

**Predicting Financial Crises:
The (Statistical) Significance of the Signals Approach**

*Makram El-Shagi
Tobias Knedlik
Gregor von Schweinitz*

April 2012

No. 3

IWH-Diskussionspapiere

IWH Discussion Papers

Authors: *Makram El-Shagi*
Department of Macroeconomics
E-mail: Makram.El-Shagi@iwh-halle.de
Phone: +49 345 7753 835

Tobias Knedlik
Department of Macroeconomics
E-mail: Tobias.Knedlik@iwh-halle.de
Phone: +49 345 7753 740

Gregor von Schweinitz
Martin-Luther-University Halle-Wittenberg and
Halle Institute for Economic Research
Department of Macroeconomics
E-mail: Gregorvon.Schweinitz@iwh-halle.de
Phone: +49 345 7753 744

The responsibility for discussion papers lies solely with the individual authors. The views expressed herein do not necessarily represent those of the IWH. The papers represent preliminary work and are circulated to encourage discussion with the authors. Citation of the discussion papers should account for their provisional character; a revised version may be available directly from the authors.

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

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

Editor:
HALLE INSTITUTE FOR ECONOMIC RESEARCH – IWH
The IWH is a member of the Leibniz Association.

Address: Kleine Maerkerstrasse 8, D-06108 Halle (Saale), Germany
Postal Address: P.O. Box 11 03 61, D-06017 Halle (Saale), Germany
Phone: +49 345 7753 60
Fax: +49 345 7753 820
Internet: <http://www.iwh-halle.de>

ISSN 1860-5303 (Print)
ISSN 2194-2188 (Online)

Predicting Financial Crises: The (Statistical) Significance of the Signals Approach

Abstract

The signals approach as an early warning system has been fairly successful in detecting crises, but it has so far failed to gain popularity in the scientific community because it does not distinguish between randomly achieved in-sample fit and true predictive power. To overcome this obstacle, we test the null hypothesis of no correlation between indicators and crisis probability in three applications of the signals approach to different crisis types. To that end, we propose bootstraps specifically tailored to the characteristics of the respective datasets. We find (1) that previous applications of the signals approach yield economically meaningful and statistically significant results and (2) that composite indicators aggregating information contained in individual indicators add value to the signals approach, even where most individual indicators are not statistically significant on their own.

Keywords: early warning system, signals approach, bootstrap

JEL Classification: C15, E60, F01

Krisenvorhersage: Die (statistische) Signifikanz des Signalansatzes

Zusammenfassung

Während Signalansätze im Rahmen von Frühwarnsystemen recht erfolgreich angewendet werden, fehlt es ihnen bislang an akademischem Ansehen, da sie nicht zwischen der möglicherweise zufällig entstehenden und tatsächlich bestehenden Anpassungsgüte innerhalb einer Stichprobe unterscheiden können. Zur Überwindung dieses Problems testen wir für drei Anwendungsbeispiele des Signalansatzes die Nullhypothese, dass die Indikatoren und die Krisenwahrscheinlichkeit nicht korreliert sind. Dazu verwenden wir Bootstrap-Methoden, die individuell auf die Eigenschaften des Datensatzes angepasst sind. Die zentralen Ergebnisse der Untersuchung sind: (1) Die untersuchten Anwendungen des Signalansatzes liefern ökonomisch sinnvolle und statistisch signifikante Ergebnisse. (2) Die Verwendung von Gesamtindikatoren, die Informationen aus den einzelnen Indikatoren zusammenfassen, ist selbst dann sinnvoll, wenn die meisten Einzelindikatoren keine signifikanten Ergebnisse erzielen.

Schlagwörter: Frühwarnsysteme, Signalansatz, Bootstrap-Methoden

JEL-Klassifikation: C15, E60, F01

1 Introduction

The use of the signals approach or indicators approach to construct early warning systems for financial crises goes back to the seminal papers of Kaminsky et al. (1998) and Kaminsky and Reinhart (1999). The signals approach is widely used to forecast financial crises. Current applications include early warning systems for currency crises (Edison, 2003), banking crises (Borio and Drehmann, 2009), asset price bubbles (Alessi and Detken, 2011), and debt crises (Knedlik and von Schweinitz, 2011). It has also been used by the International Monetary Fund and other policy makers. The approach has the advantage of simplicity and traceability, but it has been criticized for being a nonparametric approach that does not allow users to derive crisis probabilities directly and to judge the significance of individual indicators in a composite indicator, or the significance of a composite indicator as such. The analysis reported in this paper focuses on the statistical significance of indicators derived by means of the signals approach. In particular, we test whether the calculated maximum utilities or minimum noise-to-signal ratios of individual indicators and/or composite indicators could be the result of a random process or actually have statistical power. We use datasets from previous studies that apply the signals approach. To test the significance of the signals approach, we include studies addressing different types of crises. We use data from Kaminsky and Reinhart (1999) to cover currency and banking crises, from Alessi and Detken (2011) to cover asset price bubbles, and from Knedlik and von Schweinitz (2011) to cover sovereign debt crises. By decomposing the statistical components of the time series and bootstrapping the sample, we simulate random time series with similar characteristics as the original data. Thereby, we are able to reproduce the distribution of the performance of both individual indicators and the composite indicator under the null hypothesis of no predictive power of the considered indicators. The results show that the previous applications of the signals approach yield significant results. Most notably, composite indicators can have high (and highly significant) predictive power even when they are based on mediocre individual indicators.

The remainder of the paper is structured as follows: the next section gives an introduction to the literature of early warning systems and relates the signals approach to this literature, describes the signals approach with the focus on the optimization methods used in the calibration of thresholds beyond which signals are issued by the indicators, and introduces the papers that we consider to test the significance of the signals approach; in Section 3, we present our methodology and testing procedure, and in Section 4 the results are set out, before recapping the conclusions in the final section.

2 The signals approach and alternatives

2.1 Early warning systems

Approaches to early warning systems vary with regard to the techniques employed. The signals approach is among the first and most widely used techniques. Two other prominent approaches are binary-choice models and the Markov-switching approach.¹

The signals approach, introduced in the seminal paper by Kaminsky and Reinhart (1999), is a nonparametric threshold approach. An indicator is read as sending a warning signal for an (upcoming) crisis if it exceeds a threshold that is chosen to optimize the predictability of a crisis. Various indicators can be summarized to form a composite indicator (Kaminsky, 1999). Applications in the literature include Berg and Pattillo (1999); Brüggemann and Linne (2002); Edison (2003).

Similar to the signals approach, binary choice models aim to predict a binary crisis variable (see e.g. Frankel and Rose, 1996; Berg and Pattillo, 1999; Kamin et al., 2001; Kumar et al., 2003; Bussière and Fratzscher, 2006). Binary choice models fit a specific stable relation between the variation in a set of indicator variables and a latent variable that translates to crisis probability at different points in time. However, establishing a stable relation between indicators and crises requires quite rigid assumptions, including a latent variable that is both linearly dependent on the indicators and strictly monotonously related to crisis probability.

Unlike the other approaches, the Markov-switching approach does not depend on an a priori definition of crises. Instead, it assumes that there are two regimes of the economy that exhibit different characteristics with regard to the model parameters, where one regime is interpreted as a

¹See Abiad (2003) for a more detailed survey on Early-Warning Systems.

crisis state and the other as a non-crisis state. The identification of crisis periods is done simultaneously with the parameter estimation of the regimes. The probability of a regime switch can be directly interpreted as crisis probability. Early papers on regime-switching models (including the seminal paper by Hamilton (1989)) deal with time-constant crisis probabilities, whilst the applications for crisis forecasts usually employ versions with time-varying crisis probabilities (Filardo (1994); Diebold et al. (1994)). In some studies, the Markov-switching approach yields a better forecasting performance than binary choice models or the signals approach (Abiad, 2003; Mariano et al., 2004; Kittelmann et al., 2006; Knedlik and Scheufele, 2008). However, since crises are not defined, the economic interpretation of the regimes is arbitrary. In addition to this drawback, applications of the Markov-switching approach can only deal with a limited number of (representative) variables. In view of the fact that in policy advice, the objective of crisis prediction tools is usually the monitoring of a broad set of macroeconomic risks, the applicability of Markov-switching models is fairly limited.

Given the limitations of the alleged refinements in crisis prediction, the original signals approach still offers substantial added value. First, it allows for strong nonlinearities in the relation between indicators and the probability of a crisis, without any need for highly restrictive assumptions such as the specific functional form of the nonlinearity that is required for many nonlinear techniques. The second advantage is that it is straightforward to see what causes a signal issued by the composite indicator, because the composite signal is constructed from signals sent by individual indicators. This characteristic is particularly important, if the crisis forecast is meant to allow political decision-makers to identify the sources of the current risk of distress, enabling them to respond appropriately. However, due to the simplicity of this approach, it is difficult to judge to what extent the in-sample predictive power of the individual indicators (and thus that of the composite indicator) is driven by chance (Berg and Pattillo, 1999). Moreover, due to the rare occurrence of crises, out-of-sample analysis is strongly limited. Thus, it is even more important to assess the in-sample performance of the signals approach properly in order to guarantee that the results reflect a relationship between indicators and crises that can be assumed to persist in the future.

In the current paper, we overcome this drawback and show that the in-sample predictive power of the signals approach is actually highly significant for several types of crisis in varying samples.

2.2 The signals approach

The signals approach builds on the observation that there is a risk-free range of fluctuation for most macroeconomic time series around their long-run equilibrium, but that strong deviations from that average often indicate looming crises.

