Bank of Finland

BoF Economics Review

4 • 2019

Banking crisis prediction with differenced relative credit

Karlo Kauko, Bank of Finland
Eero Tölö, Bank of Finland

Abstract

Indicators based on the ratio of credit to GDP have been found to be highly useful predictors of banking crises. We study the difference in this ratio as an early warning indicator. We test a large number of different versions of the differenced credit-to-GDP ratio with data on Euro area members. The optimal time interval of the difference is about two years. Using the moving average of GDP instead of the latest annual data has little impact on forecasting performance. The indicator is a particularly promising choice at relatively short forecasting horizons, such as two or three years.

Keywords: banking crises, early warning indicators, differenced relative credit, credit intensity, countercyclical capita buffer

JEL G01, G17, G28

1. Introduction

What kinds of economic and financial phenomena are regularly observable in a national economy before the country is hit by a major banking crisis? This question is of general interest, but it has become even more topical in the last decade, especially since the publication of the BIS Consultative Document (BIS 2010) and the introduction of the Countercyclical Capital Buffer in various jurisdictions. Regulators are now mandated to increase the capital adequacy requirements imposed on banks if the risk of a banking crisis seems heightened. Thus, research on early warning signs of banking crises has become highly policy relevant.

Research on the early warning signals of banking crises is older than regulations on the Countercyclical Capital Buffer. A handful of pioneering econometric studies, such as Kaminsky and Reinhart (1999) and Demirgüç-Kunt and Detragiache (1998), were published in the late 1990s. Since these early publications in this field, there seems to be an increasing consensus that excessive lending is a leading indicator of future problems, possibly the most important one. As to more recent research, at least Antunes et al (2018) and Schularick and Taylor (2012) have presented further evidence on its impact.

Yet, it is still not obvious what would be the best empirical specification of excessive lending. There seems to be no consensus yet of an optimal indicator of excessive lending derived from an established theory. What can be done in practice instead, one can systematically test the performance of alternative indicators.

Borio and Lowe (2002) concluded that one could use the trend deviation of the credit-to-GDP ratio. These authors estimated the trend by using the Hodrick-Prescott (HP) filter. Because the trend deviation was back-tested with data that would have been available on real time, the filter was run for each data point separately, i.e. not using the data for subsequent periods. This is normally referred to as the one-sided HP-filter.

Subsequent authors have confirmed that one version of this trend deviation, the so-called Basel gap, clearly outperforms many other indicators as an early warning signal of banking crises (see e.g. Detken et al 2014 or Alessi and Detken 2018). The Basel gap has gained official recognition. Both the Basel Committee (2010) and the European Systemic Risk Board in its Recommendation 2014/1 have promoted the use of the Basel gap as a benchmark guide for decision making. However, Repullo and Saurina (2011) have argued that the gap tends to reduce capital requirements mainly when GDP growth is high, potentially exacerbating the inherent pro-cyclicality of risk-sensitive bank capital regulations. The mechanistic interpretation of the gap can lead to absurd conclusions in the aftermath of an abnormally strong boom-bust cycle; recent history of Iceland would make a good example[1].

The Basel gap cannot be calculated in a meaningful way if suitable time series are not available for a relatively long period, such as two decades or more. In some countries data series on the loan stock may be too short. In some cases there have been abrupt crises or structural changes, making it problematic to calculate trend values for the present situation with data from the distant past.

In addition to the trend deviation of the relative size of the credit stock, the growth of the loan stock may also be an important factor to be monitored. A simple annual percentage growth of credit aggregates performs relatively poorly as an early warning indicator (see e.g. Detken et al 2014 or Alessi and Detken 2018, p 222). This may not be surprising. Basically, the percentage growth is the difference of the credit stock divided by the past value of the credit stock, i.e. 100 (ΔCt)/Ct-1. Dividing the difference of the loan stock by the past value of the loan stock would be meaningful if any given amount of additional credit, say one billion, were less dangerous if the pre-existing stock of credit were large. This would seem counterintuitive, and the opposite is more likely. In reality, additional loans are probably increasingly dangerous when the amount of pre-existing credit has already been unsustainable; there would be more and more over-indebted agents in the economy.

But what would make a good denominator for the difference of the credit stock (ΔC_t), if its own past value is not a promising candidate? The GDP is an obvious alternative, simply because it is closely related to the debt servicing ability of debtors. Kauko (2012) may have been

[1] See e.g. the Recommendation of the Icelandic Financial Stability Council from 22nd January 2016

the first to present this approach and an indicator based on it. This indicator was named "KK" (after the initials of the proposer) by Tölö et al (2018) in a systematic test of different early warning signals of banking crises. Detken et al (2014) call a slightly modified version of this indicator "differenced relative credit". Other variations have been called "credit intensity" (Castro et al 2014) and "moving average of simple slope" (Gonzalez et al 2017). In this paper, the term "differenced relative credit" is used. This indicator is used for policy purposes by e.g. the Bank of Finland.

Both Detken & al (2014) and Tölö et al (2018) have presented systematic comparisons on the crisis prediction ability of indicators that aim to measure excessive credit growth. Both comparisons conclude that both the Basel gap and the differenced relative credit are among the best crisis predictors. In a later study, Lo Duca et al (2017, p 41-42) have found that certain versions of the differenced relative credit indicator clearly outperform the Basel gap. Thus, it may be used either in addition to the Basel gap, or even instead of it, in both academic research and in macroprudential policy making.

