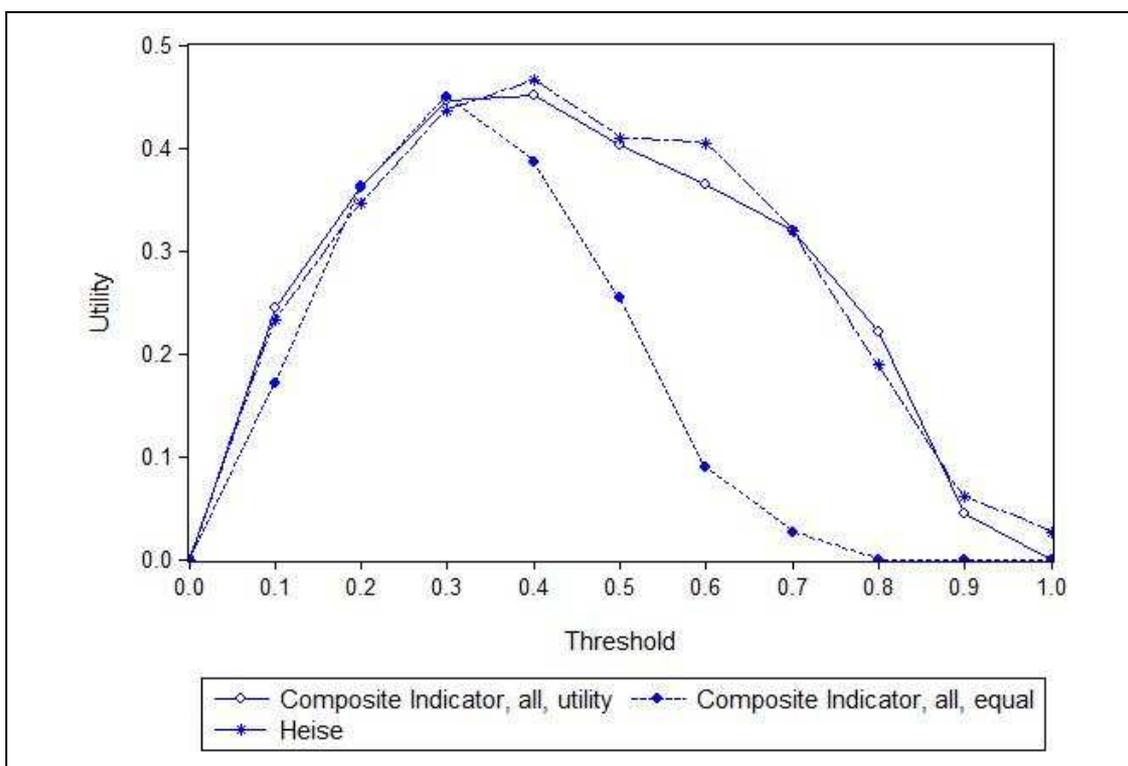
28   5 out of of 19 indicators are excluded in the utility weighted indicator set due to negative utility, but fully counted in the equally weighted composite indicator. This leads to this strong reduction of indicator values.

In figures 6 to 8 we show the utility, the probability of correct crisis forecasts $\left(P(\mathcal{C}|\mathcal{S})\right)$ and the probability of having a crisis within the next 24 month given the value of the composite indicator is in a predefined interval $P(\mathcal{C}|CI_t^k)$ for the three indicators (CI^{all} (utility and equal weights) and the Heise indicator). We find that the main difference between the Heise indicator and utility weighted CI^{all} is that the latter is more stable. The equally weighted indicator, however, reaches its maximal utility already at a threshold of 0.3 (compared to 0.4 for the others). This faster increase is due to the lower values of the indicator mentioned above. The same feature can also be observed for $P(\mathcal{C}|\mathcal{S})$ and $P(\mathcal{C}|CI_t^k)$ in figures 7 and 8. Because of this, we propose 0.3 as a threshold for CI^{all}, based on equal weights.

In addition to this general comparability to the Heise indicator, the two indicators show the expected three attributes.

