The Accuracy of a Forecast Targeting Central Bank

Nina Skrove Falch and Ragnar Nymoen
University of Oslo

Abstract This paper evaluates inflation forecasts made by Norges Bank which is a successful forecast targeting central bank. It is expected that Norges Bank produces inflation forecasts that are on average better than other forecasts, both 'naive' forecasts, and forecasts from econometric models outside the central bank. We find that the superiority of the bank's forecast cannot be asserted, when compared with genuine *ex-ante* real time forecasts from an independent econometric model. The 1-step Monetary Policy Report forecasts are preferable to the 1-step forecasts from the outside model, but for the policy relevant horizons (e.g., 4 quarters ahead), the forecasts from the outsider model are preferred with a wider margin. An explanation in terms of too high speed of adjustment to the inflation target is supported by the evidence. Norges Bank's forecasts are better than the 'naive' forecasts. Norwegian inflation appears to be predictable, despite the reduction in inflation persistence that has taken place over the last two decades.

Correspondence Ragnar Nymoen, University of Oslo,Department of Economics, P.O. Box 1095 Blindern, N-0317 Oslo, Norway.; e-mail: ragnar.nymoen@econ.uio.no

1   Introduction

Over the past two decades, several countries have adopted inflation targeting as the framework for monetary policy. Inflation targeting means that the central bank's inflation forecast is the target of interest rate setting, see i.e., Svensson (1997). It would be favourable for forecast targeting if it could be asserted that the central bank's forecasting model is a good approximation to the true inflation process in the economy and that inflation is stationary and without intermittent regime shifts. Under these assumptions, there would be very few inflation forecast failures.

A forecast failure occurs when inflation outcomes are outside the forecast uncertainty intervals. Failures are not rare in economic forecasting, see Hendry (2001), and inflation forecasting is no exception, Nymoen (2005). The assumptions about model correctness and of stationary inflation may therefore have low relevance for practical forecast targeting.

Although inflation forecasting central banks must face up to the possibility of forecast failure, an occasional failure is not necessarily destructive to an inflation targeting regime. Woodford (2007) argues that if inflation is under-predicted for several periods, the forecast targeting bank can nevertheless "error correct" without loss of credibility by setting an interest rate which is forecasted to generate too low inflation for a period ahead. As a result, over an extended policy horizon, and with no new breaks in that period, the inflation rate may come out on target. Realistically however, breaks can occur at short intervals and can create repeated forecast failures which may wear down the belief in the inflation forecasts, and they may complicate policy decisions (see e..g, Section 3 below).

In Norway, the central bank's inflation forecast has been the operational target of monetary policy since March 2001, and Norges Bank is regarded as a leading practitioner. For example, Woodford (2007) argues that the forecast targeting done by Norges Bank may be an example to be followed by the *US Federal Reserve*.[1] On the other hand, reviewers of Norwegian monetary policy have emphasized the need to evaluate the bank's forecasts and to work towards a clearer understanding

---

[1] "*A forecast-targeting procedure similar to that of the Norges Bank could plausibly be introduced as a framework intended to ensure that policy conforms to the mandates of the Federal Reserve Act and to make this conformity more evident to Congress and to the public.*" Woodford (2007, p 22-23).

of why relatively large inflation errors still occur, see Bjørnland et al. (2004, Ch 4.1 - 4.2) and Juel et al. (2008, Ch 4). The premise is of course that the policy decisions (interest rate setting) depends on the quality of the forecasts. The recognition of the importance of forecast accuracy is also shown by Norges Banks own annual evaluation of their forecasts, and by the priority given to data assembly and to the development of new forecasting methods.[2]

In this paper, we first review the main hindrances for accuracy in inflation forecasting, see Section 2. Existing discussions focus mainly on data measurement problems and the real time aspects of forecasting, e.g., Orphanides (2003), Svensson and Woodford (2005). Experience and research show that there are other sources of forecast failure that can be equally important though. First, even for a correctly specified model, the occurrence of intermittent structural breaks in the economy will generate forecast failures see e.g., Clements and Hendry (1999, 2008). Second, in practice, model specification plays a large role in forecast targeting, see e.g., Bårdsen et al. (2003) and Akram and Nymoen (2009).

In particular, structural breaks in the means of economic theoretical equilibrium relationships are damaging for macroeconomic forecasts. Section 3, shows, by a simple example, that the policy consequences of such a break depend on the dynamics of the transmission mechanism and the degree of gradualism in the central bank's interest rate setting.

In Section 4 we review some important developments in inflation forecasting in Norges Bank. We present Monetary Policy Reports (MPRs) forecasts from the period 2002q1-2009q4, and give a brief account of the development of forecasting models in Section 5.[3] This provides the backdrop for the main contribution of the paper, which is a comparison of the MPR inflation forecasts with alternative real-time *ex ante inflation* forecasts. In Section 6, we present briefly the alternative forecast which are, first, real time forecasts from an outside Norges Bank econometric model and, second, univariate inflation forecasts.

---

[2] Norges Bank's own evaluations are found in the the first Monetary Policy Report each year. Norges Bank's *Annual Report 2008* (Ch 1) contains a comprehensive account of forecasting methods used. Model developments are summarized in Annual Report 2007 , see http://www.norges-bank.no.

[3] As a matter of fact, the name of the publication changed from *Inflation Report* to *Monetary Policy Report* in 2007. For simplicity, we refer to all reports as MPRs, despite the anachronism.

The comparison with the forecasts from the outside econometric model is relevant because that model contains knowledge about the inflation process that Norges Bank may have lost when they changed modelling policy, on the advise of the independent expert group called Norges Bank Watch 2002, Svensson et al. (2002). Whether this change had any cost in terms of forecasting ability is however an empirical question, which we attempt to answer below. The relevance of the univariate forecast stems from the insight that simple forecasting methods may succeed exactly when econometric forecasting models fail in their forecasts. Therefore they can function as robust forecasting devices, see e.g., Clements and Hendry (2008). They also provide the standard baseline in the assesment of the forecasts from DSGE models, see e.g. Edge and Gürkaynak (2010). The results of the comparison are presented in Section 6.3, while Section 7 contains our conclusion.

2 Why Is Accurate Inflation Forecasting Difficult?

The theory of economic forecasting builds on the conceptual difference between the forecasting model and the true data generating process (DGP for short). In textbook expositions, it is assumed that the forecasting model corresponds to the DGP, expect for an unknown disturbance term. The parameters of the model/DGP can be assumed to be unknown since parameter estimation does not represent any principal difficulty for forecasting. This is because, with suitable assumptions about the distribution of the disturbances, conventional statistical measures of uncertainty, like prediction intervals or forecasts "fans", are valid and can be used to illustrate the realistic uncertainty of the inflation forecasts. This, it seems, was also one premise for choosing the inflation forecast as the target for practical monetary policy, in order to be able to monitor policy performance when lags in the transmission mechanism are a reality, see Svensson (1997) and Clarida et al. (1999).

The problem with the textbook assumption is that it does do not match up with the realities of macroeconomic forecasting. Specifically, if the assumption about model correctness really was relevant, forecast failures would be rare, which they are not. The weakest point in the theory is the (often implicit) assumption about

parameter constancy and stationarity. In practical forecasting we cannot expect the parameters to remain constant over the forecasting period—structural breaks are likely to occur when we forecast a changing economy.

As discussed by Clements and Hendry (1999), a frequent source of forecast failure is a regime shift in the forecast period, i.e., *after* the preparation of the forecasts. Since there is no way of anticipating them, it is unavoidable that *post* forecast breaks damage forecasts from time to time. The task is then to be able to detect the nature of the regime shift as quickly as possible, in order to avoid repeated unnecessary forecast failure even after the break has become part of the information set of the (next) updated forecast. Failing to pick up a *before* forecast structural break may be due to low statistical power of tests of parameter instability. There are also practical circumstances that complicate and delay the detection of regime shifts. For example, there is usually uncertainty about the quality of the provisional data for the period that initialize the forecasts, making it difficult to assess the significance of a structural change. Hence both *post* and *pre* forecast structural breaks are realistic aspects of real life forecasting situations of the type faced by inflation targeters. In particular, one should seek forecasting models and tools which help cultivate an adaptive forecasting process. The literature on forecasting and model evaluation provide several guidelines, see e.g., Bates and Granger (1969), Hendry (2001) and Granger (1999).

Realistically, the preferable model specification to use for a given observable sample is also unknown *a priori*. However, the link between model misspecification and forecast failure is not always as straight-forward as one would first believe. The complicating factor is again non-stationarity, regime-shifts and structural change. For example, a time series model in terms of the change in the rate of inflation—a random walk —adapts quickly to regime shifts, and is immune to *pre* forecast structural breaks, even though it is blatantly mis-specified over the historical data period, see Clements and Hendry (1999, Ch 5). In terms of forecasting vocabulary, the random walk model has automatic "intercept correction" to *pre* forecast breaks, making it a robust forecasting method. An economic forecasting model is less adaptable. In order to avoid forecast failure after a (big) structural break the model forecasts must be corrected by the user until the change can be "build into" the model structure. As noted above, this can be time-consuming, for example because it may be difficult to understand the nature

and consequences of the break on the basis of only a few data observations. Bård-sen and Nymoen (2009) discuss how fast the forecasts from a macroeconometric model can be expected to adapt to breaks and compare them with robust forecasts.

The relevance of breaks, and the practical problems of incorporating the effects of breaks into a model structure, provide a rationale for considering ensemble forecasts, where forecast from different forecasting mechanism and methods are averaged together. Norges Bank now relies less on its monetary policy model, and more on ensemble inflation forecasts, than they used to do. Interestingly this is motivated by the recognition that the true data generating process is unknown, see Gerdrup et al. (2009).

