JEL Classification: G01, G21, G28.
Keywords: early warning indicators, banking crises, credit-to-GDP gap

Forecast-Augmented Credit-to-GDP Gap as an Early Warning Indicator of Banking Crises

Adam GERŠL - Charles University, Institute of Economic Studies, Czech Republic (adam.gersl@gmail.com) *corresponding author*

Thomas MITTERLING - Vienna University of Economics and Business, Department of Economics, Austria

Abstract

This paper explores whether augmenting the credit-to-GDP series with a forecast improves the early warning property of the credit-to-GDP gap – a frequently used indicator of excessive credit expansions calculated using the one-sided Hodrick-Prescott filter. Improving the early warning property of this indicator is extremely important as policymakers frequently rely on it when deciding about macroprudential policy interventions such as when calibrating the Basel III countercyclical capital buffer or other macroprudential instruments. Using data for 56 countries over 1950-2016, we simulate in a quasi-real-time setting how different types of forecasts would have changed the gap. We build simple statistical forecasts, more complex economic forecasts based on regression models estimated in real-time with IMF country-specific WEO macro projections used as inputs, and plausible credit cycle corrections. The early warning power of alternative credit-to-GDP gaps is tested within the ROC/AUROC framework. Our results indicate that for advanced markets, none of the adjustments can beat the simple one-sided filter, but for emerging markets, the correction-adjusted gaps could improve the signalling power.

1. Introduction

Macroprudential authorities charged with the objective to safeguard financial stability often monitor indicators capturing excessive credit expansion. One of the most popular ones is the credit-to-GDP gap, i.e. the difference between the observed and the trend credit-to-GDP ratio. This indicator is known for its good early warning properties for future financial crises (BCBS 2010a; Drehmann et al. 2010) given that periods of very strong credit growth are often associated with a build-up of systemic risk. This indicator has also been proposed by the Basel Committee for Banking Supervision (BCBS) as the main variable underlying the calibration of the countercyclical capital buffer (CCyB) in Basel III (BCBS 2010b). To calculate the trend credit-to-GDP ratio, the Basel Committee suggests that a Hodrick-Prescott (HP) filter with a high degree of smoothing (λ = 400,000) is applied on quarterly data. While many countries do calculate the credit-to-GDP gap following the BCBS

https://doi.org/10.32065/CJEF.2021.04.04

Authors would like to thank Jesus Crespo Cuaresma, two anonymous referees, and participants to the Economists Meetings at the Joint Vienna Institute for useful comments to previous versions of this article. Adam Geršl acknowledges the support of the Czech Science Agency, Grant No: 18-05244S. Usual disclaimers apply.

guidance, they often rely on a much wider range of other suitable indicators when calibrating the final CCyB rate.

When the gap is calculated in real-time, the trend value for the last observation is estimated using data only to the left side of the last observation, i.e., with a one-sided HP filter. This trend value will be revised ex-post once future data becomes available (in which case a two-sided filter will be applied to this particular observation). Thus, the last observation has a major effect on the trend estimation, leading to the well-documented end-point bias of the HP filter discussed in the 1990s when this method was frequently used to estimate output gaps (Apel, Hansen, and Lindberg 1996; St-Amant and Van Norden 1997). The end-point bias is an important issue in macroprudential policy. During periods of strong credit growth, when macroprudential policy should tighten to tame accumulation of systemic risk, the trend credit-to-GDP would be typically upward biased and thus the credit-to-GDP gap downward biased, underestimating the degree of excessive credit expansion and thus the need for policy intervention. While the one-sided gap measure still retains its good early warning properties in comparison with alternative indicators, Edge and Meisenzahl (2011) document that ex-post revisions to the credit-to-GDP gap can be sizable and as large as the gap itself. Drehmann and Tsatsaronis (2014) show that if we knew perfectly the future path of credit-to-GDP, the gap indicator's signaling power (estimated ex post through a two-sided filter) would increase considerably.

The recommended way to deal with the end-point bias problem is to extend the series with a forecast to try to get as close as possible to the future observed values of the filtered variable and perform a quasi-two-sided filtering to estimate the trend values (Kaiser and Maravall 1999; Kim and Newbold 2005). However, as financial cycles are long and the recommended smoothing parameter is high, the forecast needs to be of a medium-term nature (several years) to have a noticeable effect on the estimated trend. Moreover, should the forecast only extend a running credit boom further to the future (a probable result of any forecast in good times given the strong persistence and supportive macro-financial feedback), the end-point bias could be even magnified.

In our paper, we explore whether extending the credit-to-GDP series by various types of forecasts and calculating the credit-to-GDP gap using the quasi-two-sided HP filter removes the end-point bias and improves the early warning property of this indicator in terms of better signalling future banking crises. We build upon Gerdrup, Kvinlog and Schaanning (2013) who were the first to show that modifying the calculation of the trend by extending the credit-to-GDP data by a very simple forecast improves the early warning properties of the gap measure in case of Norway, one of the first countries that implemented the CCyB in 2013. In comparison to them, we focus on a sample of 56 countries over the period 1950-2016 and produce a number of alternative real-time forecasts for credit-to-GDP.

We test three types of forecasts. First, we build pure statistical forecasts such as an autoregressive forecast and a linear trend, which typically offer good short-term fit. Second, we use economic forecasts that are based on regression models estimated in quasi-real-time, using at that time available IMF World Economic Outlook forecasts for GDP, inflation, and current account balance to produce credit-to-GDP forecasts over a horizon of 5 years. By relying on economic fundamentals, these forecasts might be promising in offering a good medium-term fit, bringing the projected credit-to-GDP path closer to its future outturn over several years. Third, following the approach by Gerdrup, Kvinlog and Schaanning (2013), we also experiment with somewhat non-intuitive, but plausible projections, assuming that the credit-to-GDP ratio will either stabilize and stay at a constant level or reverse its expansion and decline at an inverse linear trend. These projections build in a certain correction in the credit developments and simulate a credit cycle pattern, with the advantage to artificially increase the credit-to-GDP gap in times of credit booms. We compare all adjusted gaps with the three benchmarks: the one-sided gap, the two-sided gap, and a perfect foresight-based gap (if we knew the reality for the next 5 years).

We evaluate the fit of the various forecasts with standard statistics and test the early warning properties of the alternative gaps using a database of banking crises by Laeven and Valencia (2013) within the ROC/AUROC framework. Our analysis suggests that for advanced economies, neither of the adjustments significantly improves the early warning properties of the simple one-sided credit-to-GDP gap – a puzzle given that the simple unadjusted gaps are clearly biased. However, for emerging markets, we find that the correction-augmented credit-to-GDP gaps – which often do not offer any good fit and are to be interpreted more as artificial adjustments rather than true forecasts – tend to improve the signalling power, albeit at the margin of statistical significance.

The paper is organized as follows. Section 2 briefly discusses the related literature, while Section 3 describes the data. Section 4 explains the filtering with the Hodrick-Prescott filter, Section 5 present the methodology, and Section 6 the results. Section 7 concludes the paper.

2. Review of Literature

This paper relates to three main areas of literature. The first one discusses the suitability of the credit-to-GDP gap to guide the calibration of the Basel III CCyB. The original BIS research that started with Borio and Lowe (2002) and developed further in the context of the preparation of the CCyB in Drehmann et al. (2010) and Drehmann, Borio, and Tsatsaronis (2011) showed that the credit-to-GDP gap is the best early warning indicator of future banking crises – at least for a sample of advanced countries – when compared to a wide range of other system-wide aggregate macro-financial variables or banking sector indicators. This is in line with existing evidence that banking crises in both advanced and emerging markets tend to get preceded by periods of strong credit growth (Schularick and Taylor 2012; Dell'Ariccia et al. 2012). Thus, credit-based indicators in general – not only the credit-to-GDP gap, but also changes in credit-to-GDP or a simple credit growth – may capture well a build-up of financial vulnerabilities and an increased risk of crises (Geršl and Jašová 2018).

On the other hand, a number of papers criticized the Basel guidance along several dimensions. Geršl and Seidler (2015) argue that the credit-to-GDP gap estimated with an HP filter with such a large λ may work for advanced countries as it requires at least 20 years of reliable data in quarterly frequency. However, this is a challenge for emerging markets with frequent structural changes and problematic data quality. Moreover, the typically strong financial deepening observed in emerging markets may influence the trend given the end-point bias of the HP filter, underestimating the gap. Repullo and Saurina (2011) claim that GDP growth works better than the credit-to-GDP gap as a signaling variable of excessive macro-financial imbalances. Using data for EU countries, Detken et al. (2014) found that overall signaling performance improves when the credit-to-GDP gap is combined with other variables such as property price-to-income ratio, property price gaps, or the debt service-to-income ratio.

The second block to which our research relates focuses on constructing and analyzing early warning systems. Candelon, Dumitrescu, and Hurlin (2012) propose a new toolbox to evaluate early warning indicators (EWIs) incorporating the receiver operating characteristics (ROC) curve and comparison tests based on the area under the curve (AUROC or AUC). Based on these methods Drehmann and Juselius (2014) present EWIs on banking crises in 26 countries, again finding the Basel III credit-to-GDP gap to be the best indicator over longer horizons.

