ANO 2006/2

Oslo

8 March, 2006

Working Paper

Economics Department

Forecasting inflation with an uncertain output gap

by

Hilde C. Bjørnland, Leif Brubakk and Anne Sofie Jore

NORGES BANK

Working papers fra Norges Bank kan bestilles over e-post:
posten@norges-bank.no
eller ved henvendelse til:
Norges Bank, Abonnementsservice
Postboks 1179 Sentrum
0107 Oslo
Telefon 22 31 63 83, Telefaks 22 41 31 05

Fra 1999 og senere er publikasjonene tilgjengelige som pdf-filer
på www.norges-bank.no, under "Publikasjoner".

Working papers inneholder forskningsarbeider
og utredninger som vanligvis
ikke har fått sin endelige form.
Hensikten er blant annet at forfatteren
kan motta kommentarer fra kolleger
og andre interesserte.

Synspunkter og konklusjoner i arbeidene
står for forfatternes regning.

Working papers from Norges Bank can be ordered by e-mail:
posten@norges-bank.no
or from Norges Bank, Subscription service,
P.O.Box. 1179 Sentrum
N-0107Oslo, Norway.
Tel. +47 22 31 63 83, Fax. +47 22 41 31 05

Working papers from 1999 onwards are available as pdf-files on the bank's
web site: www.norges-bank.no, under "Publications".

Norges Bank's working papers present
research projects and reports
(not usually in their final form)
and are intended inter alia to enable
the author to benefit from the comments
of colleagues and other interested parties.

Views and conclusions expressed in working papers are
the responsibility of the authors alone.

ISSN 0801-2504 (printed) 1502-8143 (online)

ISBN 82-7553-346-5 (printed), 82-7553-347-3 (online)

Forthcoming in Empirical Economics

Forecasting inflation with an uncertain output gap*

Hilde C. Bjørnland[a,b], Leif Brubakk[b] and Anne Sofie Jore[b]

21 August 2007

Abstract

The output gap is a crucial concept in the monetary policy framework, indicating demand pressure that generates inflation. However, its definition and estimation raise a number of theoretical and empirical questions. This paper evaluates a series of univariate and multivariate methods for extracting the output gap in Norway, and compares their value added in predicting inflation. We find that models including the output gap have better predictive power than models based on alternative indicators, and they forecast significantly better than simple benchmark models. At the longer forecast horizons, multivariate measures of the output gap perform better than the univariate gaps.

Keywords: Output gap, forecast, Phillips curve, forecast combination.

JEL-codes: C32, E31, E32, E37

* Comments from two anonymous referees, Q. Farooq Akram, Tommy Sveen, Ken West, Fredrik Wulfsberg and seminar participants in Norges Bank are gratefully acknowledged. All mistakes remain our own. The views expressed are those of the authors and do not necessarily represent those of Norges Bank

[a] *Corresponding author*. Norwegian School of Management (BI). E-mail: hilde.c.bjornland@bi.no
[b] Norges Bank

1 Introduction

The output gap - measuring the deviation of output from its potential - is a crucial concept in the monetary policy framework, indicating demand pressure that generates inflation. Because the output gap will have an effect on inflation, an optimal inflation-targeting policy implies a monetary policy response to the output gap. Such a policy response will help stabilize inflation as well as output, as pointed out by Svensson (1997, 2000) and Rotemberg and Woodford (1997). Many central banks that have announced inflation-targeting policies, therefore attempt to stabilize both inflation and the output gap.

The output gap is also an important variable in itself, as a measure of economic fluctuations. Over time, economic resources are utilized efficiently when economic growth is stable and the output gap remains close to zero (or output close to potential output). At this level, employment growth and unemployment will also be stable.

Despite the output gap's central role in monetary policy making, its definition and estimation raise a number of theoretical and empirical questions. Ever since Nelson and Plosser (1982) failed to reject the hypothesis of a unit root in macroeconomic time series, the long run trend in output can no longer be treated as deterministic. Given the uncertainties associated with the estimation of a stochastic trend, measuring potential output (and the output gap) with any degree of accuracy has proved to be difficult.

The uncertainties surrounding the measurement of potential output and the output gap has also direct and strong implications on optimal monetary policy, as pointed out by Rudebusch (2002), Smets (2002) and Ehrmann and Smets (2003). In particular, they show that the optimal weight to place on output stabilisation for the monetary policymaker declines when the output gap is poorly measured. In addition, there is also added uncertainty from the fact that real-time data on output are preliminary and subjected to substantial subsequent revisions, as emphasized by Orphanides (2001) for U.S. data. The mismeasurement of the output gap in real time represents a major problem for the implementation of policy strategies that rely on information about the current output gap, as pointed out by Orphanides and van Norden (2002) and Orphanides (2003).

A key aspect in all of these investigations is the recognition that policymakers may be uncertain regarding the true data-generating processes describing the output gap and the extent of the mismeasurement problem. As a result, standard applications of certainty equivalence based on the classical linear-quadratic-Gaussian control problem do not apply.[1] Hence, simple monetary policy rules based on the output gap may not be robust to output gap uncertainty.

[1] See Svensson and Woodford (2003) for a recent exposition of certainty equivalence in the absence of any model uncertainty.

There have been a variety of suggestions in the literature on how to mitigate the problem of output gap mismeasurement for monetary policy decisions, by placing less weight on the "uncertain" output gap, replacing the gap with the change in output, ignoring the gap fully by relying exclusively on past and future inflation rates, or aiming directly at stabilizing the nominal income growth, see for instance McCallum (1998, 2001), Orphanides et al. (2000), Rudebusch (2002), Leitemo and Lønning (2006) and Spencer (2004) among many others.

Although the mismeasurement of the output gap based on an inappropriate detrending method is a general problem (see for instance Canova, 1998; Bjørnland, 2000), the mismeasurement of the output gap due to data revisions and lack of hindsight may not necessarily be so. In particular, Gruen et al. (2005) find real-time output gap estimates for Australia which are unbiased and highly correlated with final estimates derived with the latest data and the benefit of hindsight. Similar results are also found in Rünstler (2002) for the Euro area and to a certain degree in Bernhardsen et al. (2004) for Norway when they estimate the real-time output gap using multivariate models.

This paper sets out to evaluate a series of methods for extracting the output gap using Norwegian quarterly data. The different methods range from simple univariate detrending methods to more elaborate multivariate models. Given the uncertainties of real time estimates, in particular for the univariate detrending methods, we argue that as a minimum criteria the output gaps should display a high degree of coherence with other indicators of economic activity that are not (or less) revised in real time. However, as optimal monetary policy is essentially about forecasting inflation (see Svensson and Woodford, 2005), the usefulness of the output gap should ultimately be addressed in terms of its value added in forecasting inflation. In the main analysis, we will use the Phillips curve, which relates inflation to real activity, as the maintained theory of inflation. As Gerlach and Svensson (2003), we will attribute greater importance to the output gap if it is a good predictor of future inflation. However, a general impression from the literature is that there does not seem to be one indicator or variable that is superior in forecasting inflation (see the discussion in Clark and McCracken, 2006, and the references therein). In order to obtain more robust inflation forecasts, we therefore also consider averaging of forecasts, using equal weight averaging and Bayesian model averaging.

The paper is organized as follows. In Section 2, the different methods are put forward and applied to the Norwegian data. Section 3 evaluates the alternative output gaps in terms of statistical properties and coherence with alternative measures of the business cycle less subject to data revisions. The different output gaps (as well as the alternative measures of the business cycle) are finally evaluated in Section 4 by their value added in predicting inflation, using Phillips curve type inflation equations. Section 5 presents our conclusions.

2 Methods for estimating the output gap

An obvious question when a time series is characterised with a unit root, is how one can distinguish the permanent (trend) component from the transitory (cyclical) component in the data. In particular, the issue of detrending becomes non-trivial when the trend can no longer be treated as deterministic. However, Beveridge and Nelson (1981) have shown that any non-stationary process can in fact be decomposed into a permanent and a transitory component, with plausible statistical properties. The issue to consider is what kind of structural relationship and driving forces one should assume for the different components, as different assumptions may produce different values in the trend-cycle decomposition. Furthermore, historical estimates of the output gap might also change when data are revised and new information emerges. The problem of data revisions applies to both actual and potential output, implying uncertainty concerning both components. In the following, we refer to the output gap as

$$ygap_t = y_t - y^*_t \tag{1}$$

The variables are expressed in logarithms, with the output gap, $ygap_t$, being the percentage deviation between actual output (y_t) and potential output (y^*_t). Below we review and apply some univariate and multivariate methods for estimating the output gap in Norway.

2.1 Univariate methods

Univariate methods use information in the time series itself (here, mainland GDP) to estimate the output gap.[2] Three examples will be reviewed here.