The current version of the Basel gap is based on a certain parametrisation. This parameter value was obtained by systematic testing of different options (see Drehmann et al 2010). The differenced relative credit indicator has not yet been calibrated in a similar way, by systematically experimenting with various options. This paper aims to fill the gap. Moreover, the predictive power of differenced relative credit is compared to the Basel gap at different forecasting horizons.

The differenced relative credit indicator cannot be calculated unless one makes four choices.

1. One must choose the difference length, i.e. should one calculate the difference between the latest credit data and the respective number observed a year earlier? Or would it be advisable to use the difference between latest credit data and a much older observation on the same variable?
2. One must choose the number of quarterly GDP observations to be included in the denominator. One can use the sum or moving average of quarterly observations, but no theory tells us the optimal length of this window.
3. As will be seen in Section 2.1, one must choose a functional form for the indicator.
4. One must choose which price indices (if any) to use for deflating credit and GDP data.

As will be seen, the indicator seems relatively robust to minor changes in parametrisation, implying that choosing any given option is not likely to significantly reduce the usefulness or the indicator relative to the best performing options. With this sample, it is found that using a long moving average window for the GDP does not significantly improve the forecasting power. The optimal difference length is probably longer than one year. However, the forecasting ability of the indicator is not particularly sensitive to minor changes in these details.

Moreover, it is found that the differenced relative credit indicator outperforms the Basel gap at short forecasting horizons, such as two years. Instead, the Basel gap is a superior early warning indicator at longer forecasting horizons.

The rest of the paper is organised as follows. Section two describes the method and the data. Sections three and four present results. Section five presents a robustness test. Section six concludes and discusses the findings.

2. The method and the data

2.1. Calculating the indicator – options

The differenced relative credit indicator can be calculated in different ways. Ideally, the choice of specification should be derived from theory, but there are few theories to rely on. Thus, one must rely on intuition and take an experimental approach and test what works. There are at least four dimensions, namely the formula for calculating the indicator, the moving average span of the GDP, the difference length and the possible use of price indices.

As suggested by Kauko (2012), the differenced relative credit can be calculated with two different formulas. The first functional form is now defined as

$$x_1 = \frac{L_t - L_{t-\beta}}{(\frac{1}{\alpha})\sum_{j=0}^{\alpha-1} y_{t-j}} \quad . \tag{1}$$

where L is the loan stock and y is the quarterly GDP. Alpha is the moving average length in quarters and beta the difference length. The second version of the differenced relative credit indicator is now defined as

$$x_2 = \frac{L_t}{(\frac{1}{\alpha})\sum_{j=0}^{\alpha-1} y_{t-j}} - \frac{L_{t-\beta}}{(\frac{1}{\alpha})\sum_{j=0}^{\alpha-1} y_{t-j-\beta}} \ . \tag{2}$$

In both functional forms, the denominator(s) is (are) the moving average of quarterly GDP, not the latest observation. Cyclical variations in the denominator can be eliminated by calculating the average over several years. In principle, it would be possible to use sophisticated methods to identify the trend value, but from the point of view of practical purposes, the simple moving average is probably a sufficient measure, at least if one assumes relatively regular business cycles. Since Burns and Mitchell (1946), it has been commonplace to assume that the typical business cycle length would be eight years. However, several recent contributions indicate that the strongest GDP cyclicality occurs at somewhat shorter frequencies, possibly at 5-7 years (see e.g. Groth et al 2015, Verona 2016, Schüler et al 2017, Jagric 2003). Thus, the maximum moving average span is now limited to seven years (28 quarters). Because the GDP is measured in levels and not in differences, it typically remains relatively stable for several quarters. In order to limit the number of minor variations to be tested, the moving average is always calculated with a number of quarters divisible by four. This will also eliminate the risk that some seasonal variation not eliminated by the deseasonalisation method distorts the findings.

The difference length (beta) is another key choice to be made. It is difficult to deduct any theoretically correct or logical value for this parameter. Difference lengths shorter than four quarters are not tested; they may be dominated by random short-term variation. The maximum length is set rather arbitrarily at three years. As will be seen in the following, there is at least a local optimum at a shorter length in each specification. The credit aggregate is a stock variable, and it is not likely to undergo much seasonal variation. Because it is measured in differences and may be volatile in the short term, all options between four and 12 quarters will be tested.

Moreover, there are open questions concerning the use of price indices. The most obvious alternative is not to use any; one would relate nominal credit to nominal GDP. Another possibility is to apply the CPI to the credit aggregate in order to measure its development in real terms. In this case, the GDP should also be deflated by a suitable price index. Normally, the GDP is deflated by its own deflator, but this may bias the credit-to-GDP ratio, in case the two indices differ substantially. Thus, it can be meaningful to deflate both variables by the CPI. All these three options will be tested.

There are two alternative functional forms, seven possible moving average spans, nine possible difference lengths and three possible combinations of price indices, yielding 378 possible indicator values for each country and quarter. In the sequel, all of them will be tested.

2.2. The data

This study uses a panel data sample consisting of all current euro area members.

The crisis database is taken from Lo Duca et al (2017, p. 53, Table C1). There is no uniform, objective criterion for a banking crisis. The table is based on ESCB Heads of Research subjective assessments, and different group members may have used different criteria in crisis identification. In total, there are 22 different crises. Unfortunately, many of them are national manifestations of the 2008 crisis, implying that not all of them are genuinely independent events. Both problems are common to many contributions in this field of literature.