Figure 7:
Probability of correct crisis forecasts $\left(P(\mathcal{C}|\mathcal{S})\right)$ for the composite indicators CI^{all} (utility and equal weights) and from the Heise set.
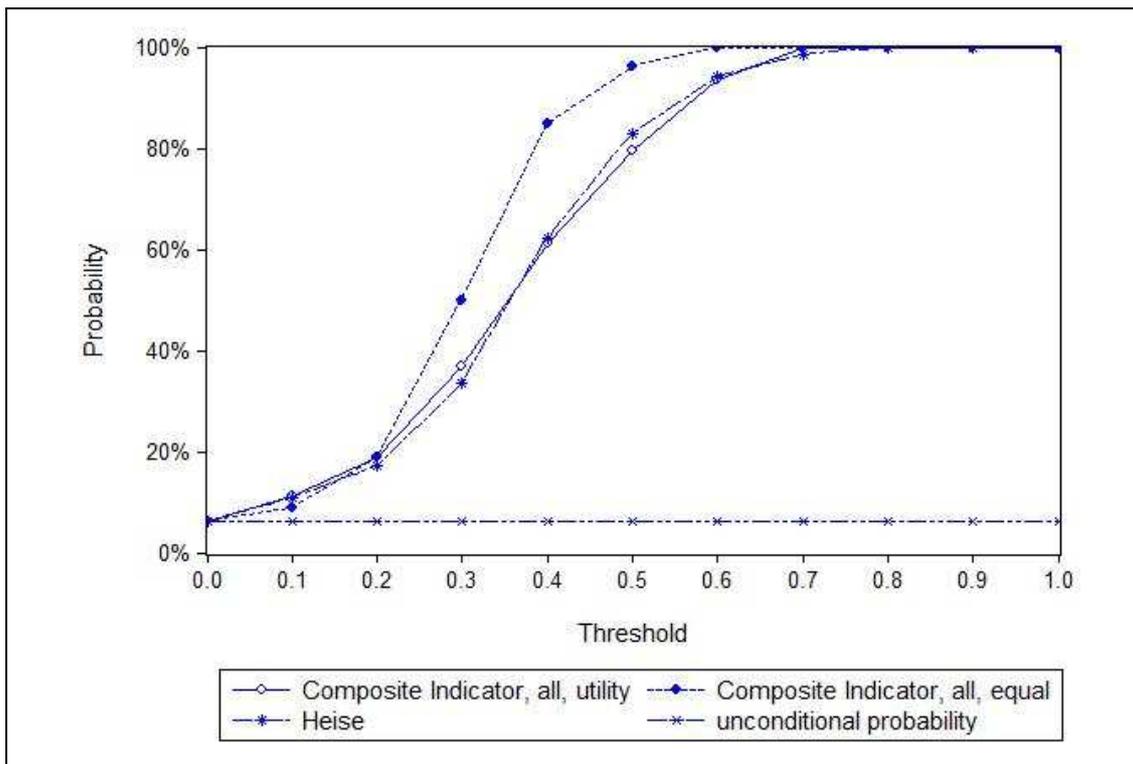
Figure 8:
Probability of having a crisis within the next 24 month given the value of the composite indicator is in a predefined interval ($P(\mathcal{C}|CI_t^k)$) for the composite indicators CI^{all} (utility and equal weights) and from the Heise set.
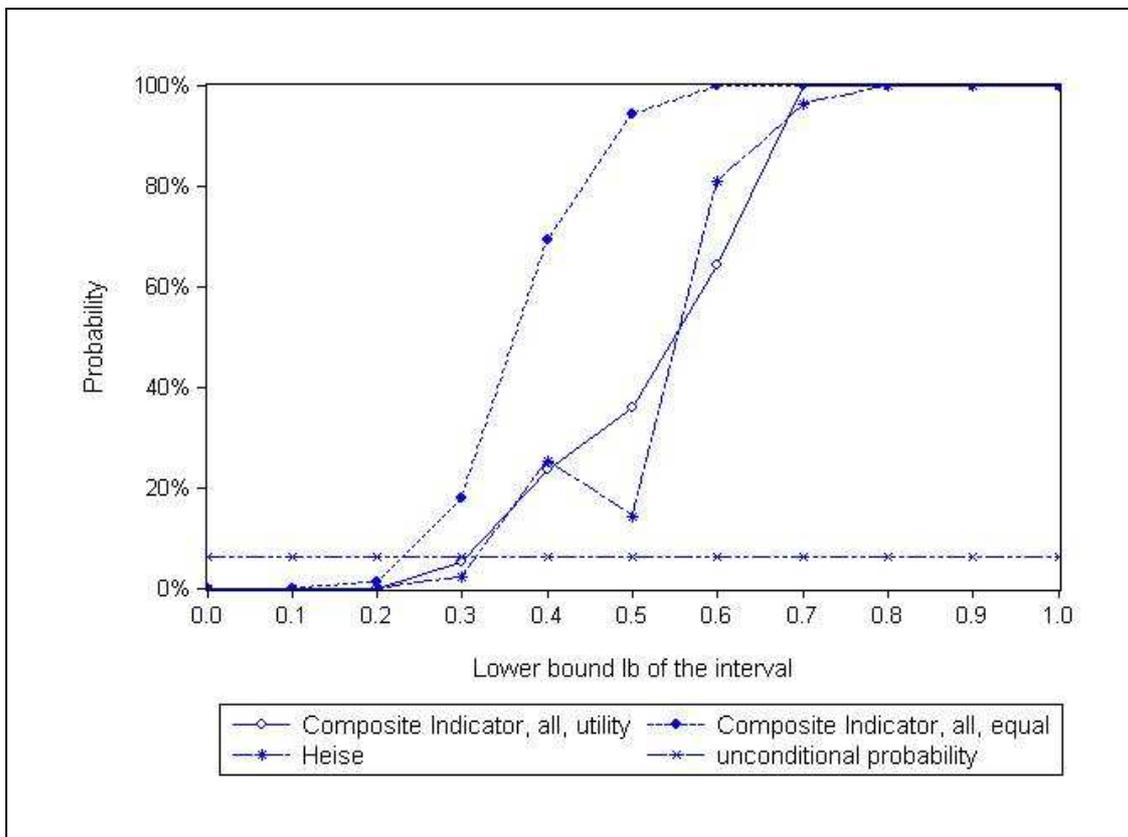
We already see from the above figures that both composite indicators are not too different from the Heise indicator. This result is found again if we look at the different performance measures in table 4. There is, however, one astounding results to be found in these measures. Moving from utility weights to equal weights, we would expect especially the maximal utility to go down, because high performing indicators get then the same weight as low performing indicators. The near-equal utility (at the maxima) points to both the high quality of the equally weighted indicator CI^{all} and the stability of the result.

