Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

INSTITUT FÜR HÖHERE STUDIEN
INSTITUTE FOR ADVANCED STUDIES
Vienna

TU WIEN
CONTINUING EDUCATION CENTER

MSc Economics

Indicators for Banking Crises in Europe

A Master's Thesis submitted for the degree of
"Master of Science"

supervised by

Univ. Prof. Dr. Robert M. Kunst

Lisa Windsteiger

0256941

Vienna, June 2011

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

I am grateful to Robert Kunst for supervising my thesis. Thanks also to my family for their support and especially to my boyfriend for his understanding and for always reminding me of what is really important in life.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Institut für Höhere Studien
Institute for Advanced Studies
Vienna

TU Wien
Continuing Education Center

MSc Economics

Affidavit

I, Lisa Windsteiger

hereby declare

that I am the sole author of the present Master's Thesis,

Indicators for Banking Crises in Europe

52 pages, bound, and that I have not used any source or tool other than those referenced or any other illicit aid or tool, and that I have not prior to this date submitted this Master's Thesis as an examination paper in any form in Austria or abroad.

Vienna, 14 June, 2011

Signature

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Contents

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

List of Figures

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

List of Tables

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Abstract

This paper examines the quality of five macroeconomic variables as indicators for banking crises in Europe. The variables analyzed are the M2-multiplier, the ratio of domestic credit to real GDP, output, stock indices and the domestic real interest rate. The first part of the paper closely follows the methodologies employed in Kaminsky and Reinhart (1999). This means that first, the average behaviour of the potential indicator variables during crisis times is examined across Europe and then the resulting insights are used to construct indicators for banking crises. However, where the methodology of Kaminsky and Reinhart turns out to be unsuccessful for the present sample, alternative methods are introduced and discussed. Moreover, various criteria to compare the quality of the indicators are proposed. Besides, composite indicators are defined and assessed. As the focus of this paper is on Europe, the sample used to construct the indicators contains only Western and Northern European Countries. Additionally, also for CEE countries the macroeconomic environment of banking crises and the performance of the indicators is analyzed and compared to the findings for Western and Northern Europe. In the last part of the paper, the quality of the indicator variables is examined in the context of a logit model. It turns out that stock indices and output are good single-variable indicator variables, both for Western/Northern Europe and for CEE, while for the former area, they are unfortunately not very leading. Output, the M2-multiplier and stock indices seem to evolve similarly before crises in the two areas, wheras the ratio of domestic credit to real GDP and the real interest rate show different pre-crisis behaviour in Western/Northern Europe and CEE. Generally, the performance of the variables as indicators can be improved by constructing composite resp. two-variable indicators. The right combinations of variables lead to a precision of the prediction and a considerable reduction of error rates. In the estimation of the logit model, it turns out that a specification that includes only output, stock indices and the real interest rate and their respective one-period lags as explanatory variables yields the best results.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

1 Introduction

Up to the 1980s, when financial liberalization advanced worldwide, banking crises were relatively rare and hence their causes were of limited interest to both policy makers and academics. However, in the 1980s and early 1990s, a number of countries, both developing and industrialized, experienced systemic banking crises. Consequently, understanding the macroeconomic surroundings of banking crises became more relevant.

A systemic banking crisis is harmful to the economy for various reasons. It may lead to a credit crunch, which causes investment and consumption slumps and can force sound firms into bankruptcy. Furthermore, as the confidence in domestic financial institutions is deeply shaken, it can result in a decline in domestic credit and large capital outflow. Additionally, not only unsustainable financial institutions but also viable banks may cease to operate. Systemic banking crises will lead to rescue operations by policy makers like expansionary monetary policy or massive bailouts. However, these measures are very costly. Loose monetary policy can be inflationary and financial aid for failed banks will help also inefficient institutions to remain in business. Besides, bailouts cause moral hazard problems, as the expectance of future bailouts reduces the incentives for sustainable risk management in banks. Therefore, it is important to understand the mechanisms leading up to a banking crisis to recognize dangerous macroeconomic developments early and thereby enable policy makers to take precautionary action to prevent the occurrence of systemic banking crises. Hence, since the 1980s, the literature on banking crises, and in particular on their macroeconomic surroundings and leading indicators, has flourished.

The methodologies used to examine banking crises have mostly been developed (or at least have first been used) in the context of analyzing currency crises. The empirical literature on balance-of-payment and banking crises can roughly be classified into three methodological categories. The first group of studies examines the behaviour of economic variables in a group of countries in the period leading up to and immediately following a crisis. This is compared either to the average behaviour of the variables in the same countries during non-crisis times or to the average evolution of the variables in other countries where no crisis has happened and which serve as some kind of control group. The first part of Kaminsky and Reinhart (1999) belongs to this branch of literature.

In the same paper, Kaminsky and Reinhart (1999) introduce the indicator approach to examine currency and banking crises. The idea behind this approach is to use macroeconomic variables to define indicators which give "warning" signals of a crisis if the value of the variable crosses some critical threshold. By comparing the timing of the signals to the actual crisis dates (ex post), the ability of the indicators to predict crises can be assessed. Applying this methodology was pioneering in this area and has met broad response in the empirical literature on banking crises.[1]

The third approach used in the literature is to estimate the probability of occurrence

[1] see for instance Borio and Lowe (2002), (2004)

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

of a banking crisis in the context of a parametric model, mostly a logit model, where the explanatory variables comprise potential indicator variables.[2]

The present paper combines these three methodologies to examine the macroeconomic developments around systemic banking crises in Europe. More explicitly, the first and the second part of the paper will closely follow Kaminsky and Reinhart's analysis with respect to the methodologies employed. This means that first, the average behaviour of potential indicator variables during crisis times is examined across Europe and then the resulting insights are used to construct indicators for banking crises. However, there are some important differences to the Kaminsky/Reinhart paper. First, the group of countries examined by Kaminsky and Reinhart consists of a few industrial countries, but the majority are developing countries. Contrary, the focus of this paper is on Europe. The sample contains only Western and Northern European countries, mostly members of the Eurozone. Additionally, the macroeconomic environment of banking crises in CEE countries is analyzed and compared to the findings for Western and Northern Europe. As the sample period includes the most recent crises in 2008, the study is very up-to-date. Second, where the methodology of Kaminsky and Reinhart turns out to be rather unsuccessful for the present sample, alternative methods are introduced and discussed. Besides, some extensions proposed by Kaminsky and Reinhart and partly put into practice in their following work (e.g. Kaminsky (1998)) are added. Moreover, by comparing the indicators to each other with respect to various features that seem desirable, like leadingness or persistence of the signals, their relative usefulness for policy making is assessed. Furthermore, parts of Kaminsky and Reinhart's methodology that seem to be imprecise or inconsistent are pointed out and discussed. The last part of the paper will be the estimation of a logit model using the previously analyzed indicators as explanatory variables.

Hence, this paper presents a very detailed analysis of the macroeconomic surroundings of banking crises in Europe. Moreover, as the indicator qualities of the examined variables are examined, its findings can be a starting point for the construction of a banking crisis warning system for Europe, which could be valuable for European policy makers like the ECB.

The rest of the paper is organized as follows: Chapter 2 presents a short overview of the theory on banking crises and their causes. Chapter 3 lists the countries that were included in the present study and the data that was used and in chapter 4, a descriptive analysis of the macroeconomic developments around banking crises in Europe is performed. In chapter 5 the indicators are constructed and compared. Chapter 6 shows and discusses the results of the logit model estimation, while chapter 7 summarizes the findings and chapter 8 concludes.

[2] see for instance Hardy and Pazarbasioglu (1998) and Demirgüc-Kunt and Detragiache (1998)

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

2. The Theory

The major part of a bank's liabilities are short-term deposits, while the asset side consists mainly of short- and long-term loans. If the value of the assets falls below the value of liabilities, a bank becomes insolvent. As the value of assets may decrease because debtors are unable to service their loans, banks (should) try to minimize their exposure to credit risk. This can be done by screening prospective borrowers, by diversification of the portfolio across regions and sectors and by demanding collateral. However, diligent screening of loan applicants can be costly and there is no guarantee that a project that seems to be profitable ex ante will actually succeed. Diversification may turn out to be difficult for regional institutions and for banks who specialize in specific sectors. Finally, monitoring the collateral can be elaborate and the values often fluctuate. Hence, the banks' exposure to credit risk cannot be avoided to a certain degree.

A systemic banking crisis arises if a critical proportion of banks in a country experience significant loan losses relative to their capital. Thus, theory would suggest that shocks that deteriorate the economic performance of debtors and cannot be ameliorated by risk diversification increase the probability of the occurrence of banking crises and that less capitalized banking systems will be especially vulnerable. The shocks can be isolated, like for instance an unexpectedly large increase in demand for currency that causes illiquidity in the banking industry and results in a banking panic, but also slumps in output growth and asset price busts can trigger crises. As the crises in Latin America in the 1980s have shown, financial liberalization, lack of banking supervision and official deposit guarantees can be an explosive mixture. The resulting excessive commercial bank lending and asset price booms put the banking system into distress if the economic growth slows down and lead to a surge in non-performing loans and stock and real estate market busts. A vicious circle of bank insolvencies, credit crunch and recession may arise. As the recession unfolds and more and more financial institutions falter, uncertainty about the quality of individual assets may cause a large withdrawal of deposits and even more bank crashes.

Increases in the short-term interest rate can cause the balance sheets of banks to deteriorate for two main reasons. The asset side of banks usually consists to a large extent of loans with fixed interest rates, while the rate for depositors is adjusted. Thus, the interest rate paid on bank assets may fall below the rate paid on liabilities. But even if interest rate changes can be passed on also on the asset side, higher real interest rates resulting from an increase in the short-term rate can lead to a higher fraction of non-perfoming loans. Therefore, high real interest rates can cause systemic banking sector distress.

While there is no consensus in the literature as to what are the ultimate causes of banking crises, the various symptoms of crises — anomalous behaviour of certain macroeconomic variables — can be studied. The common features of the developments around banking crises across Europe can be identified and will be the basis of the

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

following analyses and the construction of the indicators.

3 The Data

3.1 Countries and time periods

Countries analyzed are Austria, Belgium, Denmark, France, Germany, Ireland, Norway, Netherlands, Portugal, Spain, Sweden, Switzerland, UK, Bulgaria, Croatia, Czech Republic, Hungary, Poland, Slovak Republic, Slovenia and the Ukraine. All of these countries have experienced at least one banking crisis during the studied time period. The studied time period however varies for every country and every macroeconomic variable, depending on data availability. Where possible, the period was chosen to be from 1980 to 2011. For some CEE countries, data was only available from the early 1990s, and for Euro area countries the monetary data was obtained from 1999 onwards. The exact definitions of the indicator variables and the data sources are documented in the appendix. Table 23 in the appendix lists the maximum time spans, for which data was available, for every country. For the descriptive analysis and the indicator approach, monthly data was used, while for the estimation of the logit model in the last part of the paper, I use annual data.

3.2 Banking crisis data

Kaminsky and Reinhart determine the dates of the banking crises at the beginning of their paper. They use a combination of different information from various sources to decide, whether an economic situation in a country is classified as a crisis. As this vast amount of information and data is not available to me and to avoid discussions about criteria for banking crises, I refrain from this and use a classification by the IMF instead. More precisely, I refer to the paper by Laeven and Valencia (2010). They use a combination of quantitative data (including data on large scale bank failures, non-performing loans, occurrence of deposit runs, government interventions) and subjective assessment to identify systemic banking crises around the world from 1970 to 2010. The method is very similar to the way Kaminsky and Reinhart determine the banking crises dates, and for the countries and the time periods where the Kaminsky/Reinhart- and the IMF- analysis overlap, they basically have the same results. The appendix lists the timing of the banking crises for the analyzed countries (Table 22).

4 Descriptive Analysis

In the first part of their paper, Kaminsky and Reinhart perform a descriptive analysis (often referred to as *event study methodology* in financial crisis literature) on 16 macroeconomic variables to assess their potential quality as leading indicators. They compare the behaviour of the variables in a time interval before and after the beginning of a crisis and during tranquil times. More specifically, except for interest-rate variables,

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

which are in levels, they look at the average difference of the twelve-month percent changes around the beginning of a crisis relative to tranquil times. The idea behind this analysis is to get insight into the macroeconomic developments around banking crises and see, which variables show anomalous behaviour and would thus be potential indicator variables for crises.

