Empirical comparison of inflation models’ forecast accuracy

Øyvind Eitrheim, Tore Anders Husebø and Ragnar Nymoen

2nd July 2000

Abstract

Forecasting inflation is important in practical monetary policy, and increasingly so in countries who have adopted inflation targeting as their operative monetary policy regime. Inflation models are however contested. Rival models are shown to have different forecasting properties over a period that covers a change in regime in the Norwegian economy—from high to low inflation. There is no ground for choosing a policy model on the basis of forecast performance alone, but different models can be useful for different purposes. For example, we find that simple forecasting rules based on differencing are relatively robust to structural breaks. Thus they contain information that can be used to intercept-correct a larger model that contains policy relevant causal information.

*The first author is head of research and the second is staff economist in Norges Bank (The Central Bank of Norway). The third author is professor of economics at the University of Oslo and special consultant in Norges Bank. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting those of Norges Bank. We would like to thank Michael P. Clements for advise and suggestions. Please address correspondence to the third author: Ragnar Nymoen, University of Oslo, Department of Economics. P.O. Box 1095 Blindern. N-0317 Oslo, Norway. Internet: ragnar.nymoen@econ.uio.no.
1 Introduction

Producers and consumers of empirical models take a shared interest in comparison of model forecasts. As pointed out by Granger (1990), consumers of models care about out-of-sample model properties and consequently put weight on comparisons of model forecasts. Since producers of models in turn wish to influence the beliefs of model consumers, comparison of model forecasts provides an important interface between producers and consumers of empirical models.

A comparison of forecast from empirical models can be based on “raw” model forecasts (e.g. provided by the producers) or on the published forecasts that include the effects of judgmental corrections (intercept corrections). In this paper the aim is to focus on the mapping from model specification to forecast properties, so we consider the raw model forecast, without the intervening corrections made by forecasters. Another issue is that any comparison will inevitably consider only a sub-set of macroeconomic variables, and the choice of variables will influence the outcome of the comparison. This problem extends to whether it is levels variables that are forecasted or their growth rates, since the ranking of forecast accuracy may depend on which linear transformation one uses, see Clements and Hendry (1993), Clements and Hendry (1998, Chapter 3).

In the following, these issues are pushed somewhat in the background by focusing on inflation forecasting. The vector of variables that enter the comparison is the “typical” list of variables that are forecasted in the central banks’ economic bulletins and outlooks. The typical list include a limited number of annual growth rates: inflation (in Norway, this is CPI inflation), wage growth, import price growth, GDP growth, growth in consumer expenditure, housing price growth (asset inflation), but also at least one levels variable, namely the rate of unemployment.

In the 1990s, inflation targeting has emerged as a candidate intermediate target for monetary policy. An explicit inflation target for monetary policy means a quantified inflation target, e.g. 2 per cent per year, and a tolerance interval around it of (for example) ±1 percentage point, and that the central bank is given full control of monetary instruments. New Zealand and Canada are the pioneering countries. Sweden moved to inflation targeting in 1993, and since 1997 an inflation target has represented the nominal anchor of the UK economy.

As emphasized in Svensson (1997b), an explicit inflation target implies that the central bank’s conditional forecasts 1-2 years ahead become the intermediate target of monetary policy. If the inflation forecast is sufficiently close to the target, the policy instruments (a short-term interest rate) is left unaltered. If the forecasted rate of inflation is higher (lower) than the target, monetary instruments are changed until the revised forecast is close to the inflation target. In such instances the properties of the forecasting model (the dynamic multipliers) can have a large influence on how much the interest rate is changed.

It is seen that with the conditional inflation forecasts as the operational target of monetary policy, there is an unusually strong linkage between forecasting and policy analysis. Decisions are more explicitly forward looking than in other instances of macroeconomic policy making, where the assessment of the current economic situation plays a prominent role. That said, inflation forecasts are also important elements in policy discussions also under exchange rate “regimes” other than explicit inflation targeting. Thus, Svensson (1997a) notes that inflation forecasts prepared by Norges Bank are “more explicit and detailed than Sveriges Riksbank’s forecast”,
even though Norway has no formal inflation target.

In this chapter we compare forecasts that made for policy purposes. Section 2 gives the economic background and discusses the relationship between model specification and inflation forecasting and section 3 provides an empirical investigation example of how the models perform econometrically and in forecasting inflation. In section 4 we outline four different forecasting models for the Norwegian economy. First, we present two large scale macroeconometric models which contain monetary policy channels (transmission mechanisms) linking monetary policy instruments like short run money market interest rates and exchange rates to other economic variables (e.g., those listed above), through causal mechanisms which allow for monetary policy analysis. Second, we also look at two simple non-causal forecasting models in differences. Section 5 discuss the forecasting properties of the four models using stochastic simulation.

2 Model specification and inflation forecasting

One of the inflation models with the longest track record is the Phillips curve. It was integrated into macroeconometric models in the 1970s. In the mid 1980s, however, the Phillips curve approach has been challenged by a model consisting of a negative relationship between the level of the real wage and the rate of unemployment, dubbed the wage curve by Blanchflower and Oswald (1994), together with firms’ price setting schedule. The wage curve is consistent with a wide range of economic theories, see Blanchard and Katz (1997), but its original impact among European economists was due the explicit treatment of union behaviour and imperfectly competitive product markets, see Layard and Nickell (1986), Rowlatt (1987), Hoel and Nymoen (1988). Because the modern theory of wage and price setting recognizes the importance of imperfect competition on both product and labour markets, we refer to this class of models as the Imperfect Competition Model—ICM hereafter.

In equation (2.1) $pc_t$ denotes the log of the consumer price index in period $t$. $\Delta$ is the first difference operator, so $\Delta pc_t$ is CPI-inflation:

$$
\Delta pc_t = \gamma_1(u_{t-1} - u^n) + \gamma_2'(L)z_t - \alpha_n \sum_{i=1}^n EC_{it-1} + \varepsilon_t
$$

The first term on the right hand is the (log of) the rate of unemployment ($u_{t-1}$) minus its (unconditional) mean $u^n$, hence $E[u_{t-1} - u^n] = 0$. It represents “excess demand” in the labour market and how wage increases are transmitted on to CPI-inflation. Empirical equations often include a product market output-gap variable alongside the unemployment term. However, this variable is not needed in order to discriminate between theories, and is omitted from (2.1)

The term $z_t$ is a vector of variables that enter in differenced form and $\gamma_2'(L)$ is the corresponding vector polynomial with coefficients. Typical elements in this part of the model are the lagged rate of inflation, i.e., $\Delta pc_{t-1}$, the rate of change in import prices and changes in indirect tax-rates.