Hodrick-Prescott filter (HP)

The Hodrick-Prescott filter extracts the value of potential output that minimises the difference between actual output and potential output, while imposing constraints on the extent to which growth in potential output can vary, see Kydland and Prescott (1990) for details. A smoothing parameter (λ), that takes values between zero and infinity, determines the extent of permissible variations in potential growth. λ is determined outside the model. Here we follow international practice and specify $\lambda = 1600$ (see Kydland and Prescott 1990).

Band-Pass filter (BP)

The basic idea behind band pass filtering is to extract information regarding the frequencies of interest.[3] For the purpose of measuring the cyclical component of GDP, this would typically be the business cycle frequencies. Following Burns and Mitchell (1946), we define the business cycles as fluctuations lasting from 6 to 32 quarters. Fluctuations with a higher frequency are considered as irregular or seasonal, whereas fluctuations with a lower frequency are attributed to movements in the trend or potential GDP. To approximate an

[2] We have used seasonally-adjusted GDP figures for the period 1978Q1 to 2004Q2.
[3] See e.g. Hamilton (1994) for an introduction to frequency domain analysis.

optimal filter (that requires an infinite number of data points), we use the Band pass filter developed by Baxter and King (1999)[4].

Univariate unobserved component (UC)

The unobserved components method assumes a relationship between an observed variable and certain unobserved components such as the output gap. This requires a specification of the time series process underlying the unobservable variable. Both the unobservable and observed variables are then modelled and estimated with "maximum likelihood" using the Kalman filter. Here we follow Harvey (1985) and Clark (1987), and specify a simple UC model as a local linear trend model. That is, in addition to the postulated relationship in (1), we assume that potential output follows a random walk with stochastic drift, driven by random and normally distributed residuals that are independent of each other. This specification places few constraints on permitted variations in the unobservable potential output. The output gap is assumed to follow an autoregressive AR(2) process.

2.2 Multivariate methods

Multivariate models explore the relationships between GDP and other observable variables. Three different methods are presented here.

Production function (PF)

The production function models the supply side of the economy, where output is determined by available technology, and the input factors labour and capital. Potential output refers to the level of output consistent with input factors at their potential levels. The difference between actual and potential output is interpreted as the output gap. Here we assume that the aggregated production function for the economy can be expressed as a standard Cobb-Douglas production function.[5] Total factor productivity is calculated as the residuals from this equation using the least-squares method. The potential levels of labour, capital and total factor productivity are then used to estimate potential output. We assume that potential use of labour depends on the labour force, working hours per employee and equilibrium unemployment. Potential capital stock is assumed to be equal to actual capital stock.[6]

Multivariate unobserved component (MVUC)

The univariate unobserved components model can be expanded by including a number of variables that are assumed to contain information about the output gap. For instance, Scott

---

[4] A problem with this filter is that we will loose 12 observations at the start and end of the sample. Here, we follow Stock and Watson (1998) and extend the output series with forecasts from an AR(4) model. Alternatively we could have used the one sided filter in Christiano and Fitzgerald (1999).

[5] We follow the approach described in Frøyland and Nymoen (2000) and estimate a production function for the sectors manufacturing, construction, services and distributive trades. These sectors account for about ¾ of output in mainland Norway.

[6] The values for the factor income shares are set to 2/3 for labour and 1/3 for capital, see the Ministry of Finance (1997). Equilibrium unemployment and the potential levels of total factor productivity, the labour force and working hours are calculated using the HP filter. However, allowing for a reasonable range of variation for λ, potential output is not affected to any substantial extent.

(2000) extends the univariate model with an equation linking inflation to the output gap and by adding capacity utilisation as an observable. The relationship between the unemployment rate and the output gap given by the Okun's law are typically explored, see Okun (1962).

In the present study, we build on among others Apel and Jansson (1999), and propose a model with output, inflation and the unemployment rate as observables:

Observation equations:

$$\Delta y_t = \Delta y_t^* + ygap_t - ygap_{t-1} \tag{2}$$

$$\pi_t = \alpha_{11}\pi_{t-1} + \alpha_{12}\pi_{t-2} + \beta_{11} ygap_{t-1} + \varepsilon_{2,t} \tag{3}$$

$$u_t - u_t^* = \alpha_{21}(u_{t-1} - u_{t-1}^*) + \beta_{21} ygap_{t-1} + \varepsilon_{3,t} \tag{4}$$

State equations:

$$ygap_t = \psi_{11} ygap_{t-1} + \psi_{12} ygap_{t-2} + \upsilon_{1,t} \tag{5}$$

$$\Delta y_t^* = \Delta y_{t-1}^* + \mu_{t-1} + \upsilon_{2,t} \tag{6}$$

$$\mu_t = \mu_{t-1} + \upsilon_{3,t} \tag{7}$$

$$u_t^* = u_{t-1}^* + \gamma_{t-1} + \upsilon_{4,t} \tag{8}$$

$$\gamma_t = \gamma_{t-1} + \upsilon_{5,t} \tag{9}$$

where (2) is an identity which simply states that the growth rate of output is equal to the growth in potential output plus the change in the output gap. Equation (3) can be interpreted as a Philips curve, linking domestic inflation, π_t, to the output gap. A version of Okun's law is given in (4), where u_t denotes the unemployment rate and u_t^* refers to the NAIRU, which is assumed to be a latent variable. We assume that the output gap can be represented by an AR(2) process, given in (5). Equation (6) specifies the growth in potential output as a random walk with a stochastic drift, μ_t, given by (7). This is a rather flexible specification that allows for mean shifts in the growth rate of potential output. The process for the NAIRU is determined by equations (8) and (9). We assume that all the error terms are iid and normally distributed. Using matrix notation, the model can be written in state space form. The model is estimated with Maximum Likelihood using the Kalman filter.

We use quarterly data for the period 1981q3 to 2004q2 to estimate the model. The output data refers to GDP for mainland Norway, which excludes the oil sector. Unemployment data are taken from the quarterly labour force survey (LFS). Domestic inflation[7] is CPI-ATE inflation excluding imported goods[8]. Table A.1 in appendix A reports estimation results. All

[7] The inflation series was detrended prior to estimation, by using an HP filter with λ equal to 40000, in order to make it stationary.
[8] CPI-ATE is the consumer price index (CPI) adjusted for tax changes and excluding energy products, by delivery sector (published by Statistics Norway). To construct a measure of *domestic* inflation, the

parameters have the expected signs. Furthermore, with the exception of some of the estimated standard deviations of the error terms, all parameters are significantly different from zero at the 5% level.

Structural vector autoregression (SVAR)

The SVAR method is an alternative way of using information inherent in a number of variables to estimate the output gap. Identification is based on Blanchard and Quah (1989), which showed how one can impose long run restrictions in a bivariate model in output and unemployment, to identify permanent and transitory components of output.

In the following we augment the bivariate model of Blanchard and Quah, to also include domestic inflation. This allows us to identify three different shocks: two demand shocks (nominal and real demand) and one supply shock. We assume that neither of the demand shocks can have a long run effect on unemployment. However, to distinguish between the two demand shocks, we assume that only the nominal demand shock is restricted from affecting output in the long run. This allows us to investigate the possibility that one of the demand shocks (real demand) can have a more persistent effect on output than the other, although without changing the unemployment rate permanently as a result. These assumptions may allow for the interpretation of the real demand shock as a preference shock and the nominal demand shock as a monetary policy shock; see Gali and Rabanal (2004) for further discussion.[9] Finally, the aggregate supply shock is allowed to have a long-term effect on output and unemployment. Since the unemployment rate has increased in the course of our estimation period and is perceived to be nonstationary, it is reasonable to assume that the supply shock can affect equilibrium unemployment over time. Note that as inflation is perceived to be stationary, none of the shocks can affect inflation permanently.

Let z_t be a vector with the three stationary variables $z_t = (\Delta u_t, \Delta y_t, \Delta p_t)'$ where Δ denotes quarterly changes, u_t is the unemployment rate, y_t is GDP and π_t ($=\Delta p_t$) is domestic inflation. The moving average representation containing the vector of original structural disturbances (ε_t) can be found as $z_t = B(L)\varepsilon_t$. Let the ε_t's be normalized so they all have unit variance. From this, the matrix of long run multipliers can be written as

prices of goods that are predominantly imported (cars, clothes etc.) are removed from CPI-ATE. This leaves approximately 70 % of the prices that are used to construct CPI-ATE.

