2010 | 03

Working Paper

Monetary Policy Department

Simple rules versus optimal policy: what fits?

By Ida Wolden Bache, Leif Brubakk and Junior Maih

NORGES BANK

Working papers fra Norges Bank, fra 1992/1 til 2009/2 kan bestilles over e-post:
servicesenter@norges-bank.no
eller ved henvendelse til: Norges Bank, Abonnementsservice
Postboks 1179 Sentrum
0107 Oslo
Telefon 22 31 63 83, Telefaks 22 41 31 05

Fra 1999 og senere er publikasjonene tilgjengelige på www.norges-bank.no

Working papers inneholder forskningsarbeider og utredninger som vanligvis ikke har fått sin e4ndelige form.
Hensikten er blant annet at forfatteren kan motta kommentarer fra kolleger og andre interesserte.
Synspunkter og konklusjoner i arbeidene står for forfatternes regning.

Working papers from Norges Bank, from1992/1 to 2009/2 can be ordered by e-mail:
servicesenter@norges-bank.no
or from Norges Bank, Subscription service,
P.O.Box. 1179 Sentrum
N-0107Oslo, Norway.
Tel. +47 22 31 63 83, Fax. +47 22 41 31 05

Working papers from 1999 onwards are available on www.norges-bank.no

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

ISSN 1502-8143 (online)
ISBN 978-82-7553-546-5 (online)

Simple rules versus optimal policy: what fits?*

Ida Wolden Bache†, Leif Brubakk and Junior Maih
Monetary Policy Wing, Norges Bank (Central Bank of Norway)

This version: 7 April 2010
First version: 4 June 2008

Abstract

We estimate a small open-economy DSGE model for Norway with two specifications of monetary policy: a simple instrument rule and optimal policy based on an intertemporal loss function. The empirical fit of the model with optimal policy is as good as the model with a simple rule. This result is robust to allowing for misspecification following the DSGE-VAR approach proposed by Del Negro and Schorfheide (2004). The interest rate forecasts from the DSGE-VARs are close to Norges Bank's official forecasts since 2005. One interpretation is that the DSGE-VAR approximates the judgment imposed by the policymakers in the forecasting process.

Keywords: DSGE models, forecasting, optimal monetary policy

JEL classification: C53, E52

*We are grateful for comments from Jesper Lindé, Marco Del Negro, Øistein Røisland, Tommy Sveen, Shaun Vahey, Lutz Weinke and participants at the CEF conference in Paris 2008, the ESEM 2008 meeting in Milan, the Third annual Dynare conference in Boston 2008, the Norges Bank Workshop on Optimal Monetary Policy in November 2008 and the National Bank of Poland Workshop on Experiences and Challenges of Forecasting at Central Banks November 2009. The views in this paper are our own and should not be interpreted as reflecting the views of Norges Bank.

†Corresponding author. Address: Monetary Policy Department, Norges Bank (Central Bank of Norway). E-mail: ida-wolden.bache@norges-bank.no

1 Introduction

The purpose of this paper is to compare the empirical merits of different approaches to modelling monetary policy within the context of a dynamic stochastic general equilibrium (DSGE) model. To this end we evaluate a New Keynesian small open economy model estimated on Norwegian data under alternative specifications of monetary policy. We believe the case of Norway to be of general interest. First, to our knowledge, Norges Bank is the only central bank that has stated publicly that it uses 'optimal' policy as a normative benchmark for monetary policy (see e.g., Holmsen et al. (2007)). Second, since 2005 Norges Bank has published its own interest rate projections along with forecasts of other key macrovariables.

In most DSGE models, the central bank is assumed to set the interest rate according to a simple instrument rule (e.g., a Taylor rule). In addition to computational simplicity, one reason behind the popularity of this approach is that simple instrument rules have been shown to give a reasonable empirical description of actual monetary policy in many countries. Moreover, simple rules are perceived to be more robust in that they perform reasonably well in terms of welfare across models.

An alternative approach is to assume that monetary policy is conducted optimally. By optimally, we mean that the central bank chooses the interest path that minimizes an intertemporal loss function. The optimal policy approach gives a more symmetric treatment of central bank and private sector behaviour and, moreover, allows the central bank to make efficient use of all relevant information. As pointed out by Svensson (2003), it seems somewhat odd to assume a priori that the central bank has a less sophisticated approach to optimization than the private agents. Finally, the optimizing framework appears to be more in line with the way monetary policy is actually conducted in most developed countries.

From an empirical point of view it is not obvious which of the two policy assumptions provides the most plausible account of the data. There are two opposing mechanisms at play. On the one hand, the optimal policy framework is more flexible than the simple instrument rule in the sense that the implied interest rate rule contains a larger set of variables than the simple instrument rule. However, this flexibility comes at the cost of introducing a new set of restrictions on the reduced form solution of the model, restrictions that could potentially be at odds with the data.

The estimated model is similar in size and structure to NEMO, the model that is used as the core model in the policy process in Norges Bank[1], and thus constitutes a real world example of empirical interest. The model is estimated using Bayesian techniques on data for the Norwegian mainland economy over the period 1987Q1–2007Q4. We consider two different specifications of monetary policy: a simple instrument rule and optimal policy based on a loss function that is consistent with the monetary policy remit. The different specifications of the model are compared using both in-sample and out-of-sample measures of fit, where the latter exercise is based on recursive forecasts from 1998Q1 to 2007Q4. We

[1]See Monetary Policy Report 3/07 (available at www.norges-bank.no).

stress the out-of-sample forecasting properties of the models for two reasons. First, model comparisons based on Bayesian measures of in-sample fit can be problematic (see e.g., Sims (2003)). Second, and more importantly, forecasting is a key activity of an inflation targeting central bank. Hence, in practical policy work, models are ultimately judged by their forecasting properties. In order to shed some light on the accuracy of the pure model projections, we also compare the model forecasts of the interest rate and inflation to the official forecasts actually published by Norges Bank from 2005Q4 onwards.[2]

There exists a small, but increasing literature estimating New Keynesian models with optimal monetary policy. Dennis (2004) jointly estimates the parameters in the central bank's objective function and the parameters in the optimizing constraints in a New Keynesian model of the US economy, under the assumption that monetary policy is conducted optimally under discretion. In two recent papers Ilbas (2008a) and Ilbas (2008b) use a Bayesian approach to estimate the monetary policy preferences in New Keynesian closed-economy models for the euro area and the US assuming that the central bank minimises an intertemporal loss function under commitment. Adolfson et al. (2009) estimate an operational medium-scale, small open economy DSGE model for the Swedish economy and compare the in-sample fit of models with alternative assumptions about monetary policy. We supplement and add to their results by also considering the out-of-sample forecasting performance of the models.

Our findings can be summarised as follows. First, the in-sample fit of the model with optimal policy is superior to the model with a simple instrument rule. However, in terms of forecasting accuracy, which is our favoured measure of model fit, the models perform about equally well. Turning to the absolute performance, the estimated models significantly overshoots both the actual outcomes and the official Norges Bank forecasts for inflation and the interest rate. This overshooting is more pronounced for the model with optimal policy than in the model with a simple instrument rule, reflecting in part the fact that optimal policy is solved under the assumption of timeless commitment. Interestingly, the parameter estimates appear to be quite robust to the choice of monetary policy. Thus, it would be tempting to conclude that the model parameters in this sense are structural. However, this would ignore the issue of misspecification.

