Working Paper Series

No 114 / February 2021

Financial crises, macroprudential policy and the reliability of credit-to-GDP gaps

by
Piergiorgio Alessandri
Pierluigi Bologna
Maddalena Galardo

ESRB
European Systemic Risk Board
European System of Financial Supervision

Abstract

The Basel III regulation explicitly prescribes the use of Hodrick-Prescott filters to estimate credit cycles and calibrate countercyclical capital buffers. However, the filter has been found to suffer from large ex-post revisions, raising concerns on its fitness for policy use. To investigate this problem we study credit cycles in a panel of 26 countries between 1971 and 2018. We reach two conclusions. The bad news is that the limitations of the one-side HP filter are serious and pervasive. The good news is that they can be easily mitigated. The filtering errors are persistent and hence predictable. This can be exploited to construct real-time estimates of the cycle that are less subject to ex-post revisions, forecast financial crises more reliably, and stimulate the build-up of bank capital before a crisis.

JEL classification: E32; G01; G21; G28.

Keywords: Hodrick-Prescott filter, credit cycle, macroprudential policy

1 Introduction

Recessions following financial crises are twice as costly than normal business cycles downturns (Jordà et al. (2011) and Schularick and Taylor (2012)).The procyclical behavior of the financial sector played a critical role in amplifying the impact of the Great Financial Crisis of 2008-2009 and its aftershocks, including the European sovereign debt crisis of 2011-2012. In response to it, regulators created a new set of countercyclical policy tools that should push banks to build up precautionary capital buffers in 'good times' and release them in 'bad times', rendering the financial sector more stable and the supply of credit less volatile. The countercyclical capital buffer (CCyB) introduced in Basel III follows this logic and is intended to play a pivotal role in protecting the banking sector from boom-and-bust credit cycles, European Parliament (2013). However, countercyclical measures are only as good as the financial cycle estimates they rely on. A key question for policymakers is thus how this can or should be measured. Can credit bubbles be identified in real-time? And how to form a view on whether credit is too high, too low or about right given the needs of the real economy?

Early research identified a promising option in the cyclical component extracted from credit-to-GDP ratios by means of a one-side Hodrick-Prescott (HP) filter: the resulting "credit gap" appeared to be at once a powerful predictor of financial crises and an intuitive and robust tool to measure financial imbalances (Drehmann et al. (2010)). Credit gaps were thus explicitly introduced in the regulatory package of Basel III (BCBS (2011)). Subsequent investigations cast doubts on the validity of this approach. Edge and Meisenzahl (2011) and Alessandri et al. (2015) document that the credit gap estimates based on the HP filter are subject to large ex-post revisions, with dramatic policy implications. In particular, the 'false positives' generated by an overly volatile filter would have caused historically a number of unnecessary tightening in capital requirements. Darracq Pariès et al. (2019) emphasizes the opposite problem, namely that the Basel gap might be biased downwards after a prolonged credit boom insofar as the boom causes an upward bias in the estimated trend component. A more systematic and drastic critique of HP filtering is laid out by Hamilton (2018), who concludes that HP filters should have no place in a macroeconomist's toolbox. In practice, HP-based credit gaps play an important role for many of the authorities that fall under the remit of the BCBS. Hence, this methodological debate has important implications for the concrete management of bank capital buffers and for financial stability around the world.

In this paper, we provide new evidence on the issue by studying the behavior of credit gaps in 26 countries between 1971 and 2018. Our first contribution is to show that the shortcomings of the one-side HP filter are not only quantitatively significant but also extremely pervasive, both across countries and over time. The ex-post corrections to the HP-filtered gap can be as large as the gap itself, rendering the filter effectively useless in real-time. Our second contribution is to demonstrate that they are not lethal. The filtering errors are highly persistent and hence predictable. This opens the way to a simple now-casting procedure that allows policymakers to obtain better estimates of the credit gap without departing from the Basel III prescription. In a nutshell, this consists of: (i) estimating a sequence of filtering errors, i.e. historical discrepancies between one-side and two-side estimates of the cycle; (ii) forecasting this discrepancy to obtain an estimate of its (unobserved) current value; and (iii) using this forecast, or nowcast, to correct the one-side estimate of the cycle. This procedure delivers credit cycle estimates that are less volatile, less subject to revisions and more correlated with financial crises than those obtained from the plain HP filter. When used as an input for the Basel III policy rule, they also generate higher capital requirements at the onset of the financial crises included in our sample. Importantly, the procedure is easy to implement and fully consistent with Basel III. In its simplest form, it can be implemented using exclusively the quarterly credit-to-GDP series used for the Basel gap. Our main conclusion is not that HP filters provide the "best" possible estimates of the credit cycle or the most reliable warnings on the likelihood of a financial crisis, but rather that these estimates are broadly fit for policy use and can be easily refined without radical departures from the Basel prescriptions.

Economists have been aware of the unreliability of the HP filter at least since Orphanides and Norden (2002). In the financial stability arena, the limitations of the Basel credit gap have spurred the creation of a wide range of alternative indicators of financial imbalances and/or crisis prediction methods. Jordà et al. (2017) argue that credit cycles are not necessarily longer than business cycles and that they should be measured scaling credit by population rather than GDP. Hamilton (2018) proposes a general alternative to HP filtering based on linear projections; Drehmann and Yetman (2018) find that this approach performs poorly in the case of credit-to-GDP ratios and that the Basel credit gap is not easily beaten by alternative measures based on different filtering techniques. Baba et al. (2020) study multivariate filters where credit cycles are estimated jointly along with cyclical imbalances in output and interest rates. Two radically different approaches are put forward by Alessi and Detken (2018) and Adrian et al. (2019), who use respectively decision trees and predictive distributions on future output growth to measure systemic risk in the financial sector. Greenwood et al. (2020) provide further evidence on predictability, showing that financial crises in the post-war era have been systematically anticipated by strong growth in credit and asset prices. Our work contributes to this literature and it is largely complementary to these papers. Instead of exploring alternative modeling strategies, we focus on investigating what can be done to fix the shortcomings of the HP filter without departing from the Basel framework.

The remainder of the paper is organized as follows. Section 2 describes the data. Section 3 presents new stylized facts on the behavior of one-side and two-side HP filters in a panel of 26 countries between 1971 and 2018. Section 4 introduces the "simple fix" to correct the endpoint problems of the HP filters. Section 5 tests the performance of our strategy using a range of alternative criteria, including the volatility and persistence of the estimated credit cycle, its capability to predict historical crisis episodes, and its implications for the calibration of the countercyclical capital buffer in good and bad times. Section 6 concludes.

2 Data

The bulk of our empirical analysis is based on aggregate credit-to-GDP ratios obtained from the BIS website (`https://www.bis.org/statistics/index.htm`).

The series published by the BIS cover 44 economies starting at the earliest in 1961. The credit-to-GDP ratio for each country is based on total credit to non-financial corporations (both privately- and publicly-owned), households and non-profit institutions, provided by all sources, including domestic banks, other domestic financial corporations, non-financial corporations, and non-residents.[1] We do not use real-time data: our analysis focuses on the (in)accuracy of the one-side filter relative to its two-side counterpart, and it abstracts from the estimation noise associated to data revisions. Edge and Meisenzahl (2011) find revisions to be small and inconsequential in the case of the US. To use reasonably long times series we restrict the sample to economies for which data are available for at least 40 years. The resulting dataset is sufficiently broad to allow us to draw general conclusions on the efficacy of our estimation procedure: it spans the period between 1971Q1 and 2018Q4 and it covers 26 countries, including the G8 economies, most European countries and a broad selection of emerging markets. Descriptive statistics on the credit-to-GDP ratios are provided in Table 1.

In Section 5 we also run predictive regressions using a set of (discrete) "systemic crises". These are obtained from three complementary sources. The first one is the ESRB-European Financial Crises database documented in Lo Duca et al. (2017).[2] The other ones, which cover also non-European countries, are Laeven and Valencia (2020) and Jordà et al. (2017). For countries that appear in more than one dataset, we count as crises all periods (i.e. quarters) that are classified as such by at least one of the sources considered. This simple, agnostic approach allows us to bypass the problem of taking a stance on each specific historical episode. It is also conservative, in the sense that it captures all the periods of financial distress that a (risk-averse) macroprudential authority would have presumably wished to face with a strongly capitalized banking sector. The merged dataset allows us to study 37 crisis episodes for 22 countries between 1971 and 2018.[3]

[1] See `https://www.bis.org/statistics/about_credit_stats.htm?m=6|380` for further details.

[2] Data are available at `https://www.esrb.europa.eu/pub/financial-crises/html/index.en.html`

[3] Table A.1 reports the crisis periods country by country. Both Laeven and Valencia (2020) and Jordà et al. (2017) do not report periods of crisis for Canada, New Zeland, Singapore, and South Africa.

3 Stylized facts on HP filters and credit gaps

Does the one-side HP filter give a misleading picture of the credit cycle? And what are the main features of the discrepancies between the filtered (computed real-time by using one-side HP filter) and smoothed (computed ex-post by using two-side HP filter) estimates of the cycle?

To answer these questions we present new estimates of the credit cycle in a panel of 26 countries, examining the stylized facts on the relation between filtered estimates, smoothed estimates and ex-post revisions (defined as the difference between smoothed and filtered estimates). Table 2 reports some basic summary statistics.[4] Columns 1 to 3 of the table show the mean of the filter, the smoother and the revision for each country. The average revision is negative in almost all cases, suggesting that the filter tends to generate an upward bias in the estimated credit gap. They are also very large, consistently with Edge and Meisenzahl (2011) and Alessandri et al. (2015). The averages reported in the bottom row show that in the pooled data the revision turns out to be of the same order of magnitude as the cycle itself.

The filter is by construction more volatile than the smoother. The standard deviations reported in columns 4 and 5 show that this excess volatility is modest on average, at roughly 80 basis points, but exhibits significant variation across countries: it exceeds 2 percentage points, for instance, in the case of Ireland or Japan. The most striking result is shown in the last column of the table: the correlation between real-time estimates and ex-post revisions is large and negative in all countries, with a cross-sectional average of -0.49. This confirms that the one-side filter suffers from a systematic overshooting problem: it delivers estimates of the credit cycle that are too large in absolute terms, and hence systematically revised downwards when more data becomes available. This feature is persistent over time as shown in Figure 1. The problems this causes to policymakers are obvious. Given the magnitude of the revisions (column 3), the negative correlation

[4]Note that the calculations are based on the entire sample period 1970Q1-2018Q4: the smoothing revisions would have *not* been available in real-time. The question of how the patterns in the data can be exploited in real-time will be tackled in Sections 4 and 5.

implies that even drawing a conclusion on whether the cycle is expansionary or contractionary at a given point in time is intrinsically very difficult. The analysis above exploits the full sample period 1971-2018. By using all available observations we obtain the best two-sided estimate of the credit cycle that can be obtained through HP filtering, but we may also introduce two distortions. The estimates might be (i) too volatile at the beginning of the sample, where filter and smoother can be unstable; and (ii) biased at the end of the sample, where the smoother converges to the filter. To get around this issue we recalculate the evaluation statistics only for the middle section of the sample, i.e. 1981Q1-2007Q4. Removing the first 10 years should be sufficient for the convergence of the smoother (Geršl and Seidler (2015) and Drehmann and Juselius (2014)), while a sample ending in 2007 removes the potential influence of the Great Financial Crisis and limits the end-of sample bias of the two-side HP filter estimates. The results, reported in Table 3, are consistent with the findings for the full sample: the revision is even larger, with a cross-sectional average of -3.34 $vis - \grave{a} - vis$ -1.11 for the full sample (Table 2). Its correlation with the filter is negative in all but one country, with an average value of approximately -0.4 and a maximum of -0.8.

4 A simple strategy to fix the endpoint problem

The evidence in Section 3 shows that the ex-post revisions to a credit-to-GDP gap obtained from one-side HP filter are about as large as the gap itself. In this respect, the results obtained by Edge and Meisenzahl (2011) and Alessandri et al. (2015) respectively for the US and Italy extend easily to all developed and emerging economies included in our panel. We now demonstrate that – precisely *because* they are large and persistent – the one-side HP filtering errors are also relatively easy to forecast based on the past history of the revision. The forecasts can then be used to correct the filter, obtaining a measure of the credit cycle that is both empirically credible and consistent with the logic of Basel III. We illustrate the methodology in this section and assess its performance in Section 5.

To fix ideas, we define as $F_{t|t}$ the one-side (filter) estimate of the credit gap at time t, which is based on a time-t information set, and as $S_{t|T}$ the two-side (smoother) estimate of the same credit gap, which relies instead on information up to time $T > t$.[5] Smoother and filter diverge significantly inside the sample ($S_{\tau|T} \neq F_{\tau|t}$ for $\tau \neq T$), but converge by construction at the end of the sample, where the hindsight advantage of the smoother disappears ($\lim_{t \to T} S_{t|T} = F_{t|t}$). The procedure proposed by Alessandri et al. (2015) exploits the information contained in the history of the filter and the smoother to improve its performance at the $t = T$ boundary. It can be summarized as follows:

1. Estimate $F_{t|t}$ and $S_{t|T}$ using all information available until today T
2. Calculate a series of smoothing (ex post) corrections: $C_t = S_{t|T} - F_{t|t}$
3. Using a truncated sample $\mathbb{C}_{t-h} = (C_1, ..., C_{T-h})$ and a generic model $\mathcal{M}$, generate a nowcast of the current correction $\widehat{C}_T = \mathcal{M}(X_T, \mathbb{C}_{T-h})$, where X_T denotes additional information available at time T.
4. Revise the current value of the filter using the estimated correction: $\widehat{F}_{T|T} = F_{T|T} + \widehat{C}_T$.