Let C denote a binary crisis variable,² which equals one in crisis periods, and \bar{C} be a corresponding matrix indicating the periods where indicators should issue a signal (warning variable), that is:

$$\bar{C}_{t,n} = \begin{cases} 1, & \text{if } \exists k \in \{0, \dots, h\} : C_{t+k,n} = 1 \\ 0, & \text{else} \end{cases}, \quad (1)$$

where $t \in \{1, \dots, T\}$ is the time index, $n \in \{1, \dots, N\}$ the country index, and $h \in \mathbb{N}$ the early warning horizon inside which a signal should be sent. The early warning horizon is chosen by economic reasoning, taking into account the type of crisis and the time that is necessary for any action taken to have an effect on the economy. Since it is assumed that the relationship between the warning variable \bar{C} and the indicator I is extremely nonlinear, the signals approach aims to identify a threshold Q_I for every indicator I where a sudden increase in crisis probability is observed.

Using the threshold,³ every indicator is translated into a binary signal S_I :

$$S_{I,t,n} = 1_{I_{t,n} \geq Q_I}. \quad (2)$$

The combination of \bar{C} and the signaling variable S_I results in a division of the observations into four subsets A , B , C , and D as described in Table 1 for every indicator.⁴

²We generally denote a stochastic process by X and its value taken in time t in country n by $X_{t,n}$.

³The equation describes the standard application of the signals approach where the thresholds are constant. There are, however, applications where the threshold varies over time, across countries or both.

⁴Usually there is a fifth set containing excluded observations. These are periods directly after a crisis when indicator values have not yet returned to normal behavior, or periods at the end of the sample when it is unknown if a crisis will follow.

	$\bar{C}_{t,n}$	1	0
$S_{I,t,n}$		A	B
	1	A	B
	0	C	D

Table 1: Combination of crisis and signaling variable. The letters A , B , C , and D denote both the sets themselves and the number of elements in the respective sets.

From this division different quality measures can be derived. Most common are the noise-to-signal ratio (Kaminsky and Reinhart, 1999) and a utility function based on Type I and Type II errors (Bussière and Fratzscher, 2008; Alessi and Detken, 2011):

$$NSR = \frac{B}{B+D} \bigg/ \frac{A}{A+C}, \quad (3)$$

$$U(\theta) = \min(\theta, 1 - \theta) - \theta \frac{C}{A+C} - (1 - \theta) \frac{B}{B+D}. \quad (4)$$

The first papers on the signals approach use the noise-to-signal ratio (NSR) to assess the indicators' quality, but more recent papers employ utility. The advantage of utility optimization over NSR optimization is that utility optimization clearly defines a marginal rate of substitution between between Type I and Type II errors. Since the preferred marginal rate of substitution remains unrestricted in NSR optimization, it sometimes produces a corner solution where only noise is minimized, almost ignoring the failure in signaling, as shown in Figure A.1. The additional parameter $\theta \in (0, 1)$ in equation (4) captures decision-makers' preference for avoiding either Type I or Type II errors (Bussière and Fratzscher, 2008; Alessi and Detken, 2011). The threshold Q_I is set either to minimize the NSR or to maximize utility. An indicator with $NSR > 1$ is ignored, since a better NSR can be achieved by choosing the lowest possible value as a threshold and letting the indicator send a signal in every period. An indicator with $U(\theta) < 0$ is ignored for similar reasons.

While the quality measure of an individual indicator assesses how important instabilities in a particular area of the economy are, policymakers are frequently interested in the overall assessment of the current risk in the economy, which can be achieved by combining individual signals into a composite indicator $CI = \sum_{m=1}^M \omega_m S_{I_m}$, as proposed by Kaminsky (1999), where M refers to the number of indicators.

In the literature, a number of different weighting schemes for w_m have been employed; most notably, equal weights, normalized inverse noise-to-signal ratios and utilities.⁵ The resulting composite indicator can be treated (almost) as if it were a normal indicator. There is, however, one caveat. Because the composite indicator is a linear combination of only few binary variables, it is discrete rather than continuous. Hence, it is often impossible to distinguish between different thresholds that are close to each other. Although an optimum threshold could be derived, it is more common to set a threshold arbitrarily close to 0.4 or to provide conditional crisis probabilities for particular ranges of the composite indicator.

2.3 Three contributions on the signals approach

We add an evaluation of statistical significance to three previous contributions that cover a broad range of possible applications of the signals approach. They capture both different types of crises and countries, showing that our approach leads to a general reassessment of the signals approach. Below, we briefly summarize the data used in those papers. For more information, refer to Table A.5 in the annex.

The first paper we analyze is that by Kaminsky and Reinhart (1999), where the signals approach is first proposed and applied to a total of 76 currency and 26 banking crises in 15 developing and 5 industrial countries between 1970 and 1995. Currency crises are identified by an index of currency market turbulence, while banking crises are identified qualitatively by means of an assessment of public support of the banking sector. Their study considers only the beginning of the crises, thereby limiting the length of a crisis (as considered for prediction) to one month. According to Kaminsky and Reinhart, currency crises should be indicated starting 24 months before the crisis,

⁵There are also some measures that try to capture either the duration or the severity of a signal (Kaminsky, 1999; Brüggemann and Linne, 2002). However, this partly removes the key advantage of the signals approach, namely that it does not rely on a specific model.

up to the first month of the crisis. Banking crises should be indicated starting 12 months before the crisis and ending 12 months after the crisis. While there is limited cross-country correlation, the crises are scattered over time and regions. 16 different monthly indicators capturing monetary stress, vulnerability of the banking sector, prices and competitiveness are employed as potential indicators. The paper has been reproduced several times (for example by Kaminsky et al. (1998); Edison (2003); Peng and Bajona (2008)) with wider datasets. We use the dataset prolonged until June 2003, which is publicly available for Kaminsky (2006), where only currency crises are examined. This dataset contains 112 currency and 41 banking crises. Following Kaminsky (2006), we employ banking crises as an additional indicator for currency crises and extend this principle by also using currency crises as a possible indicator for banking crises in our application.

The second paper, by Alessi and Detken (2011), applies the signals approach to costly asset price booms in 18 OECD countries. Asset price booms are identified by high growth of asset prices in at least four consecutive quarters. High- and low-cost booms are distinguished by using deviations from potential growth in the years following the boom. The authors identify 29 high-cost and 16 low-cost booms in the time from 1970 to 2007. These booms occur in three waves, and consist mainly of high-cost booms in the last two of these waves. An additional wave – the recent financial crisis – at the end of their sample is identified, but is not used in the analysis because at the time of writing it was unclear whether the boom before the financial crisis would have to be classified as a high- or low-cost boom. Because Alessi and Detken predict an asset price boom rather than the subsequent drop in GDP, and political action might still be taken after the boom has started, they do not only consider the beginning of the boom but also its first few quarters. Thus, they allow for the possibility that early warning signals can be issued both in the six quarters before the event (i.e. the boom) and the first three boom quarters. For the prediction of one of those quarters, 89 indicators constructed from different transformations of 18 underlying quarterly indicators are used. Because of the confidentiality of equity, housing and aggregate asset prices provided by the BIS to Alessi and Detken, we are only able to use 15 of these 18 variables. Since some transformations are based on multivariate systems, we only have a total of 50 instead of 89 different indicators.⁶ Most of the 15 underlying indicators are available from the OECD Economic Outlook and Main Economic Indicators; domestic credit can be found in the IMF’s International Financial Statistics (IFS); private credit, corrected for structural breaks, was provided by the authors.

The third paper, by Knedlik and von Schweinitz (2011), is on the European debt crisis. It analyzes 11 countries in the European Monetary Union (EMU), from the introduction of the Euro until November 2011. A public debt crisis is defined using the spreads of government bond yields over the yield of the average AAA-rated country in the EMU. Using this criterion, a crisis is identified in 5 countries at the end of the observation period. Like Kaminsky and Reinhart (1999), Knedlik and von Schweinitz only consider the outbreak of a crisis, using an early warning horizon of 24 months. However, in contrast to Kaminsky and Reinhart (1999), the crises Knedlik and von Schweinitz identify are not scattered over time. A total of 20 monthly indicators out of 5 different categories (fiscal indicators, competitiveness and domestic demand, asset prices, labor, private and foreign debt) are used. All of them are publicly available from EuroStat, ECB, the OECD Main Economic Indicators, the IMF’s IFS and Morgan Stanley Capital International (MSCI).

3 Dealing with uncertainty in the signals approach

The quality of the signals approach is hard to judge from an econometric perspective. Because thresholds are adjusted to create an optimal fit of individual indicators, even randomly chosen indicators that are not causally related to the crisis probability may produce considerable in-sample fit by chance.

Whether or not a set of indicators performs well in-sample albeit having no predictive power, strongly depends on the data-generating process underlying the chosen indicators. For example, it is unlikely that a white noise process or (more generally) a strongly mean reverting process produces high utility by chance, since the signals approach usually requires a good indicator to send a signal for a number of consecutive periods. To account for the data structure, we thus propose a bootstrap to evaluate the distribution of the performance under the null hypothesis that the indicators do not truly explain crises. To do so, we can either resample the indicators and match them with the original crisis data, or resample the occurrence of crises, matching the

⁶We also do not consider their “global” indicators, as these indicators do not differentiate between single countries.

counterfactual crises with the original indicator data. Depending on the structure of the indicator data and the crisis, one of these approaches might be more feasible than the other.

Because Kaminsky and Reinhart (1999) mostly cover emerging markets that display highly volatile macroeconomic indicators, it is hard to find a statistical model that can reproduce the original indicator structure. Thus, it seems more appropriate to resample the occurrence of crises. To capture the cross-country correlations and dynamics of crisis propagation, we propose an adapted moving block bootstrap (Künsch (1989); Fitzenberger (1997)) for this purpose, as described in Subsection 3.1.1.

Alessi and Detken (2011), who cover an OECD sample, use indicators that are more stable than those used by Kaminsky and Reinhart, but their sample includes a high number of indicators that are missing for extended periods at the beginning and end of the sample in some countries. This makes the estimation of cross-country and cross-indicator correlations unfeasible. Thus, we again propose to resort to resampling the crises. However, this requires a different approach from that indicated above, since crises continue over several periods. A moving block bootstrap would produce more crises that begin or end in the same period in a large number of countries than found in the true data, thereby causing inappropriately high cross-country correlation. Instead, we develop a conditional probability bootstrap, as described in Subsection 3.1.2.