Uncertainty about model specification and of *pre* and *post* forecasts structural breaks are probably the main obstacles to inflation forecast accuracy. This does not deny that there are addition problems as well. Three sources that have been more in the forefront of the debate than structural breaks are: data revisions, adjustments to model forecasts and the projections of non-modelled variables in the forecasting model. Inflation forecasts and monetary policy decisions are made in "real time", so if important variables are inaccurately measured in real time, inflation forecasts will suffer by the conditioning on mismeasured initial conditions. As noted above, Woodford (2007) and others argue that the real-time data problem is the main hindrance for accurate inflation forecasting. Progress has been made with methods that can mitigate this problem. Factor models and the use of large data sets to nowcast the output-gap have for example shown promising results, see e.g., Aastveit (2010).

3   Can Forecast Errors Harm Policy?

When central banks set interest rates so that the inflation forecast is in accordance with the inflation target, there is a danger that structural breaks in the monetary policy transmission mechanism will also affect interest rate setting. But in which way, and with what potentially harmful consequences, depends both on the operational aspects of inflation targeting and the nature of the transmission mechanism.

In order to simplify as much as possible, we omit all other variables than the policy instrument. Hence, we consider the simple dynamic model:

$$\pi_t = \delta + \alpha \pi_{t-1} + \beta_1 i_t + \beta_2 i_{t-1} + \varepsilon_t, \ \ t = 1,2,3...,T, \tag{1}$$
$$-1 < \alpha < 1, \beta_1 + \beta_2 < 0,$$

where π_t denotes the rate of inflation, and i_t is the interest rate. ε_t is stochastic and normally distributed with a constant variance and zero autocorrelation.

Suppose, for simplicity, that the central bank has chosen a 2-period horizon—for the time being we may think of the period as annual. The forecasts are prepared conditional on period T information, so

$$\hat{\pi}_{T+1|T} = \delta + \alpha \pi_T + \beta_1 i_{T+1|T} + \beta_2 i_{T|T} \tag{2}$$
$$\hat{\pi}_{T+2|T} = \delta(1+\alpha) + \alpha^2 \pi_T + \alpha \beta_1 i_{T+1|T} + \alpha \beta_2 i_{T|T} + \beta_1 i_{T+2|T} + \beta_2 i_{T+1|T} \tag{3}$$

give the first and second year forecasts. There are two degrees of freedom if the bank chooses to attain the target π^* in period 2, and $i_{T+1|T}$ and/or $i_{T|T}$ can therefore be set to (help) attain other priorities.

For simplicity, set $i_{T+1|T}$ and $i_{T|T}$ to some autonomous level, represented by 0.[4] In this case, the interest rate path becomes

$$i_{T|T} = 0$$
$$i_{T+1|T} = 0 \tag{4}$$
$$i_{T+2|T} = \frac{1}{\beta_1} \left\{ -\mu + \alpha^2(\mu - \pi_T) + \pi^* \right\}$$

where μ denotes the long run mean of inflation. Assuming that the parameters are constant over the forecast period, this path will secure that $\pi_{T+2|T}$ is equal to π^* on average. This is the benign stationary case with no regime shifts in the forecast period. On the other hand, if μ increases to a higher level μ' in period $T+1$ and $T+2$, the forecasts $\hat{\pi}_{T+1|T}$ and $\hat{\pi}_{T+2|T}$ will turn out to be too low, and if large

---

[4] i.e.,we interpret i_t as the deviation from mean, hence the interest rate in period T and the planned interest rate in $T+1$ are both equal to the mean.

enough, the errors will constitute a forecast failure. However, the forecast failure is not too worrying since only the announced future interest rate $i_{T+2|T}$ is affected. With the 'policy rule' in (4), a future interest rate is planned to be changed in such a way that the inflation target is reached in the second year. Today's interest rate $i_{T|T}$ is not affected, and the planned $i_{T+2|T}$ can always be replaced by $i_{T+2|T+1}$ in the next forecast round. Thus, there seems to be negligible damage on today's policy associated with a poor forecast.

Note that the apparent "policy irrelevance" of forecast failure depends crucially on the transmission mechanism, namely that the transmission of interest rate changes on inflation is sufficiently fast. Formally, unless $\beta_1 < 0$, the interest rate two years ahead cannot be used to bring the forecasted rate of inflation in line with the target. Central banks typically regard the transmission mechanism to be dynamic, and Norges Bank's view is that *'Monetary policy influences the economy with long and variable lags'*.[5] This is probably the main reason for choosing a different operational procedure in practise, namely to move $\hat{\pi}_{T+1|T}$ in the direction of the target by changing the current interest rate, $i_{T|T}$. As a rule, a central bank's policy will therefore be to change the interest rate gradually. In our simplified model, we can represent gradual instrument adjustment by:

$$i_{T+j-1|T} = \gamma i_{T+j|T}, \quad j = 1, 2, ...h, 0 \leq \gamma < 1$$

which gives the interest rate path:

$$i_{T|T} = \frac{\gamma^2}{B} \left\{ -\mu + \alpha^2(\mu - \pi_T) + \pi^* \right\}$$
$$i_{T+1|T} = \frac{\gamma}{B} \left\{ -\mu + \alpha^2(\mu - \pi_T) + \pi^* \right\} \tag{5}$$
$$i_{T+2|T} = \frac{1}{B} \left\{ -\mu + \alpha^2(\mu - \pi_T) + \pi^* \right\}$$

where $B < 0$ is a function of α, β_1, β_2 and γ. Clearly, with gradual interest rate adjustment, the event of a new mean inflation rate μ' in period 2 will not only cause a forecast failure, it will also imply that today's interest rate $i_{T|T}$ ought to have been set differently.

---

[5] see, http://www.norges-bank.no/english/monetary_policy/in_norway.html#horizon

In sum, long lags in the transmission mechanism, and gradualism in interest rate setting imply that inflation forecast failures also means errors in interest rate setting. We have considered a *post* forecast structural break, but the same analysis applies to *pre a* forecast structural break and real-time data measurement problems that are not taken into account by the forecast targeter.

4   Norges Bank's Inflation Forecasts since 2002

On 29 March 2001 Norway formally introduced an inflation targeting monetary policy regime. The operational measure of core inflation has been based on the consumer price index adjusted for the influence of energy prices and indirect taxes (CPI-ATE) and the target was set to 2.5 %. CPI-ATE, is published by Statistics Norway together with the headline CPI. Recently, Norges Bank has changed to a new price index called CPI-EX. This price index is not produced by Statistics Norway, but by Norges Bank. CPI-EX allows for permanent, but not temporary effects of energy price changes. Inflation forecasts are published three times a year in the central bank's monetary policy reports. The change in the operational definition of core inflation means that the main forecasts in the monetary policy reports are for the CPI-EX index. However, forecasts for CPI-ATE inflation are still included, but only for the first quarters ahead.[6]

As discussed by Akram and Nymoen (2009), the horizon for monetary policy is relevant for optimal interest rate setting in an inflation targeting regime. Since interest rate setting is linked to the inflation forecasts, and vice versa, any changes in the policy horizon is also likely to affect the inflation forecasts. Specifically, since all end-of-period forecasts are 2.5 %, a short policy horizon will imply forecasts that converge more quickly to the target than will be the case for a longer policy horizon. Initially Norges Bank operated inflation targeting with a 2-year forecast horizon. In the summer of 2004, the policy horizon was changed to 1-3 years. Judging from the graphs, this instigated a change in the following MPR

---

[6]   Monetary Policy Report 3/09 only contains forecast for 2009q3, 2010q1 and 2010q2.

The CPI-EX was introduced in Monetary Policy Report 2/08, see the box on page 54 titled "Underlying inflation", http://www.norges-bank.no/templates/article____69444.aspx

inflation forecasts, which seem to have been geared towards 2.5% over a longer period than before.

Norges Bank's forecasts are published in fan-charts where the wideness of the bands represents 30%, 60% and 90% probabilities for future inflation rates. The forecasted uncertainty is particularly relevant when assessing forecast performance. Actual inflation rates outside the 90% bands represent *forecast failures*, since they are inflation outcomes that were judged to be highly unlikely in Norges Bank's analysis of the Norwegian inflation mechanism. Forecast failures are not uncommon in economics and can sometimes be used constructively to increase knowledge, see e.g., Eitrheim et al. (2002). That said, in a practical forecasting situation there is a premium on avoiding forecast failures, and Section 3 notes the specific relevance of this for inflation forecast targeting.

Figure 1 and 2 shows the inflation forecasts from the 21 Monetary Policy Reports published in the period 2002-2008.[7] In each panel there are graphs for the dynamic inflation forecasts together with the 90% forecast confidence bounds, and also the actual inflation rate.

In Figure 1 there are several examples of forecast failure. For example in MPR 2/02, the first four inflation outcomes are covered by the forecast confidence interval, but the continued fall in inflation in 2003 (the second year of the forecast horizon) led to a forecast failure. The forecast failure became even more evident in the two other forecasting rounds from 2002, and also the three forecasts produced in 2003 predicted significantly higher inflation than the actual outcome. Specifically, the forecast confidence interval of MPR 3/03 did not even cover the actual inflation in the first forecast period.

