

No. 13 | 2013

Staff Memo

Financial Stability

Key indicators for a countercyclical capital buffer in Norway - Trends and uncertainty

Karsten Gerdrup, Aslak Bakke Kvinlog and Eric Schaanning

Staff Memos present reports and documentation written by staff members and affiliates of Norges Bank, the central bank of Norway. Views and conclusions expressed in Staff Memos should not be taken to represent the views of Norges Bank.

© 2013 Norges Bank

The text may be quoted or referred to, provided that due acknowledgement is given to source.

Staff Memo inneholder utredninger og dokumentasjon skrevet av Norges Banks ansatte og andre forfattere tilknyttet Norges Bank. Synspunkter og konklusjoner i arbeidene er ikke nødvendigvis representative for Norges Banks.

© 2013 Norges Bank

Det kan siteres fra eller henvises til dette arbeid, gitt at forfatter og Norges Bank oppgis som kilde.

ISSN 1504-2596 (online only)

ISBN 978-82-7553-758-2 (online only)

Key indicators for a countercyclical capital buffer in Norway - Trends and Uncertainty *

Karsten Gerdrup[†] Aslak Bakke Kvinlog[‡] Eric Schaanning[§]

June 19, 2013

Abstract

The credit-to-GDP gap has a prominent role in the Basel Committee's framework for a countercyclical capital buffer under Basel III. The Committee uses a one-sided Hodrick-Prescott filter to calculate the trend of credit-to-GDP. In this paper we suggest applying the filter to a sample of the indicator where the historical observations have been augmented with forecasts of the indicator. This may provide a more robust estimate of the gap (deviation of indicator from its trend) and thereby a more reliable early warning of a crisis. We analyse Norges Bank's four key indicators for identifying a build-up of imbalances: credit-to-GDP, house prices-to-income, real commercial property prices and the wholesale funding ratio of Norwegian credit institutions. We find that we can reduce revisions in the gaps and improve their signalling quality as indicated by a ROC/AUC analysis even by using forecasts based on a relatively simple method. The forecast is an average of the quarterly indicator variables over the last 4 quarters and the forecast horizon is 5 years.

Keywords: Countercyclical capital buffer; filters; macroprudential policy; AUC; ROC

*We thank Farooq Akram, Ida Wolden Bache, Mikael Juselius, Kjersti-Gro Lindquist and Ingvild Svendsen for helpful comments. The views expressed in this paper are those of the authors and should not be attributed to Norges Bank.

[†] *Corresponding author:* Norges Bank, Email: karsten.gerdrup@norges-bank.no

[‡] Norges Bank, Email: aslak-bakke.kvinlog@norges-bank.no

[§] Imperial College London and Norges Bank, Email: eschaanning@gmail.no

1 Introduction

The credit-to-GDP indicator has a prominent role in the Basel Committee’s framework for a countercyclical capital buffer (CCB) under Basel III. The Committee uses a one-sided Hodrick-Prescott (HP) filter with a high degree of smoothing ($\lambda = 400,000$) to compute the trend in this indicator, see [Basel Committee \(2010\)](#). The advantage of HP-filtering is that more recent observations are given higher weights, which can be an effective means of capturing structural breaks. The basic idea is that when the deviation between the indicator and trend is large, i.e. the gap¹ is high, this may signal a financial crisis a few years ahead and should therefore trigger a response from policymakers to increase banks’ resilience to adverse shocks.

The one-sided version of the HP filter is applied because it only uses data available when macroprudential policy decisions actually may be made. However, the HP filter using all available information (i.e. two-sided) provides a more precise estimate of the trend. [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. The difference between the ex post estimates of the gap based on all information (two-sided filter) and the real-time estimates of the gap (one-sided filter using vintage data) is in their paper used as a measure of revisions. We follow these authors in using the difference between ex post and real-time² estimates of the trend as a measure of revisions. We investigate whether we can provide a more robust estimate of the trend and thereby a more reliable early warning of a crisis by applying the HP filter to a sample of

¹We use ‘gaps’ and ‘cycles’ interchangeably.

²However, due to a lack of vintage data, we limit the analysis to quasi real-time estimates of the gap where we roll over the endpoints of the latest available vintage.

the indicator where the historical observations have been augmented with forecasts of the indicator. By providing more robust gap indicators the risk of errors in the conduct of macroprudential policy may be reduced.

To this end, we analyse Norges Bank's four key indicators for identifying a build-up of imbalances (see [Norges Bank \(2013\)](#)): credit-to-GDP, house prices-to-income, real commercial property prices and the wholesale funding ratio of Norwegian credit institutions. The four gap indicators have provided early warning signals of the banking crisis in Norway (1988-93) and the latest financial crisis (2008-09).³ Even though the financial crisis was not triggered by domestic conditions, banks were still vulnerable prior to the crisis, and the Norwegian authorities had to implement measures to improve access to funding and strengthen banks' solvency. As a basis for the advice on the CCB, Norges Bank will analyse developments in the key indicators and compare these indicators with their corresponding trends.

The paper is organized as follows. In Section 2, we present the HP filter and investigate whether a λ of 400,000 provides a suitable characterization of financial cycles also in Norway. We also explain how the one-sided HP filter is constructed and how it can be applied to historical data augmented with forecasts. In Section 3, we evaluate the gap indicators by comparing revisions to the gaps using different forecasting methods, and by comparing their signalling quality via a ROC/AUC (receiver operating characteristic curve/area under the curve), following the approach in [Drehmann and Juselius \(2013\)](#). We conclude in Section 4.

³Studies using historical data for Norway back to the end of the 1800s also find support for using such indicators, see [Gerdrup \(2003\)](#) and [Riiser \(2005\)](#).

2 Calculation of trends

2.1 Financial cycles

Cycles can be calculated as the deviation between an indicator and its trend. To identify financial cycles, trend calculations should ideally incorporate a view on the sustainable growth path of relevant indicators of credit and financial assets. Our understanding of such growth paths is incomplete, but statistical filters can be useful tools.

The HP filter was originally developed by [Hodrick and Prescott \(1981\)](#)⁴ to facilitate the analysis of fluctuations in economic activity. The authors proposed a method for decomposing a data series into trend and cycle components, i.e. converting a high-frequency series, y_t , into a low-frequency series, μ_t (see also [King and Rebelo \(1993\)](#)). The filter uses all available information, both historical and future data, to estimate the historical trend at each point in time and yields an optimal decomposition of a time series into orthogonal components. This result does not hold for the most recent time periods, see e.g. [Baxter and King \(1999\)](#), but we leave the discussion of the real-time properties of the HP filter to the next section. The filter includes a parameter, λ , which determines the smoothness of the output series. Mathematically the HP filter finds the trend series (μ_t) which minimises the following sum for a given value of λ :

$$\min_{\{\mu_t\}_{t=0}^T} \left(\sum_{t=0}^T (y_t - \mu_t)^2 + \lambda \sum_{t=1}^{T-1} ((\mu_{t+1} - \mu_t) - (\mu_t - \mu_{t-1}))^2 \right)$$

⁴A later version of the paper was published in [Hodrick and Prescott \(1997\)](#)

A higher value of λ implies a higher degree of smoothing. The Basel Committee, and more recently [Borio \(2012\)](#), suggested using $\lambda = 400,000$ to estimate the trend in credit-to-GDP. [Figure 1](#) shows the medium- to long-term financial cycles of Norges Bank's four key indicators using this λ . We use data from 1975Q4 for credit-to-GDP, from 1981Q2 for commercial property prices, from 1978Q4 for house prices-to-disposable income, and from 1975Q4 for banks' wholesale funding ratio. To improve our understanding on the length of these cycles, we also include cycles calculated with the band pass filter in this figure.⁵ The band pass filter enables us to isolate certain frequencies of a time series, and here we have chosen frequencies ranging from 8-30 years.⁶

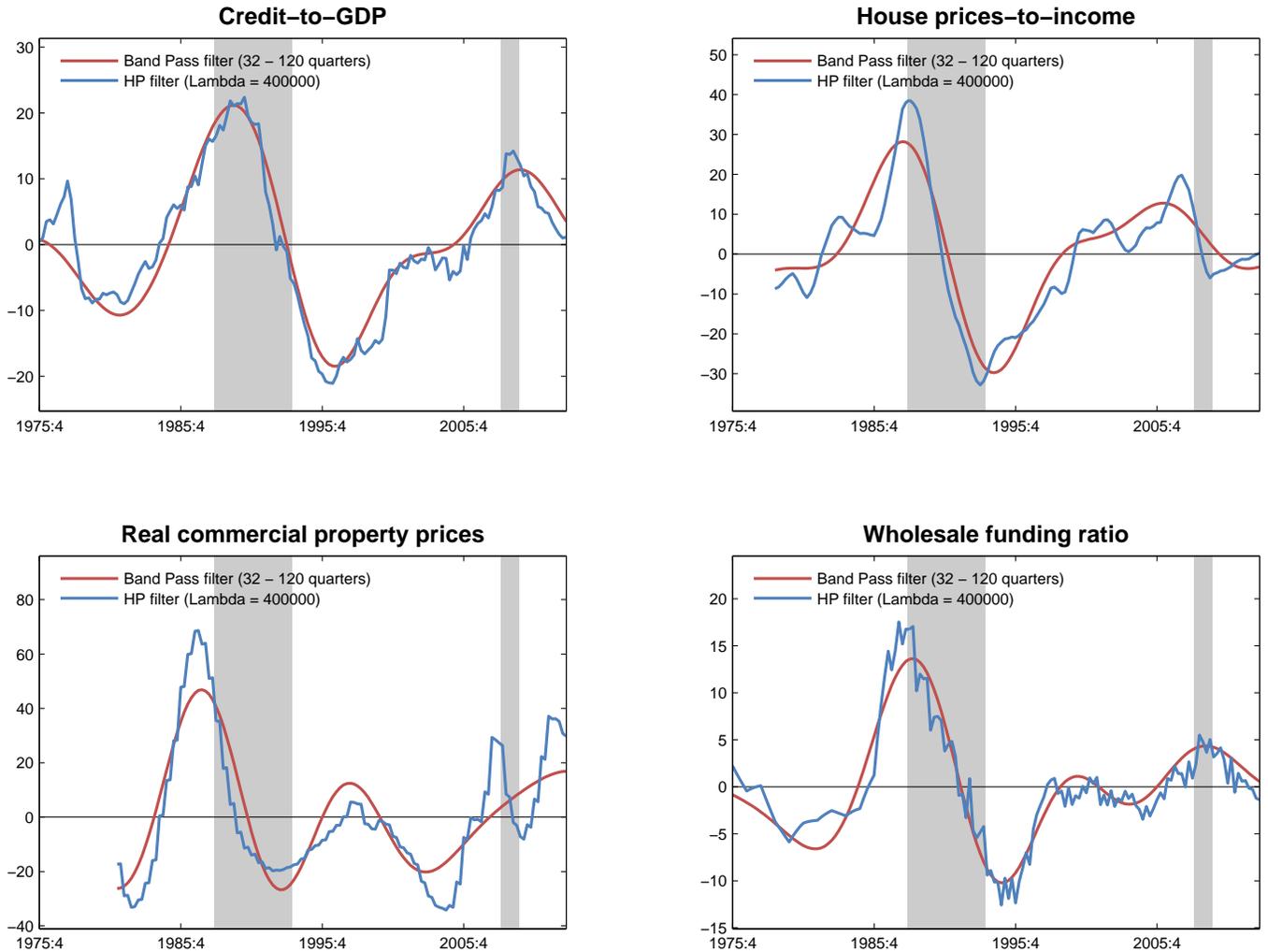
[Figure 1](#) shows that the medium- to long-term cycles build up gradually before the banking crisis (1988-93), and also in the years leading up to the recent financial crisis (2008-09). The financial cycles peaked around the onset of the two crisis episodes. The credit-to-GDP gap peaked during the crises because growth in GDP slowed down faster than credit. The trough between these two peaks was reached around 1993-95 for all the four indicators, which corresponds to a period where the Norwegian economy was gradually picking up after the banking crisis.

For comparison we also estimate cycles at business cycle frequencies. [Figure 2](#) shows cycles where we have used $\lambda = 3,000$, corresponding to cycles lasting 1.5-8 years as indicated by the band pass filter. At these frequencies, the indicators are

⁵See [Christiano and Fitzgerald \(1999\)](#) for more information on this filter

⁶ $\lambda = 400,000$ corresponds to a length of financial cycles that are approximately 3-4 times that of business cycles. In business cycle analysis for the U.S. economy, a λ value of 1,600 is often applied to quarterly data. [Ravn and Uhlig \(2002\)](#) have developed a method where they set λ equal to 1,600 multiplied by the fourth power of the frequency rate (the ratio of the desired frequency to the frequency in the business cycle analysis), e.g. $1,600 \times 4^4 \approx 400,000$.