In the above exercise, we implicitly assume that, under each of the two approaches to modelling monetary policy, the resulting theoretical model provides an accurate probabilistic description of our data. This is obviously a strong assumption. Despite the recent progress in getting DSGE models to fit the data (see e.g., Smets & Wouters (2004), Edge et al. (2010), Adolfson et al. (2007b) and Adolfson et al. (2007c)), potential model misspecification remains a key concern. As discussed in Del Negro & Schorfheide (2009), the issue of model misspecification in DSGE models can be approached in a number of ways. The more practical approach, favoured by most central banks, is to account for model misspecification by adding significant amounts of judgment to the forecasts from their core models. In order to investigate the importance of model misspecification, we employ the DSGE-VAR approach proposed by Del Negro & Schorfheide (2004). In their framework,

[2]This coincides with the quarter where the Norges Bank first started publishing its interest rate paths.

the DSGE model is used as a prior to inform the parameters of an unrestricted vector autoregression (VAR). The idea is to impose some of the structure from the theoretical model on the less dogmatic data representation provided by the VAR. The DSGE-VAR approach still produces estimates of the parameters in the DSGE model that can be compared to those obtained using the traditional full-information approach. In some sense, the DSGE-VAR captures the dichotomy between model and judgement in practical policy work. One of the questions we ask in this paper is to what extent accounting for misspecification affects the parameter estimates and the forecasts from the models.

In a related paper, Adjemian et al. (2008) use the DSGE-VAR framework to compare the in-sample fit of a closed economy DSGE model for the US economy when monetary policy is conducted optimally under commitment and when the central bank follows a Taylor-type rule. Recognising that in-sample model comparison within a Bayesian framework can be problematic, we extend their results by considering the out-of-sample forecasting performance of the two models. Our recursive estimation procedure has the added advantage that it allows us to investigate the stability of the parameters over time. Our paper also differs from Adjemian et al. (2008) in that we assume the same set of stochastic disturbances in the two models, which makes the model comparison more transparent.

Based on the marginal data densities, we find that the data clearly favour the DSGE-VAR model with optimal policy. This runs contrary to the findings in Adjemian et al. (2008). However, as in our benchmark case, the forecasting performance of the two DSGE-VAR models are almost identical. Interestingly, allowing for misspecification brings the projected interest and inflation paths from the DSGE-VARs much closer to both the actual outcomes and Norges Bank's official interest rate forecasts. In this sense, the DSGE-VAR models can be said to better capture the judgment imposed by the policymakers.

The remainder of the paper is organized as follows. In section 2 we give a brief description of the DSGE model used in the empirical exercise. In section 3 we present the estimation strategy and the empirical results for the two specifications of the DSGE model. In section 4 we discuss the results obtained when using the DSGE-VAR approach. Section 5 concludes the paper.

2 The DSGE model

The benchmark DSGE model used in the forecasting exercise is a medium-scale New Keynesian open economy model. The theoretical framework builds on the New Open Economy Macroeconomics (NOEM) literature (see e.g., Lane (2001) for a survey) as well as the closed economy models in e.g., Christiano et al. (2005) and Smets & Wouters (2003), and is similar in structure to existing open-economy models such as the Global Economy Model (GEM) model at the International Monetary Fund and the model developed in Adolfson et al. (2007a).[3]

The economy has two production sectors. Firms in the intermediate goods sector produce differentiated goods for sale in monopolistically competitive markets at home and

[3] We refer to Brubakk et al. (2006) for a more thorough discussion of the model and literature references.

abroad, using labour and capital as inputs. Firms in the perfectly competitive final goods sector combine domestically produced and imported intermediate goods into an aggregate good that can be used for private consumption, private investment and government spending. The household sector consists of a continuum of infinitely-lived households that consume the final good, work and save in domestic and foreign bonds. The model incorporates real rigidities in the form of habit persistence in consumption, variable capacity utilisation of capital and investment adjustment costs, and nominal rigidities in the form of local currency price stickiness and nominal wage stickiness. The model is closed by assuming that domestic households pay a debt-elastic premium on the foreign interest rate when investing in foreign bonds. The model evolves around a balanced growth path as determined by a permanent technology shock. The fiscal authority runs a balanced budget each period, and we consider two alternative specifications of monetary policy. The exogenous foreign variables are assumed to follow autoregressive processes.

Final goods sector The perfectly competitive final goods sector consists of a continuum of final good producers indexed by $x \in [0, 1]$ that aggregates composite domestic intermediate goods, Q, and imports, M, using a constant elasticity of substitution (CES) technology:

$$A_t(x) = \left[\eta^{\frac{1}{\mu}} Q_t(x)^{1-\frac{1}{\mu}} + (1-\eta)^{\frac{1}{\mu}} M_t(x)^{1-\frac{1}{\mu}}\right]^{\frac{\mu}{\mu-1}}, \tag{1}$$

The degree of substitutability between the composite domestic and imported goods is determined by the parameter $\mu > 0$, whereas η $(0 \leq \eta \leq 1)$ measures the steady-state share of domestic intermediates in the final good for the case where relative prices are equal to 1.

The composite good $Q(x)$ is an index of differentiated domestic intermediate goods, produced by a continuum of firms $h \in [0, 1]$:

$$Q_t(x) = \left[\int_0^1 Q_t(h, x)^{1-\frac{1}{\theta_t}} \, dh\right]^{\frac{\theta_t}{\theta_t - 1}}, \tag{2}$$

where the time-varying elasticity of substitution between domestic intermediates is captured by θ_t and evolves according to:

$$\ln\left(\frac{\theta_t}{\theta}\right) = \lambda_\theta \ln\left(\frac{\theta_{t-1}}{\theta}\right) + \varepsilon_t^\theta, \qquad 0 \leq \lambda_\theta < 1, \quad \varepsilon_t^\theta \sim iid\left(0, \sigma_\theta^2\right) \tag{3}$$

where $\theta > 1$ is the steady-state value.

Similarly, the composite imported good is a CES aggregate of differentiated import goods indexed by $f \in [0, 1]$:

$$M_t(x) = \left[\int_0^1 M_t(f, x)^{1-\frac{1}{\theta^f}} \, df\right]^{\frac{\theta^f}{\theta^f - 1}}, \tag{4}$$

where $\theta^f > 1$ is the steady-state elasticity of substitution between imported goods.