Following Kaminsky and Reinhart, this paper examines the evolution of five potential indicator variables in a time interval of 18 months before and after the beginning of a banking crisis and looks at the average difference of the twelve-month growth rates between tranquil and crisis periods. The first two analyzed variables, the M2-multiplier and the ratio of domestic credit to real GDP, are informative of the financial sector. The M2-multiplier is defined as the ratio of the monetary aggregate M2 to base money. Examining the evolution of this variable around crisis times in Western/Northern Europe yields a picture very similar to its equivalent in Kaminsky and Reinhart. The solid line in Figure (1) shows that the twelve-month growth of the M2-multiplier is on average up to 20 percent higher than normal in the months before a crisis. The dashed lines above and below depict the standard deviation. The result is in line with what we would expect from theory: The higher the ratio of M2 to base money, the more susceptible is the banking system to confidence crises and bank runs. Moreover, Kaminsky and Reinhart argue that financial liberalization is likely to be accompanied by reductions in reserve requirements. Hence, the M2-multiplier can also be viewed as reflecting the degree of financial liberalization, which in turn often precedes banking crises. For the

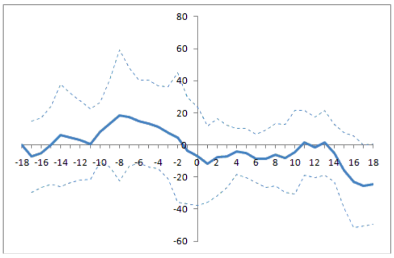

Figure 1: Evolution of the M2-Multiplier during crisis times in Western/Northern Europe

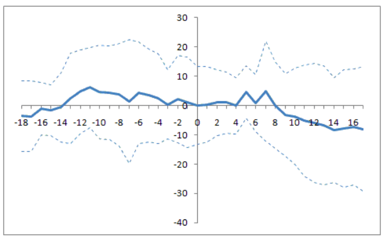

Figure 2: Evolution of the M2-Multiplier during crisis times in CEE

CEE countries the picture (Figure (2)) looks similar, although the average difference of the growth rates between normal and crisis times seems to be smaller. Note however,

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

that the significance of the difference cannot be assessed in this framework.

The next variable under study is the ratio of domestic credit to real GDP. Kaminsky and Reinhart examine this variable, because theory suggests that a credit boom will put the banking system into distress and can therefore lead to banking crises. For example, Calomiris and Gorton (1991) find that crises are often preceded by recessions following a period of high credit expansion. As the recession comes along, depositors, when reassessing the risk of bank debt, have little information about the quality of indiviual assets and the general uncertainty and instability often leads to a huge amount of withdrawals from banks. For Western and Northern Europe, it can be concluded that on average the ratio grows higher than normal for the whole interval before and after a crisis (Figure (3)). This fits theory quite well: before the crisis the ratio is high due to a credit boom and if the amount of borrowing declines later (credit bust) the ratio is high, because output falls relatively more than credit. However, the standard deviation is huge. The reason for this is that the evolution of this variable is very inhomogeneous across countries. For most countries the growth rate is above average, but for some countries, like Norway and the UK, the situation is exactly the opposite. For CEE,

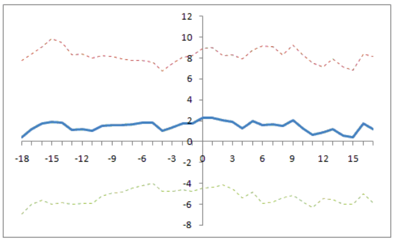

Figure 3: Evolution of Domestic Credit/Real GDP during crisis times in Western/Northern Europe

the picture looks more like a credit boom and bust, but the standard deviation is also very high (Figure (4)). This leads to the conclusion that the ratio of domestic credit to output might not be a very good indicator for banking crises in Europe.

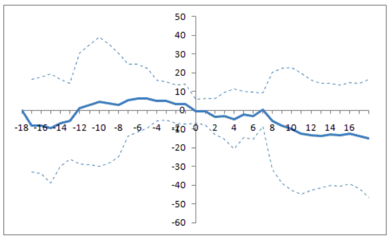

Figure 4: Evolution of Domestic Credit/GDP during crisis times in CEE

Next, the real sector is analyzed: Figures (5) and (6) show the evolution of output (resp. real GDP) and stock indices in the months before a crisis in Western Europe. These pictures fit the story that slumps in economic activity and asset prices often precede banking crises. For CEE the pictures look similar (Figures (7) and (8)), though

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

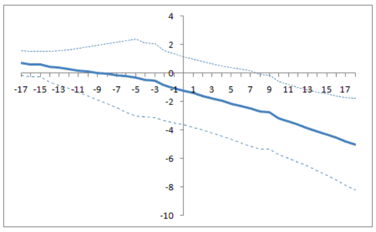

Figure 5: Evolution of output during crisis times in Western/Northern Europe

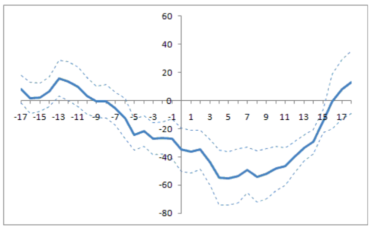

Figure 6: Evolution of stock indices during crisis times in Western/Northern Europe

the standard deviations seem higher.

Finally, analyzing the evolution of the domestic real interest rate during crisis times, it turns out that it is on average up to two percentage points higher than normal in Western/Northern Europe (Figure 9). This is what we would expect, as high real interest rates could be the sign of a liquidity crunch precedeing a slowdown in economic activity and banking distress. For CEE the shape of the picture looks similar, although the percentage point difference is much higher (Figure 10). The main reason for this is the high level of real interest rates in Ukraine during crisis times.

5 The indicator approach

5.1 What Kaminsky and Reinhart do

At the end of their paper, Kaminsky and Reinhart introduce a methodology, which has a long history in the literature on business-cycle turning points, to develop leading indicators for banking and currency crises. This approach requires two main steps.

First, a reasonable time interval around the beginning of a banking crisis needs to be defined in which the indicator should give a signal. For banking crises, Kaminsky and Reinhart choose this interval to be twelve months before and after the beginning of the crisis. Their argument for choosing this interval is that the peak of a banking crisis is usually several months after the beginning of the crisis. Therefore, the events at the beginning of a banking crisis are often not viewed as being systemic and not treated seriously enough by policy makers. Hence, a signal coming from an indicator after the beginning of a crisis could still be valuable in the sense that it could help to clarify the situation for policy makers. For Kaminsky and Reinhart, an indicator correctly calls a

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

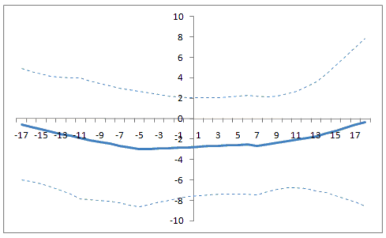

Figure 7: Evolution of output during crisis times in CEE

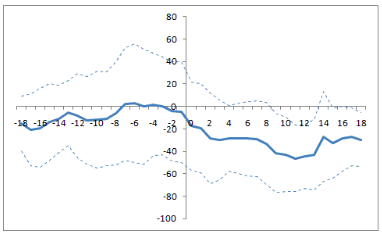

Figure 8: Evolution of stock indices during crisis times in CEE

crisis if it issues a signal at least once in the defined time interval.

After choosing the interval, which is common for all indicators, the second main step in the analysis is to define, when an indicator is seen as issuing a signal. This basically amounts to finding an optimal cutoff-point in the (sample) distribution of the economic variable. The indicator is set to 0 if the variable is below the threshold and becomes 1 above the threshold (or the other way round for variables for which small values are typical before crises[3]). Kaminsky and Reinhart choose the optimal threshold on an indicator-to-indicator basis.[4] Of course, as they point out, there is a trade-off in selecting the optimal threshold for an indicator. If the threshold is very low, the indicator will correctly signal many crises, but it will also signal a lot of crises that do not occur. On the other hand, if the threshold is very high, the indicator will rarely give false alarms, at the expense of missing also many crises that DO occur. Therefore, their criterion is to minimize the noise-to-signal ratio of an indicator on a grid of potential thresholds. This criterion will be explained in detail below. The optimal threshold of an indicator is then some percentile of the sample distribution of the indicator. Note that the percentile value for every indicator is common across countries, but the numerical value of the threshold can (and most likely will) be different for each country. Having selected the optimal threshold this way, Kaminsky and Reinhart compare the indicators by looking at the percentage of crises they signal. For banking crises, they find that especially the real sector plays an important role, with output and stock prices calling 89 resp. 91 percent of the crises. Additionally, high real interest rates seem to be a

[3]Whether small or large values of the indicator variable are typical before the occurrence of a crisis can be deducted from the descriptive analysis of the previous section.

[4]Again, for most of the variables they do not look at levels, but at the twelve-month percentage changes.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

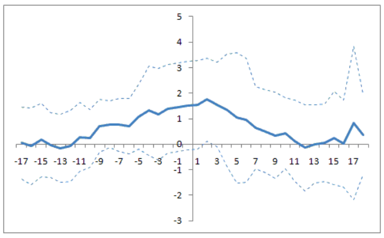

Figure 9: Evolution of the domestic real interest rate during crisis times in Western/Northern Europe

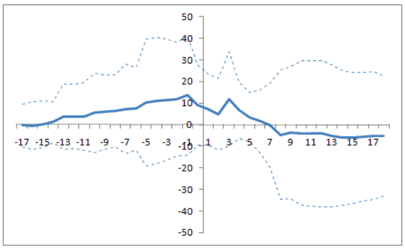

Figure 10: Evolution of the domestic real interest rate during crisis times in CEE

good indicator, with 100 percent of the crises correctly signalled.

5.2 Own analysis

In this section the macroeconomic variables examined above are used to construct leading indicators for banking crises and examine their quality. The first analysis is completely analogous to the methodology of Kaminsky and Reinhart. Especially, for every indicator the cutoff-point in the percentile is selected that minimizes the noise-to-signal ratio. However, the results that are obtained are not very satisfactory (Table 1). Except for the M2-multiplier and the real interest rate, the optimal thresholds are very extreme percentiles, and the percentage of crises correctly called is not overwhelming.

At this point it seems instructive to explain in detail the criterion used for choosing the optimal threshold, namely minimizing the noise-to-signal ratio. If the threshold is very low (or very high for certain indicators), then the indicator will issue many signals and therefore correctly signal a high percentage of the crises, but the "noise", so the number of false alarms, will also be high. Setting the threshold to a very extreme value

Indicator	critical region	percentage signalled	noise-to-signal ratio
M2-multiplier	≥ 0.92	50	0.79
Domestic credit/GDP	≥ 0.998	35.7	0.3
Real GDP	≤ 0.002	66.67	0.041
Stock indices	≤ 0.02	84.6	0.1
Real interest rate	≥ 0.72	100	0.64

Table 1: Results for minimized noise-to-signal ratio

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

will reduce the number of false alarms, at the expense of missing many crises that do occur. For every indicator variable and every threshold one can construct the following table:

	Crisis occurs in the interval ± 12 months of the signal	Crisis doesn't occur in the interval ± 12 months of the signal
Indicator issues a signal	A	B
Indicator doesn't issue a signal	C	D

The months in which the indicator gives a signal and the crisis occurs within the defined interval are counted in cell A. If no crisis occurs after the signal, this is the case of false alarm and counted in cell B. If the indicator doesn't signal but a crisis occurs within the defined time frame this is counted in cell C. Every month the indicator doesn't signal and no crisis occurs is counted in D. A good indicator would therefore have many entries in A and D and few in the off-diagonal cells. The noise-to-signal ratio as defined by Kaminsky and Reinhart is then $\frac{\frac{B}{B+D}}{\frac{A}{A+C}}$, so the ratio of wrong signals to all possible wrong signals divided by the ratio of correct signals to all possible correct signals. Kaminsky herself (1998) mentions, that there is no general approach to choosing the optimal threshold. Kaminsky and Reinhart point to two drawbacks of their criterion: First, if policy makers react to the signal of an indicator and successfully prevent the occurrence of a crisis after the signal, then the signal is counted as a false alarm and the indicator might have a very high noise-to-signal ratio, although its signals are very accurate. Second, all signals that occur within the specified time frame before and after a crisis are treated the same, but of course the signals before the occurrence of a crisis are more valuable to policy makers. Indicators which give many signals in the twelve months before a crisis are therefore more desirable than those who mostly signal after the beginning of a crisis, although the noise-to-signal ratio might be the same or even smaller.

These are certainly drawbacks one has to keep in mind when using the criterion. However, Kaminsky and Reinhart do not even mention its largest disadvantage: When applying it to choose the optimal cutoff-value for indicators this criterion doesn't put equal weight on minimizing $\frac{B}{B+D}$ and $\frac{C}{A+C}$. Why not? This is easy to see: Note that $\frac{B}{B+D}$ and $\frac{C}{A+C}$ are, loosely speaking, inversely connected. $\frac{B}{B+D}$ can be made small by choosing a very extreme percentile threshold, but then there will be very few correct crises signals, which means that $\frac{C}{A+C}$ will be relatively high.