Provided the difference between filter and smoother is predictable, the inaccuracy of the filter at the end of the sample (where policymakers need it the most) can be reduced by adding a model-based prediction of the as-yet-unobserved smoothing correction. Dropping some observations in step (3) is important because owing to the gradual convergence between filter and smoother, the correction drops mechanically to zero as t approaches T.

The choice of the horizon h, the model $\mathcal{M}$ and the auxiliary information set X can of course be important in practice, and is far from trivial. In what follows, as in Alessandri et al. (2015), we deliberately stick to two extremely naive forecasting models, namely (1) a random walk where C_t is simply held constant at some past

[5] We rely throughout the analysis on a smoothing parameter λ=400,000, as prescribed by the Basel agreements. The reason is, once again, that we intend to focus on the endpoint problems of the filter selected by the regulator rather than its general performance.

value (RW), and (2) a distributed lag equation where $\widehat{C}_t$ is predicted based on its own lags and the lags of the one-side filter ($ARDL$):

$$C_{c,t} = C_{c,t-h} \tag{1}$$

$$C_{c,t} = \alpha + \sum_{i=1}^{p} \beta_i C_{c,t-h-i} + \sum_{j=1}^{k} \gamma_i F_{c,t-j|t-j} + \gamma_c + \epsilon_{c,t} \tag{2}$$

These models are appealing for two reasons. First, they can be tested and used at no cost by any policy authority that calculates HP-filtered gaps as part of its risk assessment analysis. Second, they clearly provide a lower bound on what the procedure might be able to achieve. It is highly likely that by using additional macro-financial indicators, or factors that summarize large-dimension information sets, one could obtain significant accuracy improvements. In this respect, our results show that the procedure can work *even if* the modeling choices are quite clearly stacked against it.

We estimate RW and $ARDL$ using a panel specification with country fixed effects, γ_c. We let the horizon h vary from 4 to 20 quarters, and calculate for each model the average root mean square error (RMSE) across countries between 1981 and 2007.[6] The RMSEs of the models are reported in Table 4.[7] For both RW and $ARDL$ the errors are minimized at h^*=6 quarters, with $ARDL$ performing generally better than RW. We thus use equations (1) and (2) with $h = 6$ to calculate $\widehat{C}_t$ and $\widehat{F}_{t|t}$ in the remainder of the paper. Notice that we abstract from country specificities (except insofar as these are captured by the country fixed effects) and we use a unique horizon to generate the forecasts for all countries. Both choices are likely to further bias the results of the analysis

[6]The errors of interest are the discrepancies between the real-time estimate of the correction delivered by a given model and the corresponding full-sample HP estimate, i.e. $\widehat{C}_t - C_{t-h|T}$. Both terms are country-specific, and the first one is also model- and horizon-specific. The RMSEs are calculated (for each model and horizon) by averaging over countries and time periods.

[7]Results of the estimation of equation (2) are shown in Table A.2.

against the proposed procedure.[8] Importantly, we construct estimates of $F_{t|t}$, $S_{t|T}$, and $\widehat{C}_t$ using all available information (1971-2018) but we evaluate the models over a restricted 1981-2007 sample. This is to get rid of the two aforementioned potential distortions, namely the volatility of the estimates at the beginning of the sample (where the filter $F_{t|t}$ and the smoother $S_{t|T}$ have not converged yet) and the bias at the end of the sample (where $S_{t|T}$ and $F_{t|t}$ estimates converge by construction). A key issue is, of course, the choice of sensible and economically relevant evaluation criteria for the final outcome of the forecasting procedure. In what sense should $\widehat{F}_{t|t}$ work better than $F_{t|t}$? In Section 5 we explore sequentially three types of validation. We start by checking whether the adjustment delivers a better approximation of the full-sample estimate of the credit cycle (5.1). If (by some sensible metric) $||\widehat{F}_{t|t} - S_{t|T}|| < ||F_{t|t} - S_{t|T}||$, we can conclude that the correction is successful in bringing the real-time estimates of the credit cycle closer to those that can be computed with hindsight $T - t$ periods later. We then take a broader economic perspective, and test whether $\widehat{F}_{t|t}$ is a better predictor of financial crises than $F_{t|t}$ (5.2). Finally, we look at the policy implications of the adjustment, comparing the properties of capital buffers calibrated using alternatively $\widehat{F}_{t|t}$ or $F_{t|t}$ (5.3).

5 Does the fix work?

This section discusses our key empirical findings. The results are organized around three questions: (i) does the strategy described in Section 4 ameliorate the end-point problems of the HP filter? (ii) Does it allow us to predict financial crises more accurately? (iii) Does it yield economically sensible prescriptions on capital buffers? To simplify the notation, from now on we label the filter (one-side estimator), the adjusted filter (one-side estimator plus real-time correction) and the smoother (two-side estimator) respectively as F_t, F_t^* and S_t.[9] The first question requires a statistical comparison of F_t and F_t^* against S_t. The following two can

[8] For instance, a country-specific analysis would result for Italy in choosing $h^* = 8$ to minimize the errors.

[9] We refer the reader to Section 4 for formal definitions of these terms.

be answered (i) by examining the pattern of F_t and F_t^* around the crises in our sample, and (ii) by using them to calculate capital buffers through a mechanical macroprudential policy rule.

5.1 Approximating the full-sample estimates in real time

This section checks to what extent the now-casting procedure sketched in the previous section can be used to improve the performance of the HP filter in estimating credit cycles. To this end, we use four evaluation criteria based on correlations, volatilities, similarity, and synchronicity of the raw and adjusted filters (F_t, F_t^*) $vis - \grave{a} - vis$ the smoother (S_t). Our first evaluation criterion is correlation: a successful procedure generates $corr(F_t^*, S_t) > corr(F_t, S_t)$, with $corr(F_t^*, S_t) \to 1$ for a perfect correction. In practice, however, a better approximation of the sign and size of the cyclical component is also important. In fact, countercyclical capital requirements should be (i) activated during credit booms, and (ii) calibrated based on the magnitude of the credit imbalances. We measure these properties in terms of synchronicity and similarity: $Sync(F_t, S_t) = (F_t S_t)/|F_t S_t|$, $Sym = -|F_t - S_t|/|F_t + S_t|$. Synchronicity ranges between 1 and -1, while similarity between 0 and -N. In both cases, F_t^* should deliver higher values than F_t in order to be useful. Our experience at the Bank of Italy, and informal exchanges with macroprudential experts at other institutions, suggest that policymakers also care about volatility; as in other policy areas, volatile indicators can translate into volatile policy decisions and weaken credibility. We thus also compute volatility ratios. Given that the filter is by construction more volatile than the smoother, a successful adjustment procedure should deliver $\sigma_{F^*}/\sigma_S < \sigma_F/\sigma_S$, with a limiting value of 1 for an optimal adjustment.

Table 5 shows the results of the evaluation exercise. For each criterion, we report the results for the unadjusted filter (F_t) and the adjusted series obtained using the ARDL and the random walk model (F_t^{ARDL}, F_t^{RW}).[10]

[10] In Table A.4 we show that the results are broadly similar if the evaluation is carried out using the restricted sample 1981-2007, which excludes the initial observations (for which the estimates might be unstable) and the Global Financial Crisis.

The average behavior of the three filters in the panel is broadly similar, but there are clear signs that the adjustment improves the performance of F_t along some of the dimensions of interest. On average, F_t^{ARDL} works better than F_t according to the three evaluation criteria: it is more correlated with the smoother (column 3), more similar (column 6) and better synchronized (column 9). By contrast, F_t^{RW} is more similar but less correlated to the smoothed estimate than F_t. The discrepancies between filters are on average fairly small. A closer look at the country-level results reveals that these patterns are common to most economies. In terms of correlation, for instance, F_t^{ARDL} outperforms the unadjusted filter in all but four countries; based on synchronicity and similarity it performs better in all but six and eight countries respectively.

The volatility results in columns 10-12 raise an interesting issue. The adjusted series are clearly successful in reducing the well-known (and heavily criticized) volatility of the HP filter. The excess volatility of F_t relative to S_t is apparent for all countries in our sample, reaching peaks of 18–20% in the case of Italy, Japan and the UK. F_t^{*ARDL} and F_t^{*RW} deliver without exceptions volatility ratios that are lower than one. In this respect, both estimators seem capable of reducing the noise caused by the end-of-sample behavior of the HP filter. However, the figures may raise an opposite concern – namely that their volatility is too low, and that useful information is being discarded along with the filtering noise. This result is not surprising, as well-calibrated forecasts are typically less volatile than their targets. Its implications hinge critically on the objective of the forecasting exercise and the loss function of the forecaster. In the present context, the risk is that the 'false positives' generated by F_t are simply replaced by the 'false negatives' generated by F_t^*. In the next two subsections we test and rule out this possibility, showing that the adjusted filter predicts financial crises better than the unadjusted one.

5.2 Predicting the outbreak of financial crises

This section checks whether, besides providing a better picture of the credit cycle as demonstrated in Section 5.1, the adjusted filter F_t^* is also a better predictor of financial crises. The question is whether the additional information provided by the correction term C_t=F_t^*-F_t is valuable from a forecasting perspective. We compare the in-sample performance of raw and adjusted filters using a quarterly panel that includes 22 countries and 37 crises episodes that took place between 1971 and 2018.[11] For the sake of parsimony, in the following, we will focus only on the correction estimated using the ARDL model of equation 2.

We estimate logit models of the following type:

$$Pr(Crisis_{c,t} = 1) = \mathcal{F}(\alpha + \beta' X_{c,t-4} + \gamma_c), \tag{3}$$

where the dependent variable $Crisis_{c,t}$ is a country-specific crisis dummy and $\mathcal{F}$ denotes the cumulative distribution function for the logistic distribution. The predictors $X_{c,t}$ include alternatively $F_{c,t}$, the pair $[F_{c,t}\ \widehat{C_{c,t}}]$ or the smoother $S_{c,t}$.Although the smoothed estimate of the credit cycle is not available to forecasters in real-time, its behavior in the logit model is informative from our perspective. If $S_{c,t}$ is a good proxy of systemic risk, and if the estimated correction $\widehat{C_{c,t}}$ works as intended, then by using the correction we should be able to narrow the gap between the performance of the specification based on $F_{c,t}$ and that based on $S_{c,t}$. In all models, the predictors are lagged four quarters and the equation includes country fixed effects (γ_c). The estimated regressions are displayed in Table 6. The coefficients are significant and positively-signed in all specifications: this confirms that a rise in the credit gap – however measured – generally anticipates a financial crisis. The regression in column 2, where filter and correction appear as regressors, suggests that the predictive power of the correction is of the same order of magnitude as that of the filter itself. The benchmark specification reported in column 3 confirms that the correlation with future crises is even higher for the smoother, which outperforms the other filters for all the three information criteria. The one-side filter is the worst performer while the adjusted falls in between filter and smoother, as one would expect a priori.[12]

[11] Canada, New Zealand, Singapore, and South Africa are excluded from the panel because of the lack of information about financial crises. A full description of the crisis episodes is provided in Table A.1 of the Annex.

[12] Following Drehmann and Juselius (2014) we also compare the area under the ROC curve for the three filters, Figure A.4: the smoother reaches the higher value, 0.86, the one-side filter the lower, 0.73, but the adjustment ameliorates the performance, 0.77.

In Table 7 we evaluate the crisis predictions obtained with specification 3 (including alternatively $F_{c,t}$, $[F_{c,t}\ \widehat{C_{c,t}}]$, and $S_{c,t}$) in terms of sensitivity and specificity. These represent the probability of an observation being classified respectively as a crisis (not crisis) conditional on it being (not being) part of a crisis episode. In other words, they are the ratios of true positives and true negatives generated by each of the models. Notice that the model based on the ex-post smoother $S_{c,t}$ dominates the alternatives across criteria and thresholds. However, this is not available to policymakers in real-time.

In columns 1 and 2 we calculate the ratios assuming that an early warning is issued when the probability of a crisis is above 30 percent. In this case, using the adjusted filter increases the sensitivity by 29 percent compared to the unadjusted filter. The specificity of the two models is virtually identical, although using the adjusted filter still delivers a marginal (2 percent) gain. The remaining columns show that these results are robust to switching from a 30% to a 35% or 40% warning threshold. This is confirmed also in Figure 2 and Figure 3 which suggest that augmenting the one-side filter by the correction moves the performance of the filter closer to that of the smoother across the full distribution. The improvement is fully driven by the reduction in the probability of classifying a crisis as a normal period. The adjustment reduces the noise caused by the end-of-sample behavior of the one side filter, as shown in section 5.1, without entailing the loss of relevant information.[13]

5.3 Calibrating countercyclical capital buffers

The final aim of the credit-to-GDP gap is to be translated in a percentage of the bank risk-weighted assets (RWA) to set a benchmark buffer rate, according to the rule suggested by BCBS (2011).