The seventh panel shows that the forecasted zero rate of inflation for 2004(1) in MPR 1/04 turned out to be very accurate. The change from the MPR 3/04 forecast is evident, and can be seen as an adaptation to a lower inflation level. That process continued in MPR 2/04, where the effect of the lengthening of the forecasting and policy horizon mentioned above is clearly visible. Although also

---

[7] The last Monetary Policy Report in Figure 1 is MPR 3/08. The reports from 2009 are omitted because the forecasts there cover only a short horizon, and the prediction intervals are also incomplete or missing. This reflects Norges Bank's change to a new operational definiton of inflation, as explained in the main text. The available CPI-ATE forecasts from MPR 1/09, MPR 2/09 and MPR 3/09 are used in the comparison of mean forecast errors and mean square forecast errors though.
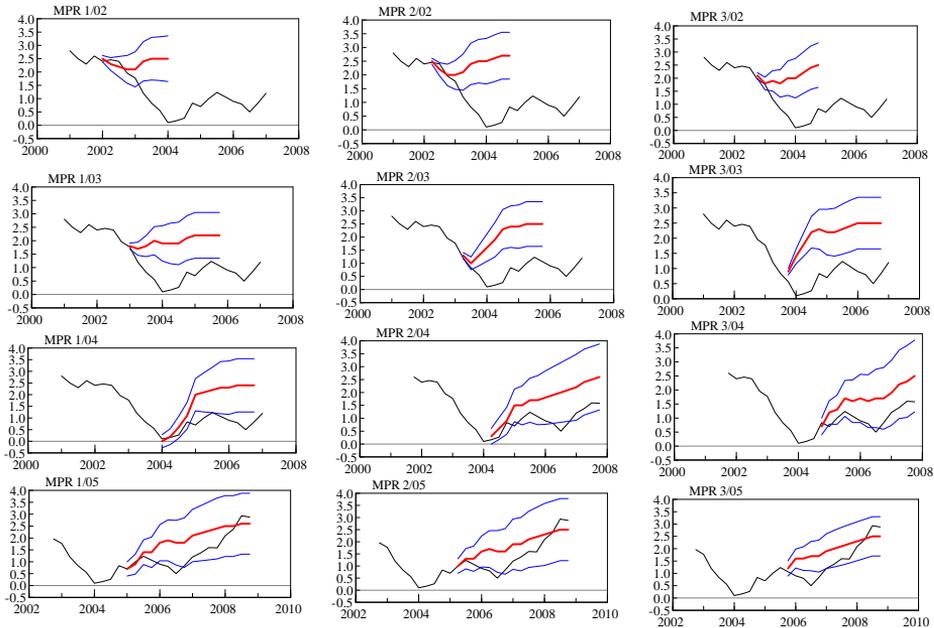
Figure 1: Actual CPI-ATE inflation rate and Norges Bank's inflation forecasts (thicker line) and the 90% confidence regions. MPR 1/02 to MPR 3/05.

the MPR 2/04 forecasts are too high, only one of them represents (formally) a forecast failure. The last four panels in Figure 1 show many of the same features. The 1-step, and sometimes also the 2-step forecasts are accurate, but otherwise the forecasted inflation rate is too high. The MPR 1/05-3/05 forecasts for the end-of-horizon are accurate though, as actual inflation was a little higher than 2.5 %.

Norges Bank's inflation forecasts from 2006, 2007 and 2008 are shown in Figure 2. Compared to the first group of forecasts, these graphs shows a more balanced picture with positive and negative forecast errors. The uncertainty bands also seem to be better calibrated to the real uncertainty facing the inflation fore-caster. It might be noted that the "zero uncertainty" for some of the 1-step fore-
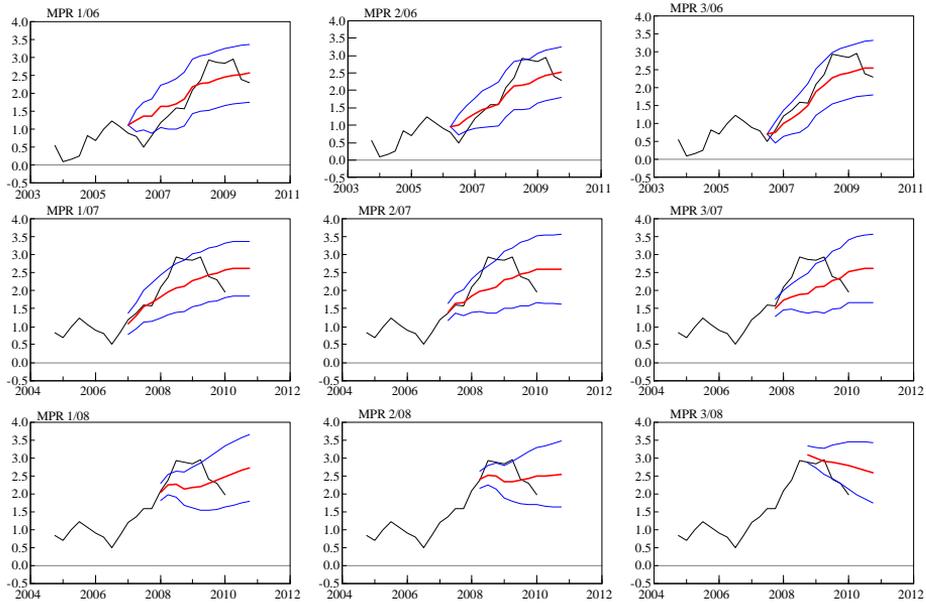
Figure 2: Actual CPI-ATE inflation rate and Norges Bank's inflation forecasts (thicker line) and the 90% confidence regions. MPR 1/06 to MPR 3/08.

casts, notably from MPR 1/06, MPR 2/0 and 1/03 are reproduced here as they appear in the data. We see that in the later reports the practice of reporting uncertainty also for the 1-step forecasts is taken up again.

5   Macroeconomic Forecasting Models Used by Norges Bank

The macroeconometric models used by Norges Bank appear to have developed considerably during the decade of inflation targeting. Following the advise in Svensson et al. (2002), Norges Bank developed a simple New Keynesian model to aid monetary policy documented in Husebø et al. (2004). This first model to aid

inflation targeting was dubbed *Model 1* by Norges Bank. Further model development, led to the model currently used by Norges Bank which a dynamic stochastic equilibrium model (DSGE) called NEMO (Norwegian Economic Model), see Brubakk et al. (2006).

As documented in Juel et al. (2008, Ch. 4), *NEMO* and its predecessor represent the main conceptual reference framework for Norges Banks forecasting and policy evaluation.[8] More recently, the role of *NEMO* for the short-term forecasts (defined as 1-4 quarters ahead) has been reduced, as Norges Bank now relies on ensemble forecasts for these horizons, see Gerdrup et al. (2009). The ensemble used for forecasting core inflation contains 167 models, *NEMO* being one of them. However, for horizons 1 to 4 years ahead, NEMO the forecasts carry full weight, see Olsen (2011).

Other aspects of Norges Bank's forecasts have developed as well. For example, until Monetary Policy Report of 3/05, the inflation forecasts were conditional on an exogenously given interest rate path which implied that the target could be met (credibly) at that interest rate path.[9] Starting with MPR 1/06, Norges Bank has published the interest rate forecast together with the inflation forecasts. The interpretation is that the inflation path and the interest rate path are consistent and that the interest rate path "gives" the inflation path.

Without becoming too speculative, it seems possible that Norges Bank's record in inflation forecast accuracy can be related to the dominant features of the models that have been used in the period that we consider. Specifically, the first operational model, *Model 1*, was a stylized representation. For example there was no direct effect of foreign inflation on Norwegian inflation. The only channel for "imported inflation" was the real exchange rate, which was one of two forcing variables in the Phillips curve equation of the model (the other one being the output-gap). The econometric assessment of this model in Nymoen and Tveter

---

[8] As noted above, this is representative of forecast targeting central banks. Hammond (2010) Table C shows that 20 out of 27 inflation targeting central banks either use or are developing a DSGE type model for forecasting and policy analysis.

[9] Alternatively, one can say, as in Juel et al. (2008), that these forecasts were not meant as conditional inflation forecasts for the policy relevant horizon, but that would suggest that Norges Bank failed to follow the main principle of inflation targeting, namely that the conditonal inflation forecast *is* the target of monetary policy.

(2007) showed that this model failed on standard econometric tests, and that the omission of a import price growth (expected or lagged) could explain the failure. Together with a high speed of adjustment to the inflation target in *Model 1*, the omission of the imported inflation channel goes some way towards explaining the forecast failures in Figure 1.

NEMO became the operational model during 2008 and is a fully fledged DSGE model with forward dated variables in all the behavioural equations. Brubakk and Sveen (2009) show *ex post* forecasts from *NEMO* for the period 1998q4-2006q4.[10] For the periods that overlap with Figure 1 and 2, the *NEMO* forecasts show a clear resemblance to the MPR forecasts shown above. Specifically, also the ex post *NEMO* forecasts seem to adjust too fast towards the inflation target after the shock to inflation in 2003. The significant weight on the forward component in the Phillips curve may contribute to a high degree of price flexibility in the model solution, which in turn may make the NEMO forecasts misspecified for periods that follow after a shock to inflation.

Another potentially important misspecification relates to wage setting and its role in the wage-price spiral. The wage equation in *NEMO* has no compensation effect for increases in the cost of living (i.e., CPI inflation), and there is no effect from the rate or unemployment (or vacancies). Both of which existing empirical models of Norwegian wage formation have shown is important, see e.g., Juel et al. (2008). The absence of an inflation effect in wage setting means that there is no genuine wage-price spiral in *NEMO*. This feature may also be part of the explanation for why the inflation forecasts of *NEMO* react relatively little to foreign nominal shocks, and why the forecasts adjust quickly towards the target.