What is the smallest value the noise-to-signal ratio can obtain? $\frac{B}{B+D}$ and $\frac{A}{A+C}$ are between zero and one, so the smallest possible value of the ratio would be zero. When will the ratio be zero or close to zero? $\frac{A}{A+C}$ (which is equal to $1 - \frac{C}{A+C}$) can be at most one, if $\frac{C}{A+C}$ is zero, and this would minimize the ratio for a constant $\frac{B}{B+D}$. But the ratio is zero only if $\frac{B}{B+D}$ is zero, and for all values of $\frac{A}{A+C}$ except for $\frac{A}{A+C} = 0$. This

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Indicator	$\frac{B}{B+D}$	$\frac{C}{A+C}$
M2-multiplier	0.078	0.9
Domestic Credit/GDP	0.0043	0.986
Real GDP	0.0008	0.97
Stock indices	0.01	0.98
Real interest rate	0.26	0.59

Table 2: Errors with minimized noise-to-signal ratio

means that the minimum in the interval $[0,1]\times[0,1]$ is obtained if and only if $\frac{B}{B+D}=0$, but for all values of $\frac{A}{A+C}$ except 0. So even if one could freely choose $\frac{B}{B+D}$ and $\frac{C}{A+C}$ in the interval $[0,1]$, there would be a tendency towards minimizing $\frac{B}{B+D}$, while more or less ignoring $\frac{C}{A+C}$ as long as it isn't 1 or very close to 1. But in the context here, this tendency is even more present. Why? Well, let us look how small $\frac{B}{B+D}$ can get. In the denominator we have all possible wrong signals, so this is the number of months where no crisis occurs within twelve months or has occured in the previous twelve months summed up over countries. This is a very large number, so if B is very small — and it can be made very small by choosing a very extreme threshold — $\frac{B}{B+D}$ can get very small. On the other hand, how small can $\frac{C}{A+C}$ get? In the denominator we have the sum of all possible correct crisis signals, so this is the sum of all months in the crisis time interval, summed up over all countries. But of course, most of the time we are in a tranquil period, so $A+C$ is much smaller than $B+D$. The minimum value of $\frac{C}{A+C}$ is also 0, but only if there is never a single signal missed by the indicator, which would mean that the threshold has to be very lax. But the next smallest possible value (if $C=1$, so one signal is missed) would already be much higher than the value of $\frac{B}{B+D}$ if $B=1$, because the denominator of the latter is higher.[5] Therefore, it is easier to minimize the noise-to-signal ratio by making $\frac{B}{B+D}$ very small. And indeed, the optimal cutoff-point that I get is always such, that $\frac{B}{B+D}$ is much smaller than $\frac{C}{A+C}$, as can be seen in Table 2.

It seems that for the analysis of Kaminsky and Reinhart this wasn't such a problem, because they had more crises relative to tranquil times and because their data was more heterogeneous (very heterogeneous countries), so that they couldn't make $\frac{B}{B+D}$ so small relative to $\frac{C}{A+C}$. Therefore, they also didn't get such extreme optimal thresholds. However, presumably they also get much smaller values for $\frac{B}{B+D}$ than for $\frac{C}{A+C}$. Kaminsky and Reinhart (1999) do not report $\frac{B}{B+D}$ and $\frac{C}{A+C}$, but I redid the analysis for some indicator variables using their data (their countries and time periods) and got the same results for the noise-to-signal ratio and the percentage of crises correctly

[5] Additionally, note that for some indicator variables the behavior of the variable is not the same for all countries in the sample, like for example for the ratio of domestic credit to GDP. For most countries the twelve-month growth of the variable is above normal during crises times, but for UK and Norway it is below normal. Therefore, no matter how low the threshold is set, A (and thus also $\frac{A}{A+C}$) will never be very high, so it is even more obvious that minimizing the noise-to-signal ratio will boil down to basically minimizing $\frac{B}{B+D}$, which is achieved by choosing a very extreme threshold and thereby minimizing the number of signals in general and hence also the number of false alarms (note that $A+B$ can never be 0 with the way the indicators are defined).

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

signalled, and the resulting errors were such that $\frac{B}{B+D}$ was always smaller than $\frac{C}{A+C}$. Besides, the errors reported in Kaminsky (1998) point to this problem as well.

Thus, minimizing the noise-to-signal ratio in this context puts more weight on minimizing $\frac{B}{B+D}$ than on minimizing $\frac{C}{A+C}$. But what is the optimal criterion for choosing the cutoff-point for the indicators? What exactly should be minimized depends on the loss function of the policy maker and on the measures the policy maker plans to take in response to indicators signalling a banking crisis in the future. If a central bank intends to react with monetary policy measures, the asymmetric nature of the costs of policy errors is not negligible. The bank might want to be pretty sure that there is really the danger of a banking crisis, because the risk of destabilizing the economy is high. Therefore the central bank might prefer to get very few false signals, at the expense of missing some crises, so there would be more weight on $\frac{B}{B+D}$ than on $\frac{C}{A+C}$ in the loss function. On the other hand, as described above, the rescue measures needed after the occurrence of a banking crisis could be much more expensive than a maybe unnecessary intervention. In addition, if the policy measure chosen by the central bank is to strengthen the prudential framework, the costs of reacting if there is no crisis are small relative to the costs of missing a crisis, so there would be more weight on $\frac{C}{A+C}$ in the loss function.

Hence, it is not at all clear that $\frac{B}{B+D}$ and $\frac{C}{A+C}$ should be equally weighted. However, it seems as if the tendency towards minimizing $\frac{B}{B+D}$ in Kaminsky and Reinhart was not deliberate, at least it wasn't mentioned anywhere. Bearing the above considerations in mind and noting that we do not know the loss function of the policy maker, I would propose an alternative criterion for choosing the optimal threshold, which gives equal weight to $\frac{B}{B+D}$ and $\frac{C}{A+C}$. To explain this criterion, it might be fruitful to move to the area of binary classification tests respectively diagnostic tests.

In medicine, diagnostic tests are conducted to decide, whether a patient suffers from an illness or not. Designing the tests results in choosing an optimal cutoff-value of a variable that is analyzed, above which the test result is positive. Hence, there is a trade-off between minimizing the number of false positives (the test outcome is positive, although the patient doesn't suffer from the illness) and false negatives. Similar to above, the following table can be constructed (which is also called the confusion matrix):

	Condition positive	Condition negative
Test outcome positive	True positives	False positives
Test outcome negative	False negatives	True negatives

Of course it is obvious that the problem of designing an optimal diagnostic test is analogous to the problem of finding optimal indicators for crises. The "illness" in this case would be an upcoming crisis, and a positive test outcome would be the signal of an indicator. For every indicator (and every cutoff-point in the sample distribution) one could then construct a table like the one above.

What kind of criteria are used in the literature to assert the quality of such diagnostic tests? Two important measures are the sensitivity and the specificity of the test,

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

where[6]

$$\text{sensitivity} = \frac{\sum \text{True positives}}{\sum \text{True positives} + \sum \text{False negatives}}$$

and

$$\text{specificity} = \frac{\sum \text{True negatives}}{\sum \text{True negatives} + \sum \text{False positives}}.$$

It is immediate to see that the sensitivity corresponds to $\frac{A}{A+C}$ from above, while the specificity corresponds to $\frac{B}{B+D}$. In the literature of binary classification tests, $\frac{A}{A+C}$ is also called the true positive rate (TPR) and $\frac{B}{B+D}$ is the false positive rate (FPR). A test with a high true positive rate, when negative, "rules out" disease, while a highly specific test, when positive, can be regarded as true positive. As noted above, sensitivity and specificity are inversely related: the higher the sensitivity of a test, the lower the specificity. This relationship can be depicted by a so-called *Receiver operating characteristic (ROC) curve*. This curve is a graph of sensitivity (y-axis) against 1-specificity (x-axis) for all possible values of the threshold. A test with high sensitivity corresponds to a high y-value, while high specificity means a low x-value. Which point on the ROC-curve is optimal depends on the utility function of the policy maker. If the utilities for policy makers of getting false positives (U_{FP}), true positives (U_{TP}), false negatives (U_{FN}) and true negatives (U_{TN}) are known, the expected utility from a diagnostic test can be computed as

$$E(U) = P \cdot [TPR \cdot U_{TP} + (1 - TPR).U_{FN}] + (1 - P) \cdot [FPR \cdot U_{FP} + (1 - FPR) \cdot U_{TN}],$$

where P is the prevalence of a crisis. If P is 0.5 and high sensitivity and high specificity are equally desirable the problem boils down to choosing the threshold such that $TPR+ (1 - FPR)$, or equally $\frac{A}{A+C} + \frac{D}{B+D}$, is maximized. This is the same as maximizing the sum of sensitivity and specificity, or minimizing $\frac{B}{B+D} + \frac{C}{A+C}$. Subsequently, the optimal cutoff-value of the indicator would be the one which yields the point closest to the upper left corner of the diagram (with respect to the L_1-norm).[7] Graphically, this point can be obtained by shifting the 45-degree line upwards until it tangents the ROC-curve. If sensitivity and specificity are not equally weighted, or if the prevalence is not equal to

[6]Other informative measures are the positive predictive value (precision rate, PPV) and the negative predictive value (NPV), where

$$\text{PPV} = \frac{\sum \text{True positives}}{\sum \text{True positives} + \sum \text{False positives}}$$

and

$$\text{NPV} = \frac{\sum \text{True negatives}}{\sum \text{True negatives} + \sum \text{False negatives}}.$$

Thus, the PPV reflects the probablity that a positive result is a true positive result, while the NPV does the same for negative results. These measures are problematic however, as they depend on the prevalence of the illness $\left(= \frac{\sum \text{True positives} + \sum \text{False negatives}}{\sum \text{all cases}}\right)$, or in our case on the prevalence of a crisis. Therefore, these measures are not used here to decide on the optimal threshold, but they appear in chapter 5.4, when the indicators are compared.

[7]This point therefore maximizes Youden's J-Statistic, which is commonly used in the literature to assess the performance of a diagnostic test and is equal to *Sensitivity + Specificity - 1.*

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Indicator	critical region	percentage signalled	$\frac{B}{B+D}$	$\frac{C}{A+C}$
M2-multiplier	≥ 0.91	50	0.089	0.889
Domestic Credit/GDP	≥ 0.68	85.7	0.29	0.5
Real GDP	≤ 0.20	86.7	0.15	0.4
Stock Indices	≤ 0.23	100	0.2	0.547
Real interest rate	≥ 0.70	100	0.28	0.569

Table 3: Results with new criterion for the optimal threshold

0.5, the slope of the tangent on the optimal point will be different.

In the subsequent analysis, the optimal threshold is chosen such that it maximizes $\frac{A}{A+C}+\frac{D}{B+D}$. This can be justified by the following: The prevalence of a crisis is hopefully not 0.5, but the costs of missing a crisis are very likely to be higher than the costs of an unnecessary policy intervention. Especially after the last period of banking crises, policy makers might therefore be keen on detecting a large fraction of banking crises in advance, and therefore give equal weight to $\frac{A}{A+C}$ and $\frac{D}{B+D}$, although banking crises do not occur that often. Applying this criterion for choosing the optimal thresholds yields more sensible results (Table 3).

Judging the indicators by how many crises they signal correctly, the M2-multiplier is the worst, with only 50% of the crises called and a very high $\frac{C}{A+C}$, which means that there are a lot of missed signals. The stock indices and the real interest rate indicator would perform best, with 100% of the crises correctly signalled. However, note that $\frac{C}{A+C}$ in this case is also relatively high, which means that although the indicator calls every crisis, it doesn't signal very often during crisis periods, which might be undesirable.

5.3 Type I and Type II error

Kaminsky and Reinhart like to see the choice of an optimal threshold in the framework of a statistical test. If the null hypothesis is that we are in a tranquil period, the cutoff-region of the sample distribution of the indicator variable would correspond to the critical region of the test, the region in which the null hypothesis is rejected. For Kaminsky and Reinhart, the size of this rejection region is equal to the probability of making the type I error (rejecting the null when it is true), denoted as α. But there is a little inconsistency in this. Unless only non-crises observations are regarded, the sample distribution of an indicator variable is no consistent estimator for the distribution of the variable under the null of a tranquil state. It rather converges to a mixture of the distribution under the null and under the alternative (=crisis state). Therefore, the size of the cutoff-region does not correspond to α, the probability of rejecting the null when it is true. To obtain a test of size α, we would have to leave aside all crisis observations and compute the threshold percentile only from the sample distribution of the non-crisis observations, which is obviously not what Kaminsky and Reinhart do — they use all observations. Hence, if their method of finding the optimal threshold is put in the framework of a statistical test, this threshold is not equal to the size of the test. However, the procedure produces valuable results and sensible indicators, so there

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

is no reason not to apply it. Nevertheless, it is imprecise to say that the cutoff-region is equal to α, the size of the test.

If we see the threshold optimization problem in the framework of a binary classification test as described above and think of the confusion matrix, it is easy to see that α and β (which is the probability of making the type II error, so to wrongly accept the null) are closely linked to sensitivity and specificity. More precisely, a highly specific test has a low type I error rate, while high sensitivity corresponds to a low type II error rate.[8]

This gives rise to another important issue. There seems to be some inconsistency in Kaminsky and Reinhart concerning the measurement of A and C, which becomes clearer in the above framework of diagnostic tests. If an indicator correctly signals a crisis, as long as at least one signal is issued in a crisis interval, then crises where no signal is issued during the interval should be counted in C (= false negatives) and crises where at least one signal is issued in A (= true positives). Kaminsky and Reinhart however count in A all signals that are issued during crisis intervals, so if an indicator issues five signals during one crisis period, this is counted five times in A. On the other hand, if an indicator issues only one signal during a crisis interval (note that the crisis is correctly signalled according to Kaminsky and Reinhart) then every month during the crisis period where the indicator doesn't signal is counted in C. Therefore, Kaminsky and Reinhart's $\frac{C}{A+C}$ doesn't really represent the fraction of missed (= not signalled) crises, it is rather a measure of how often (= how many months) the indicator doesn't issue a signal during crisis intervals. Of course, this measure is also informative of the quality of the indicator, as an indicator can be viewed as being good if it issues many signals during a crisis period, because then the information is clearer. However, if Kaminsky and Reinhart define that a crisis is correctly signalled already with one signal, then the method of counting A and C which is proposed here is more appropriate and $\frac{C}{A+C}$ would really reflect the probability of making the type II error then.