The BCBS (2011) recommends that the accumulation period of the CCyB should be such that (i) the buffer rate is at the maximum of 2.5 percent of RWA

[13] The advantage gained by adjusting the filter becomes smaller when we try to predict the crisis two years ahead, see Annex A.2.

prior to a major crisis and (ii) banks are given one year to build up the CCyB. The CCyB should not reach the maximum too early or too late. With these guidelines in mind, the BCBS calibrated the CCyB based on the distribution of the filter. The rule stipulates that the CCyB should be activated when the filter gap exceeds 2 percentage points, so to avoid accumulating positive buffers in normal times, and peak when the gap reaches 10 percentage points (the level of the filter typically observed before the major systemic crises).[14] The credit-to-GDP gap estimates based on the adjusted filter has different sample statistics. To set up a fair comparison with the filter, we identify for our indicator a lower bound that is consistent with the 2 percentage points stated for the filter, a pre-crisis maximum and a slope using the same criteria employed by BCBS. For our sample of countries, 5 years prior to a crisis the 2 percentage points activation gap identified for the filter corresponds to the first quartile of the distribution, namely nearly three-fourths of the countries would have started to accumulate the buffer. The corresponding activation threshold for the adjusted filter has to be set to zero (Table 8) while the level of the adjusted filter that is typically observed one year before a crisis is 5 percent. The resulting piece-wise linear rules for the calculation of the benchmark for the CCyB is the following:[15]

1. Adjusted filter
    - $CCyB_t = 0$ if $F_t^{ARDL} \leq 0\%$
    - $CCyB_t = 0.5 * F_t^{ARDL}$ if $0\% < F_t^{ARDL} < 5\%$
    - $CCyB_t = 2.5\%$ if $F_t^{ARDL} > 5\%$

Table 9 shows that the buffer rate implied by the adjusted filter is on average higher than that based on the one-side filter and the BCBS proposed calibration. As suggested by the last columns of Table 9 the adjusted buffer is at the maximum a year before the crisis more often than the one-side buffer. To further explore this issue, we conduct a horse race to explicitly test which estimate of the credit-to-GDP

[14] The CCyB rule is described in detail in Annex A.1. We refer the reader to BCBS (2010) for details on the identification and the sample used by the Basel Committee on Banking Supervision.

[15] The accumulation rule defined by the BCBS for the one-side gap is reported in Annex A.1

gap does a better job in maximizing the CCyB rate one year before a crisis starts. This is done by regressing the probability of a crisis $Pr(Crisis_{c,t} = 1)$ on a dummy variable taking value 1 when the CCyB is at the maximum level, i.e. 2.5 %.

$$Pr(Crisis_{c,t} = 1) = \mathcal{F}(\alpha + \beta \widehat{CCyB}_{c,t-4} + \gamma_c) \tag{4}$$

$$\widehat{CCyB}_{c,t} \equiv I(CCyB_{c,t} \geq 2.5) \tag{5}$$

where the buffer $CCyB_{c,t}$ is defined alternatively based on F_t and F_t^{ARDL}. The regression results presented in Table 10 show that the CCyB calibrated on the basis of the adjusted filter has a higher probability to be at its maximum at the onset of a crisis.[16]

6 Conclusions

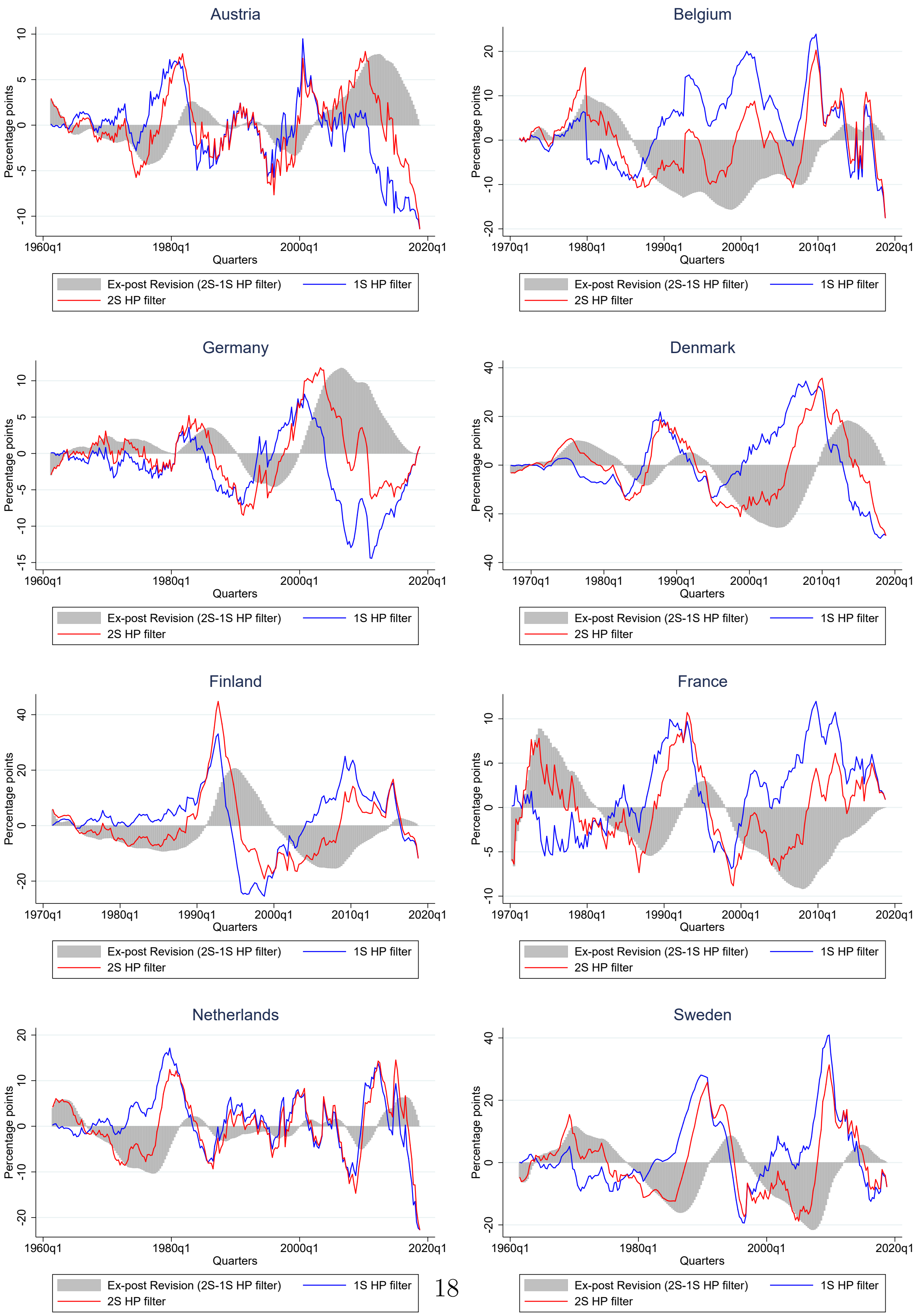
Financial crises are often anticipated by unsustainable credit booms and followed by dramatic credit contractions. To mitigate the volatility of the financial sector, regulators have introduced in the Basel III reform package a countercyclical capital requirement that should be tightened in good times and relaxed in bad times, so to stabilize both bank balance sheets and the supply of credit to the real economy. Measuring credit cycles has thus become a critical task for macroprudential authorities around the world. The task, however, is as problematic as it is important. The Basel regulation explicitly prescribes the use of the one side HP filter to extract the cyclical component on national credit aggregates, but the filter has been found to suffer from large ex-post revisions, leaving the authorities doubtful on its fitness for policy-making.

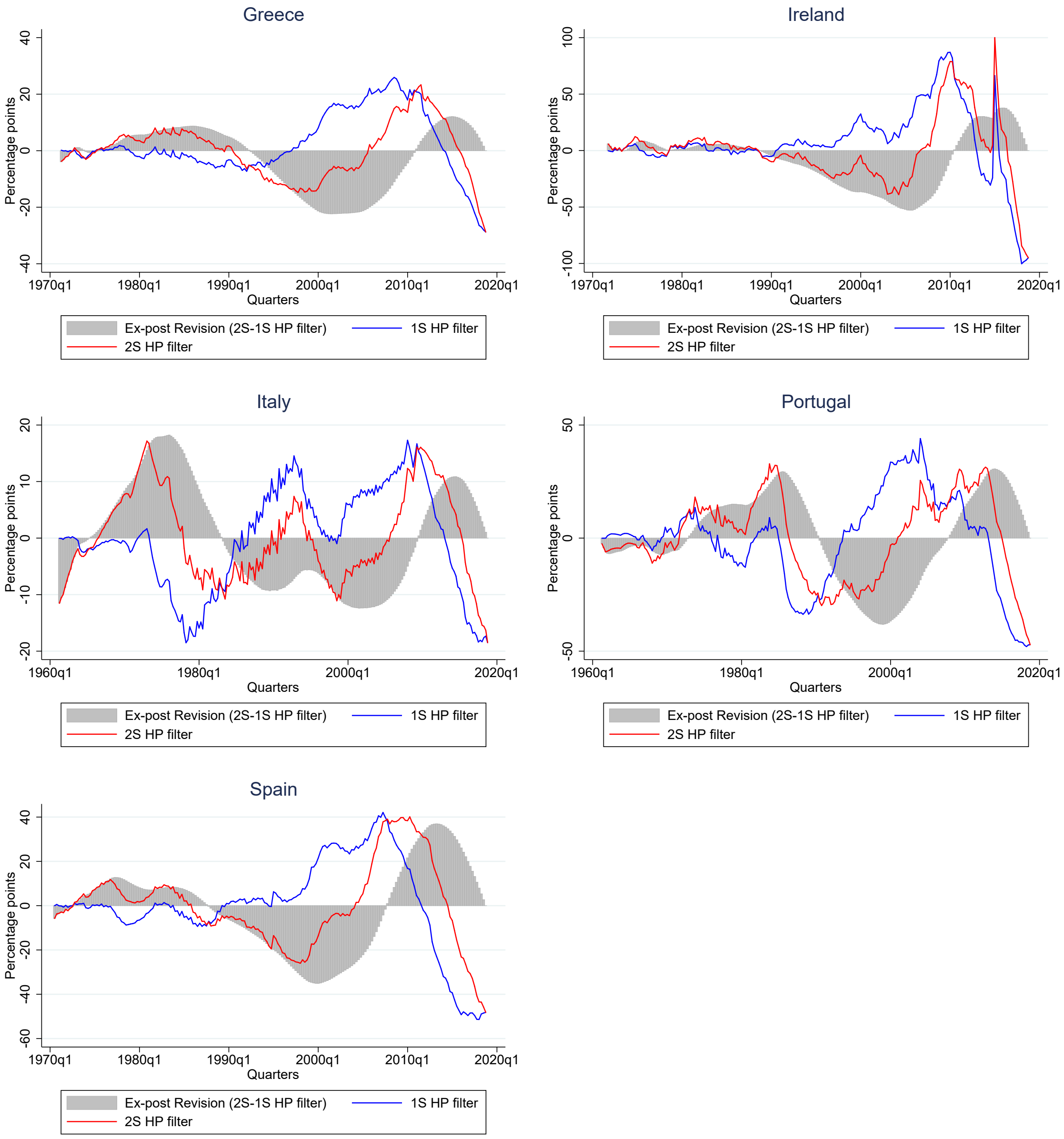
This paper provides both a new assessment of the problem and an intuitive, readily-implementable solution to it. Our study of a panel of 26 advanced and emerging countries over nearly 40 years reveals that the concerns about the one side HP filter are well-grounded. The limitations of the filter occasionally identified in previous studies are geographically pervasive, and the filtering errors can be as

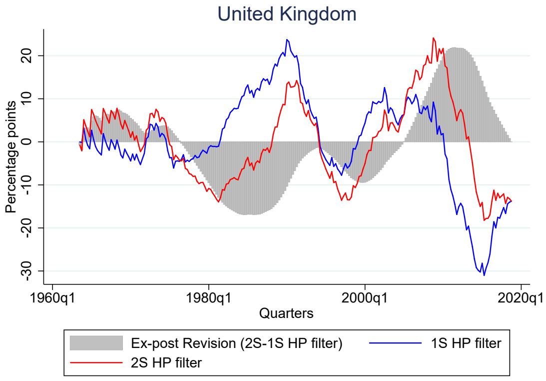
[16]The results are confirmed using the ROC curve, Figure A.5 in the annex.