The third stream of literature to which our paper contributes is statistical filtering of macro-financial variables in order to decompose a series into trend and cycle components and, in particular, how to deal with the end-point bias. Most of the literature here is on business cycles (Baxter and King 1999), although with the growing interest in analyzing financial cycles, new contributions are emerging also in this area. Edge and Meisenzahl (2011) show that the end-point bias in credit-to-GDP gap estimations is sizable, as ex-post revisions to the gap can be very large. Drehmann and Tsatsaronis (2014) show that if we knew perfectly the future path of credit-to-GDP, the gap indicator's signaling power (estimated through a two-sided filter) as measured by the area under the ROC curve (AUC) would increase considerably. Kaiser and Maravall (1999) and Mise, Kim and Newbold (2005) show that augmenting the time series with a forecast produced by a simple ARIMA model decreases the end-point bias. Based on these findings, Gerdrup, Kvinlog and Schaanning (2013) amend historical observations with a very simple forecast (assuming that credit-to-GDP will stay at a level equal to the average of the last four quarterly observations for the next 5 years) in order to provide more robust estimates of the trend and the cycle component. For Norway, they show that using the forecast-augmented HP filtering improves the signaling quality of the credit-to-GDP gap as well as of three other macrofinancial indicators.

Hamilton (2018) argues that the HP filter should never be used, as it leads to spurious dynamics. While Drehmann and Yetman (2018) agree with Hamilton's criticism, they show that the credit-to-GDP gap based on HP filtering is not outperformed by any other gap. Wolf, Mokinski and Schüler (2020) explore the properties of the one- and two-sided HP filter. Based on the differences they propose adjustments to the one-sided HP filter to align its properties with the two-sided filter.

3. Data

We use quarterly data on 58 advanced (21) and emerging (37) economies collected from the IMF, BIS and national central banks. We tried to get the longest possible time series for the main variable of interest, i.e. the (bank) credit-to-GDP ratio, for each country to maximize the number of observations for our analysis. Data availability differs across countries – for some, mostly advanced economies, the data starts as early as in the 1950s or 1960s, while for others, especially Emerging Europe countries, the data starts in the 1990s or early 2000s, although other (non-European) emerging markets' data start often in 1970s or 1980s. Tables A1 and A2 in the Appendix show the start of the series for all countries in our dataset. The last observation available is either 2015q4 or 2016q1. Such a long time span makes sure that there is at least one – but mostly two or more – financial cycle for each country covered by our data.

In 2016 the IMF released a dataset with historical World Economic Outlook (WEO) annual forecasts for three different variables, namely real GDP growth, CPI inflation and the current account balance as a percentage of GDP. The dataset contains forecasts from all WEO publications since 1990. For each publication the dataset contains two years of historical and six years of forecast data (for the year of publication and next 5 years). Details on the data availability for different countries can be found in the Appendix. In order to obtain quarterly data we use the annual values of real GDP growth and CPI inflation for each quarter of the corresponding year but linearly interpolate the current account balance, using the annual value as end-of-year value.

We complement the bank credit-to-GDP series and the IMF WEO forecasts with historical macroeconomic data on GDP growth, inflation, and current account balance. These are used to estimate regression models "in real-time" to link the bank credit-to-GDP development to these macroeconomic factors. The estimated coefficients are then applied onto the IMF WEO forecast data to generate projections of credit-to-GDP.

Data on banking crises is mainly based on Laeven and Valencia (2013) and Babecky et al. (2014), with additions from Mehrez and Kaufmann (2000) and Kaminsky and Reinhart (1999). Tables A3 and A4 in the Appendix show the crises included in our analysis. Interestingly, each country experienced a crisis during the time analyzed in our research, with some countries having two or more crises (such Argentina). In advanced markets, we have 33 crises recorded, but only 30 crises are used in our early warning analysis (for 3 crises, there was no data on the credit-to-GDP ratio before the crises). In emerging markets, the database includes 46 crises, but only 17 of them are used (for the same reason as in advanced economies – missing data on the credit-to-GDP ratio before the crises).

4. Calculating Credit-to-GDP Gap with Hodrick-Prescott Filter

The Hodrick-Prescott (HP) filter, developed by Hodrick and Prescott (1997), is a standard tool to decompose an economic time series into its trend and cyclical component. The resulting trend series minimizes the following sum for a given value of λ:

$$\min_{\tau}\left(\sum_{t=1}^{T}(y_t-\tau_t)^2+\lambda\sum_{t=2}^{T-1}\left[(\tau_{t+1}-\tau_t)-(\tau_t-\tau_{t-1})\right]^2\right) \tag{1}$$

Higher values of λ imply a higher degree of smoothing. In order to estimate the trend credit-to-GDP ratio, the Basel Committee proposed an one-sided HP filter with λ = 400,000 to calculate the credit-to-GDP gap as a guidance for calibrating the countercyclical capital buffer. As financial cycles are longer than business cycles, Drehmann et al. (2010) explain why λ needs to be higher than the traditional 1,600 for business cycle detrending. They show how to derive the value of 400,000 by using the Ravn-Uhlig formula (Ravn and Uhlig 2002), assuming that the financial cycle is about 4-times longer than the business cycle. The existing research indicates that the difference between the duration of the financial and business cycles might not be that long for all countries (Claessens, Kose and Terrones 2011). However, Drehmann et al. (2010) find that out of alternative calibrations of λ, assuming that the financial cycle is as long as the business cycle (λ = 1, 600), twice as long (λ = 25,000), three times as long (λ = 100,000) and four times as long (λ = 400,000), the highest λ led to the credit-to-GDP gap with the best early warning properties.

Estimating the trend values in real-time (i.e. for the last observation for which data is available) and using the calculated gap as guide for policy is a challenge for policymakers, as towards the end of a data series the estimated trend is uncertain and will be revised once new observations are added. If the data used to estimate the trend is only available to the left of the observation concerned, we talk about the one-sided HP filter. While this is inevitably the case for the last available observation at the end of the series, one can run an one-sided HP filter also recursively on the whole series, simulating what would have been the gap value "in real-time", i.e. if we were at that particular period and had only data up to that point in time.

Depending on the value of the smoothing parameter λ, incoming observations will lead to revisions of the estimated trends in the past. With a high value of λ = 400,000, these revisions can affect observations up to 20 years back in time, often changing what was thought to be a positive gap into a negative gap and the other way round, as documented by Edge and Meisenzahl (2011). Thus, the one-sided filter suffers from the end-point bias, which is the difference in the estimated trend with the one-sided filter and the two-sided filter, once enough data to the right of the observation concerned is available.

While in principle the trend can be estimated with as low as three observations, to capture some meaningful economic trends, we set the minimum number of observations (*minT*) before the HP filtering is applied to 20 quarters. Contrary to Gerdrup, Kvinlog and Schaanning (2013) we discard the first 19 (= *minT* − 1) values at the beginning of the series. The one-sided HP-filtered credit-to-GDP gap is used as our main benchmark, as this is what policymakers always have at their disposal based on the Basel Committee guidance.

Even if unavailable for the policymaker in real-time, we are also calculating and using the two-sided HP filter as another benchmark in order to illustrate what would have happened if we had known the future perfectly. Clearly, this is not attainable in reality, but as we are testing various types of forecasts to project the future as close as possible to the (only ex-post available) outturns, we are de facto attempting to reach the ideal of the two-sided filter. As shown by Drehmann and Tsatsaronis (2014), signaling properties of the two-sided HP filter-based gaps are better than the one-sided-based gaps.

Finally, we estimate the credit-to-GDP trends by applying the so-called "quasi-two-sided" HP filter. In each period, we extend the credit-to-GDP series by a forecast over a horizon H, which we set to be 20 quarters, i.e. 5 years. Then, we apply the two-sided filter to the extended series and recover the trend value for the period concerned (which was the last period with actual data).

5. Methodology

We use three different types of forecast schemes in order to extend the credit-to-GDP time series for each country and each point in time over a horizon *H* in order to be able to run the quasi-two-sided HP filter. In other words, we simulate the policymakers' situation in real-time in terms of data availability and the possibility to amend the time series with a forecast prepared out of the set of information available at a point in time. In the following equations *h* takes on values between 1 and 20 (= *H*) quarters ahead. Following other literature in this area, we opt to show the results for all these horizons, but we acknowledge that the typical horizon for policy purposes would be between 4Q to 8Q, i.e. around 6Q ahead.

First, we produce two simple statistical forecasts, namely a rolling linear forecast and an autoregressive forecast. The rolling linear forecast assumes that the indicator follows a linear trend. The coefficients of the regression are recursively updated using the observations of the last four quarters. This linear trend is then used to forecast the indicator for the next five years. In case of the autoregressive forecast an AR(4) process is applied to the 4-quarter change of the credit-to-GDP ratio, i.e. the 4-quarter change is regressed on its four last values.

Rolling linear forecast: $$\Delta\hat{y}_{t+h,t+h-4} = \alpha + \beta^{\mathsf{T}} \times \Delta y_{t-3:t} \tag{2}$$

Autoregressive forecast: $$\hat{y}_{t+h} = \alpha_{t-3:t} + \beta_{t-3:t} \times (t+h) \tag{3}$$

Second, we create two "economic" forecasts, both based on regressing the (4-quarter change in the) credit-to-GDP on macroeconomic factors and using the IMF WEO forecast vintage available at that point in time for the given country as input into the estimated model to forecast credit-to-GDP over the next five years. The WEO1 forecast is using only the real GDP growth, while the WEO2 is using the real GDP growth, inflation, and current account balance.