Opposed to the applications described above, Knedlik and von Schweinitz (2011) use a sample where crisis resampling is inappropriate. Their data only covers a single block of crises close to the end of the sample. However, because their sample only covers highly developed members of the European Union, most data are easily available for the entire sample and the data-generating process of the indicators is relatively stable. We are thus able to resample the indicators rather than the crises, again providing counterfactual simulations where indicators and crises are independent, while still retaining similar statistical features concerning cross-country correlations and dynamics. We use the econometric model described in Subsection 3.2.1.⁷

We always use 10 000 bootstrap samples of either crisis or indicator data that are subsequently used with the remaining original data (i.e. either original indicators or crises) in the signals approach. Both the thresholds of individual indicators and the weights of the individual indicators in the composite indicator are determined separately for each counterfactual sample, producing a set of 10 000 quality measures (utility or noise to signal ratio) under the null hypothesis. This results in a simulated probability distribution for the quality measures that is then used to assess the significance of the signals approach.

3.1 Resampling crises

Our resampling aims to preserve the unconditional crisis probability in the whole sample, and the crisis probability conditional on current and recent crises in all countries. Our first resampling method achieves this objective by roughly following a conventional moving blocks bootstrap. However, since we do not want to carry over country-specific crisis probabilities, we add a second layer of resampling. The second resampling method relies on a set of conditional probabilities obtained from the true data to resample counterfactual crises with similar statistical properties.

3.1.1 A moving block bootstrap

Let us recall our binary crisis $T \times N$ matrix C where

$$C_{t,n} = \begin{cases} 0, & \text{if there is no crisis in country } n \text{ in period } t \\ 1, & \text{if there is a crisis in country } n \text{ in period } t \end{cases}. \quad (5)$$

We generate $T+1-b$ blocks B of consecutive rows of blocklength b from our matrix C , resulting in a sequence of blocks defined by:

$$B_t = \begin{bmatrix} C_{t,1} & \dots & C_{t,N} \\ \vdots & \ddots & \\ C_{t+b-1,1} & & C_{t+b-1,N} \end{bmatrix}. \quad (6)$$

⁷Since we are resampling indicators, one could argue that we do not capture the true model (in terms of parameter estimates and included variables). However, for our purpose, this is of limited importance. We are more interested in the true statistical properties when we want to test our null hypothesis of no correlation between crises and indicators.

Resampling is achieved by randomly drawing a number of blocks (with replacement) that is sufficient⁸ to create a new counterfactual version of matrix C , denoted by \hat{C} . We then add a second layer of resampling, drawing N columns from \hat{C} (with replacement), to create a final bootstrap crisis sample C^* .

In the application to currency and banking crises, we choose a blocklength b of 12 months.

3.1.2 The conditional probability bootstrap

Starting from our binary crisis matrix C , as defined in equation (5), we can compute the probability of a crisis in any country at time t , conditional on the occurrence of a crisis in this country at time $t - 1$, and the number of countries being in the state of crisis in period $t - 1$. Since we are more interested in the statistical properties of our resampled matrix C , rather than the question of which countries play a specific role in the transmission of crises, we assume that these probabilities are identical across countries.⁹ That is, we have a set of $2(N + 1)$ probabilities defined by:

$$p_{\chi,n} = p(C_{t,i} = 1 | C_{t-1,i} = \chi \wedge \sum_{j=1}^N C_{t-1,j} = n), \quad (7)$$

where χ is 0 or 1 and n is an integer between 0 and N , capturing the number of countries in the state of crisis in the previous period. If $p_{\chi,n}$ cannot be defined from the sample for a certain n , we replace $p_{\chi,n}$ by $p_{\chi,\tilde{n}}$, \tilde{n} being the number of crises as close to n as possible, where a probability $p_{\chi,\tilde{n}}$ can be computed from the sample. Situations where this replacement is necessary occur if the simulation produces a number of crises in a single period that have never been observed in the data.

Each repetition of the bootstrap is initialized using a random row from C as the first period of the resampled crisis matrix. Starting from this counterfactual crisis situation in the first period, the following periods are randomly generated, using the probabilities defined in equation (7).¹⁰ As above, this results in a bootstrap crisis sample C^* .

3.2 Resampling the indicators

3.2.1 The panel VAR bootstrap

We assume that every indicator I_m , $m \in \{1, \dots, M\}$ has a global and a national component. Both follow an individual VAR process.

Roughly following the model structure that underlies the panel unit root test accounting for cross sectional dependencies proposed by Bai and Ng (2005), we define

$$f_t = [f_{1,t} \quad f_{2,t} \quad \dots \quad f_{M,t}], \quad (8)$$

where $f_{m,t}$ is the first principal component of the matrix of indicator m :

$$I_m = \begin{bmatrix} I_{m,1,1} & \dots & I_{m,1,N} \\ \vdots & \ddots & \vdots \\ I_{m,T,1} & & I_{m,T,N} \end{bmatrix} \quad (9)$$

at time t .

Defining a vector of values of all indicators at a given time t in country n ,

$$I_{t,n} = [I_{1,t,n} \quad I_{2,t,n} \quad \dots \quad I_{M,t,n}], \quad (10)$$

we can then write:

⁸To be precise, $\lceil T/b \rceil$ blocks are needed. Surplus periods are cut.

⁹This assumption should capture the main dynamics if countries are highly integrated. If, however, bilateral financial or economic ties between countries prevail, we observe repetitive patterns of certain country groups that experience crises jointly. In these situations, our simplification might prove unfeasible.

¹⁰Because the sample used by Alessi and Detken (2011) does not include crises of less than four periods, we slightly adjust the simulation to produce corresponding results. If a counterfactual crisis would end after one period, the crisis is immediately removed from the sample. Crises that continue over three periods are prolonged to cover a fourth period without randomization. Crises that continue for two periods and are about to end from the random drawing, are either removed from the sample (with probability q) or extended to the following two periods (with a probability of $1 - q$), thereby overwriting the original random draw. The probability q is calibrated to produce an unconditional crisis probability that matches the true data.

$$I_{t,n} = f_t P_n + \eta_{t,n}, \quad (11)$$

where P_n is a diagonal matrix produced from the individual inverse loading matrices of the principal components decompositions and $\eta_{t,n}$ is the vector of national contributions to the indicator vector.

Both the global and individual components are assumed to follow a VAR process, i.e.

$$f_t = L(A)f_t + v_{t,n} \quad (12)$$

$$\eta_{t,n} = L(B)\eta_{t,n} + c_n + \varepsilon_{t,n}, \quad (13)$$

where c_n capture national effects; $v_{t,n}$ and $\varepsilon_{t,n}$ are the respective errors.

The national processes are estimated using a panel VAR. Since $T > 30$ for all examples, we can forgo GMM estimation and employ OLS, as shown by Maddala and Wu (1999). Their conclusion, originally for the application in single equations, can be directly applied to this VAR because the VAR can be estimated blockwise.

Lag orders of the global VAR and the national level panel VAR are determined separately using the Schwartz criterion. The residuals used when resampling are drawn from all the residuals, because we want to exclude cross-country heteroscedacity. We simulate $T + 100$ periods of data in each iteration, thereby allowing to discard the first 100 periods to get rid of the impact of the starting values. This results in a bootstrap indicator sample I^* , containing all indicators, which is used together with the original crisis matrix C in order to test our null hypothesis.

3.2.2 Evaluation techniques

The above described methodology of bootstrapping crisis dates and indicators allows us to test the significance of the signals approach. To do this we rely on three criteria. First, we test the null hypothesis that the utility or noise-to-signal ratio for individual indicators could have been achieved by chance. Second, we employ a Fisher test (based on bootstrapped distribution functions)¹¹ to test the null hypothesis that, in a set of indicators, the utility or noise-to-signal ratio of all the individual indicators could have been the result of a random process. Third, we test for the null hypothesis that a composite indicator (based on the noise-to-signal ratio or utility optimization) of the whole set of indicators in one application could have been the result of a random process.

4 Results

The discussion of the results is divided into four subsections. In the first three, we reevaluate the economic findings of Kaminsky and Reinhart (1999), Alessi and Detken (2011) and Knedlik and von Schweinitz (2011) and assess their statistical significance, concentrating on specific findings. Against this background, we then present general results on the application of the signals approach in the fourth subsection. The results of the different tests for four applications are presented in Tables A.1 to A.4 in the annex.¹²

4.1 Currency and banking crises (Kaminsky and Reinhart, 1999)

Our results concerning the application of the signals approach to currency and banking crises are shown in Tables A.1 and A.2 in the annex. Unlike Kaminsky and Reinhart (1999), who use NSR optimization in their paper, we show results for both utility and NSR optimization.

Our findings mostly confirm the results reported by Kaminsky and Reinhart (1999). The seven best performing indicators in the application to currency crises found by Kaminsky and Reinhart (1999) are the real exchange rate, stock prices, exports, output, the deficit (% GDP), M2 (% reserves) and reserves. The same indicators are also among the best and most significant when using utility optimization. However, when we apply NSR optimization as Kaminsky and Reinhart do, we find neither deficit nor output to be significant at the five percent level. Deficit and output also have a considerably higher NSR than in the original paper. However, other indicators improve

¹¹The original test (Fisher, 1932) has been rediscovered by Maddala and Wu (1999), who proposed the application to panel unit-root testing.