The above observations can also be related to the general discussion about the relationship between the forecast performance of the typical DSGE model and its structure, see e.g. Rubaszek and Skrzypczynski (2008), Wang (2009) and Cristoffel et al. (2010). A DSGE model is parsimonious in terms of parameters, which is due to the theoretically motivated specification and several cross-equation restrictions. If the implied restrictions are not "very false" the parsimony may help the forecast performance, for example if it purges the model for spurious trends. However, the DSGE model's theoretically imposed balanced growth path implies tight

---

[10] See Chart 6, p. 46, in Brubakk and Sveen (2009).

restrictions of the mean of the growth rates of the endogenous variables that are forecasted. Specifically, the model assumes common growth rates for the subset of real and nominal variables, respectively. As shown in Cristoffel et al. (2010), there can be detrimental effects of the trend restrictions on dynamic forecasts over a relatively short horizon, e.g., 1-8 quarters.

6   Relative Forecast Accuracy

As noted and discussed above, the MPR forecast errors bear the marks of slow adaptation to a reduction of the inflation rate that became manifest during 2003. After 2004 there are fewer very large forecast errors. The first period with large forecast errors was analysed in Nymoen (2005), where it was shown how an econometric inflation model produced better forecasts. The explanation offered was that explanatory factors like imported inflation and labour market pressure were better represented in the econometric model than in the monetary policy model used by Norges Bank. These econometric inflation forecasts were however *ex post*, since the econometric model was designed late in the year 2003, after the biggest forecast failures had occurred. In the period 2004-2009 we have however published forecasts from the same model and these forecasts are directly comparable to the Monetary Policy Report forecasts from the same period.

Below, we compare the MPR forecasts with the forecast from the outsider econometric model which is presented briefly in Section 6.1. The second group of contesting forecasts use univariate time series models, which are introduced in Section 6.2. Section 6.3 shows the comparison by graphs and Section 6.4 contains formal tests of the equality of prediction mean squared errors.

6.1   An Outside Norges Bank Econometric Model

Real time forecasts from the macroeconometric model in Nymoen (2005) have been published on the internet since June 2004 and have been dubbed *Automatized Inflation Forecasts* (AIF).[11] The AIF forecasting model consists of an econometric "inflation function" for Norway and equations for the each of the ex-

---

[11] The url for the AIF project is http://folk.uio.no/rnymoen/forecast_air_index.html.
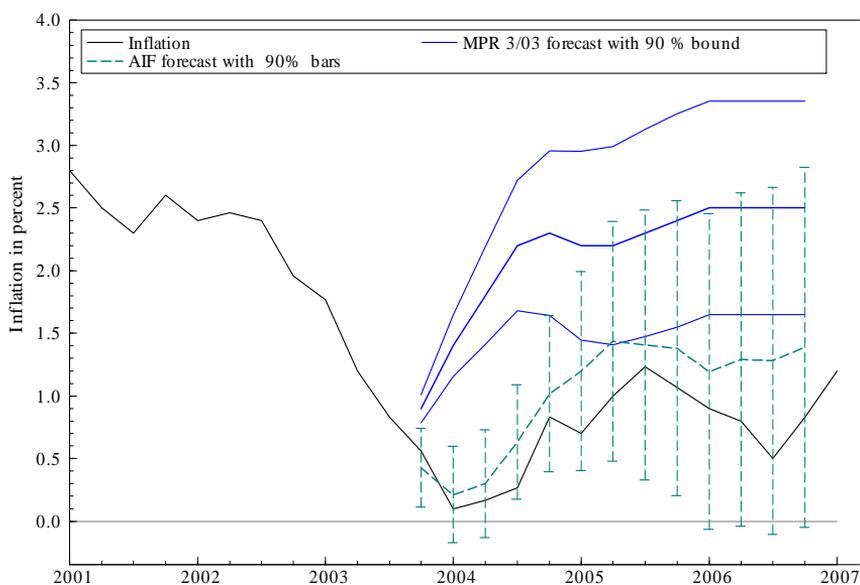
Figure 3: Ex post AIF forecasts for 2003q4 - 2006q4, with 90 % confidence regions shown as bars (in green), shown together with MPR 3/03 forecasts with 90 % bounds (in blue). The sample period for the estimation of the parameters of the AIF model ends in 2003q2. MPR 3/03 is also conditional on the information in 2003q3.

planatory variables. It is a closed system of equations and after estimation of the model's parameters, dynamic inflation forecasts are easy to produce (hence "Automatized").

In more detail, AIF consists of an equation for the rate of inflation (CPI-ATE), and 8 equations which are needed to forecast the following variables: the rate of unemployment, productivity growth, the logarithms of the nominal and real exchange rates, foreign inflation, domestic and foreign interest rates and the logarithm of the oil price. The model is presented in appendix A.

Like any forecasting model, the AIF forecasting model has evolved over time. Model maintenance is necessary for having a reasonably well adapted econometric forecasting system, but there is of course no guarantee that the model's fore-

casts are not damaged by structural breaks, see e.g., Eitrheim et al. (2002). The latest version of the AIF model has retained the same ability to outperform Norges Bank's forecasts *ex post*, see Figure 3. The graph shows AIF *ex post* forecasts for the period 2002q4-2006q4, together with the MPR 3/03 forecasts.

In Section 6.3 we investigate whether this result emerges also after 2003. For that purpose we compare forecasts from all MPRs in the period 2004-2009 to genuine real time forecasts from the AIF forecasting model that have been published on the internet during the period 2004-2009.

6.2 Univariate Forecasts

In addition to the AIF forecasts, we also compare the MPR forecasts with three simple univariate forecasting methods. The first is a random-walk model of inflation (M1 below) where a unit root is imposed, so that the full value of the previous inflation rate is retained in the forecast. This ensures a form of error-correction in the forecasts, so that the random walk model (M1) may succeed exactly in the situations where econometric forecasting models fail, see e.g. Clements and Hendry (1999, Ch. 5). One particularly relevant case is when there is a structural break in the mean of inflation. As this break moves from being a *post-forecast* break to becoming a *pre-forecast* break, the M1 forecast will adapt perfectly to the new mean of inflation. It follows that the differenced data model M1 may potentially have smaller and less systematic errors than the forecasts of an econometric model like AIF, and from Norges Bank's macroeconomic model NEMO which is also an equilibrium-correction model.

As noted above there was a marked drop in the mean of inflation in 2003. After 2004, inflation again has a positive drift, and M1 will then have a tendency to miss the evolution of inflation by one quarter. However, two other univariate models, which are almost as easy to use as M1, may be better suited to compete with the MPR forecasts over that period, They are the random-walk with drift, which we dub M2, and the autoregressive model, M3 below.

In M2 we add an estimated intercept to the random-walk model, so that it may forecast a trend in inflation. In M3 an autoregressive coefficient for the lagged inflation rate is estimated, so the unit roots of M1 and M2 is not imposed. M3 may be expected to forecast well when the inflation rate is relatively stationary.

The M1-M3 forecasts are updated each quarter. For the M1 forecasts we only need to update the initial value. For M2 and M3, we use a rolling 40 quarter sample for the estimation of the parameters. For example: the M2 and M3 forecasts for 2004q1-2006q1 use data from 1994q1 to 2003q4 for estimation, while the 2007q1-2009q4 forecasts use a sample from 1997q1 to 2006q4. To summarise, we will assess the forecasts of the following univariate methods:

- M1: The rate of inflation in period t is equal to the inflation rate in period $t-1$.

- M2: As M1, but with an estimated intercept added.

- M3: As M2, but an autoregressive coefficient is estimated (instead of being restricted to one, as in M1 and M2).

The CPI-ATE based inflation data are never revised, and it is therefore relevant to compare the M1-M3 forecasts with the MPR forecasts for the whole period of inflation targeting.

6.3 Forecast Comparison

As mentioned above, we compare the MPR inflation forecasts with real time forecasts from the AIF macroeconometric model, and with forecast from the three univariate methods M1-M3. We start with the period from 2004 until 2009 since the first AIF forecast were published on the internet in June 2004.[12]

Panel a) of Figure 4 shows the mean forecast errors, MFE, for the MPR and the AIF forecasts from one to twelve quarters ahead. The length of the forecast horizon is along the horizontal axis. Hence, the start of the graph marked MPR in panel a) gives the MFE for *Horizon 1* as the average of all 1-step ahead forecast errors in the Monetary Policy Reports from 1/04 to 3/09. Then follows the average of all 2-step ahead forecast errors, for the same sequence of MPR forecasts, and so on, until *Horizon 12*. The graph marked AIF in panel a) has a similar interpretation: it shows the average forecast errors from all published AIFs, as a function