WEO forecasts: $$\Delta\hat{y}_{t+h,t+h-4} = \alpha + \gamma^{\mathsf{T}} \times X_{t-3:t} \tag{4}$$

Third, we experiment with somewhat non-intuitive, but plausible "technical" forecasts that build in a certain correction in the credit developments to account for a typical credit cycle. Two of these assume that credit-to-GDP will stabilize and stay at an unchanged level over the next 5 years – either at the average value of the last four quarters as in Gerdrup, Kvinlog and Schaanning (2013) (rolling average forecast) or at a level of the last observation (constant forecast). The third projection expects a more radical correction in which credit-to-GDP reverses its expansion and declines at the same rate as projected by the linear forecast (only with a negative sign). This artificial simulation of a typical credit cycle pattern would increase the credit-to-GDP gap in times of credit booms as it naturally brings down the estimated trend, decreasing the end-point bias.

Average forecast: $$\hat{y}_{t+h} = \frac{1}{4}\sum_{s=t-3}^{t} y_s \qquad (5)$$

Constant forecast: $$\hat{y}_{t+h} = y_t \qquad (6)$$

Inverse linear forecast: $$\hat{y}_{t+h} = \alpha_{t-3:t} - \beta_{t-3:t} \cdot (t+h) \qquad (7)$$

We also create one additional benchmark, namely the perfect forecast represented by Equation 8. In this case, we use the actual values of future observations as forecast values. However, the resulting credit-to-GDP gap will not be identical with the one estimated by the "true" two-sided filter as here we only use the actual values over the forecasting horizon of 5 years.

Perfect forecast: $$\hat{y}_{t+h} = y_{t+h} \qquad (8)$$

Figure 1 shows an example of all forecasts (apart from the perfect forecast, which coincides with the actual development) for Spain in 2008Q4, just at the beginning of the financial crisis there. The charts nicely illustrate the pros and cons of various forecasting schemes when faced with a typical credit cycle development. First, the linear forecast is too optimistic, extending the credit boom, which Spain experienced since late 1990s, further into the future. The autoregressive approach works quite well over a horizon of about 1-2 years but fails to account for the decline of the time series in the medium-term. Using a regression-based model estimated on the data until 2008Q4 and at that time available vintage of the IMF WEO forecast (the October 2008 WEO) did not help either to capture the reversal in Spain's credit cycle, mainly due to the IMF WEO being still optimistic in 2008 as to the macroeconomic developments in Spain during the time of the Global Financial Crisis. They are even more optimistic than the AR process! Only the three correction-based projections have been somewhat able to get closer to the real evolution.

Figure 1 Illustration of Various Technical (Left-Hand Side) and Economic (Right-Hand Side) Forecast Schemes for Credit-to-GDP Ratio for Spain.

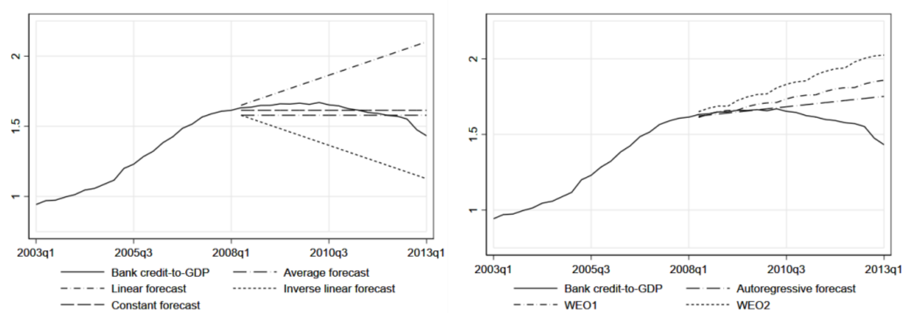
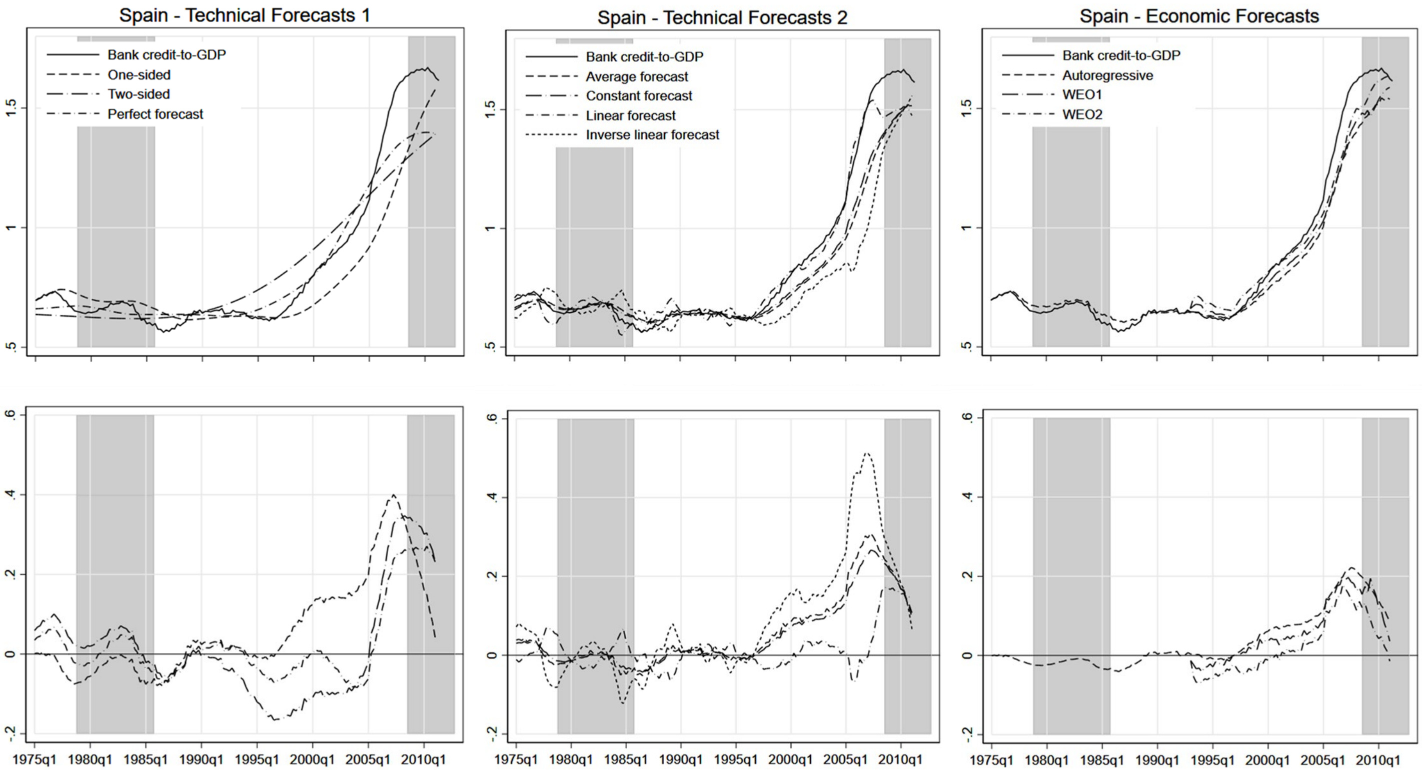


Figure 2 Trend Estimates (Top) for the Credit-to-GDP Ratio and Corresponding Credit-to-GDP Gaps (Bottom) for Spain Based on Various Technical (Left and Middle Part) and Economic (Right Part) Forecast Schemes.



Notes: Gray areas indicate banking crises.

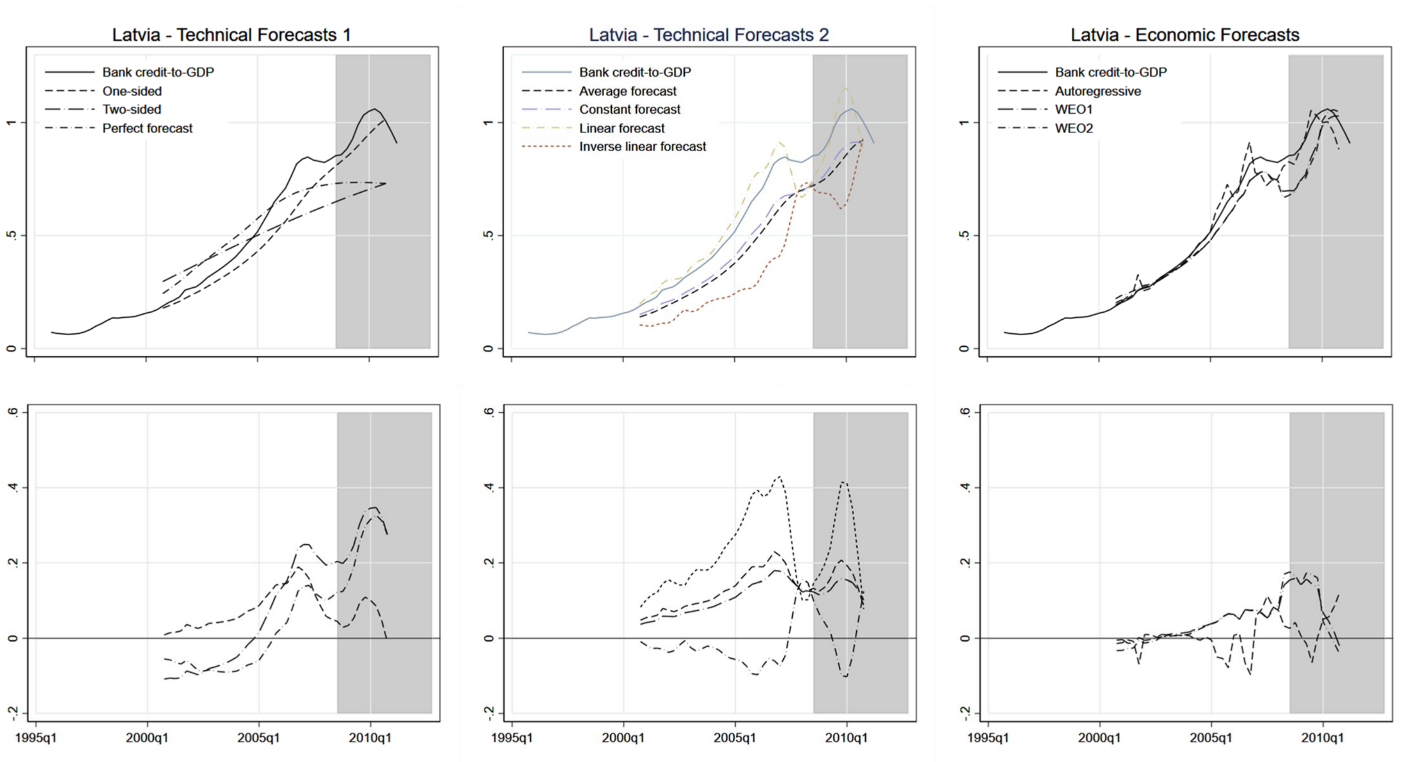
The left and middle parts of Figure 2 shows the trend and gap estimates for Spain based on the various technical forecast schemes "in real-time" for all periods of time (apart from the two-sided HP trend, which is using all data up to 2015, and the perfect forecast, which is using true outturns as forecast values). The left-hand side charts cover all the benchmarks, while the charts in the middle depict the technical forecasts, including those including some implicit corrections.