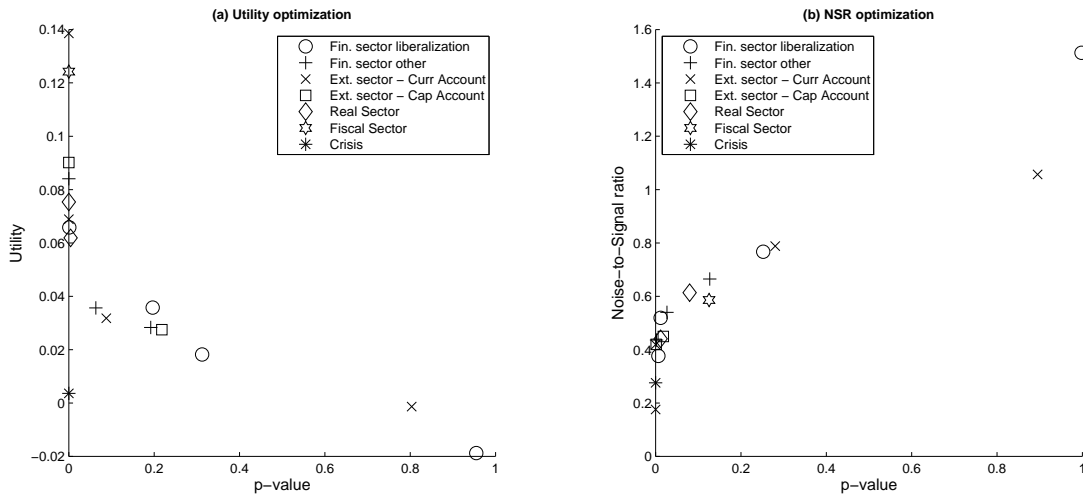
¹²The utilities are normally calculated for $\theta = 0.5$ and – for the composite indicators – with a threshold of 0.4. Only for the results of Alessi and Detken (2011), we choose $\theta = 0.4$ in line with these authors. The row of the “Fisher test statistic” contains the value of the Fisher test and its p-value for the two optimization methods.

substantially compared to the original paper, for example, real interest rates and the real interest rate differential are highly significant, both displaying a substantially lower NSR than in the original sample, which ends in 1995.¹³

As noted above, we also include banking crises as a possible indicator. Although the utility of banking crisis is close to zero, it is highly significant, as can also be seen from the outlier in Figure 1(a). This low utility is mostly due to the rare occurrence of banking crises in the sample. If there were as many banking crises as currency crises, the ratio of potential signals to required signals (24 periods before the crisis and the crisis itself) would be $1/25$. That is, even if there was no noise at all, but every upcoming currency crisis was announced by a banking crisis, the utility would still only be $0.02 = 0.5 - 0.5 \left(\frac{24}{25} + 0 \right)$. Even though banking crises have low utility as an indicator, this result (especially the significance of the indicator) can be seen as a confirmation of the findings of Kaminsky and Reinhart, who argue that banking crises frequently cause currency crises. Nevertheless, banking crises would only play a negligible role in a composite indicator using utility optimization.

Since utility optimization seems to be superior to NSR optimization as described in Section 2.2, it might be a promising approach to include the significance of indicators in the weighting scheme of a composite indicator in order to overcome the above mentioned problem. In addition, a moving average of indicators might be used in the composite indicator in order to compensate for the very short duration of signals sent by the banking crisis indicator.

Even without this possible improvement, the composite indicator obtained by both optimization schemes is better (in terms of the same measure as used in the scheme) than all the individual indicators, and is highly significant, as shown in Figure 2. We see that the same does not necessarily hold true when looking at the respective other measure: the utility of the composite indicator obtained by NSR optimization is both low and insignificant. This is explained by the high thresholds of individual indicators' leading to low signal ratios (and correspondingly high Type I error probabilities).¹⁴

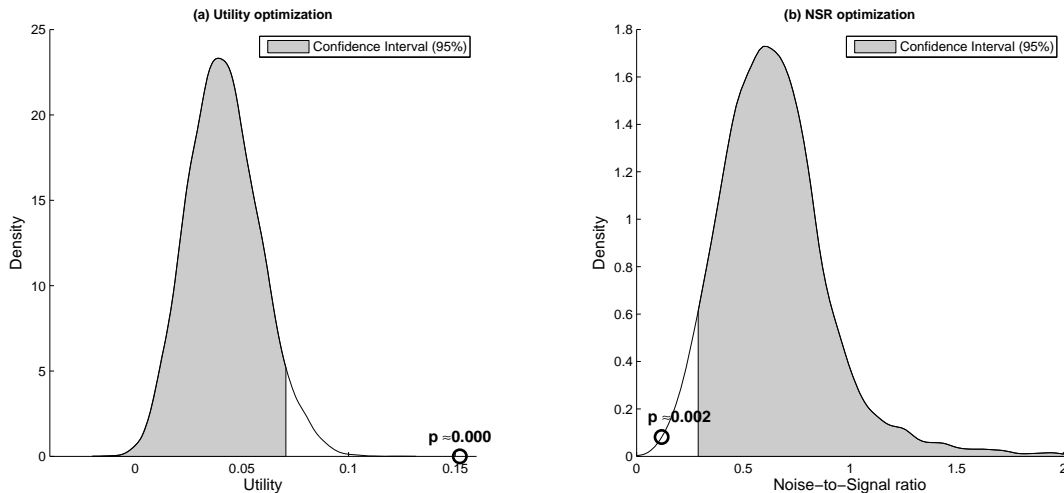


Note: The 16 indicators are grouped according to Kaminsky and Reinhart.

Figure 1: Significance and quality measure of individual indicators in the currency crises application

¹³The difference in results between the original paper and our recreation is only partly due to our extension of the sample. Even when we apply the signals approach to the original sample (ending in 1995), we find slightly different results, probably due to revisions included in the data.

¹⁴See also Figure A.1(b) in the annex.



Note: The density is derived under the null hypothesis, the circle indicates the respective quality measure of the original sample.

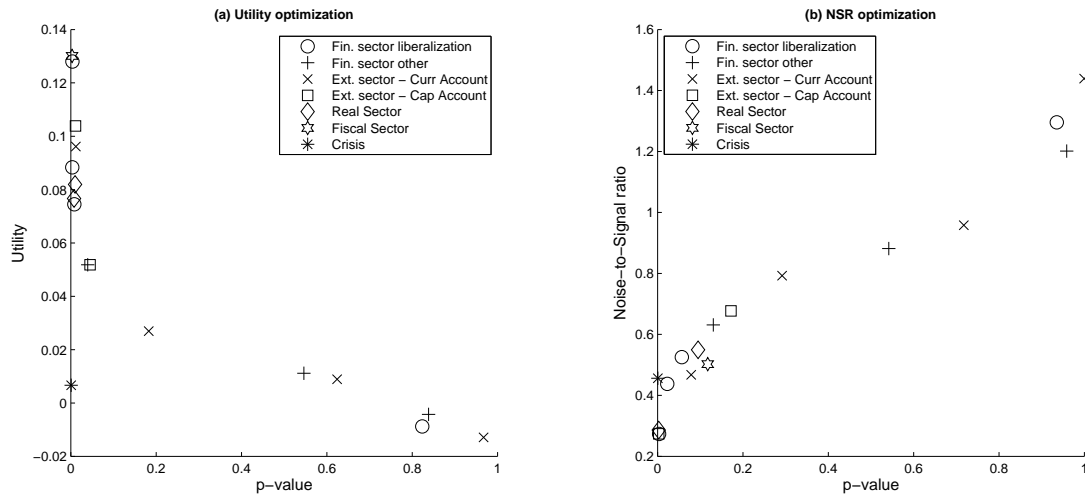
Figure 2: Density of utility and NSR in the currency crises application

In the utility optimization in the case of banking crises, the three indicators that perform best and are the most significant for both sample sizes are the deficit (% GDP), the real interest rate and the real interest rate differential. In the NSR optimization, however, the best indicator from the original sample period (real exchange rate) performs comparably poor, while the real interest rate and the real interest rate differential become the two indicators that perform best when the full sample (ending in 2003) is used. This is partly explained by their good performance in the Asian crisis, which is only included for out-of-sample analysis by Kaminsky and Reinhart (1999).

Currency crises seem to “predict” banking crises about as well as banking crises predict currency crises. However, this is mostly due to the fact that the “early warning” horizon continues until 12 months after the crisis breakout in the banking crisis application.¹⁵ Generally, we find a substantially different distribution of p-values of the considered indicators when we look at utility and NSR optimization respectively. If we divide the indicators into groups, “good” (p-value smaller than 5%), “mediocre” (p-value between 5% and 20%) and “bad” (p-value above 20%), the utility optimization results in a bipartition, as shown in Figure 3(a). A large group is good (11/17), a smaller one “bad” (5/17), while only exports belong to the “mediocre” group, with a p-value around 18%. By contrast, the NSR optimization yields three groups of nearly equal size. This, combined with the fact that the indicators in the “mediocre” group have only average NSR and a combined weight of 37% in the composite indicator, leads to the insignificance of the composite indicator using NSR optimization (and a worse NSR than that of its best components) as presented in Figure 4(b). The same result holds for the utility of the composite indicator in the NSR optimization. Real interest rates, real interest rate differentials and stock prices produce better NSRs than the composite indicator does when forecasting banking crises. However, the utility of the utility-optimized composite indicator is better than that of the best individual indicator by itself.