Table 4:
Quality measures at optimal thresholds for composite indicators, calculated from all individual indicators using utility and equal weights. Measures using a threshold of 0.4 for CI^{all} based on equal weights given for comparison.

	CI^{all}, Utility	CI^{all}, Equal ($Q_{CI^{all}} = 0.3$)	CI^{all}, Equal ($Q_{CI^{all}} = 0.4$)
Utility	0.4507	0.4492	0.3855
Probability of correct crisis forecast ($P(\mathcal{C}|\mathcal{S})$)	61.04%	50.00%	84.78%
Probability of correct crisis/non-crisis period forecast ($\frac{A+D}{A+B+C+D}$)	96.01%	93.96%	97.83%
Quadratic probability score (QPS)	0.0288	0.0349	0.0349

Because the indicator CI^{all}, based on equal weights, avoids (at least partly) the problem, that its measured performance is only a past performance while still obtaining similar results to its utility weighted counterpart, we propose this indicator as a system to monitor macroeconomic imbalances in the EMU.

7 Robustness Analysis

In this section we test if our results are stable under different specifications. In a first part the reaction of individual indicators to different crisis definitions as mentioned in section 2 are tested. This test should establish that the imbalances we found to be valid under the crisis definition used above are really the problems that lead to the current public debt crisis in some countries of the EMU. In a second part we will present results for exogenous (politically determined) thresholds for some indicators, namely the criteria of the SGP for public debt and deficit as well as the inflation target of the ECB. Until now, the thresholds were calculated in a utility-maximizing way. The results obtained with these thresholds can be strengthened further if they do not differ too much from those obtained by using fixed thresholds. In a third part we present the performance of the composite indicators of the different sets, given different weighting schemes. This way we can further strengthen the argument that an equally weighted composite indicator may be much better for the detection of a future crisis while still not losing too much explanatory content for the past.

Effect of different crisis definitions

We could have a problem with endogeneity, meaning that the quality of the individual indicators depends too strongly on the crisis definition. Then different definitions of crises (resulting in different starting dates and slightly differing crisis countries) should result in strongly differing optimal thresholds. We point out that this is not the case. It will be seen that the signals sent by the individual indicators in most cases do not depend on the exact timing of a crisis. This leads to the conclusion that indeed the observed imbalances are the reason for the current debt crises in several countries of the EMU. Two alternative crisis definitions, as already mentioned in section 2, are:

1) A crisis occurs if the spread $S_t^k = R_t^k - R_t^{AAA} > 94.7bp$ exceeds the 95%-quantile of the empirical distribution of S, see Pescatori and Sy (2004)[29]. This definition leads to earlier crisis dates and an additional country with an occurring crisis: Italy.

2) A crisis occurs in the month, where the IMF approved non-concessional loans in excess of 100% of quota to that country, Manasse and Roubini (2008). This definition leads to only three crisis countries with very late occurrences, since (until the time of writing) only Greece, Ireland and Portugal applied for IMF loans.

In table 5, we list the beginnings of crises in the affected countries, including the standard definition used until now.[30] Table 6 presents the optimal thresholds and corresponding utilities for the different individual indicators given the three different crisis definitions.

Table 5:
Beginning of crises with different crises definitions.

	Greece	Ireland	Portugal	Spain	Italy
Standard	01.01.2010	01.06.2010	01.07.2010	01.06.2010	
Quantile	01.11.2008	01.02.2009	01.05.2010	01.06.2010	01.06.2010
IMF loans	01.05.2010	01.12.2010	01.05.2011		

---

[29] Due to extreme outcomes, the 90%-quantile of a gamma distribution fitted to the data contains the 95%-quantile of the empirical distribution in its 95%-confidence interval. Pescatori and Sy (2004) show that their chosen crisis measure corresponds as well to the 90%-quantile of their dataset.

[30] The reason for Portugal having a crisis one month after Spain after the standard definition and one month before using the quantile definition is due to Portuguese spreads lying very close to, but below the standard threshold in May and June 2010.

Table 6:
Utilities and optimal thresholds of the individual indicators for different crises definitions in comparison.