Intermediate goods sector Each intermediate goods firm h is assumed to produce a differentiated good $T_t(h)$ for sale in domestic and foreign markets using the following CES production function:

$$T_t(h) = \left[(1-\alpha)^{\frac{1}{\xi}} \left(Z_t z_t^L l_t(h) \right)^{1-\frac{1}{\xi}} + \alpha^{\frac{1}{\xi}} \overline{K}_t(h)^{1-\frac{1}{\xi}} \right]^{\frac{\xi}{\xi-1}}, \tag{5}$$

where $\alpha \in [0,1]$ is the capital share and ξ denotes the elasticity of substitution between labour and capital. The variables $l_t(h)$ and $\overline{K}_t(h)$ denote, respectively, hours used and effective capital of firm h in period t. There are two exogenous shocks to productivity in the model: Z_t refers to an exogenous permanent (level) technology process, which grows at the gross rate π_t^z, whereas z_t^L denotes a temporary (stationary) shock to productivity (or labour utilization). The technology processes are modelled as

$$\ln(Z_t) = \ln(Z_{t-1}) + \ln(\pi^z) + \ln\left(\frac{\pi_t^z}{\pi^z}\right), \tag{6}$$

where

$$\ln\left(\frac{\pi_t^z}{\pi^z}\right) = \lambda_z \ln\left(\frac{\pi_{t-1}^z}{\pi^z}\right) + \varepsilon_t^z, \qquad 0 \le \lambda_z < 1, \quad \varepsilon_t^z \sim iid\left(0, \sigma_z^2\right), \tag{7}$$

and

$$\ln\left(\frac{z_t^L}{z^L}\right) = \lambda_L \ln\left(\frac{z_{t-1}^L}{z^L}\right) + \varepsilon_t^L, \qquad 0 \le \lambda_L < 1, \quad \varepsilon_t^L \sim iid\left(0, \sigma_L^2\right). \tag{8}$$

The variable $K_t(h)$ is defined as firm h's capital stock that is chosen in period t and becomes productive in period $t+1$. Firm h's *effective* capital in period t is related to the capital stock that was chosen in period $t-1$ by

$$\overline{K}_t(h) = u_t(h) K_{t-1}(h), \tag{9}$$

where $u_t(h)$ is the endogenous rate of capital utilization. When adjusting the utilization rate the firm incurs a cost of $\gamma_t^u(h)$ units of final goods per unit of capital. The cost function is

$$\gamma_t^u(h) = \phi_1^u \left(e^{\phi_2^u (u_t(h)-1)} - 1 \right), \tag{10}$$

where ϕ_1^u and ϕ_2^u are parameters determining the cost of deviating from the steady state utilization rate. The steady state utilization rate is normalized to one.[4]

Firm h's law of motion for physical capital reads:

$$K_t(h) = (1-\delta) K_{t-1}(h) + \kappa_t(h) K_{t-1}(h), \tag{11}$$

where $\delta \in [0,1]$ is the rate of depreciation and $\kappa_t(h)$ denotes capital adjustment costs. The adjustment costs take the following form:

[4] Note that ϕ_1^u is not a free parameter. It is set to ensure that the marginal cost of utilisation is equal to the rental rate of capital in steady-state.

$$\kappa_t(h) = \frac{I_t(h)}{K_{t-1}(h)} - \frac{\phi_1^I}{2}\left[\left(\frac{I_t(h)}{K_{t-1}(h)} - \left(\frac{I}{K} + z_t^I\right)\right)\right]^2$$
$$-\frac{\phi_2^I}{2}\left(\frac{I_t(h)}{K_{t-1}(h)} - \frac{I_{t-1}}{K_{t-2}}\right)^2, \tag{12}$$

where I_t denotes investment and z_t^I is an investment shock[5] that evolves according to

$$\ln\left(\frac{z_t^I}{z^I}\right) = \lambda_I \ln\left(\frac{z_{t-1}^I}{z^I}\right) + \varepsilon_t^I, \qquad 0 \leq \lambda_I < 1, \quad \varepsilon_t^I \sim iid\left(0, \sigma_I^2\right). \tag{13}$$

The labour input is a CES aggregate of hours supplied by a continuum of infinitely-lived households indexed by $j \in [0,1]$:

$$l_t(h) = \left[\int_0^1 l_t(h,j)^{1-\frac{1}{\psi_t}} dj\right]^{\frac{\psi_t}{\psi_t - 1}}, \tag{14}$$

where ψ_t denotes the elasticity of substitution between different types of labour that evolves according to:

$$\ln\left(\frac{\psi_t}{\psi}\right) = \lambda_\psi \ln\left(\frac{\psi_{t-1}}{\psi}\right) + \varepsilon_t^\psi, \qquad 0 \leq \lambda_\psi < 1, \quad \varepsilon_t^\psi \sim iid\left(0, \sigma_\psi^2\right). \tag{15}$$

Firms sell their goods in markets characterised by monopolistic competition. International goods markets are segmented and firms set prices in the local currency of the buyer. An individual firm h charges $P_t^Q(h)$ in the home market and $P_t^{Mf}(h)$ abroad, where the latter is denoted in foreign currency. Nominal price stickiness is modelled by assuming that firms face quadratic costs of adjusting prices,

$$\gamma_t^{PQ}(h) \equiv \frac{\phi^Q}{2}\left[\frac{P_t^Q(h)}{\pi P_{t-1}^Q(h)} - 1\right] \tag{16}$$

$$\gamma_t^{PMf}(h) \equiv \frac{\phi^{Mf}}{2}\left[\frac{P_t^{Mf}(h)}{\pi P_{t-1}^{Mf}(h)} - 1\right] \tag{17}$$

in the domestic and foreign market, respectively and π denotes the steady-state inflation rate in the domestic economy. In every period cash-flows are paid out to the households as dividends.

Firms choose hours, capital[6], investment, the utilization rate and prices to maximize the present discounted value of cash-flows, adjusted for the intangible cost of changing prices, taking into account the law of motion for capital, and demand both at home and

[5] This shock could e.g., represent changes in the relative price of consumption and investment.

[6] Capital is firm-specific, but since all firms are identical and there is no price dispersion this assumption does not affect the linearised dynamics of the model.

abroad, $T_t^D(h)$. The latter is given by:

$$T_t^D(h) = \int_0^1 Q_t(h,x)dx + \int_0^1 M_t^f(h,x^f)dx^f \tag{18}$$

Households The period utility function is additively separable in consumption and leisure. The lifetime expected utility of household j is:

$$U_t(j) = E_t \sum_{i=0}^{\infty} \beta^i \left[z_{t+i}^u u\left(C_{t+i}(j)\right) - v\left(l_{t+i}(j)\right)\right], \tag{19}$$

where C denotes consumption, l is hours worked and β is the discount factor $0 < \beta < 1$. The consumption preference shock, z_t^u, evolves according to

$$\ln\left(\frac{z_t^u}{z^u}\right) = \lambda_u \ln\left(\frac{z_{t-1}^u}{z^u}\right) + \varepsilon_t^u, \qquad 0 \le \lambda_u < 1, \quad \varepsilon_t^u \sim iid\left(0, \sigma_u^2\right). \tag{20}$$

The current period utility functions for consumption and labour choices, $u(C_t(j))$ and $v(l_t(j))$, are

$$u\left(C_t(j)\right) = \left(1 - b^c/\pi^z\right) \ln\left[\frac{\left(C_t(j) - b^c C_{t-1}\right)}{1 - b^c/\pi^z}\right], \tag{21}$$

and

$$v\left(l_t(j)\right) = \frac{1}{1+\zeta} l_t(j)^{1+\zeta}. \tag{22}$$

where the degree of external habit persistence in consumption is governed by the parameter b^c $(0 < b^c < 1)$ [and the disutility of supplying labour is governed by the parameter $\zeta > 0$.]