[9] It may also be that real demand shocks like government consumption/investment can change potential output, due to changes in capital accumulation. This effect may, however, be expected to be small, since capital accumulation is slow, and with little consequences for long run unemployment.

$$\begin{bmatrix} \Delta u \\ \Delta y \\ \Delta p \end{bmatrix}_t = \begin{bmatrix} B_{11}(1) & B_{12}(1) & B_{13}(1) \\ B_{21}(1) & B_{22}(1) & B_{23}(1) \\ B_{31}(1) & B_{32}(1) & B_{33}(1) \end{bmatrix} \begin{bmatrix} \varepsilon^{AS} \\ \varepsilon^{RD} \\ \varepsilon^{ND} \end{bmatrix}_t \tag{10}$$

where ε_t^{AS} is the aggregate supply shock, ε_t^{RD} is the real demand shock, ε_t^{ND} is the nominal demand (i.e. monetary policy) shock and $B(1) = \sum_{j=0}^{\infty} B_j$ indicate the long run matrix of B(L). The restriction that none of the demand shocks can affect the unemployment rate permanently implies that $B_{12}(1)=B_{13}(1)=0$. Furthermore, the restriction that nominal demand shocks can not affect GDP permanently entails that $B_{23}(1)=0$.

Based on the above identification, GDP can now potentially be split into two different components; a component determined by shocks that have a permanent effect on the supply side of the economy, and a component determined by shocks that only affect demand in the short term. The first component represents potential GDP and will consist of the accumulated supply shocks, while the latter can be interpreted as the output gap and will consist of the accumulated nominal demand shocks. For the third shock, the real demand shock that can potentially affect output in the long run, we assume that it contributes to the output gap the first two years (business cycle frequencies), whereas any remaining effect will contribute to developments in potential output. To find the short run contribution, we calculate the eight quarter forecast error of output that is due to the real demand shock. By focusing on the eight quarter forecast error, we emphasise the contribution to the business cycle frequencies. Hence, the output gap consists of accumulated nominal demand shocks and the eight quarter forecast error of output due to real demand shocks. However, as it turns out, the real demand shocks account for relatively little of the variation in the output gap.[10]

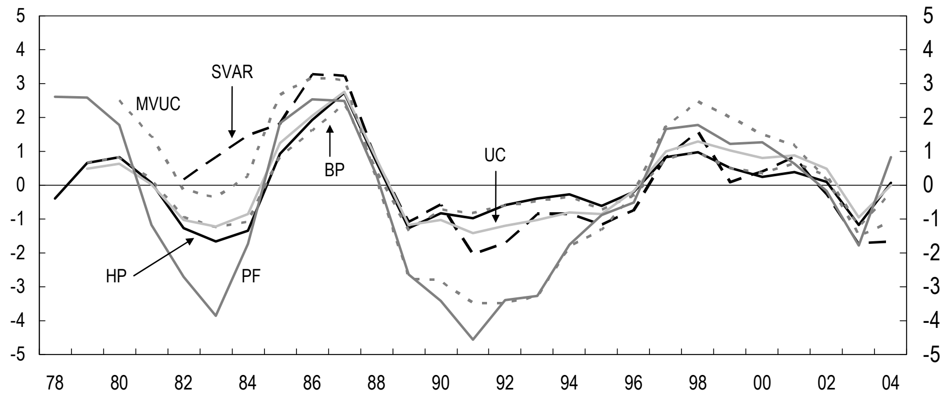
We use quarterly data for GDP, unemployment and domestic inflation over the period 1981q1 to 2004q2 to estimate the model (see the MVUC method for data descriptions). However, some initial values are lost due to the aggregation of shocks, so that the output gap will be available from 1982q4. Based on a set of information criteria, the VAR model is estimated with 5 lags. With 5 lags, the model satisfies a series of goodness-of-fit properties. The impulse responses seem consistent with theory predictions (and can be obtained from the authors on request). In particular, it turns out that the effect of the nominal and the real demand shocks on GDP will eventually die out, although the real demand shocks at a slower pace than the nominal demand shocks.

[10] Appendix A compares the output gap using our preferred SVAR model to a bivariate SVAR model in output and inflation; identified by assuming that only one shock (aggregate demand) has no long run effect on output (as in Blanchard and Quah, 1989). The chart shows that the output gaps move closely together over the sample, although in some periods (in particular at the end of the sample) there are some observed differences.

3 Comparison of output gaps

This section presents a set of statistics that illustrate the properties of the different output gaps. Chart 1, below, shows the various output gaps over time. All calculations are based on quarterly data. However, for ease of illustration, we plot annual figures. The different output gaps describe main economic fluctuations in Norway as they are commonly referred to, with two downturns in the 1980s, an upturn from the mid-1990s and a downturn again from the end-1990s. The PF method differs from the other methods in estimating a considerably more negative output gap during the downturn in the early 1980s. Both the MVUC and PF method also estimate a more severe downturn at the beginning of the 1990s than the other methods. From the mid 1990s, the output gaps seem to "co-move" as they are in more agreement regarding the state of the cycle. This is not surprising, as the period from 1993 (when the exchange rate was floated) seems to be relative stable with few structural breaks, and with monetary policy being broadly consistent with a Taylor rule (see Olsen et al. 2003). Hence, there might be less divergence as to how the different methods separate the trend from the cycle. Note, however, that for the two-sided HP and BP filters, the estimate for the output gap will be particularly uncertain towards the end of the sample.

Chart 1 Output gaps[1]. All methods. Percentage of potential GDP.



[1] The output gaps are calculated using quarterly data. For ease of exposition, annual aggregates are shown in the chart.

Tables 1 to 3 contain statistical summaries of the different output gaps for the period 1982 to 2004. Table 1 first compares some key properties of the gaps. One reasonable criterion is that the average value of the output gap should be close to zero over time. This seems to be the case for all the output gaps except the PF gap, that has an average value of -0.7. The PF output gap also displays the highest standard deviation, closely followed by the MVUC output gap (2.55 and 2.43 respectively). At the other end, the band pass output gap has a standard deviation of 1.22. A general observation is that the univariate gaps move closer to zero and have smaller standard deviations than the multivariate gaps. However, we have no objective criteria to determine whether an output gap "behaves reasonably", other than indicating that the output gaps should not be "too wide" or "too narrow".

Table 1 Statistical summary for the output gap, 1982:4 to 2004:2

Method	HP	BP	UC	PF	MVUC	SVAR
Average	-0.02	-0.03	0.06	-0.69	-0.23	0.14
Standard deviation	1.31	1.22	1.40	2.55	2.43	1.58
Lowest value	-2.87	-2.58	-2.35	-5.69	-4.63	-2.84
Highest value	3.82	3.47	3.81	5.14	4.68	4.01

Table 2 Correlation between output gaps calculated by different methods, 1982:4 to 2004:2

Method	HP	BP	UC	PF	MVUC	SVAR
HP	1.00	0.99	0.95	0.67	0.80	0.71
BP		1.00	0.96	0.66	0.80	0.72
UC			1.00	0.77	0.92	0.78
PF				1.00	0.83	0.66
MVUC					1.00	0.77
SVAR						1.00

Table 3 Concordance in business cycles, 1982:4 to 2004:2

Method	HP	BP	UC	PF	MVUC	SVAR
HP	1.00	0.95	0.91	0.86	0.82	0.74
BP		1.00	0.89	0.89	0.84	0.78
UC			1.00	0.89	0.86	0.78
PF				1.00	0.86	0.80
MVUC					1.00	0.83
SVAR						1.00

Table 2 shows the correlation coefficients between the different methods. As expected from looking at the charts, the correlation between the alternative output gaps is generally high, particularly between the univariate methods. The correlation coefficients are lowest between the PF and either the SVAR, BP or HP method.

Table 3 contains a measure of concordance in business cycles, i.e. the proportion of time that the cycles of two series spend in the same phase, see McDermott and Scott (2000). This is of particular interest in analyses where the focus is on the sign of the gap and not necessarily its magnitude. Table 3 confirms the impression from the charts and Table 2 that the alternative methods provide close descriptions of cyclical developments.

It is also interesting to investigate whether the different output gaps yield the same conclusion as to the dates for the different turning points in the cycle. Table 4 shows the various turning points suggested by the different gaps. We define a peak/trough as the quarter the output gap reaches its highest/lowest value within a period generally regarded as an upturn/downturn.

Table 4 Turning points

Method Period	HP	BP	UC	PF	MVUC	SVAR
Upturn mid-1980s	1987q2	1987q2	1987q2	1987q3	1987q2	1987q1
Downturn early 1990s	1989q3	1989q3	1990q4	1991q3	1991q4	1991q4
Upturn late 1990s	1997q4	1997q4	1997q4	1997q1	1997q4	1998q4
Downturn early 2000s	2003q1	2003q1	2003q1	2003q2	2003q2	2003q1

The different output gaps are in relative agreement in suggesting that the upturn in the mid-1980s peaked in the first part of 1987.[11] This is in line with the general perception of the business cycle (see for example Bjørnland (2000) and Johansen and Eika (2000)). However, the output gaps pinpoint different dates for the trough in the early 1990s. The HP and BP gaps date the turning point as early as 1989q3, while the MVUC and SVAR gaps indicate 1991q4. However, this period of low growth lasted for around 8 years. Pinpointing the trough may therefore be subject to coincidental quarterly variations. In the subsequent upturn, most output gaps indicate a peak in 1997q4, while the PF gap finds the top to be three quarters earlier and the SVAR method one year later. Finally, all output gaps concur in that the final downturn ended in the first half of 2003.