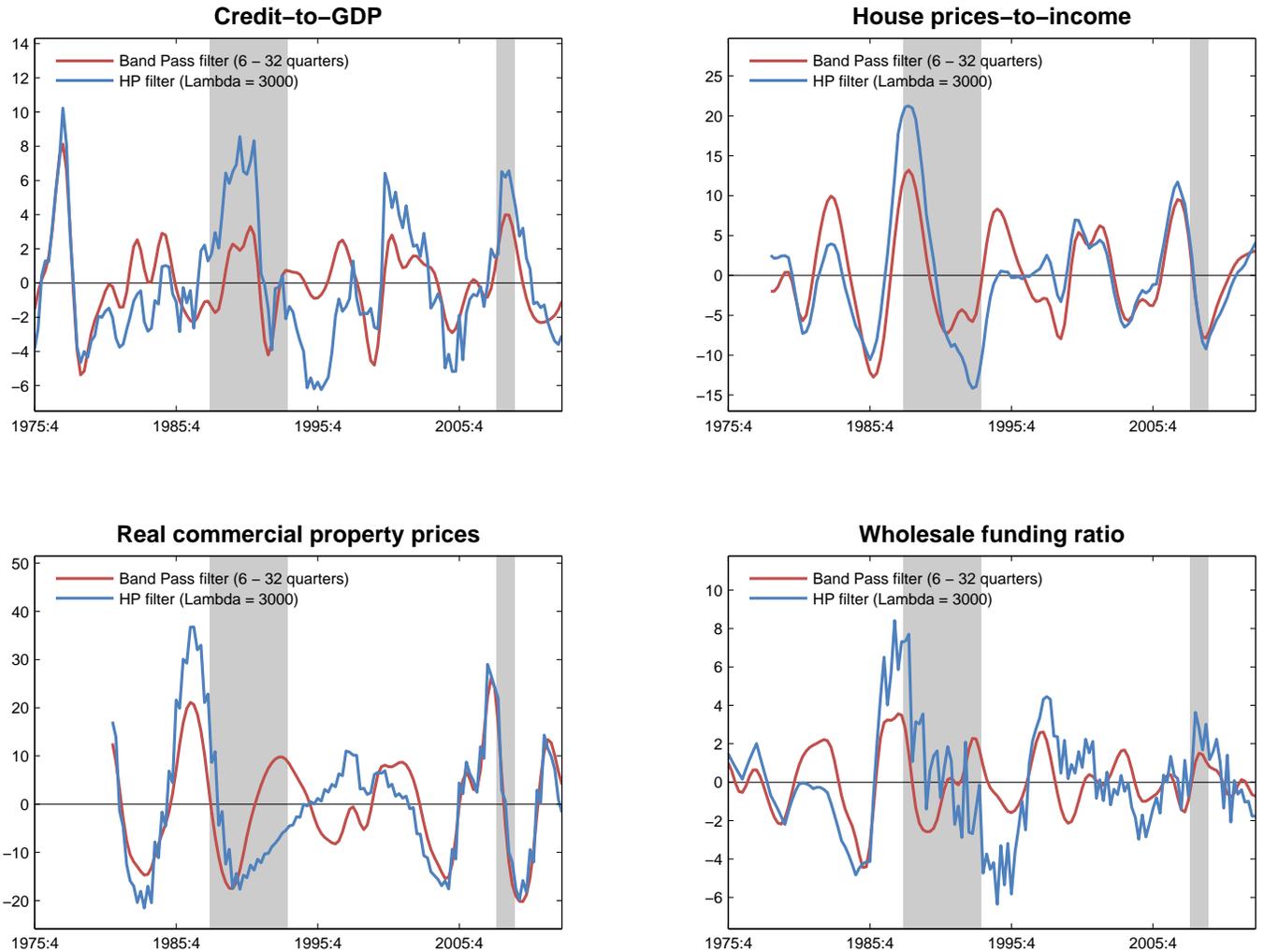
Figure 1. *Financial cycles*



Note: Credit-to-GDP is the sum of C3 non-financial corporations in mainland Norway and C2 households and measured in per cent of mainland GDP. Gap is measured as deviation of this indicator from trend in percentage points. House prices are measured in per cent of disposable income (indexed, 1998Q4=100). The corresponding gap is measured as this indicator in per cent of trend. Commercial property prices are deflated using mainland GDP price deflator (indexed, 1998=100). The corresponding gap is measured as this indicator in per cent of trend. Banks' wholesale funding is measured in per cent of total assets. Gap is measured as deviation of this indicator from trend in percentage points.

Sources: Statistics Norway, Norwegian Association of Real Estate Agents (NEF), Eiendomsmeglerforetakenes Forening (EFF), Finn.no, Eiendomsverdi, IMF, Dagens Næringsliv, OPAK and Norges Bank

Figure 2. *Cycles at business cycle frequency*



Note: Credit-to-GDP is the sum of C3 non-financial corporations in mainland Norway and C2 households and measured in per cent of mainland GDP. Gap is measured as deviation of this indicator from trend in percentage points. House prices are measured in per cent of disposable income (indexed, 1998Q4=100). The corresponding gap is measured as this indicator in per cent of trend. Commercial property prices are deflated using mainland GDP price deflator (indexed, 1998=100). The corresponding gap is measured as this indicator in per cent of trend. Banks' wholesale funding is measured in per cent of total assets. Gap is measured as deviation of this indicator from trend in percentage points.

Sources: Statistics Norway, Norwegian Association of Real Estate Agents (NEF), Eiendomsmeglerforetakenes Forening (EFF), Finn.no, Eiendomsverdi, IMF, Dagens Næringsliv, OPAK and Norges Bank

not useful in providing signals of future crises/stress. They give too many signals, and they do not tend to build up in due time prior to crises. Based on this exercise we believe that $\lambda = 400,000$ is appropriate also for Norwegian data.

2.2 One-sided HP filter

Towards the end of a data series, the estimate of a HP trend is more uncertain because the information set used in the estimation is smaller and little or no future data are available. As the information set is augmented with more data, historical trend estimates are revised and improved. With a λ of 400,000, there will be revisions of the trend up to 20 years back in time each time a new quarter is added to the observation period. The trend can also be revised due to data revisions. For example, statistical agencies typically revise historical GDP series as new information becomes available or methods improve.

To evaluate gap indicators for macroprudential policy, we need to calculate the real-time values of the gaps. While techniques other than the HP filter could be used to estimate these trends, such as the linear trend method or moving average filtering, our aim is to investigate whether we can improve upon the one-sided HP filter. We follow [Drehman et al. \(2011\)](#) and [Basel Committee \(2010\)](#) in using recursively estimated HP trends in order to express what the trend would be at any point in time during the observation period (a one-sided filter).

The two-sided filter coincides with the one-sided filter at the end of the observation period, which means that the one-sided filter consists of all the endpoints from a two-sided filter. However, we need a minimum of observations ($minT$) in order to

calculate a trend. To calculate the one-sided filter, we first calculate the HP trend for each column vector from column $minT$ in the following data matrix:

$$Y = \begin{bmatrix} y_1 & y_1 & \cdots & y_1 & y_1 & y_1 & \cdots & y_1 \\ & y_2 & \cdots & y_2 & y_2 & y_2 & \cdots & y_2 \\ & & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ & & & y_{minT-1} & \vdots & \vdots & \ddots & \vdots \\ & & & & y_{minT} & \vdots & \ddots & \vdots \\ & & & & & y_{minT+1} & \ddots & \vdots \\ & & & & & & \ddots & \vdots \\ & & & & & & & y_T \end{bmatrix}$$

In the matrix y_t is the value of the indicator at time t . All the two-sided trend series (i.e. the HP filter applied to the last $T - minT + 1$ columns in the Y matrix) as well as the one-sided trend (bold symbols) are shown in the following matrix:

$$TREND = \begin{bmatrix} \boldsymbol{\mu}_{1,minT} & \mu_{1,minT+1} & \cdots & \mu_{1,T} \\ \boldsymbol{\mu}_{2,minT} & \mu_{2,minT+1} & \cdots & \mu_{2,T} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{\mu}_{minT,minT} & \vdots & \ddots & \vdots \\ & \boldsymbol{\mu}_{minT+1,minT+1} & \ddots & \vdots \\ & & \ddots & \vdots \\ & & & \boldsymbol{\mu}_{T,T} \end{bmatrix}$$

Thus, the final trend series consists of the first column vector and all the following endpoints:

$$TREND^{O-S} = (\mu_{1,minT}, \mu_{2,minT}, \dots, \mu_{minT,minT}, \mu_{minT+1,minT+1}, \dots, \mu_{T,T}).$$

2.3 HP filter with recursive forecasts

A well-known technique for making the trend estimation more stable and less sensitive to strong variation of the indicators towards the end of the observation period is to extend the observation period with a forecast over a certain horizon. [Kaiser and Maravall \(1999\)](#) have shown that the use of forecast-augmented series in the HP filter can reduce the revision errors of the most recent cyclical components. [Mise et al. \(2005\)](#) confirm this result. [Aastveit and Trovik \(2008\)](#) augment data with forecasts based on a factor model in order to reduce endpoint uncertainty when estimating the output gap for Norway.

Using forecast-augmented data in the filter implies that the trend will be affected in part by the historical series up to the time of calculation and in part by the forecast. The weight of the forecast will, among other things, depend on the forecast horizon. The mathematical formulation of this problem is:

$$\min_{\{\mu_t\}_{t=0}^{T+H}} \left(\sum_{t=0}^{T+H} (y_t - \mu_t)^2 + \lambda \sum_{t=1}^{T-1+H} ((\mu_{t+1} - \mu_t) - (\mu_t - \mu_{t-1}))^2 \right)$$

In this case the time series of the indicator consists of historical observations up to time T , and forecasts for the period $T + 1$ to $T + H$, where H is the forecast horizon.

In matrix notation we can show the procedure by extending each column vector in Y with forecasts (indicated in bold symbols) except for the first $(\min T - 1)$ columns:

$$Yf = \begin{bmatrix} y_1 & y_1 & \cdots & y_1 & y_1 & & y_1 & \cdots & y_1 \\ & y_2 & \cdots & y_2 & y_2 & & y_2 & \cdots & y_2 \\ & & \ddots & \vdots & \vdots & & \vdots & \ddots & \vdots \\ & & & y_{\min T-1} & \vdots & & \vdots & \ddots & \vdots \\ & & & & y_{\min T} & & \vdots & \ddots & \vdots \\ & & & & \mathbf{y}_{\min T+1, \min T} & & y_{\min T+1} & \ddots & \vdots \\ & & & & \vdots & & \mathbf{y}_{\min T+2, \min T+1} & \ddots & \vdots \\ & & & \mathbf{y}_{\min T+H, \min T} & & & \vdots & \ddots & \vdots \\ & & & & & & \mathbf{y}_{\min T+1+H, \min T+1} & \ddots & \vdots \\ & & & & & & & \ddots & \vdots \\ & & & & & & & & y_T \\ & & & & & & & & \mathbf{y}_{T+1, T} \\ & & & & & & & & \vdots \\ & & & & & & & & \mathbf{y}_{T+H, T} \end{bmatrix}$$

All the two-sided trend series (in all $T - \min T + 1$ columns) as well as the one-sided trend (bold symbols) are shown in the following matrix:

$$TRENDf = \begin{bmatrix} \mathbf{\mu}_{1, \min T} & \mu_{1, \min T+1} & \cdots & \mu_{1, T} \\ \mathbf{\mu}_{2, \min T} & \mu_{2, \min T+1} & \cdots & \mu_{2, T} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\mu}_{\min T, \min T} & \vdots & \ddots & \vdots \\ \mu_{\min T+1, \min T} & \mathbf{\mu}_{\min T+1, \min T+1} & \ddots & \vdots \\ \vdots & \mu_{\min T+2, \min T+1} & \ddots & \vdots \\ \mu_{\min T+H, \min T} & \vdots & \ddots & \vdots \\ & \mu_{\min T+1+H, \min T+1} & \ddots & \vdots \\ & & \ddots & \vdots \\ & & & \mathbf{\mu}_{T, T} \\ & & & \mu_{T+1, T} \\ & & & \vdots \\ & & & \mu_{T+H, T} \end{bmatrix}$$

The one-sided filter estimates of the trend can, as before, be put together as

follows (indicated in bold symbols in the matrix):

$$TRENDf^{O-S} = (\mu_{1,minT}, \mu_{2,minT}, \dots, \mu_{minT,minT}, \mu_{minT+1,minT+1}, \dots, \mu_{T,T}).$$

The forecast must be mechanical in nature, since there is no structural model available that is recursively estimated over this time period. We will in Section 3 evaluate the trends using a HP filter and three different forecasting schemes in addition to the one-sided HP filter without any forecasts. The three different forecasting schemes for horizons $h = 1, 2, \dots, H$ can be formulated as follows (recursively estimated for all $t = minT, \dots, T$):

$$\textit{Rolling average forecast} : y_{t+h} = \frac{1}{4} \sum_{s=t-3}^t y_s \quad (1)$$