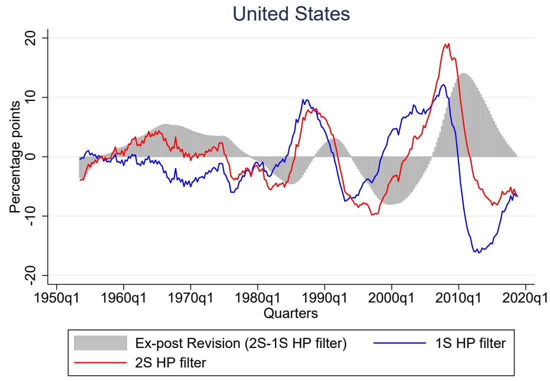
large as the credit cycle itself, rendering the real-time estimate virtually useless. However, the problem can easily be mitigated. Besides being large these errors are highly persistent and hence predictable. We show that this predictability can be exploited to construct an alternative real-time estimate of the 'credit gap' that has a number of appealing features. Relative to the HP filter, this estimate (i) is more correlated with the smoothed (full-sample) estimate of the credit cycle (i.e. the one based on the two side HP filter); (ii) is a better predictor of financial crises; and (iii) generates capital buffers that are higher at the onset of a crisis. The CCyB based on the adjusted filter also provides authorities more macroprudential space as the releasable buffer is on average higher compared to the one based on the filter. In a context characterized by a constrained monetary policy, macroprudential tools become particularly important. The estimation is technically trivial and requires no additional data beyond those used for the standard HP filter. Furthermore, the method can be easily tailored to the needs of different authorities. Our general conclusion is that, with minor variations relative to the original Basel III recipe, HP-filtered credit cycles remain a useful risk assessment tool for macroprudential authorities.

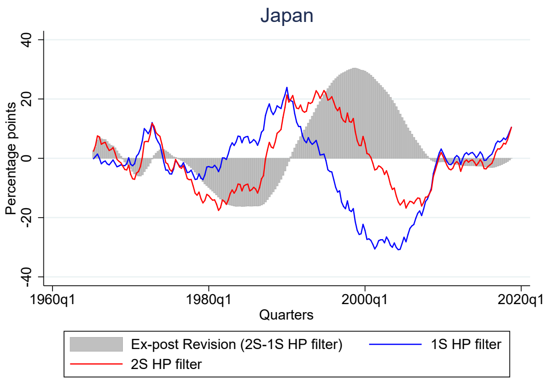
Figure 1: Credit-to-GDP gap: comparing two-side and one-side Hp filtering

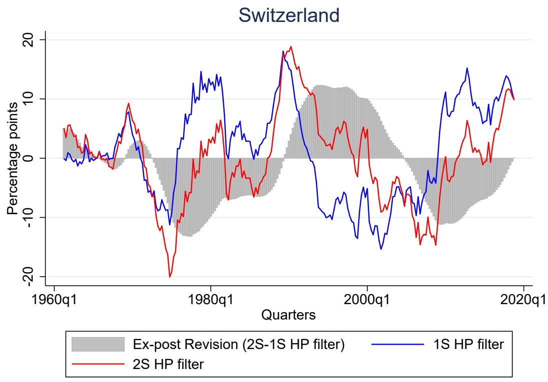


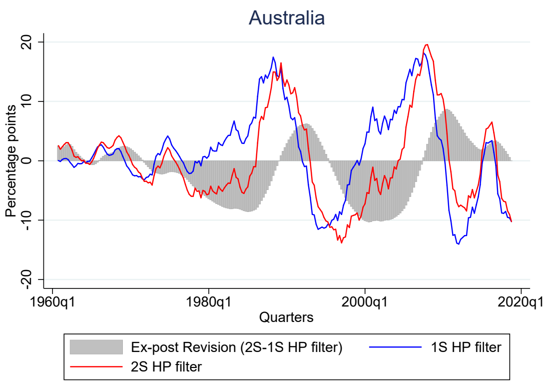

Greece
Percentage points
Quarters
1970q1
1980q1
1990q1
2000q1
2010q1
2020q1
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter
Ireland
Percentage points
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter
Italy
Percentage points
Quarters
1960q1
1980q1
2000q1
2020q1
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter
Portugal
Percentage points
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter
Spain
Percentage points
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter

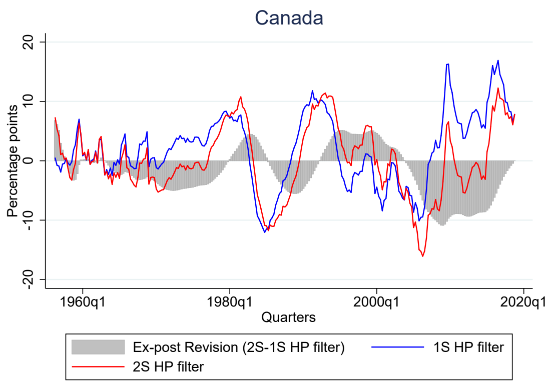

United Kingdom
Percentage points
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter

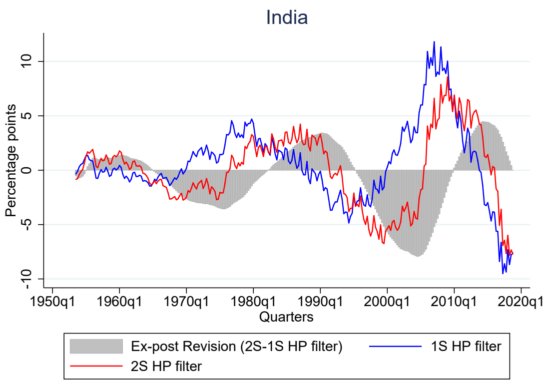

United States
Percentage points
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter

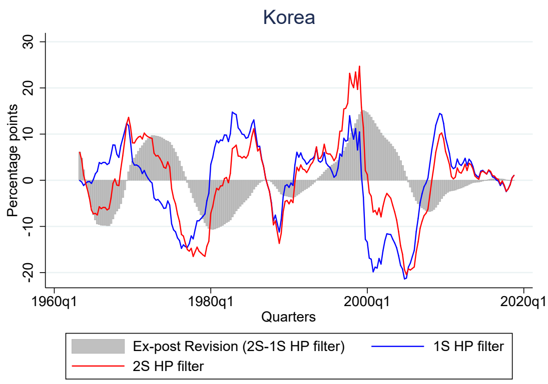

Japan
Percentage points
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter


Switzerland
Percentage points
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter


Australia
Percentage points
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter


Canada
Percentage points
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter


India
Percentage points
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter


Korea
Percentage points
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter

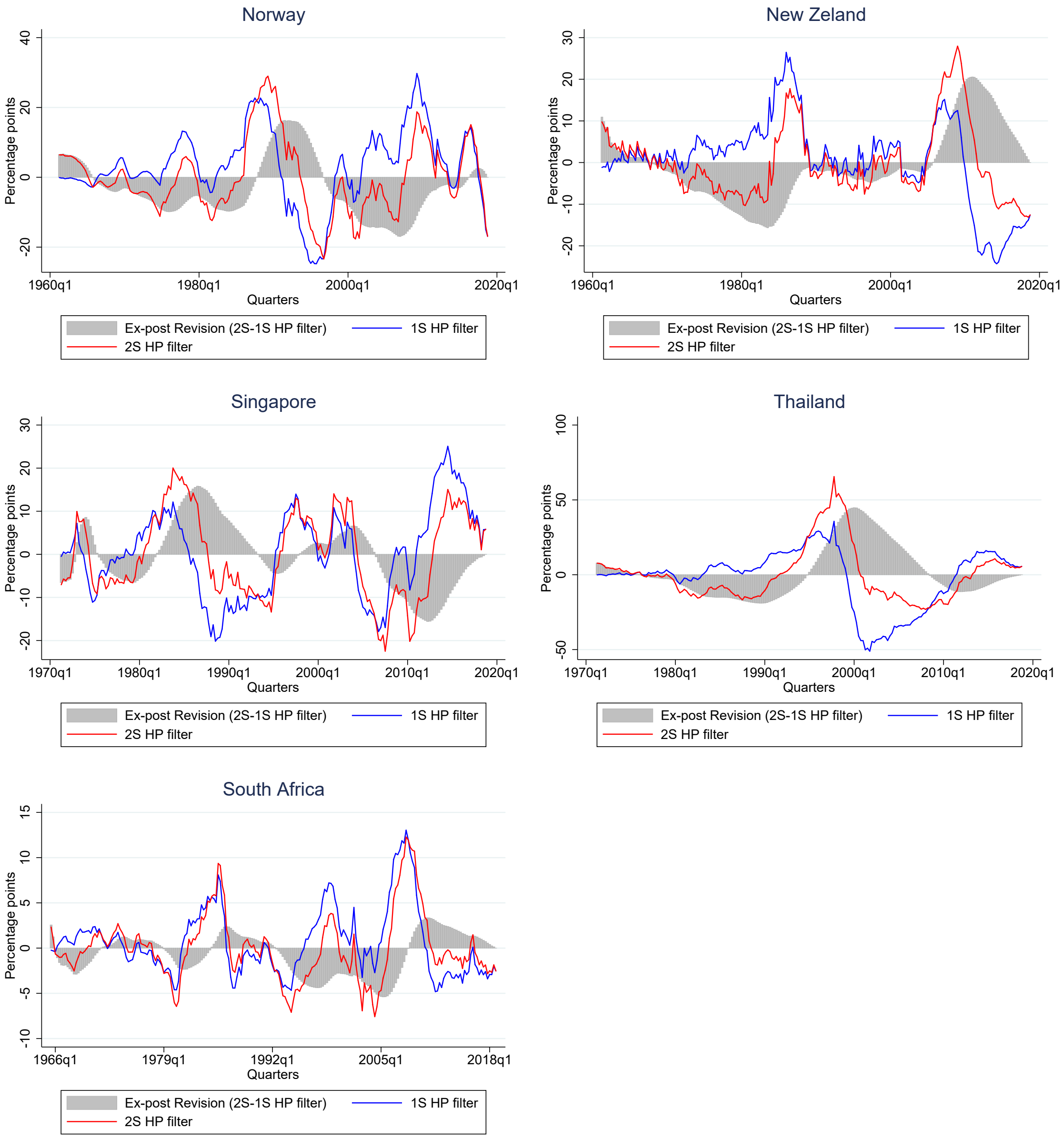
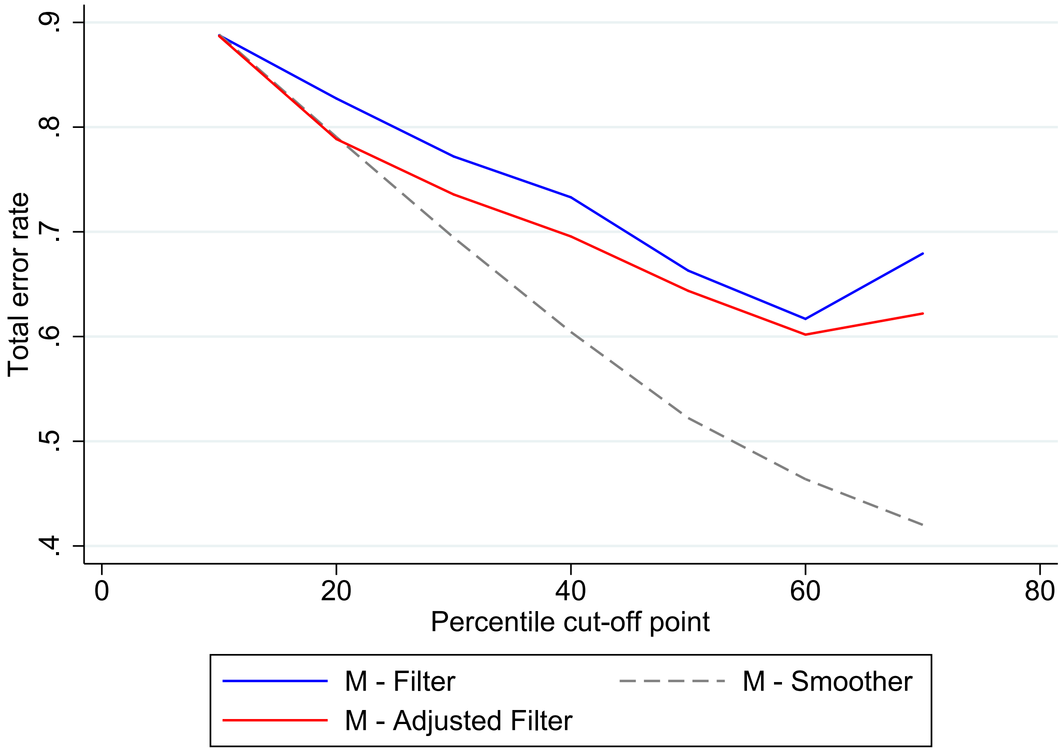
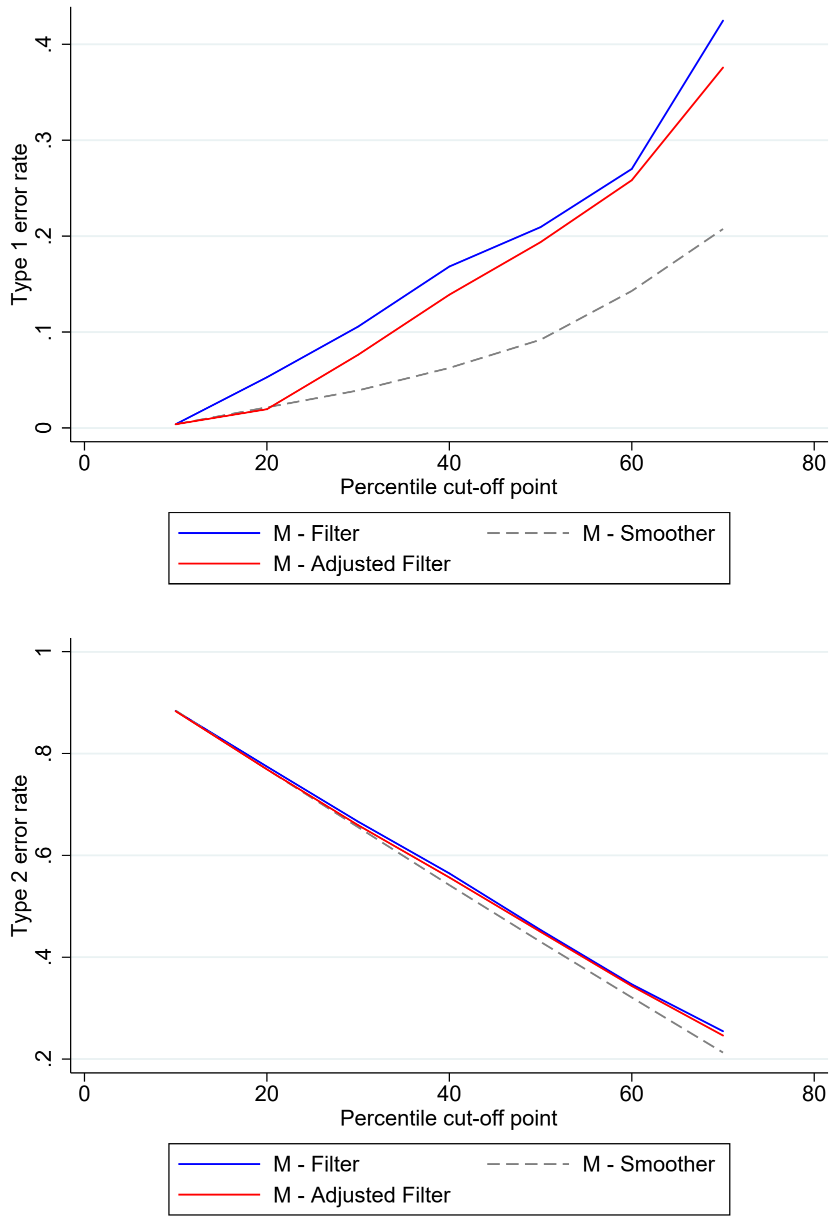