The rest of the data are from the ECB Statistical Warehouse. The loan stock includes mere bank loans to the non-financial private sector. Previous research, such as Detken et al (2014), has demonstrated that indicators based on bank loans instead of total credit are superior in crisis forecasting. Loan stock data refers to the end of period and it is not seasonally adjusted. The CPI is the overall harmonised index (2015=100). The GDP series are also deseasonalised. The GDP deflators use the base year 2010. Previous national currencies are converted into euro at a fixed parity.

Whenever possible, the data covers the period 1970-2017. However, in many cases there are data availability problems and a shorter period must be chosen. The sample is described in Table 1. The data range in Table 1 indicates the period for which the data are available.

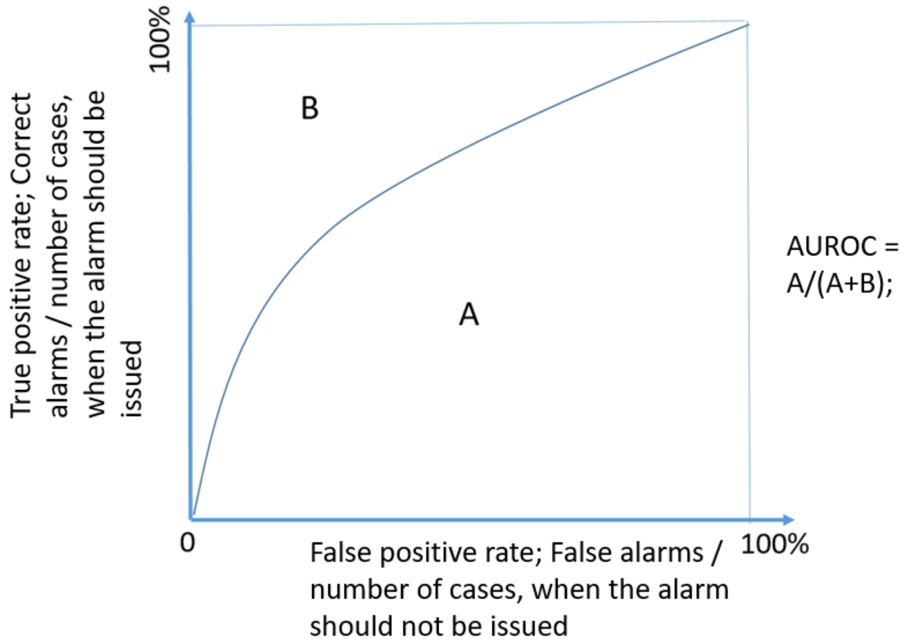
Table 1. The sample

Country	From	To	Crisis periods
Austria	1970q4	2017q3	2007/4 - 2017/3
Belgium	1970q4	2017q2	2007/4 - 2012/2
Cyprus	2005q4	2017q2	2011/2 - 2016 /1
Estonia	2008q1	2017q2	(none)
Finland	1974q1	2017q2	1991/3 - 1996/4
France	1970q4	2017q3	1991/2 - 1995/1; 2008/2 - 2009/4
Germany	1970q4	2017q3	1974/2 - 1974/4; 2001/1 - 2003/4 ; 2007/3 - 2013/2
Greece	1970q4	2017q2	2010/2 - 2017/2
Ireland	1971q2	2017q2	2008/3 - 2013/4
Italy	1974q4	2017q2	1991/3 - 1997/4; 2011/3 - 2013/4
Latvia	2003q1	2017q2	2008/4 - 2010/3
Lithuania	2004q2	2017q3	2008/4 - 2009/4
Luxembourg	1997q3	2017q2	2008/1 - 2010/4
Malta	2005q1	2017q2	(none)
Netherlands	1970q4	2017q3	2008/1 - 2013/1
Portugal	1970q4	2017q2	1983/1 - 1985/1; 2008/4 - 2015/4
Slovakia	2006q1	2017q3	(none)
Slovenia	2004q1	2017q2	2009/4 - 2014/4
Spain	1970q4	2017q2	1978/1 - 1985/3; 2009/1 - 2013/4

2.3. The assessment method

Following Detken et al (2014), the performance of an indicator is evaluated by its AUROC value, a method originally developed for medical science by Hanley and McNeil (1982). The method can be briefly explained as follows. Whenever an indicator is used as the sole source of information to forecast a binary variable, such as the occurrence of a crisis, one must choose a threshold. If and only if the value of the indicator signals at least a certain risk, a crisis is predicted. If the threshold is far too high (or low, if low indicator values signal heightened risk), there will be no false alarms, but each crisis will be missed. If the threshold value is set to the opposite extreme, each crisis is correctly forecasted, but the number of false alarms turns out to be very large. Thus, when different threshold levels are tested, the number of correctly predicted crises is an increasing function of the number of false positives. The Receiver Operating Characteristic (ROC) curve plots the false positive rate against the true positive rate for every possible threshold value. AUROC is the area below this curve. The AUROC value of a useless indicator is 0.5. If a perfect indicator existed, it would have the AUROC value 1. (See Chart 1) As to a more detailed description of the method and its properties, see e.g. Berge and Jordà (2011).