	Utility			Optimal threshold		
Crisis Definition	**Standard**	**Quantile**	**IMF loans**	**Standard**	**Quantile**	**IMF loans**
Government deficit	0.4315	0.3338	0.4900	87%	83%	93%
Unemployment rate	0.3771	0.2334	0.3747	93%	85%	91%
Non-MFI debt	0.3367	0.2666	0.2894	88%	89%	91%
Household debt	0.3318	0.2629	0.2920	84%	86%	91%
Current account	0.3241	0.3677	0.2692	68%	68%	58%
Private debt	0.3153	0.2406	0.2590	84%	84%	85%
Unit labor cost	0.2506	0.2127	0.2306	86%	68%	86%
Foreign asset	0.1926	0.2475	0.1008	78%	67%	80%
Trade share	0.1861	0.1318	0.1844	85%	94%	78%
Labor-force participation	0.1858	0.2027	0.1843	64%	65%	89%
Labor productivity	0.1620	0.0991	0.1560	50%	74%	51%
Domestic demand	0.1506	0.1396	0.2661	90%	87%	92%
Government debt	0.0865	0.1329	0.2528	71%	84%	70%
ULC-competitiveness	0.0471	0.1109	0.0100	89%	66%	90%
Inflation	0.0319	0.0828	0.0234	91%	66%	91%
HICP-competitiveness	-0.0163	0.0356	-0.0265	95%	70%	95%
GDP-deflator-competitiveness	-0.0263	0.0562	-0.0258	95%	56%	95%
Interest payments	-0.0269	-0.0278	-0.0263	95%	95%	95%
Asset prices	-0.0267	-0.0274	-0.0261	95%	95%	95%
Property prices	-0.0273	-0.0282	-0.0266	95%	95%	95%

We see that indeed optimal indicator thresholds do hardly depend on the different crisis definitions. There are, however, some exceptions. We find dampened signals towards the end of the observation period inducing lower optimal thresholds for the crisis definition using IMF loans (GDP-competitiveness and asset prices for Portugal, current account for Ireland). We also find effects on indicators which perform better than average

in the countries that are in- or excluded, depending on the crisis definition. Labor-force participation does not send many signals in Spain. With the crisis definition of IMF loans, this results in much higher optimal thresholds. Inflation and foreign assets are better in Italy than in the other crises countries, leading to much lower optimal thresholds with crises defined by quantiles.

As we find that the optimal thresholds of the indicators do not depend on the crisis definition, we can deduce that the indicators chosen in the four analyzed proposals indeed reflect macroeconomic imbalances amounting to risks for the stability of public debt as presented by the utilities in table 2.

Exogenous thresholds

As an alternative to the endogenously defined optimal thresholds, one could also use exogenously given thresholds, e.g. for reasons of simplicity or transparency. Such thresholds exist for government deficit (3%) and government debt (60%) in the SGP, and for inflation (2%)[31] as a target of the ECB. In table 7, we compare the performance for these three indicators using standard (optimized utility in the standard crisis definition) thresholds and the exogenously defined ones. We see, that the optimal thresholds are much lower, especially the optimal quantiles of public debt and price development are below 50% (meaning, that in more than 54% of the periods signals would be sent). Correspondingly, utility and probability of a crisis given a signal $P(\mathcal{C}|\mathcal{S})$ are lower for all three indicators. However, the rules of the SGP and of the ECB are concerned with the long-term stability. The proposals for the detection of macroeconomic imbalances are concerned with the medium-term occurrence of a crisis, thus they should sent signals at more critical values.

Table 7:
Comparison of indicator performance for standard (optimized) and exogenous thresholds

	Quantile		Absolute Threshold		Utility		Probability of correct crisis forecast $P(\mathcal{C}|\mathcal{S})$	
Threshold Determination	Standard	Exogenous	Standard	Exogenous	Standard	Exogenous	Standard	Exogenous
Government deficit	87%	65%	5.09%	3.00%	0.43154	0.33961	52.25%	20.42%
Government debt	71%	46%	71.01%	60.00%	0.08653	0.04883	11.87%	8.94%
Inflation	91%	46%	3.84%	2.00%	0.03194	- 0.10360	10.07%	3.88%

---

31   We approximate the target of below but close to 2 % by 2 %.

Different weighting schemes

The final robustness check concerns the construction of the composite indicators. In section 6 we already used equal weights for the composite indicator, based on all individual indicators. Here, we are going to present the composite indicators of the four proposed sets, weighted equally and by inverse noise-to-signal ratio (NSR). The idea behind the usage of equal weights – as described above – is the uncertainty about the origins of future crisis. The reason for testing weights according to inverse NSR is slightly different. We established in section 3 the comparability of NSR and utility. Thus, there should not be huge differences between the composite indicators based on these measures.[32]

Table 8 contains the utility, the probability of a crisis given a signal $P(\mathcal{C}|\mathcal{S})$, the share of correctly signaled periods $\frac{A+D}{A+B+C+D}$ and the quadratic probability score for both weighting schemes and the four composite indicators. For the equally weighted composite indicator, we find comparable results for the BMWi and Heise set. These two sets incorporate indicators with quite balanced utilities, leading to a small difference to the utility weighted composite indicators. There are huge differences especially for the ECB, but also for the EC set, since these sets include most of the individual indicators with negative utility (competitiveness indicators, asset and property prices). Since these are fully included in the equally weighted composite indicator, they are responsible for the large drop in utility – reaching even negative utility for the ECB set.[33] On the other hand, for the composite indicator, based on inverse NSR-weights, the significant changes are to be found in the BMWi and Heise indicator. The difference originates in the now huge share of the unemployment rate (Heise: 34% compared to 13%, BMWi: 60% compared to 29% in the utility weighted indicator) as a component of the composite indicators. As can be seen in table 2, the unemployment rate has by far the lowest share of type-II errors, leading to an extremely high inverse NSR. This decreases utility of the composite indicators, but increases strongly the probability of a crisis given a signal (since the indicator of unemployment rate rarely sends false signals).