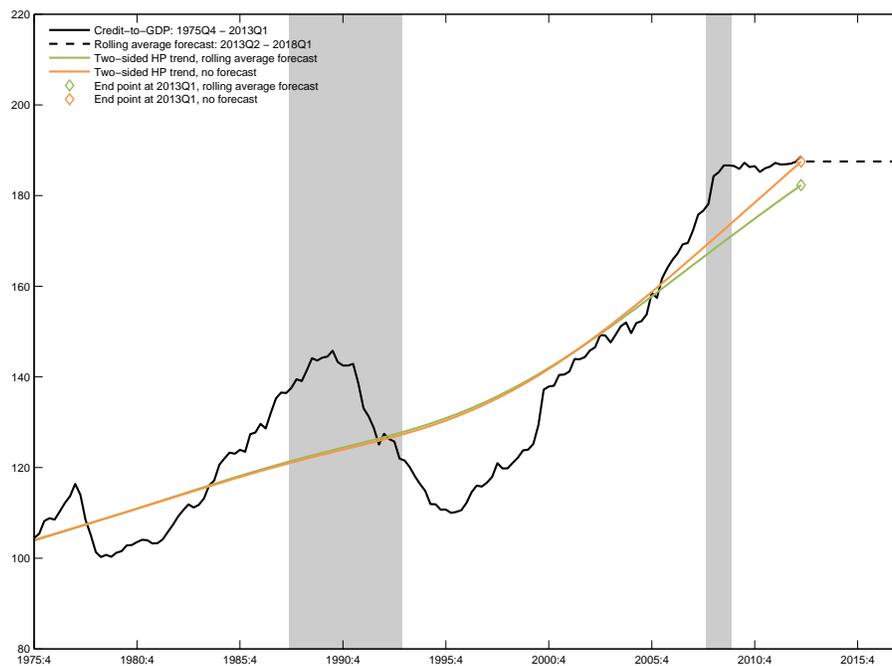
$$\textit{Linear forecast} : y_{t+h} = \alpha_{1:t} + \beta_{1:t} \times (t + h) \quad (2)$$

$$\textit{Rolling linear forecast} : y_{t+h} = \alpha_{t-20:t} + \beta_{t-20:t} \times (t + h) \quad (3)$$

In the first method, we assume that the indicator remains at the same level in the forecast period as at the end of the observation period. To avoid excessive weighting of variations in single observations, an average for the preceding four quarters is used. The next method assumes that the indicator follows a linear trend. The coefficients α and β are recursively updated using data from the start of the observation period up to time t . The third method also assumes that the indicator follows a linear trend, but in this case the coefficients α and β are estimated on data for the last 20 quarters (rolling sample) to take into account possible structural breaks. The coefficients in method 2 and 3 are estimated using ordinary least squares. The

forecast horizon (H) used is always 20 quarters.⁷ Figure 3 depicts an example of two-sided trend calculations and end-points with and without a simple forecast method (rolling average). The figure depicts the last column in matrix $TRENDf$ (with forecast) and $TREND$ (without forecasts) for credit-to-GDP up to time T .

Figure 3. *Example of two-sided HP filtered trends and end-points. Credit-to-GDP augmented with rolling average forecast and no forecasts. Per cent*



Note: Credit-to-GDP is the sum of C3 non-financial corporations in mainland Norway and C2 households and measured in per cent of mainland GDP.

Sources: Statistics Norway, IMF and Norges Bank

⁷The estimated trends change little when the forecast horizon is extended.

3 Evaluating gaps

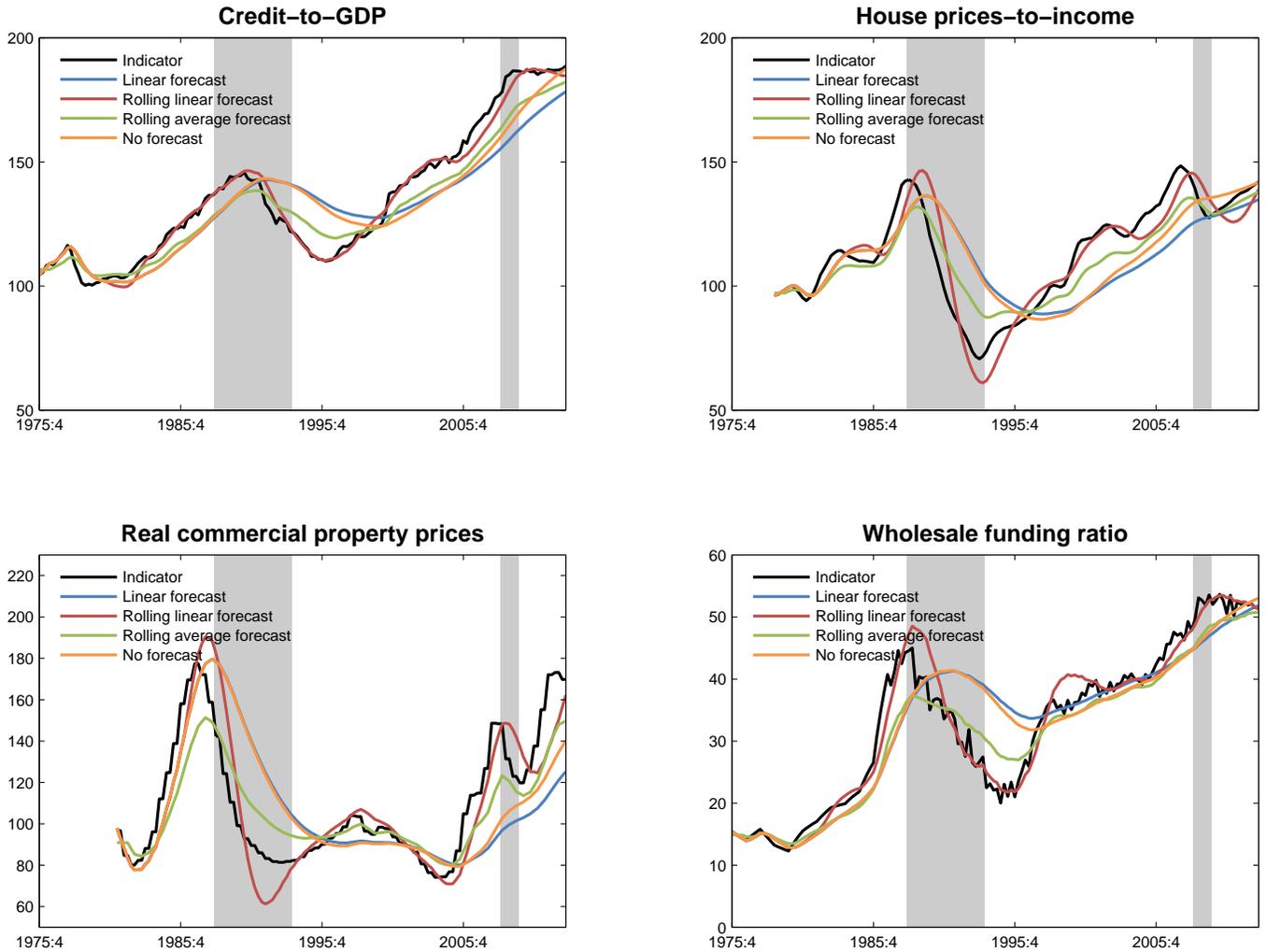
We use the difference between ex post and quasi real-time estimates of the trend as a measure of revisions. We should in principle take into account that a HP trend also may change due to data revisions. A lack of vintage data for a longer period makes this difficult to accomplish. However, analyses on output gap estimates using vintage data for Norway (see [Bernhardsen et al. \(2005\)](#)) and for the US (see [Orphanides and van Norden \(2002\)](#)) show that revisions due to end-of-sample uncertainty is the primary source of revisions in measured output gaps, and thus much larger than data revisions. [Edge and Meisenzahl \(2011\)](#) find the same result when measuring credit-to-GDP gaps in real time, and find furthermore that ex post revisions to the gap are sizable and as large as the gap itself. This motivates finding robust forecasting schemes in order to reduce this uncertainty.

In [Section 3.1](#) we calculate the revisions in the gap indicators for the different forecast methods. In [Section 3.2](#) we look at the gap indicators' ability to predict the two mentioned episodes of financial crisis or stress: the banking crisis 1988-93, and the financial crisis 2008-09. To do so, we evaluate the gap indicators' signalling quality via a ROC/AUC analysis.

3.1 Trend estimate uncertainty

[Figure 4](#) depicts each key indicator for the CCB in Norway alongside the alternative trends. An important difference between the one-sided HP filter applied on historical data only and the ones applied on historical data augmented with forecasts, is that the first appears to lag the indicators to a greater extent. This means that the peaks

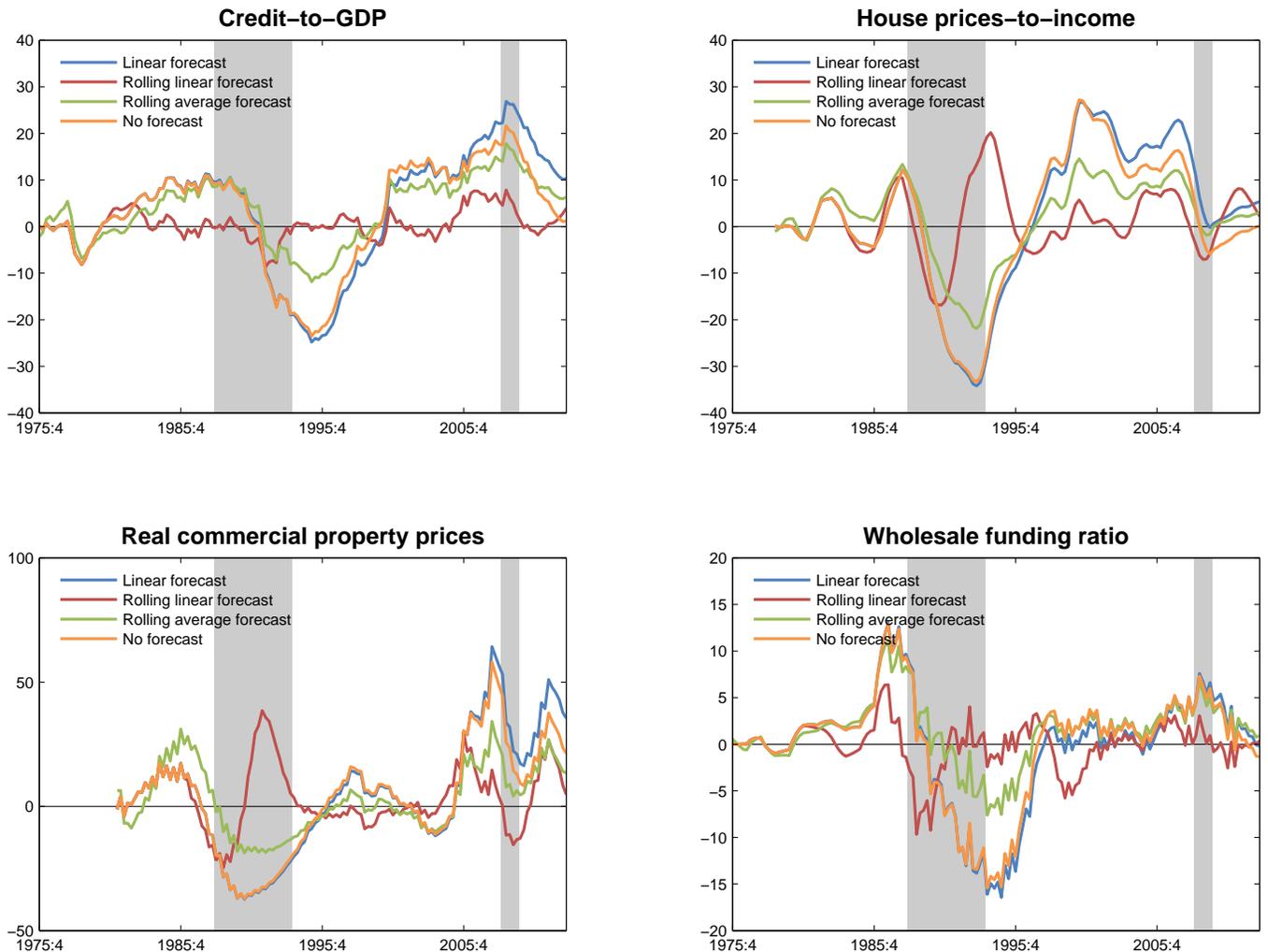
Figure 4. *Key indicators and trend estimates*



Note: Credit-to-GDP is the sum of C3 non-financial corporations in mainland Norway and C2 households and measured in per cent of mainland GDP. House prices are measured in per cent of disposable income (indexed, 1998Q4=100). Commercial property prices are deflated using mainland GDP price deflator (indexed, 1998=100). Banks' wholesale funding is measured in per cent of total assets.

Sources: Statistics Norway, Norwegian Association of Real Estate Agents (NEF), Eiendomsmeglerforetakenes Forening (EFF), Finn.no, Eiendomsverdi, IMF, Dagens Næringsliv OPAK and Norges Bank

Figure 5. *Key indicator gaps*



Note: Credit-to-GDP is the sum of C3 non-financial corporations in mainland Norway and C2 households and measured in per cent of mainland GDP. Gap is measured as deviation of this indicator from trend in percentage points. House prices are measured in per cent of disposable income (indexed, 1998Q4=100). The corresponding gap is measured as this indicator in per cent of trend. Commercial property prices are deflated using mainland GDP price deflator (indexed, 1998=100). The corresponding gap is measured as this indicator in per cent of trend. Banks' wholesale funding is measured in per cent of total assets. Gap is measured as deviation of this indicator from trend in percentage points.