---

[12] The AIF project have been discussed by Norges Bank Watch (2008), see Juel et al. (2008), who conclude that it represents a relevant alterantive to the MPR forecating method.
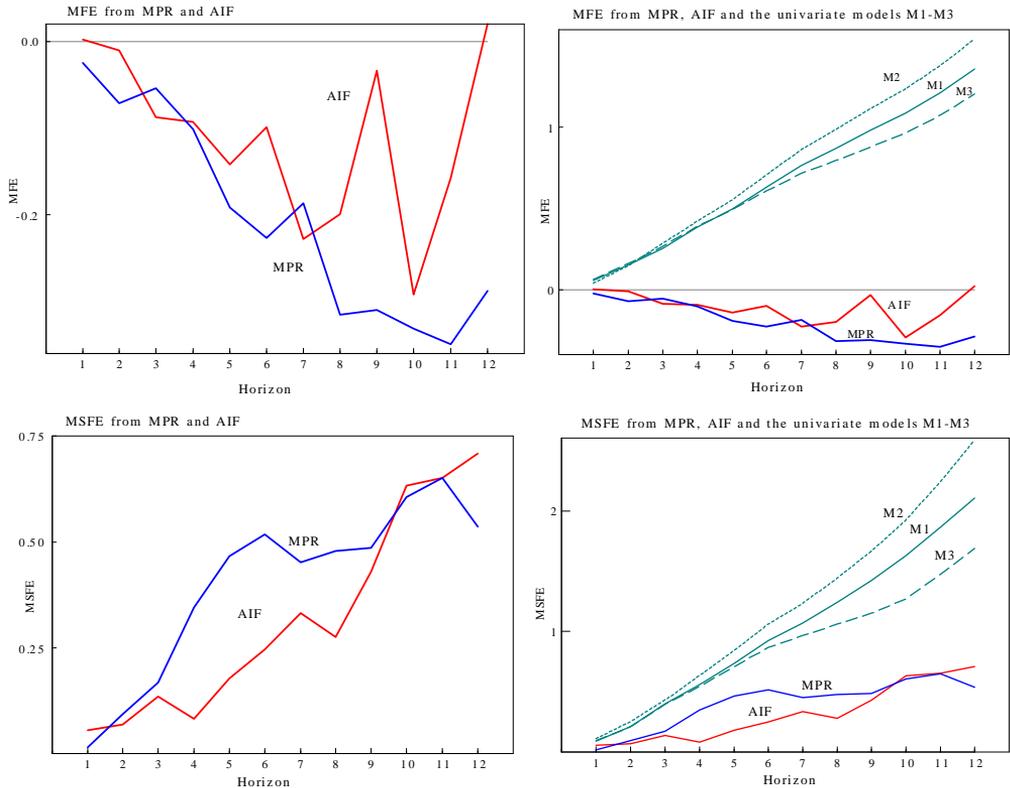
Figure 4: Mean forecast errors (MFE) and mean square forecast errors (MSFE) for core inflation (annual rate of change in CPI-ATE) for MPR (Norges Bank) and AIF (outside econometric model) forecasts. The MPR and AIF forecasts were published in the period 2004-2009. Panel a). and c) show the MFE and MSFE for forecast horizons 1-12 for MPR and AIF. Panel b) and d) also show the MFEs and MSFEs for the random walk (M1), the random walk with drift (M2) and the autoregressive model (M3).

of the forecast horizon. The last observation of inflation used in the underlying calculations is from the first quarter of 2010.

A negative MFE means that the inflation forecasts were on average higher than the actual inflation rates in the period. For example a horizon of four quarters the AIF forecasts were on average 0.2 percentage points too high. The biases are small to begin with, for the short forecasting horizons. There is a tendency that the size of the bias grows with the length of the forecasting horizon, in particular this is true for the MPR forecasts. The MPRs have smaller bias than AIF for horizons 3, 4 and 7. For the other horizons, the AIF average forecast errors are smaller than the MPR average errors.

Panel b) also shows average forecast errors, but we have here included the MFEs for the three univariate methods. M1-M3. As noted, M1 is a random walk model for the rate of inflation, M2 is a random walk with drift, and M3 is an autoregressive model for inflation. The 1-step ahead forecasts of M1-M3 have small and positive biases. The biases increase with the length of the forecast horizon—and much faster than for MPR and AIF. It is interesting to note that, because of the sign difference, an ensemble forecast made up of MPR, AIF and M1-M3 could have had a near zero mean forecast error.

Panel c) shows the mean squared forecast errors, MSFE for MPR forecasts and AIF forecasts, as a function of the forecast horizon. Given a symmetric loss function, which is relevant for inflation targeting, the forecast with the lowest MSFE may be regarded as the preferred forecast. The MPRs from 2004 to 2009 have 1-step inflation forecasts with a lower MSFE than the 1-step AIFs from the same period. However, for the 2-step forecasts, the roles are reversed. Importantly, for MSFEs for the 4-to-8-step forecasts are much larger for MPR than for the AIF. This result supports the hypothesis formulated above, about the speed of adjustment in inflation towards the target being overestimated in the central bank forecasting model. The result is the more noteworthy since it is found for a sample where almost all the MPR forecasts were produced after the extension of the monetary policy horizon in 2004, which resulted in slower speed of adjustment in the MPR forecast, as we have seen.

The last panel d) in Figure 4 shows the MSFE graphs for the three univariate forecast M1-M3 for comparison. All three are clearly inferior to both the MPR forecast and the AIF forecasts. The M3 forecasts, from the autoregressive inflation
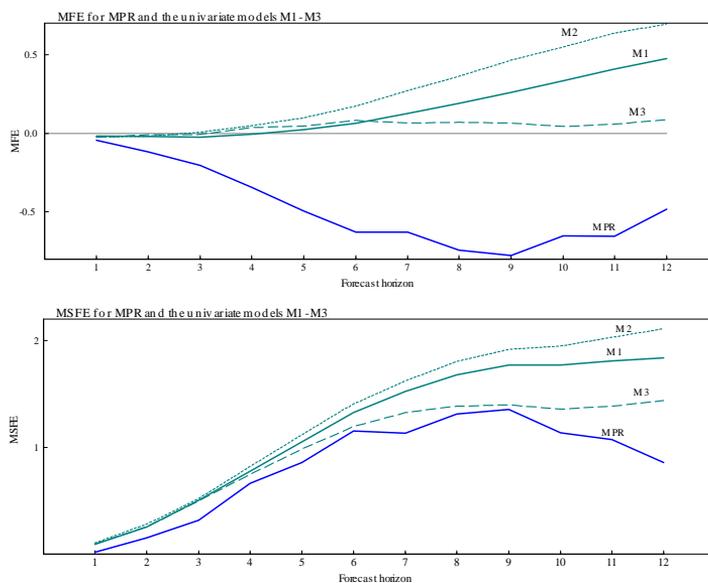
Figure 5: Mean forecast errors (MFE) and mean square forecast errors (MSFE) for core inflation (annual rate of change in CPI-ATE) for MPR (Norges Bank) and the three univariate models: random walk (M1), random walk with drift (M2) and autoregressive model (M3). The MPR forecasts were published in the period 2002-2009. Panel a. shows the MFEs for forecast horizons 1-12, and panel b) shows the MSFEs.

model, is the best of the three. Hence the robust feature of M1 is not enough to make it perform nearly as well as MPR over this forecast period.

In Figure 5 we compare the forecast accuracy of the MPR forecasts from the longer period 2002-2009, with the corresponding forecasts from the three univariate models. This means that the year 2003, which was difficult to forecast, is now part of the sample. The graphs show that both the average forecast errors and the MSFEs for the central bank forecasts are much larger in magnitude when 2003 is included in the evaluation period. In contrast, the average forecast errors of the univariate forecasts are smaller than in the assessment that omitted 2003, confirm-

ing that they are relatively robust to changes in the mean of the forecasted variable. In particular M3 has a low bias, measured by MFE, for all forecast horizons.

The second panel of Figure 5, with the MSFEs, shows that there is less absolute improvement for M1-M3 compared to Figure 4, panel d). However, because the MPR forecasts are poorer than on the sample that omits 2003, the relative accuracy of the univariate forecast are much closer to MPR than in the comparison we saw in Figure 4.

6.4 Testing the Equality of Mean Squared Inflation Forecast Errors

As the graphs show, some of the differences between the forecasts errors seem to be large. On the other hand, there is a good deal of variability, and given the low number of observations, it is an open issues whether the differences are statistically significant. To investigate this more formally, we test a hypothesis of equality of the mean squared errors.

In order to test the equality between MPR and AIF forecast errors, we must first tackle the complication that there are three MPR inflation forecasts each year, but only two AIFs. The tests in Table 1 are based on comparison of the prediction errors of the first AIF and MPRs each year (AIF-1 was published in January/February each year, and MPR-1 in March), while the second AIF forecast (from July) has been compared with the average of the squared errors of MPR-2 (June) and MPR-3 (October). The tests of equality of forecast accuracy between the MPR forecasts and the univariate forecasts is based on comparing the three annual MPRs with the three first univariate forecasts. In order to retain a necessary minimum of observations, we test the equality of forecast accuracy for horizons $h = 1, 2, \ldots, 6$.

We concentrate on the years from 2004 to 2009, which means that only genuine real time AIF forecasts error are used in the calculations. For comparison, the observations for the tests of equality of the MPR and univariate forecast errors are taken from the same period. Judging from Figure 4, panel b), we can use M1 as a representative for the three univariate forecasts: The means of the M1 squared prediction errors are only a little larger than the M2 and M3 errors (for the horizons that we consider). Specifically, unless we reject the null of "equal

Table 1: Tests of equality of prediction mean squared errors from Monetary Policy Report (MPR) forecasts, and from the two alternatives: the econometric model (AIF) and the random walk (M1). Original (DM) and modified Diebold-Mariano (MDM) tests. The numbers in parenthesis below each test are *p*-values using t_{n-1} critical values.

		MPR *vs* AIF				MPR *vs* M1	
h	*n*	DM-test	MDM-test	*h*	*n*	DM-test	MDM-test
1	11	-2.08 (0.064)	-1.98 (0.076)	1	18	-2.49 (0.023)	-2.42 (0.027)
2	11	0.63 (0.543)	0.55 (0.594)	2	17	-1.54 (0.143)	-1.41 (0.178)
3	11	0.95 (0.365)	0.74 (0.476)	3	17	-2.15 (0.047)	-1.83 (0.086)
4	9	3.57 (0.005)	2.43 (0.036)	4	15	-1.24 (0.234)	-0.96 (0.352)
5	9	5.85 (0.000)	2.90 (0.020)	5	15	-0.58 (0.571)	-0.42 (0.680)
6	9	10.33 (0.000)	3.98 (0.004)	6	13	-0.93 (0.369)	-0.57 (0.578)

Note: For MPR *vs* AIF, the tests for $h = 4$
and $h = 3$ truncate the autocovariances at $k = 3$ and $k = 2$, respectively.

precision" between MPR and M1 at a low *p*-value, it is unlikely that there is any formal support for superiority of MPR forecasts over M2 and M3 forecasts.