Summing up, all the output gaps co-move over the cycle, displaying relatively high correlation coefficients. However, the amplitude varies substantially between the different output gap measures, with the univariate gaps displaying volatility in the lower end. The dating of the turning points are also in general agreement between the models, except for the turning points associated with the prolonged recovery in the 1990s.

3.1 Alternative indicators

Most indicators of economic activity like GDP and its components are revised over time, sometimes substantially. Given the uncertainties of real time estimates, in particular for the univariate detrending methods, we argue that as a minimum criteria the output gaps should display a high degree of coherence with indicators of economic activity that are not revised in real time, or at least subject to only minor revisions.

The Industrial Confidence Index (ICI) published by Statistics Norway is such a variable. While this indicator is not revised, except for revisions due to changes in seasonal factors, it only covers manufacturing industry. Nevertheless, it may be a good indicator of business cycle conditions.

[11] We have not included the trough in the early 1980s since calculations of the output gap using the SVAR method starts in 1982.

The unemployment rate is an alternative indicator of economic activity not affected by revisions[12]. However, as the unemployment rate has increased over time, we need to measure the unemployment rate as a deviation from a natural rate ("trend"), i.e. the unemployment gap (UGAP). This involves the issue of de-trending again. As it turns out, the unemployment rate only changes gradually and very smoothly, implying that the different methods provide very similar pattern for the unemployment gap. For simplicity, the UGAP is therefore calculated by smoothing the unemployment rate (taken from the labour force survey (LFS)) by a Hodrick Prescott filter with λ=40000. The series is identical to the unemployment gap used in the PF method.

In tables 5 and 6 we show correlations and concordance between the output gaps and the ICI and the UGAP[13], respectively, for the period 1988:1-2004:2. We have chosen to start in 1988 here and in the subsequent analysis, as this is the first observation available for ICI.

Table 5 Correlation between output gaps and different indicators, 1988:1 to 2004:2

Method	ICI	UGAP
HP	0.28	0.65
BP	0.27	0.65
UC	0.29	0.77
PF	0.37	0.69
MVUC	0.39	0.75
SVAR	0.23	0.71

Table 6 Concordance between output gaps and different indicators, 1988:1 to 2004:2

Method	ICI	UGAP
HP	0.56	0.82
BP	0.56	0.82
UC	0.55	0.88
PF	0.55	0.86
MVUC	0.58	0.94
SVAR	0.55	0.86

Correlations between ICI and the output gaps are low, ranging from 0.23 to 0.39. This may be due to the nature of this indicator, which reflects only one sector of the economy. Another explanation is the much larger and more irregular fluctuations in the ICI compared to fluctuations in the output gaps. Concordance is less affected by irregular fluctuations from one quarter to the next; hence concordances between ICI and the output gaps in table 6

[12] From time to time, the calculation method has changed. This has not altered the general development in the series.
[13] For ease of exposition, we multiplied UGAP by (-1) before calculating correlation and concordance.

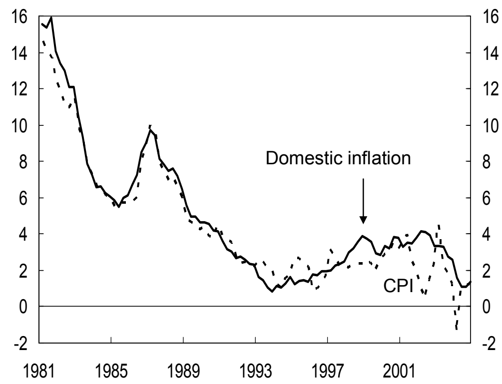
indicate a closer relationship than the correlations do. Concordances with the different output gaps are of the same magnitude, varying from 0.55 to 0.58.

With regard to the UGAP, correlation and concordance with the output gaps are much higher. Correlations lie in the area 0.65 to 0.77, while concordances vary from 0.82 to 0.94. The correlation and concordance measures are highest for the multivariate gaps (in particular for MVUC) as well as for UC, making these gaps slightly more reliable with regard to assessing the current economic situation.[14]

4 Forecasting inflation

We now proceed to test to what extent the various estimates of the output gap contribute to any value added over past inflation rates in predicting inflation. As our preferred measure of inflation, we use quarterly changes in the prices of goods and services produced domestically. We refer to this measure as domestic inflation[15]. We focus on domestic inflation, as import prices are less likely to be influenced by the domestic output gap. Further, domestic inflation is one of the measures of underlying inflation that the monetary authorities in Norway assess when conducting monetary policy. Chart 2 graphs both domestic inflation and total CPI. Since the late 1990s, prices of imported goods have fallen, mainly due to increased trade with China and other emerging markets. As a result, overall inflation was pushed downwards.

Chart 2 Headline inflation (CPI) and domestic inflation[1]



[1] Domestically produced consumer goods and services, adjusted for tax changes and excluding energy products.

[14] Previous studies have indicated that unemployment might be lagging the business cycle (Bjørnland 2000). However, here we find the correlation coefficient between the output gap and the unemployment gap to be largest when we investigate contemporaneous relationships, and not when the unemployment gap is lagging the cycle.
[15] See footnote 8 for an explanation.

To investigate the role of the output gaps in predicting inflation, we estimate a forecasting equation for inflation that includes the output gap as an explanatory variable. We then determine if the output gap contains additional information compared to a benchmark autoregressive (AR) model. To evaluate the forecast we first compare the root mean square forecast error (RMSE) at different horizons. In the next section, we finally investigate whether differences in forecasting performance of competing models are significant, using the Diebold and Mariano (1995) and West (1996) (DMW henceforth) test statistics.

The output gap is of course not the only information used to gauge the future path of inflation. In the recent literature on inflation forecasting, large sets of competing explanatory variables are typically considered, see for instance Stock and Watson (2004). In addition to the ICI and the UGAP discussed above, whose real-time properties are more accurate than the output gaps', we therefore also consider some alternative variables that may be equally useful as the output gap in predicting inflation. These variables are, however, also subject to revisions of varying degrees. The full set of alternative variables is listed in Table B.1 in appendix B. It includes variables reflecting pressures in the labour market (i.e. employment and unemployment data), as well as more direct wage pressure indicators. All of these are useful indicators that the Central Bank regularly monitors to assess potential future inflation pressure. In addition we also investigate the usefulness in employing (changes in) GDP directly, rather than the output gap.

Throughout the analysis, we will use a simple Phillips curve relationship to describe the dependence between domestic inflation and a given indicator (see Orphanides and van Norden, 2005). Denoting a given single indicator at time t as I_t, which could be any of the six output gaps or any one of the alternative variables, we have:

$$\pi_{t+h}^{h} = \alpha + \sum_{j=1}^{n} \beta_j \pi_{t-j}^{1} + \sum_{j=1}^{m} \lambda_j I_{t-j} + \varepsilon_{t+h}, \tag{11}$$

where π_{t+h}^{h} is domestic inflation over h quarters ending in quarter $t+h$. For example, $h=4$ is the year-on-year inflation and $h=8$ is inflation measured over 2 years. π_{t-j}^{1} is the quarter-on-quarter inflation. α, β and λ are coefficients and ε is a white noise residual. Inflation h quarters ahead is expressed as a linear function of past inflation and output gaps. In the estimation we keep the number of lags fixed; $n = 8$ and $m = 4$. The model is estimated up to time $t\text{-}1$, producing forecasts for the period $t+4$ and $t+8$. Parameters are then updated recursively, adding a new observation to the sample. The 4 quarter forecasts start in 1996:4 and ends in 2005:3. This leaves us with 36 forecasts for 4 quarter inflation. The 8 quarter forecasts start in 1997:4 and end in 2006:3 (36 forecasts). The forecasts are compared to an AR model, where we assume $\lambda_j = 0$, for j = 1,..,4.

A general impression from the literature is that there does not seem to be one indicator or variable that is superior in forecasting inflation (see e.g. Banjeree and Marcellino, 2006). These results seem to hold independent of country and time period under investigation. In order to obtain robust inflation forecasts, some kind of information pooling may therefore be useful. In particular, Stock and Watson (2003, 2004) argue that the best predictive performance is obtained by constructing forecasts from a large set of single-indicator models and simply averaging these forecasts. This would offset bias and reduce forecast error variance. However, as pointed out by Wright (2003), the conclusion that equal weighted averaging gives the best forecast may not hold in general. He instead proposes to use Bayesian model averaging, which implies that the weights assigned to the different models are updated for each forecasting period, based on the model posterior probabilities. Hence, in order to shed some light on the overall forecasting performance of the single-indicator-models, we also report RMSEs based on model averaging, using both equal weights and Bayesian updating. The Bayesian model averaging approach is explained in detail in appendix C.