Using this new measure of $\frac{C}{A+C}$, the gridsearch to identify the optimal threshold can be repeated, where the old measure of $\frac{C}{A+C}$ is now replaced by the "false negatives", so the fraction of crises that are missed by the indicator, which can easily be seen to be one minus the percentage of crises correctly signalled. The result is displayed in Table 4. In the following comparative analysis this specification of the indicators will be used.

5.4 Comparing the indicators

After defining the indicators as above, where the threshold is set according to the explained criterion, there are many ways to assess the quality of the individual indicators and to compare them to each other. One way is to simply look at the percentage of crises they signal. If interest exclusively focused on this number, then the real inter-

[8] If the sample is typical for the population distribution, $\frac{C}{A+C}$ is a good estimate for β and $\frac{B}{B+D}$ for α.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Indicator	critical region	percentage signalled	$\frac{B}{B+D}$	$\frac{C}{A+C}$
M2-multiplier	≥ 0.83	92.9	0.172	0.071
Domestic Credit/GDP	≥ 0.84	85.7	0.146	0.143
Real GDP	≤ 0.11	86.7	0.074	0.134
Stock Indices	≤ 0.1	92.3	0.083	0.077
Real interest rate	≥ 0.73	100	0.252	0

Table 4: Final results

Indicator	Early signal percentage
M2-multiplier	56.9%
Domestic Credit/GDP	59%
Real GDP	13.84%
Stock Indices	4.6%
Real Interest Rate	69%

Table 5: Leading quality of the indicators

est rate would be the best indicator variable, followed by the M2-multiplier and the stock indices. Bear in mind, however, that the correctly signalled crises are only one side of the medal, and that a high percentage of those goes hand in hand with a high percentage of false alarms. For instance, for the real interest rate, in more than one quarter of the tranquil months a false alarm is issued. Clearly, such an indicator is very undesirable. Then again, remember the caveat mentioned at the beginning of this chapter: the high percentage of false alarms of this indicator could also be due to successful preemptive policy interventions in the face of financial instability. Therefore, an indicator that issues a lot of false alarms should not be repudiated straightaway. The real interest rate, for instance, correctly signals every crisis in the sample. This means that the interest rate is always above average before a crisis, which is quite a strong finding.

5.4.1 How leading are the indicators?

Another way to compare the indicators is to look at the timing of the correctly issued signals. As mentioned above, early signals might be more desirable than signals that are issued after the beginning of a crisis. One can therefore rank the indicators according to the fraction of correct signals that they issue in the months before the crisis starts. This yields the results displayed in Table 5.

The real interest rate seems to be the most leading indicator, with almost 70% of the correct signals in the early half of the crisis period. Contrary, real GDP and stock indices signal mainly after the beginning of a crisis, which might also hint to the fact that there might be some degree of causality in the reverse direction, meaning that banking distress causes real GDP and stock indices to (further) decline. The lack of leading quality of the stock indices is especially interesting, as it is often argued that asset markets are very forward-looking and every information about the future is immediately reflected in the prices. The result here suggests that rather the opposite is

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Indicator	Average number of signals
M2-multiplier	5
Domestic Credit/GDP	7.4
Real GDP	10
Stock Indices	7.25
Real Interest Rate	9.71

Table 6: Persistence of the signals

true in practice, and that stock prices do not lead but rather follow the developments in the banking sector.

5.4.2 Persistence of the signals

As mentioned above, the more signals an indicator issues during a crisis period, the better. Consequently, an interesting way to compare the indicators is to look at the persistence of their signals. Table 6 reports the average number of months in which the indicator issues a signal during a crisis period if the indicator correctly signals this crisis.

We see that with respect to persistence, the real GDP and the real interest rate indicator definitely perform best. However, note that the main reason for the good performance of the real interest rate indicator is, that it generally issues a lot of signals. It correctly signals all crises, and also with a very high degree of persistence, but, as will be seen in detail later, a large fraction of the signals are false alarm. The drawback of the real GDP indicator, on the other hand, is that, although it signals very often during a crisis period, most of its signals lie in the second half of the period, so the alarm may come "too late". The M2-multiplier doesn't signal very often during crisis periods, but on the other hand it often issues at least one signal, such that very few crises are missed.

5.4.3 Matthew's correlation coefficient

Matthew's correlation coefficient is a measure of the quality of binary classification tests closely related to ROC-curves and can be used even if the classes are of very different sizes. It is defined as

$$\mathrm{MCC} = \frac{a \cdot d - b \cdot c}{\sqrt{(a+b) \cdot (a+c) \cdot (b+d) \cdot (d+c)}},$$

ranges from -1 to 1 and is basically a correlation coefficient between the observed and the predicted binary classifications. The closer the coefficient is to 1, the better the test. A coefficient of 0 represents an average random prediction and -1 an inverse prediction. For the indicator variables we get the result shown in Table 7.

According to the MCC, the real GDP indicator is the best classifier. However, none of the indicators achieves a very high coefficient. The especially low value for the

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Indicator	MCC
M2-multiplier	0.175
Domestic Credit/GDP	0.17
Real GDP	0.228
Stock Indices	0.2
Real Interest Rate	0.13

Table 7: Matthew's correlation coefficient

real interest rate reflects the noisiness of the indicator. It signals very often, and as a banking crisis is a relatively rare event, the real interest rate manages to call every crisis and therefore the fraction of missed crises is zero. However, the noise is substantial, as on average every fourth month during tranquil times a false alarm is issued. This is punished here by the MCC.

5.4.4 Comparing ROC-curves

The shape of the ROC-curve of each indicator is infomative about the general aptitude of the variable as an indicator for banking crises. Again, the curve is a graph of sensitivity against 1-specificity, which means in the present analysis a graph of $\frac{A}{A+C}$ (y-axis) against $\frac{B}{B+D}$ (x-axis). The area under the ROC-curve shows, how well the variable under study can distinguish between a crisis and tranquil times. For instance, if the area under the ROC-curve is 0.8, this means that a randomly selected observation of the variable in a crisis interval has a more extreme value than a randomly selected observation in tranquil times with a probability of 80%. If the examined variable cannot distinguish between tranquil and crisis times, the area equals 0.5. This is the worst possible value one can get and means that one might as well flip a coin to distinguish between tranquil and crisis times. An ideal ROC-curve would be steeply increasing in the beginning and very flat at the end and the area under the ROC-curve should be huge.

Figures (11) - (15) show the ROC-curves of the indicator variables (for the last analysis, with the "correct" counting of true positives and false negatives) together with the 45-degree line (which is the ROC-curve that a completely random indicator would yield). It can be concluded that for all of them the area under the curve is larger than 0.5. Note that for all variables except the M2-multiplier the curve starts at a value of $\frac{A}{A+C}$ much larger than 0. This is due to the way the indicators were set up. As they signal at least once in every country during the sample period by definition,[9] $A+B$ will never be zero and therefore, if the few signals that are issued with the tight thresholds are correct, this already results in a relatively large value of $\frac{A}{A+C}$ because $A+C$ is small. Comparing the shape of the ROC-curves, the stock indices variable shows the most favourable shape, followed by real GDP. This confirms the results in

[9] As only countries which have experienced crises during the sample period are examined, it is sensible that the indicator variables issue at least one signal once in every country. However, this requirement is not crucial for the analysis.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

terms of the errors that we get with the optimal threshold.

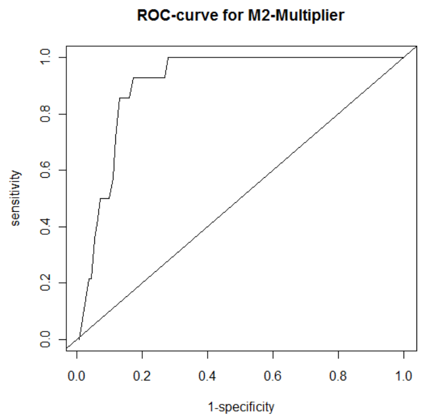

Figure 11: ROC-curve for the M2-Multiplier

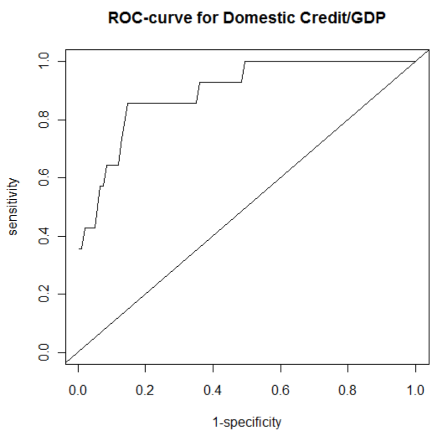

Figure 12: ROC-curve for Domestic Credit/Real GDP

5.4.5 Conditional vs. unconditional probabilities

Further information on the quality of individual indicators can be obtained by comparing the in-sample probability of occurrence of a crisis conditional on the signal of an indicator to the unconditional in-sample probability of a banking crisis.[10] For the conditional probabilities we have

$$\text{P}(\text{crisis occurs} \mid \text{indicator signals}) = \frac{\text{Number of months where indicator signals and crisis period}}{\text{Number of months where indicator signals}},$$

[10] It is immediate to see that the probability of a crisis conditional on the signal of an indicator is equal to the positive predictive value, while the unconditional probability corresponds to the prevalence of a crisis.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

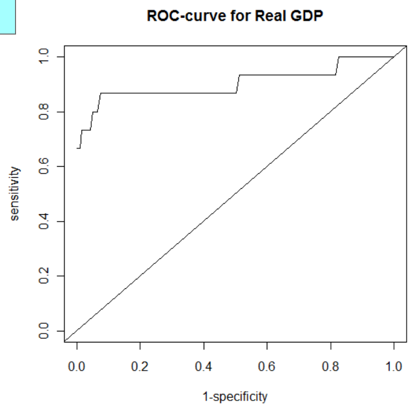

Figure 13: ROC-curve for Real GDP

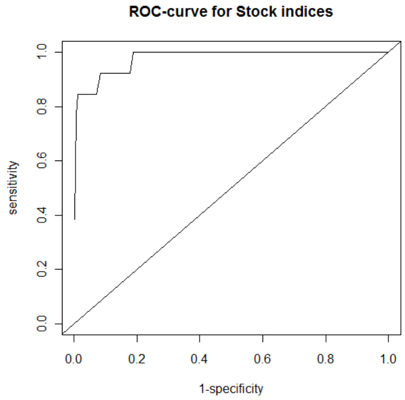

Figure 14: ROC-curve for Stock indices

whereas for the unconditional probabilies we simply get[11]

$$\text{P(crisis occurs)} = \frac{\text{Number of crisis period months}}{\text{Total number of months}}.$$

The results of this comparison are displayed in Table 8.

As can be seen, the conditional probabilities are always higher than the unconditional ones, which means that the signals of the indicators are definitely an information gain. However, for some indicators the difference is more pronounced than for others. A signal from the real GDP indicator increases the probability of a crisis by more than 28% (percentage points), while a signal from the M2-multiplier indicator yields a surprisingly small information gain and increases the probability only by 1% (percentage point). The reason for this is that the M2-multiplier indicator as defined above signals a high percentage of crises, but the number of signals issued during crisis times is small. On the other hand, it issues a wrong signal in 17% of the tranquil months. This

[11]Note that this number varies for every indicator, because sometimes different time spans were analyzed for different variables due to data availability.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

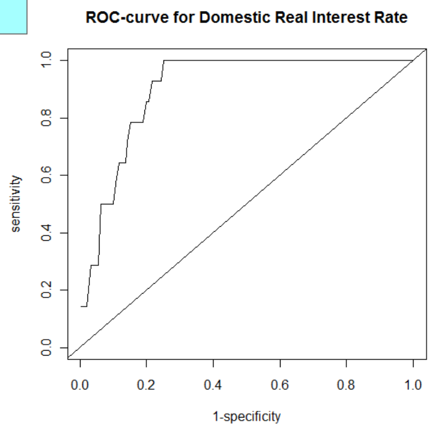

Figure 15: ROC-curve for Domestic Real Interest Rate

Indicator	Unconditional probability	Conditional probability
M2-multiplier	16.6%	17.66%
Domestic Credit/GDP	15.77%	24.3%
Real GDP	13.23%	41.67%
Stock Indices	9.9%	26.77%
Real Interest Rate	13.16%	18.92%

Table 8: Conditional and unconditional crisis probability

brings about that more than 82% of its signals are false alarms, and so a signal from the indicator is quite likely to be wrong. The result for the real interest rate is not compelling either — more than 80% of the signals are wrong. Generally, note that, due to the low crisis prevalence, for all indicators the conditional probability is below 50%, which means that it is still more likely that we are in a tranquil state if the indicator signals. Therefore, it is not advisable to rely on the information of one indicator only, but to take into account also the information from the other indicators. This leads to the idea of composite indicators, which is elaborated in detail in chapter 5.6.