Norway
Percentage points
-20
0
20
40
1960q1
1980q1
2000q1
2020q1
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter
New Zeland
Percentage points
-20
-10
0
10
20
30
1960q1
1980q1
2000q1
2020q1
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter
Singapore
Percentage points
-20
-10
0
10
20
30
1970q1
1980q1
1990q1
2000q1
2010q1
2020q1
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter
Thailand
Percentage points
-50
0
50
100
1970q1
1980q1
1990q1
2000q1
2010q1
2020q1
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter
South Africa
Percentage points
-10
-5
0
5
10
15
1966q1
1979q1
1992q1
2005q1
2018q1
Quarters
Ex-post Revision (2S-1S HP filter)
1S HP filter
2S HP filter

Figure 2: Predicted crises classification rate (1 year before crisis)



Notes: This figure plots the incidence of Type 1 errors (classifying a crisis as normal period) and Type 2 errors (classifying a normal period as crisis) according to model-predicted probability. For a given percentile, we report the frequency of Type 1 and Type 2 errors. For example, at the 50th percentile, we consider the incidence of Type 1 and Type 2 errors if all observations with a predicted probability above the 50th percentile are classified as crisis and all others as normal periods. The horizontal axis presents the percentile cut-off points while the vertical axis the sum of Type 1 and Type 2 errors. Type 1 and Type 2 errors are reported separately in Figure 3.

Figure 3: Predicted crises classification rate (1 year before crisis)



Notes: The upper panel plots the incidence of Type 1 errors (classifying a crisis as normal period). The bottom panel reports Type 2 errors (classifying a normal period as crisis). For a given percentile cut-off, the frequency of Type 1 or Type 2 errors is reported.

Table 1: Credit-to-GDP ratio

Country	Obs.	Mean	Median	Std	Min	Max
Austria	191	106.32	105	30.62	50.2	147.9
Australia	191	126.79	119.6	44.9	61.2	204
Belgium	191	129.76	108.8	50.4	77.2	232.6
Canada	191	143.65	146.7	32.91	89.9	217.8
Switzerland	191	180.96	190.7	33.48	113.2	246.9
Germany	191	107.74	107.5	13.02	79.9	132.1
Denmark	191	164.62	146.3	47.78	110.5	254.6
Spain	191	116.82	82.3	52.06	67.8	218.1
Finland	191	124.87	118.4	33.95	79.8	194.6
France	191	136.86	130.6	29.29	99.7	201
United Kingdom	191	120.06	117	47.84	55.1	195.3
Greece	191	65.73	46.2	34.7	34.3	133.5
Ireland	189	145.19	87.3	93.1	66.2	398.6
India	191	35.1	29	14.89	14	62
Italy	191	83.19	75.6	24.16	51.7	127.3
Japan	191	170.11	163.4	25.68	125	218.2
Korea	191	127.01	138.6	45.72	51.3	199.4
Netherlands	191	182.08	179.1	67.66	64.4	293.9
Norway	191	168.47	163.1	40.68	110.9	255.8
New Zeland	191	116.7	114.3	54.14	28.6	201.4
Portugal	191	138.88	119.1	46.73	76.8	231.5
Sweden	191	154.59	146.5	47.1	100.7	246.4
Singapore	191	115.82	116.2	25.92	66.1	171
Thailand	191	88.95	94.5	38.48	27.4	181.9
United States	191	126.61	123	23.5	91.6	170.1
South Africa	191	59.84	56.6	8.22	47.1	79

Notes: Averages for the full sample 1971q1-2018q4.

Table 2: Comparing smoothed and filtered estimates - Full sample (1971q1-2018q4)

Country	CtGDP gap 1S	CtGDP gap 2S	Revision	SD CtGDP gap 1S	SD CtGDP gap 2S	SD 1S/SD 2S	Corr(1S; Revision)
Austria	-0.83	-0.03	0.8	4.23	3.92	1.08	-0.48
Australia	1.71	-0.36	-2.07	8.21	8.04	1.02	-0.37
Belgium	3.46	0	-3.46	8.13	7.06	1.15	-0.59
Canada	2.34	0.14	-2.2	6.77	6.64	1.02	-0.38
Switzerland	1.35	-0.62	-1.97	8.83	8.03	1.1	-0.55
Germany	-2.47	0.02	2.49	4.99	4.71	1.06	-0.5
Denmark	1.74	0.14	-1.6	14.76	13.9	1.06	-0.48
Spain	0.73	0.13	-0.6	20.95	18.67	1.12	-0.58
Finland	2.35	-0.12	-2.47	11.59	11.5	1.01	-0.39
France	1.83	0.13	-1.7	4.68	4.25	1.1	-0.57
United Kingdom	0.58	-0.62	-1.2	11.95	10.1	1.18	-0.62
Greece	2.43	0.06	-2.37	11.19	9.68	1.16	-0.62
Ireland	7.47	-0.07	-7.54	30.69	28.64	1.07	-0.47
India	1.04	0.1	-0.95	4.13	3.84	1.07	-0.52
Italy	0.15	0.14	0	9.94	8.31	1.2	-0.65
Japan	-3.25	-0.07	3.18	13.39	11.2	1.2	-0.66
Korea	-0.82	-0.24	0.58	9.12	8.97	1.02	-0.37
Netherlands	1.4	-0.23	-1.64	7.02	6.93	1.01	-0.31
Norway	3.76	-0.29	-4.06	11.59	11.4	1.02	-0.4
New Zeland	0.72	-0.59	-1.32	10.84	9.26	1.17	-0.59
Portugal	-0.61	1	1.61	22.15	19.72	1.12	-0.59
Sweden	3.56	-0.59	-4.15	12.51	11.83	1.06	-0.43
Singapore	0.32	0.05	-0.27	10.04	9.81	1.02	-0.41
Thailand	-1.71	-0.09	1.62	19.16	17.78	1.08	-0.53
United States	-1.13	-0.26	0.87	7.17	6.8	1.05	-0.46
South Africa	0.59	0.01	-0.59	3.92	3.88	1.01	-0.32
Average	1.03	-0.09	-1.11	11.08	10.19	1.08	-0.49

Notes: Averages for the full sample 1971q1-2018q4. CtGDP gap 1S refers to the gap estimated by one-side HP filter (filter), CtGDP gap 2S refers to the gap estimated by two-side HP filter (smoother), Revision=(CtGDP gap 2S-CtGDP gap 1S), and Corr(1S; Revision) is the correlation between the filtered estimate of the cycle and the ex-post revision.

Table 3: Comparing smoothed and filtered estimates - Short sample (1981q1-2007q4)

Country	CtGDP gap 1S	CtGDP gap 2S	Revision	SD CtGDP gap 1S	SD CtGDP gap 2S	SD 1S/SD 2S	Corr(1S; 2S-1S)
Austria	-0.13	0.07	0.2	3.05	3.37	0.91	-0.11
Australia	4.45	0.17	-4.29	8.47	8.92	0.95	-0.23
Belgium	4.48	-3.07	-7.55	8.04	5.43	1.48	-0.75
Canada	-1.25	-0.93	0.32	6.38	7.53	0.85	-0.04
Switzerland	-1.83	0.74	2.56	8.43	8.15	1.03	-0.48
Germany	-0.69	1.28	1.96	4.82	5.6	0.86	-0.36
Denmark	5.12	-3.89	-9.01	12.78	12.28	1.04	-0.46
Spain	9.36	-3.9	-13.26	14.52	14.05	1.03	-0.54
Finland	0.1	-1.64	-1.74	13.5	14.3	0.94	-0.33
France	1.63	-1.35	-2.98	4.22	4.72	0.89	-0.24
United Kingdom	7.47	0.03	-7.44	7.57	8.86	0.86	-0.26
Greece	3.5	-2.98	-6.48	9.43	7.4	1.28	-0.8
Ireland	10.7	-9.27	-19.97	13.72	13.24	1.04	-0.77
India	0.86	-0.63	-1.49	3.7	3.58	1.03	-0.58
Italy	4.46	-3.15	-7.61	6.71	4.79	1.4	-0.71
Japan	-5.91	1.68	7.6	16.84	13.55	1.24	-0.68
Korea	-0.91	0.34	1.25	10.43	9.94	1.05	-0.39
Netherlands	0.37	-0.16	-0.53	4.58	4.68	0.98	-0.12
Norway	1.65	-1.38	-3.03	12.91	13.44	0.96	-0.37
New Zeland	5.24	0.92	-4.33	7.9	7.94	1	-0.33
Portugal	4.74	-3.23	-7.97	22.5	19.42	1.16	-0.63
Sweden	4.58	-3.43	-8.01	11.11	12.41	0.89	-0.23
Singapore	-2.74	1.28	4.02	9.98	10.57	0.94	-0.19
Thailand	-4.22	1.05	5.27	24.07	22.32	1.08	-0.54
United States	2.38	-0.16	-2.53	5.6	6.73	0.83	0.02
South Africa	1.6	-0.26	-1.86	3.91	4.02	0.97	-0.24
Average	2.12	-1.23	-3.34	9.82	9.51	1.03	-0.4

Notes: Averages for a short sample 1981q1-2007q4. CtGDP gap 1S refers to the gap estimated by one-side HP filter (filter), CtGDP gap 2S refers to the gap estimated by two-side HP filter (smoother), Revision=(CtGDP gap 2S-CtGDP gap 1S), and Corr(1S; Revision) is the correlation between the filtered estimate of the cycle and the ex-post revision.

Table 4: Root Mean Square Error

horizon h								
4	6	8	10	12	14	16	18	20
				RW Model				
11.35	11.14	11.14	11.31	11.58	11.93	12.32	12.73	13.13
				ARDL Model				
8.2	8.17	8.29	8.35	8.4	8.44			

Notes: Root mean square errors for the estimation of the ex-post revision. The target revision is calculated on ex-post estimation of the smoother based on the sample 1971-20018 while, for the model assessment the reference sample is 1981q1-2007q4.