---

[32] There will be differences in case of low type-II errors (denominator of the inverse NSR), since this results in extremely (in the limit infinitely) high weights.

[33] Arguably, the highest utility for the ECB indicator using equal weights is reached at a threshold of 0.1 and not at 0.4 due to the huge number of missing signals at larger thresholds. We present here the figures for a threshold of 0.4 for reasons of comparability.

Table 8:
Different quality measures for the composite indicators of the proposed sets, calculated using equal and inverse-NSR weights

	Weights	BMWi	ECB	EC	Heise
Utility	equal	0.4410	-0.0138	0.2194	0.4341
	inverse NSR	0.3710	0.3194	0.3802	0.4054
Probability of correct crisis forecast ($P(\mathcal{C}|\mathcal{S})$)	equal	41.45%	2.63%	27.18%	58.33%
	inverse NSR	73.08%	30.24%	35.68%	73.45%
Probability of correct crisis/non-crisis period forecast ($\frac{A+D}{A+B+C+D}$)	equal	91.55%	89.61%	88.59%	95.53%
	inverse NSR	96.86%	88.04%	89.79%	97.16%
QPS	equal	0.0498	0.1016	0.0970	0.0340
	inverse NSR	0.0376	0.0834	0.0700	0.0239

Using these variations we find that the results for our individual and composite indicators do not change dramatically. Strong changes could only be observed for different weighting schemes. However, as they were mostly restricted to the unbalanced composite indicators of the ECB and EC (which are more the bivariate signal of a single indicator than a proper diversified composite indicator), this strengthens the analysis of the indicators and the validity of the approach.

8   Summary and Conclusion

The current problems of some countries of the European Monetary Union to service their debts, manifested in high spreads of interest yields on government bonds, lead to the widespread wish of a stronger surveillance of macroeconomic imbalances. This surveillance is supposed to go far beyond the criteria of government debt and deficit of the stability and growth pact.

We analyzed four existing proposals for possible indicators of macroeconomic imbalances in this paper. Using the framework of a signals approach we test weather these indicators of macroeconomic imbalances could be used as an early-warning system for public debt crises. We found out that – consistent with economic theory – most indicators were helpful in signaling the current public debt crises. For the indicators that did not give valid signals before the current crisis, the missing signals could be explained by either the new-economy-bubble, the exchange rate development of the Euro, or the convergence of interest rates on public debt after the introduction of the euro. All other indicators showed a slow building-up of macroeconomic imbalances in very different

areas of the economies (covered by different indicators). The culminated imbalances lead to economic problems triggered by the financial crisis that we could observe mainly in the four crisis countries (Greece, Ireland, Portugal, and Spain). The results for single indicators were stable under different crisis definitions. This leads to the conclusion, that they reflect relevant economic imbalances and problems that lead to deteriorating public debts.

Going further from the single indicators, we analyzed composite indicators for every proposed indicator set. Not surprisingly, the in-sample forecast quality (as measured in terms of utility) of the composite indicators is better than that of almost all single indicators.[34] A further improvement can be achieved by combining all indicators of all different sets into one single composite indicator. Doing this, we had to be careful to keep a certain balance between the five different categories. The composite indicator CI^{all}, introduced in section 6, outperformed (or has been as good as) all other composite indicators not only in terms of utility, but also in terms of different other quality measures. This recommended indicator also reflects the desire of ECB (2010) to ideally use only one indicator to have condensed information about the amount of macroeconomic imbalances in Europe while still capturing as many different developing imbalances as possible.

In addition to the optimal (in terms of utility) construction of composite indicators, we also showed the results obtained by using different weighting schemes. This robustness check is necessary since we observed debt crises in only four out of eleven countries since the introduction of the Euro. This means that there might be as well other sorts of imbalances leading to public debt crises in the future. The purely theoretical reason for other weighting schemes (in absence of any knowledge of future developments, one would choose equal weights for every indicator) is strongly supported by the results of the indicator CI^{all}, calculated with equal weights and presented in section 6.

Because of our ignorance of the exact location, where the next crisis will start to develop and unfold, we propose to combine as many meaningful single indicators as possible in one composite indicator using equal weights. We also suggest undertaking in-depth quantitative and qualitative analysis of country cases if a risk is signaled. This will hopefully help to issue warnings and allows for time to address potential problems before they culminate in a fully grown crisis like the current ones.