Another informative measure would be the in-sample probability of being in a tranquil period, conditional on getting no signal from the indicator. We have that

$$\text{P(no crisis occurs} \mid \text{indicator doesn't signal)} =$$
$$= \frac{\text{Number of months where indicator doesn't signal and no crisis period}}{\text{Number of months where indicator doesn't signal}}$$

and

$$\text{P(no crisis occurs)} = \frac{\text{Number of tranquil months}}{\text{Total number of months}}.$$

For the five indicators, the result in Table 9 is obtained.

Again, we see that the M2-multiplier is not very informative in this respect, the conditional and the unconditional probability of not being in a crisis period are almost the same. The reason for this is that the M2-multiplier generally issues very few signals

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Indicator	Unconditional probability	Conditional probability
M2-multiplier	83.4%	83.6%
Domestic Credit/GDP	84.22%	85.88%
Real GDP	86.76%	90.28%
Stock Indices	90.1%	92%
Real Interest Rate	86.8%	89%

Table 9: Conditional and unconditional probability of no crisis

Indicator	critical region	percentage signalled	$\frac{B}{B+D}$	$\frac{C}{A+C}$
M2-multiplier	≥ 0.83	87.5	0.155	0.125
Domestic Credit/GDP	≥ 0.84	50	0.169	0.5
Real GDP	≤ 0.11	77.78	0.083	0.22
Stock Indices	≤ 0.1	83.34	0.081	0.167
Real Interest Rate	≥ 0.73	77.78	0.265	0.22

Table 10: Results for the indicators in CEE

during a crisis period (however, it often issues at least one signal, such that a lot of crises are correctly signalled). Hence, the number of correct signals is relatively low and the number of crisis months in which no signal is issued is relatively high. Therefore, the fact that no signal is issued by the M2-multiplier doesn't increase the probability of being in a tranquil state a lot. In contrast, for the real GDP and the real interest rate indicator, which both signal very often during crisis periods, the probability of being in a tranquil state is considerably higher conditional on not obtaining a signal.

5.5 Performance of the indicators out-of-sample for CEE

Next, the performance of the indicators for CEE is examined (out-of-sample). I determine $\frac{B}{B+D}$, $\frac{C}{A+C}$ and the percentage of crises correctly signalled for the optimal thresholds that result from the previous analysis. Additionally, I check whether for CEE the optimal threshold according to the criterion from above differs from the critical region for Western/Northern Europe. Note however, that I cannot say anything about the significance of the difference. The results are summarized in Tables 10 - 12.

In general, from these results it can be concluded that the M2-multiplier, real GDP and stock indices indicators perform quite well with the optimal thresholds from Western/Northern Europe, while the ratio of domestic credit to GDP and the real interest rate seem to behave quite differently in the two areas. With respect to early signals,

Indicator	optimal critical region
M2-multiplier	≥ 0.81 (100% signalled)
Domestic Credit/GDP	≥ 0.81 (75% signalled)
Real GDP	≤ 0.08 (77.78% signalled)
Stock Indices	≤ 0.2 (100% signalled)
Real Interest Rate	≥ 0.65(100% signalled)

Table 11: Optimal thresholds for CEE

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Indicator	early signals
M2-multiplier	55.77%
Domestic Credit/GDP	47.83%
Real GDP	38.57%
Stock Indices	55.5%
Real Interest Rate	55.7%

Table 12: Leading quality of the indicators in CEE

the difference between the indicator variables is not as pronounced as for Western and Northern Europe. It seems that the variables are generally more leading in CEE, with the M2-multiplier and the real interest rate sending the highest proportion of early signals, while real GDP issues the latest signals. What is striking is, that the stock indices variable, which is the least leading indicator for Western and Northern Europe, issues a relatively large fraction of signals in the early half of the crisis period in CEE.

The ROC-curves again show, how good the indicator variables can differentiate between crisis and tranquil times. From these curves one can clearly see, that the M2-

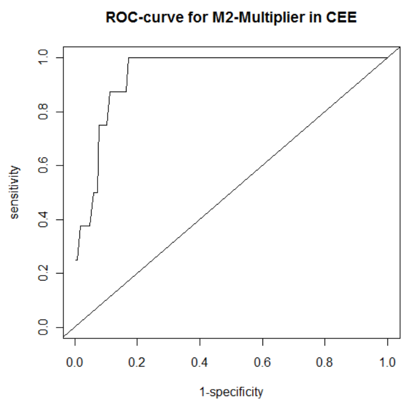

Figure 16: ROC-curve for M2-multiplier in CEE

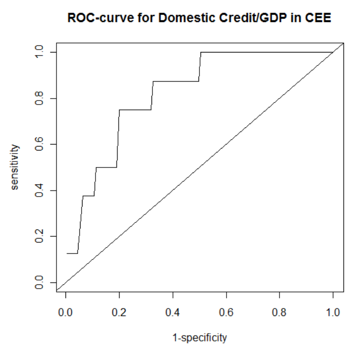

Figure 17: ROC-curve for Domestic Credit/GDP in CEE

multiplier and the stock indices are generally good indicator variables for CEE, while

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

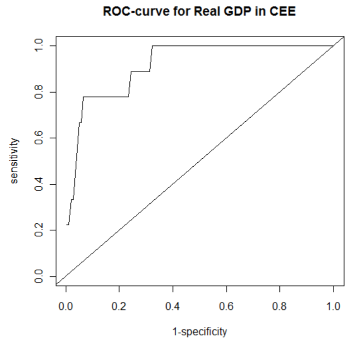

Figure 18: ROC-curve for Real GDP in CEE

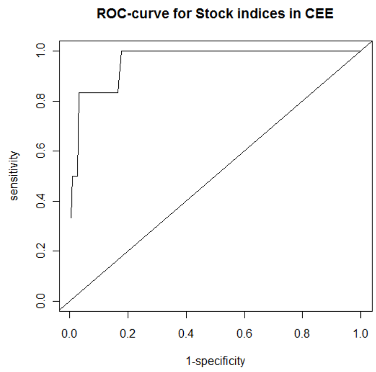

Figure 19: ROC-curve for Stock Indices in CEE

the shape of the ROC-curves of the domestic credit to GDP ratio and the real interest rate are less favourable.

Additional information is gained by looking at the conditional and unconditional probabilities of crisis and tranquil times. Table 13 shows the probability of being in a crisis period conditional on a signal from the indicator, Table 14 shows the probability of being in a tranquil period conditional on getting no signal from the respective indicator.

Like in Western and Northern Europe, the probability of being in a crisis period if the indicator signals is highest for real GDP. As was already expected from the above analysis, the ratio of domestic credit to GDP and the real interest rate do not perform very well with the Western thresholds. A signal from the real interest rate doesn't

Indicator	Unconditional probability	Conditional probability
M2-multiplier	14.58%	22.3%
Domestic Credit/GDP	15%	10%
Real GDP	12.7%	35.35%
Stock Indices	9.2%	27.69%
Real Interest Rate	12.22%	14.37%

Table 13: Conditional and unconditional crisis probability in CEE

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

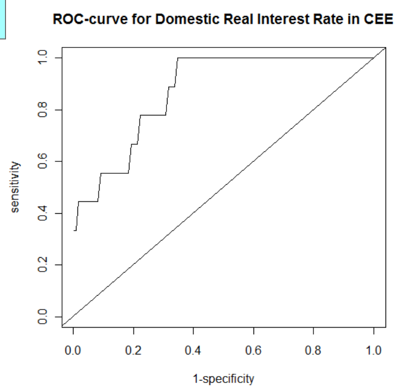

Figure 20: ROC-curve for Domestic Real Interest Rate in CEE

Indicator	Unconditional probability	Conditional probability
M2-multiplier	85.4%	87%
Domestic Credit/GDP	85%	84.2%
Real GDP	87.3%	90.15%
Stock Indices	90.8%	92.9%
Real Interest Rate	88.57%	87.77%

Table 14: Conditional and unconditional probability of no crisis in CEE

change the probability of a crisis much and the indicator is less informative than for Western and Northern Europe. The domestic credit/GDP ratio is by far the worst indicator variable — a signal from this indicator actually decreases the probability of being in a crisis. Hence, signals of this indicator are mainly noise and shouldn't be viewed as crisis warnings.

These results provide strong evidence that the ratio of domestic credit to GDP and the real interest rate behave differently in CEE compared to Western and Northern Europe during crisis times and should in any case not be used with the thresholds from Western Europe. Actually, the ROC-curves of these variables convey even more information. Their shapes suggest that the indicators cannot differentiate between crisis and tranquil times very well. This may be because the variables do not show a very homogeneous behaviour across CEE countries and are therefore not very suitable as crisis indicators for this region.

5.6 Composite Indicators

5.6.1 A crisis warning system

Having found the optimal thresholds for the individual, single-variable indicators, it is possible to combine the information from all variables into a composite indicator. Kaminsky and Reinhart don't go this far, but Kaminsky (1998) defines the composite

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

	percentage correctly signalled	$\frac{B}{B+D}$	$\frac{C}{A+C}$
$I_t \geq 1$	100%	0.5	0
$I_t \geq 2$	100%	0.13	0
$I_t \geq 3$	91.67%	0.026	0.083
$I_t \geq 4$	16.7%	0.0064	0.83
$I_t \geq 5$	0%	0.0008	1

Table 15: Performance of different specifications of a composite indicator

indicator I_t as

$$I_t = \sum_{j=1}^{n} S_t^j,$$

where n is the number of single-variable indicators and S_t^j is equal to 1 if indicator j issues a signal in period t and 0 else.

For my analysis, this means that the value of the composite indicator in period t ranges from 0 to 5, depending on the number of indicators that issue a signal. Now we can determine for every j from 1 to 5 the predictive quality of the composite indicator if we define it as issuing a signal if $I_t \geq j$. This yields the results documented in Table 15.

From this table, the trade-off between $\frac{B}{B+D}$ and $\frac{C}{A+C}$ is clearly visible. If the composite indicator is defined as giving a signal as long as one of the individual indicators signals, every crisis is called, but the indicator signals in 50% of the tranquil months. Increasing the number to two is unambiguously better: again all of the crises are signalled, but now the percentage of false alarms during non-crisis months is only 13%. If the criterion from above (minimize $\frac{B}{B+D} + \frac{C}{A+C}$) is applied to choose the optimal indicator, this would be the specification in the third line, where a signal is issued if $I_t \geq 3$. Now, only 91.67% of the crises are correctly predicted, but the percentage of false alarms during tranquil times has been considerably reduced to 2.6%. Moving to the specification $I_t \geq 4$, $\frac{B}{B+D}$ becomes very small, but 83% of the crises are missed. Finally, with $I_t \geq 5$, $\frac{B}{B+D}$ is negligible, but none of the crises is predicted, which is definitely a very undesirable result.

This composite indicator can be used as a crisis warning system. More specifically, we can compute the probability of being in a crisis period (resp. in a period leading up to a crisis) conditional on the indicator having a specific value. We have

$$\mathrm{P}(\text{crisis occurs} \mid I_t \geq j) = \frac{\text{Number of months where } I_t \geq j \text{ and crisis period}}{\text{Number of months where } I_t \geq j}$$

and

$$\mathrm{P}(\text{crisis occurs} \mid I_t = j) = \frac{\text{Number of months where } I_t = j \text{ and crisis period}}{\text{Number of months where } I_t = j}.$$

The results are shown in Table 16.

Some additional explanation of the table is needed here. Going from $I_t \geq 1$ up to $I_t \geq 3$, the conditional probabilities increase, but then they start decreasing again.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

	conditional probability of crisis		conditional probability of crisis
$I_t \geq 1$	27.45%	$I_t = 1$	17.79%
$I_t \geq 2$	45.25%	$I_t = 2$	37.96%
$I_t \geq 3$	62.92%	$I_t = 3$	68.35%
$I_t \geq 4$	20%	$I_t = 4$	22.22%
$I_t \geq 5$	0%	$I_t = 5$	0%

Table 16: Conditional and unconditional crisis probability for the composite indicator

This is puzzling at first sight, because one might think that the more indicators signal, the more likely it is that a crisis occurs. However, the result that is found here is due to the different timing of the individual indicators. As some of them, like the real interest rate or the ratio of domestic credit to GDP, are very leading, and some, like stock indices and real GDP, typically signal in the second half of the crisis interval, it is usually not the case that all of them issue a correct signal in the same month. If this happens, which is very rare (there is only one month in the sample for which $I_t = 5$), it is mainly a coincidence and rather a sign of noise during tranquil times than a sign of an upcoming crisis. Hence, a value of the composite indicator which is larger or equal to four shouldn't be regarded as a warning signal, it is more likely to be just a wrong signal.

Note that the unconditional probability of the occurrence of a crisis is just

$$\mathrm{P(crisis\ occurs)} = \frac{\text{Number of crisis period months}}{\text{Total number of months}},$$

which is in our case 18.64%. As the conditional probabilities are generally higher (except for $I_t = 1$), looking at the value of the combined indicator can definitely be informative for policymakers. Especially if the combined indicator is exactly equal to 3, the probability of a crisis is 68.35%, which is definitely much higher than the unconditional probability of 18.64% and will therefore be a serious sign of distress.