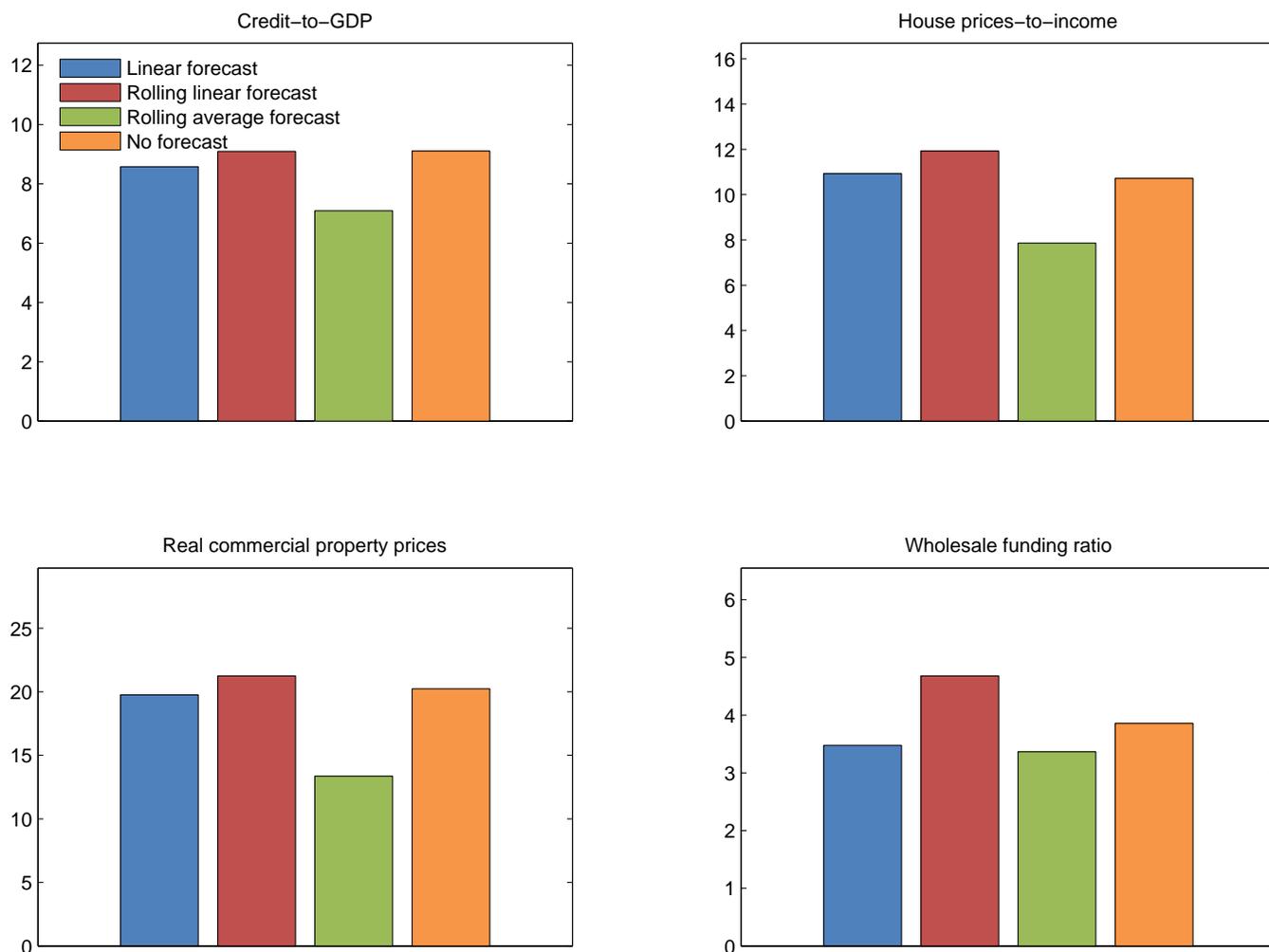
*Sources: Statistics Norway, Norwegian Association of Real Estate Agents (NEF), Eien-
domsmeglerforetakenes Forening (EFF), Finn.no, Eiendomsverdi, IMF, Dagens Næringsliv
OPAK and Norges Bank*

and troughs of the trend are not aligned, i.e. the turning points occur at different dates than the indicators. The corresponding gaps are shown in Figure 5. The gaps follow each other, but there are some interesting differences between the methods: First, the amplitude of the gaps are larger for the method with no forecasts and the one with linear forecasts. The reason is the tendency of the trend to lag the actual indicator, which creates a wide gap after a turning-point in the indicator. Thus, the filter applied on data that are not augmented with appropriate forecasts may create phase-shifts.⁸ The linear forecast method provides trends and gaps that are close to the method with no forecasts. Second, the rolling linear forecast method yields gaps with a small amplitude, and provides no signals prior to the banking crisis. The reason is that this method extrapolates the recent development in the indicator, making the trend track the actual development in the indicator. Third, the rolling average forecast method seems to provide consistently good signals prior to the two crisis episodes for all variables.

A measure of revisions in these gaps are shown in Figure 6 and 7. The first of these shows the mean of the absolute difference between the two-sided HP gap and the HP gap computed for each of the forecasting methods, i.e. the difference between the gap based on the trend in the last column in matrix $TRENDf$ and the gap based on the one-sided trend ($TRENDf^{O-S}$) (see previous section). Figure 6 indicates that the rolling average forecast method provides the lowest revisions in the gap estimate, while the methods with rolling linear forecast and no forecast

⁸Indeed, one of the criteria for an ideal band-pass filter, according to [Baxter and King \(1999\)](#), is that it should not introduce phase shift, i.e. that it should not alter the timing relationships between series at any frequency.

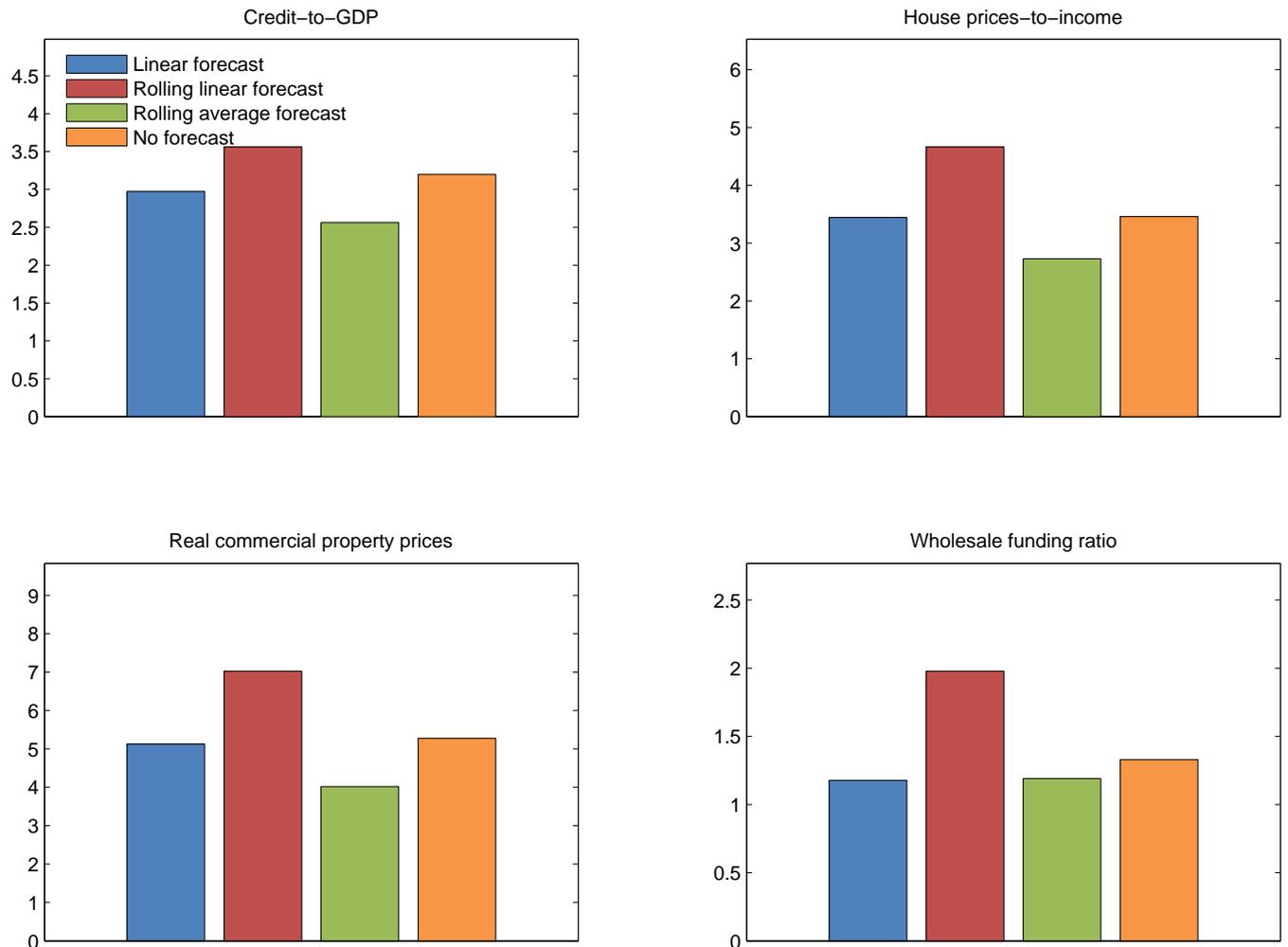
Figure 6. Mean of distance between gap indicators computed with one-sided and two-sided HP trend. Percentage points



Note: A low value means that the trend is revised to a smaller extent from the first estimate (i.e. one-sided) to the most recent estimate (i.e. last two-sided trend).

Sources: Statistics Norway, Norwegian Association of Real Estate Agents (NEF), Eiendomsmeglerforetakenes Forening (EFF), Finn.no, Eiendomsverdi, IMF, Dagens Næringsliv, OPAK and Norges Bank

Figure 7. *Standard deviation of distance between gap indicators computed for all two-sided HP trends*



Note: A low value means that the variation in the estimates of the trend is small.

Sources: Statistics Norway, Norwegian Association of Real Estate Agents (NEF), Eiendomsmeglerforetakenes Forening (EFF), Finn.no, Eiendomsverdi, IMF, Dagens Næringsliv, OPAK and Norges Bank

provide the largest revisions on average. The average standard deviation of revisions for each method is shown in Figure 7. This is technically done by considering the variation in all the two-sided trend estimates for each quarter, i.e. the variation in each row in matrix $TRENDf$. This figure furthermore indicates that the method with a rolling average forecast has the lowest uncertainty. Figures depicting all the two-sided trends and gaps for each variable and forecasting method are shown in the Appendix, see Figures 10-17.

3.2 Evaluation of trend computation methods via ROC and AUC

In this section, we evaluate the different trend computation methods by comparing the quality of the signal that the respective gap indicators produce. To this end, we evaluate the statistical properties of the gap indicators using a ROC/AUC (receiver operating characteristic curve/area under the curve) analysis, following the approach in Drehmann and Juselius (2013)⁹.

3.2.1 A closer look at the ROC and AUC methodology

The method takes as a starting point that the economy can be in two states: A normal state and a (pre-crisis) boom. For a given value of the indicators, we want to evaluate the probability of the economy being in one or the other state. In the first state, no imbalances are building up in the financial system, whereas in the latter state, a tighter macroprudential policy may be appropriate. Early warning

⁹A more detailed exposition on ROC/AUC analyses can be found in Pepe et al. (2009) and Cohen et al. (2009). See Berge and Òscar Jordà (2011) for an economic application.

indicators are used to signal the state of the economy, but this signal is often noisy. The intuition is that the higher the value of the indicator, the more likely we believe the economy to be in a boom that could lead to a crisis. When the probability of being in a (pre-crisis) boom is sufficiently high, policymakers may want to take corrective policy actions, e.g. by increasing the CCB.

Concretely, one chooses a threshold θ such that if the signal S of the early warning indicator lies above the threshold, then the economy is believed to be in a boom, and if it is below, the economy is believed to be in its normal state. This leads to four possible outcomes, summarised in the following table:

Signal (S) \ Real state (R)	Pre-crisis boom (R=1)	Normal state (R=0)
Pre-crisis boom ($S \geq \theta$)	True Positive (TP)	False Positive (FP)
Normal state ($S < \theta$)	False Negative (FN)	True Negative (TN)

When the threshold for a given indicator is very low, that indicator will issue many signals and capture almost all crises. Thus the True Positive Rate (TPR), given by

$$TPR_{\theta} = \mathbb{P}(S \geq \theta | R = 1) = \frac{TP}{TP + FN},$$

is close to unity. However, the indicator will also sound many false alarms, which is reflected by the False Positive Rate (FPR), given by

$$FPR_{\theta} = \mathbb{P}(S \geq \theta | R = 0) = \frac{FP}{FP + TN},$$

also being close to unity.