Table 1 contains the outcome of the Diebold-Mariano test (DM) due to Diebold and Mariano (1995), and the modified test (MDM) proposed by Harvey and Newbold (1997). The basic entities in both tests are pairs of *h*-steps ahead forecast errors, e.g., e_1 and e_2. The tests are based on the observed sample mean

$$\bar{d} = \frac{1}{n} \sum_{\tau=1}^{n} (e_{1\tau}^2 - e_{2\tau}^2) \tag{6}$$

where n denotes the numbers of pairs of forecast errors (the number of observations in the test). For $h > 1$, the series made up of $d_\tau = e_{1\tau}^2 - e_{2\tau}^2$ is likely to be autocorrelated, and the DM test therefore takes into account that the variance of the mean $\bar{d}$ is affected by the autocovariance of d_τ.

Diebold and Mariano (1995) showed simulation evidence suggesting that their test (DM) may be oversized (i.e., rejecting too often when the null hypothesis of no difference is in fact true) in the case of two-steps ahead forecasts, $h = 2$. Harvey and Newbold (1997) proposed a modification that makes use of an approximately unbiased estimator of the variance of $\bar{d}$. The simulation results in Harvey and Newbold (1997) also suggests strongly that the size properties of both the DM and the MDM tests are improved by using critical values from the Student's t distribution with $n - 1$ degrees of freedom rather than from the standard normal distribution. In line with this, the p-values reported in parenthesis below each entry in our table are from the t distribution.

In Table 1, negative values mean that the MPR squared errors are on average smaller than the alternative set of forecasts. We see that if a 5 % significance level is used, the 1-step forecasts ($h = 1$) of the MPR are more accurate than the random-walk forecasts, and with at 10 % level, the $h = 1$ MPR inflation forecasts are also more precise than the AIF.

As one would expect from the graphs of the MSFEs above, the mean of the loss differential series d_τ becomes positive for $h > 1$ in the test of MPR against AIF. For $h = 2$ and $h = 3$ the differences are insignificant, but for $h = 4, 5, 6$ the tests indicate the we can reject the null hypothesis of equality of squared errors, and accept the alternative that the AIF forecasts are more accurate. There are two factors that explain this outcome. First, the sizeable differences for these horizons noted above, and second that there are negative sample autocovariances in the relevant d_τ series. For the test of MPR against M1, the situation is that the numerical value $\bar{d}$, although sometimes large, is swamped by large estimated $Var(\bar{d})$'s, which is due to dominating positive autovariances in the d_τ series that constitute these tests. That said, for $h = 3$ the DM test is nevertheless significant in favour of MPR at the 5 % significance level, and the MDM test is significant at the 10 % level.

7 Conclusion and Discussion

Norges Bank is a leading forecast targeting central bank, and it is reasonable to expect that Norges Bank produces inflation forecasts that regularly outperform both 'naive' forecasts and forecasts from econometric models that are specified and maintained outside the central bank. When it comes to forecasting inflation, the forecast targeting central bank may be expected to know the most and to be best.

It is surprising therefore that the superiority of the Monetary Policy Report inflation forecasts cannot be asserted. This is the conclusion when MPR forecasts over the period 2004-2009 are compared with genuine *ex ante* real time forecasts from an outsider model called AIF. The 1-step MPR forecasts are preferable to the 1-step forecasts from AIF, but for several of the policy relevant horizons (e.g., 4 to 8 quarters ahead), the evidence goes in favour of the outsider model. Because there are fewer AIF forecast to average over, there will be more uncertainty about the forecast accuracy of AIF than about the MPR forecasts. This also affects the formal tests of equality of mean squared forecast errors, which nevertheless give some indication of significantly larger MPR errors for forecast horizons that are in the middle of the policy horizon. The informal evidence in the graphs also show that the situation has been quite stable since the comparisons started in 2004, which indicate that there are structural elements in MPR forecasting system that makes it difficult for Norges Bank's forecasters to win a contest like this one— which on the face of it should be an easy victory.

Based on the properties of the AIF model, one can hypothesize that the tendency of the MPR forecasts to adjust too quickly to the target will damage the forecast accuracy of the MPR. This hypothesis is supported by the evidence. The opposite hypothesis, that the AIF forecasts will suffer because it is not well adapted to forecast targeting regime that has been in operation since 2001, is not supported by the evidence presented above.

Following the earlier difficulties with forecasting inflation in 2003, Norges Bank has moved in the direction of ensemble forecast, and thereby Norges Bank place less weight than intended on the quantitative theoretical macro model NEMO. Also our results confirm that there can be gains from forecast averaging.

Even very inaccurate forecasts, like those from a random walk, would have positive weights, since they correct a bias in the two discipline forecasts.

We have compared forecasts for a period when Norwegian inflation was running at a much lower level than in the three last decades of the previuos millenium. This is the same development as in the US, and in other western economies. According to Stock and Watson (2007), the U.S. rate of inflation has become more difficult to predict after the Great Moderation. This is because the inflation process has become markedly less persistent than before, and is now dominated by short-run fluctuations. The findings in Edge and Gürkaynak (2010) are in line with this, since they find that their DSGE model is very poor at forecasting US inflation, but so are all the other approaches that they use (uni- and multivariate time series models).

As just noted, also Norwegian inflation has become less persistent than before. Nevertheless, our results show that in particular the forecasts from the econometric model AIR, but also MPR, produced better forecasts than the 'naive' univariate forecasting methods. This indicates that one way to 'maintain predictability' for a series that has become less persistent than before, is to keep the information from the period when persistence was high in the information set that underlies the forecasting model. Hence, there is an argument for using rather long time series in inflation forecasting, possibly from different regimes. We leave the further analysis of this conjecture, and how it can be reconciled with need to avoid that structural breaks damage the forecasts, to future work.

A   The AIF Forecasting Model

Our approach is to represent the effects of the labour market, the market for foreign exchange and the price setting in domestic product markets in a compact econometric model of Norwegian. We build on earlier modelling projects like the one documented in Bårdsen et al. (2003). The model can also be seen as a condensed version of the incomplete competition model of wage and price dynamics, see Bårdsen and Nymoen (2003).

Since the purpose the model is to forecast the annual rate of inflation we have modelled the fourth difference of the natural logarithm of CPI-ATE. It is denoted

Δp_t in this appendix. As mentioned in the main text, the specification of the "inflation function" has evolved during the time period of the AIF project. That said, the rate of unemployment (U_t), the logarithm of the real exchange rate (rex_t), foreign inflation $\Delta_4 p_{*t}$ and interest rates, domestic (i_t) and foreign (i_{*t}), have consistently made up the set of explanatory variables. In this appendix we document the current version of the model

A.1 The Inflation Equation

The inflation equation, used in the AIF forecasts published in February 2010,[13] is reported in equation (7).

$$
\begin{aligned}
\Delta_4 p_t \;=\; & \underset{(0.05)}{0.717}\ \Delta_4 p_{t-1} - \underset{(0.07)}{0.251}\ \Delta_4 p_{t-3} + \underset{(0.05)}{0.250}\ \Delta_4 p_{t-4} \\[4pt]
& + \underset{(0.0015)}{0.017} + \underset{(0.07)}{0.971}\ D_{p,t} \\[4pt]
& - \underset{(0.003)}{0.0019}\ U_{t-2} - \underset{(0.004)}{0.0016}\ U_{t-4} \\[4pt]
& - \underset{(0.018)}{0.073}\ (i_{t-4} - i_{*t-4}) - \underset{(0.024)}{0.115}\ \Delta i_{t-4} + \underset{(0.024)}{0.223}\ \Delta i_{*,t-4} \\[4pt]
& + \underset{(0.020)}{0.110}\ \Delta_4 p_{*t-2} + \underset{(0.004)}{0.032}\ rex_{t-4}
\end{aligned}
\tag{7}
$$

$$
\begin{aligned}
T = 1983q1 - 2009q4 = 108 & \qquad OLS \\
\hat{\sigma} \quad = \quad 0.19\% & \\
F_{AR1-5}(5,91) = 1.0125[0.42] & \qquad \chi^2_{\text{normality}}(2) = 0.10836[0.95] \\
F_{ARCH\ 1-4}(4,100) = 0.79160[0.53] & \qquad F_{x^2}(22,85) = 0.56915[0.93] \\
F_{x_i x_j}(76,31) = 0.61974[0.95] & \qquad F_{RESET}(2,94) = 2.1603[0.12]
\end{aligned}
$$

On the right hand side of equation (7) we first have the three autoregressive terms $\Delta_4 p_{t-1}$, $\Delta_4 p_{t-3}$ and $\Delta_4 p_{t-4}$. Conventional standard errors are in brackets below each coefficient, showing that they are statistically significant different from zero. The huge positive coefficient for $\Delta_4 p_{t-1}$ is not surprising. Since we model the annual rate of inflation, it represents inflation persistence. The third and the

---

[13] See http://folk.uio.no/rnymoen/forecast_air_index.html.

fourth lags of inflation are also significant, but the almost identically sized coefficients have opposite signs, showing that these autoregressive terms capture more high-frequency dynamics than the first lag.
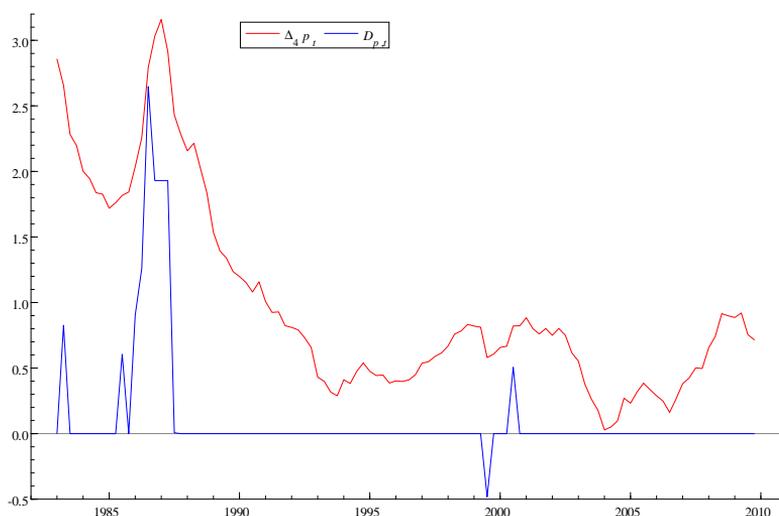

Figure 6: $\Delta_4 p_t$ and the composite structural break dummy $D_{p,t}$. (The dummy has been rescaled to match the range of $\Delta_4 p_t$).