Each household is the monopolistic supplier of a differentiated labour input and sets the nominal wage subject to the labour demand of intermediate goods firms and subject to quadratic costs of adjustment, γ^W:

$$\gamma_t^W(j) \equiv \frac{\phi^W}{2}\left[\frac{W_t(j)/W_{t-1}(j)}{W_{t-1}/W_{t-2}} - 1\right]^2 \tag{23}$$

where W_t is the nominal wage rate.

The flow budget constraint for household j is:

$$\begin{aligned} P_t C_t(j) + S_t B_{H,t}^f(j) + B_t(j) \le {} & W_t(j)\, l_t(j) \left[1 - \gamma_t^W(j)\right] \\ & + \left[1 - \gamma_{t-1}^{B^f}\right]\left(1 + r_{t-1}^f\right) S_t B_{H,t-1}^f(j) \\ & + \left(1 + r_{t-1}\right) B_{t-1}(j) + DIV_t(j) - TAX_t(j), \end{aligned} \tag{24}$$

where S_t is the nominal exchange rate, $B_t(j)$ and $B_{H,t}^f(j)$ are household j's end of period t holdings of domestic and foreign bonds, respectively. Only the latter are traded internationally. The domestic short-term nominal interest rate is denoted by r_t, and the nominal return on foreign bonds is r_t^f. The variable DIV includes all profits from intermediate goods firms and nominal wage adjustment costs, which are rebated in a lump-sum fashion.

Finally, home agents pay lump-sum (non-distortionary) net taxes, TAX_t, denominated in home currency.

A financial intermediation cost, γ^{Bf}, is introduced to guarantee that aggregate net foreign assets follow a stationary process. This cost depends on the average net foreign asset position of the domestic economy. The intermediation cost takes the following form[7]

$$\gamma_t^{Bf} = \phi^{B1} \frac{\exp\left(\phi^{B2}\left(\frac{S_t B_{H,t}^f}{P_t Z_t}\right)\right) - 1}{\exp\left(\phi^{B2}\left(\frac{S_t B_{H,t}^f}{P_t Z_t}\right)\right) + 1} + z_t^B, \tag{25}$$

where $0 \leq \phi^{B1} \leq 1$ and $\phi^{B2} > 0$. The exogenous 'risk premium', z_t^B, evolves according to

$$\ln\left(\frac{z_t^B}{z^B}\right) = \lambda_B \ln\left(\frac{z_{t-1}^B}{z^B}\right) + \varepsilon_t^B, \qquad 0 \leq \lambda_B < 1, \quad \varepsilon_t^B \sim iid\left(0, \sigma_B^2\right). \tag{26}$$

Government The government purchases final goods financed through a lump-sum tax. Real government spending (adjusted for productivity), $g_t \equiv G_t/Z_t$, is modelled as a first-order autoregressive process

$$\ln\left(\frac{g_t}{g}\right) = \lambda_G \ln\left(\frac{g_{t-1}}{g}\right) + \varepsilon_t^G, \qquad 0 \leq \lambda_G < 1, \quad \varepsilon_t^G \sim iid\left(0, \sigma_G^2\right) \tag{27}$$

where G_t is real per capita government spending.

The central bank sets a short-term nominal interest rate, r_t^*. We consider two alternative specifications of monetary policy. First, we assume that the behaviour of the central bank can be represented by a simple instrument rule. Specifically, the central bank sets the interest rate according to a rule which in its log-linearised version takes the form

$$r_t^* = \omega_r r_{t-1}^* + (1 - \omega_r)\left[\omega_\pi \pi_t + \omega_y gdp_t + \omega_{rer} rer_t\right], \tag{28}$$

where π_t is the aggregate inflation rate, and rer_t is the real exchange rate defined as $\ln\left(S_t P_t^f / P_t\right)$. The parameter $\omega_r \in [0, 1\rangle$ determines the degree of interest rate smoothing. Output (gdp_t) is measured in deviation from the stochastic productivity trend[8], the remaining variables are in deviation from their steady-state levels.

The alternative assumption about monetary policy is that the central bank sets the interest rate to minimise the intertemporal loss function.

$$E_t \sum_{i=0}^{\infty} \beta^i \left[\pi_{t+i}^2 + \omega_y^2 \left(gdp_{t+i}\right)^2 + \omega_{\Delta r}(r_{t+i}^* - r_{t+i-1}^*)^2\right]. \tag{29}$$

As argued by e.g., Holmsen et al. (2007) including this interest rate changes in the loss function is necessary in order to produce interest rate paths that do not look immediately

[7] See e.g., Laxton & Pesenti (2003) for a discussion of this specification of the intermediation cost.

[8] Empirically, and under both assumptions about monetary policy, this measure of the output gap turns out to be quite similar to the output gap obtained using a standard Hodrick-Prescott filter which again resembles the preferred measure of the output gap published by Norges Bank.

unacceptable. Adolfson et al. (2009) also include an interest-rate smoothing term in their loss function. The central bank minimises the loss function subject to the log-linearised first-order conditions of the private sector and the exogenous shock processes.

For both specifications of monetary policy we assume that the interest rate that enters into the decisions of households and firms, r_t, equals the interest rate set by the monetary policy authority, r_t^*, plus a shock, z_t^r, that is

$$r_t = r_t^* + z_t^r \tag{30}$$

where

$$\ln\left(\frac{z_t^r}{z^r}\right) = \lambda_r \ln\left(\frac{z_{t-1}^r}{z^r}\right) + \varepsilon_t^r, \qquad 0 \leq \lambda_r < 1, \quad \varepsilon_t^r \sim iid\left(0, \sigma_r^2\right) \tag{31}$$

This shock could be interpreted e.g., as variations in the banks interest rate margins or in the spread between the key policy rate and the short-term interest rate in the money market.

Note that we depart from the conventional set-up by excluding the monetary policy shock from the instrument rule. This reflects the fact that there is no obvious equivalent to the monetary policy shock in the model with optimal policy. In order to make the model comparison as transparent as possible, we would like the two models to include the same number of stochastic shocks. Instead of taking out the monetary policy shock from the instrument rule model, we could of course have added some sort of 'monetary policy' shock to the optimal policy model (as in e.g., Adjemian et al. (2008)). The problem, however, is that there is no unique way of doing this.