The case where the remaining coefficients in the equation are zero, i.e. $\alpha_1 = ... = \alpha_n = 0$, corresponds to the Phillips-curve model. The Phillips-curve is an

---

1Wallis (1993) gives an excellent exposition, emphasizing the impact on macroeconometric modelling and policy analysis.
important model in current macroeconomics, for example in the theory of monetary policy as laid out in Clarida and Gertler (1999), and it dominates the theoretical literature on inflation targeting, see Svensson (2000). The Bank of England (1999) includes Phillips-curve models in their suite of models for monetary policy. Mervyn King, the Deputy Governor of the Bank of England puts it quite explicitly: ‘..the concept of a natural rate of unemployment, and the existence of a vertical long-run Phillips curve, are crucial to the framework of monetary policy.’

The empirical literature shows that the Phillips curve holds its ground when tested on US data—see Fuhrer (1995), Gordon (1997), Gali and Gertler (1999), and Blanchard and Katz (1999). Studies from Europe usually conclude differently: The preferred models tend to imply a negative relationship between the real wage level and the rate of unemployment, see e.g. Drèze and Bean (1990, Table 1.4), OECD (1997, Table 1.A.1), Wallis (1993) and Rødseth and Nymoen (1999). These findings are consistent with the seminal paper of Sargan (1964), and the later theoretical developments leading to the ICM class of wage and price setting equations.

The negative relationship between the real wage level and the rate of unemployment, which is also referred to in the literature as the *wage-curve*, see Blanchflower and Oswald (1994), can be incorporated in (2.1) in the following way: Let $w_t$ denote the log of the nominal wage rate in period $t$, so $w_t - p_t$ is the real-wage, then the equilibrium correction term $EC_{1,t}$ can be specified as

$$EC_{1,t} = w_t - pc_t + \beta_{11}u_t - \mu_1, \quad \text{with } \alpha_1 \geq 0 \text{ and } \beta_{11} \geq 0. \quad (2.2)$$

Thus, the future rate of inflation is influenced by a situation “today” in which the real-wage is high relative to the “equilibrium” real-wage $\beta_{11}u_T + \mu_1$. The parameter $\mu_1$ denotes the (long run) mean of the relationship, i.e. $E[EC_{1,t}] = 0$. Disequilibria in firms’ price setting have similar implications, and the inflation equation therefore contains a second term $EC_{2,t}$ that relates the price level in period $T$ to the equilibrium price level. Hence, in a simple specification,

$$EC_{2,t} = (pc_t - w_t) - \beta_{12}(pi_t - pc_t) + \mu_2 \quad \text{with } \alpha_2 \geq 0 \text{ and } \beta_{21} \geq 0, \quad (2.3)$$

where $pi_t$ is the log of the import price index.

It is seen that both the Phillips-curve and the ICM theories focus on the labour and product markets. Although, in theory disequilibria in other markets, such as the money market and the market for foreign exchange might have predictive power for inflation, we only consider the case of $n = 2$ in equation (2.1). Thus, the Phillips-curve equation takes the simple form

$$\Delta pc_t = \gamma_1(u_{t-1} - u^n) + \gamma_2'(L)\Delta z_t + \varepsilon_{phil,t}, \quad (2.4)$$

while an inflation equation consistent with the ICM is given by

$$\Delta pc_t = \gamma_2'(L)\Delta z_t + \alpha_1 EC_{1,t-1} + \alpha_2 EC_{2,t-1} + \varepsilon_{ICM,t}. \quad (2.5)$$

In (2.5) the unemployment term $(u_{t-1} - u^n)$ is omitted because, according to theory, unemployment creates inflation via wage-setting. Thus, if the wage-curve implicit in (2.2) is correctly modelled, there is no additional predictive power arising from an inclusion of $(u_{t-1} - u^n)$ in the equation, see Kolsrud and Nymoen (1998) for a discussion.

\footnote{King (1998, p.12)}
The two models take different views on the causal mechanisms in the inflation process, and as a result, they can lead to conflicting policy recommendations. Consider for example a situation where both models forecast a rise in inflation, e.g. because of a sudden rise in domestic spending. Based on the Phillips-curve, one might recommend a rise in the Bank’s interest rate, since otherwise it would take several periods before unemployment rises enough to curb inflation. Moreover, since any departure of unemployment from its natural-rate is temporary, one might as well provoke a temporary rise in \( u_t \), “today”, in order to cut-off inflation pressure directly. The ICM model points to other mechanism than unemployment that can stabilize inflation. For example, in an open economy, a rise in inflation leads to a fall in profits, and wage claims will be reduced as a result even at the going rate of unemployment. According, to the ICM it is also possible that a rise in the interest rate has lasting effects on the rate of unemployment, i.e., the natural-rate may not be invariant to policy changes, see Kolsrud and Nymoen (1998), Bårdsen et al. (2000). Thus, the recommendation could be a more moderate rise in the interest rate.

However, equation (2.4) and (2.5) have in common that they include causal information about the effects of other variables on inflation. In this respect they stand apart from univariate time series models which only include causal information in the form of lagged values of inflation itself. A simple example is given by

\[
\Delta p c_t = \gamma_0 + \gamma_3 p \Delta p c_{t-1} + \varepsilon_{AR,t},
\]

which is a 1th order autoregressive model of inflation. Finally, we consider forecasting models of the random-walk type, i.e.,

\[ \Delta \Delta p c_t = \varepsilon_{dAR,t}. \]  

We use the acronym \( dAR \) for the disturbance in (2.7), since an interpretation of (2.7) is that it is an autoregression in differences, obtained by setting \( \gamma_3 p = 1 \) in (2.6).

A common tread in many published evaluations of forecasts is the use of time-series models as a benchmark for comparison with forecasts derived from large-scale econometric systems of equations, see Granger and Newbold (1986, Chapter 9.4) for a survey. The finding that the benchmark-models often outperformed the econometric models represents an important puzzle that has not been fully resolved until recently, by the work of Michael Clements and David Hendry. In short, the solution lies in the insight that e.g., the random-walk model is not the “naive” forecasting tool that it appears at first sight. Instead, its’ forecasts are relatively robust to some types of structural changes that occurs frequently in practice, and that are damaging to forecasts derived from econometric models.