The second line of (7) shows the deterministic part of the equation, consisting of an intercept and a composite dummy $D_{p,t}$. The dummy represents the estimated inflation effects of currency devaluations that occurred during the 1980's, but also the cost-push effect of the reduction in the length of the working week in January 1987, see Nymoen (1989). The dummies were obtained by the utility for automatic detection of breaks in *PcGive 13*, see Doornik and Hendry (2009a). Figure 6 shows $D_{p,t}$ together with $\Delta_4 p_t$. As expected, the structural break dummies are mainly located in the 1980s. There is however a drop in inflation in 1999q3 which is detected as a break, and a corresponding positive dummy coefficient a year later in 2000q3. There are no detected breaks at the time of the formal introduction of

inflation targeting in March 2001 or to events later in the sample period (notably the 'financial crisis').

The remaining part of equation (7) contains the economic explanatory variables. First, the rate of unemployment (U_t) enters with its second and fourth lag. This effect may reflect both that price setters adjust their mark-up with the business cycle (U_t is highly correlated with changes in GDP in Norway), but also indirect effects working through wage setting. Specifically, existing research has established an effect from unemployment on wages, both in a Phillips curve framework and in wage equations that can be rationalized from bargaining theory, see Bårdsen et al. (2005, Chapter 3-6) for an overview. The numerically and statistically significant effects in equation (7) are therefore consistent with several existing studies.

We interpret the fourth line in equation (7) as representing the inflationary impulses from the market for foreign exchange: Under quite general assumptions, the difference between the interest rate in the domestic, (i_t) and foreign money markets ($i_{*,t}$) affects the nominal exchange rate, see Bårdsen and Nymoen (2009) for an analysis. In equation (7) the interest rate differential manifests itself quite clearly (at the fourth lag) although we have estimated the two coefficients freely in this case.

In the fifth line of the inflation equation, we first include the direct effects of imported inflation, as measured by $\Delta_4 p_{*,t-2}$ the annual rate of change in foreign consumer prices, in foreign currency. Finally, the fourth lag of the logarithm of the real exchange rate (rex_{t-4}) is highly significant with a positive coefficient. Usually, a real depreciation of the currency, corresponding to a higher rex, will lead to a degree of 'internal revaluation' after a period of time, which is a relevant interpretation of the significance of rex_{t-4} in equation (7).

Below the equation we report the sample size (number of quarterly observations) by T, and the residual standard error in percent is denoted by $\hat{\sigma}$. We also report a set of mis-specification tests. As indicated by the notation, the normality test is a Chi-square test, while the others are F-distributed under their respective null hypotheses. They are: $F_{AR\,1-5}$, autoregressive residual autocorrelation, ARCH heteroskedastisity, $F_{ARCH\,1-4}$, and heteroskedasticity due to squares, F_{x^2}, and cross products of the regressors $F_{x_i x_j}$. Finally F_{RESET} tests the null hypothesis of no functional form mis-specification, see Doornik and Hendry (2009a).

Table 2: Marginal and definition equations of the AIF foecasting model. All equation are estimated by OLS . The sample period is 1983q1-2009q4)

The exchange rate

$$\Delta e_t = \underset{(0.3897)}{1.1} \; \Delta_4^2 p_t - \underset{(0.139)}{0.56} \; [i_t - i_{*t}] - \underset{(0.04)}{0.220} \; rex_{-1} - \underset{(0.0019)}{0.002} \; U_{t-1}$$

The rate of unemployment

$$U_t = \underset{(0.02)}{0.88} \; U_{t-1} - \underset{(1.04)}{4.62} \; [i_{t-3} - \Delta_4 p_{t-3}]$$

Foreign inflation

$$\Delta_4 p_{*t} = \underset{(0.08)}{1.32} \; \Delta_4 p_{*t-1} - \underset{(0.08)}{0.33} \; \Delta_4 p_{*t-2} + \underset{(0.002)}{0.012} \; \Delta_4 p_{ot} - \underset{(.002)}{0.013} \; \Delta_4 p_{ot-1}$$

Foreign interest rate

$$i_{*t} = \underset{(0.03)}{0.90} \; i_{*t-1} + \underset{(0.13)}{0.37} \; \Delta_4 p_{*t} - \underset{(0.13)}{0.17} \; \Delta_4 p_{*t-1} + \underset{(0.007)}{0.0013} \; \Delta_4 p_{ot}$$

Domestic interest rate

$$\Delta i_t = \underset{(0.06}{0.19} \; \Delta i_{t-1} - \underset{(0.04)}{0.23} \; i_{t-1} + \underset{(0.20)}{0.75} \; \Delta_4 p_t - \underset{(0.21)}{0.23} \; \Delta_4 p_{t-1}$$

$$- \underset{(0.001)}{0.003} \; U_t + \underset{(0.001)}{0.002} \; U_{t-1} + \underset{(0.09)}{0.15} \; i_{*t-1}$$

Oil price

$$\Delta_4 p_{ot-1} = \underset{(0.09)}{1.13908} \; \Delta_4 p_{ot-1} - \underset{(0.09)}{0.518669} \; \Delta_4 p_{ot-1}$$

A.2 Marginal Equations

All the stochastic variables in equation (7) are in turn modelled by econometric relationship, so that the AIFs are based on estimation and dynamic simulation of a system: There is one equation for each of the 5 explanatory variables in (7), and in addition the annual rate of change in the price of oil ($\Delta_4 p_{ot}$) is modelled because it is used in the equation for $\Delta_4 p_{*t}$, and the nominal exchange rate (e_t) is also modelled in order to forecast the real-exchange rate (*rex*) within the forecasting system.

Equations that make-up the AIF forecasting system together with the inflation equation (7), is given in Table 2. In order to highlight the relationships between variables, the equations are given in stylized form, for example without intercepts and dummies for structural breaks. The structural breaks are however found by the same method as explained in connection with the inflation equation.

Equation (7) and Table 2 is a system of 11 equations. The 11 endogenous variables are: $\Delta_4 p_t$, $\Delta_4^2 p_t$, $\Delta_4 p_{*t}$, $\Delta_4 p_{ot}$, Δe_t, $\Delta_4 e_t$, rex_t, i_t, Δi_t, i_{*t}, and U_t.

A.3 Data Definitions and Sources

The forecasting model employs seasonally unadjusted data. Unless another source is given, all data are taken from the RIMINI database in Norges Bank (The Central Bank of Norway). For each RIMINI-variable, the corresponding name in the RIMINI-database is given by an entry [RIMINI: variable name] at the end of the description. The RIMINI identifier is from Rikmodnotat 405, Norges Bank, Research department, 9th December 2002. Each variable's name in the PcGive batch file that generate the AIF forecasts is also given, with transformation indicated when relevant.

The variables used in the forecasting equation for the rate of inflation are defined as follows:

P Consumer price index adjusted for the influence of energy prices and indirect taxes, CPI-ATE. 1991=1. [RIMINI: CPIJAE].{PcGive: $cpijae = \ln(CPIJAE)$}.

P_c Consumer price index (CPI). 1991=1. [RIMINI: CPI].{PcGive: $p = \ln(CPI)$}.

P_* Consumer prices abroad in foreign currency. 1991=1. [RIMINI: PCK-ONK].{PcGive: $pck = \ln(PCKONK)$}.

P_o Oil price index (USD).[RIMINI: SPOILUSD].{PcGive: $oil = \ln(PCKONK)$}.

E Trade weighted nominal value of the krone based on import-shares of trading countries. [RIMINI: CPIVAL].{PcGive: $v = \ln(CPIVAL)$}.

Z Value added per man hour at factor costs in the Norwegian non-tradables sector, fixed baseyear (1991) prices. Mill. NOK. [RIMINI: ZYI].{PcGive: $zyi = \ln(ZYI)$}.

i Money market interest rate (3 month Euro-krone interest rate). [RIMINI: RS].{PcGive: RSH}.

i_* ECU interest rate. For the period 1967q1-1986q3: Effective interest rate on foreign bonds, NOK-basket weighted. [RIMINI: R.BKUR] For the period 1986q4-1996q4: ECU weighted effective rate on foreign bonds. [RIMINI: REC/100].{PcGive: RSW}.

U Labour force survey (AKU) rate of unemployment. [RIMINI: UAKU2].{PcGive: $UAKU2$}.

REX Real effective exchange rate. $REX = EP_*/P_C$.