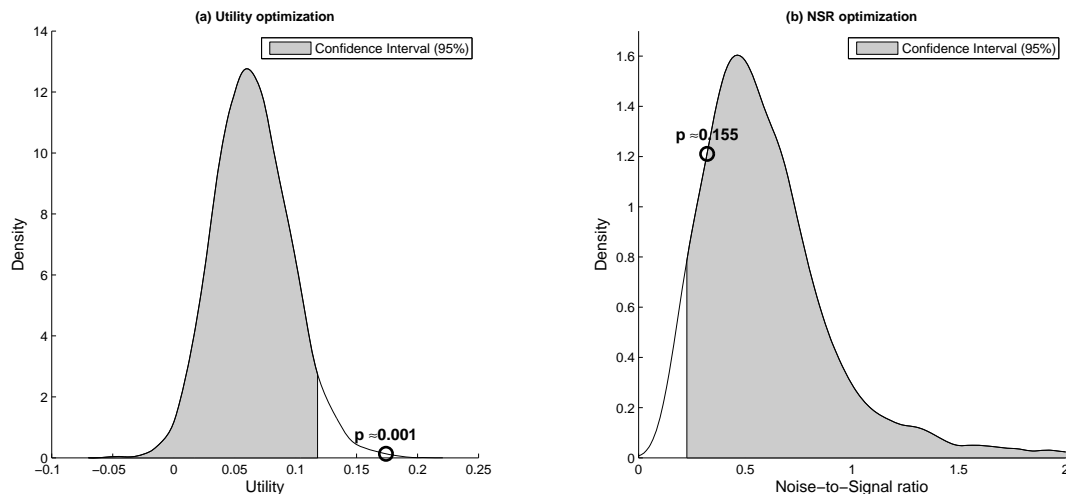
Given the large number of individual indicators in the applications to currency and banking crises that are highly significant, it is hardly surprising that the Fisher test strongly rejects the null hypothesis that no indicators are truly correlated to crises.

¹⁵While 24 banking crises announce upcoming currency crises, only ten currency crises actually precede banking crises within a 24 month window. However, 21 currency crises occur in a 25 month window centered at the outbreak of a banking crisis.



Note: The 16 indicators are grouped according to Kaminsky and Reinhart.

Figure 3: Significance and quality measure of individual indicators in the banking crisis application



Note: The density is derived under the null hypothesis, the circle indicates the respective quality measure of the original sample.

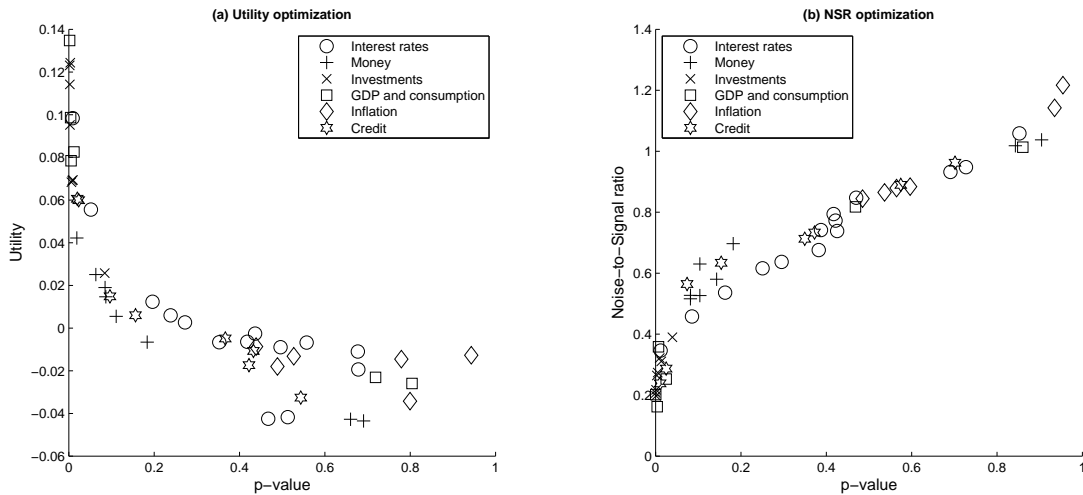
Figure 4: Density of utility and NSR in the banking crisis application

4.2 Costly asset price booms (Alessi and Detken, 2011)

One of the most interesting features of the study by Alessi and Detken (2011) is that the authors do not select one preferred transformation for each macroeconomic variable, but use every variable in a number of different transformations, producing groups of highly related indicators. Some groups of indicators have near-equal performance (e.g. investment and consumption), while others display a large variety of performance measures and p-values (e.g. money indicators such as M3), as can be seen in Figure 5. As in the original paper, we find a large number of indicators that perform poorly. However, their low utility (and high NSR) is partly due to the long early warning horizon. This also has an effect on the significance of the individual indicators: using the classification of significance introduced in Section 4.1, 27 out of 50 indicators are “bad” for both optimization rules, while eight of them are “mediocre”.

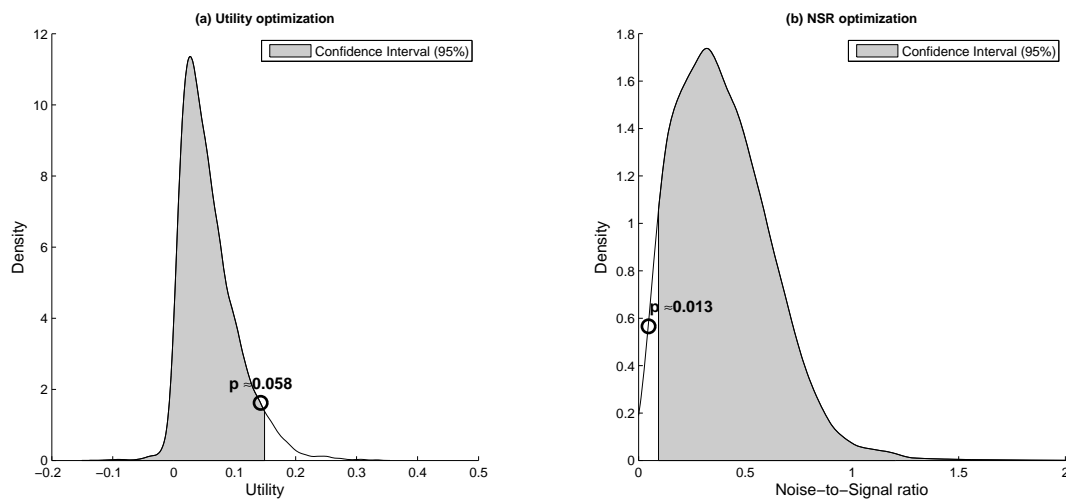
Using different transformations of a single variable increases the number of indicators, and

hence the number of significant indicators, substantially, but it adds little to the significance of the whole set in the Fisher test and to the significance of the composite indicators. This is due to two characteristics of the indicator set. In the large group of insignificant indicators, many have a utility below zero (25 indicators) or an NSR near or above one (19 indicators have an NSR above 0.8). Thus, the composite indicator is constructed from relatively few individual indicators: Transformations of housing investment, investment and the real GDP have a share of 72% and 62% in the utility and NSR-optimized composite indicator, respectively. This concentration on the correlated transformation of only three individual indicators explains why, despite the large number of indicators, the Fisher test and composite indicators are not significant at the one percent level (compare Figure 6).



Note: The 50 indicators have been grouped to enhance readability.

Figure 5: Significance and quality measure of individual indicators in the costly asset price boom application



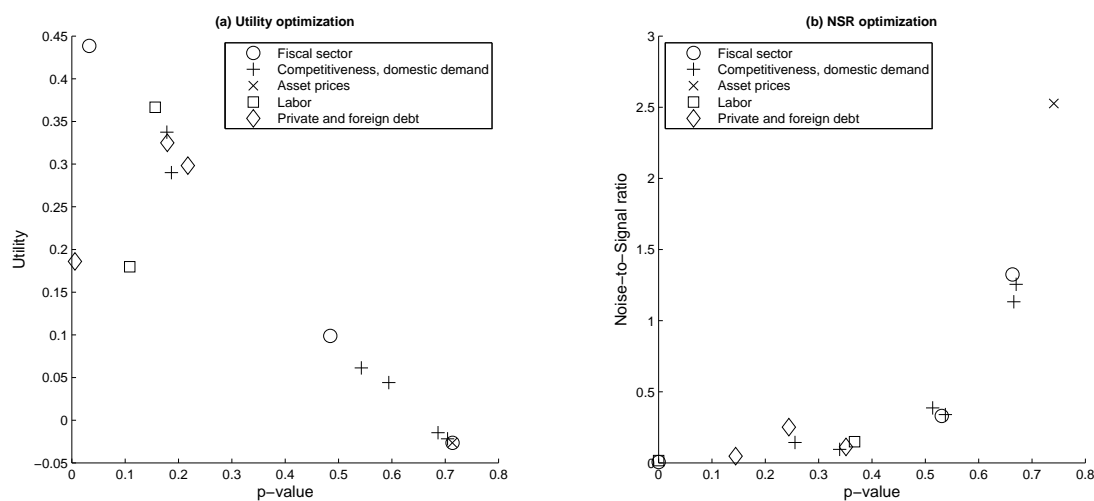
Note: The density is derived under the null hypothesis, the circle indicates the respective quality measure of the original sample.

Figure 6: Density of utility and NSR in the costly asset price boom application

4.3 Public debt crises in Europe (Knedlik and von Schweinitz, 2011)

The results of the fourth application – public debt crises in a monetary union – are shown in Table A.4. We were unable to exercise the bootstrap for five out of the original 20 indicators, due to the low number of data points available for these variables, as all indicators need to be available at all times in all countries to estimate the the VAR. From the remaining indicators, in every optimization scheme, only two are significant. The group of “slightly insignificant” indicators consists of five indicators in the case of utility optimization, and one indicator in the case of the NSR optimization, while all the other indicators are “bad”, as can also be seen in Figure 7. The most significant indicator in the utility optimization, namely foreign assets, has a rather low utility, which is explained by the persistence of that indicator. In our dataset, all countries have either a positive (countries without a crisis) or a negative (countries with a crisis) foreign asset position over almost the entire period. Thus, even if there are many false signals over the first part of the observation sample, this indicator can hardly yield better utilities in a bootstrap capturing this characteristic. This argument does not hold true for the NSR, where a low signal ratio in a resampling can easily be offset by a low noise ratio.