Chart 1. Calculating AUROC values



The signal value for each quarter and country is now considered a separate forecast. The forecasting horizon is 5-12 quarters. If the indicator issues an alarm, and a banking crisis breaks out in 5-12 quarters, the prediction is considered correct. Signals observed during crisis periods are excluded. Moreover, if a crisis begins in less than five quarters, the prediction is excluded from the assessment.

3. Results

A computer code went through the above described data, calculated the 378 indicator values for each quarterly observation, went through possible threshold levels and calculated the AUROC statistics for each indicator. All the 378 possible AUROC values are presented in Tables 2 and 3. AUROC statistics for a slightly modified version of the Basel gap are also calculated; the only difference between the standard Basel gap and the one tested here is that the indicator is based on bank loans to the private sector only, not on total credit, which would also include e.g. corporate bonds. The intention is to identify the best way to derive early warning signals from credit and GDP data, not to test which credit definition should be used. The AUROC value for this version of the Basel gap is 0.810. It would be possible to use a complete credit data sample, but the idea is to test nothing but the differences in the predictive power of different ways to calculate indicators based on the same data set.

As can be seen in Table 2, with this data, the first functional form (1) of the differenced relative credit indicator generally outperforms both the Basel gap and the second functional form. When one compares the average of the 189 versions of the first version to the Basel gap, i.e. the average of averages of panels 2a, 2b and 2c, the difference is about 0.02. If one compares the best outcomes, the difference is much larger. In fact, this might be the correct way to do it because even the Basel gap parametrisation, i.e. the value of lambda, was optimised with a sample that largely overlaps with this one. With this sample, the simple correlation between the Basel gap and the best performing parametrisation of X_1 is about 0.68.

The second functional form (2) of the indicator yields results that are, on average, weaker than the accuracy of the Basel gap, yet with an extremely narrow margin, the average of averages of panels 3a, 3b and 3c being 0.804.

Moreover, one obvious finding is that the parametrisations, i.e. values of alpha and beta, have a relatively minor impact on accuracy. This robustness is probably an encouraging finding. Choosing the wrong parametrisation does not render the indicator useless in e.g. setting the countercyclical capital requirement. There is no reliable formula for calculating the precise statistical significance of these values. However, in the light of bootstrapping results, the robust standard errors adjusted for clustering for each of these estimates is about 0.04. Thus, it is relatively safe to say that most AUROC values reported in Tables 2 and 3 do not statistically significantly differ from each other.

Taking the moving average of the GDP over five or seven years is unnecessary. If anything, a much shorter time span yields better results, although the difference is insignificant. These findings indicate that normally, one could use the average GDP of one to three years. This applies to both functional forms. Not eliminating cyclical variation from the denominator may be advisable. This observation may be related to fundamental questions on the nature, statistical properties and even the existence of the business cycle as a truly cyclical phenomenon (see e.g. Aslanidis and Fountas 2014). Possibly the indicator derives useful information from the GDP slow-down, if cyclical variations are not eliminated. It may also be related to inflation; as can be seen in panel 3c, the moving average length is of very little relevance, when both the GDP and the credit aggregate are deflated by the CPI.

As to the length of the difference, one year is probably not enough. In each case, it is optimal to use 7-10 quarters. Interestingly, the optimal difference length is often longer than the optimal moving average span.

As to nominal versus real data, there is a clear difference between the two functional forms. The first functional form works best with nominal data whereas the second formula yields better results with deflated data. In fact, with nominal data, the second functional form performs clearly weaker than the Basel gap. Applying the CPI to both credit and GDP seems a better choice than using different price indices.

These results are largely driven by the crisis of the years 2007-2008. However, the fit of the first version of the indicator would have been good even before this crisis. There are six pre 2007 crises where the data allows to calculate the value of X_1 eight quarters before a crisis breakout, using a two years difference and a three years moving average span of the GDP. These indicator values range from 0.607 to 1.274; even the lowest of these indicator values is higher than about 70% of all observations in the data. If the indicator were useless, the probability of all these indicator values falling into the highest 30% is less than 1 per mille.

Table 2. AUROC values for different specifications of the differenced relative credit indicator, 19 euro area countries

First version of the indicator (X1), 5-20 quarters horizon

Nominal data only

a)

		Alpha						
		4	8	12	16	20	24	28
Beta	4	0.821	0.825	0.821	0.825	0.824	0.822	0.811
	5	0.835	0.837	0.835	0.836	0.834	0.831	0.817
	6	0.851	0.853	0.851	0.851	0.849	0.843	0.827
	7	0.855	0.857	0.858	0.858	0.856	0.846	0.828
	8	0.856	0.862	**0.863**	0.862	0.859	0.844	0.826
	9	0.852	0.861	**0.863**	0.862	0.856	0.839	0.82
	10	0.848	0.86	0.863	0.86	0.851	0.833	0.814
	11	0.842	0.855	0.859	0.855	0.844	0.823	0.803
	12	0.837	0.852	0.855	0.85	0.836	0.813	0.793
Average	0.842							
Maximum	0.863							

GDP deflator; loans CPI

b)