References

Aastveit, K. A. (2010). *Empirical Studies of Business Cycles and Monetary Policy Regimes in Real Time*. Ph.D. thesis, Department of Economics, Univeristy of Oslo.

Akram, Q. F., and Nymoen, R. (2009). Model Selection for Monetary Policy Analysis—How Important is Empirical Validity? *Oxford Bulletin of Economics and Statistics*, 71: 35–68. URL http://ideas.repec.org/a/bla/obuest/v71y2009i1p35-68.html.

Bårdsen, G., Eitrheim, Ø., Jansen, E. S., and Nymoen, R. (2005). *The Econometrics of Macroeconomic Modelling*. Oxford: Oxford University Press.

Bårdsen, G., Jansen, E. S., and Nymoen, R. (2003). Econometric Inflation Targeting. *Econometrics Journal*, 6: 429–460. URL http://www.feweb.vu.nl/econometriclinks/journal/volume6/BaardsenJansenNymoen/ectjBardsen-2.pdf.

Bårdsen, G., and Nymoen, R. (2003). Testing Steady-State Implications for the NAIRU. *Review of Economics and Statistics*, 85: 1070–1075. URL http://www.jstor.org/stable/3211827.

Bårdsen, G., and Nymoen, R. (2009). Macroeconometric Modelling for Policy. In M. T.C., and K. Patterson (Eds.), *Palgrave Handbook of Econometrics Vol 2*, chapter 17, pages 851—916. Palgrave Mac-Millan.

Bates, J. M., and Granger, C. (1969). The Combination of Forecasts. *Operations Research Quarterly*, 20: 451–468. URL http://www.jstor.org/stable/3008764.

Bjørnland, H. C., Ekeli, T., Geraats, P., and Leitemo, K. (2004). Norges Bank Watch 2004. An Independent Review of Monetary Policy and Institutions in Norway. Discussion paper, Centre for Monetary Economics at the Norwegian School of Management, BI. URL http://www2.bi.no/cme/NBW%202004.pdf.

Brubakk, L., Husebø, T., Maih, J., Olsen, K., and Østnor, M. (2006). Finding NEMO: Documentation of the Norwegian Economy Model.

Staff Memo 06, Norges Bank. URL http://www.norges-bank.no/en/about/published/publications/staff-memo/staff-memo-2006/finding-nemo-documentation-of-the-norwegian-economy-model/.

Brubakk, L., and Sveen, T. (2009). NEMO-A New Macro Model for Forecasting and Monetary Policy Analysis. *Economic Bulletin*, 80(1): 39–47. URL http://www.norges-bank.no/en/about/published/publications/economic-bulletin/economic-bulletin-12008/nemo--a-new-macro-model-for-forecasting-and-monetary-policy-analysis/.

Clarida, R., Gali, J., and Gertler, M. (1999). The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature*, 37(4): 1661–1707.

Clements, M., and Hendry, D. F. (2008). Economic Forecasting in a Changing World. *Capitalism and Society*, 3(2): 1–17. URL http://wrap.warwick.ac.uk/890/.

Clements, M. P., and Hendry, D. F. (1999). *Forecasting Non-stationary Economic Time Series*. Cambridge, MA: The MIT Press.

Cristoffel, K., Coenen, G., and Warne, A. (2010). Forecasting With DSGE Models. Working Paper Series 1185, European Central Bank. Forthcoming in 'Oxford Handbook' on Economic Forecasting, edited by Michael P. Clements and David F. Hendry.

Diebold, F. X., and Mariano, R. (1995). Comparing Predictive Accuray. *Journal of Business and Economic Statistics*, 13: 253–263. URL http://www.jstor.org/stable/1392185.

Doornik, J. A., and Hendry, D. F. (2009a). *Empirical Econometric Modelling PcGive 13. Volume 1*. London: Timberlake Consultants.

Doornik, J. A., and Hendry, D. F. (2009b). *Modelling Dynamic Systems PcGive 13. Volume 2*. London: Timberlake Consultants.

Edge, R., and Gürkaynak, R. (2010). How Useful Are Estimated DSGE Model Forecasts for Central Bankers? *Brookings Papers on Economic Activity*, 2: 209–259. URL http://ideas.repec.org/a/bin/bpeajo/v41y2010i2010-02p209-259.html.

Eitrheim, Ø., Jansen, E. S., and Nymoen, R. (2002). Progress from Forecast Failure: The Norwegian Consumption Function. *Econometrics Journal*, 5: 60–64. URL http://ideas.repec.org/a/ect/emjrnl/v5y2002i1p40-64.html. Forthcoming.

Gerdrup, K., Jore, A., Smith, C., and Thorsrud, L. (2009). Evaluating Ensemble Density Combination—Forecasting GDP and Inflation. Working Paper 19, Norges Bank. URL http://www.norges-bank.no/upload/english/publications/working%20papers/2009/norges_bank_working_paper_2009_19.pdf.

Granger, C. W. J. (Ed.) (1999). *Empirical Modeling in Economics. Specification and Evaluation*. Cambridge: Cambridge Universty Press.

Hammond, G. (2010). *State of the Arte of Inflation Targeting*. Bank of England. CCBS Handbook No. 29. http://www.bankofengland.co.uk/education/ccbs/handbooks/ccbshb29.htm.

Harvey, S. L., D., and Newbold, P. (1997). Testing the Equality of Mean Squared Errors. *International Journal of Forecasting*, 13: 281–291. URL http://ideas.repec.org/a/eee/intfor/v13y1997i2p281-291.html.

Hendry, D. F. (2001). How Economists Forecast. In N. R. Ericsson, and D. F. Hendry (Eds.), *Understanding Economic Forecasts*, chapter 2, pages 15–41. Cambridge, Massachusetts: The MIT Press.

Husebø, T., McCaw, S., Olsen, K., and Røisland, Ø. (2004). A Small Calibrated Macromodel to Support Inflation Targeting at Norges Bank. Staff Memo 3, Norges Bank, Oslo. URL http://www.norges-bank.no/Upload/Publikasjoner/Staff%20Memo/2004/memo-2004-03.pdf.

Juel, S., Molnar, K., and Røed, K. (2008). Norges Bank Watch 2008. An Independent Review of Monetary Policymaking in Norway. Discussion paper,

Centre for Monetary Economics at the Norwegian School of Management, BI. URL http://www.bi.no/cmeFiles/NBW2008%20report.pdf.

Nymoen, R. (1989). Wages and the Length of the Working Day. A Empirical Test Based on Norwegian Quarterly Manufacturing Data. *Scandinavian Journal of Economics*, 91: 599–612. URL http://ideas.repec.org/a/bla/scandj/v91y1989i3p599-612.html.

Nymoen, R. (2005). Evaluating a Central Bank's Recent Forecast Failure. Memorandum No 22/05, Department of Economics, University of Oslo.

Nymoen, R., and Tveter, E. (2007). Er Norges Banks pengepolitiske modell en god nok modell for norsk økonomi? *Økonomisk Forum*, 61(5): 45–54.

Olsen, Ø. (2011). Use of models and economic theory in Norges Bank. The Scweigaard Lecture 2011. Lecture by Governor Øystein Olsen at the Department of Economics, University of Oslo. http://www.norges-bank.no/en/about/published/speeches/2011/schweigaard-lecture/.

Orphanides, A. (2003). Monetary Policy Evaluation with Noisy Information. *Journal of Monetary Economics*, 50(3): 605–631. URL http://ideas.repec.org/a/eee/moneco/v50y2003i3p605-631.html.

Rubaszek, M., and Skrzypczynski, P. (2008). On the Forecasting Performance of a Small Scale DSGE Model. *International Journal of Forecasting*, 24: 498–512. URL http://ideas.repec.org/a/eee/intfor/v24y2008i3p498-512.html.

Stock, J. H., and Watson, M. W. (2007). Why Has U.S. Inflation Become Harder to Forecast ? *Journal of Money, Credit and Banking*, 39(1): 3–33. URL http://ideas.repec.org/a/mcb/jmoncb/v39y2007is1p3-33.html.

Svensson, L. (1997). Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets. *European Economic Review*, 41: 1111–1146. URL http://ideas.repec.org/a/eee/eecrev/v41y1997i6p1111-1146.html.

Svensson, L., Houg, K., Solheim, H. O. A., and Steigum, E. (2002). Norges Bank Watch 2002. An Independent Review of Monetary Policy and Institutions

in Norway. Discussion paper, Centre for Monetary Economics at the Norwegian School of Management, BI. URL http://www2.bi.no/cme/NBW2002.pdf. Http://www2.bi.no/cme/NBW2002.pdf.

Svensson, L. E. O., and Woodford, M. (2005). Implementing Optimal Policy through Inflation-Forecast Targeting. In B. B.S., and M. Woodford (Eds.), *The Inflation Targeting Debate*, chapter 2. University of Chicago Press. URL http://ideas.repec.org/h/nbr/nberch/9556.html.

Wang, M. C. (2009). Comparing the DSGE model with the Factor Model: An Out of Sample Forecasting Experiment. *Journal of Forecasting*, 28: 167–182. URL http://econpapers.repec.org/article/jofjforec/v_3a28_3ay_3a2009_3ai_3a2_3ap_3a167-182.htm.

Woodford, M. (2007). The Case for Forecast Targeting as a Monetary Policy Strategy. *Journal of Economic Perspectives*, 21: 3–24. URL http://ideas.repec.org/a/aea/jecper/v21y2007i4p3-24.html.

Please note:

You are most sincerely encouraged to participate in the open assessment of this article. You can do so by either recommending the article or by posting your comments.

Please go to:

http://dx.doi.org/10.5018/economics-ejournal.ja.2011-15

The Editor