The one-sided filter confirms its good signaling power by opening the gap as early as in early 2000s, with a strong increase between 2005 and 2008. Here, even the two-sided filter or perfect forecast would not beat the simple benchmark, although both signal excessive credit expansion at least from 2005 on. The linear forecast would not help at all: by extending the credit boom pattern, it actually makes the end-point bias problem worse by moving the trend up and decreasing the gap. On the other hand, the correction-based projections decrease the trend somewhat, with inverse linear approach leading to the lowest trend and highest gaps. All in all, however, the one-sided gap would still remain difficult to beat.

When moving to the AR process and both economic forecasts (right part of Figure 2), the autoregressive forecast does a better job than both WEO-based forecasts in the late 1990s and early 2000s when the speed of financial deepening was strong but not as high as after 2005, but it would not beat the simple one-sided filter. In times of high credit growth, these types of forecasts, even if elaborate and based on regression models, would typically extend the observed speed of credit expansion to additional years, not getting rid of the end-point bias.

To illustrate the use of all forecast schemes for an emerging market with a shorter data availability but fast financial deepening, Figure 3 shows the trends and the gaps for Latvia, a country that went through a credit boom-and-bust cycle in the 2000s. Also, here the one-sided filter works quite well, opening the gap as early as in early 2000s once the speed of the credit-to-GDP increase started to accelerate. However, in the last two years before the crisis, with credit-to-GDP stabilizing around 80 percent, the one-sided filter is dominated by the long expansionary phase and the trend stays elevated, decreasing the gap. This contrasts with the two-sided filer, which, taking into account the correction in 2009-2013, better captures the true size of the gap in 2006-2007. A perfect foresight, i.e. knowing the development for the nearest five years, would help in those critical years, too. All correction-based projections would help in this respect. An opening of the gap in the case of linear forecast is caused by the small decrease in credit-to-GDP just before the crisis, but otherwise this method would not indicate any excessive credit expansion until 2007. The two WEO-based forecasts perform somewhat better here than in the case of Spain, leading to a lower (but still positive) gap compared to the one-sided trend in 2002-2006, but spiking too late – only in the last year before the crisis. Finally, the AR process leads to a quite volatile trend, but given the observed small correction in 2006-2007, it indirectly builds in a continuation of the correction, decreasing the trend somewhat and increasing the gap in 2007. This however reverses once the credit-to-GDP resumes its growth.

Figure 3 Trend Estimates (Top) for the Credit to GDP Ratio and Corresponding Credit to GDP Gaps (Bottom) for Latvia based on Various Technical (Left and Middle Part) and Economic (Right Part) Forecast Schemes.



Notes: Gray areas indicate banking crises.

Acknowledging the fact that knowing the future credit-to-GDP development would generally improve the early warning property of the gap (as two-sided gaps appear to have a better signaling property than one-sided gaps as reported by Drehmann and Tsatsaronis (2014), having a forecast with a low forecast error should bring us closer to the two-sided gap benchmark. To evaluate the various forecasts, we use two traditional indicators: the root mean square error (RMSE) and the mean absolute error (MAE). Both statistics are calculated in two ways. First, we follow the usual approach of calculating these indicators for each h-period ahead forecast separately across the sample with T observations, illustrated by the following formulas:

$$RMSE_h = \sqrt{\frac{\sum_{t=1}^{T} (\hat{y}_{t+h} - y_{t+h})^2}{T}} \tag{9}$$

$$MAE_h = \frac{\sum_{t=1}^{T} |\hat{y}_{t+h} - y_{t+h}|}{T} \tag{10}$$

Second, we also calculate the indicators for the forecast horizon H as a whole, which takes into account all horizons and can be interpreted as an average RMSE or MAE across all horizons:

$$RMSE = \frac{\sum_{t=1}^{T} \sqrt{\frac{\sum_{h=1}^{H} (\hat{y}_{t+h} - y_{t+h})^2}{H}}}{T} \tag{11}$$

$$MAE = \frac{\sum_{t=1}^{T} \sum_{h=1}^{H} |\hat{y}_{t+h} - y_{t+h}|}{HT} \tag{12}$$

Finally, to evaluate the predictive abilities of the various credit-to-GDP gaps, we follow Candelon, Dumitrescu and Hurlin (2012) and calculate the area under the ROC curve (AUROC/AUC). The ROC curve is based on the confusion matrix (Table 1), which assumes two possible states of the world, crisis ($D = 1$) and no crisis/calm times ($D = 0$), and two states of the signal given whether our variables of interest turns “on” ($S = 1$) if it exceeds the threshold θ, or remains “off” ($S = 0$) otherwise.

Table 1 Signals and Crises

		State of the world	
	Signal	*Crisis (D = 1)*	*No crisis (D = 0)*
Signal	On (S = 1)	True Positive (TP)	False Positive (FP) *Type 2 error*
	Off (S = 0)	False Negative (FN) *Type 1 error*	True Negative (TN)

The numbers in the matrix, such as the True Positives or False Positives, are dependent on the threshold θ. The lower θ the more crises can be correctly identified (higher TP), but at the same time there will also be more false alarms (higher FP). On the contrary, high values for θ lead to very few false alarms, but a lot of crises will be missed.

The predictive abilities of a signal are analyzed by looking at the trade-off between the true positive rate (TPR) and the false positive rate (FPR) for all possible values of θ, whereby the TPR and FPR are defined as:

$$TPR(\theta) = P(S = 1|D = 1) \tag{13}$$

$$FPR(\theta) = P(S = 1|D = 0) \tag{14}$$

While the optimal value for θ would depend on the preferences of the policymaker, i.e. the benefits of correctly identified crises and the cost of false alarms, the advantage of the ROC/AUC framework is that it doesn't require estimating the optimal threshold. Instead, the ROC curve plots the TPR against the FPR for all possible values for θ. Indicators with a better trade-off between TPR and FPR will have the curve more bowed in the north-west direction (with FPR on the horizontal axis), having much higher TPR than FPR for most values of θ. This trade-off can be quantified by the area under the ROC curve (AUROC/AUC), which would move between 0 and 1 as the natural limit for both TPR and FPR is 1. The higher AUC, the better signaling property does an indicator have, with AUC of 1 being a fully informative signal while an AUC of 0.5 representing an uninformative signal (i.e. toss of a coin). For details and use in the early warning literature, see e.g., Gersl and Jasova (2018).

6. Results

We first present the evaluation of the forecast accuracy for credit-to-GDP ratio, after which we analyze the signaling performance of the various credit-to-GDP gaps with respect to banking crises using the ROC/AUC framework. The objective is to compare the various forecasting schemes, looking at which of them would score well along both the accuracy and the early warning signals. Given that the IMF WEO forecasts used in generating our two economic forecasts are only available since 1990, while credit-to-GDP data is in many cases available from 1960s or 1970s, we proceed in two steps to rule out any effect that the number of observations could have. First, we report the results of technical forecasts for the whole sample (including those projections that build in a correction, i.e. linear, average, constant, inverse linear, and autoregressive). Second, we report the results for the economic forecasts on a smaller sample (since 1990), but for a proper comparison we also recalculate all the technical forecasts for the same sample.

6.1 Forecast Performance

Table 2 shows the accuracy of different types of technical forecasts as measured by the RMSE for selected horizons (4-, 8-, 12- and 20-quarters ahead) and the whole horizon H.[1]

Table 2 Forecast Performance - Technical Forecasts I

	Linear forecast	*Average forecast*	*Constant forecast*	*Inverse linear forecast*	*Autoregressive forecast*
h = 4	0.037	0.056	0.041	0.072	**0.034**
h = 8	0.078	0.084	0.071	0.132	**0.066**
h = 12	0.111	0.114	0.100	0.193	**0.099**
h = 16	0.155	0.142	0.132	0.250	**0.131**
h = 20	0.197	0.163	**0.150**	0.303	0.162
H	0.065	0.070	0.063	0.116	**0.059**

Notes: Forecast performance for different horizons h, with the last row with H showing the average forecast performance over all horizons. The results for the model with the best performance for a given horizon are shown in bold.