5.7 Two-variable indicators

For the composite indicators of the previous section, it doesn't matter which variables issue a signal, only the overall number counts. Another possible method to obtain better indicators than the single-variable ones is to use combinations of two variables as indicators. The advantage compared to the composite indicators is, that signals from the two-variable indicators are more informative about the macroecomic surroundings, as it is taken into account exactly which variables signal. Moreover, two simultaneous signals from indicator variables with a similar timing are most likely more instructive than simultaneous signals from all the indicators with a very heterogeneous timing. In combining the indicators, it seems unreasonable to stick to the cutoff-percentiles that were found to be optimal for the single-variable indicators, because most likely if we look at two variables jointly, other percentiles will yield better results. Therefore, in the following analysis I allow the thresholds to differ from the optimal thresholds of the

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

single-variable indicators.

Table 17 and 18 display the quality of the resulting new indicators, which are the combination of two variables, with respect to correctly predicting crises, error rates and early signal percentage. The indicator is defined as issuing a signal if both variables lie above some percentiles of their sample distributions and the optimal threshold (which can be a different one for each variable) is determined by a gridsearch over critical regions up to a maximum of 30 percent.[12] The results are quite satisfactory: for some combinations, the percentage of correctly predicted crises is huge, while the errors can be considerably reduced compared to the single-variable indicators.

As can be seen from the tables, the indicator which combines stock indices and real GDP correctly calls 91.7% of the crises, with very small error rates. This corroborates the hypothesis that banking crises are often preceded by slumps in economic activity and asset prices. However, the leading quality of this indicator is terrible, with all of the signals lying in the second half of the crisis period. This result doesn't come unexpectedly, as both individual variables, and especially the stock indices, show a very bad leading quality.

Combining stock indices and the M2-multiplier works well in terms of predicting crises and also the combination of domestic credit/GDP and stock indices seems to be very good. This provides evidence for the assumption that the banking system is more vulnerable to asset price busts in times of a credit boom. The indicator which is a combination of domestic credit/GDP and real GDP performs quite well, which also confirms our previous claim that excessive credit growth followed by a recession can lead to banking distress. Again, the leading quality of these indicators is not overwhelming.

With respect to early warnings, it is not surprising that combining two indicator variables that are very leading yields satisfactory results. The combination of M2-multiplier and the real interest rate, 80% of the signals are early. On the other hand, the indicator misses half of the crises. For the combination of the real interest rate with domestic credit/GDP, 82% of the issued signals lie in the time intervals before the beginning of a crisis. The percentage of correctly predicted crises is also quite high for this indicator. This supports the hypothesis that high real interest rates can put the banking system into distress especially during a credit boom.

As for the single-variable indicators, also the comparison of the two-variable indicators can be extended to a further dimension: the conditional crisis and non-crisis probability. How likely is it to be really in a tranquil state, if the respective indicator doesn't signal, and how high is the probability that a signal issued by an indicator is true? This analysis yields some interesting additional insights into the informative content of the indicators. The results are shown in table 19 (the unconditional probabilites are in brackets). It is immediate to see, that for most variable combinations the probabilities that a signal is correct are substantially higher than for the single-variable indicators. With respect to this criterion again the combination of stock indices and

[12] The optimality criterion is, as before, to maximize the sum of specificity and sensitivity.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Indicator	critical region	percentage signalled
M2-multiplier and Domestic Credit/GDP	$\geq 0.71 \wedge \geq 0.71$	70%
M2-multiplier and Real interest rate	$\geq 0.72 \wedge \geq 0.74$	50%
Domestic Credit/GDP and Stock Indices	$\geq 0.83 \wedge \leq 0.16$	77%
Domestic Credit/GDP and Real GDP	$\geq 0.76 \wedge \leq 0.08$	84.6%
Domestic Credit/GDP and Real Interest Rate	$\geq 0.78 \wedge \geq 0.76$	75%
Real Interest Rate and Real GDP	$\geq 0.73 \wedge \leq 0.15$	66.7%
Stock Indices and Real GDP	$\leq 0.02 \wedge \leq 0.14$	91.7%
Stock Indices and Real Interest Rate	$\leq 0.3 \wedge \geq 0.78$	50%
Stock Indices and M2-multiplier	$\leq 0.3 \wedge \geq 0.71$	91.7%
M2-multiplier and Real GDP	$\geq 0.71 \wedge \leq 0.2$	91.7%

Table 17: Performance of two-variable indicators

Indicator	$\frac{B}{B+D}$	$\frac{C}{A+C}$	early signals
M2-multiplier and Domestic Credit/GDP	0.097	0.3	36.7%
M2-multiplier and Real interest rate	0.064	0.5	80%
Domestic Credit/GDP and Stock Indices	0.02	0.23	7%
Domestic Credit/GDP and Real GDP	0.014	0.15	20%
Domestic Credit/Real GDP and Real Interest Rate	0.05	0.25	82%
Real Interest Rate and Real GDP	0.026	0.34	43.8%
Stock Indices and Real GDP	0.003	0.084	0%
Stock Indices and Real Interest Rate	0.074	0.5	15%
Stock Indices and M2-multiplier	0.086	0.084	13.4%
M2-multiplier and Real GDP	0.045	0.084	21.6%

Table 18: Error rates and early signal percentage of two-variable indicators

real GDP performs best. A signal from this indicator raises the probability of being in a crisis period to almost 90%, so it can be regarded really as a serious sign of alarm. Also the combination of domestic credit/GDP and real GDP is a very reliable indicator in this sense — 75% of the signals are correct. Besides, for this indicator the probability of not being in a crisis conditional on not getting a signal is the highest, with 85%. Overall, it seems that combinations of variables, which have already a high positve predictive value as single-variable indicators, yield the most favourable conditional probabilities, whereas the worst indicator in this respect is definitely the combination of M2-multiplier and real interest rate, which have both very low conditional probabilities as individual indicators. A signal from this joint indicator actually decreases the probability of being in a crisis period. The same holds for the combinations M2-multiplier and domestic credit/GDP and stock indices and real interest rate. These joint indicators generally do not perform very well, with relatively high error rates, and therefore do not seem to be very useful for policy makers.

5.7.1 Two-variable indicators for CEE

Table 20 shows the performance of the two-variable indicators as defined above for CEE. Combining the M2-multiplier with the stock indices and with real GDP yields the best

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Indicator	conditional crisis probability	conditional non-crisis probability
M2-multiplier and Domestic Credit/GDP	17.65%(17.76)	82.22%(82.24)
M2-multiplier and Real interest rate	10.98%(18.64)	80.88%(81.36)
Domestic Credit/GDP and Stock Indices	49.25%(18.64)	83.32%(81.36)
Domestic Credit/GDP and Real GDP	75.61%(17.76)	85.06%(82.24)
Domestic Credit/GDP and Real Interest Rate	40.9%(18.64)	83.07%(81.36)
Real Interest Rate and Real GDP	49.23%(18.64)	82.7%(81.36)
Stock Indices and Real GDP	88.24%(18.64)	82.92%(81.36)
Stock Indices and Real Interest Rate	17.7%(18.64)	81.29%(81.36)
Stock Indices and M2-multiplier	21.73%(18.64)	81.66%(81.36)
M2-multiplier and Real GDP	39.36%(18.64)	82.7%(81.36)

Table 19: Conditional crisis and non-crisis probabilities (unconditional probabilities in brackets)

results in terms of correctly signalled crises. This is not surprising, as these indicator variables work well as single-variable indicators already. The combinations also perform well in Western and Northern Europe, so it seems that these joint indicators are suitable for Europe in general. Besides, in CEE they are even relatively leading (contrary to Western and Northern Europe).

The combination of stock indices and the real interest rate also works well, with 80% of the crises correctly signalled. Hence, the performance of this indicator is much better than for Western and Northern Europe, where it only signals 50% of the crises.

The worst indicator is definitely the combination of domestic credit/GDP and the real interest rate. It calls only 20% of the crises correctly and none of the signals are in the early half of the crisis period. However, the result was to be expected, as the two single-variable indicators do not perform very well for CEE, either.

With respect to leading quality, it is remarkable that the combination of domestic credit/GDP and stock indices is by far the most leading indicator in CEE, with more than 80% of the signals in the early crisis months, whereas this indicator has only 7% early signals in Western and Northern Europe. This is mainly a result of the better leading quality of the stock indices for CEE. In contrast, the combination of domestic credit/GDP and the real interest rate issues no early signals in CEE, while for Western and Northern Europe it is the most leading indicator, with an early signal percentage of 82%. Again, as for the single-variable indicators, the difference between indicators with respect to early signals is not as pronounced as for Western and Northern Europe and the indicators generally seem to be more leading in CEE.

Table 21 shows the conditional crisis and non-crisis probabilities for the indicators. The combinations of stock indices and of the M2-multiplier with real GDP yield the highest positive predictive values with 43.75% and 43.24%, respectively. The latter indicator also exhibits the highest probability of being in a tranquil state, conditional on not getting a signal. Not surprisingly, the combination of domestic credit/GDP and the real interest rate performs worst also with respect to this criterion. A signal from

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Indicator	$\frac{B}{B+D}$	$\frac{C}{A+C}$	percentage of crises correctly signalled	early signals
M2-multiplier and Domestic Credit/GDP	0.067	0.6	40%	40%
M2-multiplier and Real interest rate	0.05	0.4	60%	28.57%
Domestic Credit/GDP and Stock Indices	0.03	0.6	40%	83.3%
Domestic Credit/GDP and Real GDP	0.023	0.6	40%	40%
Domestic Credit/Real GDP and Real Interest Rate	0.06	0.8	20%	0%
Real Interest Rate and Real GDP	0.038	0.4	60%	31.25%
Stock Indices and Real GDP	0.01	0.6	40%	42.86%
Stock Indices and Real Interest Rate	0.07	0.2	80%	37.5%
Stock Indices and M2-multiplier	0.08	0	100%	39.13%
M2-multiplier and Real GDP	0.05	0	100%	53%

Table 20: Performance of two-variable indicators for CEE

Indicator	conditional crisis probability	conditional non-crisis probability
M2-multiplier and Domestic Credit/GDP	7.93%(10.35)	89.47%(89.65)
M2-multiplier and Real interest rate	13.73%(10.35)	89.85%(89.65)
Domestic Credit/GDP and Stock Indices	17.65%(10.35)	89.93%(89.65)
Domestic Credit/GDP and Real GDP	20.83%(10.35)	89.93%(89.65)
Domestic Credit/GDP and Real Interest Rate	3.8%(10.35)	89.26%(89.65)
Real Interest Rate and Real GDP	34.04%(10.35)	90.93%(89.65)
Stock Indices and Real GDP	43.75%(10.35)	90.24%(89.65)
Stock Indices and Real Interest Rate	21.62%(10.35)	90.64%(89.65)
Stock Indices and M2-multiplier	24.21%(10.35)	91.25%(89.65)
M2-multiplier and Real GDP	43.24%(10.35)	92.54%(89.65)

Table 21: Conditional crisis and non-crisis probabilities in CEE (unconditional probabilities in brackets)

this indicator actually decreases the probability that a crisis occurs. This also holds for the combination of the M2-multiplier and domestic credit/GDP. Clearly, as the ratio of domestic credit to GDP is a very poor indicator variable for CEE, combinations where this variable is involved do not yield very satisfactory results, either. However, in general jointly using the information of two individual indicators seems to be a promising method also for CEE. For some combinations, the predictive quality and the precision can be considerably increased compared to the single-variable indicators.

6 Logit Model approach

Alternative to the non-parametric approach from above, the relationship of macroeconomic variables and the risk of a banking crisis can be examined in the context of a logit model. It is designed to identify the conditions under which one observes a banking crisis. The dependent variable y takes on the value 1 if there is a banking crisis and 0 in tranquil times.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

The explanatory variables $\mathbf{x}$ explain y such that

$$\begin{aligned}\Pr(y &= 1|\mathbf{x}) = F(\mathbf{x}, \beta) \\ \Pr(y &= 0|\mathbf{x}) = 1 - F(\mathbf{x}, \beta),\end{aligned}$$

where the logistic distribution is used to model the probabilities,

$$\Pr(y = 1|\mathbf{x}) = \Lambda(\mathbf{x}'\beta) = \frac{e^{\mathbf{x}'\beta}}{1 + e^{\mathbf{x}'\beta}}.$$

To stay consistent with the previous analysis, the indicator variables from above are used as explanatory variables, namely the M2-multiplier, the ratio of domestic credit to real GDP, real GDP itself, stock indices and the real interest rate. Again, except for the real interest rate, I look at twelve-month growth rates. Contrary to before, where high-frequency data was used, annual data is analyzed here and one lag is included to account for dynamic interaction effects of the explanatory variables.

When panel data is used, country-specific fixed effects are often included in the regression. In the context of a logit model, this is not without complications. Especially, it is only possible if the examined event (here: the banking crisis) happens in one and only one period per country.[13] However, for the analyzed sample this is generally not the case. There are some countries which have experienced more than one banking crisis during the observed time span and some countries, which have experienced no crisis at all (because the time interval for which all indicator variables are available jointly doesn't cover a banking crisis period) and serve as a useful control group. Hence, including country-specific fixed effects would mean to exclude a large amount of available information. Therefore, the model is estimated here using the full sample but without fixed effects.