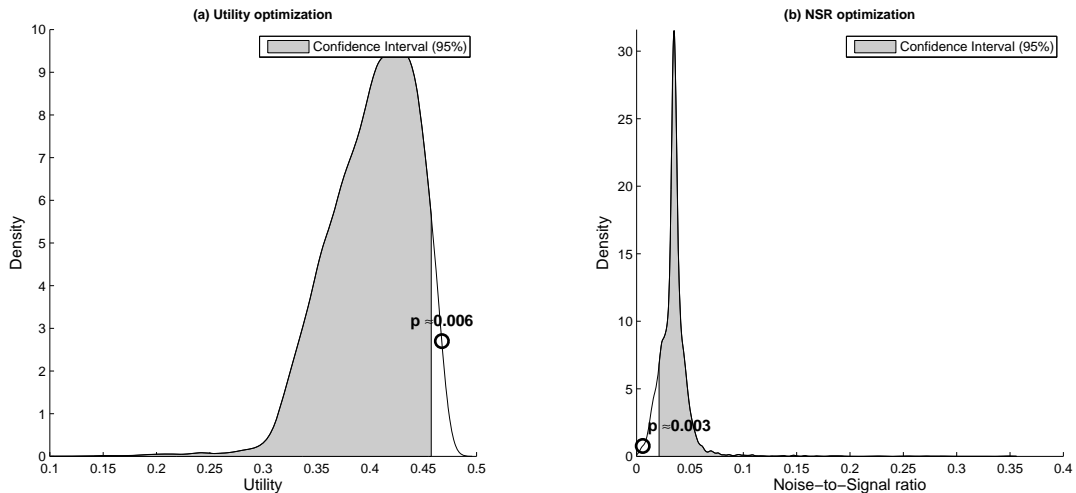
As the specific p-values of the significant indicators are the main components of the Fisher test statistic, we observe that the null hypothesis of jointly randomly distributed indicators can be strongly rejected for the NSR optimization, because of government deficit and the unemployment rate, whose significance is undistinguishable from 0. By contrast, this is not possible for the utility optimization, due to more evenly distributed p-values. However, the composite indicator in both cases is extremely good (bearing in mind that a utility of 0.5 and NSR of 0 are the best results that can be achieved) and highly significant. The relation between the composite indicator and the subsequent crisis seems to be especially strong in the case of the NSR, where the bootstrap (meaning: randomness) resulted in 55% of the repetitions in an NSR of infinity. Thus the composite indicator, if it were random, would have been unlikely to have sent a signal¹⁶ in the 24 months before the crises concerned.



Note: The 16 indicators are grouped according to Knedlik and von Schweinitz.

Figure 7: Significance and quality measure of individual indicators in the debt crisis application

¹⁶The composite indicator sends a signal if it exceeds the threshold of 0.4.



Note: The density is derived under the null hypothesis, the circle indicates the respective quality measure of the original sample.

Figure 8: Density of utility and NSR in the debt crisis application

4.4 General results

First, our results show that the key findings of the three applications of the signals approach that we recreated in this paper are significant. This applies to both the general predictive power of the approach, for example, through a composite indicator, and the most important individual indicators (compare the last row “Composite Indicator” in Tables A.1 to A.4 and Figures 2, 4, 6 and 8).

Considering the four applications in two versions each (utility maximization and the noise-to-signal ratio minimization) we find that in five out of eight cases the composite indicator is significant at a one percent error probability level, two further versions are significant at a ten percent level. In only one case (the banking crisis application using the noise-to-signal ratio minimization) do we have to accept a higher error probability of 17.3%. The highly significant results of most of the composite indicators show that the signals approach can provide a major contribution to early warning systems.

Second, we find that using composite indicators adds substantial value to the signals approach, especially if a large number of truly different indicators is used. Even if there is little information in the individual indicators, the composite indicator can be a good crisis predictor, since the simultaneous movement of a large number of economic indicators provides additional information that can be exploited. Knedlik and von Schweinitz (2011) provide an excellent composite indicator, although the null hypothesis that no individual indicator is correlated to crises cannot be rejected using a Fisher test. Similarly, although this result is not quite as strong, we find the utility of the composite indicator to exceed the utility of the best individual indicator in Kaminsky and Reinhart (1999). Only in Alessi and Detken (2011) is the added value of the composite indicator rather limited. However, this is mostly due to the fact that they focus on finding the best individual indicator by trying a large number of transformations of possibly important variables, rather than testing a large number of variables that provide truly different information on the economy. Thus, in their application, the composite indicator aggregates indicators that provide similar information, and cannot rely on untypical joint movements of several indicators.

Third, while the best indicators in each considered application are highly significant, there is no perfect correlation between the chosen quality measure (NSR or utility) and significance. Due to different statistical properties of the indicators, the extent of in-sample fit that can be produced by chance differs widely. This becomes clear when one looks at the quality measures and significance levels in Tables A.1 to A.4. To demonstrate this point, we show that the cumulative distribution functions (CDFs) for the performance of the individual indicators under the null hypothesis may cross, as shown, for example, in Figure 9 for three indicators from Alessi and Detken (2011).¹⁷

¹⁷Actually, we show the counter cumulative distribution functions because they show directly the declining p-value

Since the CDFs are not identical even in the same application, the sequential arrangement of utility (or NSR) and significance is not necessarily the same. Comparing the applications with each other provides even stronger evidence of the fact that we cannot define a generally valid threshold for a “good” indicator. For example, while Knedlik and von Schweinitz (2011) clearly have the indicators with the highest utility, most of them are not significant or are barely significant.

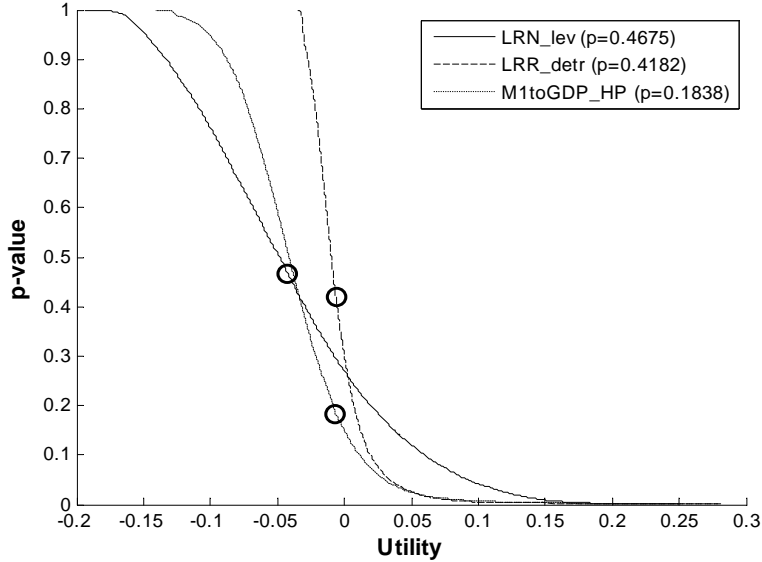


Figure 9: Utility and significance for three different indicators. Circles show the real values.

5 Conclusions

We augment four previous applications of the signals approach by adding an evaluation of significance based on a bootstrap approach. For all applications considered, we find that the major findings are indeed significant. This justifies a general reassessment of the usefulness of the signals approach. The signals approach holds a number of advantages and features favorable characteristics for early warning systems compared to alternative approaches. In particular, the signals approaches’ ability to provide both a composite crisis indicator and a more detailed picture when one looks at the individual indicators enables policy makers to get an impression of general risk and to identify areas where there is a need for action to be taken.

For those interested in an early warning system rather than in just the best individual indicator, we propose using a broad set of indicators, covering a wide range of economic issues, to create a composite indicator based on utility optimization. The results achieved in this study when using this method produce composite indicators that are stable over time and add substantial forecasting power to the best individual indicator, even if the performance of many individual indicators is rather poor.

Acknowledgements

The authors are indebted to Andrew Hughes Hallet and Carsten Detken for their valuable comments and suggestions. We also thank Katja Drechsel and Herbert Buscher for comments and discussions.

for increasing utility.

References

References

- Abiad, A., 2003. Early Warning Systems: a Survey and a Regime-Switching Approach. IMF Working Paper 32.
- Alessi, L., Detken, C., 2011. Quasi Real Time Early Warning Indicators for Costly Asset Price Boom/Bust Cycles: A Role for Global Liquidity. *European Journal of Political Economy* 27 (3), 520–533.
- Bai, J., Ng, S., 2005. Tests for Skewness, Kurtosis, and Normality for Time Series Data. *Journal of Business and Economic Statistics* 23 (1), 49–60.
- Berg, A., Pattillo, C., 1999. Predicting Currency Crises: The Indicators Approach and an Alternative. *Journal of International Money and Finance* 18 (4), 561–586.
- Borio, C., Drehmann, M., March 2009. Assessing the Risk of Banking Crises—Revisited. *BIS Quarterly Review*, 1–18.
- Brüggemann, A., 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 157.
- Bussière, M., Fratzscher, M., 2006. Towards a New Early Warning System of Financial Crises. *Journal of International Money and Finance* 25 (6), 953–973.
- Bussière, M., Fratzscher, M., 2008. Low Probability, High Impact: Policy Making and Extreme Events. *Journal of Policy Modeling* 30 (1), 111–121.
- Diebold, F. X., Lee, J.-H., Weinbach, G. C., 1994. Regime Switching with Time-Varying Transition Probabilities. In: Hargreaves, C. (Ed.), *Nonstationary Time Series Analysis and Cointegration*. Advanced Texts in Econometrics. Oxford University Press., pp. 283–302.
- Edison, H., 2003. Do Indicators of Financial Crises Work? An Evaluation of an Early Warning System. *International Journal of Finance and Economics* 8 (1), 11–53.
- Filardo, A., 1994. Business Cycle Phases and Their Transitional Dynamics. *Journal of Business and Economic Statistics* 12, 299–308.
- Fisher, S., 1932. *Statistical Methods for Research Workers*. No. 5. Genesis Publishing Pvt Ltd.
- Fitzenberger, B., 1997. The Moving Blocks Bootstrap and Robust Inference for Linear Least Squares and Quantile Regressions. *Journal of Econometrics* 82 (2), 235–287.
- Frankel, J. A., Rose, A. K., 1996. Currency Crashes in Emerging Markets: An Empirical Treatment. *Journal of International Economics* 41 (3-4), 351–366.
- Hamilton, J. D., 1989. A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle. *Econometrica* 57, 357–384.
- Kamin, S., Schindler, J., Samuel, S., 2001. The Contribution of Domestic and External Factors to Emerging Market Devaluation Crises: An Early Warning Systems Approach. FRB International Finance Discussion Paper 711.
- Kaminsky, G. L., 1999. Currency and Banking Crises: The Early Warnings of Distress. IMF Working Paper 99/178.
- Kaminsky, G. L., 2006. Currency Crises: Are they all the same? *Journal of International Money and Finance* 25 (3), 503–527.
- Kaminsky, G. L., Lizondo, S., Reinhart, C. M., 1998. Leading Indicators of Currency Crises. IMF Staff Papers 45 (1), 1–48.
- Kaminsky, G. L., Reinhart, C. M., 1999. The Twin Crises: the Causes of Banking and Balance-of-Payments Problems. *American Economic Review* 89 (3), 473–500.