Chart 3 and 4 graph the RMSE for the 4 and 8 quarter horizon respectively, for the benchmark model, the six models containing the different measures of the output gap as additional explanatory variable and the models including some of the alternative indicators. For a comprehensive list of RMSEs for the estimated models, see appendix B below.[16]

Chart 3 shows that all output gap models do better in terms of RMSE than the benchmark AR model. This is in contrast to many studies that find simple AR models to forecast better than output gap based models, see for instance Cecchetti et al. (2000), Camba-Mendez and Rodriguez-Palenzuela (2003) and Billmeier (2004).

Furthermore, the results indicate that the models using the multivariate gaps forecast slightly better at the 4 quarter horizon than the models based on the univariate gaps, with the exception of the HP gap, which does as well as the SVAR gap. The best forecasting performance, measured by RMSE, is obtained by employing the PF output gap, followed closely by models using the MVUC output gap. However, the differences between models with the alternative gaps are not large (RMSE varies from 0.87-1.01), suggesting essentially that all output gap based models perform better than benchmark ARs.

Some of the alternative indicators also do relatively well at the 4-quarters horizon, compared to the simple AR model. The best alternative indicator for predicting inflation at the 4-quarter horizon is the unemployment gap (UGAP), followed closely by the industrial confidence indicator (ICI).

[16] Appendix B also graphs the inflation forecasts from the benchmark model and the three forecasts with the lowest RMSE, together with actual inflation.

Chart 3 RMSE. 4 quarter out-of-sample forecasts. Alternative explanatory variables. 1996Q4 – 2005Q3

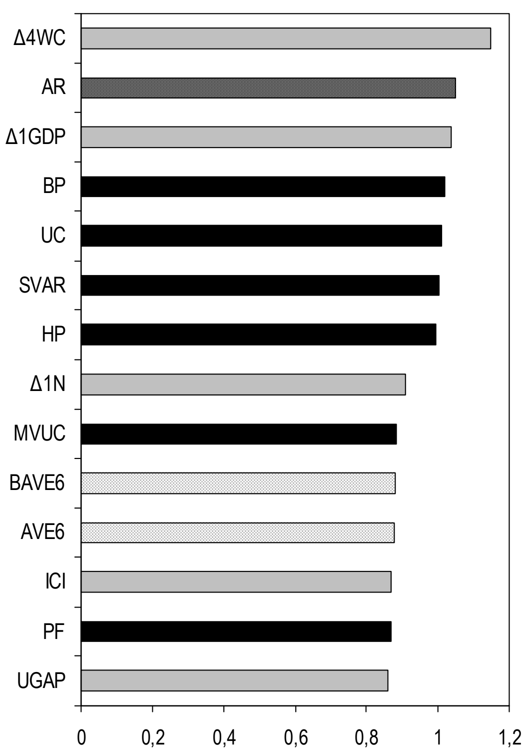
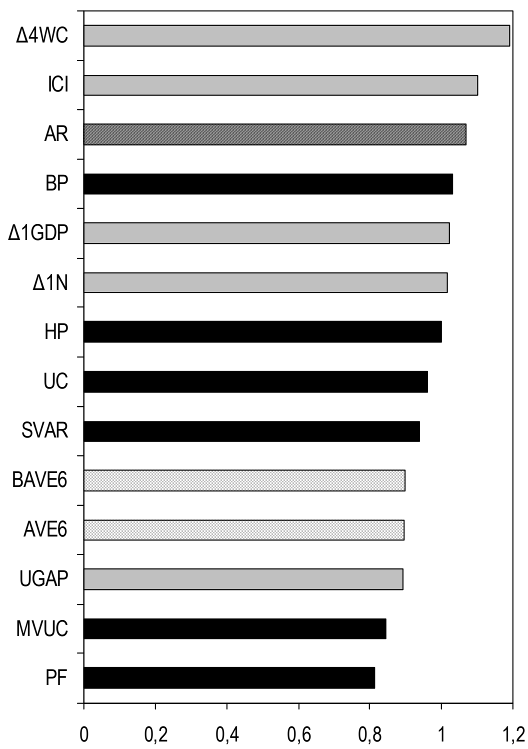


Chart 4 RMSE. 8 quarter out-of-sample forecasts. Alternative explanatory variables. 1997Q4 – 2006Q3.



Note: Bandpass filter (BP), Univariate unobserved component (UC), Hodrick Prescott filter (HP), Structural vector autoregression (SVAR), Multivariate unobserved component (MVUC), Production function (PF), Benchmark autoregressive model (AR), Unemployment gap (UGAP), Industrial Confidence Index (ICI), Wage Costs Mainland Norway, year-on-year % change (Δ4WC), GDP Mainland Norway, quarter-on-quarter % change (Δ1GDP), Employment in persons, Mainland Norway quarter-on-quarter % change (Δ1N), Average of forecasts from models with output gap (AVE6) and Bayesian average of forecasts from models with output gap (BAVE6).

The output gap based models seem to have more predicting power relative to the AR model at the 8 quarter forecasting horizon. RMSE increases for the AR model while it decreases for the three multivariate gap models. The fact that the information content in the output gaps is more important when predicting inflation at the longer horizons is not surprising. Inflation is usually lagging the output gap by 1.5-2 years during a normal business cycle (see Bjørnland, 2000, among others). Hence, the information content in the gaps will provide most value added in predicting inflation at the longer horizons. The *multivariate* gaps now seem to outperform all the univariate gaps in predicting inflation. This is interesting, since the multivariate methods rely on a wider information set than the univariate gaps. This may prove to be useful when forecasting at longer horizons. Regarding the multivariate gaps, the PF gap has the smallest forecasting error, followed closely by the MVUC and the SVAR gap. [17]

When compared to the alternative indicators, the RMSEs are now more spread out, indicating that the information content is more varied. The UGAP still outperforms the other alternative gaps, but is now beaten by both the PF and the MVUC gap. However, the general impression is that the output gaps do a superior job in forecasting inflation relative to most alternative indicators over the 8 quarter horizon (being centered low in Chart 4). The ICI is no longer among the best alternative indicators, implying that it is not as useful as the multivariate output gaps in predicting inflation at the longer horizons.

Note that neither the first nor the fourth differences of the unemployment rate do well in predicting inflation at either the 4- or 8-quarter horizon (see appendix B). On the other hand, the unemployment gap is among the indicators with the best predictive abilities, emphasizing that it is the *level* (relative to some natural rate) and not the *change* in unemployment that is the most relevant variable when predicting future inflation. Interestingly, the unemployment gap is also strongly correlated with the output gap over the sample, making it a useful indicator in real time.

Finally, our results indicate that both simple and Bayesian averaging produce forecast errors that are smaller than the median forecast error, both for the full set of indicators and the subset including only the output gaps. Furthermore, the ranking of these average forecasts appear to be relatively constant over different forecasting horizons. However, we do not reach the strong conclusions found in Stock and Watson (2003, 2004) and Wright (2003), claiming that model averaging yields superior inflation forecasts in terms of RMSE. One reason for this could be that we consider a rather limited set of indicators. Still, our results indicate that forecast averaging could be a robust approach when pooling various sources of information.

[17] We have also assessed the forecast performance of a model using a naïve forecast, predicting a flat profile of inflation over the horizon. 4 step naïve forecasts fit the data slightly poorer than the AR-model, measured by RMSE. 8 step naïve forecasts, however, did better than the AR-model, but not any better than models that include the alternative output gap in the Phillips curve.

4.1 Forecast evaluation

Finally, we employ the DMW test to explore whether the improvement in forecast accuracy reported above is statistically significant. More specifically, we test for statistical differences in the forecasting performance of competing models by comparing the squared forecast errors of the models. We will assess whether the inflation rate predicted by adding each output gap to the Phillips curve relation (equation 11) above, is significantly different from the benchmark autoregressive (AR) forecast itself.

Table 7 presents the DMW test statistic for the forecasts to be equally accurate as the benchmark forecast, with corresponding p-values. Failure to reject the null hypothesis implies that the inclusion of the output gap measure does not improve the AR model significantly. The DMW test statistic may be computed as follows:

$$DMW = \frac{\bar{d}}{\sqrt{\dfrac{2\pi\hat{f}(0)}{T}}}$$

where $\bar{d}$ is the mean of the difference in squared forecast errors between the two models that is compared, and $\hat{f}(0)$ is an estimator of its spectral density of frequency zero. Here we use the standard Newey-West robust estimator of the long run variance of $\bar{d}$.