Clearly, the longer the horizon, the larger the error, going from about 3 percentage points in 1-year ahead forecasts up to more than 15 percentage points over the 5-year horizon. While the autoregressive forecast performs best on average and over horizons of up to 12 quarters, the constant forecast performs slightly better over the longest horizon of 20 quarters (the average forecast has a somewhat lower fit over the long horizon that the constant forecast but still comparable to the autoregressive forecast). This may capture the fact that over very long horizons, assuming an unchanged level of credit-to-GDP is actually a simplified way how to account in for the credit cycle, i.e. that over the medium term either credit-to-GDP will grow but correct again, or decline but recover in the future – something the AR process cannot account for. This result is broadly in line with Gerdrup, Kvinlog and Schaanning (2013) who find that the rolling average forecast performs quite well, although they have not analyzed a forecast generated by an autoregressive process. The linear forecast is comparable to the AR forecast over short horizons, but deteriorates over longer horizons, as it typically assumes that the credit grows at the same speed, while the AR would typically lead to a deceleration of the speed, which is often more probable over medium terms. As expected, the inverse linear forecast, building in a reversal of the trend, performs far worse than the other forecasts.

Table 3 shows the forecast performance separately for advanced (AE) and emerging market economies (EME). In general, the different forecast schemes perform better in case of advanced economies, but the overall picture remains unchanged.

[1] Results for the MAE are similar and can be received from the authors upon request.

Table 3 Forecast Performance - Technical Forecasts II

	Linear forecast		*Average forecast*		*Constant forecast*		*Inverse linear forecast*		*Autoregressive forecast*	
	AE	***EME***	***AE***	***EME***	***AE***	***EME***	***AE***	***EME***	***AE***	***EME***
h = 4	0.032	0.046	0.047	0.064	0.034	0.045	0.060	0.084	**0.030**	**0.042**
h = 8	0.061	0.085	0.071	0.100	0.061	0.081	0.111	0.160	**0.054**	**0.077**
h = 12	0.088	0.132	0.093	0.131	0.083	0.114	0.159	0.231	**0.077**	**0.111**
h = 16	0.119	0.169	0.112	0.152	0.104	0.136	0.203	0.297	**0.094**	**0.140**
h = 20	0.155	0.216	0.131	0.175	0.123	0.163	0.243	0.357	**0.110**	**0.173**
H	0.056	0.081	0.060	0.094	0.054	0.083	0.101	0.151	**0.050**	**0.077**

Notes: Entries in the left-hand subcolumns refer to the root mean squared error (RMSE) for advanced economies (AE), while entries in the right-hand subcolumn refer to the RMSE for emerging market economies (EME). The rows show the forecast performance for different horizons h, with the last row with H showing the average forecast performance over all horizons. The results for the model with the best performance for a given horizon and country group are shown in bold.

The autoregressive forecast performs best in most cases, but in case of emerging market economies it is again outperformed by the constant forecast. This is supportive of the observation that emerging markets' credit cycles might be more pronounced, with credit corrections more frequent than in advanced markets and thus an assumption of an unchanged level of credit-to-GDP more appropriate in certain cases.

Table 4 shows the performance of the two economic forecasts (WEO1 and WEO2) in comparison to the accuracy of technical forecasts, which have been recalculated for the new smaller sample. The new forecasts based on economic models perform quite well, but they cannot beat the autoregressive forecast over the short term or the constant forecast over the medium and long term. Interestingly, the simpler of the two (WEO1 with the real GDP growth as the only macroeconomic factor) tends to perform slightly better than the model with three macroeconomic factors. In the case of the whole forecast horizon H, the autoregressive forecast is the best, followed by the linear and constant forecasts, the two WEO-based forecasts, and with some distance the inverse linear forecast.

Table 4 Forecast Performance - Economic Forecasts I

	Linear forecast	*Average forecast*	*Constant forecast*	*Inverse linear forecast*	*Autoregressive forecast*	*WEO1*	*WEO2*
h = 4	0.041	0.058	0.042	0.077	**0.037**	0.041	0.044
h = 8	0.079	0.090	0.074	0.148	**0.070**	0.081	0.085
h = 12	0.115	0.116	**0.103**	0.217	0.104	0.116	0.113
h = 16	0.156	0.146	**0.132**	0.278	0.137	0.145	0.143
h = 20	0.202	0.166	**0.154**	0.333	0.165	0.168	0.175
H	0.080	0.095	0.084	0.155	**0.079**	0.089	0.090

Notes: Forecast performance for different horizons h, with the last row with H showing the average forecast performance over all horizons. The results for the model with the best performance for a given horizon are shown in bold.

Table 5 Forecast Performance - Economic Forecasts II

	Linear forecast		***Average forecast***		***Constant forecast***		***Inverse linear forecast***		***Autoregressive forecast***		***WEO1***		***WEO2***	
	AE	*EME*	*AE*	*EME*	*AE*	*EME*	*AE*	*EME*	*AE*	*EME*	*AE*	*EME*	*AE*	*EME*
h = *4*	0.041	**0.041**	0.052	0.062	0.039	0.045	0.066	0.082	**0.034**	0.042	0.036	0.051	0.038	0.051
h = *8*	0.075	0.083	0.079	0.096	0.069	0.079	0.121	0.158	**0.056**	**0.074**	0.057	0.090	0.058	0.094
h = *12*	0.102	0.128	0.102	0.126	0.096	**0.109**	0.170	0.227	**0.078**	0.115	0.081	0.132	0.085	0.139
h = *16*	0.132	0.169	0.127	0.150	0.121	**0.137**	0.216	0.291	**0.103**	0.146	0.106	0.166	0.116	0.165
h = *20*	0.169	0.215	0.154	0.176	0.147	**0.161**	0.253	0.349	**0.126**	0.174	0.129	0.205	0.150	0.221
H	0.072	0.088	0.087	0.103	0.078	0.092	0.146	0.163	**0.068**	**0.091**	0.078	0.100	0.081	0.101

Notes: Entries in the left-hand subcolumns refer to the root mean squared error (RMSE) for advanced economies (AE), while entries in the right-hand subcolumn refer to the RMSE for emerging market economies (EME). The rows show the forecast performance for different horizons h, with the last row with H showing the average forecast performance over all horizons. The results for the model with the best performance for a given horizon and country group are shown in bold.

The picture changes slightly when looking at the forecasting accuracy separately for advanced and emerging markets. Table 5 shows that while the autoregressive forecast performs best in case of advanced economies over all horizons, the results for EMEs are more mixed. In case of very short horizons, the linear forecast slightly outperforms the autoregressive one, while in case of medium to long horizons between 12 and 20 quarters, the constant forecast performs best. The economic forecasts perform better for advanced economies than for emerging markets but cannot beat many other forecasts.

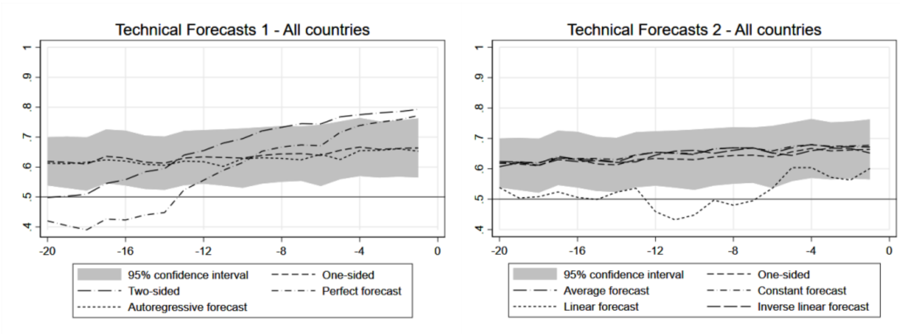
6.2 Signaling Performance

What effect do the different forecasts have on the signaling performance of the resulting credit-to-GDP gaps if we extend the credit-to-GDP series by such forecasts and perform a quasi-two-sided HP filtering? In this section, we report the AUCs calculated for alternative signaling horizons (from 20 to 1 quarter before a crisis starts) for the benchmarks (one-sided, two-sided, perfect forecast) and the various forecast schemes in the form of charts. The reported results also include the 95% confidence interval around the AUC of the simple one-sided HP-based gap to evaluate the statistical significance of other credit-to-GDP gaps in comparison with the benchmark.[2]

Figure 4 shows the AUCs of the credit-to-GDP gaps augmented by the technical forecasts using the larger sample. Our findings confirm what Drehmann and Tsatsaronis (2014) report – knowing the whole future development of credit-to-GDP would significantly increase the AUC over the horizons up to 2 years, while knowing the next 5 years would significantly improve the AUC at least over a 1-year horizon, in both cases from about 65% to almost 80%. Unfortunately, the technical forecasts would not improve the AUCs – most of them are around (such as AR) or only slightly above the one-sided benchmark, comfortably within the 95% confidence interval. Interestingly, the inverse linear forecast – the one with the worst accuracy and actually not a forecast at all – performs reasonably well, with AUCs around the one-sided gap. This indicates that building in a correction in the credit cycle is not as bad idea as it seems for getting a reasonable early warning indicator. The worst performance is achieved by the linear forecast, which is especially bad for horizons over 6 quarters. This is in line with our expectations. If we assume that credit-to-GDP continues to grow especially in cases where there has already been a long-lasting boom (and where a correction is more probable, eventually), using a linear forecast would amplify rather than reduce the end-point bias.