On the other hand, when the threshold is high, an indicator will provide few

signals and may thus miss many crises (TPR close to zero), but the rate of false signals will also be low (FPR close to zero).

A False Positive corresponds to the classical statistical Type I error, i.e. a false alarm or sounding a crisis when there actually is none. A False Negative corresponds to the statistical Type II error, which characterises a miss. The True Positive and True Negative cases correspond to the situations in which the signal correctly identified the pre-crisis boom and the normal state respectively.

To construct the ROC curve, one iterates over all possible thresholds θ , and for each θ plots TPR_θ against FPR_θ . For $\theta = -\infty$, we have $TPR = 1$ and $FPR = 1$, because the indicator issues a signal all the time and thereby picks up all crises, but at the same time raises numerous false alarms. For $\theta = +\infty$, we get $TPR = 0$ and $FPR = 0$, because no signal at all will be issued. Any threshold θ between these extremes quantifies the trade-off between the TPR and the FPR . A signal S is said to be fully informative, if there exists a threshold θ for S such that $TPR_\theta = 1$ and $FPR_\theta = 0$. A signal is uninformative if the ROC curve lies on the diagonal. For an uninformative signal, the False Positive Rate increases just as quickly as the True Positive Rate.

If one has reasonable estimates of the costs and benefits of macroprudential policy interventions, then one can pick a point on the ROC curve that corresponds to a specific threshold which realizes the optimal trade-off between missing crises (TPR) and imposing unnecessary tight macroprudential policy actions (FPR). For example, if the costs of imposing stricter macroprudential standards are low even in normal states of the economy while costs of crises are high, then the policymaker would

prefer an indicator with a high TPR , which misses very few crises, at the expense of ringing a few false alarms, i.e. having a higher FPR .

However, as [Drehmann and Juselius \(2013\)](#) suggest, in the absence of the cost estimates and the preferences of policymakers, one can alternatively calculate the quality of the indicators for all possible thresholds by calculating the area under the ROC curve (AUC). A high value of the AUC means that the indicator consistently provides more precise signals, independent of the specific threshold that is chosen.

AUC is the area under the ROC curve and given by

$$AUC_h(S) = \int_0^1 ROC_S^h(x) dx,$$

where S indicates a signal issued h quarters before a crisis. It holds that $0 \leq AUC \leq 1$. For a fully informative signal, we have $AUC = 1$, for a fully uninformative signal, we have $AUC = 0.5$. Note that $AUC = 0.5$ does not necessarily imply that the signal is fully uninformative.

The index h , in $AUC_h(S)$, indicates that only observations in quarter h are used to compute the TPR . Following the general literature, signals issued during – and two years after – a crisis are not evaluated at all. Hence, for every early warning indicator S_i , we compute $AUC_h(S_i)$ for all quarters $h \in H = \{-20, \dots, -1\}$, i.e. with the crisis occurring at $h=0$. When evaluating $AUC_h(S_i)$ for a specific quarter h , the TPR (True Positive Rate) is computed using only the signals issued in quarter h . The FPR is evaluated using all signals outside of H except for the periods excluded above.

To provide an illustration of ROC curves and the corresponding AUC values, we

use the credit-to-GDP gap indicator to plot the *ROC* curves and compute *AUC* for all four trend computation methods over four different horizons before the onset of crises, see Figure 8. In Figure 9 we plot the AUC_h values as functions of the horizon ($h \in \{-20, \dots, -1\}$) for all four trend computation methods and all four indicators.

3.2.2 *AUC* results

We use three statistical criteria for evaluating the *AUC* of the gap indicators as depicted in Figure 9: timing, consistency, and relative performance. First, timing has to be appropriate. An indicator should give a signal early enough (long enough horizon) for policymakers to act and banks to react, but not too early since this could make policy decisions seem less relevant for the public. Next, the indicators should be stable and provide consistent signals. This is important since the gaps are based on uncertain trend calculations. Finally, we compare the gap indicators' relative performance.

The three criteria can be described more precisely as follows:¹⁰

Criterion 1 - Timing: S_i has the right timing if:

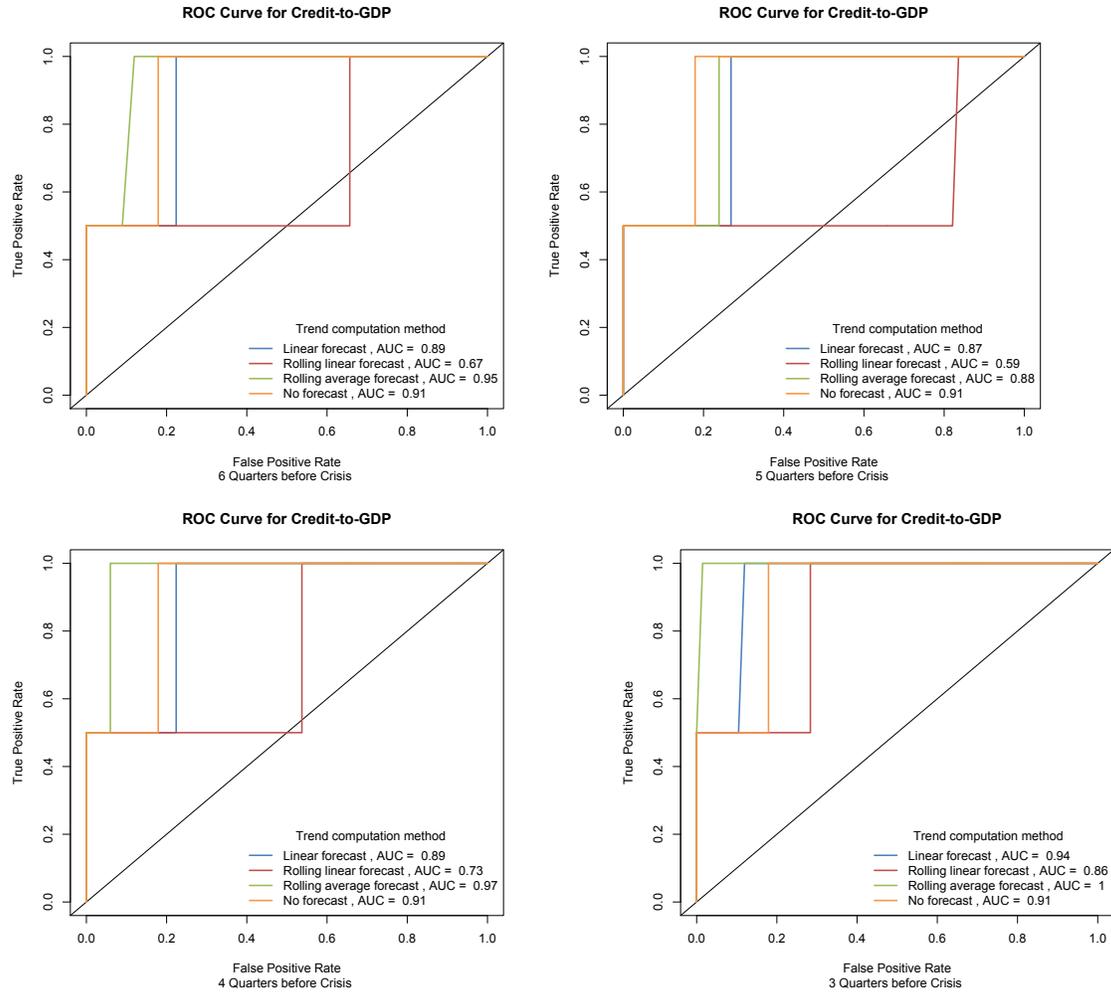
$$AUC_h(S_i) > 0.5 \text{ for some horizon } h \in [-20, -6]$$

¹⁰Drehmann and Juselius (2013) use a different criterion 2 called *stability*, which puts particular emphasis on how the indicator behaves six quarters before the onset of the crisis, relative to other quarters. They define S_i to be stable if:

$$AUC_{-6-j}(S_i) \leq AUC_{-6}(S_i) \leq AUC_{-6+k}(S_i) \text{ for } j = 1, \dots, 14 \text{ and } k = 1, \dots, 5$$

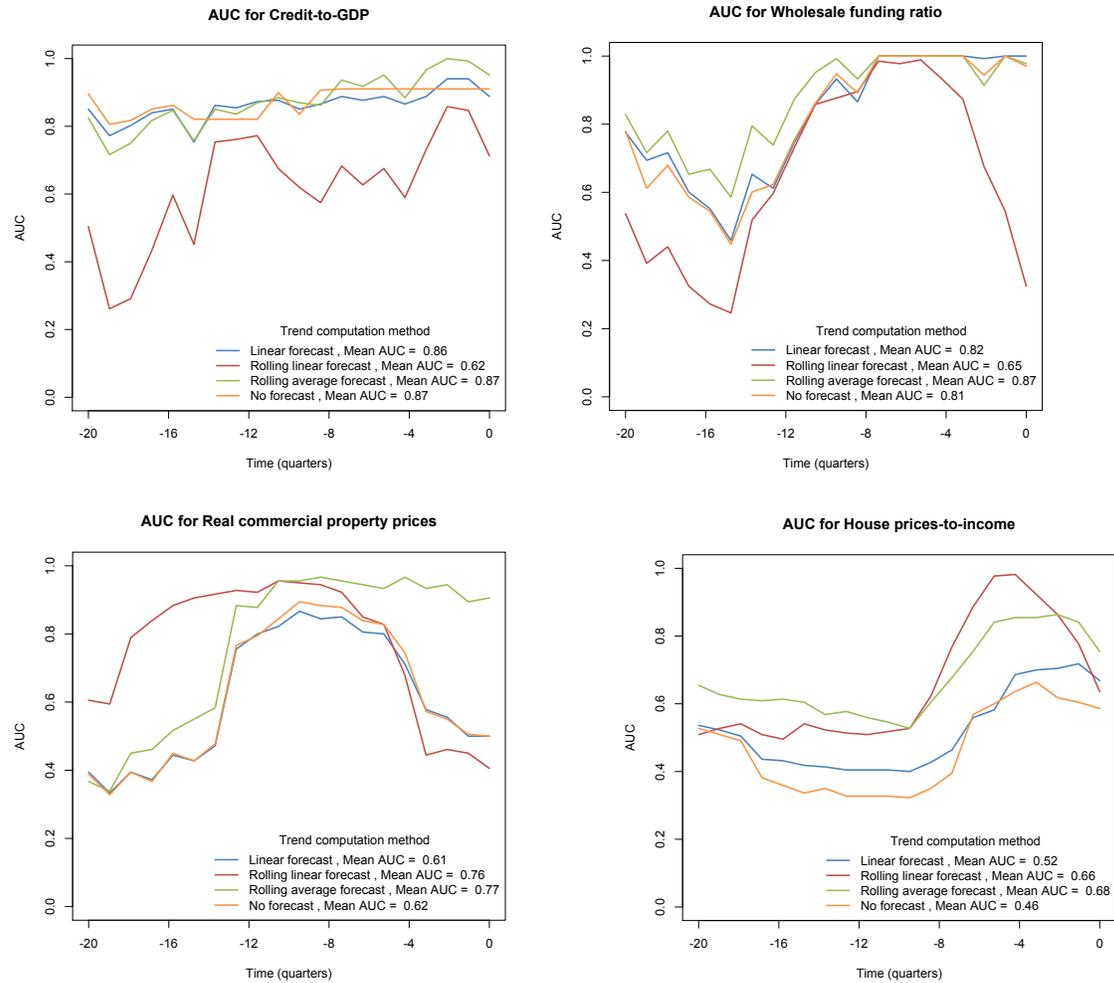
However, we want to take a stricter stance. Also, this definition does not exclude indicators that present decreasing *AUC* behaviour over the isolated horizons $h \in \{-20, -7\}$ or $h \in \{-5, -1\}$. Therefore, we require the *AUC* to be an increasing function of the horizons h .

Figure 8. Visualising ROC and $AUC_h(S)$ for $h \in \{-6, -5, -4, -3\}$ quarters and all four gap computation methods for the credit-to-GDP gap indicator (S)



Sources: Statistics Norway, IMF and Norges Bank

Figure 9. *AUC of the four trend computation methods on four different key indicators for macroprudential policy*



Sources: Statistics Norway, Norwegian Association of Real Estate Agents (NEF), Eiendomsmeglerforetakenes Forening (EFF), Finn.no, Eiendomsverdi, IMF, Dagens Næringsliv, OPAK and Norges Bank

Criterion 2 - Consistency: S_i is consistent over horizon H if

$$AUC_{h-1}(S_i) \leq AUC_h(S_i) \text{ for } h \in H$$

Criterion 3 - Superiority: S_i is superior to S_j over a horizon H if

$$AUC_h(S_i) \geq AUC_h(S_j) \text{ for } h \in H$$

Based on the three criteria mentioned above, we summarize the behaviour of the different trend computations methods in Table 1. A more detailed discussion of the three criteria for each indicator can be found in the Appendix.

Indicator \ Criterion	Timing	Consistency	Superiority (ranking)
Rolling average forecast	Very good	Good	1.
Linear forecast	Good	Fair	2.
No forecast	Fair	Fair	3.
Rolling linear forecast	Fair	Poor	4.