Foreign variables The foreign variables that enter the model are the real marginal cost of foreign firms, mc_t^f, the output gap, y_t^f, the interest rate r_t^f and the inflation rate π_t^f. There are two shocks originating in the foreign economy.[9] Specifically, foreign marginal costs and the output gap in the foreign economy are assumed to follow first-order autoregressive processes

$$\ln\left(\frac{mc_t^f}{mc^f}\right) = \lambda_{mc^f} \ln\left(\frac{mc_{t-1}^f}{mc^f}\right) + \varepsilon_t^{mc^f}, \qquad 0 \leq \lambda_{mc^f} < 1, \quad \varepsilon_t^{mc^f} \sim iid\left(0, \sigma_{mc^f}^2\right) \tag{32}$$

$$y_t^f = \lambda_{y^f} y_{t-1}^f + \varepsilon_t^{y^f}, \qquad 0 \leq \lambda_{y^f} < 1, \quad \varepsilon_t^{y^f} \sim iid\left(0, \sigma_{y^f}^2\right) \tag{33}$$

Model solution To solve the model we first transform the model into a stationary representation by detrending the relevant real variables by the permanent technology shock. Next, we take a first-order approximation (in logs) of the equilibrium conditions around the steady-state. In the computation of the optimal policy we treat the model as exactly linear. Following the exposition in Juillard & Pelgrin (2007), the equilibrium conditions of

[9] As we have not included shocks to foreign inflation or the interest rate in the model, we cannot separate the risk premium shock in the UIP condition from a foreign interest rate shock. Moreover, all movements in the real exchange rate will be attributed to shocks affecting the nominal exchange rate or the domestic inflation rate.

the model can be written

$$F_+ E_t x_{t+1} + F_0 x_t + F_- x_{t-1} + G r_t^* + H\varepsilon_t = 0 \tag{34}$$

where x_t is a vector of endogenous variables, r_t^* is the key policy rate and ε_t is the vector of white noise disturbances. Letting $z_t = \begin{bmatrix} x_t' & r_t^* \end{bmatrix}'$ we can rewrite the intertemporal loss function (29) as

$$\frac{1}{2} E_0 \sum_{t=0}^{\infty} \beta^t z_t' W z_t \tag{35}$$

the Lagrangian of the optimal policy problem can be expressed as

$$L = E_0 \sum_{t=0}^{\infty} \beta^t \begin{bmatrix} \frac{1}{2}\left(x_t' W_{xx} x_t + 2x_t' W_{xr} r_t^* + r_t^{*\prime} W_{rr} r_t^*\right) + \\ \lambda_t' \left(F_+ E_t x_{t+1} + F_0 x_t + F_- x_{t-1} + G r_t^* + H\varepsilon_t\right) \end{bmatrix} \tag{36}$$

or, alternatively, in matrix form, as

$$L = E_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2} z_t' W z_t + \lambda_t' \left(\begin{bmatrix} F_+ & 0 \end{bmatrix} E_t z_{t+1} + \begin{bmatrix} F_0 & G \end{bmatrix} z_t + \begin{bmatrix} F_- & 0 \end{bmatrix} z_{t-1} + H\varepsilon_t \right) \right] \tag{37}$$

The first-order conditions can be written as

$$W z_t + \beta^{-1} \begin{bmatrix} F_+' \\ 0' \end{bmatrix} \lambda_{t-1} + \begin{bmatrix} F_0' \\ B' \end{bmatrix} \lambda_t + \beta \begin{bmatrix} F_-' \\ 0' \end{bmatrix} E_t \lambda_{t+1} = 0 \tag{38}$$

and

$$\begin{bmatrix} F_+ & 0 \end{bmatrix} E_t z_{t+1} + \begin{bmatrix} F_0 & G \end{bmatrix} z_t + \begin{bmatrix} F_- & 0 \end{bmatrix} z_{t-1} + H\varepsilon_t = 0 \tag{39}$$

with $\lambda_0 = 0$ and x_0 given. Inspection of equations (38) and (39) reveals that this is a linear rational expectations model expanded with difference equations for the Lagrange multipliers that can be solved using standard techniques.

Notice that the optimal commitment rule involves treating the first period differently from subsequent periods. When setting the interest rate in the first period, the policy maker takes the expectations of the private sector as given and is not constrained by any previous commitments. This is reflected in the initial value of the Lagrange multiplier being zero. The optimal commitment policy is time-inconsistent; for all periods $t > 0$ the policy maker will have an incentive to deviate from the previously announced path and exploit the private sector expectations. To overcome this 'initial value' problem Woodford (1999) proposes instead that the policy maker behaves as if the commitment to the optimal policy was made far in the past. This approach is referred to as 'timeless perspective commitment'. To compute optimal policy projections under commitment in a timeless perspective one must provide initial values for the Lagrange multipliers. See Juillard & Pelgrin (2007), Ilbas (2008a) and Adolfson et al. (2009) for alternative methods to compute these initial values.

In this paper we simply assume that monetary policy has been conducted optimally under commitment since the start of the estimation period and that the central bank never re-optimizes. The unobserved state variables, including the Lagrange multipliers, are initialized at zero which correspond to the steady-state values of the variables. When the effect of the initial conditions have died out, the optimal commitment policy will coincide with the timelessly optimal policy. Following the suggestion in Ilbas (2008a), we also experimented with using a presample approach to initialise the multipliers. Our experience is that the estimation results are not much affected by how we initialize the multipliers in the estimation.

Adopting the notation in Fernández-Villaverde et al. (2007), the transition equations describing the model solution can be expressed in state-space form as

$$\begin{aligned} Z_{t+1} &= A(\theta)Z_t + B(\theta)\varepsilon_t \\ Y_t &= C(\theta)Z_t + D(\theta)\varepsilon_t \end{aligned} \tag{40}$$

where Y_t is a $k \times 1$ vector of variables observed by the econometrician. In the case of optimal commitment policies, the state vector Z_t will contain the Lagrange multipliers associated with the behavioural equations of the private sector and the structural shock processes. The matrices A, B, C and D are non-linear functions of the structural parameters in the DSGE model as represented by the vector θ. In this paper we focus on the case with an equal number of shocks and observable variables so that D is square and invertible.

In the DSGE-VAR approach, the finite-order VAR approximation to the DSGE model plays a key role. Fernández-Villaverde et al. (2007) show that iff the eigenvalues of $A - BD^{-1}C$ are strictly less than one in modulus, Y_t has an infinite-order VAR representation given by:[10]

$$Y_t = \sum_{j=1}^{\infty} C(A - BD^{-1}C)^{j-1}BD^{-1}Y_{t-j} + D\varepsilon_t \tag{41}$$

In general, a finite-order VAR is not an exact representation of the linearised DSGE model. Specifically, the finite order VAR approximation will only be exact if all the endogenous state variables are observable and included in the VAR (see e.g., Ravenna (2007)). The rate at which the autoregressive coefficients converge to zero is determined by the largest eigenvalue of $A - BD^{-1}C$. If this eigenvalue is close to unity, a low order VAR is likely to be a poor approximation to the infinite-order VAR implied by the DSGE model.

2.1 Empirical results

This section documents the estimation results for the two DSGE models that differ only in their assumptions regarding monetary policy. First, we describe the data and the estimation method. Then we document estimation results based on the full sample, before we turn to the out-of-sample forecasting exercise.

[10]If one or more of the eigenvalues of $A - BD^{-1}C$ are exactly equal to one in modulus, Y_t does not have a VAR representation, i.e., the autoregressive coefficients do not converge to zero as the number of lags tends to infinity. Often, roots on the unit circle indicate that the observables have been overdifferenced. Fernandez-Villaverde et al (2007) refer to this as a 'benign borderline case'.