The vector β reflects the impact of changes in $\mathbf{x}$ on the probability that $y = 1$. A positive coefficient β_i on a particular explanatory variable $\mathbf{x}_i$ means the greater is the realization of $\mathbf{x}_i$, the more probable is the occurrence of $y = 1$ rather than 0. Therefore, β is informative about the influence of the explanatory variables (the indicators) on the probability that a crisis happens. The model can be estimated by maximum likelihood. Once the parameters are estimated, it is possible to calculate the probabilites of occurrence of a banking crisis ($y = 1$) both in-sample and out-of-sample.

Judged against the indicator approach above, this method has one main advantage: it combines the information from all the indicator variables into one number — the probability of the occurrence of a banking crisis. All indicators are considered simultaneously and the variables that do not contribute information that is independent of the information of the other variables are disregarded. On the other hand, the quality of every individual indicator variable cannot be assessed in a satisfactory way. In the regression, the variable is either significant or not, but it is hard to compare two variables to each other. Of course, the higher the level of significance, the more reliable the

[13] see Greene (2002)

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

variable seems to be as an indicator. However, one cannot say whether the strength of the indicator lies in correctly calling many crises at the expense of sending a lot of false alarms, or whether it misses a lot of crises but seldomly issues false alarms. Besides, in this framework it is hard to say which variable shows abnormal behaviour, making it difficult for policy makers to decide on preemptive action. Also, due to the nonlinear specification, the interpretation of the estimated coefficients is not straightforward and, unlike in linear models, the marginal effect of a change in one indicator variable on the probability of a banking crisis depends on the other explanatory variables. In general,

$$\frac{\partial E[y|\mathbf{x}]}{\partial \mathbf{x}} = \left\{ \frac{dF(\mathbf{x}'\beta)}{d(\mathbf{x}'\beta)} \right\} \beta = f(\mathbf{x}'\beta)\beta.$$

For the logit model we have

$$\frac{d\Lambda(\mathbf{x}'\beta)}{d(\mathbf{x}'\beta)} = \frac{e^{\mathbf{x}'\beta}}{(1+e^{\mathbf{x}'\beta})^2} = \Lambda(\mathbf{x}'\beta)[1-\Lambda(\mathbf{x}'\beta)],$$

therefore we get

$$\frac{\partial E[y|\mathbf{x}]}{\partial \mathbf{x}} = \Lambda(\mathbf{x}'\beta)[1-\Lambda(\mathbf{x}'\beta)]\beta.$$

It is immediate to see that these marginal effects vary with the values of $\mathbf{x}$.

In the following, different specifications of a logit model are estimated for the data from Western/Northern Europe by maximum likelihood. Wald-tests are performed to determine the significance level of every individual explanatory variable. Additionally, the quality of the model is assessed by looking at the Akaike Information Criterion (AIC). Furthermore, the fit of the model is described by the percentage of correct predictions of crisis/non-crisis periods. Finally, I analyze the suitability of the model for CEE by looking at the percentage of crisis/non-crisis periods correctly predicted out-of-sample.

Estimating the model with all five indicator variables and a one-period lag used as explanatory variables yields the following result:

```
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-1.9805  -0.2896  -0.1150  -0.0438   1.8955

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.855840   1.216465  -3.170  0.00153 **
m2mult1     -0.008708   0.019540  -0.446  0.65584
lagm2mult   -0.015382   0.013975  -1.101  0.27104
domcred1     0.121074   0.083014   1.458  0.14471
lagdomcred  -0.096518   0.079970  -1.207  0.22746
realgdp1    -0.520445   0.253554  -2.053  0.04011 *
lagrealgdp   0.207977   0.251032   0.828  0.40740
```

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

```
stockind1    -0.026576   0.028937  -0.918  0.35839
lagstockind   0.079883   0.028246   2.828  0.00468 **
realint1     -1.330909   0.605719  -2.197  0.02800 *
lagrealint    1.032406   0.499334   2.068  0.03868 *
---
Signif. codes:  "***" 0.001  "**" 0.01  "*" 0.05  "." 0.1  " " 1

AIC: 66.338

Number of Fisher Scoring iterations: 7

Number of observations: 127
% total correct: 89.8
% crises correct: 18.2
% no-crises correct: 96.5

Out-of-sample prediction (CEE):
Number of observations: 78
% total correct: 70.5
% crises correct: 0
% no-crises correct: 73.4
```

Concerning the estimation results for individual variables, we find that real GDP enters the regression with a negative coefficient, as we would expect, both from theory and from the previous analysis, and the coefficient is significant at the 5% level. The coefficient on the stock index variable is not significant at conventional levels, but it carries the expected sign. The coefficient on the lag is significant and enters with a positive sign. This fits the story that an asset price bust following a period of high stock returns often causes banking distress. The coefficients for the M2-multiplier are not significant and besides they are negative, which contradicts the findings from the previous analysis. Also, the coefficient on the domestic credit/GDP ratio is not significant, but it is close to being significant and carries the correct sign. Concerning the real interest rate, the outcome is puzzling at first. The coefficient on the real interest rate is significant but carries the "wrong" sign according to the above analysis. However, the coefficient on the lag is positive and also significant. This fits the result from the previous chapter that the real interest rate is a very leading indicator variable. The negative coefficient on the current real interest rate could be due to policy interventions that are implemented in the face of an impending crisis, which is accompanied by a recession. To combat the economic downturn, the nominal interest rate might be lowered, leading also to a decrease in the real interest rate.

We see that for non-crisis periods reasonable predictive power has been obtained. More than 96% of the non-crisis periods are predicted correctly. Nevertheless, the predictive power for the (relatively few) crisis periods is not very good — only 18% of those are predicted correctly. Out-of-sample, 70% of the observations are predicted correctly, for non-crisis times this number is 73%. However, none of the crises in CEE

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

was predicted correctly.

Next, the model is estimated without including the M2-multiplier:

```
Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.87306  -0.29198  -0.12504  -0.04993   2.04975

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.69794    1.14574  -3.228  0.00125 **
domcred1     0.07007    0.07374   0.950  0.34199
lagdomcred  -0.07105    0.07586  -0.937  0.34895
realgdp1    -0.52050    0.23954  -2.173  0.02979 *
lagrealgdp   0.23798    0.24478   0.972  0.33094
stockind1   -0.02733    0.02855  -0.957  0.33848
lagstockind  0.07217    0.02649   2.724  0.00645 **
realint1    -1.11642    0.54197  -2.060  0.03940 *
lagrealint   0.84128    0.44541   1.889  0.05892 .
---
Signif. codes:  "***" 0.001  "**" 0.01  "*" 0.05  "." 0.1  " " 1

AIC: 63.844

Number of Fisher Scoring iterations: 7

Number of observations: 127
% total correct: 90.6
% crises correct: 27.3
% no-crises correct: 96.6

Out-of-sample prediction (CEE)
Number of observations: 78
% total correct: 73
% crises correct: 0
% no-crises correct: 76
```

The AIC of the model has decreased, most of the coefficients have gained significance and all coefficients except the one for the real interest rate carry the expected sign now. The rate of crises correctly predicted in-sample has increased to 27%, and also the predictive quality out-of-sample has improved.

Finally, the model is estimated without the M2-multiplier and domestic credit:

```
Deviance Residuals:
     Min        1Q    Median        3Q       Max
-1.73094  -0.30002  -0.15930  -0.03919   2.40998
```

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

```
Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -3.46884    1.00168  -3.463 0.000534 ***
realgdp1    -0.52776    0.23290  -2.266 0.023445 *
lagrealgdp   0.20102    0.22530   0.892 0.372271
stockind1   -0.02693    0.02649  -1.016 0.309433
lagstockind  0.07305    0.02508   2.913 0.003582 **
realint1    -1.04975    0.50916  -2.062 0.039233 *
lagrealint   0.75043    0.41383   1.813 0.069770 .
---
Signif. codes:  "***" 0.001  "**" 0.01  "*" 0.05  "." 0.1  " " 1

AIC: 61.317

Number of Fisher Scoring iterations: 7

Number of observations: 127
% total correct: 91.3
% crises correct: 27.3
% no-crises correct: 97.4

Out-of-sample prediction (CEE)
Number of observations: 78
% total correct: 96.1
% crises correct: 0
% no-crises correct: 100
```

Both the significance levels of the coefficients and the predictive quality have improved further.

Note, that for determining the predictive quality of the models, the probability of occurrence of a crisis was computed and if this probability exceeded 50%, this was regarded as the prediction of a crisis. The vector of predicted crises was then compared to the actual crisis vector. Defining the prediction of a crisis this way seems plausible: if the probability of a crisis is higher than 50%, it is more likely that a crisis will occur, than that it will not occur. However, as the prevalence of a crisis is relatively small, policy makers might see signs of alarm and the need for interventions already if the predicted crisis probability is below 50%. The coefficients that result from the logit model estimation yield the crisis probabilities, while it is in fact left open to the policy makers to decide, when to step in. Therefore, the above computed predictive qualities of the models should just be seen as benchmark values. Whether a lot of crises are missed or a high percentage of false alarms arises, depends in the logit model on the assessment of the policy maker. If the same loss function as for the indicator approach is assumed, which gives equal weight to false alarms and missed crises, the optimal probability threshold for policy intervention would be much lower than 50%. For instance, for the specification without the M2-multiplier and domestic credit, the

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

probability calling for policy action would be 13%. This is only a little bit higher than the unconditional crisis probability, which is here 9.1%. With this low probability threshold, the total percentage of correct predictions (crisis and non-crisis) is 86.81% and hence lower than with the 50%-threshold. However, the percentage of correctly predicted crises increases to 91%, while the percentage of correctly predicted non-crisis periods decreases to 86.6%.

As already explained above, the model cannot be estimated with fixed effects in this context. However, random effects can be included. The estimation yields the following results:[14]

```
                 Value Std.Error  DF   t-value p-value
(Intercept) -12.703527 2.8220184 105 -4.501575  0.0000
m2mult1      -0.028465 0.0155128 105 -1.834904  0.0694
lagm2mult    -0.072512 0.0166086 105 -4.365952  0.0000
domcred1      0.312507 0.0824766 105  3.789037  0.0003
lagdomcred   -0.581912 0.0998736 105 -5.826480  0.0000
realgdp1     -1.219760 0.2392287 105 -5.098720  0.0000
lagrealgdp    0.632174 0.2394542 105  2.640064  0.0096
stockind1    -0.123404 0.0328857 105 -3.752513  0.0003
lagstockind   0.297666 0.0498357 105  5.972949  0.0000
realint1     -3.898681 0.7518035 105 -5.185771  0.0000
lagrealint    4.082654 0.7214277 105  5.659131  0.0000

Number of observations: 127
Number of groups: 12

% total correct: 93.7
% crises correct: 54.54
% no-crises correct: 97.4
```

Although the magnitudes of the coefficients have changed, the signs remain the same. However, all coefficients are now highly significant (except for the coefficient of the M2-multiplier, which is only significant at the ten percent level) and the predictive performance of the model has improved — more than half of the crises are correctly predicted.

In the interpretation of the results of the logit model estimation, some caveats have to be considered. One issue, that was already shortly mentioned concerning the indicator approach, but of course applies to the logit approach as well, is the problem of how to account for the possibility of effective policy interventions. In the face of an upcoming crisis, an indicator might issue a (correct) signal, which is counted in

[14] Of course, if random effects for the specific countries are estimated, out-of-sample prediction for CEE doesn't make much sense anymore.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

the analysis as a false alarm, because policy makers have successfully reacted to the signs of distress and prevented the crisis. The same holds for the logit model: the probability of occurrence of a crisis might be high, but a crisis doesn't occur, due to preemptive policy measures. Additionally, these policy measures could be mirrored in the data. For instance, as described above, the central bank might lower the nominal interest rate to combat a recession precedeing a banking crisis, and this might lead to a decline in the real interest rate. The negative coefficient on the current real interest rate in the logit model might therefore not reflect the effect of the real interest rate on the probability of occurrence of a banking crisis. Rather, it could result from already implemented policy measures to combat banking crises and its concomitants. Besides, for the logit model a further problem arises. A banking crisis in one country may affect the stability of the banking system in another country. This spillover effect might lead to cross-sectional dependence in the response variable. Hence, the results of the logit model estimations should be interpreted with caution.

7 Summary

This paper analyzes the macroeconomic surroundings of banking crises in Europe using two different approaches: the non-parametric indicator approach and the estimation of a logit model. Both methods yield informative results, but from the point of view of policy makers, the indicator approach seems to be the more promising direction for further analyses. There are two main reasons for this. The first is a rather technical one: as explained above, it is very hard to assess the quality of the individual indicator variables in the context of a logit model. The variables are either significant or not, but it is hard to compare significant variables to each other and we do not see, how good the individual variables can discern crises and tranquil times. Second, also a well-specified model, which includes only the best indicator variables, provides limited information for policy makers. The coefficients estimated in the logit model indicate by their sign, how developments in the respective variables affect the probability of a banking crisis. Furthermore, the parameters can be used to compute the probability of occurrence of a crisis. However, this probability will most of the time not be very helpful, as it is hard to find out, which of the variables behave abnormally. The probability alone doesn't tell which sector of the economy is getting out of hand, therefore no clear policy recommendation will come out of the model.