---

34 This is not entirely true for the sets of the ECB and the EC, since they rely heavily on competitiveness indicators that could not send any valid signals before the public debt crises. Thus, these sets depend almost completely on the current account deficit (in case of the EC set) and government debt. Since the values of the indicators are translated in a bivariate signal, the ECB and EC composite indicator is nearly bivariate itself.

References

Alessi, L. and Detken, C. (2011) 'Quasi real time early warning indicators for costly asset price boom/bust cycles: A role for global liquidity'. European Journal of Political Economy, in press.

Beetsma, R. and Uhlig, H. (1999) 'An Analysis of the Stability and Growth Pact', Economic Journal, Vol. 109, No. 458, pp. 546-571.

Bhattacharyay, B. (2009) 'Towards a Macroprudential Surveillance and Remedial Policy Formulation System for Monitoring Financial Crisis'. CESifo Working Paper, No. 2803.

Brüggemann, A. and Linne, T. (2002) ‚Are the Central and Eastern European Transition Countries still Vulnerable to a Financial Crisis? Results from the Signals Approach'. IWH-Discussion Paper, No. 157.

Bundesministerium für Wirtschaft und Technologie (2010) ‚Ein neues Verfahren für die wirtschaftspolitische Koordinierung in Europa'. Schlaglichter der Wirtschaftspolitik, Monatsbericht August 2010, Berlin: BMWi.

Brier, G.W. (1950) 'Verification of forecasts expressed in terms of probability'. Monthly weather review, Vol. 78, No. 1, pp. 1-3.

Bussière, M and Fratzscher, M. (2008) 'Low probability, high impact: Policy making and extreme events'. Journal of Policy Modeling Vol. 30, No. 1, pp. 111-121.

Council of the Euopean Union (2011) 'State of play concerning the Economic Governance Package, 7579/11'. 11 March 2011. Available at: «http://register.consilium.europa.eu/pdf/en/11/st07/st07579.en11.pdf».

Demirgüç-Kunt, A. and Detragiache, E. (1999) 'Monitoring Banking Sector Fragility: A Multivariate Logit Approach with an Application to the 1996-97 Banking Crises'. Policy Research Working Paper 2085 (Washington, DC: World Bank).

Detragiache, E. and Spilimbergo, A. (2001) 'Crises and Liquidity – Evidence and Interpretation'. IMF Working Paper, No. 01/2.

Edison, H.J. (2003) 'Do indicators of financial crises work? An evaluation of an early warning system'. International Journal of Finance and Economics, Vol. 8, No. 1, pp. 11-53.

European Commission (2010a) 'EU economic governance: the Commission delivers a comprehensive package of legislative measures, IP/10/1199'. 29 September 2010. Available at: «http://europa.eu/rapid/pressReleasesAction.do?reference=IP/10/1199&format=HTML&aged=0&language=EN&guiLanguage=en».

European Commission (2010b) 'Economic governance package (2): Preventing and correcting macroeconomic imbalances, MEMO/10/454'. 29 September 2010. Available at: «http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/10/454&format=HTML&aged=0&language=EN&guiLanguage=en»

European Commission (2010c) 'Economic governance package (3): Chronology and overview of the new framework of surveillance and enforcement, MEMO/10/456'. 29 September 2010. Available at:

«http://europa.eu/rapid/pressReleasesAction.do?reference=MEMO/10/456&format=HTML&aged=0&language=EN&guiLanguage=en».

European Council (2010) 'Strengthening Economic Governance in the EU, Report of the Task Force to the European Council (van Rompuy Task Force)'. 21 Oktober 2010.   Available   at: «http://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/117429.pdf».

European Central Bank (2010) 'Reinforcing Economic Governance in the Euro Area'. 10 Juni 2010. Available at:

«http://www.ecb.int/pub/pdf/other/reinforcingeconomicgovernanceintheeuroarean.pdf»

Gregoriou, A., Kontonikas, A. and Montagnoli, A. (2011) 'Euro Area Inflation Differentials: Unit Roots and Nonlinear Adjustment'. Journal of Common Market Studies, Vol. 49, Nr. 3, pp. 525-540.

Von Hagen, J. and Hammond, G.W. (1998) 'Regional Insurance against Asymmetric Shocks: An Empirical Study for the European Community'. The Manchester School, Vol. 66, No. 3, pp 331-53.

Heinen (2011) ,Macroeconomic coordination. What can a scoreboard approach achieve?' EU-Monitor, No. 78, pp. 1-19.

Heise, M. (2011) ,Notwendigkeit und Ausgestaltung makroökonomischer Überwachung im Euroraum' Wirtschaftsdienst, Vol. 91, No. 1, pp. 22-30.

Heise, M., Lindeck, S., Petersen, A.-K., Schneider, R. and Voegeli, K. (2010) 'The Euro Monitor 2010 - Indicators for Balanced Growth'. 25. Oktober 2010. Available at: «https://www.allianz.com/de/economic_research/publikationen/spezialthemen/monitor2010.html».