		4	8	12	16	20	24	28
Beta	4	0.811	0.812	0.803	0.8	0.797	0.803	0.804
	5	0.826	0.824	0.815	0.811	0.809	0.814	0.814
	6	0.84	0.836	0.828	0.823	0.822	0.827	0.827
	7	**0.842**	0.839	0.833	0.828	0.827	0.832	0.83
	8	**0.842**	0.841	0.835	0.831	0.832	0.835	0.832
	9	0.839	0.839	0.834	0.831	0.833	0.834	0.83
	10	0.835	0.836	0.833	0.831	0.832	0.833	0.828
	11	0.829	0.832	0.831	0.829	0.829	0.828	0.823
	12	0.827	0.831	0.83	0.829	0.828	0.825	0.819
Average	0.827							
Maximum	0.842							

CPI to both variables

c)

		4	8	12	16	20	24	28
Beta	4	0.812	0.817	0.805	0.802	0.801	0.807	0.808
	5	0.828	0.828	0.818	0.813	0.813	0.819	0.818
	6	0.846	0.841	0.832	0.827	0.828	0.833	0.831
	7	0.849	0.844	0.837	0.833	0.835	0.839	0.836
	8	**0.85**	0.846	0.84	0.837	0.839	0.843	0.838
	9	0.844	0.843	0.84	0.838	0.842	0.843	0.838
	10	0.84	0.841	0.839	0.839	0.842	0.843	0.837
	11	0.833	0.837	0.837	0.838	0.839	0.839	0.832
	12	0.831	0.836	0.838	0.839	0.839	0.836	0.829
Average	0.833							
Maximum	0.850							

Basel gap AUROC 0.81

Table 3. AUROC values, 19 Euro area countries

Second version of the indicator (X2), 5-20 quarters horizon

Nominal data only

a)		Alpha 4	8	12	16	20	24	28
Beta	4	0.776	0.769	0.761	0.753	0.745	0.738	0.732
	5	0.785	0.777	0.769	0.761	0.753	0.745	0.739
	6	0.796	0.787	0.778	0.769	0.76	0.752	0.745
	7	**0.798**	0.789	0.779	0.77	0.76	0.752	0.744
	8	**0.798**	0.788	0.779	0.768	0.758	0.75	0.742
	9	0.795	0.785	0.776	0.765	0.755	0.747	0.739
	10	0.792	0.782	0.772	0.762	0.752	0.743	0.735
	11	0.787	0.777	0.767	0.756	0.747	0.738	0.73
	12	0.784	0.773	0.762	0.752	0.742	0.733	0.725
Average	0.763							
Maximum	0.798							

GDP deflator; loans CPI

b)		4	8	12	16	20	24	28
Beta	4	0.799	0.799	0.799	0.798	0.797	0.796	0.796
	5	0.81	0.81	0.81	0.809	0.808	0.807	0.806
	6	0.822	0.822	0.822	0.821	0.82	0.819	0.819
	7	0.828	0.828	0.828	0.827	0.826	0.825	0.824
	8	0.831	0.831	0.831	0.83	0.829	0.828	0.827
	9	**0.832**	**0.832**	**0.832**	0.83	0.829	0.828	0.827
	10	0.831	0.831	0.831	0.83	0.829	0.828	0.827
	11	0.83	0.83	0.83	0.829	0.828	0.827	0.826
	12	0.831	0.831	0.831	0.829	0.828	0.827	0.826
Average	0.822							
Maximum	0.832							

CPI to both variables

c)		4	8	12	16	20	24	28
	4	0.804	0.803	0.802	0.801	0.801	0.8	0.799
Beta	5	0.815	0.814	0.813	0.812	0.812	0.811	0.81
	6	0.829	0.827	0.826	0.825	0.824	0.823	0.822
	7	0.834	0.833	0.832	0.831	0.83	0.829	0.828
	8	0.838	0.837	0.835	0.834	0.833	0.832	0.831
	9	**0.839**	0.837	0.836	0.835	0.833	0.832	0.831
	10	**0.839**	0.837	0.836	0.835	0.834	0.833	0.832
	11	0.837	0.836	0.835	0.833	0.832	0.831	0.83
	12	0.837	0.836	0.835	0.834	0.833	0.832	0.83
Average	0.827							
Maximum	0.839							

Basel gap AUROC 0.81

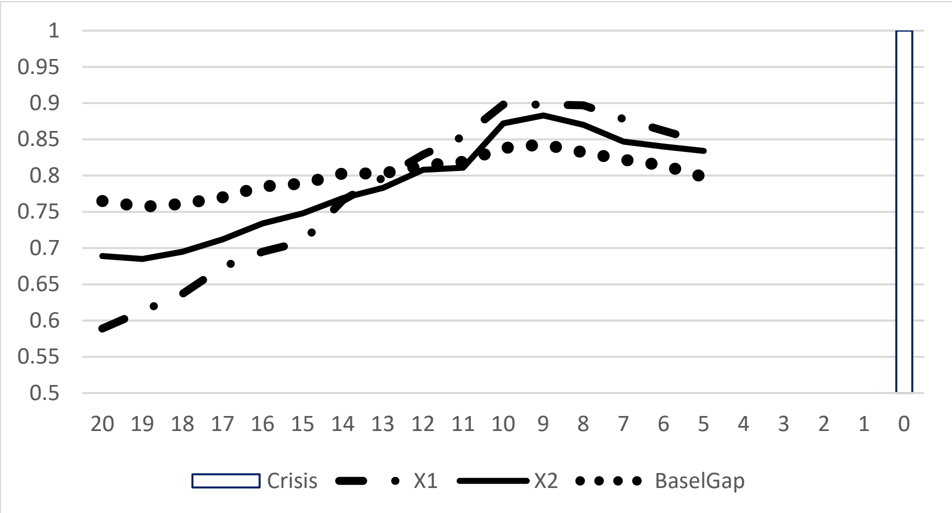
4. Forecasting horizons

Results presented in Section 3 are based on a rather wide forecasting horizon. In this section, the predictive power of the indicators is tested with different forecasting horizons. For each horizon N, an alarm is considered a true positive if a crisis breaks out precisely N quarters later. If there is an earlier or later crisis outbreak, the indicator value is not taken into account in assessments.