Table 5: Reliability Criteria - Full sample

Country	Correlation			Similarity			Syncronicity			Volatility		
	F_t	F_t^{*RW}	F_t^{*ARDL}	F_t	F_t^{*RW}	F_t^{*ARDL}	F_t	F_t^{*RW}	F_t^{*ARDL}	F_t	F_t^{*RW}	F_t^{*ARDL}
Australia	0.76	0.7	0.84	-3.38	-2.31	-2.36	0.3	0.24	0.39	1.02	0.74	0.65
Austria	0.68	0.65	0.69	-1.64	-4.71	-2.23	0.5	0.44	0.37	1.07	0.78	0.66
Belgium	0.53	0.5	0.38	-6.61	-6.54	-2.28	0.24	0.16	0.09	1.15	0.85	0.73
Canada	0.74	0.69	0.78	-2.79	-2.38	-2.54	0.31	0.3	0.43	1.02	0.75	0.64
Denmark	0.64	0.59	0.75	-10.03	-1.71	-5.88	0.44	0.37	0.49	1.06	0.75	0.68
Finland	0.71	0.65	0.77	-13.31	-5.2	-2.12	0.2	0.14	0.4	1.01	0.73	0.66
France	0.51	0.51	0.59	-6.09	-3.98	-1	0.22	0.23	0.56	1.11	0.74	0.82
Germany	0.59	0.54	0.68	-1.3	-6.67	-1.52	0.51	0.49	0.49	1.06	0.76	0.69
Greece	0.45	0.41	0.53	-14.69	-2.89	-41.28	0.09	-0.03	-0.01	1.16	0.82	0.75
India	0.59	0.55	0.71	-2.25	-8.38	-10.36	0.33	0.3	0.38	1.08	0.77	0.68
Ireland	0.67	0.64	0.68	-7.84	-3.27	-5.07	-0.02	-0.13	-0.07	1.07	0.8	0.62
Italy	0.43	0.4	0.55	-6.56	-5.99	-2.68	0.08	0.08	0.16	1.24	0.86	0.81
Japan	0.36	0.3	0.46	-2.43	-5.02	-3.59	0.3	0.21	0.43	1.19	0.82	0.82
Korea	0.76	0.7	0.82	-1.65	-1.35	-0.77	0.74	0.62	0.74	1.04	0.74	0.65
Netherlands	0.83	0.79	0.83	-4.54	-2.51	-3.87	0.54	0.46	0.6	1.03	0.78	0.56
New Zeland	0.57	0.54	0.61	-2.87	-4.12	-1.36	0.28	0.26	0.36	1.17	0.84	0.76
Norway	0.71	0.66	0.75	-1.84	-4.17	-2.13	0.23	0.22	0.38	1.02	0.75	0.63
Portugal	0.49	0.45	0.59	-2.17	-1.69	-1.49	0.3	0.28	0.32	1.12	0.79	0.72
Singapore	0.71	0.6	0.54	-0.72	-1.3	-1.84	0.59	0.36	0.7	1.03	0.74	0.88
South Africa	0.81	0.75	0.78	-4.18	-3.27	-1.37	0.49	0.51	0.46	1.01	0.75	0.61
Spain	0.5	0.47	0.62	-9.96	-3.23	-13.16	-0.11	-0.14	0.02	1.12	0.79	0.72
Sweden	0.74	0.69	0.8	-22.79	-1.28	-3.41	0.36	0.34	0.56	1.05	0.75	0.68
Switzerland	0.54	0.49	0.6	-3.11	-2.74	-2.19	0.22	0.21	0.36	1.09	0.78	0.71
Thailand	0.56	0.51	0.67	-38.77	-1.89	-1.8	0.36	0.32	0.42	1.08	0.76	0.72
United Kingdom	0.51	0.46	0.6	-1.6	-4.12	-2.14	0.48	0.44	0.47	1.18	0.83	0.79
United States	0.68	0.62	0.79	-1.45	-1.89	-1.87	0.56	0.5	0.72	1.05	0.74	0.69
Average	0.62	0.57	0.67	-6.69	-3.56	-4.61	0.33	0.28	0.39	1.09	0.78	0.7

Notes: Averages for the full sample 1971q1-2018q4. The first three columns report the correlation with the smoother. Synchronicity ranges between 1 and -1 and refers to $Sync(F_t, S_t) = (F_t S_t)/|F_t S_t|$. Similarity ranges between 0 and -N and refers to $Sym = -|F_t - S_t|/|F_t + S_t|$.

Table 6: Predicting Crisis 1 year before

VARIABLES	$Pr(Crisis_{c,t} = 1)$		
$F_{c,t-4}$	0.0380***	0.114***	
	(0.00395)	(0.00814)	
$\widehat{C}_{c,t-4}$		0.174***	
		(0.0151)	
$S_{c,t-4}$			0.125***
			(0.00628)
Constant	-1.430***	-1.270***	-1.547***
	(0.191)	(0.192)	(0.194)
Observations	3899	3899	3899
AIC	0.728	0.692	0.579
BIC	-29258	-29391	-29837
Deviance	2791	2649	2212

Notes: The model is estimated using a generalized linear model for binomial outcome. All the regressions include country fixed effects. The last rows report three different measures of fit: (i) *AIC* refer to the Akaike information criterion, (ii) *BIC* to the Bayesian information criterion and, (iii) *Deviance* measures the distance of the fitted model with respect to an abstract model that fits perfectly the sample assigning probability 0 or 1 based on the actual value, larger is the deviance the lower the fit. Standard errors are reported in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table 7: Forecast evaluation (1 year before the crisis)

Filter	$Pr(Dep = 1) > 30\%$		$Pr(Dep = 1) > 35\%$		$Pr(Dep = 1) > 40\%$	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
F_t	0.41	0.86	0.42	0.86	0.39	0.85
F_t $\hat{C}_t$	0.53	0.88	0.54	0.87	0.53	0.86
$S_{t\|T}$	0.61	0.92	0.65	0.91	0.7	0.9
Gain	0.12	0.02	0.12	0.01	0.14	0.01
Percent Gain	29	2	29	1	36	1

Notes: Calculation based on the sample from 1971q1 to 2018q4. Sensitivity measures the percentage of predicted crisis that were crisis. Specificity refers to the percentage of non-crisis periods correctly identified (computed as one minus the portion of periods predicted as non-crisis that were crisis).

Table 8: Calibrating countercyclical capital buffers

	Years before a banking crisis									
	5 years		4 years		3 years		2 years		1 year	
	mean	p 25	mean	p 25	mean	p 25	mean	p 25	mean	p 25
Adjusted Filter	3.3	0.2	3.7	0.4	4.2	0.3	4.7	0.9	5.4	0.9
Filter	6.1	1.8	7.6	2.6	8.3	2.3	10.2	0.7	10.1	3.9

Notes: The table reports the average development of the credit-to-GDP gap for the five years prior to the outbreak of 37 crises. The gaps are calculated using two different filters: the adjusted and the one-side filter.

Table 9: Countercyclical capital buffers - Comparing Filtering

Country	CCyB		Std(CCyB)		$I(CCyB = 2.5\|_{Crisis_{t+4}=1})$	
	F_t	F_t^{ARDL}	F_t	F_t^{ARDL}	F_t	F_t^{ARDL}
Australia	0.77	0.84	0.99	1.04	1	1
Austria	0.18	0.21	0.44	0.46	0	0
Belgium	0.94	1.14	1.01	1.11	0	0
Canada	0.86	0.96	0.96	1.01	0	0
Denmark	0.88	0.94	1.12	1.13	1	1
Finland	0.91	1.09	1.07	1.09	1	1
France	0.6	0.79	0.84	0.96	0	0
Germany	0.15	0.18	0.39	0.45	0	0
Greece	0.76	0.86	1.12	1.14	1	1
India	0.35	0.44	0.66	0.73	0	0
Ireland	1.11	1.3	1.11	1.13	1	1
Italy	0.88	0.97	1.06	1.12	1	2
Japan	0.56	0.67	0.9	1.01	0	0
Korea	0.66	0.7	0.93	0.93	0	0
Netherlands	0.64	0.67	0.89	0.9	0	0
New Zeland	0.74	0.79	0.95	0.97	0	0
Norway	1.11	1.18	1.09	1.11	1	2
Portugal	0.98	1.05	1.12	1.16	1	1
Singapore	0.77	0.86	1.01	1.09	0	0
South Africa	0.32	0.37	0.68	0.67	0	0
Spain	0.74	0.86	1.09	1.09	1	1
Sweden	0.87	0.9	1.11	1.12	2	2
Switzerland	0.97	1.01	1.09	1.12	1	1
Thailand	0.94	0.99	1.1	1.11	1	1
United Kingdom	0.95	1.02	1.07	1.11	1	2
United States	0.54	0.6	0.84	0.93	1	1
Average	0.74	0.82	0.95	0.99	0.58	0.69

Notes: Quarterly data from 1971q1-2018q4. $I(CCyB = 2.5|_{Crisis_{t+4}=1})$ is an indicator variable that equals one if the estimated gap suggests to set the CCyB to the upper bound a year before a crisis occurred.

Table 10: Benchmark buffer guide at the maximum 1 year before the crisis

VARIABLES	$Pr(Crisis_{c,t} = 1)$	
$\widehat{CCyB(F)}_{c,t}$	0.999***	
	(0.115)	
$\widehat{CCyB(F^{ARDL})}_{c,t}$		1.449***
		(0.111)
Constant	-1.443***	-1.443***
	(0.191)	(0.191)
Observations	3899	3899
AIC	0.735	0.710
BIC	-29229	-29327
Deviance	2820	2722

Notes: The model is estimated using a generalized linear model for binomial outcome. All the regressions include country fixed effects. The last rows report to three different measures of fit: (i) *AIC* refer to the Akaike information criterion, (ii) *BIC* to the Bayesian information criterion and, (iii) *Deviance* measures the distance of the fitted model with respect to an abstract model that fits perfectly the sample assigning probability 0 or 1 based on the actual value, larger is the deviance the lower the fit. Standard errors are reported in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

References

Adrian, Tobias, Nina Boyarchenko, and Domenico Giannone, "Vulnerable growth," *American Economic Review*, 2019, *109* (4), 1263–89.

Alessandri, Piergiorgio, Pierluigi Bologna, Roberta Fiori, and Enrico Sette, "A note on the implementation of the countercyclical capital buffer in Italy," *Bank of Italy Occasional Paper*, 2015, (278).

Alessi, Lucia and Carsten Detken, "Identifying excessive credit growth and leverage," *Journal of Financial Stability*, 2018, *35*, 215–225.

Baba, Chikako, Salvatore Dell'Erba, Enrica Detragiache, Olamide Harrison, Aiko Mineshima, Anvar Musayev, and Asghar Shahmoradi, "How Should Credit Gaps Be Measured? An Application to European Countries," *IMF Working Papers*, 2020, (20/6).

BCBS, "Guidance for national authorities operating the countercyclical capital buffer," *https://www.bis.org/publ/bcbs187.htm*, 2010.

_ , "Basel III: A global regulatory framework for more resilient banks and banking systems, Revised version," *https://www.bis.org/publ/bcbs189.htm*, 2011.

Darracq Pariès, Matthieu, Stephan Fahr, and Christoffer Kok, "Macroprudential space and current policy trade-offs in the euro area," *Published as part of the ECB Financial Stability Review, https://www.ecb.europa.eu/pub/fsr/special/html/ecb.fsrart201805_2.en.html*, 2019, (May).

Drehmann, Mathias and James Yetman, "Why you should use the Hodrick-Prescott filter–at least to generate credit gaps," *BIS Working Paper*, 2018.

_ **and Mikael Juselius**, "Evaluating early warning indicators of banking crises: Satisfying policy requirements," *International Journal of Forecasting*, 2014, *30* (3), 759–780.

— , **Claudio EV Borio, Leonardo Gambacorta, Gabriel Jimenez, and Carlos Trucharte**, "Countercyclical capital buffers: exploring options," *BIS working paper*, 2010.

Edge, Rochelle M and Ralf R Meisenzahl, "The unreliability of credit-to-GDP ratio gaps in real-time: Implications for countercyclical capital buffers," *International Journal of Central Banking*, 2011, *7* (4), 261–298.

European Parliament, "Directive 2013/36/EU of the European Parliament and the council," *https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32013L0036*, June 2013.

Geršl, Adam and Jakub Seidler, "Countercyclical capital buffers and Credit-to-GDP gaps: simulation for central, eastern, and southeastern Europe," *Eastern European Economics*, 2015, *53* (6), 439–465.

Greenwood, Robin, Samuel G. Hanson, Andrei Shleifer, and Jakob Ahm Sørensen, "Predictable Financial Crises," *NBER Working Paper*, 2020, *June* (27396).

Hamilton, James D, "Why you should never use the Hodrick-Prescott filter," *Review of Economics and Statistics*, 2018, *100* (5), 831–843.

Jordà, Òscar, Moritz Schularick, and Alan M Taylor, "Financial crises, credit booms, and external imbalances: 140 years of lessons," *IMF Economic Review*, 2011, *59* (2), 340–378.

— , — , **and** — , "Macrofinancial history and the new business cycle facts," *NBER macroeconomics annual*, 2017, *31* (1), 213–263.

Laeven, Luc and Fabian Valencia, "Systemic Banking Crises Database II," *IMF Economic Review*, 2020, *68*, 1–55.

Lo Duca, Marco, Anne Koban, Marisa Basten, Elias Bengtsson, Benjamin Klaus, Piotr Kusmierczyk, Jan Hannes Lang, Carsten Detken, and Tuomas A Peltonen, "A new database for financial crises in European countries: ECB/ESRB EU crises database," *ECB Occasional Paper Series*, 2017, (194).

Orphanides, Athanasios and Simon van Norden, "The unreliability of output-gap estimates in real time," *Review of economics and statistics*, 2002, *84* (4), 569–583.

Schularick, Moritz and Alan M Taylor, "Credit booms gone bust: Monetary policy, leverage cycles, and financial crises, 1870-2008," *American Economic Review*, 2012, *102* (2), 1029–61.