- Kittelman, K., Tirpak, M., Schweickert, M., Vinhas De Souza, L., 2006. From Transition Crises to Macroeconomic Stability? Lessons from a Crises Early Warning System for Eastern European and CIS Countries. *Comparative Economic Studies* 48 (3), 410–437.
- Knedlik, T., Scheufele, R., 2008. Forecasting Currency Crises: Which Methods signaled the South African Crisis of June 2006? *South African Journal of Economics* 76 (3), 367–383.
- Knedlik, T., von Schweinitz, G., 2011. Macroeconomic Imbalances as Indicators for Debt Crises in Europa. IWH Discussion Papers 12/2011.
- Künsch, H. R., 1989. The Jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics* 17 (3), 1217–1241.
- Kumar, M., Moorthy, U., Perraudin, W., 2003. Predicting Emerging Market Currency Crashes. *Journal of Empirical Finance* 10 (4), 427–454.
- Maddala, G., Wu, S., 1999. A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test. *Oxford Bulletin of Economics and Statistics* 61 (S1), 631–652.
- Mariano, S., Gultekin, B., Ozmucur, S., Shabbir, T., Alper, C., 2004. Prediction of Currency Crises: Case of Turkey. *Review of Middle East Economics and Finance* 2 (2), 87–107.
- Peng, D., Bajona, C., 2008. China’s Vulnerability to Currency Crisis: A KLR Signals Approach. *China Economic Review* 19 (2), 138–151.

Annex

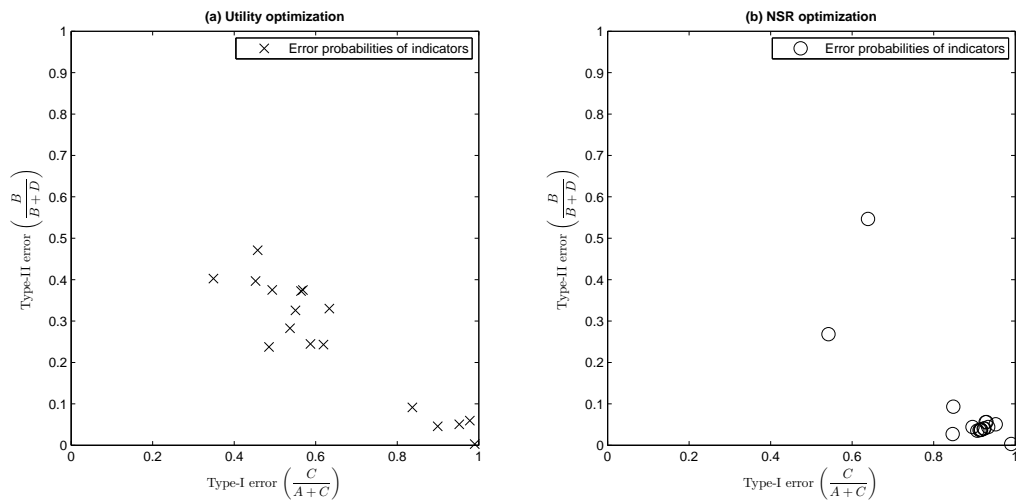


Figure A.1: Type I and Type II error probabilities of individual indicators in the application to currency crises (Kaminsky and Reinhart, 1999)

Table A.1: Results for the currency crisis application

	Utility optimized				NSR optimized			
	Utility	p	NSR	p	Utility	p	NSR	p
M2 multiplier	0.018	0.312	0.901	0.419	0.008	0.470	0.767	0.252
Domestic credit (%GDP)	0.066	0.001	0.740	0.061	0.019	0.157	0.519	0.012
Real interest-rate	0.036	0.197	0.561	0.017	0.029	0.127	0.377	0.007
Lending-deposit rate ratio	-0.019	0.955	2.712	0.966	-0.093	0.999	1.513	0.998
Excess M1 balances	0.028	0.192	0.869	0.335	0.011	0.388	0.665	0.127
M2 (%reserves)	0.084	0.000	0.593	0.007	0.024	0.053	0.436	0.002
Bank deposits	0.036	0.063	0.868	0.324	0.017	0.154	0.540	0.027
Exports	0.069	0.000	0.638	0.002	0.030	0.007	0.419	0.000
Imports	-0.001	0.803	1.057	0.849	-0.001	0.800	1.057	0.895
Terms of trade	0.032	0.088	0.854	0.263	0.008	0.493	0.788	0.280
Real exchange rate	0.139	0.000	0.461	0.004	0.063	0.002	0.176	0.000
Reserves	0.090	0.000	0.610	0.005	0.025	0.050	0.420	0.001
Real interest-rate differential	0.027	0.218	0.454	0.003	0.023	0.144	0.449	0.018
Output	0.075	0.000	0.725	0.063	0.029	0.018	0.614	0.080
Stock prices	0.062	0.004	0.725	0.106	0.023	0.081	0.441	0.012
Deficit (%GDP)	0.124	0.000	0.618	0.042	0.095	0.000	0.586	0.125
Banking crisis	0.004	0.000	0.276	0.000	0.004	0.000	0.276	0.000
Fisher test statistic	498.2	0.002					303.5	0.002
Composite Indicator	0.152	0.000	0.393	0.002	0.009	0.273	0.117	0.002

Note: This and the following tables show the quality measures of the original samples as well as the significance obtained by the bootstrap. The first columns show values obtained from utility optimization, the last columns those for noise-to-signal ratio optimization.

Table A.2: Results for the banking crisis application

	Utility optimized				NSR optimized			
	Utility	p	NSR	p	Utility	p	NSR	p
M2 multiplier	0.075	0.008	0.670	0.078	0.028	0.129	0.437	0.023
Domestic credit (%GDP)	0.088	0.004	0.629	0.040	0.020	0.270	0.525	0.057
Real interest-rate	0.128	0.004	0.549	0.045	0.048	0.104	0.273	0.004
Lending-deposit rate ratio	-0.009	0.824	1.387	0.858	-0.054	0.946	1.296	0.936
Excess M1 balances	-0.004	0.838	1.201	0.888	-0.004	0.828	1.201	0.959
M2 (%reserves)	0.052	0.040	0.825	0.279	0.014	0.390	0.631	0.131
Bank deposits	0.011	0.546	0.883	0.481	0.011	0.489	0.882	0.542
Exports	0.027	0.183	0.902	0.438	0.014	0.257	0.792	0.292
Imports	-0.013	0.967	1.946	0.974	-0.079	1.000	1.439	1.000
Terms of trade	0.009	0.624	0.958	0.704	0.009	0.551	0.958	0.718
Real exchange rate	0.096	0.012	0.553	0.055	0.070	0.014	0.468	0.079
Reserves	0.052	0.045	0.728	0.106	0.011	0.442	0.677	0.172
Real interest-rate differential	0.104	0.011	0.633	0.106	0.057	0.045	0.274	0.003
Output	0.077	0.008	0.652	0.073	0.026	0.128	0.549	0.095
Stock prices	0.082	0.010	0.475	0.014	0.047	0.035	0.286	0.003
Deficit (%GDP)	0.130	0.003	0.635	0.115	0.030	0.231	0.502	0.118
Currency crisis	0.007	0.001	0.456	0.001	0.007	0.001	0.456	0.001
Fisher test statistic	113.7	0.002					89.0	0.004
Composite Indicator	0.174	0.001	0.444	0.049	0.015	0.386	0.322	0.155