Note, however, that the use of DMW statistics may provide non-normal critical values for asymptotic inference if the two models being compared are nested. However, Clark and McCracken (2001) find that the limiting distribution of these statistics is non-pivotal for forecast horizons greater than one period, and is therefore less of a problem here (see also the discussion in Orphanides and van Norden, 2005).[18]

The results from the DMW test confirm the results discussed above. At the 4-quarter horizon, only the PF gap performs significantly better than the benchmark model. At the 8-quarter horizon, however, all the output gap based forecasts are significantly better than the AR forecast. This is interesting, as a general finding in the literature has been that models with additional explanatory variables has tended to produce forecasts not significantly different from more parsimonious benchmark models (such as the AR model used here). However, the results may be even sharper than indicated, as Clark and West (2006) have suggested that by reducing the noise inherent in less parsimonious models, the RMSE from these models will be reduced relatively to the parsimonious benchmark.[19]

[18] Note that Ashley (2003) has argued that more than 100 observations are necessary to establish significant differences in predictive accuracy across models. Hence, with few observations, our results should be taken with some caution.

[19] Under the null that the smaller parsimonious model generates the data, the larger model will tend to produce noise into the forecasts by estimating parameters whose population values are zero. By subtracting off the average squared value of differences in forecasts, Clark and West (2006) suggest a

We believe that our results can be explained by two factors, sample stability (inflation has been relatively stable and predictable over this period) and the fact that we focus on domestic inflation in our forecasting exercise. In particular, for a small open economy, the output gap will signal pressures that will eventually feed into the *domestic* component of inflation. There is less reason to believe that the output gap can explain imported inflation in any consistent way. This can be confirmed by replacing domestic inflation with total inflation (CPI in Chart 2 above) in the forecast equation. By doing so, we find that forecasts from models including the output gaps may no longer be significantly different from the simple benchmark AR model.

Table 7 Test for significant differences in the forecasting performance of models including output gaps with an AR model for inflation. Diebold-Mariano-West test. P-values in parenthesis

Method	4-quarter 1996:4-2005:3	8-quarter 1997:4-2006:3
HP	-1.22 (0.112)	-3.95 (0.001)
BP	-0.86 (0.196)	-2.27 (0.012)
UC	-0.43 (0.333)	-2.86 (0.002)
PF	-2.39 (0.008)	-3.42 (0.001)
MVUC	-1.49 (0.068)	-3.03 (0.001)
SVAR	-0.39 (0.347)	-2.79 (0.003)

5 Conclusion

This paper evaluates a series of univariate and multivariate methods for extracting the output gap in Norway based on a set of commonly used criteria, inter alia their ability to forecast domestic inflation. The output gap based forecasts are compared both to forecasts from models using alternative indicators, and simple benchmark models.

The results illustrate that the various output gaps share some important similarities, as there is a high degree of correlation between the gaps. However, the multivariate methods display the highest correlation with other indicators of economic activity that are not (or less) revised in real time, making them more reliable with regard to assessing the current economic situation.

With regard to the usefulness for predicting inflation, all the output gaps provide information about future inflation beyond what is found in past inflation rates. This is important news for the policymakers, and a relatively unique finding in the literature. We argue that the finding

way to construct a test statistic that is much better approximated by a normal distribution than the DMW test statistic.

is due to a series of factors, of which the fact that we focus on domestic inflation in the forecast evaluation is among the most important.

In addition, the output gap based forecast models generally outperform models using alternative indicators, at both the 4 and 8 quarter horizon. One exception is the unemployment gap, which does as well as many other output gaps in predicting inflation. Hence, assessment of pressures in the economy based on the uncertain output gap could benefit from being supplemented with alternative indicators like the unemployment gap.

Furthermore, models including multivariate output gaps outperform models based on univariate output gaps with regard to predicting. The multivariate gaps also do relatively better than the benchmark model at the 8-quarter horizon than at the 4-quarter horizon, indicating that fundamentals matter more for inflation forecasting at longer horizons.

Finally, the results suggest that model averaging can be a useful approach in order to do inflation forecasting based on an uncertain output gap. Both simple and Bayesian model averaging produce forecast errors that are smaller than the median error, making them a robust way of combining different sources of information for forecasting purposes. However, our results do not allow us to discriminate between the two.

References:

Apel, M. and P. Jansson (1999) System estimates of potential output and the NAIRU. Empirical Economics 24: 373-388.

Ashley, R. (2003) Statistically significant forecasting improvements: How much out-of-sample data is likely necessary?. International Journal of Forecasting 13: 229-239.

Banjeree, A and M. Marcellino (2006) Are There any Reliable Leading Indicators for US Inflation and GDP? International Journal of Forecasting 22: 137-151

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

Bernhardsen, T., Ø. Eitrheim, A.S. Jore and Ø. Røisland (2004) Real-time Data for Norway: Challenges for Monetary Policy. Discussion Paper 26/2004, Deutsche Bundesbank.

Beveridge S. and C.R. Nelson (1981) A New Approach to the Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the 'Business Cycle'. Journal of Monetary Economics 7: 151-174.

Billmeier, A. (2004): Ghostbusting: Which Output Gap Measure Really Matters? IMF Working Paper 04/146.

Bjørnland, H. (2000) Detrending methods and stylized facts of business cycles in Norway - an international comparison. Empirical Economics 25: 369-392.

Blanchard, O.J and D. Quah (1989) The Dynamic Effects of Aggregate Demand and Supply Disturbances, American Economic Review 79: 655-673.

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

Camba-Mendez, G. and D. Rodriguez-Palenzuela (2003) Assessment criteria for output gap estimates. Economic Modelling 20: 529-562.

Canova, F. (1998) Detrending and business cycle facts. Journal of Monetary Economics 41: 475-512.

Cecchetti, S. G., R. S. Chu, and C. Steindel (2000) The Unreliability of Inflation Indicators. Current Issues in Economics and Finance 4/6, Federal Reserve Bank of New York.

Christiano, L.J. and T.J. Fitzgerald (1999) The Band Pass Filter. NBER Working Paper No. W7257.

Clark, P. (1987) The cyclical component of U.S. economic activity. Quarterly Journal of Economics 102: 797-814.

Clark, T.E. and M.W. McCracken (2006) The Predictive Content of the Output Gap for Inflation: Resolving In-Sample and Out-of-Sample Evidence. Journal of Money Credit and Banking 38: 1127-1148.

Clark, T.E. and M.W. McCracken (2001) Tests of Equal Forecast Accuracy and Encompassing for Nested Models, Journal of Econometrics, 105: 85-110.

Clark, T.E. and K.D. West (2006) Approximately Normal Tests for Equal Predictive Accuracy in Nested Models. (Forthcoming), Journal of Econometrics.

Diebold, F.X. and R.S. Mariano (1995) Comparing Predictive Accuracy. Journal of Business and Economic Statistics 13: 134-144.

Ehrmann, M. and F. Smets (2003) Uncertain potential output: implications for monetary policy. Journal of Economic Dynamics and Control 27: 1611-1638.

Frøyland, E. and R. Nymoen (2000) Output gap in the Norwegian economy – different methodologies, same result?. Economic Bulletin 2: 46-52.

Gali, J. and P. Rabanal (2004) Technology Shocks and Aggregate Fluctuations: How Well Does the RBC Model Fit Postwar U.S. Data?. In Gertler, M. and K. Rogoff, (eds.) NBER Macroeconomics Annual, 225-288.

Gerlach, S. and L.E.O. Svensson (2003) Money and Inflation in the Euro Area: A Case for Monetary Indicators?. Journal of Monetary Economics 50: 1649-1672.

Gruen, D., Robinson T. and A. Stone (2005) Output Gaps in Real Time: How Reliable are they?. The Economic Record 81: 6-18.

Hamilton, J.D. (1994) Time Series Analysis. Princeton University Press, Princeton, New Jersey.

Harvey, A. C. (1985) Trends and cycles in macroeconomic time series. Journal of Business and Economics Statistics 3: 216-227.

Johansen, P. R. and T. Eika (2000) Drivkrefter bak konjunkturforløpet på 1990-tallet (driving forces behind cyclical developments in the 1990s). Economic Survey 6/2000, Statistics Norway.

Kydland, F. E. and E. C. Prescott (1990) Business Cycles: Real Facts and a Monetary Myth. Federal Reserve Bank of Minneapolis Quarterly Review. Spring: 3-18.

Okun, A.M. (1962) Potential GNP: Its Measurement and Significance. Proceedings of the Business and Economic Statistics Section of the American Statistical Association, 98-103.

Leamer, E.E. (1978) Regression Selection Strategies and Revealed Priors. Journal of the American Statistical Association 73: 580-587.

Leitemo, K. and I. Lønning (2006) Simple Monetary Policymaking without the Output Gap. Journal of Money, Credit, and Banking 38: 1619-1640.