Equations (2.4)-(2.7) are special cases of the general equation (2.1). Assume now that the disturbances \( \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T \) in (2.1) represent an innovation process relative to the information available at time \( T \), denoted \( \mathcal{I}_T \). A strategy of choosing a forecasting model for period \( T + 1 \) is to first estimate (2.1) on the sample \( (t = 1, 2, \ldots, T) \) and then test the validity of the restrictions of the different models. Only the model which is valid reductions of (2.1) will have disturbance that are innovations relative to the information set \( \mathcal{I}_T \). The conditional mean of that congruent model the predictor of inflation in period \( T + 1 \) that has minimum mean-squared forecast error (MMSFE), see e.g. Clements and Hendry (1998, Chapter 2.7).
However, this result implicitly assumes that the process that we forecast is stable over the forecast horizon. But experience tells us that parameters frequently change. Suppose for example that a "regime shift" occurs in period $T + 1$. The congruent model then no longer reflects the true process in the forecast period, and thus we cannot use the above theorem to show that its conditional mean is the MMSFE forecast. Conversely, simple univariate models like (2.7) are unlikely to be congruent representations. Despite this, they offer some degree of protection against damage caused by non-stationarities.

Clements and Hendry (1998), (1999) have developed the theory of forecasting economic time series to account for instability and non-stationarity in the processes. One important results is that there is no way of knowing a priori which model will have the best forecast properties, an econometric model that include relevant causal information or a simple random walk model like (2.7), see e.g., Clements and Hendry (1998, Chapter 2.9). In our case, the econometric ICM may be the encompassing model, but if the parameters of the inflation process alters in the forecast period, its forecasts may still compare unfavourably with the forecasts of simple time series models.

As an example, suppose that possibility that the parameters of the wage curve $\mu_1$ changes in period $T + 1$, i.e., there is a shock to wage-setting. The ICM-forecast $\mathbb{E}[\Delta pc_{T+2} \mid \mathcal{I}_{T+1}]$ then becomes biased. Moreover, in the case that $\mu_1$ changes within sample (in period $T$ say), and that change is undetected by the forecaster, the ICM model will produce a biased forecast for period $T + 1$, while the random-walk forecast from (2.7) may be unbiased. Thus the better model in terms of causal information actually loses in a comparison to the random walk on this measure of forecast accuracy. Moreover, it is also possible that the Phillips-curve model outperforms the ICM. The reason is that the Phillips-curve shares the 1-step forecast properties of the random walk in this case, since it omits the $EC_{1, t-1}$ term that is affected by the structural break.

So far we have only considered 1-step forecasts, but it is clear that the same issues arise for dynamic multi-step forecasts. For example, in the Phillips-curve and in the ICM model, $u_{T+1}$ has to be forecasted in order to calculate $\mathbb{E}[\Delta pc_{T+2} \mid \mathcal{I}_T]$. Hence a larger econometric model is needed for forecasting, and if a structural break occurs in those parts of the economy that determine $u_{T+1}$, the inflation forecast is damaged. Conversely, simple univariate forecasting tools are by construction insulated from structural changes elsewhere in the system. Also, in that sense they produce robust forecasts.

Systems of equation that are developed with econometric methods are referred to as equilibrium-correcting models, EqCMs. The generalization of the simple autoregressive inflation models in (2.6) and (2.7) are system of equation that only use differences of the data, without equilibrium correction terms, i.e., VARs in differences (DV$s$, Clements and Hendry (1999, Chapter 5)) and double-differences (DDV$s$).

The Phillips-curve and the ICM are examples of EqCMs. At first sight it may seem that this tag only applies to the ICM, since the terms $EC_{1, t-1}$ and $EC_{2, t-1}$ are omitted from the Phillips-curve. However, the Phillips-curve conveys the alternative view that inflation is stabilized by $u_t - u^a \rightarrow 0$ in steady-state. Thus the equilibrating term is the rate of unemployment itself, rather than $EC_{1, t-1}$ and $EC_{2, t-1}$.

In the following, these issues are investigated by comparison of the forecasts of the different model specifications: ICM, Phillips-curve, univariate autoregression in differences and double differencing forecasting rules. In the next section we compare
small scale models of inflation are evaluated econometrically prior to forecast comparison. In section 4 we investigate different version of the macroeconomic models used by the Central Bank of Norway.

3 Two inflation models

In this section we compare forecast of the two contending inflation quarterly models over the period 1995.1-1998.4. The estimation sample is 1968.1-1994.4. The sample-split coincides with an important change in the Norwegian economy: The move from a high-inflation regime to a new regime with low and stable inflation. The means of the annual CPI growth rate of are 6.7% (estimation sample) and 2.1% (forecast period). The corresponding standard deviations are 3% and 0.6%. Thus the experiment is relevant for elucidating how well the different models forecast the new regime, conditional on the old regime.

A VAR for the five endogenous variables \( \Delta w_t, \Delta pc_t, \Delta p_t, \Delta y_t, \Delta u_t \) was estimated with data 1995.1-1998.4 \((N = 108)\). The definitions of the variables are:

- \( w_t = \log \) of nominal wage cost per man hour in Norwegian manufacturing
- \( y_t = \log \) of manufacturing value-added (fixed prices) per man hour.
- \( p_t = \log \) of the manufacturing value-added deflator.
- \( u_t = \log \) of the rate of unemployment.
- \( pc = \log \) of the consumer price index.

The equilibrium correction mechanism in equations (3.1) and (3.2) were taken as known, thus.

\[
\begin{align*}
w_t &= p_t + y_t - 0.08u_t + EC_{1,t}, \\
pct &= 0.6(w - y)_t + 0.4pi_t + \tau_3t + EC_{2,t}.
\end{align*}
\]

\(pi\) is log of the import price index of manufactures, and \(\tau_3t\) is an indirect tax-rate. These two equations are the empirical counterparts to (2.2) and (2.3). The estimates are taken from Bjørnstad and Nymoen (1999), and are also consistent with the findings for annual data in Johansen (1995). Equation (3.2) is from Bårdsen et al. (1998).4

The unrestricted system is large (43 coefficients in each equation), due to 4. order dynamics and the inclusion of 18 non-modelled variables, e.g., current and lagged import price growth \(\Delta pi_t\), the change in normal working hours in manufacturing \(\Delta h_t\), a variable that captures the coverage of labour markets programmes \((prog_t)\), and changes in the payroll and indirect tax-rates \((\Delta \tau_{1,t} \text{ and } \Delta \tau_{3,t})\). The deterministic terms include intercepts, three centered seasonal dummies, incomes policy dummies for 1979 and 1988, and a VAT dummy for the 1970q1. Finally, dummies that capture both deterministic shifts in the mean of the rate of unemployment as well as a changing seasonal pattern, see Akram (1999).