Table 1. Summary of the performance of the different trend computation methods relative to the three criteria on all indicators of Figure 9.

All in all, the signalling quality of the gap indicators that are constructed with the HP rolling average forecast are the best for all four indicators. Also, in the three years before the onset of crises, the AUC values lie on average well above 0.80. These values look very promising and attest to the good predictive qualities of the indicators in general. Nonetheless, given the small number of crises that our sample contains, one must be careful not to over-interpret these results.

4 Conclusion

In this paper we propose to modify the simple, one-sided Hodrick-Prescott filter used by the Basel Committee to derive a credit-to-GDP gap. We find that augmenting historical observations with forecasts based on a simple method before applying the HP filter improves the properties of the gap indicators. The evaluation of the gaps is done in part by comparing the revisions to the gaps using different forecasting methods, and in part by comparing their signalling quality via a ROC/AUC (receiver operating characteristic curve/area under the curve) analysis of the different gap indicators. Our proposed forecasting method is a simple average of the indicator values over the last four quarters and a forecast horizon of 5 years. Using a forecast formula for calculating trends and gaps is a novel feature of Norges Bank's approach to identifying financial imbalances in the conduct of macroprudential policy.

References

- Aastveit, K. A. and T. G. Trovik (2008, December). Estimating the output gap in real time: A factor model approach. Working Paper 2008/23, Norges Bank.
- Basel Committee (2010, December). Guidance for national authorities operating the countercyclical capital buffer. Technical report, Basel Committee on Banking Supervision.
- Baxter, M. and R. G. King (1999, November). Measuring business cycles: Approximate band-pass filters for economic time series. *The Review of Economics and Statistics* 81(4), 575–593.
- Berge, T. J. and Òscar Jordà (2011, April). Evaluating the classification of economic activity into recessions and expansions. *American Economic Journal: Macroeconomics* 3(2), 246–77.
- Bernhardsen, T., O. Eitrheim, A. S. Jore, and O. Roisland (2005, December). Real-time data for Norway: Challenges for monetary policy. *The North American Journal of Economics and Finance* 16(3), 333–349.
- Borio, C. (2012, December). The financial cycle and macroeconomics: What have we learnt? Working Paper 2012/395, Bank for International Settlements.
- Christiano, L. J. and T. J. Fitzgerald (1999). The band pass filter. Working Paper 1999/9906, Federal Reserve Bank of Cleveland.
- Cohen, J., S. Garman, and W. Gorr (2009, July). Empirical calibration of time series

- monitoring methods using receiver operating characteristic curves. *International Journal of Forecasting* 25(3), 484–497.
- Drehman, M., C. Borio, and K. Tsatsaronis (2011, November). Anchoring counter-cyclical capital buffers: the role of credit aggregates. Working Paper 2011/355, Bank for International Settlements.
- Drehmann, M. and M. Juselius (2013). Evaluating EWIs for banking crises - satisfying policy requirements. mimeo, Bank for International Settlements.
- Edge, R. M. and R. R. Meisenzahl (2011, December). 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. (2003, October). Three episodes of financial fragility in Norway since the 1890s. Working Paper 2003/142, Bank for International Settlements.
- Hodrick, R. J. and E. Prescott (1981, May). Post-war u.s. business cycles: An empirical investigation. Discussion Papers 451, Northwestern University, Center for Mathematical Studies in Economics and Management Science.
- Hodrick, R. J. and E. C. Prescott (1997, February). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit and Banking* 29(1), 1–16.
- Kaiser, R. and A. Maravall (1999). Estimation of the business cycle: A modified Hodrick-Prescott filter. *Spanish Economic Review* 1(2), 175–206.
- King, R. G. and S. T. Rebelo (1993). Low frequency filtering and real business cycles. *Journal of Economic Dynamics and Control* 17(1-2), 207 – 231.

- Mise, E., T.-H. Kim, and P. Newbold (2005, March). On suboptimality of the Hodrick-Prescott filter at time series endpoints. *Journal of Macroeconomics* 27(1), 53–67.
- Norges Bank (2013). Criteria for an appropriate countercyclical capital buffer. Norges Bank Papers 2013/1, Norges Bank.
- Orphanides, A. and S. van Norden (2002, November). The unreliability of output-gap estimates in real time. *The Review of Economics and Statistics* 84(4), 569–583.
- Pepe, M. S., G. Longton, and H. Janes (2009, March). Estimation and comparison of receiver operating characteristic curves. *Stata Journal* 9(1), 1–16.
- Ravn, M. O. and H. Uhlig (2002). On adjusting the Hodrick-Prescott filter for the frequency of observations. *The Review of Economics and Statistics* 84(2), 371–375.
- Riiser, M. (2005). House prices, equity prices, investment and credit - what do they tell us about banking crises? A historical analysis based on norwegian data. Economic Bulletin 2005/2, Norges Bank.

5 Appendix

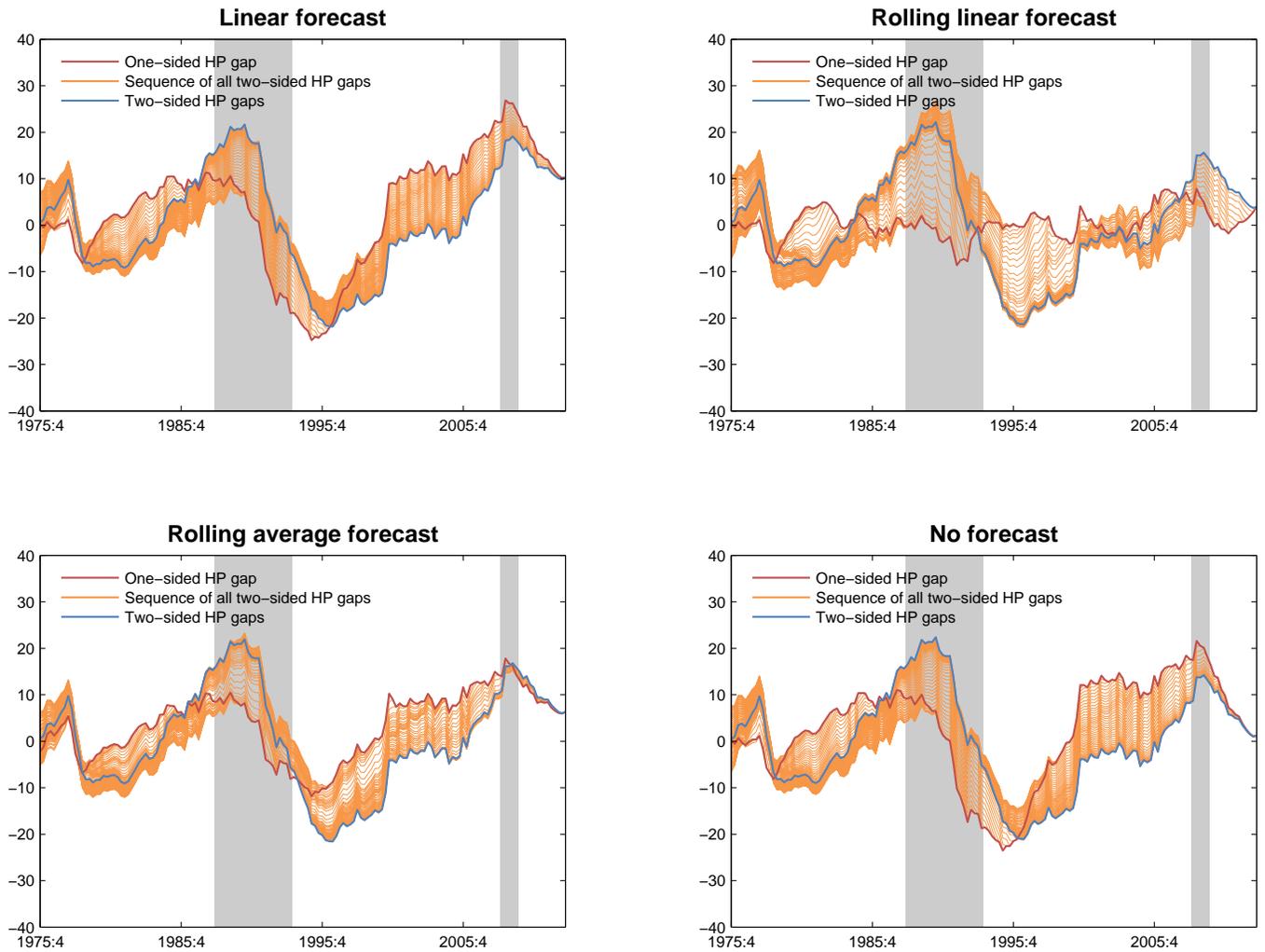
5.1 Detailed description of AUC results

- *Rolling average forecast - RAFO:*
 - Timing: Except for longer horizons for the real property prices indicators, the *AUC* computed via RAFO lies consistently above 0.5 for all indicators.
 - Consistency: Except for a minor drop in the very proximity of the onset of the crises for the house prices-to-income indicator, as well as for long time horizons for the wholesale funding ratio, RAFO fulfils this criterion reasonably well. It clearly outperforms the other three trend computation methods in this respect for the real property price indicator. The minor fluctuations as in the credit-to-GDP indicator could well be statistical artifacts.
 - Superiority: RAFO significantly outperforms the other trend computation methods on several occasions and generally performs at least as well as the others.
- *Linear forecast - LFO, and No forecast - NFO:*
 - Timing: These two trend computation methods behave almost identically in terms of their *AUC* performance. Their timing is mostly right and presents the same minor defaults as RAFO for long horizons for the real property prices indicator.

- Consistency: For the real property prices indicator, these trend computation methods present a significant inconsistency in the crisis run-up. Apart from this and a minor inconsistency at the long horizon for the wholesale funding ratio associated with all methods, the two trend computation methods perform well in this respect.
- Superiority: For the credit-to-GDP and the wholesale funding ratio indicators, these trend computation methods perform roughly equally well as RAFO. However, their performance is inferior to RAFO on the two other indicators.
- *Rolling linear forecast - RLFO*: The RLFO trend computation method performs worst according to our three criteria on all four indicators. Except for two short periods for the house prices-to-income and real property prices, it is always dominated by the other indicators and also presents significant inconsistent behaviour across time.

5.2 One-sided and all two-sided HP filters

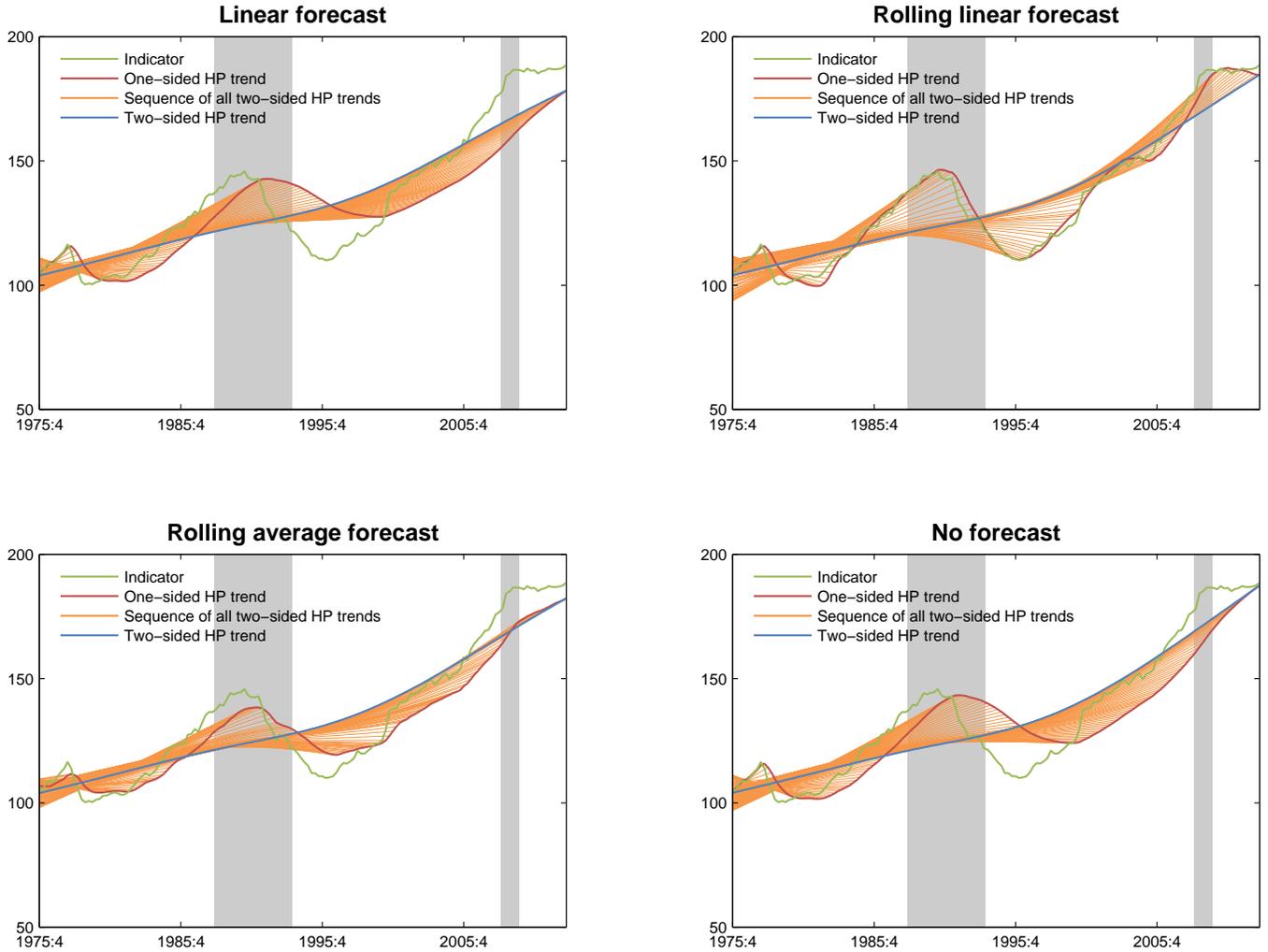
Figure 10. *Credit-to-GDP gap uncertainty*



Note: Credit-to-GDP is the sum of C3 non-financial corporations in mainland Norway and C2 households and measured in per cent of mainland GDP. Gap is measured as deviation of this indicator from trend in percentage points.