Kaminsky, G.L. and Reinhart, C.M. (1999) 'The Twin Crises: The Causes of Banking and Balance-of-Payments Problems'. American Economic Review, Vol. 89, No. 3,
pp. 473-500.

Kaminsky, G.L., Lizondo, S. and Reinhart, C.M. (1998) 'Leading Indicators of Currency Crises'. IMF Staff Paper, No. 45.

Kaminsky, G.L. (1999): 'Currency and banking crises: the early warnings of distress'. IMF Working Paper, No. 99/178.

Knedlik, T. and Scheufele, R. (2008) 'Forecasting currency crises: Which methods signaled the South African currency crisis of June 2006?'. South African Journal of Economics, Vol. 76, No. 3, pp. 367-383.

Lane, P.R. (2011) 'The Irish Crisis', CEPR Discussion Paper, No. 8287.

Manasse, P., Roubini, N. and Schimmelpfennig, A. (2003) 'Predicting Sovereign Debt Crises'. IMF Working Paper, No. 03/221.

Manasse, P. and Roubini, N. (2009) ''"Rules of Thumb" for Sovereign Debt Crises'. Journal of International Economics, Vol. 78, No. 2, pp. 192-205

Minsky, H.P. (1972) 'The Financial Instability Hypothesis'. Reprinted as Working Paper No. 74 (1992) by the Jerome Levy Economics Institute of Bard College.

Pescatori, A. and Sy, A.N.R. (2004) 'Debt Crises and the Development of International Capital Markets'. IMF Working Paper, No. 04/44.

Reinhart, C.M. (2002) 'Default, Currency Crises and Sovereign Credit Ratings'. The World Bank Economic Review, Vol. 16, No. 2, pp. 151-170.

Reinhart, C.M. and Rogoff, K.S. (2010) 'From Financial Crash to Debt Crisis'. American Economic Review, forthcoming.

Suarez, J. (2010). 'The Spanish Crisis: Background and Policy Challenges'. Moneda y Crédita, forthcoming.

Sy, A. N. R. (2004) 'Rating the Rating Agencies: Anticipating Currency Crises or Debt Crises'. Journal of Banking and Finance, Vol. 28, No. 11, pp. 2845-2867.

Annex I:

Spreads of government bonds with a maturity of 10 years over the average of a AAA-rated government bond. Crisis periods shaded, 0 spread marked by the straight line.
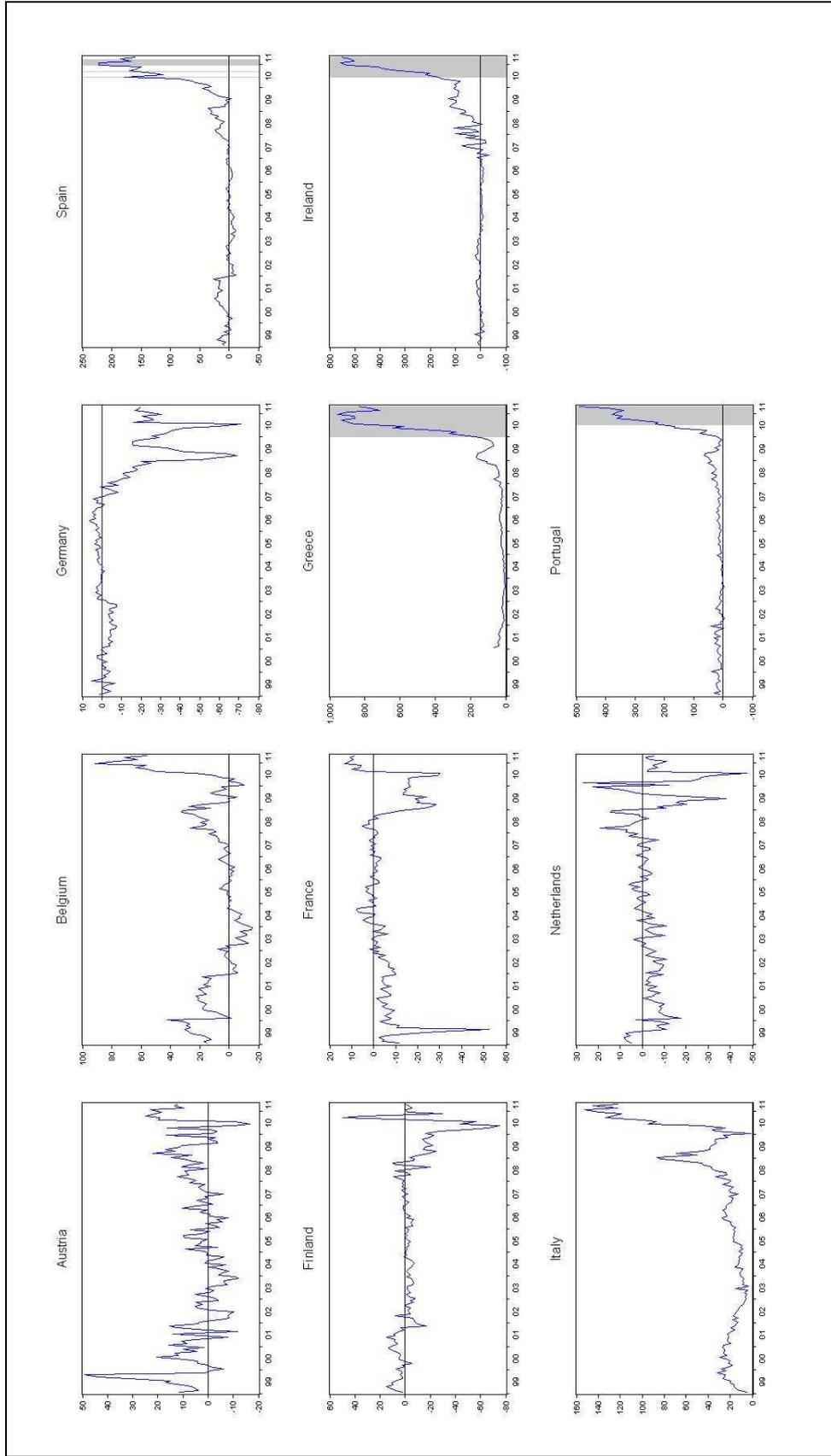
Annex II:
List of individual indicators