As to the differenced relative credit indicators, the best parametrizations identified in the previous section, i.e. values of alpha and beta, are used in the following analysis. The first functional form (X_1) is derived from nominal data and the parametrisation alpha = 12 and beta = 8. The second functional form (X_2) is derived from CPI corrected credit and GDP data and it uses the parametrisation alpha = 4 and beta = 9.

The AUROC values for both functional forms and the Basel gap are calculated separately for each forecasting horizon between 20 and five quarters. The results are plotted in Chart 2.

Chart 2. Predictive power of the different indicators at different forecasting horizons



As can be seen in Chart 2, both functional forms of the differenced relative credit indicator outperform the Basel gap if the crisis does not lie more than 5-10 quarters ahead. They work remarkably well at predicting crises if the forecasting horizon is about two years. When one applies a horizon of four to five years, the Basel gap clearly outperforms both differenced relative credit indicators. The first functional form performs remarkably poorly at long forecasting horizons. Thus, the differenced relative credit indicator seems a good short to medium term early warning signal, but it does not detect slowly growing imbalances many years in advance.

If the differenced relative credit indicator is used as a short-term early warning signal only, it is not obvious the parametrisations found optimal in Tables 2 and 3 would be suitable. As can be seen in Panel *a* of Table 4, the optimal parametrisation of the first functional form is slightly different, if the focus is solely on the two years forecasting horizon. With nominal data, the GDP should be the moving average over five years. Moreover, the indicator calculated with deflated data works remarkably well with very low values of both alpha and beta (Panel b of Table 4).

As to the second version, optimal parametrisation does not change when the forecasting horizon is limited to two years (see Panel c of Table 4). Versions with nominal data and different price indices for the two time series still underperform results with CPI deflated data; precise results are available from the authors upon request.

As a final experiment, it was systematically tested which parametrisation would yield the best results at different horizons in the case of the most promising version of the indicator, namely X_1 with nominal data. The results are presented in Table 5. As can be seen, the two years difference (beta) is not the best choice at most forecasting horizons. Interestingly, long moving averages of the GDP work best with short forecasting horizons, as if eliminating the latest business cycle developments would be useful when the crisis is already approaching.

Table 4. AUROC values for the two functional forms at the 8 quarter horizon

First version of the indicator X1, 8 quarter horizon

Nominal data only		Alpha						
a)		4	8	12	16	20	24	28
Beta	4	0.894	0.876	0.857	0.858	0.860	0.873	0.861
	5	0.891	0.867	0.855	0.856	0.859	0.874	0.861
	6	0.895	0.883	0.875	0.875	0.880	0.872	0.857
	7	0.900	0.887	0.883	0.885	0.889	0.879	0.862
	8	0.896	0.896	0.897	0.898	**0.903**	0.886	0.866
	9	0.883	0.886	0.891	0.895	0.897	0.880	0.860
	10	0.882	0.891	0.898	0.901	0.895	0.873	0.853
	11	0.879	0.889	0.896	0.899	0.889	0.866	0.844
	12	0.870	0.884	0.893	0.894	0.880	0.852	0.829
Average	0.879							
Maximum	0.903							

First version of the indicator X1, 8 quarter horizon

GDP deflator; loans CPI								
b)		4	8	12	16	20	24	28
Beta	4	**0.906**	0.888	0.868	0.856	0.851	0.866	0.865
	5	0.899	0.879	0.859	0.849	0.846	0.861	0.858
	6	0.892	0.874	0.857	0.849	0.85	0.863	0.86
	7	0.897	0.876	0.858	0.851	0.855	0.864	0.861
	8	0.886	0.873	0.861	0.857	0.864	0.871	0.866
	9	0.879	0.871	0.86	0.856	0.864	0.869	0.863
	10	0.878	0.871	0.867	0.865	0.871	0.874	0.867
	11	0.869	0.864	0.861	0.864	0.866	0.866	0.858
	12	0.857	0.854	0.856	0.858	0.859	0.859	0.852
Average	0.866							
Maximum	0.906							

Second version of the indicator (X2), 8 quarter horizon

CPI to both variables								
c)		4	8	12	16	20	24	28
	4	0.854	0.853	0.852	0.852	0.852	0.852	0.85
Beta	5	0.851	0.849	0.848	0.847	0.846	0.846	0.844
	6	0.856	0.854	0.853	0.852	0.851	0.85	0.848
	7	0.86	0.859	0.858	0.856	0.855	0.854	0.853
	8	0.869	0.866	0.864	0.863	0.862	0.862	0.86
	9	0.87	0.868	0.866	0.865	0.864	0.863	0.862
	10	**0.878**	0.876	0.873	0.873	0.872	0.87	0.869
	11	0.876	0.874	0.871	0.87	0.869	0.868	0.866
	12	0.874	0.872	0.87	0.869	0.867	0.866	0.864
Average	0.861							
Maximum	0.878							