\footnote{All results in this section were obtained by GiveWin 1.20 and PcFIML 9.20, see Doornik and Hendry (1996a) and Doornik and Hendry (1996b).}

\footnote{However, this relationship was obtained for total economy wages and productivity, and therefore cointegration may be weaker for our manufacturing wage cost data.}
Table 3.1: Diagnostics for the ICM-system and model.

<table>
<thead>
<tr>
<th>Diagnostic tests for the VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\sigma}_{\Delta w}$ = 0.0113</td>
</tr>
<tr>
<td>$\hat{\sigma}_{\Delta pc}$ = 0.0054</td>
</tr>
<tr>
<td>$\hat{\sigma}_{\Delta p}$ = 0.0299</td>
</tr>
<tr>
<td>$\hat{\sigma}_{\Delta y}$ = 0.0277</td>
</tr>
<tr>
<td>$\hat{\sigma}_{\Delta u}$ = 0.084</td>
</tr>
<tr>
<td>$vAR\ 1 - 5\ F(125, 182) = 1.94[0.53]$</td>
</tr>
<tr>
<td>$v\ Normality\ \chi^2(10) = 20.00[0.03]$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnostic tests for the ICM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overidentification $\chi^2(152) = 175.3[0.10]$</td>
</tr>
<tr>
<td>$vAR\ 1 - 5\ F(125, 334) = 0.99[0.52]$</td>
</tr>
<tr>
<td>$v\ Normality\ \chi^2(10) = 18.0[0.06]$</td>
</tr>
</tbody>
</table>

Notes:
Sample period: 1968.1-1994.4. The VAR is estimated by OLS, The model estimated by FIML, p-values in brackets.

The upper part of Table 3.1 shows residual properties of the estimated VAR. The residual standard errors of wage growth ($\hat{\sigma}_{\Delta w}$) and inflation ($\hat{\sigma}_{\Delta pc}$) are similar to what we expect from earlier studies, cf., Nymoen (1989a) and Bårdsen et al. (1999). The estimated standard errors for product price growth and productivity are both close to 3%, and they are clearly going to induce a lump of uncertainty in inflation forecasts based on this system. For the rate of unemployment, $\hat{\sigma}_{\Delta u}$ = 0.084 corresponds to a residual standard error of 0.18 percentage points (using that the sample mean of the rate of unemployment is 2.2%). There is no evidence of vector residual autocorrelation, as shown by the joint test of absence of 5th order autocorrelation. However, the vector normality test is significant at the 5% level. In the lower part of Table 3.1, Overidentification $\chi^2(152)$ shows that the overidentifying restrictions implied by the ICM-model, are jointly data acceptable. That the ICM encompasses the VAR is also confirmed by the (vector) test statistics for absence of residual autocorrelation and non-normality.

Next, Table 3.2 focuses on the FIML estimates for the wage and price equations of the ICM. The equation for manufacturing wages resembles earlier models of this variable, see Nymoen (1989a), (1989b). Structural break-dummies are present in the form of $WD_t$, which captures the effects of wage-freeze periods and the highly centralized settlements in 1988 and 1989, and a dummy for the devaluation of the Norwegian currency in May 1986. The price equation includes the devaluation dummy, and a dummy for the introduction of VAT ($i70q1$), price freeze ($i79q1$) and seasonal dummies ($S_{t,i} (i = 1, 2, 3$). When we replace $\Delta w_t$ by the right hand side of the wage equation, we obtain the empirical counterpart to equation (2.1) in section 2. The two equilibrium correction terms in particular are significant, and imposition of the Phillips-curve restrictions on the model produces a highly significant Chi-square statistic: $\chi(2) = 43.59[0.0000]$.

Next, consider a “Phillips-curve VAR”, i.e., we start from a VAR that omits the equilibrium correction terms while retaining $u_{t-1}$ as an unrestricted variable. Table 3.3 gives the diagnostics for the system and the model, and Table 3.4 shows the estimated Phillips-curve wage equation together with $\Delta pc_t$ equation.
Table 3.2: The wage and price equations of the estimated ICM.

<table>
<thead>
<tr>
<th>Equation</th>
<th>Coefficients</th>
<th>Standard Errors</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The wage equation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta w_t = 0.341 + 0.512 \Delta_3 pc_{t-1} - 0.199 \Delta_3 w_{t-1} + 0.063 \Delta_3 pi_{t-1} + 0.074 \Delta_4 y_{t-1}$</td>
<td>(0.007)</td>
<td>(0.04)</td>
</tr>
<tr>
<td></td>
<td>$+ 0.494 \Delta_1 t_{t-2} - 0.023 \Delta u_{t-1} - 0.045 WD_t$</td>
<td>(0.143)</td>
</tr>
<tr>
<td></td>
<td>$+ 0.023 \times 86 q_3_{t} - 0.078 [(w - p - y)<em>{t-4} + 0.08 u</em>{t-1}]$</td>
<td>(0.011)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_{\Delta w} = 0.0124$</td>
<td></td>
</tr>
<tr>
<td><strong>The price equation</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\Delta p_c_t = -0.065 + 0.079 \Delta w_t + 0.167 \Delta_3 pc_{t-1} - 0.041 \Delta y_{t-1}$</td>
<td>(0.015)</td>
<td>(0.025)</td>
</tr>
<tr>
<td></td>
<td>$- 0.031 [p_{t-1} - 0.6 (w_{t-1} - y_{t-2}) - 0.4 pi_{t-1} - 1.71 t_{t-1}]$</td>
<td>(0.007)</td>
</tr>
<tr>
<td></td>
<td>$+ 0.037 (\Delta pi_{t-1} + \Delta pi_{t-4}) - 0.007 \Delta u_{t-2} - 0.005 \Delta u_{t-5}$</td>
<td>(0.011)</td>
</tr>
<tr>
<td></td>
<td>$+ 0.114 \Delta 3 t_{t-2} + 0.0397 q_1 t_{t} - 0.028 i 79 q_1 t_{t}$</td>
<td>(0.034)</td>
</tr>
<tr>
<td></td>
<td>$+ 0.067 i 86 q_3_{t} + 0.022 S_{1, t} + 0.013 S_{2, t} + 0.008 S_{3, t}$</td>
<td>(0.005)</td>
</tr>
<tr>
<td></td>
<td>$\hat{\sigma}_{\Delta p_c} = 0.0049$</td>
<td></td>
</tr>
</tbody>
</table>

Notes:
Sample period: 1968.1-1994.4. Estimation by FIML, jointly with equations for $\Delta p_t$, $\Delta y_t$ and $\Delta u_t$. Standard errors are in parantheses below the estimates.