2.1.1 Data and estimation method

The model is estimated on quarterly, seasonally adjusted data for the Norwegian economy covering the period from 1987Q1 to 2007Q4. The sample period available for presample estimation is 1981Q4-1986Q4. The estimation is based on the following eleven variables: GDP, private consumption, business investment, exports, the real wage, the real exchange rate, overall inflation, imported inflation, the 3-month nominal money market rate, the overnight deposit rate (the policy rate) and hours worked. Since the model predicts that domestic GDP, consumption, investment, exports and the real wage are non-stationary, these variables are included in first differences. We take the log of the real exchange rate and hours worked.

The data series relate to the mainland economy, that is, the total economy excluding the petroleum sector. The series for GDP, exports, consumption, business investment and hours worked are measured relative to the size of the working age population (16-74 yrs.). The real wage is measured as total wage income per hour divided by the private consumption deflator. The quarterly series for growth in wage income per hour is obtained by taking a linear interpolation of the annual series from the national accounts. The nominal exchange rate is an effective import-weighted exchange rate based on the bilateral exchange rates of the Norwegian krone versus 44 countries. Consumer price inflation is measured as the total CPI adjusted for taxes and energy (CPI-ATE), and imported inflation is measured as the inflation rate for imported goods in the CPI-ATE. The money market rate is the 3 months effective nominal money market rate (NIBOR). All the series are demeaned prior to estimation.

The choice of information set is based on data availability and on the perceived quality of the data series as well as a desire to obtain good estimates of the structural parameters in the DSGE model.[11] In general, the issue of parameter identification points to including a large number of variables in the information set.[12] Within the context of a DSGE-VAR, however, the price of working with a large set of variables is that the size of the VAR becomes large relative to the sample size, resulting in imprecise estimates of the VAR parameters and wide forecast error bands. In particular, the VAR becomes much larger than what is typically used in standard forecasting applications.[13]

We estimate the DSGE models from a Bayesian perspective. The estimation of the DSGE model is based on the state-space representation (40). The likelihood function is evaluated using the Kalman filter and we use a Metropolis-Hastings (MH) algorithm to draw from the posterior distribution of the structural parameters starting from the posterior mode of the parameters computed in a first step. The full-sample results reported below are based on 3 million draws from the posterior distribution. In the forecasting experiment, the number of draws in each recursion is 100000.[14]

[11] E.g., due to perceived poor quality of the national accounts data, imports are not used as an observable variable.

[12] See e.g., the discussion in Adolfson et al. (2007a).

[13] For example, a typical VAR for a small open economy contains a measure of real activity, inflation, the exchange rate and the interest rate in addition to foreign variables.

[14] The results are obtained using Dynare (see http://www.cepremap.cnrs.fr/dynare/) and our own Matlab

The shape, the mean and the standard deviation of the prior distributions for the estimated parameters are given in tables 3 and 4. Priors for the means are partly taken directly from other studies and partly chosen in order to provide shock responses that are consistent with our prior beliefs on the transmission mechanism of the Norwegian economy. Note that we apply the same priors independent of the choice of monetary policy. This is meant to reflect the somewhat heroic assumption that these parameters are truly structural. Another way to choose the priors, would be to follow the approach of Del Negro & Schorfheide (2008a). In our case, their approach would imply having different sets of priors for the structural parameters depending on the choice of monetary policy. However, since we deal explicitly with the issue of misspecification in the DSGE-VAR set-up, it makes sense to assume that the priors on the structural parameters are independent of the policy assumptions.

Some of the parameters were fixed at the outset. This can be interpreted as a very strict prior, where all the probability mass is concentrated on a single value. The steady-state per capita growth rate π_z is calibrated to equal 2.25 per cent on an annualised basis. Based on current estimates,[15] we assume a long-run annual real interest rate of 2.5 per cent. Consistent with this, we set the discount factor β to 0.9994. The quarterly depreciation rate of capital is set to 1.8 per cent, which is in line with the recent estimates from the national accounts. The steady-state elasticity of substitution between differentiated intermediate goods, θ and θ^* is set to 6 corresponding to a price mark-up on marginal cost of 20 per cent. The home bias parameter,[16] η, is set close to 0.65 to ensure a steady state import share of roughly 30 per cent, and the elasticity of substitution between capital and labour, ξ, is set to 0.7, which yields a steady state wage income share of 0.6. The utilization cost parameter, ϕ_2^u, is set to 0.38.

Some parameters, such as the parameters related to investment costs, ϕ_1^I and the adjustment cost parameter in export prices ϕ^{Mf} turned out to be difficult to identify. Furthermore, it is not possible to identify both intermediation cost parameters ϕ_1^B and ϕ_2^B, using a first order approximation of the model. We therefore set $\phi_1^B = \phi_1^I = \phi^{Mf} = 1$.

2.1.2 Full-sample estimation results

Table 1 reports measures of the in-sample fit of the DSGE model for alternative assumptions about the conduct of monetary policy. The marginal data density is measured using the modified harmonic mean estimator proposed by Geweke (1999). A key result is that the model with a simple instrument rule is clearly dominated by the model with optimal policy in terms of in-sample fit. Hence, the implicit rule following from the assumption of optimal monetary policy appears to give a more accurate description of the way monetary policy was conducted over the sample period than does a simple instrument rule. However, this result depends to a large extent on the symmetric treatment of the shock processes in the

codes for estimating of DSGE models with optimal policy under commitment and forecasting with a DSGE-VAR.

[15] See Norges Bank's Inflation Report 2/06.

[16] This parameter represents the share of domestic intermediates in the final goods aggregate that would prevail in the hypothetical case where the prices on domestic and imported goods were equal.

two models. Including a policy shock in the simple rule brings the marginal data density much closer to the model with optimal policy. Adolfson et al. (2009) also make the point that the ranking of the models in terms of in-sample fit will depend on whether one or both of the models include a monetary policy shock. E.g., they find that when the instrument rule includes a monetary policy shock, but the model with optimal policy does not, the model with a simple instrument rule gives a better fit. Note, however, that the instrument rule considered in Adolfson et al. (2009) is somewhat more flexible than the instrument rule considered in this paper; in addition to the level variables, it includes both the change in inflation and in the growth rate of GDP.

Turning to the parameter estimates, table 2 reports the estimates of the monetary policy preferences from the DSGE model. The estimates imply a high relative weight on interest rate changes in the loss function. The posterior estimates of the remaining parameters are reported in tables 3 and 4. Comparing the DSGE models, the parameter estimates do not seem to be significantly influenced by the choice of monetary policy, consistent with the finding in Adolfson et al. (2009). This conclusion is supported by the impulse responses of the estimated shocks, which appear fairly similar. However, there is one notable exception to this conclusion. The stickiness of domestic good prices is estimated to be significantly higher in the model employing a simple instrument rule. As we shall see in the next section, this could potentially explain the differences in the forecasting properties of the two models, in particular with respect to inflation and the interest rate.