Basel gap AUROC 0.832

Table 5. Best parametrisations, X1, nominal data

Horizon	Alpha	Beta	AUROC
5	16	12	0.888
6	16	12	0.897
7	16	11	0.894
7	16	12	0.894
8	20	8	0.903
9	4	6	0.907
10	4	5	0.904
11	8	4	0.866
11	12	4	0.866
12	8	3	0.867

If two different parametrisations yield the same AUROC value, both parametrisations reported

5. A robustness test

The conclusions presented in the above sections are based on a relatively small sample, consisting of 19 countries only. Therefore, an additional robustness test with a different sample is conducted. In the following, the main conclusions are tested with a sample of 21 countries; the two samples do not overlap at all. In order to also test whether the main results are robust to changes in the method, we try a different statistical technique, namely the logit analysis. This method has the additional advantage of having standard methods to test explanatory variables' statistical significance. Crises are from the Laeven and Valencia (2012, p 24-26) database. There are 27 crises in the sample. The domestic credit to private sector (% of GDP) is from the World Bank database. In the case of Denmark, Sweden and Switzerland, missing observations were filled with data based on the Jordà-Schularick-Taylor macrohistory database[2]. The data is annual. The explanatory variables are differences between annual observations, i.e. (Credit-to-GDP year t) – (Credit-to-GDP year t-N). Conceptually, this difference is analogous to Equation 2, where alpha = 4 and beta is either 4, 8 or 12.

Results presented Tables 2 and 3 indicate that two years might be the optimal length of difference. As can be seen in Chart 2, the forecasting power of differenced relative credit is particularly strong at the two years horizon. Thus, it is tested whether the two years change in the credit-to-GDP ratio forecasts crises with a two years lag. As can be seen in Table 6, this result is confirmed. Rapid credit growth for two years (say, between years t-2 and t) significantly increase the probability to experience a crisis after two years (year t+2). This growth is not likely to lead to problems that would materialise significantly earlier or later.

[2] http://www.macrohistory.net/JST/JSTdocumentationR3.pdf

Table 6. Logit analysis of financial crisis determinants

	eq 1	eq 2	eq 3	eq4	eq5	eq6
C	-3.598	-3.592	-3.606	-3.606	-3.606	-3.556
	(-14.3)	(-13.1)	(-14.8)	(-13.9)	(-13.9)	(-14.3)
Diff2(-1)		-0.008		0.016		
		(-0.6)		(0.7)		
Diff2(-2)	0.034	0.053	0.069			0.047
	(3.1)***	(3.9)***	(3.1)***			(3.8)***
Diff2(-3)		-0.029				
		(-1.9)*				
Diff1(-1)				-0.038	0.003	-0.003
				(-2.1)**	(0.2)	(-0.2)
Diff1(-2)			0.005		0.040	
			(0.2)		(2.3)**	
Diff1(-3)					0.024	
					(1.8)*	
Diff3(-1)				0.024		
				(1.8)*		
Diff3(-2)			-0.035			
			(-2.1)**			
Diff3(-3)						-0.023
						(-2.1)**
McFadden R square	0.094	0.128	0.109	0.095	0.095	0.106

Explained variable the dummy variable "crisis", 27 crises in the sample
DiffX(-N) for year t = w(t-N) - w(t-N-X), where w(z) = credit-to-GDP year z
Switzerland, Chile, Denmark, Indonesia, Iceland, Israel, Mexico, Singapore, Uruguay, Brazil, Argentina, United States, United Kingdom, South Korea, Japan Australia, New Zealand, Norway, Sweden, Canada, Turkey
Credit-to-GDP data covers years 1969-2010 (Indonesia 1980-2010)
Annual fixed effects for years when at least one crisis
Cases where another crisis broke out 1-3 years earlier are omitted
Z stats in parentheses;
***, ** and * denote 1%, 5% and 10% significance, respectively.

6. Conclusions and discussion

This paper contributes to the growing empirical literature on early warning signals for banking crises. The focus is on different variations of one specific indicator, namely the difference of the credit to GDP ratio. The data covers bank loans in all current Euro Area countries.

When applied to bank loan data, both versions of the differenced relative credit indicator generally outperform the Basel gap at relatively short forecasting horizons, such as two years. If one tries to forecast crises several years in advance, the Basel gap outperforms the differenced relative credit indicators. This observation has obvious policy implications. The Basel gap and the differenced relative credit indicator are useful for slightly different purposes, and they should be seen as complements rather than alternative options. Because the Basel gap detects slow developments well in advance, a high value does not necessarily indicate that there is an urgent need to take action. There is an underlying trend of excessive debt growth that needs to be curtailed. If the differenced relative credit reaches an alarming level, the situation is more acute and maybe one should even prepare for crisis management.

As to constructing the indicator, it was found that the optimal difference length is longer than one year. The two years growth yields better results. This finding was corroborated with the robustness test of Section 5. Instead, it seems that using the moving average of the GDP for a period of several years adds little or no value. Using relatively recent GDP values provides indicators that require less data and perform about as well or slightly better.

As to future work, there are three obvious topics to work on.

First, despite of the robustness test of Section five, many results are based on a fairly limited number of countries and crises, and little has been said on the statistical significance of the results. It would be an important topic for future research to test whether the results apply to other country groups, such as other developed or emerging countries. It would also be interesting to know whether the results generalise to data from another era, such as early 20th century. These further analyses could shed some light on how sample-specific the main results are likely to be.