The Phillips-curve wage equation has actually a little lower estimated residual standard error than its ICM counterpart. The equation implies a natural rate of unemployment $(u^n$ in 2.4) 1.5%, which may seem to be low, but consider that the average rate of unemployment over the sample period is only 2.2%.

We first compare the forecasts of the models for the 16-quarter period from 1995.1 to 1998.4. The forecasts are dynamic and are conditional on the actual values of the non-modelled variables: Import prices, the payroll and indirect tax-rate and labour market policy stance.

In Figure 3.1 the two graphs in the first row are for the annual rate of inflation $\Delta_4 pc_t$, the ICM on the left and the Phillips-curve on the right. The two bottom graphs compare the wage growth, $\Delta_3 w_c_t$, forecasts of the two rival models. Each graph in Figure 3.1 also contains the 95% prediction intervals in the form of ±2 standard errors, as a direct measure of the uncertainty of the forecasts. The inflation forecasts differ both in terms of bias and uncertainty. The Phillips-curve systematically over-predicts the rate of inflation, and the Phillips-curve prediction intervals appear to overstate the degree of uncertainty in inflation forecasting, see Sgherri and Wallis (1999) for a similar finding on UK inflation forecast uncertainty.

We next considered the following sequence of forecasts: First the 12-period forecast for 1996.1 to 1998.4 conditional on information up to and including 1995.4, then the 8 period and 4 period forecasts.

Figure 3.2 shows the forecasts for $\Delta_4 pc_t$ over all three horizons. Both forecasts overpredict significantly in 1996q1. In that quarter there was a reduction in the excises on cars (which explains around 40 per cent of the overprediction), so this is an example of a non-constancy in the process in the forecast period. Interestingly, the 8 quarter horizon shows that both models are “back on track” in 1997q1, i.e.,
Table 3.3: Diagnostics for the system and the Phillips curve model.

<table>
<thead>
<tr>
<th>Diagnostic tests for the Phillips-curve VAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\sigma}_{\Delta w} = 0.0188$</td>
</tr>
<tr>
<td>$\hat{\sigma}_{\Delta pc} = 0.0533$</td>
</tr>
<tr>
<td>$\hat{\sigma}_{\Delta p} = 0.0302$</td>
</tr>
<tr>
<td>$\hat{\sigma}_{\Delta \mu} = 0.0276$</td>
</tr>
<tr>
<td>$\hat{\sigma}_{\Delta u} = 0.0836$</td>
</tr>
<tr>
<td>$\nu AR 1 - 5 F(125, 191) = 0.977[0.55]$</td>
</tr>
<tr>
<td>$\nu Normality \chi^2(10) = 21.84[0.02]$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Diagnostic tests for the Phillips-curve model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overidentification $\chi^2(139) = 159.1[0.11]$</td>
</tr>
<tr>
<td>$\nu AR 1 - 5 F(125, 329) = 0.977[0.55]$</td>
</tr>
<tr>
<td>$\nu Normality \chi^2(10) = 18.19[0.05]$</td>
</tr>
</tbody>
</table>

Notes: See Table 3.1.

Figure 3.1: 16 quarters dynamic forecasts for the ICM and the Phillips-curve
The wage Phillips-curve

\[ \Delta \hat{w}_t = 0.0094 + 0.503 \Delta 2pc_{t-1} - 0.195 \Delta w_{t-1} + 0.06 \Delta 3\hat{pi}_{t-1} \\
\quad + 0.395 \Delta \tau_{1_{t-2}} - 0.013 \Delta u_t - 0.0067 u_{t-1} - 1 \Delta h_t \\
\quad - 0.046 WD_t + 0.027i86q3 - 0.005 S_{1t} - 0.001 S_{2t} - 0.001 S_{3t} \\
\quad \hat{\sigma}_{\Delta w} = 0.0122 \]

The price equation

\[ \Delta \hat{pc}_t = -0.065 + 0.108 \Delta w_{tt-1} + 0.216 \Delta 3pc_{t-1} + 0.039 (\Delta \hat{pi}_{t-1} + \Delta \hat{pi}_{t-4}) \\
\quad - 0.029 \Delta^2 y_{t-1} - 0.007 \Delta u_{t-2} + 0.129 \Delta \tau_{2_{t-2}} \\
\quad + 0.038 i70q_{1t} - 0.025 i79q_{1t} + 0.013 i86q_{3t} \\
\quad - 0.021 S_{1t} - 0.012 S_{2t} - 0.010 S_{3t} \\
\quad \hat{\sigma}_{\Delta pc} = 0.0055 \]

Notes: See Table 3.2

when the excise reduction is in the conditioning information set. However, the Phillips-curve continues to overpredict, also for the 8 and 4 quarter horizons. In fact, it is also evident from the graphs that on a comparison of biases, the Phillips-curve model would be beaten by a no-change rule, i.e., a double difference “model” \( \Delta \Delta pc_t = 0 \).

In sum, although the both the ICM and the Phillips-curve appear to be congruent models within sample, their forecast properties are indeed significantly different from a user’s point of view. In this one-off “test”, the Phillips-curve is poor on forecasting the mean of inflation in the new low-inflation regime, and it overstates the uncertainty in inflation forecasting. Causal information does not necessarily lead to successful forecasts, i.e., the finding that the Phillips-curve loses to a random-walk model of inflation over the latter part of the sample. In the next two sections the empirical analysis of these issues are carried one step further when we compare forecasts derived from large-scale systems of the Norwegian economy, using stochastic simulation techniques.