Sources: Statistics Norway, IMF and Norges Bank

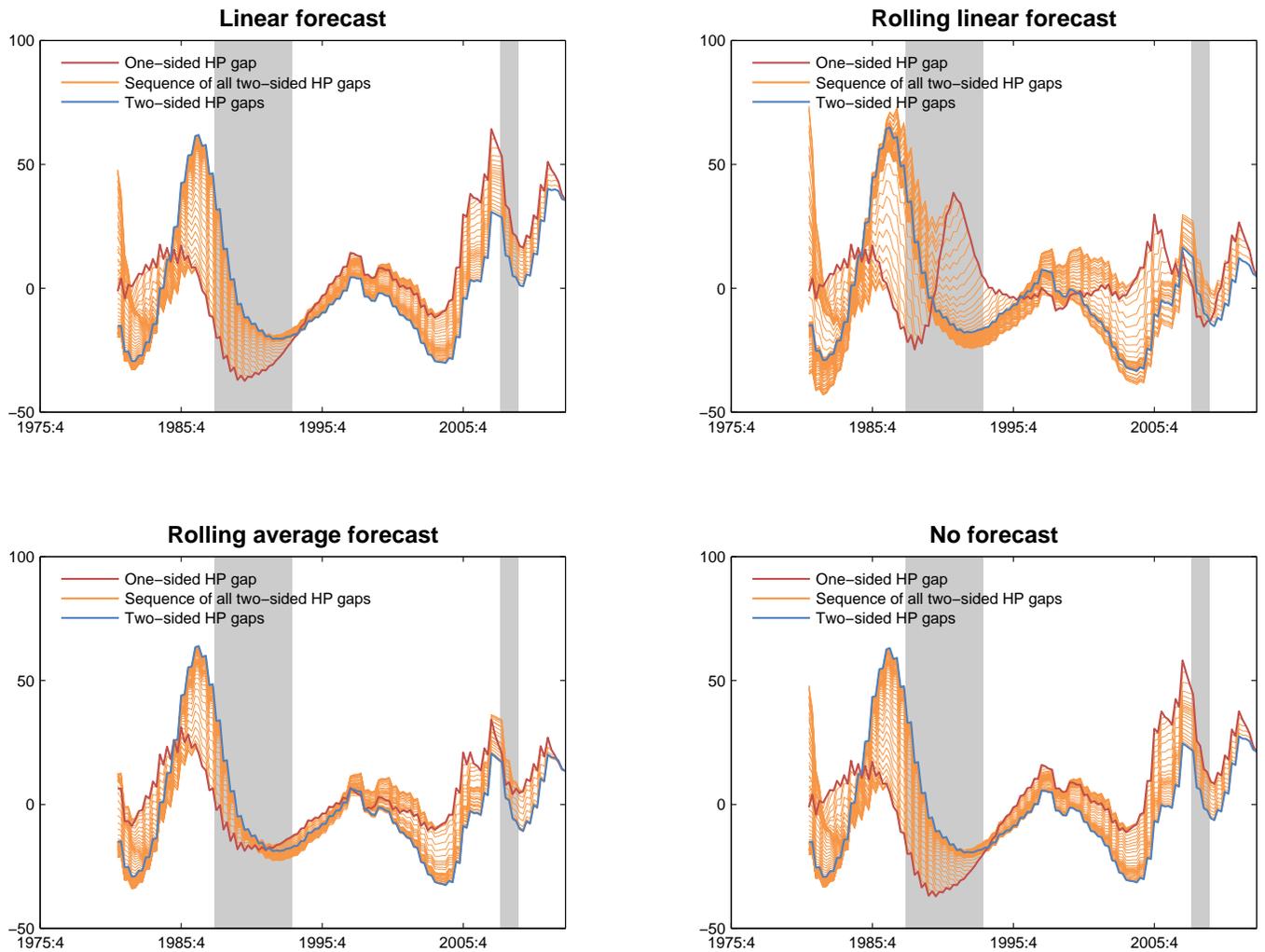
Figure 11. *Credit-to-GDP trend uncertainty*



Note: Credit-to-GDP is the sum of C3 non-financial corporations in mainland Norway and C2 households and measured in per cent of mainland GDP.

Sources: Statistics Norway, IMF and Norges Bank

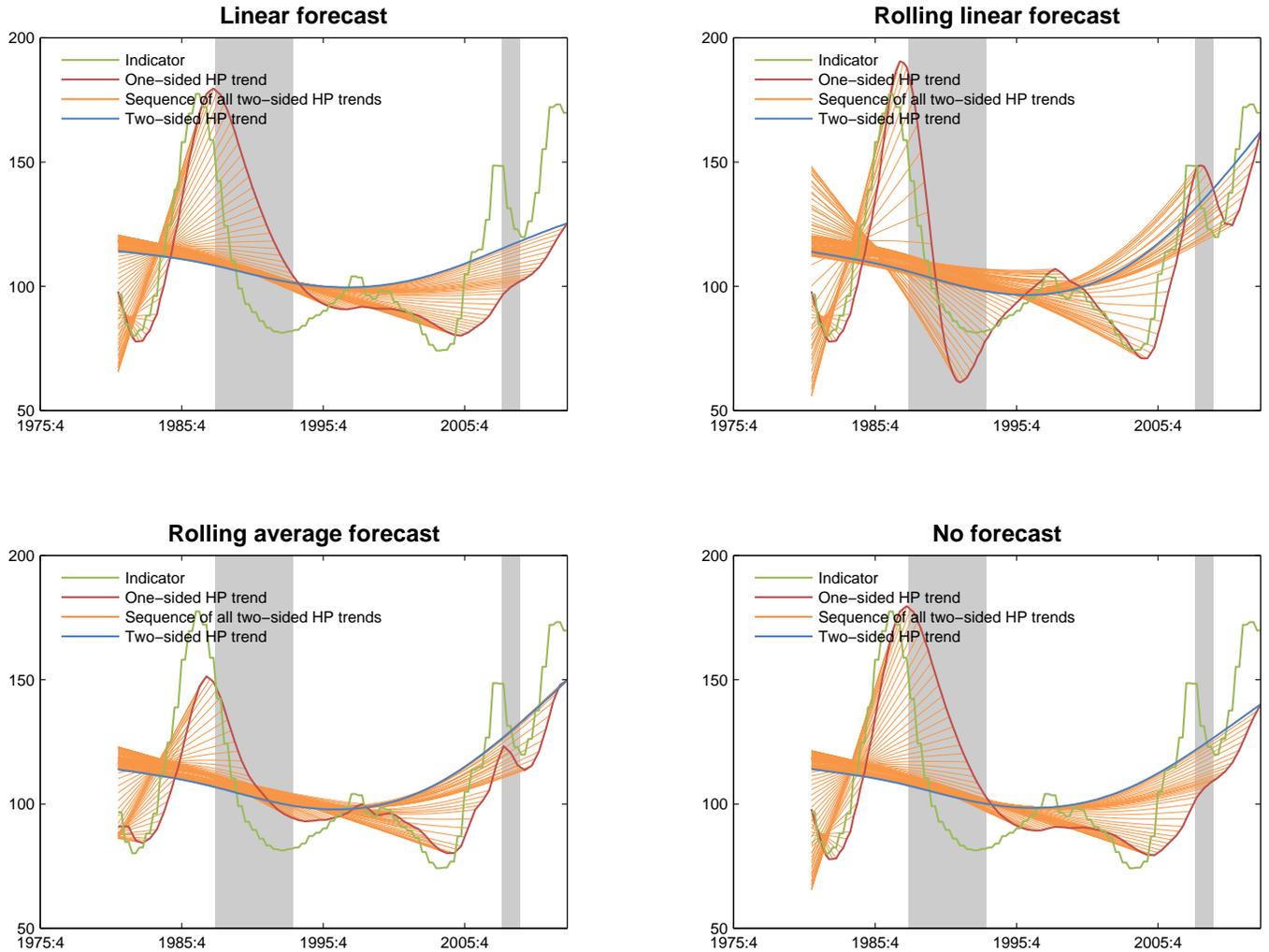
Figure 12. *Real commercial property price gap uncertainty*



Note: Commercial property prices are deflated using mainland GDP price deflator (indexed, 1998=100). The corresponding gap is measured as this indicator in per cent of trend.

Sources: Statistics Norway, IMF, Dagens Næringsliv, OPAK and Norges Bank

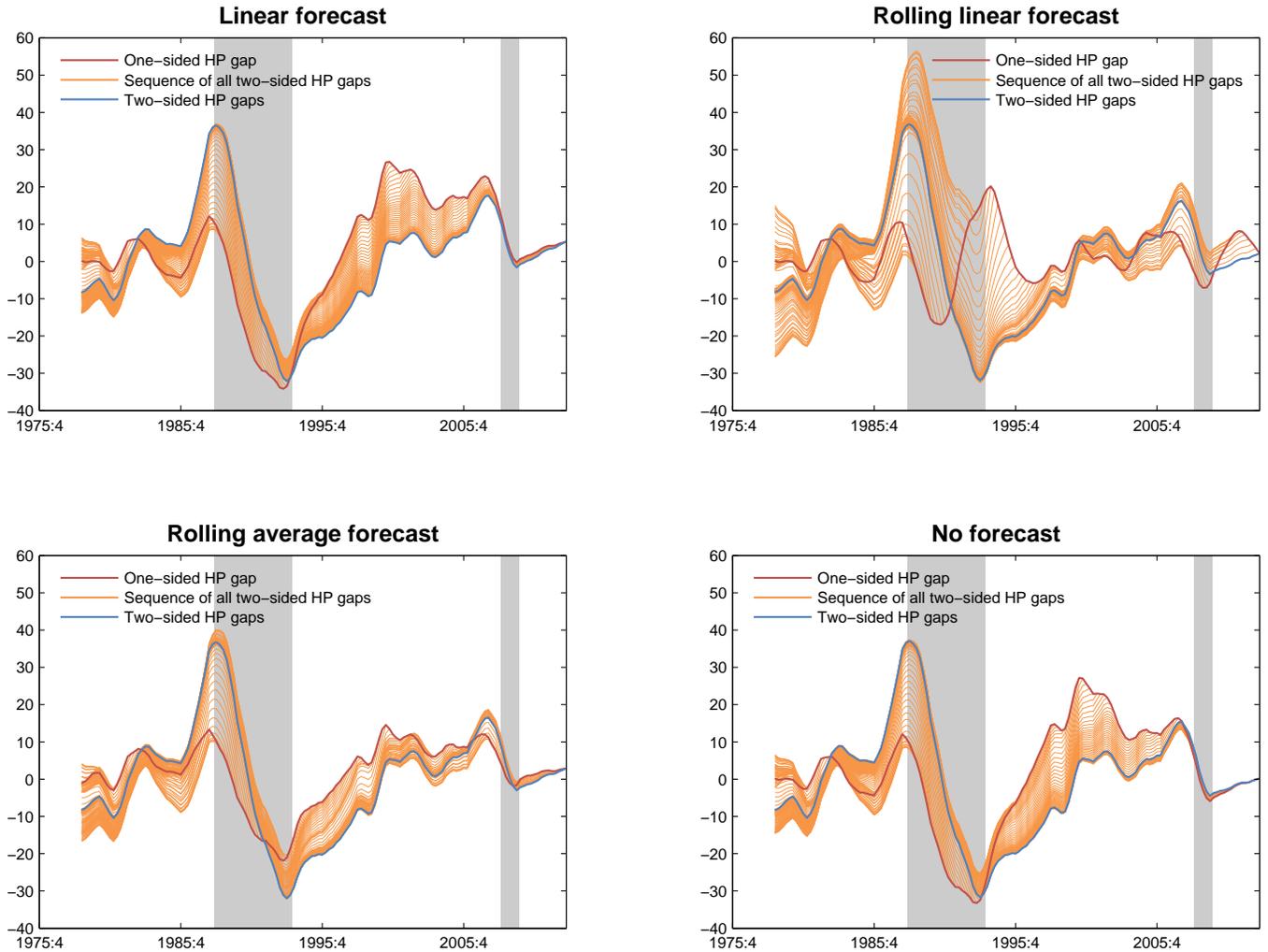
Figure 13. *Real commercial property price trend uncertainty*



Note: Commercial property prices are deflated using mainland GDP price deflator (indexed, 1998=100).