2.1.3 Forecast comparison

The forecast experiment is constructed as follows. We estimate each model on a sample period ending 1998Q4 and compute forecasts for horizons of one up to twelve quarters. We then extend the sample by one quarter, demean the data, re-estimate the models and compute new forecasts. The implicit steady-states of the variables are allowed to vary over time; we demean the data prior to estimation in each recursion. This exercise is repeated until the end of the sample. All the parameters in the DSGE model are re-estimated in each recursion. The forecasts are based on 100000 MH draws starting from the posterior mean of the previous recursion.

We measure forecasting accuracy by univariate root mean squared forecast error (RMSE). The point forecasts used to calculate the RMSEs are the posterior means of the forecast draws. Following Adolfson et al. (2007c) we also compute a measure of multivariate forecast accuracy, namely the trace of the mean squared forecast error (MSE) matrix for horizon h. The MSE matrix is denoted $\Omega_M(h)$ and is defined as

$$\Omega_M(h) = \frac{1}{N_h} \sum_{t=T}^{T+N_h-1} \left(Y_{t+h} - \widehat{Y}_{t+h|t}\right) M^{-1} \left(Y_{t+h} - \widehat{Y}_{t+h|t}\right)', \tag{42}$$

where N_h is the number of forecasts and M is a diagonal matrix with the sample variances of the variables as diagonal elements. For the variables that enter the model in growth rates, we follow Del Negro et al. (2007a) and report the RMSE for the cumulative changes in the variables.

Figure 1 plots the univariate RMSEs from the DSGE model under the different assumptions about monetary policy. The ranking of the models is less clear than was the case when using measures of in-sample fit based on the full sample. In terms of forecasting accuracy the models perform about equally well. The model with optimal policy produces more accurate forecasts of the growth rates of GDP, consumption and investment, whereas the model with a simple instrument rule produces more accurate forecasts of the inflation and interest rates. We conjecture that one reason why the model with a simple instrument rule produces more accurate forecasts of inflation and interest rates is that price stickiness parameter is estimated to be higher in this version of the model, giving rise to weaker equilibrium-correction, which is an inherent feature of both interest rates and inflation over the out-of-sample period.

As a next step we compare the model projections of inflation and the interest rate with the official Norges Bank projections. The exercise is somewhat restricted by the fact that official forecasts are only available from 2005 onwards, and the fact that forecasts are published only three times per year, however, we still believe that it provides some interesting insights. Figure 4 shows the DSGE forecasts and the official forecasts for each quarter in the period 2005q3-2008q2.[17] As is evident from the figures, both versions of the DSGE model consistently predict a sharper increase in interest rates than the official forecasts. This is especially true for the model assuming optimal policy. Furthermore, we observe that Norges Banks official forecasts are more in line with the actual interest path. However, in contrast to the DSGE models, there seems to be a slight tendency for the Norges Bank forecast to under-predict the actual interest path. The differences between the model forecasts and the official Norges Bank forecasts reflect to some extent the use of judgment and off-model considerations in arriving at the final projections. This can be interpreted as an attempt to correct for model misspecification.

3 Acknowledging model misspecification

In the above exercise, we implicitly assume that, under each of the two approaches to modelling monetary policy, the resulting theoretical model provides an accurate probabilistic description of our data. In this section, we assess the robustness of our results to model misspecification using the DSGE-VAR approach proposed by Del Negro & Schorfheide (2004). The DSGE-VAR approach allows us to relax the tight cross-equation restrictions implied by the DSGE model for the parameters in a VAR. The DSGE-VAR approach also produces estimates of the parameters in the DSGE model that can be compared to those obtained using the traditional full-information approach.

[17] Norges Bank publishes forecasts three times a year. To compare these forecasts to the forecasts from our quarterly model we have added a "synthetic" forecast round with forecasts equal to the previously published path. In general, the forecasts made by Norges Bank are made later in time and in that sense incorporate more information than the model forecasts.

3.1 The DSGE-VAR approach

As alluded to in the introduction, the basic idea of the DSGE-VAR approach is to use the DSGE model to construct prior distributions for the VAR. The starting point for the estimation is an unrestricted VAR of order p

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + u_t, \tag{43}$$

where Y_t is an $n \times 1$ vector of observables, ϕ_0 is an $n \times 1$ vector of constant terms, ϕ_i are $n \times n$ matrices of autoregressive parameters $i = 1, \ldots, p$ and $u_t \sim N(0, \Sigma_u)$. If we let the vector of regressors in the VAR be denoted $x_t = [1, y_{t-1}, y_{t-2}, \ldots, y_{t-p}]$, the VAR can be written compactly as

$$Y = X\Phi + U, \tag{44}$$

where Y is $T \times n$ with rows y_t', X is $T \times (1 + np)$ with rows x_t' , U is $T \times n$ with rows u_t' and $\Phi = \left[\phi_0', \phi_1', \ldots, \phi_p'\right]$. The likelihood function for the VAR is given by

$$\begin{aligned} p(Y|\Phi, \Sigma_u) \quad &\propto \quad |\Sigma_u|^{-T/2} \\ &\times \exp\left\{-\frac{1}{2}\mathrm{tr}\left[\Sigma_u^{-1}\left(\begin{array}{c} Y'Y - \Phi'X'Y \\ -Y'X\Phi + \Phi'X'X\Phi \end{array}\right)\right]\right\} \end{aligned} \tag{45}$$

The prior distribution for the VAR parameters proposed by Del Negro & Schorfheide (2004) is based on the VAR approximation to the DSGE model. Let Γ_{xx}^*, $\Gamma_{yy}^*, \Gamma_{xy}^*$ and Γ_{yx}^* be the theoretical second-order moments of the variables in Y and X implied by the DSGE model. Then

$$\begin{aligned} \Phi^*(\theta) &= \Gamma_{xx}^{*-1}(\theta)\Gamma_{xy}^*(\theta) \\ \Sigma_u^*(\theta) &= \Gamma_{yy}^*(\theta) - \Gamma_{yx}^*(\theta)\Gamma_{xx}^{*-1}(\theta)\Gamma_{xy}^*(\theta) \end{aligned} \tag{46}$$

can be interpreted as the probability limits of the coefficients in a VAR estimated on artificial observations generated by the DSGE model. Conditional on the vector of structural parameters in the DSGE model θ, the prior distribution for the VAR parameters $p(\Phi, \Sigma_u|\theta)$, is of the Inverted-Wishart ($\mathcal{IW}$) - Normal ($\mathcal{N}$) form

$$\begin{aligned} \Sigma_u|\theta &= \mathcal{IW}\left(\lambda T \Sigma_u^*(\theta), \lambda T - k, n\right) \\ \Phi|\Sigma_u, \theta &= \mathcal{N}\left(\Phi^*(\theta), \Sigma_u \otimes \left(\lambda T \Gamma_{xx}^*\right)^{-1}\right) \end{aligned} \tag{47}$$

where $k = 1+np$. The tightness of the prior distribution is governed by the hyperparameter $\lambda \in [0, \infty]$. This hyperparameter can be loosely interpreted as the size of the sample of artificial or dummy observations generated by the DSGE model relative to the size of the actual sample in the estimation.