4 Large scale macroeconomic models of the Norwegian economy

RIMINI\(^5\) is the Norges Bank macroeconometric model, which is routinely used for practical forecasting and policy analysis. RIMINI is a large scale model which links together several important submodels of the Norwegian economy. One of the submodels is an ICM-type wage/price submodel of the labour market similar to the model discussed in section 3, hence we adopt the same labels and denote this as an

\(^{5}\)RIMINI was originally an acronym for a model for the Real economy and Income accounts - a MINI-version. The model version used in this paper has 205 endogenous variables, although a large fraction of these are accounting identities or technical relationships creating links between variables, see Eitrheim and Nymoen (1991) for a brief documentation of a predecessor of the model.
\textbf{EqCM submodel.} Other key submodels in RIMINI (e.g. submodels for household and corporate sector behaviour, and for housing and credit markets etc.) are also represented as equilibrium correcting dynamic equations, and in the following we will label RIMINI as an \textbf{EqCM}-type model. Section 2 above pointed out that the Phillips-curve model, which is easily obtained from the ICM-model as a special case, could logically be labelled as an \textbf{EqCM} model, although with the exception of the level of unemployment all variables enter in differences only, thus also pointing in the direction of a \textbf{dVAR} model. The issues raised in section 3 on modelling considerations, emphasizing model selection criteria and their consequences, easily generalizes to a broader and more realistic modelling framework in the RIMINI-model.

In this section we will first outline the \textbf{EqCM}-type RIMINI-model, and then briefly describe how we have constructed a contending \textbf{dVAR}-type models in differences. Typically we have remodelled the econometric equations, replacing \textbf{EqCM}-type models with models in differences only (Eitrheim et al., 1999), and notwithstanding the somewhat ambiguous labelling of the Phillips curve, the ICM submodel of the labour market has been replaced by a Phillips-curve model in the \textbf{dVAR} system counterpart to RIMINI, along the same lines as in section 3. Finally, in light of the forecasting comparison in Clements and Hendry (1999) we have also developed non-causal forecasting models in first order and second order differences, which will be denoted as \textbf{dAR} and \textbf{dARr} respectively in the following.

Clements and Hendry (1999) brought out that even for very simple systems, comparing models with and without causal information it is in general difficult to predict which version of the model is going to have the smallest forecast error, where in our case \textbf{EqCM} and \textbf{dVAR} on the one hand and \textbf{dAR} and \textbf{dARr} on the other.
represent models with and without causal information respectively.

In section 5 below, we generate multi-period forecasts from the econometric model RIMINI used by Norges Bank, and compare these to the forecasts from models based on differenced data, focusing on both forecast error biases and uncertainty using stochastic simulation. The latter extends the analysis in Eitrheim et al. (1999), and we have also extended the maximum forecasting horizon from 12 to 28 quarters. As a background for the simulations, the rest of this section describes the main features of the incumbent EqCM and how we have designed the three rival dVAR forecasting systems.

4.1 The incumbent EqCM model - eRIM

The typical forecast horizon when RIMINI is used as a forecasting tool in the preparation of Norges Bank’s forecasts for the Norwegian economy is four to eight quarters in the Bank’s Inflation report, but forecasts for up to five years ahead are also published regularly as part of the assessment of the medium term outlook for the Norwegian economy. Simulations of the RIMINI model can also provide estimates of the quantitative effects on inflation, economic growth and unemployment from changes in monetary policy instruments, and the RIMINI-model is frequently used to analyse monetary policy issues. A requirement for policy analysis is that the model contains the necessary links between monetary policy instruments like interest rates and the exchange rate and other economic variables of interest. The latter is reflected in the causal information of the model structure. In addition we have to rely on invariance properties and that the monetary policy instruments work through channels which satisfy the requirements for super exogeneity (Engle et al., 1983).

The 205 equations of RIMINI (version 2.9) fall into three categories

- 26 estimated stochastic equations, representing economic behaviour.
- 146 definitional equations, e.g. national accounting identities, composition of the work-force etc.
- 33 estimated “technical” equations, e.g. price indices with different base years and equations that serve special reporting purposes (with no feedback to the rest of the model).

It is the specification of 26 stochastic equations representing economic behaviour that distinguish the models. Together they contain putative quantitative knowledge about behaviour relating to aggregate outcome, e.g. consumption, savings and household wealth; labour demand and unemployment; wage and price interactions (inflation); capital formation; foreign trade. The oil and shipping sectors are treated exogenously in the model, as are agriculture, forestry and fisheries. The rest of the private non-financial sector is divided between the manufacturing and construction sectors (producers of traded goods) and services and retail trade (producers of non-traded goods).

Seasonally unadjusted data are used for the estimation of the equations. To a large extent, macroeconomic interdependencies are contained in the dynamics of the model. For example, prices and wages are Granger-causing output, trade and employment and likewise the level of real-activity feeds back on to wage-price inflation.
The model is an open system: Examples of important non-modelled variables are the level of economic activity by trading partners, as well as inflation and wage-costs in those countries. Indicators of economic policy (the level of government expenditure, the short-term interest rate and the exchange rate), are also non-modelled and the forecasts are therefore conditional on a particular scenario for these variables. The EqCM model RIMINI will be labeled \textit{eRIM} in the following.

To provide some insights in the type of causal information which is reflected in the behavioural relationships in \textit{eRIM} (and \textit{dVARc} below), consider the link between interest rates and other economic variables in Figure 4.1. The main links between short-term interest rates and aggregated variables like output, employment and CPI inflation in RIMINI, is often denoted as the “interest rate channel” of monetary policy.
Figure 4.1: Interest rate channels in RIMINI. Effects on CPI inflation assuming constant exchange rates.
The main mechanisms of the interest rate channel in RIMINI are: A partial rise in the short-term money market interest rate (typically 3-month NOK rates) assuming fixed exchange rates, leads to an increase in banks’ borrowing and lending interest rates with a lag. Aggregate demand is influenced by the interest rate shift through several mechanisms, such as a negative effect on housing prices which (for a given stock of housing capital) causes real household wealth to decline and suppresses total consumer expenditure. Likewise, there are negative direct and indirect effects on real investments in sectors producing traded and non-traded goods and on housing investments. The housing and credit markets in RIMINI are interrelated through the housing price and household loans equations which a.o. reflects that housing capital is collateralized against household loans (mainly in private and state owned banks) and noting that the ownership rate among Norwegian households exceeds 80%. CPI inflation is reduced after a lag, mainly as a result of the effects of changes in aggregate demand on aggregate output and employment (productivity), but also as a result of changes in unit labour costs.

4.2 A full scale dVAR model - dRIMc

Because all the stochastic equations in RIMINI are in equilibrium correction form, a simple dVAR version of the model, can be obtained by omitting the equilibrium correcting terms from the equation and re-specifying all the affected equations in terms of differences alone. In our earlier paper however, we found that this left us with seriously misspecified equations a.o. due to the autocorrelation in the omitted equilibrium correcting term. This model was denoted dRIM in Eitrheim et al. (1999) and we have discarded this model in the present analysis.