Second, an international dimension would be useful. Credit cycles and financial crises seem imported phenomena in many countries (Rey 2016). Domestic and global credit booms may have different implications for financial stability at the national level.

Third, a stronger theoretical background would be welcome. To a large extent, the literature on optimal indicators for banking crisis prediction still suffers from lack of explicit theories. Although the economic concept to be measured is intuitive and broadly guided by theory, the indicators used so far are mainly empirical. Functional forms, parameter values or lag-lengths are assumed, or based on in-sample optimisation. However, if we do not fully understand why certain indicators have worked, and why some ways to measure a macrofinancial phenomenon have outperformed other alternative specifications in crisis prediction, it is difficult to say whether their predictive power will remain strong in the future.

References

Alessi, L., Detken, C. (2018) Identifying excessive credit growth and leverage; Journal of Financial Stability 38, 289–308.

Antunes, A., Bonfim, D., Monteiro, N., Rodrigues, P.M.M. (2018) Forecasting banking crises with dynamic panel probit models; International Journal of Forecasting 34, 249-275.

Aslanidis, N., Fountas, S. (2014) Is Real GDP Stationary? Evidence from a Panel Unit Root Test with Cross-Sectional Dependence and Historical Data; Empirical Economics 46, 101-108.

Basel Committee on Banking Supervision (2010) Guidance for national authorities operating the countercyclical capital buffer; Bank for International Settlements, December 2010.

Bege, T.J., Jordà, O. (2011) Evaluating the classification of economic activity into recessions and expansions; American Economic Journal: Macroeconomics 3, 246-277.

BIS (2010) Countercyclical Capital Buffer Proposal; Basel Committee on Banking Supervision, Consultative Document, July 2010.

Borio, C., Lowe, P., (2002) Assessing the risk of banking crises; BIS Q. Rev. (December), 43–54.

Burns, A. F., Mitchell, W.C. (1946) Measuring Business Cycles; NBER Books, New York.

Castro, C., Estrada, A., Martínez, J (2014) The countercyclical capital buffer in Spain: an Exploratory Analysis of Key Guiding Indicators; Financial Stability Journal no 27, Bank of Spain.

Demirgüç-Kunt, A., Detragiache, E. (1998). The determinants of banking crises in developing and developed countries; IMF Staff Papers 45, 81–109.

Detken, C., Weeken, O., Alessi, L., Bonfim, D., Boucinha, M.M., Castro, C., Frontczak, S., Giordana, G., Giese, J., Jahn, N., Kakes, J., Klaus, B., Lang, J.H., Puzanova, N., Welz, P. (2014) Operationalising the countercyclical capital buffer: indicator selection, threshold identification and calibration options; ESRB Occasional Paper Series 5.

Drehmann, M., Borio, C. Gambacorta, L., Jiménez, G., Trucharte, C (2010) Countercyclical capital buffers: exploring options; BIS Working Papers 317.

Gonzalez, R.B., Marinho, L.S.G., Vasconcellos e Lima (2017) Re-anchoring Countercyclical Capital Buffers: Bayesian Estimates and Alternatives Focusing on Credit Growth; International Journal of Forecasting 33, 1007-1024.

Groth, A., Ghil, M., Hallegatte, S., Dumas, P (2015) The role of oscillatory modes in US business cycles; OECD Journal: Journal of Business Cycle Measurement and Analysis.

Hanley, J. A.; McNeil, B.J. (1982). The Meaning and Use of the Area Under a Receiver Operatine Characteristic CROCÌ Curve; Radiology 143, 29-36.

Jagric, T. (2003) Business Cycles in Central and East European Countries; Eastern European Economics 41, 6-23.

Kaminsky, G.L., Reinhart, C.M. (1999) The twin crises: the causes of banking and balance-of-payments problems; American Econonic Review 89, 473–500.

Kauko, K (2012) Triggers for countercyclical capital buffers; BoF Online 7/2012.

Laeven, L., Valencia, F. (2012) Systemic Banking Crises Database: An Update; IMF WP/12/163.

Lo Duca, M., Koban, A.,Basten, M., Bengtsson, E., Klaus, B., Kusmierczy, P., Lang, J.H., Detken, C. (2017) A new database for financial crises in European Countries; ECB Occasional Paper Series 194.

Repullo, R., Saurina, J., (2011) The Countercyclical Capital Buffer of Basel III: A Critical Assessment; CEMFI Working Papers 1102.

Rey, H. (2016) International channels of transmission of monetary policy and the mundellian trilemma; IMF Economic Review 64, 6-35.

Schularick, M., Taylor, A.M. (2012) Credit booms gone bust: monetary policy, leverage cycles and financial crises 1870-2008; American Economic Review 102, 1029-1061.

Schüler, Y. S., Hiebert, Paul P., Peltonen, T.A, (2017) Coherent financial cycles for G-7 countries: Why extending credit can be an asset; ECB Working Paper Series 43/2017.

Tölö, E., Laakkonen, H., Kalatie, S. (2018) Evaluating Indicators for Use in Setting the Countercyclical Capital Buffer; International Journal of Central Banking 14, 51-111.

Verona, F (2016) Time–frequency characterization of the U.S. financial cycle; Economics Letters 144, 75-79.