A Annex

A.1 Basel Committee recommendations on the calculation of the countercyclical capital buffer

According to BCBS (2011) the credit-to-GDP gap is defined as the difference between an economy's aggregate credit-to-GDP ratio and its long-run trend. The long-term trend of credit-to-GDP ratio is computed with a one-side (recursive) Hodrick-Prescott filter with a smoothing parameter $\lambda = 400,000$. Credit denotes a broad measure of the stock of domestic credit to the private non-financial sector outstanding at the end of quarter t. The credit-to-GDP gap is then translated in a percentage of the bank risk-weighted assets by calculating a benchmark buffer rate based on the piece-wise linear rule:

- $CCyB_t = 0$ if $GAP_t < 2\%$
- $CCyB_t = 0.3125 * GAP_t - 0.625$ if $2\% < GAP_t < 10\%$
- $CCyB_t = 2.5\%$ if $GAP_t > 10\%$

The Credit definition suggested by BCBS (2011) includes all private credit issued by banks and non-bank financial institutions.Authorities are however allowed to use (i) additional measures of the credit gap and/or (ii) alternative de-trending methods in order to better capture the specificities of their national economies.

A.2 Does the credit gap predict financial crises 2 years ahead?

We explore whether the dominance of the adjusted filter on the one-side filter is confirmed in a model with 2 years lags by estimating the logit models given by the following equations.

$$Pr(Crisis_{c,t} = 1) = \mathcal{F}(\alpha + \beta F_{c,t-8} + \gamma_c) \qquad (6)$$

$$Pr(Crisis_{c,t} = 1) = \mathcal{F}(\alpha + \beta F^{ARDL}_{c,t-8} + \gamma_c) \tag{7}$$

Where $\mathcal{F}$ denotes the cumulative distribution function for the logistic distribution and the dependent variable $Crisis_{c,t}$ is a crisis dummy. In the explanatory variables, Equation 6 includes $F_{c,t-8}$ which is the credit-to-GDP gap estimated through a one-sided HP filter, with lag 8, and Equation 7 includes $F^{ARDL}_{c,t-8}$ which is the adjusted credit-to-GDP gap with lag 8; both equations include country fixed effects, γ_c. Estimates are reported in Table A.3. The relative performance of the two models compared to the target predictions estimated using the smoother is depicted in Figures A.6 and A.7. The advantage gained by the estimated revision, in this case, is lower then the one obtained in the 1 year ahead prediction.

Table A.1: Number of Quarters

Country	No Crisis	Crisis	Total	Periods of Crisis
Australia	178	4	185	1989q1-1989q4
Austria	148	34	185	2008q1-2016q2
Belgium	159	20	182	2008q1-2012q4
Denmark	126	56	185	1987q2-1995q1, 2008q1-2013q4
Finland	158	21	182	1991q4-1996q4
France	160	22	185	1991q3-1995q1, 2008q2-2009q4
Germany	145	37	185	1974q3-1974q4, 2001q1-2003q4, 2007q4-2013q2
Greece	148	31	182	2010q3-2018q1
India	180	2	185	1993q2-1993q3
Ireland	156	21	180	2008q4-2013q4
Italy	148	34	185	1991q4-1997q4, 2011q4-2013q4
Japan	164	18	185	1997q2-2001q3
Korea	178	4	185	1997q4-1998q3
Netherlands	161	21	185	2008q1-2013q1
Norway	173	9	185	1988q2-1988q3, 1992q1-1993q3
Portugal	145	37	185	1983q2–1985q1, 2008q4-2015q4
Spain	132	50	185	1978q1-1985q3, 2009q2-2013q4
Sweden	147	35	185	1991q1-1997q2, 2008q4-2010q4
Switzerland	171	11	185	1991q1-1991q4, 2008q4-2009q3
Thailand	165	14	182	1983q2-1983q3, 1997q4- 2000q3
United Kingdom	152	30	185	1974q1-1975q4, 1991q3- 1994q2, 2007q4-2010q1
United States	159	23	185	1984q1-1984q4, 1988q1-1988q4, 2008q1-2011q3
Total	3453	534	4053	

Notes: Quarterly data from 1971q1 to 2018q4. Crisis periods are obtained from three sources: the ESRB-European Financial Crises database for European countries, Laeven and Valencia (2020) and Jordà et al. (2017) for non-European countries. All the sources do not report periods of crisis for Canada, New Zeland, Singapore and South Africa.

Table A.2: Model Selection - ARDL

	horizon (h)					
VARIABLES	4	6	8	10	12	14
$Correction_{t-h}$	23.17***	23.96***	19.67***	17.98***	17.51***	15.26***
	(2.459)	(2.454)	(2.478)	(2.564)	(2.583)	(2.013)
$Correction_{t-h-1}$	-26.11***	-25.58***	-19.38***	-18.60***	-18.46***	-16.67***
	(2.957)	(3.754)	(2.297)	(2.393)	(2.425)	(2.040)
$Correction_{t-h-2}$	37.58*	10.03***	5.362	3.622	1.681	0.934
	(21.58)	(3.213)	(3.723)	(2.874)	(1.806)	(1.221)
$Correction_{t-h-3}$	-30.07	-6.547***	-0.784	0.175	1.495	1.788
	(19.29)	(1.476)	(2.627)	(2.317)	(1.576)	(1.195)
$Filter_{t-2}$	2.003*	0.676**	0.300	-0.135	-0.273*	-0.559***
	(1.137)	(0.315)	(0.329)	(0.233)	(0.152)	(0.102)
$Filter_{t-3}$	0.376	0.252	0.441**	0.384***	0.285***	0.232***
	(0.381)	(0.254)	(0.188)	(0.0995)	(0.0703)	(0.0490)
$Filter_{t-4}$	-0.773*	0.359	0.380***	0.410***	0.230***	0.319***
	(0.392)	(0.214)	(0.133)	(0.0932)	(0.0636)	(0.0523)
$Filter_{t-5}$	0.435	0.144	0.375*	0.283**	0.277***	0.174***
	(0.411)	(0.255)	(0.183)	(0.121)	(0.0743)	(0.0579)
$Filter_{t-6}$	-1.434	-1.209***	0.160	0.317*	0.466***	0.422***
	(1.149)	(0.330)	(0.264)	(0.185)	(0.142)	(0.108)
Constant	-2.417***	-2.437***	-2.489***	-2.521***	-2.574***	-2.636***
	(0.137)	(0.141)	(0.155)	(0.159)	(0.161)	(0.160)
Observations	2626	2626	2626	2626	2626	2626
R-squared	0.390	0.393	0.377	0.367	0.360	0.353
Number of id_n	26	26	26	26	26	26
RMSE	8.196	8.173	8.285	8.350	8.396	8.442

Notes: Equation 2 is estimated using ordinary least square with country fixed effects. The target ex-post revision is estimated on the sample 1971-20018 while for the model assessment the reference sample is 1981q1-2007q4. Robust standard errors are reported in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table A.3: Predicting Crisis 2 year before

VARIABLES	$Pr(Crisis_{c,t} = 1)$		
$F_{c,t-8}$	0.0567***	0.112***	
	(0.00466)	(0.00904)	
$\widehat{C}_{c,t-8}$		0.130***	
		(0.0173)	
$S_{c,t-8}$			0.115***
			(0.00614)
Constant	-1.089***	-0.978***	-1.252***
	(0.198)	(0.199)	(0.200)
Observations	3080	3080	3080
AIC	0.780	0.762	0.647
BIC	-22199	-22250	-22609
Deviance	2357	2298	1946

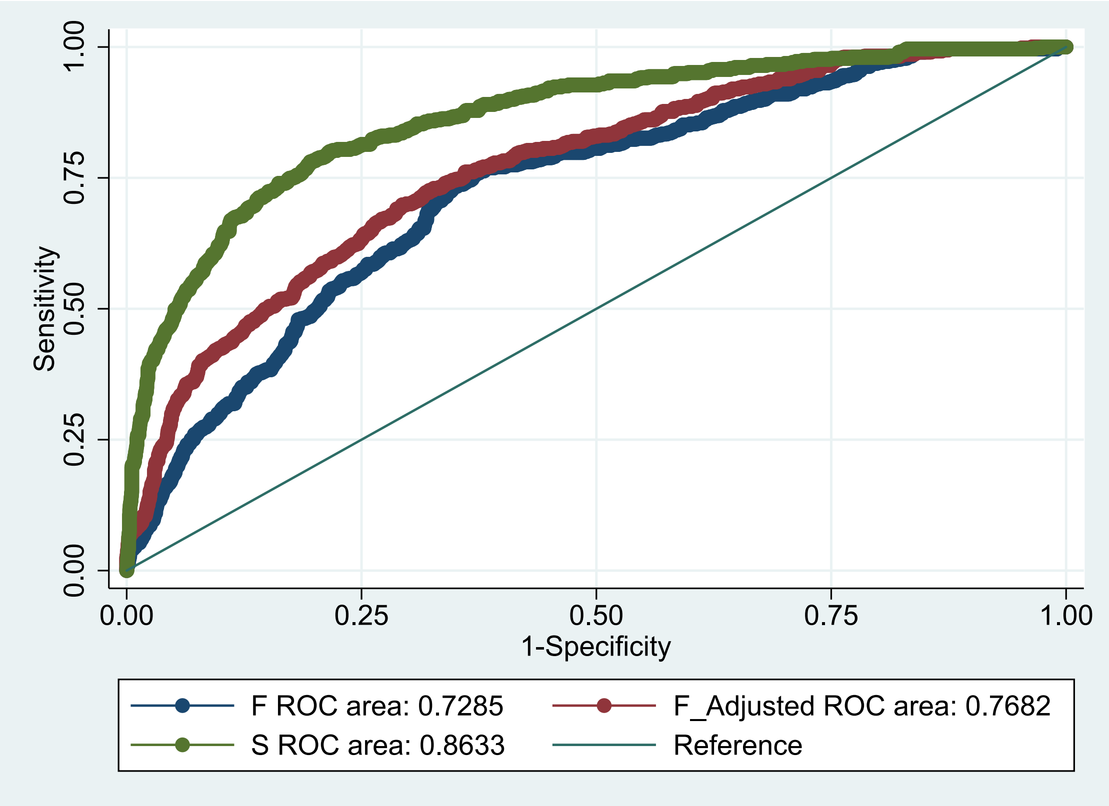
Notes: The model is estimated using a generalized linear model for binomial outcome. All the regressions include country fixed effects. The last rows report to three different measures of fit: (i) *AIC* refer to the Akaike information criterion, (ii) *BIC* to the Bayesian information criterion and, (iii) *Deviance* measures the distance of the fitted model with respect to an abstract model that fits perfectly the sample assigning probability 0 or 1 based on the actual value, larger is the deviance the lower the fit. Standard errors are reported in parentheses. $***p < 0.01$, $**p < 0.05$, $*p < 0.1$.

Table A.4: Reliability Criteria - Short sample

	Correlation			Similarity			Syncronicity			Volatility		
Country	F_t	F_t^{*RW}	F_t^{*ARDL}	F_t	F_t^{*RW}	F_t^{*ARDL}	F_t	F_t^{*RW}	F_t^{*ARDL}	F_t	F_t^{*RW}	F_t^{*ARDL}
Australia	0.81	0.76	0.91	-4.93	-2.04	-2.98	0.18	0.1	0.33	0.95	0.68	0.6
Austria	0.85	0.78	0.85	-1.3	-1.71	-2.87	0.63	0.55	0.46	0.91	0.69	0.59
Belgium	0.52	0.53	0.52	-9.93	-3.64	-2.48	-0.08	-0.07	-0.01	1.48	1	1
Canada	0.83	0.75	0.9	-2.38	-1.12	-1.01	0.55	0.53	0.66	0.85	0.62	0.53
Denmark	0.65	0.6	0.72	-2.22	-1.83	-1.96	0.42	0.35	0.53	1.04	0.74	0.64
Finland	0.69	0.63	0.8	-20.55	-5.71	-2.45	0.08	0.03	0.38	0.94	0.68	0.6
France	0.73	0.66	0.79	-9.3	-1.88	-0.64	0.16	0.1	0.59	0.89	0.64	0.58
Germany	0.54	0.47	0.75	-1.6	-2.79	-1.44	0.44	0.4	0.51	0.86	0.63	0.51
Greece	-0.06	-0.07	0.01	-3.53	-2.99	-69.4	-0.23	-0.27	-0.21	1.28	0.89	0.82
India	0.39	0.35	0.5	-2.66	-13.48	-17.17	0.29	0.25	0.29	1.03	0.75	0.62
Ireland	-0.17	-0.17	-0.27	-10.84	-4.18	-7.35	-0.4	-0.48	-0.38	1.04	0.77	0.57
Italy	0.81	0.81	0.77	-9.39	-7.55	-3.74	-0.07	-0.07	0.08	1.4	0.96	1.05
Japan	0.38	0.33	0.51	-2.46	-7.26	-1.61	0.14	0.1	0.29	1.24	0.86	0.84
Korea	0.78	0.74	0.79	-0.4	-0.51	-0.61	0.85	0.68	0.79	1.05	0.76	0.66
Netherlands	0.94	0.87	0.84	-1.71	-1.15	-6.12	0.72	0.64	0.68	0.98	0.76	0.63
New Zeland	0.77	0.75	0.69	-2.54	-3.29	-1.49	0.5	0.48	0.44	1	0.74	0.61
Norway	0.66	0.6	0.75	-2.52	-3.5	-2.87	0.01	-0.05	0.33	0.96	0.7	0.61
Portugal	0.41	0.37	0.51	-2.22	-1.56	-1.32	0.29	0.31	0.31	1.16	0.81	0.78
Singapore	0.84	0.79	0.88	-0.59	-0.96	-2.42	0.7	0.64	0.79	0.94	0.66	0.59
South Africa	0.83	0.78	0.74	-6.09	-5.01	-1.14	0.36	0.42	0.42	0.97	0.75	0.52
Spain	0.46	0.46	0.45	-14.25	-1.77	-2.47	-0.27	-0.27	-0.18	1.03	0.73	0.66
Sweden	0.74	0.67	0.85	-38.28	-1.61	-1.66	0.14	0.08	0.42	0.89	0.65	0.57
Switzerland	0.59	0.56	0.58	-4.46	-3.14	-2.89	0.1	0.1	0.14	1.03	0.74	0.71
Thailand	0.55	0.5	0.66	-64.9	-2.3	-0.91	0.25	0.23	0.42	1.08	0.76	0.7
United Kingdom	0.65	0.6	0.74	-1.78	-6.02	-1.74	0.36	0.35	0.5	0.86	0.61	0.57
United States	0.84	0.8	0.92	-1.46	-1.94	-0.49	0.57	0.5	0.81	0.83	0.59	0.53
Average	0.62	0.57	0.66	-8.55	-3.42	-5.43	0.26	0.22	0.36	1.03	0.74	0.66