[2]Data tables underlying the charts are available from the authors upon a request.

Figure 4 AUCs for Technical Forecasts for All Countries

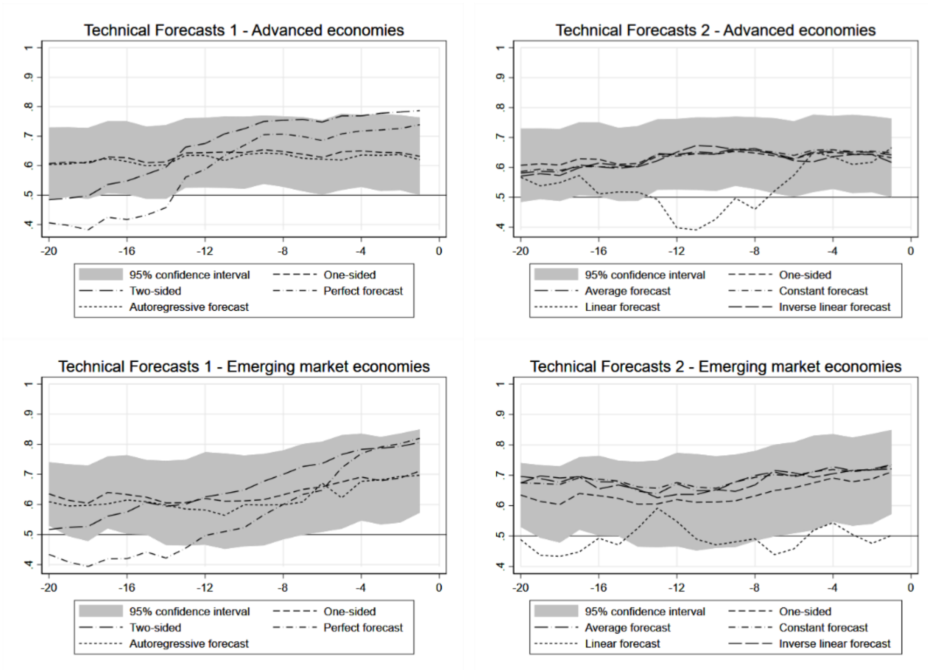


Notes: The horizontal axis denotes the forecast horizon in quarters before the crisis. The vertical axis denotes the AUC. The black horizontal line at 0.5 highlights the value of an uninformative indicator.

When looking separately at advanced economies and emerging markets, we can see that the two-sided filter-based gap has higher AUC value than the one-sided benchmark for both types of economies, but in term of statistical significance, this only holds for advanced economies and only at very short horizons of a few quarters (Figure 5). Interestingly, in our sample, the one-sided gap has somewhat higher AUC for emerging markets compared to advanced economies, a somewhat surprising finding given all the criticism of the Basel guidance for emerging markets and in contrast to Drehmann and Tsatsaronis (2014), who find that the AUC is typically higher for advanced economies. This discrepancy might be driven by the differences in countries included in the sample, slightly different coding of crises, and time coverage of the analysis.

In terms of whether applying the various technical forecast schemes improves the signaling property of the credit-to-GDP gap, we do not see any difference for the autoregressive forecast across the two groups of countries. The linear forecast clearly underperforms especially in emerging markets, given the frequent boom-bust pattern in credit evolution, while in advanced countries, this type of forecast performs relatively well (but not better than the one-sided gap) over short horizons. The remaining, correction-augmented gaps have actually higher AUCs than the one-sided benchmark in the case of emerging markets, but this difference is not statistically significant in this sample.

Figure 5 AUCs for Technical Forecasts Estimated Separately for Advanced (Top) and Emerging Market Economies (Bottom)

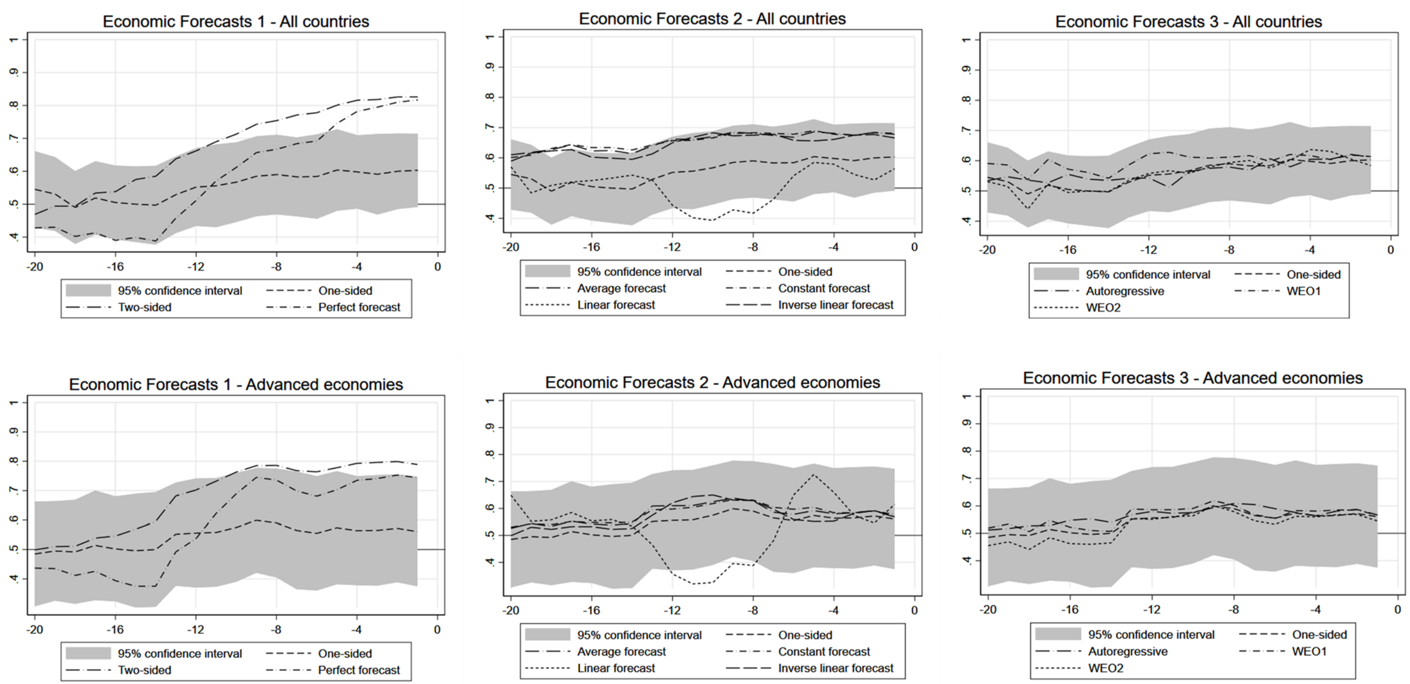


Notes: The horizontal axis denotes the forecast horizon in quarters before the crisis. The vertical axis denotes the AUC. The black horizontal line at 0.5 highlights the value of an uninformative indicator. The grey area shows the 95% confidence interval around the AUC of the one-sided gap.

Figure 6 reports the results for AUCs for all countries as well as separately for advanced and emerging market economies with the two economic forecasts added. As these results are based on a smaller sample, the AUCs for our benchmarks and for the technical forecasts were also recalculated. The smaller size also results in somewhat larger confidence intervals as captured by the broader grey bands in the following figures. Additionally, the signaling performance of the one-sided filter decreases compared to the previous sample, while the signaling performance of all correction-augmented gaps increases relative to the one-sided filter. However, their AUCs are still inside the confidence interval for horizons of up to 3 years, with only those around 4 years performing significantly better.

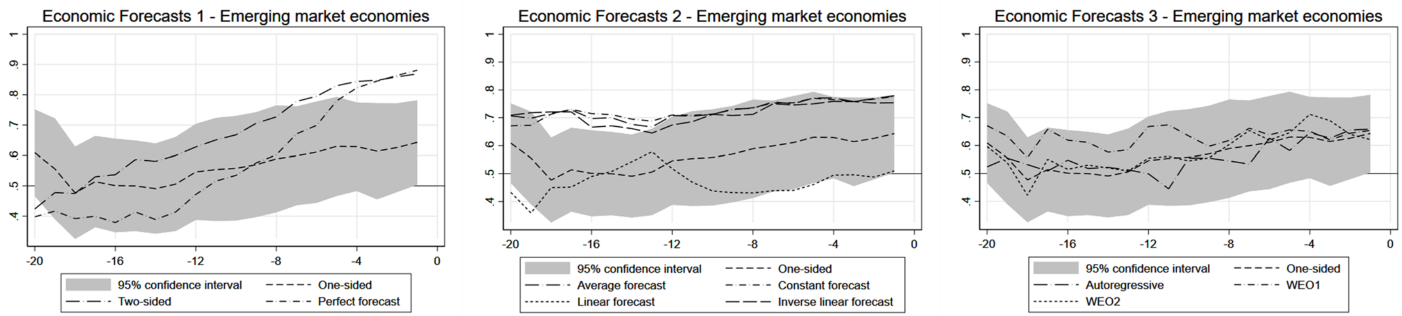
The two economic forecasts WEO1 and WEO2 (charts on the right side of the figure) perform as good as the one-sided benchmark in terms of statistical significance, with AUCs being very close to each other (and inside the confidence bands), even though the WEO1 forecast seems to perform somewhat better over longer horizons beyond 2 years.