McCallum, B.T. (1998) Issues in the Design of Monetary Policy Rules. In Taylor, J. and M. Woodford (eds.) Handbook of Macroeconomics, New York: North Holland, 1483-1530.

McCallum, B.T. (2001) Should Monetary Policy Respond Strongly to Output Gaps?. American Economic Review 91: 258-262.

McDermott, C.J and A. Scott (2000) Concordance in Business Cycles. IMF working Paper 00/37.

Ministry of Finance (1997) Fakta og analyser (facts and analyses). Annex to Report No. 4 (1996-1997) to the Storting, Long-term programme 1998-2001, 74

Nelson, C.R. and C.I. Plosser (1982) Trends and Random Walks in Macroeconomic Time Series. Journal of Monetary Economics 10: 129-162.

Orphanides, A. (2001) Monetary Policy Rules Based on Real-Time Data. American Economic Review 91: 964-985.

Orphanides, A. (2003) Monetary Policy Evaluation with Noisy Information. Journal of Monetary Economics 50: 605-631.

Orphanides, A. and S. van Norden (2002) The Unreliability of Output Gap Estimates in Real Time. Review of Economics and Statistics 84: 569-583.

Orphanides, A. and S. van Norden (2005) The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time. Journal of Money, Credit and Banking, 37: 583-601.

Orphanides, A. Porter, R.D., Reifschneider, D., Tetlow, R. and F. Finan, (2000) Errors in the measurement of the output gap and the design of monetary policy. Journal of Economics and Business 52: 117-141.

Rotenberg, J.J. and M. Woodford (1997) An Optimization-Based Econometric Model for the Evaluation of Monetary Policy. NBER Macroeconomics Annual 12: 297-346.

Rudebusch, G.D. (2002) Assessing Nominal Income Rules for Monetary Policy with Model and Data Uncertainty. Economic Journal 112: 402-432.

Rünstler G. (2002) The Information Content of Real-time Output Gap Estimates: An Application to the Euro Area. European Central Bank Working Paper No. 182.

Scott, A. (2000) Stylised facts from output gap measures. Reserve Bank of New Zealand Discussion Paper, DP2000/07.

Smets, F. (2002) Output Gap Uncertainty: Does it Matter for the Taylor Rule?. Empirical Economics 22: 113-129.

Spencer, D.E. (2004) Output Gap Uncertainty and Monetary Policy During the 1970s. Topics in Macroeconomics 4: Article 2.

Stock, J.H. and M.W. Watson (1998) Business Cycle Fluctuations in US Macroeconomic Time Series. Working Paper No. 6528, National Bureau of Economic Research.

Stock, J.H. and M.W. Watson (2003) Forecasting Output and Inflation: The Role of Asset Prices. Journal of Economic Literature, Vol. XLI, 788-829.

Stock, J.H. and M.W. Watson (2004) Combination Forecasts of Output Growth in a Seven-Country Data Set. Journal of Forecasting 23: 405-430.

Svensson, L.E.O. (1997) Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets. European Economic Review, 41: 1111-1146.

Svensson, L.E.O. (2000) Open-Economy Inflation Targeting. Journal of International Economics, 50: 155-183.

Svensson, L.E.O. and M. Woodford (2003) Indicator Variables for Optimal Policy. Journal of Monetary Economics, 50: 691-720.

Svensson, L.E.O. and M. Woodford (2005) Implementing Optimal Policy through Inflation-Forecast Targeting. In Bernanke, B.S., and M. Woodford, (eds.) The Inflation-Targeting Debate, University of Chicago Press, 19-83.

West, K.D. (1996) Asymptotic Inference about Predictive Ability. Econometrica, 64: 1067-1084.

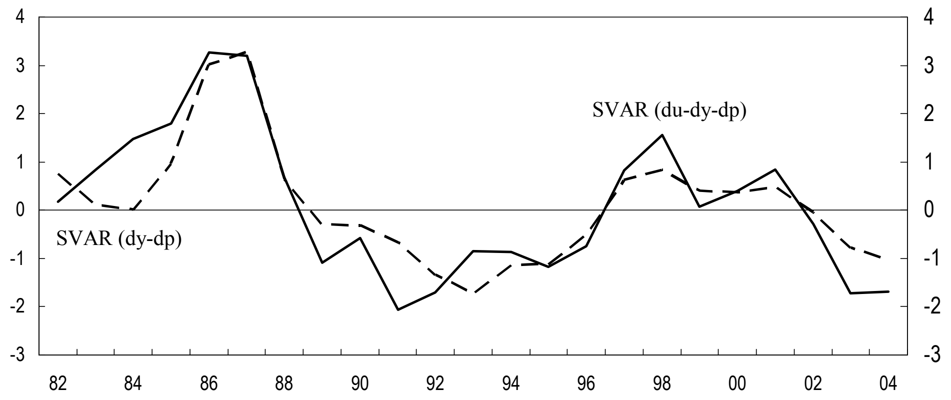
Wright, J.H. (2003) Forecasting US inflation by Bayesian Model Averaging. International Finance Discussion Paper 780, Board of Governors of the Federal Reserve System.

Appendix A

Table 8 Estimation results for Multivariate unobserved component-method (MVUC)

Parameter	Estimate	St.dev	z-Statistic
α_{11}	1.269	0.085	14.841
α_{12}	-0.400	0.088	-4.561
α_{21}	0.635	0.093	6.867
β_{11}	0.052	0.026	1.992
β_{21}	-0.159	0.027	-5.922
ψ_{11}	1.146	0.067	17.182
ψ_{12}	-0.195	0.041	-4.802
σ_{ε_2}	0.192	0.019	10.180
σ_{ε_3}	0.025	0.026	0.947
σ_{υ_1}	0.453	0.037	12.351
σ_{υ_2}	0.007	0.012	0.596
σ_{υ_3}	0.000	0.000	0.003
σ_{υ_4}	0.013	0.033	0.393
σ_{υ_5}	0.000	0.000	0.508

Chart 5 Comparing the implied output gap calculated from our preferred SVAR to the output gap from a bivariat VAR in output and inflation.



Appendix B

Table 9 List of alternative variables [1]

UGAP	Unemployment gap. LFS unemployment ratio filtered by the HP-filter (λ=40000)
Δ1U	LFS unemployment, quarter-on-quarter % change
Δ4U	LFS unemployment, year-on-year % change
ICI	Industrial Confidence Index
Δ1N	Employment Mainland Norway, quarter-on-quarter % change
Δ4N	Employment Mainland Norway, year-on-year % change
Δ1GDP	GDP Mainland Norway, quarter-on-quarter % change
Δ4GDP	GDP Mainland Norway, year-on-year % change
Δ4WC	Wage cost, growth from same quarter previous year, year-on-year % change
ULC	Unit Labour Costs

[1] Sources: Statistics Norway and own calculations

Table 10 RMSE for all indicators

4 quarter forecasts		8 quarter forecasts	
UGAP	0.8615	PF	0.8130
PF	0.8684	MVUC	0.8466
ICI	0.8702	UGAP	0.8913
AVE6 [1]	0.8786	AVE6	0.8955
BAVE6 [2]	0.8803	BAVE6	0.8976
MVUC	0.8835	Δ4N	0.9255
Δ1N	0.9106	SVAR	0.9396
Δ4N	0.9443	UC	0.9630
HP	0.9946	HP	0.9981
SVAR	1.0025	Δ1N	1.0180
UC	1.0100	Δ1Y	1.0221
BP	1.0188	Δ4Y	1.0221
Δ4U	1.0359	BP	1.0320
Δ1GDP	1.0360	ULC	1.0662
Δ1U	1.0361	Δ4U	1.0667
AR	1.0495	AR	1.0694
Δ4GDP	1.0845	Δ1U	1.0874
ULC	1.1432	ICI	1.1024
Δ4WC	1.1483	Δ4WC	1.1922

[1] Average of models with output gap
[2] Bayesian average of models with output gap

Chart 6 Forecasts and actual inflation. The first four figures show inflation forecasts over 4 quarters. The last four figures depict inflation forecasts over 8 quarters, in annualized rates. Over both horizons, actual inflation rates are shown with forecasts from the benchmark (AR) models and the three forecasts exhibiting the lowest RMSEs, respectively.