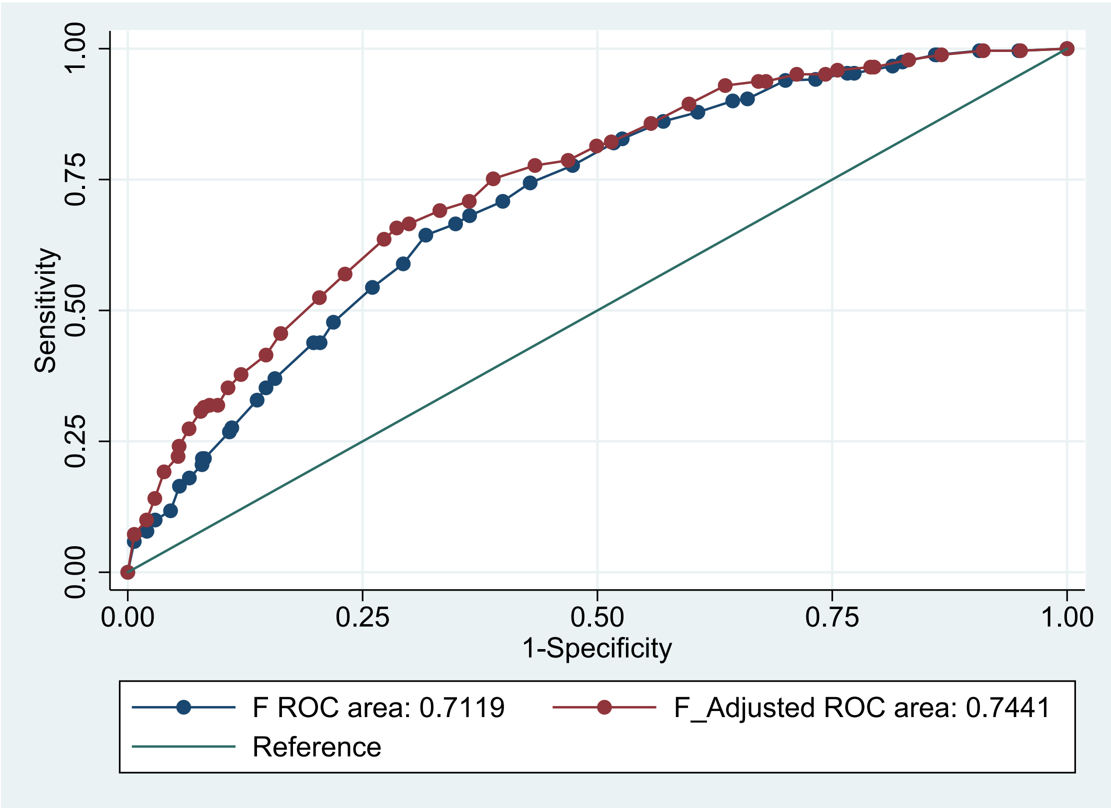
Notes: Averages for the full sample 1981q1-2007q4. The first three columns report the correlation with the smoother. Synchronicity ranges between 1 and -1 and refers to $Sync(F_t, S_t) = (F_t S_t)/|F_t S_t|$. Similarity ranges between 0 and -N and refers to $Sym = -|F_t - S_t|/|F_t + S_t|$.

Figure A.4: AUROC: Gaps relative performance in predicting crises (1 year before crisis)



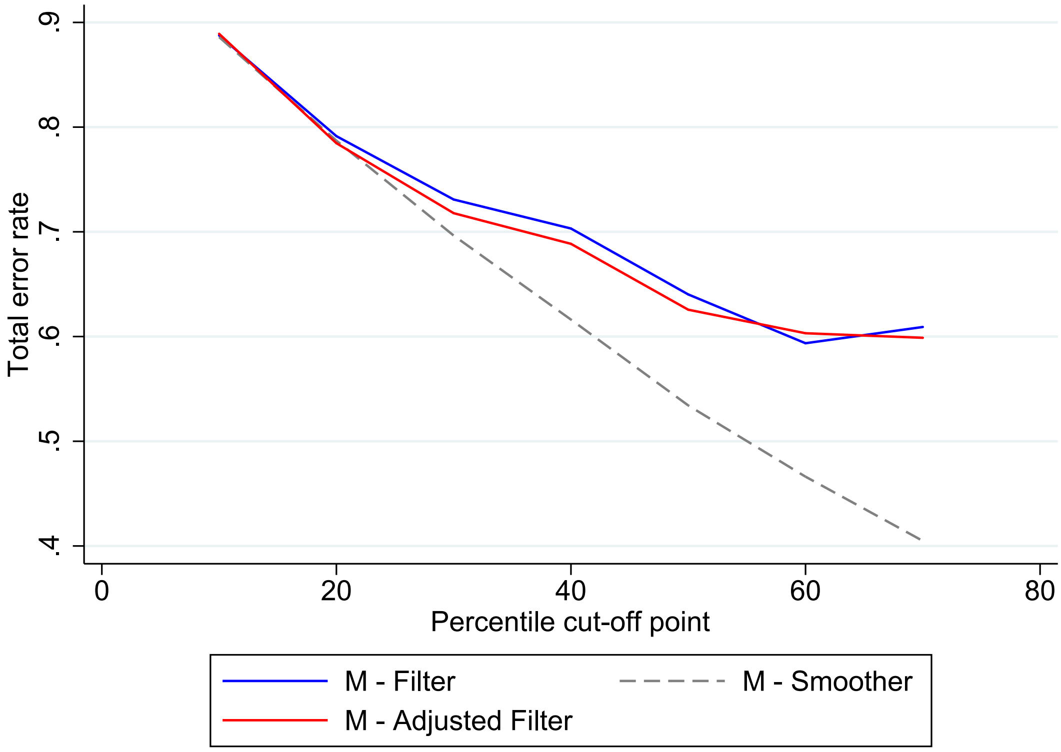
Notes: The area under the receiver operating characteristic curve (AUROC) is a summary measure of the signalling quality of indicator. Drehmann and Juselius (2014) adopt the AUROC to assess the relative performance of different early warning indicators of banking crises.

Figure A.5: AUROC: CCyB at the maximum 1 year before crises



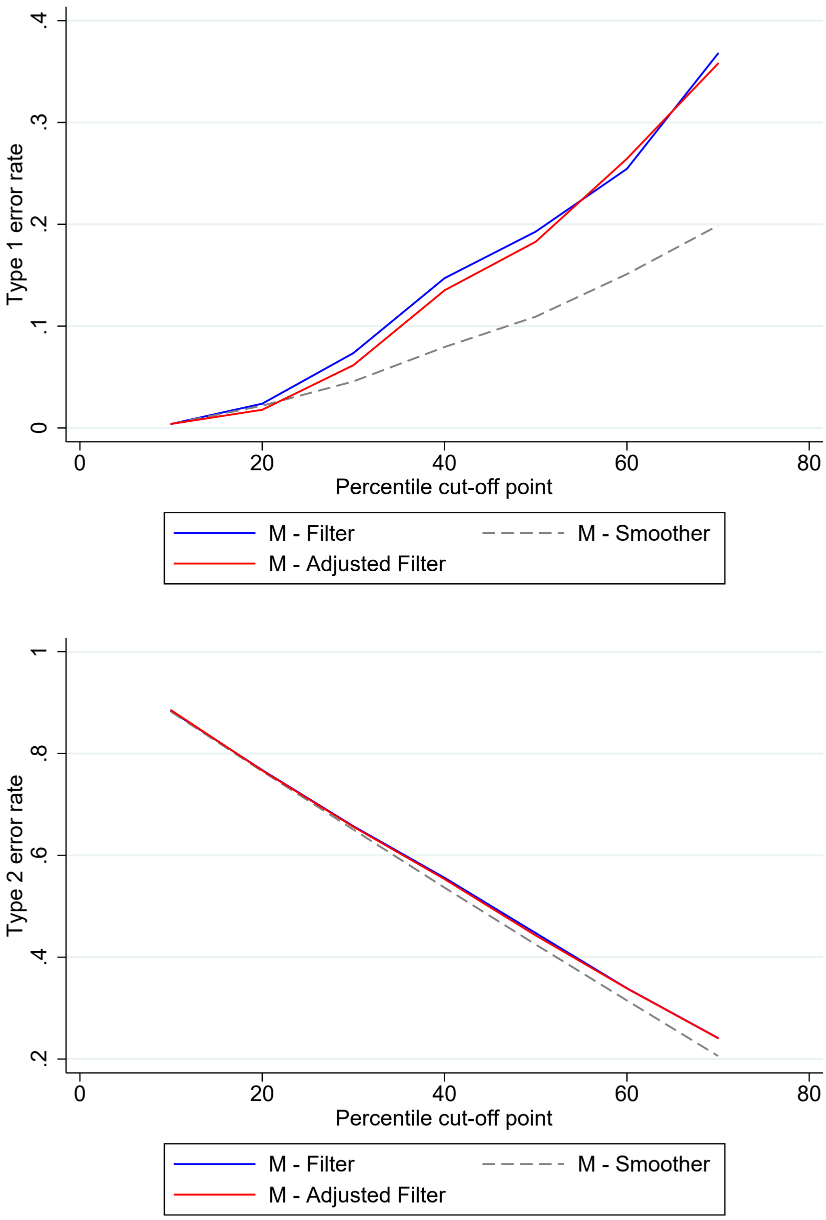
Notes: The area under the receiver operating characteristic curve (AUROC) is a summary measure of the signalling quality of indicator. Drehmann and Juselius (2014) adopt the AUROC to assess the relative performance of different early warning indicators of banking crises.

Figure A.6: Predicted crises classification rate (2 year before crisis)



Notes: This figure plots the incidence of Type 1 errors (classifying a crisis as normal period) and Type 2 errors (classifying a normal period as crisis) according to model predicted probability. For a given percentile, we report the frequency of Type 1 and Type 2 errors. For example, at the 50th percentile, we consider the incidence of Type 1 and Type 2 errors if all observations with a predicted probability above the 50th percentile are classified as crisis and all others as normal periods. The horizontal axis presents the percentile cut-off points while the vertical axis the sum of Type 1 and Type 2 errors. Type 1 and Type 2 errors are reported separately in Figure A.7.

Figure A.7: Predicted crises classification rate (2 year before crisis)



Notes: The upper panel plots the incidence of Type 1 errors (classifying a crisis as normal period). The bottom panel reports Type 2 errors (classifying a normal period as crisis). For a given percentile cut-off, the frequency of Type 1 or Type 2 errors is reported.

Imprint and acknowledgements

The authors appreciate comments and suggestions from Stijn Claessens and Luigi Federico Signorini. The views expressed are those of the authors and should not be attributed to the Bank of Italy or the Eurosystem. The authors retain all responsibility for any errors and omissions. On behalf of all authors, the corresponding author states that there is no conflict of interest.

Piergiorgio Alessandri
Banca d'Italia, Rome, Italy; email: **piergiorgio.alessandri@bancaditalia.it**

Pierluigi Bologna
Banca d'Italia, Rome, Italy; email: **pierluigi.bologna@bancaditalia.it**

Maddalena Galardo (corresponding author)
Banca d'Italia, Rome, Italy; email: **maddalena.galardo@bancaditalia.it**

Postal address	60640 Frankfurt am Main, Germany
Telephone	+49 69 1344 0
Website	**www.esrb.europa.eu**

Note:
The views expressed in ESRB Working Papers are those of the authors and do not necessarily reflect the official stance of the ESRB, its member institutions, or the institutions to which the authors are affiliated.

ISSN	2467-0677 (pdf)
ISBN	978-92-9472-178-5 (pdf)
DOI	10.2849/200497 (pdf)
EU catalogue No	DT-AD-21-001-EN-N (pdf)