Figure 6 AUCs for Technical and Economic Forecasts for All, Advanced and Emerging Market Countries



Notes: The horizontal axis denotes the forecast horizon in quarters before the crisis. The vertical axis denotes the AUC. The black horizontal line at 0.5 highlights the value of an uninformative indicator. The grey area shows the 95% confidence interval around the AUC of the one-sided gap.

Figure 6 AUCs for Technical and Economic Forecasts for All, Advanced and Emerging Market Countries Continued



Notes: The horizontal axis denotes the forecast horizon in quarters before the crisis. The vertical axis denotes the AUC. The black horizontal line at 0.5 highlights the value of an uninformative indicator. The grey area shows the 95% confidence interval around the AUC of the one-sided gap.

Interestingly, when we look at technical forecast at this smaller sample, those forecast schemes that build in a certain correction (constant, average and inverse linear) now have a higher AUC that the one-sided filter, but not with a statistically significant difference. The AR process is close to the one-sided filter, as in the previous larger sample, and the linear forecast is inferior again.

When we look separately on advanced economies and emerging markets, the results are broadly similar to the whole sample, with AUCs for both WEO forecasts not statistically different from the one-sided gap. We can also see that the higher AUCs of the three correction-augmented credit-to-GDP gaps reported for the whole set of countries are driven by emerging markets. For these forecast schemes in the case of emerging markets, the AUCs are around 75% compared to about 60% for the one-sided filter at a 6Q horizon, which is still only at the upper margin of the confidence interval of the one-sided benchmark given the larger confidence intervals. The difference becomes statistically significant only for longer horizons above 3 years.

Actually, higher AUCs for correction-augmented credit-to-GDP gaps in emerging markets have been found also in the case of the longer sample in Figure 5, although the difference in AUCs has not been so large (66% for one-sided filter versus around 70% for the correction forecasts for the 6Q horizon). As Tables A3 and A4 in the Appendix show, the shorter sample is dominated by emerging market crises (15 out of 17 crises included while for advanced countries, it is only 14 out of 30), that's why we see the higher AUCs of those forecasts also in the overall sample across all country types.

Despite the marginal statistical significance, our findings in this area suggest that for emerging market economies in credit booms, it might be a good idea to complement the traditional one-sided-filtered credit-to-GDP gap with an alternative indicator by augmenting the credit-to-GDP series with a projection that builds in a certain correction in the credit cycle over a medium term, essentially removing the end-point bias. This is actually used in some countries in practice already, such as in Lithuania.

7. Conclusions

In our research, we analyzed whether extending the credit-to-GDP series by a forecast and calculating the credit-to-GDP gap as an indicator of excessive credit expansions using the quasi-two-sided HP filter reduces the well-known end-point bias of one-sided HP filter and improves the early warning property of this indicator in terms of signaling future banking crises. This would be useful for policymakers charged with calibrating the Basel III countercyclical capital buffer, for which the Basel Committee recommended to use the credit-to-GDP gap as a conditioning variable.

We extended the work that was initiated by Gerdrup, Kvinlog and Schaanning (2013), who conducted a similar exercise for Norway. However, we focused on a large sample of 56 countries over the period 1950s-2016 and produced a number of alternative real-time forecasts for credit-to-GDP. We tested three types of forecasts: (i) simple statistical (technical) forecasts such as an autoregressive forecast and linear trend, (ii) economic forecasts based on regression models estimated in quasi-real-time and relying on at that time available vintages of the IMF WEO macroeconomic forecasts for GDP, inflation, and current account balance, and (iii) projections building in some plausible correction in the credit developments, so that credit-to-GDP would stabilize or even decrease. We compared all adjusted gaps with three benchmarks: the one-sided HP filter-based gap, the two-sided gap, and a perfect foresight-based gap, evaluated the forecast accuracy with standard statistics such as RMSE and MAE and tested their early warning properties using a database of banking crises by Laeven and Valencia (2013) within the ROC/AUC framework.

Our findings can be summarized as follows. First, forecasts with relatively good accuracy over the short and medium term, such as the linear forecast, autoregressive forecast or the economic forecasts based on the WEO projections, do not help to increase the signaling performance of the credit-to-GDP gap and do not beat the benchmark of a simple one-sided HP filter. This holds for both advanced economies and emerging markets. This is a puzzle given that the simple unadjusted gaps are clearly biased. Second, for emerging markets, we find that the correction-augmented credit-to-GDP gaps – which often have inferior forecasting accuracy – might to improve the signaling power at the margin of statistical significance. Thus, when an emerging economy is in a credit boom, the traditional one-sided HP filter could be complemented by an alternative indicator of the credit-to-GDP gap based on augmenting the credit-to-GDP series with a projection that builds in a correction in the credit cycle over a medium term.

APPENDIX

A1. Availability of Data

Table A1 Data Availability for Advanced Economies

Country		***Historical WEO data (Vintage)***			***WEO data***		
	Bank credit-to-GDP	*GDP*	*CPI*	*BCA*	*GDP*	*CPI*	*BCA*
Australia	1960q2	S1990	S1990	S1990	1980q1	1980q1	1980q4
Austria	1960q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Belgium	1970q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Canada	1955q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Denmark	1966q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Finland	1974q1	S1990	S1990	S1990	1980q1	1980q1	1980q4
France	1969q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Germany	1960q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Greece	1970q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Ireland	1971q2	S1990	S1990	S1990	1980q1	1980q1	1980q4
Italy	1974q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Japan	1963q1	S1990	S1990	S1990	1980q1	1980q1	1980q4
Netherlands	1961q1	S1990	S1990	S1990	1981q1	1981q1	1980q4
New Zealand	1960q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Norway	1960q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Portugal	1960q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Spain	1970q1	S1990	S1990	S1990	1980q1	1980q1	1980q4
Sweden	1961q1	S1990	S1990	S1990	1980q1	1980q1	1980q4
Switzerland	1960q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
United Kingdom	1963q1	S1990	S1990	S1990	1980q1	1980q1	1980q4
United States	1952q1	S1990	S1990	S1990	1980q1	1980q1	1980q4

Notes: S and F denote the spring and fall publication of the WEO, respectively. GDP = real GDP growth; CPI = CPI inflation; BCA = Current account balance in percent of GDP;

Table A2 Data Availability for Emerging Market Economies

Country		Historical WEO data (Vintage)			WEO data		
	Bank credit-to-GDP	*GDP*	*CPI*	*BCA*	*GDP*	*CPI*	*BCA*
Albania	1995q4	S1993	S1993	S1993	1980q1	1990q1	1980q4
Argentina	1984q4	S1990	S1990	S1990	1980q1	1998q1	1980q4
Armenia	2000q1	S1993	S1993	S1993	1993q1	1993q1	1992q4
Belarus	1999q4	S1993	S1993	S1993	1993q1	1993q1	1992q4
Bosnia and Herzegovina	1997q3	F1994	F1999	F2002	1997q1	1997q1	1998q4
Brazil	1995q1	S1990	S1990	S1990	1980q1	1980q1	1980q4
Bulgaria	1997q4	S1990	F1990	S1991	1980q1	1981q1	1980q4
Chile	1998q1	S1990	S1990	S1990	1980q1	1980q1	1980q4
China	1985q4	S1990	S1990	S1990	1980q1	1981q1	1997q4
Croatia	1997q4	S1994	S1994	F1994	1993q1	1993q1	1992q4
Czech Republic	1993q1	S1994	S1994	S1994	1996q1	1996q1	1995q4
Estonia	1995q4	S1993	S1993	S1993	1994q1	1994q1	1993q4
Georgia	2003q1	S1993	S1993	S1993	1995q1	1995q1	1995q4
Hong Kong	1978q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Hungary	1989q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
India	1951q2	S1990	S1990	S1990	1980q1	1980q1	1980q4
Indonesia	1976q1	S1990	S1990	S1990	1980q1	1980q1	1980q4
Israel	1990q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Latvia	1995q4	S1993	S1993	S1993	1993q1	1993q1	1992q4
Lithuania	1995q4	S1993	S1993	S1993	1996q1	1996q1	1995q4
Macedonia	1995q4	S1994	S1994	F1994	1993q1	1993q1	1992q4
Malaysia	1964q2	S1990	S1990	S1990	1980q1	1980q1	1980q4
Mexico	1980q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Moldova	2000q4	S1993	S1993	S1993	1993q1	1993q1	1992q4
Montenegro	2002q4	S2008	S2008	S2008	2001q1	2000q1	2001q4
Poland	1992q1	S1990	S1990	S1990	1980q1	1980q1	1980q4
Romania	1996q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Russia	1995q2	S1993	S1993	S1993	1993q1	1993q1	1992q4
Serbia	2002q1	F2007	F2007	F2007	1999q1	1998q1	2000q4
Slovakia	1995q4	S1994	S1994	S1994	1994q1	1994q1	1993q4
Slovenia	1995q4	S1994	S1994	F1994	1993q1	1993q1	1992q4
South Africa	1965q1	S1990	S1990	S1990	1980q1	1980q1	1980q4
South Korea	1962q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Thailand	1970q4	S1990	S1990	S1990	1980q1	1980q1	1980q4
Turkey	1986q1	S1990	S1990	S1990	1980q1	1980q1	1980q4
Ukraine	1996q4	S1993	S1993	S1993	1993q1	1993q1	1992q4

Notes: S and F denote the spring and fall publication of the WEO, respectively. GDP = real GDP growth; CPI = CPI inflation; BCA = Current account balance in percent of GDP.