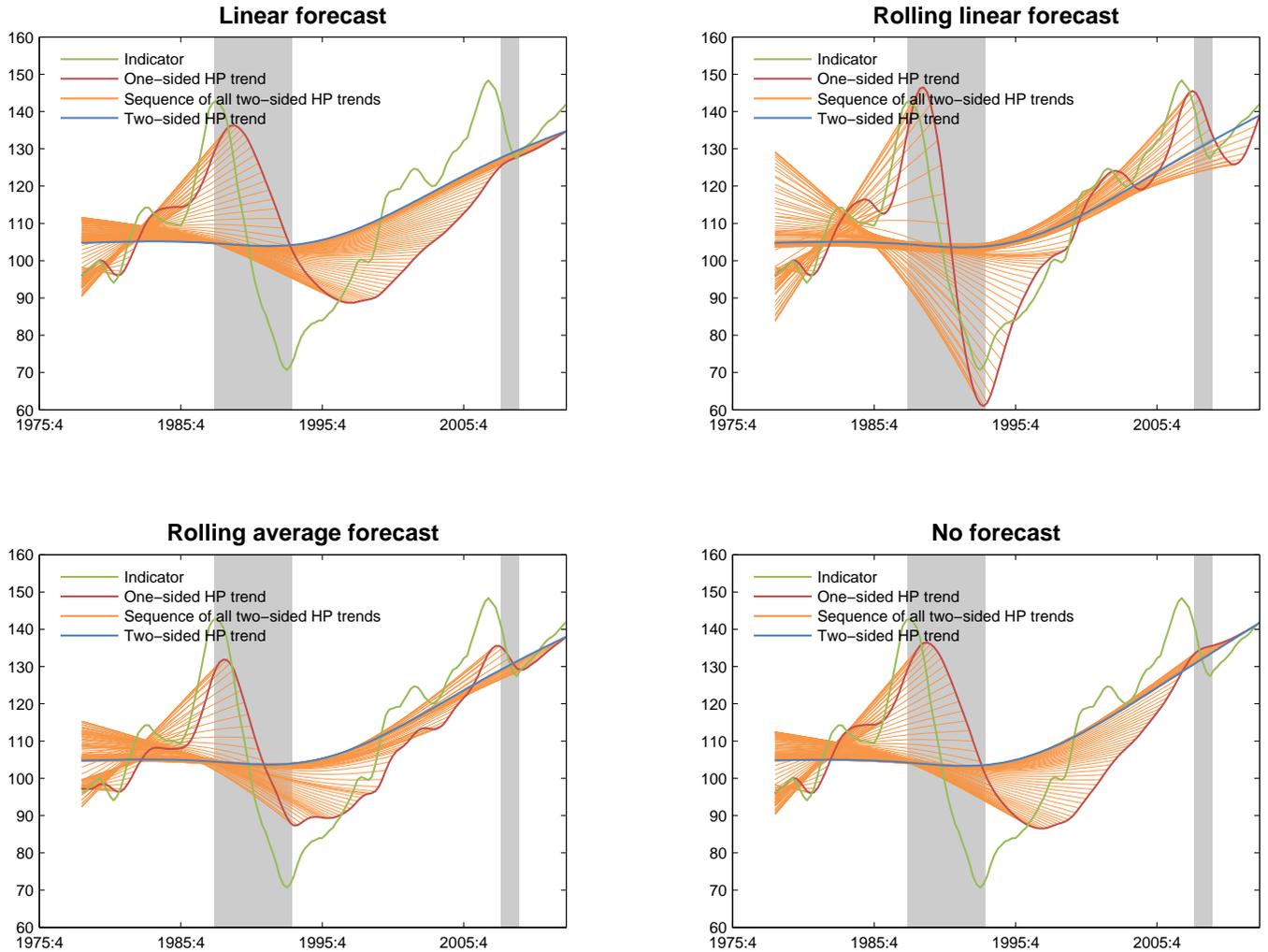
Sources: Statistics Norway, IMF, Dagens Næringsliv, OPAK and Norges Bank

Figure 14. *House prices-to-income gap uncertainty*



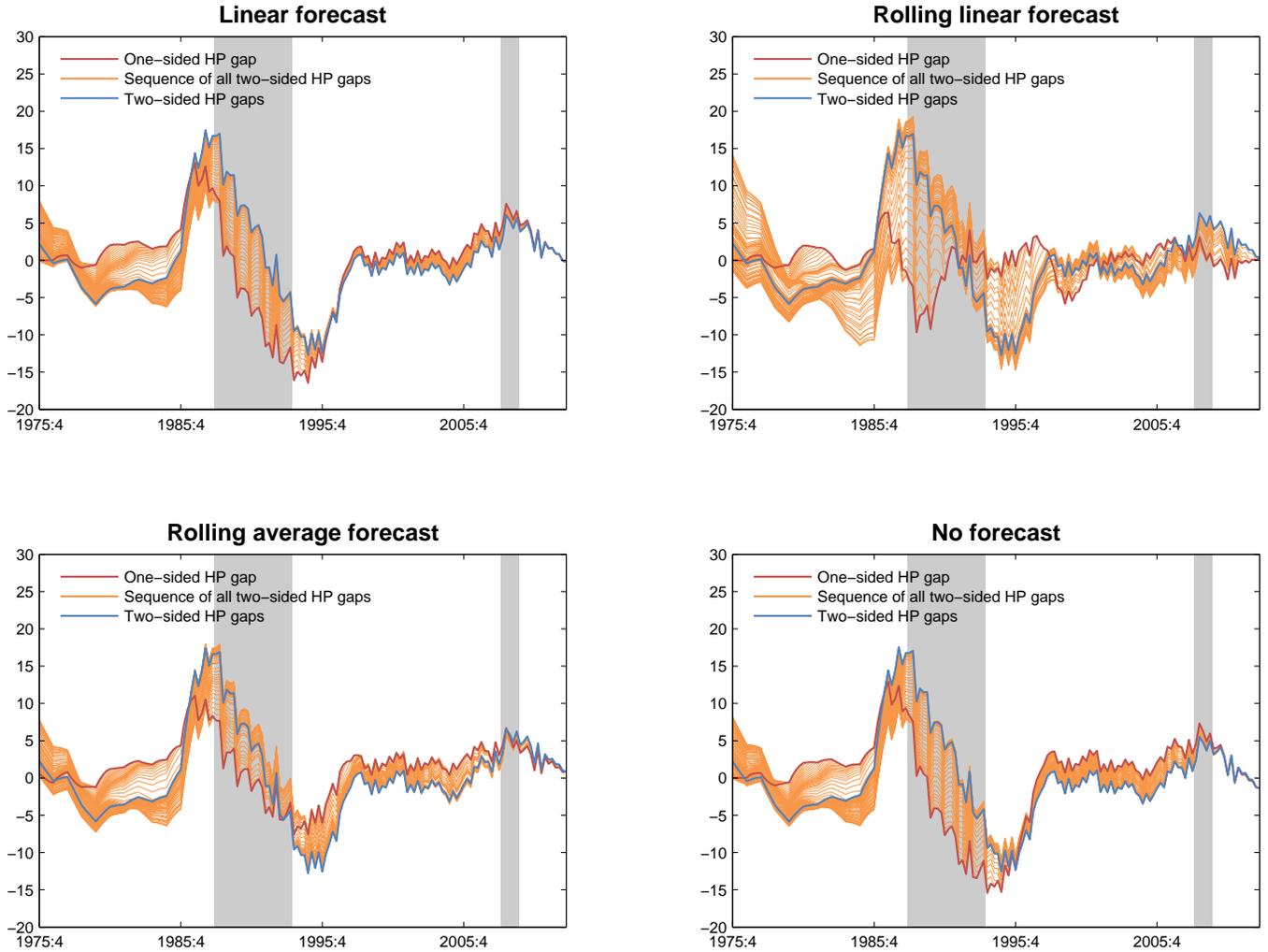
Note: House prices are measured in per cent of disposable income (indexed, 1998Q4=100). The corresponding gap is measured as this indicator in per cent of trend.
Sources: Statistics Norway, Norwegian Association of Real Estate Agents (NEF), Eiendomsmeglerforetakenes Forening (EFF), Finn.no, Eiendomsverdi and Norges Bank

Figure 15. *House prices-to-income trend uncertainty*



Note: House prices are measured in per cent of disposable income (indexed, 1998Q4=100). Sources: Statistics Norway, Norwegian Association of Real Estate Agents (NEF), Eiendomsmeglerforetakenes Forening (EFF), Finn.no, Eiendomsverdi and Norges Bank

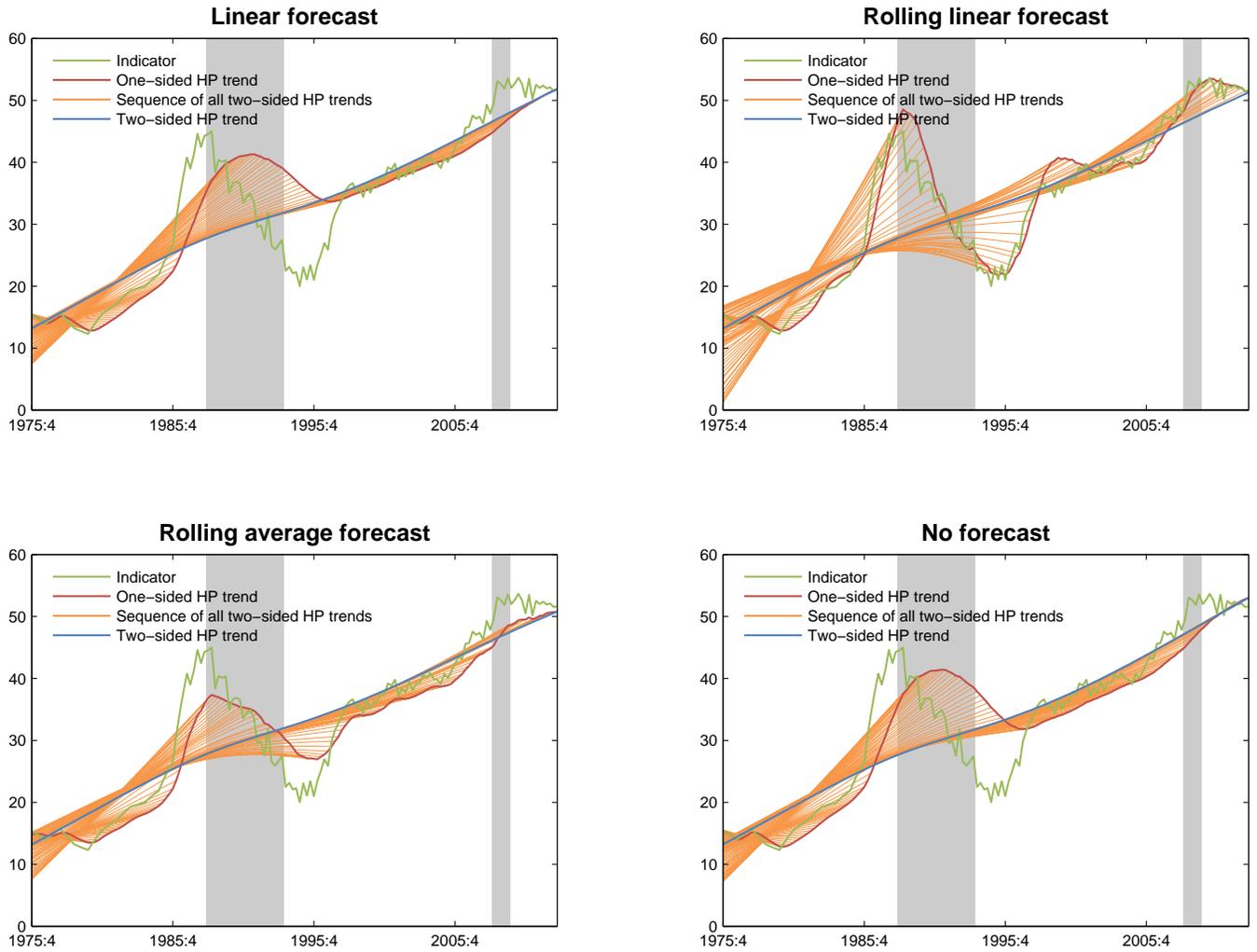
Figure 16. Wholesale funding ratio gap uncertainty



Note: Banks' wholesale funding is measured in per cent of total assets. Gap is measured as deviation of this indicator from trend in percentage points.

Source: Norges Bank

Figure 17. Wholesale funding ratio trend uncertainty



Note: Banks' wholesale funding is measured in per cent of total assets.
 Source: Norges Bank