Contrary, the indicator approach gives a very clear picture of the ability of every individual variable to predict banking crises. A signal of an indicator is very informative: we know, how noisy the indicator is and the approximate probability that a crisis occurs, if the indicator gives a signal. Besides, which variable signals can tell us, where in the economy the problem might lie. A high growth rate of the M2-multiplier, for instance, might point to the danger of a highly leveraged banking system and be a sign that more prudent regulations are needed. Besides, the information of various indicators can be combined to create a very detailed picture of the state of the economy.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

For Western and Northern Europe, the stock indices and the real GDP show very desirable features as indicators. They correctly predict a large fraction of crises, while the error rates are relatively small. Moreover, their signals are very persistent, and the crisis probability increases a lot conditional on a signal. The only drawback is that these indicators are not very leading, so a correct (and correctly interpreted) signal might actually come too late for efficient policy interventions.

The M2-multiplier also shows good indicator qualities in Western and Northern Europe. It is more noisy than the former, but on the other hand it is very leading. Furthermore, the ratio of domestic credit to GDP issues a high fraction of early signals, but it is very noisy, due to the heterogeneous behaviour of the variable across countries. The lack of ability of this variable to differentiate between crisis and tranquil times is also clearly visible from the ROC-curve.

At first sight, the real interest rate seems to be a very good indicator variable for Western and Northern Europe. It never misses a crisis, the signals are very persistent and a high fraction of the signals are in the early half of the crisis period. However, as mentioned above, the noise is substantial: On average, the indicator issues a false alarm in one out of four months during tranquil times. Then again, as illustrated before, the high percentage of false alarms could be due to successful preemptive policy measures, which is why an indicator that issues a lot of false alarms should not be rejected straightaway.

As banking crises are thankfully very rare events, a single signal, even if coming from a very good indicator, is still very likely to be false alarm. Therefore, an obvious improvement is to combine the information coming from the individual indicators and thereby make it more precise. Two ways of doing this were introduced in the present paper: composite indicators and two-variable indicators. The former approach looks at the number of single-variable indicators that issue a signal, while it is not important, exactly which of them signal. As demonstrated in section 5.6.1, this only makes sense up to a certain number, as the different timing of the individual indicators renders it very unlikely that all of them signal correctly in the same month. The latter approach looks at specific combinations of two indicator variables and is therefore more informative about the macroeconomic background that causes the signal. Besides, simultaneous signals from two variables with a similar timing are likely to be more precise and instructive than simultaneous signals from variables with a very heterogeneous timing. Not surprisingly, combining the two most precise single-variable indicators — stock indices and real GDP — yields the best result in terms of correctly predicted crises and noise. However, like the individual indicators, also the combination has a very bad leading quality. Joining each of the two variables with the M2-multiplier keeps the percentage of correctly predicted crises constant, but increases the noise. The conditional crisis probabilities decrease substantially. On the other hand, the fraction of early signals rises, although it is still small. Combinations of two relatively leading variables like the M2-multiplier, the real interest rate and domestic credit/GDP yield indicators that are again quite leading. However, combinations where the real interest

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

rate is involved are generally very noisy. This suggests once more, that this variable is not well suited as an indicator for banking crises. The remarkable performance of the combination of domestic credit/GDP and stock indices provides support for the theory that excessive credit growth renders the banking system especially vulnerable to asset price busts, while the performance of domestic credit/GDP and real GDP provides evidence for the claim that crises are often preceded by credit booms and recessions. Overall, it can be concluded that combining two variables into a new indicator can produce very satisfactory outcomes. The right combinations lead to a reduction in noise and a precision of the prediction.

The macroeconomic surroundings of banking crises in CEE were analyzed out-of-sample. The main reason for this is, that the evolution of the variables is quite heterogeneous already within the Eastern European countries, partly chaotic and often very different in size to Western and Northern Europe. Besides, for some countries and some variables, data was only available for a short, maybe non-representative time period. Including the countries in the sample for defining the indicators would have introduced a lot of noise and would have made it difficult to identify any common features of the developments of the variables across countries. Estimating the logit model for all countries jointly yields very bad results, with not a single significant coefficient. Therefore, I decided to define the indicators using only the data from Western and Northern Europe, and afterwards analyze, how well these indicators work for CEE. From the results it can be concluded that the M2-multiplier, real GDP and stock indices behave similarly in CEE and Western/Northern Europe. The ratio of domestic credit to GDP and the real interest rate seem to behave differently and are much noisier for CEE, although the error rate is quite high in Western/Northern Europe already. Also the ROC-curves of these two variables for CEE show that they are ill-suited indicators. Among the two-variable indicators it is therefore also no surprise, that combinations of the other three variables, especially stock indices and M2-multiplier and real GDP and M2-multiplier, perform best, with all crises correctly predicted and relatively little noise. Remarkably, both the single-variable and the two-variable indicators are on average more leading than in Western/Northern Europe. Especially stock indices and combinations with the stock indices variable included have a much better leading quality in CEE compared to Western and Northern Europe.

8 Conclusion

This paper examines the aptitude of five macroeconomic variables as indicators for banking crises. The indicators are compared to each other with respect to various criteria. Which variables or combinations of variables are most advantageous for policy makers cannot be assessed without knowing the utility function of the decision-making authorities. No indicator, neither single-variable nor composite, performs unanimously best. Some indicators are very leading, or they correctly signal a high percentage of crises, others issue very few false alarms or yield a high conditional crisis probability.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

Which of them are most desirable depends on the weight the policy maker assigns to the various criteria.

Although the present analysis is quite extensive, there are some important issues to keep in mind. The problems associated with the impact of policy interventions and the probability of spillover effects that cause cross-sectional dependence in the response variable of the logit model have already been mentioned. Anonther caveat concerns the usability of the findings in practice. The paper offers a very detailed insight into the macroeconomic surroundings of crises ex post, using all the data that is now available. This is certainly interesting and valuable, but to determine which indicators are really useful for policymakers, real-time data would have to be used during the whole analysis. Especially GDP-related data is often revised and most likely not correct in real time. Therefore, the values of the indicator variables at the time the policy maker has to decide whether to take action or not would differ from the values in this analysis and the predictive quality of the indicator variables would perhaps be worse than what is found here. Besides, the way the optimal thresholds are defined here is also problematic for policy applications. The optimal threshold is a certain percentile of the sample distribution, but this is computed ex post for the whole sample, using all available information. Of course, throughout the analysis we assume the distribution of the indicator variables to be stationary, which means that the sample distribution is a consistent estimate of the actual distribution, but still, in practice the numerical value of the cutoff-point will change with every month and every new observation. Therefore, to really examine the aptitude and predictive quality of the variables as real-time indicators for banking crises, the analysis would have to be redone using every month in the sample only the information that is available up to that point to determine the cutoff-value.

Further extensions of the present analysis could include other indicator variables, possibly also non-macroeconomic time series, like bank balance sheets. Moreover, the definition of a signal from an indicator could be varied. For instance, an indicator could be defined as giving a signal if the value 1 is obtained two or more times within a certain timeframe. This would make the indicators less noisy and also be more practice-oriented, as policy makers are more likely to react to persistent signals of macroeconomic imbalances. Additionally, the time interval around a crisis in which the indicators have to signal could be varied. Some policy measures take effect only with a considerable lag, which would make the early detection of crises more desirable. Hence, the time frame for correct signals could be changed from twelve months before and after the beginning of a crisis to the period from two years before the crisis until the beginning.

Finally, this paper only examines the macroeconomic developments around banking crises in Europe. The question of how policy makers should react to the warning signals has to be addressed in different frameworks. A number of issues have to be considered: When is a signal from an indicator a sufficiently severe sign of alarm for policy makers to step in? How harmful are banking crises to the economy (what is the loss function

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

of the policy maker)? What kind of action should be taken, which policy measures are most effective and appropriate? Moreover, for Europe the situation is even more complicated: what happens if there is severe evidence of an impending banking crisis in one country of the Eurozone, but not in the others? How can the central bank react? The present paper cannot answer these questions, but the results presented here provide valuable insights into the developments around banking crises in Europe and are therefore definitely a very good starting point for any analysis trying to deal with these problems.

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

References

[1] Borio Claudio and Lowe Philip (2002): Asset prices, financial and monetary stability: Exploring the nexus. BIS Working Papers No. 144

[2] Borio Claudio and Lowe Philip (2004): Securing sustainable price stability: should credit come back from the wilderness? BIS Working Papers No. 157

[3] Calomiris Charles and Gorton Gary (1991): The Origins of Banking Panics: Models, Facts and Bank Regulation. in: Financial Markets and Financial Crises. National Bureau of Economic Research, Inc., pp 109-174.

[4] Demirgüc-Kunt Asli and Detriagache Enrica (1998): The Determinants of Banking Crises in Developing and Developed Countries. IMF Staff Papers Vol. 45, No. 1

[5] Greene William H. (2002): Econometric Analysis. 5th ed., Upper Saddle River, New Jersey: Prentice Hall

[6] Hardy Daniel C. and Pazarbasioglu Ceyla (1998): Leading Indicators of Banking Crises: Was Asia Different? IMF Working Paper 98/91

[7] Kaminsky Graciela L. and Reinhart Carmen M. (1999): The Twin Crises: Causes of Banking and Balance-of-Payments Problems. American Economic Review, Vol. 89, No. 3, June, pp. 473-500.

[8] Kaminsky Graciela L. (1998): Currency and Banking Crises: The Early Warnings of Distress. International Finance Discussion Papers No. 629, October 1998

[9] Laeven Luc and F. Valencia (2010), "Resolution of Banking Crises: The Good, the Bad, and the Ugly", IMF Working Paper 10/146

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

A Data Appendix

The macroeconomic time series were obtained from the *International Financial Statistics (IFS)* database by the IMF and from the Thomson Reuters Datastream. In what follows I will give a detailed description of every indicator variable. To be consistent with Kaminsky and Reinhart, I tried to use the same datasources as they did whenever this was possible. Unless otherwise noted, twelve month percent changes where used.
M2 Multiplier: Ratio of M2 (IFS lines 34 plus 35) to base money (IFS line 14)
Domestic Credit/GDP: Ratio of real domestic credit (IFS line 52 divided by IFS line 64) to real GDP (IFS line 99B divided by IFS line 99BIP multiplied by 100), where annual real GDP was interpolated to obtain monthly data
Real GDP: see above
Stock indices: Thomson Reuters Datastream, total market index (stock market) (TOTMK+CC, CC=country code), where TOTMK was not available national or MSCI stock indices where used (Ukraine, Slovak Republic, Croatia)
Domestic Real interest rate: Deposit rate (IFS line 60) deflated using consumer prices (IFS line 64); in levels
Table 22 lists the timing of the banking crises for each country and whether the indicators have issued a signal during the crisis period (1) or not (0), or whether data was not available (NA).

Country	Crisis	M2-Multiplier	Domestic Credit/GDP	Real GDP	Stock Indices	Real Interest Rate
Austria	2008	1	1	1	1	1
Belgium	2008	1	1	1	1	1
Denmark	1987	1	1	1	NA	1
	2008	1	1	1	1	1
France	2008	1	1	1	0	1
Germany	2008	0	1	1	1	1
Ireland	2008	1	1	1	1	1
Netherlands	2008	1	1	1	1	1
Norway	1991	1	0	0	1	1
Portugal	2008	1	1	1	1	1
Spain	2008	1	1	1	1	1
Sweden	1991	NA	1	1	NA	1
	2008	1	NA	1	1	1
Switzerland	2008	1	1	1	1	1
UK	2007	1	0	0	1	NA
Bulgaria	1996	1	0	1	NA	1
Croatia	1998	1	1	1	1	1
Czech Republic	1996	1	0	1	1	1
Hungary	1992	1	0	1	0	0
	2008	1	1	1	1	1
Poland	1993	1	1	0	NA	1
Slovak Republic	1999	1	1	1	1	0
Slovenia	2008	NA	0	1	1	1
Ukraine	1999	0	NA	0	NA	1

Table 22: Crises dates and indicator signals

Die approbierte Originalversion dieser Diplom-/Masterarbeit ist an der Hauptbibliothek der Technischen Universität Wien aufgestellt (http://www.ub.tuwien.ac.at).

The approved original version of this diploma or master thesis is available at the main library of the Vienna University of Technology (http://www.ub.tuwien.ac.at/englweb/).

	Data availability	
Country	First year	Last year
Austria	1990	2011
Belgium	1990	2011
Denmark	1967	2011
France	1990	2011
Germany	1990	2011
Ireland	1990	2011
Netherlands	1990	2011
Norway	1980	2011
Portugal	1990	2011
Spain	1990	2011
Sweden	1980	2011
Switzerland	1980	2011
UK	1988	2011
Bulgaria	1990	2011
Croatia	1995	2011
Czech Republic	1994	2011
Hungary	1987	2011
Poland	1981	2011
Slovak Republic	1994	2011
Slovenia	1991	2011
Ukraine	1993	2011

Table 23: Countries and data availability