The previous paper also showed that a more interesting rival was a re-modelled version of dRIM (a similar procedure was applied in section 3). In order to make the residuals of the dVAR-equations empirically white-noise, additional terms in differences often had to be added to remove autocorrelation from the model residuals. The corrected dVAR version of RIMINI is denoted dRIMc, and, bearing in mind the potential bias in the estimation of the drift term in small samples, we have made simulations with a version of dRIMc where we systematically excluded the constant term from typical no-drift variables like unemployment rates and interest rates, see Eitrheim et al. (1999) for discussion.

Hence, the two complete system forecasting models, eRIM and the no-drift version of dRIMc, both broadly satisfy the same set of single equation model design criteria and show model residuals which are close to empirical white noise, zero mean innovations.

4.3 Difference and double difference models - dAR and dARr

Both models considered so far are “system of equations” forecasting models. For comparison, we have also prepared single equation forecasts for each variable, i.e. in line with equations (2.6) and (2.7) in section 2 (see also Clements and Hendry (1999, Chapter 5)) but allowing for higher order dynamics and seasonality. The first set of single equation forecasts is dubbed dAR, and is based on unrestricted estimation of AR(4) models, including a constant term and three seasonal dummies. Finally, we generate forecasts from $\Delta_4 \Delta \ln X_t = 0$, for each variable $X_t$ in the set of endogenous variables. This set of forecasts is called dARr, where the $r$ is a
reminder that the forecasts are based on completely restricted AR(4) processes. The univariate dARr “models” are specified without drift terms, hence their forecasts are protected against trend-misrepresentation.

5 Forecast comparisons of the large-scale models

Table 5.1 summarizes the four models of the previous section in terms of the incumbent “baseline” EqCM model and the three “rival” dVAR type models.

<table>
<thead>
<tr>
<th>Model</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>eRIM</td>
<td>26 Behavioural equations, equilibrium-correcting equations 33+146 Technical and definitional equations</td>
</tr>
<tr>
<td>1.Rival</td>
<td>dRIMc</td>
<td>26 Behavioural equations, remodelled without levels-information, restricting drift terms to zero 33+146 Technical and definitional equations</td>
</tr>
<tr>
<td>2.Rival</td>
<td>dAR</td>
<td>71 equations modelled as 4.order unrestricted AR models</td>
</tr>
<tr>
<td>3.Rival</td>
<td>dARr</td>
<td>71 equations modelled as restricted 4.order AR models, restricting drift terms to zero</td>
</tr>
</tbody>
</table>

All models that enter this exercise were estimated on a sample ending in 1991.4. The period 1992.1-1998.4 is used for forecast comparisons.

In this section we use graphs to compare forecasts of eRIM and the three dVARs: dRIMc, dAR and dARr. The graphs allow direct inspection of bias and uncertainty in the form of Monte-Carlo simulated prediction intervals. The set of variables that we consider is the rate of inflation (∆₄cpi), and some of the key factors affecting it: The annual growth in import prices ,∆₄pbi, (imported inflation), the annual growth in wage costs per hour, ∆₄wcf, (cost-push), the level of unemployment, UTOT, (labour market pressure), and annual non-oil GDP growth, ∆₄yf, (product market pressure).

We evaluate four dynamic forecasts, distinguished by the start period: The first forecast is for the whole 28 quarter horizon, so the first period being forecasted is 1992.1. The second simulation starts in 1995.1 (16 quarters horizon), the third in 1997.1 (8 quarters horizon) and the fourth in 1998.1 (4 quarters horizon). Furthermore, all forecast are conditional on the actual values of the models’ exogenous variables and the initial conditions, which of course change accordingly when we initialize the forecasts in different start periods.

The figures 5.1 - 5.5 show the results from stochastic simulations of each of the four models at the different forecasting horizons. For each of the report variables we have plotted the mean of 500 replications (using antithetic drawings) against the observed historical values. The uncertainty of the model forecasts at different horizons is illustrated by putting a 95% prediction interval around the mean. Some important theoretical properties about the second order moments of EqCM-type and differenced models are discussed in Clements and Hendry (1999, Chapter 5).
We first consider the results for annual inflation $\Delta_4 cpi$ in figure 5.1. The eRIM model seems to do better than its dVAR rivals in the first half of the 28 quarter forecasting horizon, both in terms of a smaller forecast error bias, and less forecast uncertainty. We note however that as we approach the end of the horizon eRIM starts to overpredict CPI-inflation from 1996 and onwards. As we move the starting point of the simulations forward, the simple univariate dARr model outperform all the other forecasting models in terms of the forecast error bias, cf. figure 5.1(d). Similar results were reported in Eitrheim et al. (1999) for the 1992.1-1994.4 period. They illustrate the inherent ability of the differencing models to intercept-correct for structural changes occurring after the estimation period (which ends in 1991q4 for all simulations), as well as they insulate against shocks occurring in other parts of a larger system.

While eRIM only overpredicts CPI-inflation from 1996 and onwards, the rival system forecast from dRIMc consistently overpredicts CPI-inflation over the entire forecasting horizon. This is also in line with the results from the Phillips curve model in section 3. From figure 5.1 we see that the extra unit-root assumption built into models in differences give rise to wider prediction intervals as we increase the forecast horizon, and in particular we see that this is the case for dARr-models where there is a clear trend. In contrast to the models eRIM and dRIMc however, which are designed such that the residuals are close to being zero mean innovations, to the extent the two univariate “models”, and in particular dARr, have autocorrelated residuals, the prediction intervals may be overstated Clements and Hendry (1999, Chapter 5).

Turning to the tendency to overpredict inflation, note that for eRIM and dRIMc, this can be tracked down to two channels. First through the growth in import prices (figure 5.2), and second through wage growth (figure 5.3)\(^6\). Import price growth is overpredicted by both eRIM and dRIMc in 1997-1998, as shown in figure 5.2(c) and 5.2(d), and wage growth forecasts seems to largely follow the same pattern as CPI-inflation forecasts in eRIM and dRIMc respectively. Again, the non-causal models dominate on the shortest forecast horizon, cf. figure 5.3(d). The forecasts for wage and price inflation in eRIM is consistent with the forecast of the total rate of unemployment in figure 5.4, which is underpredicted in the latter part of the simulation period.

Whereas the eRIM model underpredicts the rate of total unemployment from 1996 and onwards, this is not the case for the dRIMc model. The rate of unemployment seems to be more or less free-coupled from the wage-price formation process in dRIMc, overpredicting unemployment from 1996 and onwards in the 28 quarter forecast and being close to spot on the actual values when simulation starts in 1997q1. This is due to an insulation property in the dRIMc model, namely that forecast errors in the price/wage block of the system does not feed into the unemployment block (see Eitrheim et al. (1999) for discussion).