Category	No.	Heise	BMWi	EC	ECB	Indicator employed
Fiscal	1	Gross government debt as percent of GDP		Government debt	Indebtedness of public sector (qual.)	Government debt (gross general government debt as percent of GDP, Eurostat, quarterly)
	2	Gross government deficit/surplus as percent of GDP				Government deficit (general government net lending/net borrowing as percent of GDP, Eurostat, quarterly)
	3	General government interest payments as percent of total government expenditure				Interest payment (General government interest payments as percent of general government expenditure, Eurostat, quarterly)
	4	Required adjustment in the primary balance due to demographic ageing in percent points				X
Competitiveness and domestic demand	5	Unit labor costs, total economy, deviation from the target path of 1.5 percent rise per year in index points	Unit labor cost development (Lohnstückkosten-entwicklung)			Unit labor costs (as percent change, yoy, Eurostat, quarterly)
	6	Current account balance as percent of GDP	Current account balance (Leistungs- und Handelsbilanz-salden)	Current account balances	External imbalances (qual.)	Current account (current account balance as percent of GDP, OECD, quarterly)
	7	Global merchandise trade shares, exports, deviation from base year 2000 in percent				Trade share (percent change of the share of exports (IMF-IFS, monthly) in world trade (CPB export volume index, world, monthly))
	8	Domestic demand, average annual change over the last five years				Domestic demand (average annual change over the last five years, Eurostat, quarterly)
	9				Competitiveness index based on HICP	HICP-Competitiveness (Competitiveness index based on HICP, as percentage change, yoy, ECB, monthly)
	10			Real effective exchange rate based on GDP deflator	Competitiveness index based on GDP-deflator	GDP-deflator-Competitiveness (Competitiveness index based on GDP-deflator, as percentage change, yoy, ECB, quarterly)
	11			Real effective exchange rate based on unit labor costs	Competitiveness index based on ULC	ULC -Competitiveness (Competitiveness index based on ULC, as percentage change, yoy, ECB, quarterly)
	12				Deviations from stability-oriented sectoral/national wage developments	X

Category	No.	Heise	BMWi	EC	ECB	Indicator employed
	13				Degree of convergence (qual.)	X
Asset prices	14		Price level developments (Preisniveau-entwicklung)			Inflation (HICP, percent change, yoy, ECB, monthly)
Asset prices	15			Increases in real house prices		Property prices (residential property price index, percent change, yoy, ECB, quarterly)
Asset prices	16				Asset prices (qual.)	Asset prices (MSCI equity price index, percent change, yoy, monthly)
Labor	17	Harmonised unemployment rate percent	Unemployment rate (Arbeitslo-senquote)			Unemployment rate (harmonized unemployment rate, change, yoy, Eurostat, monthly)
Labor	18	Employment ratio, change over five years in percentage points				Labor-force participation (percent change, yoy, Eurostat, quarterly)
Labor	19	Labor productivity per person employed, average annual change over the last five years				Labor productivity (average annual change over the last five years, Eurostat, quarterly)
Labor	20	Gross inland consumption of energy divided by GDP (kilogram of oil equivalent per EUR 1000)				X
Private and foreign debt	21		Private debt (Private Ver-schuldung)	Ratio of private sector credit to GDP	Indebtedness of private sector (qual.)	Private debt (loans to MFI and non-MFI as percent of GDP, ECB, monthly)
Private and foreign debt	22	Debt-to-GDP ratio of households, change over five years in percentage points				Household debt (loans to household as percent of GDP, ECB, monthly)
Private and foreign debt	23	Debt-to-GDP of non-financial corporations, change over five years in percentage points				Non-MFI debt (loans to non-MFI as percent of GDP, ECB, monthly)
Private and foreign debt	24	Net international investment position as percent of GDP		Net foreign asset position		Foreign assets (Net foreign assets position as percent of GDP, IMF-IFS, monthly)
Private and foreign debt	25				Credit booms (qual.)	X