A2. Crises in the Sample

Table A3 Crises in Advanced Economies

Country	*Start*	*End*	*Included in signalling analysis (technical)*	*Included in signalling analysis (non-technical)*
Australia	1990q1	1992q4	x	-
Austria	2008q3	2009q4	x	x
Belgium	2008q3	2009q4	x	x
Canada	1983q1	1985q4	x	-
Denmark	1987q1	1993q4	x	-
Denmark	2008q3	2010q4	x	x
Finland	1991q3	1993q4	x	-
France	1994q1	1995q4	x	x
France	2008q3	2009q4	x	x
Germany	1977q1	1979q4	x	-
Germany	2008q3	2010q4	x	x
Greece	1991q1	1995q4	x	-
Greece	2008q3	2010q4	x	x
Iceland	1985q1	1986q4	-	-
Iceland	1993q1	1993q4	-	-
Iceland	2008q3	2010q4	-	-
Ireland	2008q3	2010q4	x	x
Italy	1990q1	1995q4	x	-
Italy	2008q3	2012q4	x	x
Japan	1991q1	2001q1	x	-
Netherlands	2008q3	2009q4	x	x
New Zealand	1987q4	1990q4	x	-
Norway	1988q4	1993q4	x	-
Portugal	2008q3	2009q4	x	x
Spain	1978q4	1985q4	x	-
Spain	2008q3	2012q4	x	x
Sweden	1991q3	1995q4	x	-
Switzerland	1991q1	1996q4	x	-
Switzerland	2007q3	2009q2	x	x
United Kingdom	1974q1	1976q4	x	-
United Kingdom	2007q3	2009q4	x	x
United States	1984q1	1991q4	x	-
United States	2007q3	2009q4	x	x

Sources: Laeven and Valencia (2013) and Babecky et al. (2014), with additions from Mehrez and Kaufmann (2000) and Kaminsky and Reinhart (1999).

Table A4 Crises in Emerging Market Economies

Country	*Start*	*End*	*Included in signalling analysis (technical)*	*Included in signalling analysis (non-technical)*
Albania	1994q1	1994q4	-	-
Argentina	1980q1	1982q4	-	-
Argentina	1989q4	1991q4	-	-
Argentina	1995q1	1995q4	x	x
Argentina	2001q4	2003q4	x	x
Armenia	1994q1	1994q4	-	-
Belarus	1995q1	1995q4	-	-
Bosnia and Herzegovina	1992q1	1996q4	-	-
Brazil	1990q1	1990q4	-	-
Brazil	1994q4	1998q4	-	-
Bulgaria	1994q1	1997q4	-	-
Chile	1976q1	1976q4	-	-
Chile	1981q1	1985q4	-	-
China	1998q1	1998q4	x	x
Croatia	1998q1	1999q4	-	-
Czech Republic	1996q2	1999q4	-	-
Estonia	1992q4	1994q4	-	-
Estonia	1998q1	1999q4	-	-
Georgia	1991q1	1995q4	-	-
Hungary	1991q1	1995q4	-	-
Hungary	2008q3	2009q4	x	x
India	1993q1	1993q4	x	-
Indonesia	1997q4	2001q4	x	x
Israel	1977q1	1985q4	-	-
Latvia	1995q2	1998q4	-	-
Latvia	2008q3	2012q4	x	x
Lithuania	1995q4	1996q4	-	-
Lithuania	2008q3	2009q4	x	x
Macedonia	1993q1	1995q4	-	-
Malaysia	1997q3	1999q4	x	x
Mexico	1981q1	1985q4	-	-
Mexico	1994q4	1996q4	x	x
Poland	1992q1	1994q4	-	-
Romania	1990q1	1992q4	-	-
Russia	1998q3	1999q4	-	-
Russia	2008q3	2009q4	x	x
Slovakia	1998q1	2002q4	-	-
Slovenia	1992q1	1992q4	-	-
Slovenia	2008q3	2013q4	x	x
South Korea	1997q3	1998q4	x	x
Thailand	1983q1	1983q4	x	-
Thailand	1997q3	2000q4	x	x
Turkey	1982q1	1984q4	-	-
Turkey	2000q4	2001q4	x	x
Ukraine	1998q3	1999q4	-	-
Ukraine	2008q1	2015q4	x	x

Sources: Laeven and Valencia (2013) and Babecky et al. (2014), with additions from Mehrez and Kaufmann (2000) and Kaminsky and Reinhart (1999).

REFERENCES

Apel M, Hansen J, Lindberg H (1996): Potential Output and Output Gap. *Quarterly Review of the Bank of Sweden*, pp. 24–35.

Babecky J et al. (2014): Banking, Debt, and Currency Crises in Developed Countries: Stylized Facts and Early Warning Indicators. *Journal of Financial Stability* 15:1–17.

Baxter M, King RG (1999): Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series. *The Review of Economics and Statistics* 81(4):575–593.

BCBS (2010a): An Assessment of the Long-Term Economic Impact of Stronger Capital and Liquidity Requirements. *Basel Committee on Banking Supervision*.

BCBS (2010b): Guidance for National Authorities Operating the Countercyclical Capital Buffer. *Basel Committee on Banking Supervision.*

Borio C, Lowe P (July 2002): Asset Prices, Financial and Monetary Stability: Exploring the Nexus. *BIS Working Papers* 114.

Candelon B, Dumitrescu EI, Hurlin C (2012): How to Evaluate an Early-Warning System: Toward a Unified Statistical Framework for Assessing Financial Crises Forecasting Methods. *IMF Economic Review* 60(1):75–113.

Claessens S, Kose A, Terrones M (2011): How Do Business and Financial Cycles Interact? *IMF Working Paper* WP/11/88.

Dell'Ariccia G et al. (2012): Policies for Macrofinancial Stability: How to Deal with Credit Booms. *IMF Staff Discussion Note* 12/06.

Detken C et al. (2014): Operationalising the Countercyclical Capital Buffer: Indicator Selection, Threshold Identification and Calibration Options. *ESRB Occasional Paper Series.*

Drehmann M, Borio CEV, Tsatsaronis K (2011): Anchoring Countercyclical Capital Buffers: the Role of Credit Aggregates. *BIS Working Papers*.

Drehmann M, Juselius M (2014): Evaluating Early Warning Indicators of Banking Crises: Satisfying Policy Requirements. *International Journal of Forecasting* 30(3):759–780.

Drehmann M, Tsatsaronis K (Mar. 2014): The Credit-to-GDP Gap and Countercyclical Capital Buffers: Questions and Answers. *BIS Quarterly Review*.

Drehmann M, Yetman J (2018): Why You Should Use the Hodrick-Prescott Filter–at Least to Generate Credit Gaps. *BIS Working Papers* 744

Drehmann M et al. (2010): Countercyclical Capital Buffers: Exploring Options. *BIS Working Papers 317*.

Edge RM, Meisenzahl RR (2011): The Unreliability of Credit-to-GDP Ratio Gaps in Real-Time: Implications for Countercyclical Capital Buffers. *International Journal of Central Banking* 7(4):261–298.

Gerdrup K, Kvinlog AB, Schaanning E (2013): Key Indicators for a Countercyclical Capital Buffer in Norway–Trends and Uncertainty. *Norges Bank Staff Memo* 13.

Geršl A, Jašová M (2018): Credit-Based Early Warning Indicators of Banking Crises in Emerging Markets. *Economic Systems* 42(1):18–31.

Geršl A, Seidler J (2015): Countercyclical Capital Buffers and Credit-to-GDP Gaps: Simulation for Central, Eastern, and Southeastern Europe. *Eastern European Economics* 53(1):439–465.

Hamilton JD (2018): Why You Should Never Use the Hodrick-Prescott Filter. *The Review of Economics and Statistics* 100(5):831–843.

Hodrick RJ, Prescott EC (1997): Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking* 29(1):1–16.

Kaiser R, Maravall A (1999): Estimation of the Business Cycle: A Modified Hodrick-Prescott filter. *Spanish Economic Review* 1:175–206.

Kaminsky GL, Reinhart CM (1999): The Twin Crises: The Causes of Banking and Balance-Of-Payments Problems. The American Economic Review 89(3):473–500.

Laeven L, Valencia F (2013): Systemic Banking Crises Database. *IMF Economic Review* 61(2)225–270.

Mehrez G, Kaufmann D (2000): Transparency, Liberalization, and Financial Crisis. *World Bank Policy Research Working Paper* 2286.

Mise E, Kim TH, Newbold P (2005): On Suboptimality of the Hodrick-Prescott Filter at Time Series Endpoints. *Journal of Macroeconomics* 27(1):53–67.

Ravn MO, Uhlig H (2002): On Adjusting the Hodrick-Prescott Filter for the Frequency of Observations. *Review of Economics and Statistics* 84(2):371–376.

Repullo R, Saurina J (2011): The Countercyclical Capital Buffer of Basel III: A Critical Assessment. *CEMFI Working Paper* 11/02.

Schularick M, Taylor AM (2012): Credit Booms Gone Bust: Monetary Policy, Leverage Cycles, and Financial Crises, 1870-2008. *American Economic Review* 102(2):1029–61.

St-Amant P, Van Norden S (1997): Measurement of the Output Gap: A discussion of recent research at the Bank of Canada. 79.

Wolf E, Mokinski F, Schüler YS (2020): On Adjusting the One-Sided Hodrick-Prescott Filter. *Deutsche Bundesbank Discussion Paper* 11.