For the annual growth in mainland GDP, see figure 5.5, eRIM outperforms the other models on the 28 quarter forecast horizon. Like the eRIM, the other system forecast, dRIMc, also benefits from the conditioning on the true values of the exogenous variables, although the dRIMc shows a tendency to underpredict output growth compared with eRIM. The two univariate forecasts underpredict output growth really badly at the 28 quarter horizon, due to the fact that both forecasts tend to

---

\(^6\)Note that the forecasts in section 3 were conditional on actual import prices.
extrapolate the low growth rates in Norway in the early 1990s. As we move to the 16 quarter forecast starting in 1995q1, the restricted univariate model $dARr$ does much better on average than both the unrestricted univariate model $dAR$ which still badly underpredicts, and the dVAR system forecast $dRIMc$ which also shows the same tendency to underpredict output growth. With the exception of a single-year instance of underpredicting output growth in 1995, the $eRIM$ continues to predict output growth relatively well. As we continue to the 8 and 4 quarter forecasts starting in 1997q1 and 1998q1, respectively, the two open system forecasts continue to benefit from their conditioning on a larger information set.

The main findings conveyed by the graphs can be summarized in the following four points. First, the simulations confirm earlier results presented in Eitrheim et al. (1999) that while $eRIM$ tends to improve the forecast on longer horizons (extending earlier result from 12 to a 28 periods), non-causal “models” like $dAR$ and in particular $dARr$ still seems to frequently outperform the forecast properties of their system counterparts on shorter horizons, thus corroborating the theoretical results in Clements and Hendry (1999) on the relative robustness of the two types of models.

Second, we have shown evidence that $dVAR$-type models offer some degree of protection against forecast failure stemming from structural breaks that have occurred prior to the forecast period. This suggests that if a structural break is suspected to have occurred, or is known from non-modelled sources, the $dVAR$ model may provide information about the magnitude of the required intercept correction.

Third, another form of robustness in $dAR$ and $dARr$ stems from the insulation of e.g., the inflation forecasts from errors in other variables like the growth rate in import prices. In other words, forecasts from models with less causal information are more robust than models with e.g., policy channels.

Fourth, the estimated second order moments seem to be in line with the forecasting theory in Clements and Hendry (1999, Chapter 5) and show typically the differences between the forecast error variances in the causal and non-causal models respectively, although it should be noted that the latter may overstate the uncertainty since the model residuals may be autocorrelated in practice.

6 Discussion

The fact that the underlying inflation process is unknown and that competing models exist, makes it attractive for forecasters to draw on several models. This is reflected by operational inflation forecasting, where a multitude of models are currently in use. Models range from large macroeconomic model to simpler Phillips-curve models, and from data based VARs to full blown theoretical optimizing models. Given that all models are necessarily partial in nature, it is easy to acknowledge the wisdom of a pragmatic and pluralist use of economic models, of which the Bank of England’s suite of model approach is one example.

However, the pluralist view has a cost in terms of lost transparency. When many models are used to forecast inflation, it may become difficult for the central bank to communicate its forecasts to the general public, the markets and the political authorities. Thus section 2 showed that it may be difficult to draw the correct policy implications from a combined forecast of two models that have different theoretical content, and that the combined forecast can be less accurate than the forecast of
Figure 5.1: Forecast comparisons over different horizons based on stochastic simulation. Annual consumer price inflation $\Delta_4\text{cpi}$

(a) 28 quarter forecast horizon, 1992.1 - 1998.4

(b) 16 quarter forecast horizon, 1995.1 - 1998.4

(c) 8 quarter forecast horizon, 1997.1 - 1998.4

(d) 4 quarter forecast horizon, 1998.1 - 1998.4
Figure 5.2: Forecast comparisons over different horizons based on stochastic simulation. Annual import price growth $\Delta p_{bi}$

(a) 28 quarter forecast horizon, 1992.1 - 1998.4

(b) 16 quarter forecast horizon, 1995.1 - 1998.4

(c) 8 quarter forecast horizon, 1997.1 - 1998.4

(d) 4 quarter forecast horizon, 1998.1 - 1998.4
Figure 5.3: Forecast comparisons over different horizons based on stochastic simulation. Annual growth in wage costs per hour $\Delta w_{cf}$

(a) 28 quarter forecast horizon, 1992.1 - 1998.4

(b) 16 quarter forecast horizon, 1995.1 - 1998.4

(c) 8 quarter forecast horizon, 1997.1 - 1998.4

(d) 4 quarter forecast horizon, 1998.1 - 1998.4
Figure 5.4: Forecast comparisons over different horizons based on stochastic simulation. 
Rate of total unemployment UTOT

(a) 28 quarter forecast horizon, 1992.1 - 1998.4

(b) 16 quarter forecast horizon, 1995.1 - 1998.4

(c) 8 quarter forecast horizon, 1997.1 - 1998.4

(d) 4 quarter forecast horizon, 1998.1 - 1998.4
Figure 5.5: Forecast comparisons over different horizons based on stochastic simulation. Annual growth in mainland GDP $\Delta_{4yf}$

(a) 28 quarter forecast horizon, 1992.1 - 1998.4

(b) 16 quarter forecast horizon, 1995.1 - 1998.4

(c) 8 quarter forecast horizon, 1997.1 - 1998.4

(d) 4 quarter forecast horizon, 1998.1 - 1998.4
the model with best econometric properties. Hence, model pluralism can be carried too far. An alternative is represented by the encompassing principle and progressive research strategies. In practice this entails a strategy where one tests competing models as thoroughly as practically feasible, and keep only the encompassing model in the suite of models. Our example in section 2 focused on competing models of the supply side, the ICM and the Phillips-curve. But several other issues in the modelling of inflation can in principle be tackled along the same line, for example the role of rational expectations, and forward looking optimizing behaviour, see e.g. Hendry (1995, chapter 14).

A more specific argument for pluralism stems from the insight that an EqCM is incapable of correcting forecasts sufficiently for the impact of parameter changes that occur prior to the start of the forecast period. In contrast, a univariate dVAR forecast moves back on track once it is conditional on the period when the parameter change took place. Section 5 of this chapter provided several empirical examples. This supports the view that an EqCM and simple dVARs could constitute a suite of models for forecasting, where the role of the dVAR is primarily to help intercept correct the policy model.

References