Table A.3: Results for the costly asset price boom application

	Utility optimized				NSR optimized			
	Utility	p	NSR	p	Utility	p	NSR	p
LRN_lev	-0.043	0.468	0.920	0.510	-0.068	0.517	0.794	0.417
LRR_lev	-0.007	0.352	0.773	0.392	-0.007	0.347	0.773	0.422
SRN_lev	-0.042	0.513	0.847	0.436	-0.042	0.463	0.847	0.470
SRR_lev	-0.011	0.677	Inf	0.773	-0.202	0.863	1.059	0.853
SPREAD_lev	-0.007	0.558	1.278	0.608	-0.058	0.621	0.932	0.691
SPREAD_lev	-0.019	0.679	1.363	0.664	-0.030	0.530	0.948	0.727
CPI_yoy	-0.013	0.943	Inf	0.730	-0.202	0.932	1.142	0.935
CONS_yoy	0.078	0.005	0.406	0.072	0.019	0.044	0.162	0.004
INV_yoy	0.095	0.003	0.400	0.011	0.066	0.004	0.264	0.003
HINV_yoy	0.068	0.006	0.421	0.028	0.041	0.012	0.322	0.010
REX_yoy	-0.009	0.439	0.899	0.460	-0.069	0.658	0.879	0.565
M1_yoy	0.042	0.019	0.527	0.071	0.042	0.013	0.527	0.082
M3_yoy	0.015	0.087	0.594	0.141	0.014	0.081	0.516	0.082
PCR_yoy	0.061	0.020	0.378	0.045	0.031	0.045	0.238	0.011
DCR_yoy	-0.005	0.367	0.726	0.359	-0.007	0.405	0.713	0.350
CPI_cum	-0.015	0.779	6.033	0.740	-0.218	0.955	1.217	0.955
CONS_cum	0.099	0.004	0.428	0.100	0.014	0.109	0.253	0.025
INV_cum	0.114	0.002	0.337	0.006	0.053	0.008	0.198	0.001
HINV_yoy	0.068	0.006	0.421	0.028	0.041	0.012	0.322	0.010
REX_yoy	-0.009	0.439	0.899	0.460	-0.069	0.658	0.879	0.565
M1_yoy	0.042	0.019	0.527	0.071	0.042	0.013	0.527	0.082
M3_yoy	0.015	0.087	0.594	0.141	0.014	0.081	0.516	0.082
PCR_yoy	0.061	0.020	0.378	0.045	0.031	0.045	0.238	0.011
DCR_yoy	-0.005	0.367	0.726	0.359	-0.007	0.405	0.713	0.350
CPI_cum	-0.015	0.779	6.033	0.740	-0.218	0.955	1.217	0.955
CONS_cum	0.099	0.004	0.428	0.100	0.014	0.109	0.253	0.025
INV_cum	0.114	0.002	0.337	0.006	0.053	0.008	0.198	0.001
HINV_cum	0.069	0.008	0.407	0.032	0.043	0.013	0.311	0.013
REX_cum	-0.013	0.527	1.021	0.556	-0.050	0.609	0.865	0.537
M1_cum	0.025	0.064	0.593	0.152	0.024	0.054	0.580	0.143
M3_cum	0.019	0.085	0.527	0.092	0.019	0.075	0.527	0.104
PCR_cum	0.060	0.023	0.464	0.107	0.024	0.078	0.287	0.025
DCR_cum	-0.011	0.433	0.882	0.487	-0.013	0.439	0.733	0.373
GDPR_detr	0.135	0.002	0.269	0.001	0.124	0.003	0.202	0.000
LRN_detr	0.056	0.052	0.543	0.144	0.042	0.060	0.458	0.085
LRR_detr	-0.006	0.418	0.863	0.457	-0.021	0.514	0.741	0.387
SRN_detr	-0.009	0.496	0.916	0.510	-0.020	0.525	0.738	0.425
SRR_detr	0.003	0.272	0.616	0.244	0.003	0.271	0.616	0.251
REX_detr	-0.018	0.489	1.004	0.535	-0.098	0.676	0.844	0.485
CONStoGDP_detr	-0.023	0.719	1.454	0.744	-0.025	0.505	0.818	0.468
INVtoGDP_detr	0.123	0.002	0.342	0.006	0.096	0.004	0.207	0.000
HINVtoGDP_detr	0.069	0.010	0.218	0.001	0.069	0.007	0.218	0.001
M1toGDP_detr	0.005	0.111	0.630	0.104	0.005	0.110	0.630	0.104
M3toGDP_detr	-0.044	0.691	1.269	0.708	-0.187	0.888	1.037	0.905
PCNtoGDP_detr	0.015	0.097	0.565	0.070	0.015	0.094	0.565	0.074
DCNtoGDP_detr	-0.017	0.422	0.890	0.474	-0.017	0.397	0.890	0.575
GDPR_HP	0.083	0.012	0.434	0.018	0.061	0.021	0.359	0.007
LRN_HP	0.098	0.009	0.383	0.015	0.097	0.007	0.346	0.012
LRR_HP	0.006	0.238	0.637	0.294	0.006	0.234	0.637	0.295
SRN_HP	0.012	0.196	0.536	0.147	0.012	0.189	0.536	0.163
SRR_HP	-0.003	0.437	0.676	0.382	-0.003	0.428	0.676	0.382
REX_HP	-0.034	0.800	1.951	0.818	-0.122	0.749	0.884	0.596
CONStoGDP_HP	-0.026	0.804	2.058	0.815	-0.187	0.897	1.013	0.861
INVtoGDP_HP	0.124	0.004	0.336	0.011	0.095	0.006	0.274	0.005
HINVtoGDP_HP	0.026	0.084	0.469	0.074	0.021	0.098	0.390	0.039
M1toGDP_HP	-0.007	0.184	0.707	0.194	-0.009	0.202	0.697	0.182
M3toGDP_HP	-0.043	0.660	1.282	0.707	-0.045	0.480	1.018	0.843
PCNtoGDP_HP	0.006	0.156	0.634	0.154	0.006	0.156	0.634	0.154
DCNtoGDP_HP	-0.033	0.544	1.094	0.606	-0.059	0.559	0.962	0.702
Fisher test statistic	230.2	0.018					233.6	0.015
Composite Indicator	0.143	0.058	0.256	0.199	0.058	0.195	0.047	0.013

Note: acronyms and abbreviations of indicators are spelled out in Table A.5.

Table A.4: Results for the debt crisis application

	Utility optimized				NSR optimized			
	Utility	p	NSR	p	Utility	p	NSR	p
Government debt (%GDP)	0.099	0.485	0.638	0.614	0.045	0.574	0.329	0.531
Government deficit (%GDP)	0.438	0.033	0.082	0.135	0.381	0.000	0.006	0.000
Interest payment (%gov expenditure)	-0.027	0.714	∞	0.714	-0.052	0.704	1.323	0.664
Unit labor costs	0.290	0.187	0.254	0.177	0.157	0.293	0.143	0.256
Current account	0.337	0.178	0.310	0.462	0.156	0.360	0.095	0.340
Unemployment rate	0.367	0.156	0.057	0.009	0.355	0.005	0.016	0.000
Labor participation rate	0.180	0.109	0.366	0.673	0.104	0.373	0.149	0.367
Private debt	0.298	0.218	0.174	0.270	0.217	0.225	0.112	0.351
Non-MFI debt	0.325	0.179	0.100	0.122	0.233	0.186	0.048	0.145
Foreign assets	0.186	0.006	0.356	0.212	0.150	0.023	0.251	0.244
Inflation	0.044	0.594	0.339	0.487	0.044	0.581	0.339	0.538
Asset prices	-0.026	0.714	∞	0.714	-0.119	0.782	2.526	0.741
HICP-competitiveness	-0.015	0.687	1.132	0.666	-0.015	0.676	1.132	0.666
GDP-deflated competitiveness	-0.022	0.705	1.490	0.673	-0.041	0.701	1.255	0.670
ULC-competitiveness	0.061	0.543	0.521	0.527	0.051	0.539	0.386	0.514
Fisher test statistic	45.0	0.123					166.7	0.002
Composite Indicator	0.468	0.006	0.065	0.154	0.381	0.004	0.006	0.003

Table A.5: Detailed description of surveyed applications

Category	Paper	Alessi and Detken (2011)	Knedlik and von Schweinitz (2011)
Countries	Kaminsky and Reinhart (1999) Argentina, Bolivia, Brazil, Chile, Colombia, Denmark, Finland, Indonesia, Israel, Malaysia, Mexico, Norway, Peru, Philippines, Spain, Sweden, Thailand, Turkey, Uruguay, Venezuela	Alessi and Detken (2011) Australia, Belgium, Canada, Denmark, Finland, France, Germany, Ireland, Italy, Japan, Netherlands, Norway, New Zealand, Spain, Sweden, Switzerland, United Kingdom, United States	Knedlik and von Schweinitz (2011) Austria, Belgium, Finland, France, Germany, Greece, Ireland, Italy, The Netherlands, Portugal, Spain
Time Period	Jan 1970 – Jun 1995	1970Q1 – 2007Q4	Jan 1999 – Nov 2011
Crisis Index	Currency crises: Exchange Rate Pressure Index $I = \frac{\Delta e}{e} - \frac{\sigma_L}{\sigma_R} \frac{\Delta R}{R}$. Banking crises: Expert judgment based on bank runs and government aids	Boom: Asset price index exceeds its filtered average + 1.75 times the standard deviation up to that time for 3 periods in a row: $I_t > HP_{I,t} + 1.75\sigma_{I,t}$	Spreads between 10 year government bond yields and the yield of the EMU-average 10 year AAA-rated government bond: $I = R_c - R_{AAA}$
Crisis Occurrence	$I > \mu_I + 3\sigma_I$	Costly boom: A boom with growth of real GDP 3% below potential growth	$I > \mu_I + 1.65\sigma_I$
Indicators	M2 multiplier, domestic credit (% GDP), real interest-rate, lending-deposit rate ratio, excess M1 balances, M2/reserves, bank deposits, exports, imports, terms of trade, real exchange rate, reserves, real interest-rate differential, output, stock returns, government deficit (% GDP)	Real GDP (GDPR), real consumption (CONS), real investment (INV), real housing investment (HINV), equity prices (QEPR), housing prices (QRPR), aggregate asset prices (QAAPR), term spread (SPREAD), real effective exchange rates (REX), real and nominal 3-month interest rates (SRR, SRN), 10-year bond yields (LRR, LRN), real M1 (M1), real M3 (M3), real private credit (PCR), real domestic credit (DCR). Variables in different transformations: yearly growth rates ($_yoy$), 6q-cumulated growth rates ($_cum$), level ($_lev$), % GDP (toGDP), detrended ($_detr_HF$)	Government debt (% GDP), government deficit (% GDP), government interest payments (% government expenditure), unit labor costs, current account (% GDP), trade share, domestic demand, competitiveness indicators (deflated with HICP, GDP deflator and ULC), inflation, property prices, asset prices, unemployment rate, labor-force participation rate, labor productivity, private debt (% GDP), household debt (% GDP), non-MFI debt (% GDP), foreign assets (% GDP)
Warning horizon	24 months (before currency crises and 12 month before and after banking crises)	6 quarters	24 months
Comments	Data set until June 2003 used instead.	Quasi real-time approach: optimal threshold based on all data, but signals in t sent only, if the quantile based on data until t has been crossed.	