Chart 6 a

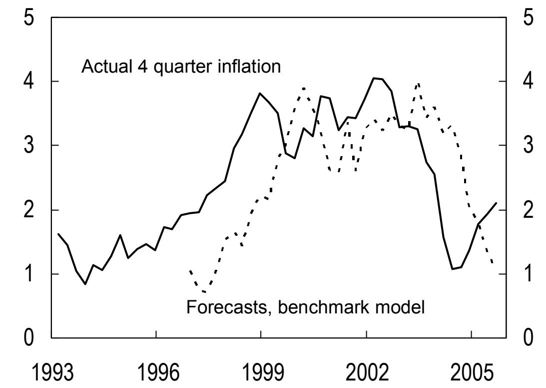


Chart 6 b

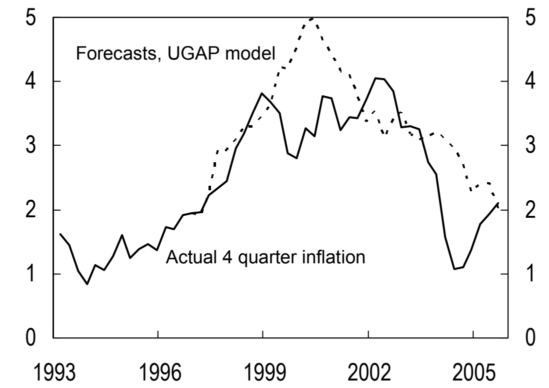


Chart 6 c

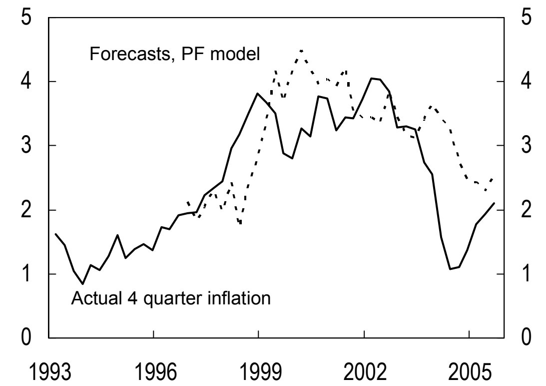


Chart 6 d

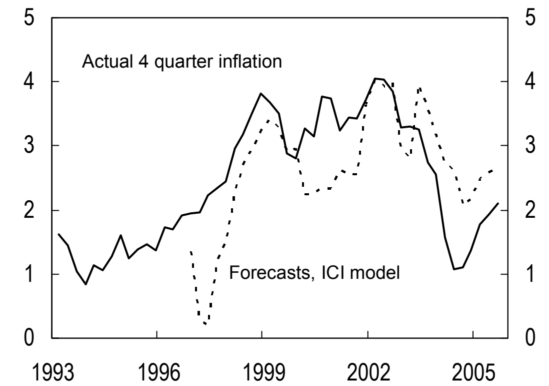
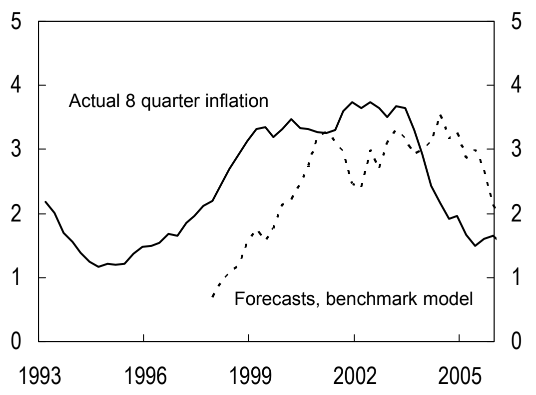


Chart 6 e | **Chart 6 f**



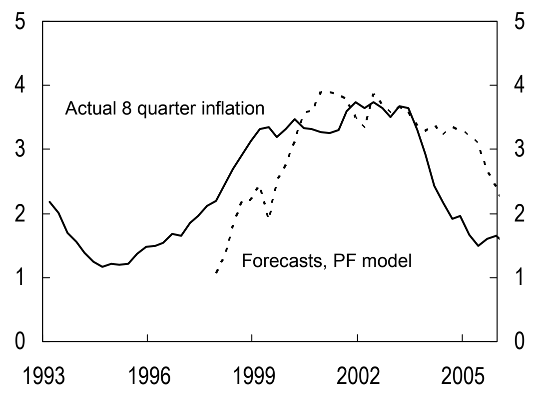
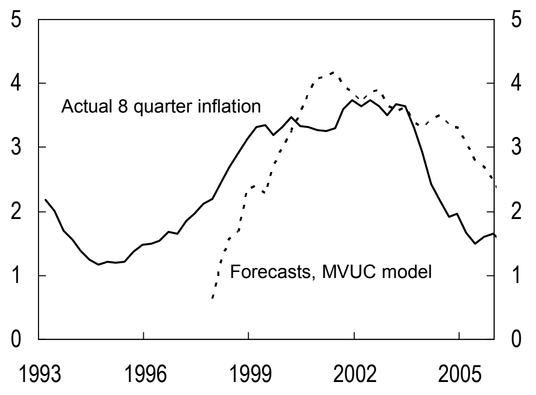
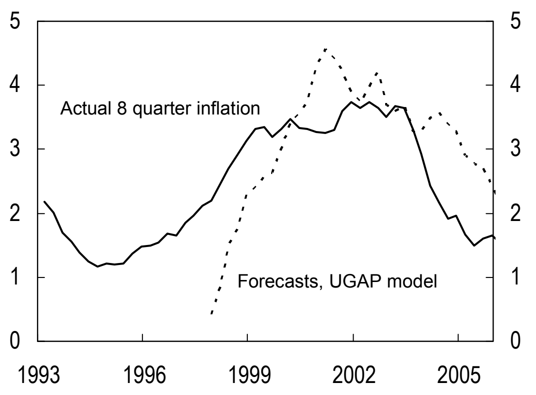


Chart 6 g | **Chart 6 h**





Appendix C

Bayesian model averaging at least goes back to Leamer (1978), and it has recently been used in many econometric applications. Wright (2003), concludes that Bayesian averaging has better forecasting properties than simple model averaging.

The starting point is a set of *n* competing models, $M_1,...,M_n$. In our case, we have *n* single-indicator models. In compact form, we can write the forecasting equation (11) as

$$Y = X_i\gamma_i + \varepsilon_i \ , \quad (i = 1,....,n)$$

where Y denotes inflation and X_i is a matrix of the different explanatory variables in model *i* (which differs from a model j only by the choice of indicator), γ_i is the corresponding parameter vector and ε_i is the vector of disturbances. The posterior probability that model *i* is the true model, is given by:

$$P(M_i \mid Y) = \frac{P(Y \mid M_i)P(M_i)}{\sum_{j=1}^{n} P(Y \mid M_j)P(M_j)} \tag{C1}$$

where $P(M_i)$ is the prior probability of M_i and

$$P(Y \mid M_i) = \int P(Y \mid \gamma_i, M_i)P(\gamma_i)d\gamma_i$$

is the marginal likelihood of M_i. $P(\gamma_i)$ denotes the prior density of the parameter vector and $P(Y \mid \gamma_i, M_i)$ is the likelihood.

A forecast *f* based on Bayesian averaging, can be written:

$$f = \sum_{i=1}^{n} P(M_i \mid Y) f_i$$

i.e. the forecast is a weighted sum of forecasts from each model, using the posterior model densities as weights.

We have assumed equal a priori weights for the different models, i.e. $P(M_i) = \frac{1}{n}$. Regarding the model parameters, the prior of γ_i is specified as a natural conjugate g-prior, whereas we assume an improper prior proportional to $1/\sigma^2$ for the variance of the error term, ε_i. This yields the following likelihood for model *i*:

$$P(Y \mid M_i) = \frac{\Gamma(T/2)}{\pi^{T/2}}(1+\delta)^{-1/2}\left[Y'Y - Y'X_i(X_i'X_i)X_i'Y\frac{\delta}{1+\delta}\right]^{-T/2} \tag{C2}$$

where δ is a shrinkage parameter. It measures the extent to which one is willing to weight the data relative to the prior. The higher is δ, the more weight is put on the data. In our exercise δ was set to 5. This is well within the range suggested by Wright (2003).

Using (C2), (C1) reduces to

$$P(M_i \mid Y) = \frac{\left[Y'Y - Y'X_i(X_i'X_i)X_i'Y\frac{\delta}{1+\delta}\right]^{-T/2}}{\sum_{j=1}^{n}\left[Y'Y - Y'X_j(X_j'X_j)X_j'Y\frac{\delta}{1+\delta}\right]^{-T/2}}$$

For each model and recursion, τ, a weight was calculated. The final weighted forecast is given by:

$$P(M_i)\left(\hat{\alpha} + \sum_{j=1}^{n}\hat{\beta}_j\pi^1_{\tau-j} + \sum_{j=1}^{m}\hat{\lambda}_j I_{i,\tau-j}\right)$$

where I_i denotes indicator i. The estimated parameters are the OLS estimates from Section 4.

Hilde C. Bjørnland, Leif Brubakk and Anne Sofie Jore: Forecasting inflation with an uncertain output gap

Working Paper 2006/2

KEYWORDS:

Output gap
Real time indicators
Forecasting
Phillips curve

Lobo Media - 35816