The posterior distribution of the VAR parameters is also of the Inverted-Wishart -

Normal form (see Del Negro & Schorfheide (2004))

$$\begin{aligned}\Sigma_u|Y,\theta &= \mathcal{IW}\left((\lambda+1)\,T\widetilde{\Sigma}_u(\theta),(1+\lambda)\,T-k,n\right) \\ \Phi|Y,\Sigma_u,\theta &= \mathcal{N}\left(\widetilde{\Phi}(\theta),\Sigma_u\otimes\left(\lambda T\Gamma^*_{xx}+X'X\right)^{-1}\right)\end{aligned} \tag{48}$$

The matrices $\widetilde{\Phi}(\theta)$ and $\widetilde{\Sigma}_u(\theta)$ have the interpretation of maximum likelihood estimates of the VAR parameters based on the combined sample of actual observations and artificial observations generated by the DSGE model, that is

$$\widetilde{\Phi}(\theta) = \left(\lambda T\Gamma^*_{xx}+X'X\right)^{-1}\left(\lambda T\Gamma^*_{xy}(\theta)+X'Y\right) \tag{49}$$

$$\begin{aligned}\widetilde{\Sigma}_u(\theta) &= \frac{1}{(\lambda+1)T}\left(\lambda T\Gamma^*_{yy}(\theta)+Y'Y\right) \\ &\quad -\frac{1}{(\lambda+1)T}\left(\lambda T\Gamma^*_{yx}(\theta)+Y'X\right)\left(\lambda T\Gamma^{*-1}_{xx}(\theta)+X'X\right)^{-1}\left(\lambda T\Gamma^*_{xy}(\theta)+X'Y\right)\end{aligned} \tag{50}$$

From the above expressions we see that if λ is small, the prior on the DSGE model restrictions is diffuse. In particular, setting $\lambda=0$ we would retrieve the unrestricted OLS estimates. Notice, however, that in order for the prior distribution (47) to be proper, λ has to take a value larger than $\lambda_{\min}=(k+n)/T$ (see e.g., Adolfson et al. (2007b)). The higher is λ, the more the VAR estimates will be tilted towards the parameters in the VAR approximation of the DSGE model ($\Phi^*(\theta)$ and $\Sigma^*_u(\theta)$). Del Negro et al. (2007a) choose λ by maximising the marginal data density $p_\lambda(Y)$ over a pre-specified grid for λ. In this paper we specify a uniform distribution for λ over the interval $[\lambda_{\min},\infty)$.

The specification of the VAR prior is completed with the specification of prior distributions for the DSGE model parameters θ. The DSGE-VAR approach allows us to draw posterior inferences about the DSGE model parameters θ. As explained by Del Negro & Schorfheide (2004), the posterior estimate of θ has the interpretation of a minimum-distance estimator, where the minimand or distance function is given by the discrepancy between the unrestricted OLS estimates of the VAR parameters and the coefficients in the VAR approximation to the DSGE model, the latter being functions of θ. Obviously, then, the posterior estimates of θ will depend on the hyperparameter λ. In the limit $\lambda\to 0$, there will not be any information about θ in $p(Y|\theta)$, and hence, the posterior estimates of θ will be equal to the prior estimates.

3.2 Empirical results

The estimation of the DSGE-VAR is based on the MH algorithm to draw from the joint posterior distribution of ϕ,Σ_u,θ described in Del Negro & Schorfheide (2004). An important modelling choice for the DSGE-VAR is the choice of lag length. As argued by Del Negro et al. (2007b) there are essentially two dimensions to the choice of lag length for a DSGE-VAR. The first dimension is related to the accuracy of the VAR approximation to the DSGE model. This suggests we choose the lag-length to minimise the approximation error, that is, to minimise the discrepancy between the dynamics of the DSGE-VAR(∞)

and the dynamics of the DSGE model. Since, in general, the accuracy of the approximation increases with lag length, this criterion points to having a fairly large number of lags. In the previous literature (see e.g., Adolfson et al. (2007b) and Del Negro et al. (2007a)), the lag-length has commonly been set to four based on this criterion. The second dimension to the choice of lag length is the empirical fit of the DSGE-VAR with the optimal value of λ,that is the DSGE-VAR($\widehat{\lambda}$). This suggests that we choose the lag length to maximise the marginal data density associated with the DSGE-VAR($\widehat{\lambda}$). As emphasized by Del Negro et al. (2007a), there is no requirement that the auxiliary model (the DSGE-VAR) nests the underlying theoretical model (the VAR approximation to the DSGE model) for the exercise to be meaningful. For our model(s), we find that the marginal data density is maximised for the model with two lags. The optimal value of the hyperparameter, λ, is smaller in the model with two lags compared with the model with four lags, however. This reflects that the gains from shrinking towards the theoretical model are smaller in the former case, since there are fewer free parameters in the VAR. Similar findings were reported by Del Negro & Schorfheide (2008b).

3.2.1 Full-sample estimation results

Table 1 reports the posterior mean of the hyperparameter λ in the DSGE-VAR and the marginal data densities for the two specifications of monetary policy. The estimated weight on the DSGE model in the DSGE-VAR is higher in the case of optimal policy than in the model with a simple rule (the posterior mean of the hyperparameter λ is 1.14 in the case of optimal policy and 0.89 in the model with a simple rule). We also see that the fit of the model is improved if we shrink the VAR parameters towards the restrictions implied by the DSGE model, or, alternatively, if we relax the DSGE model restrictions in the direction of the unrestricted VAR estimates. That is, the marginal data density is higher for the DSGE-VAR than for the DSGE model. This is true under both assumptions about monetary policy. In the next subsection we examine whether this holds true also in terms of out-of-sample forecasting performance.

Table 2 reports the estimates of the monetary policy parameters obtained using the DSGE-VAR approach. The parameters in the loss-function do not appear to be much affected by allowing for model misspecification. This does not hold true for the parameters in the simple instrument rule: the weight on inflation increases significantly once we allow for misspecification. In this sense, the optimal policy framework appear to be more robust to misspecification than a model with a simple Taylor-type rule.

As evidenced in tables 3 and 4, the estimates of the other parameters in the model differ even less than for the DSGE models. This indicates that part of the differences in the estimated DSGE parameters are due to misspecification. One way to think about this is that misspecification adds an extra source of variation to the estimated parameters. Another robust finding is that the degree of external persistence as measured by the first-order autocorrelation of the exogenous shock processes is reduced significantly once misspecification is taken into account. It is clear from table 4 that both the autocorrelation coefficient and the standard deviation of the shock processes are in general lower in the

DSGE-VAR models than the DSGE models. Hence, taking into account misspecification reduces the need for exogenous persistence.[18]

3.2.2 Forecast comparison

In addition to the DSGE-VAR forecasts we compute forecasts from a Bayesian VAR (BVAR) with a Minnesota-type prior. The prior in the BVAR will tilt the VAR towards univariate random walks of the variables in levels. The lag-length in the BVAR is set to two. All the parameters in the BVAR and DSGE-VARs, including the hyperparameter λ, are re-estimated in each recursion. The forecasts are based on 100000 MH draws starting from the posterior mean of the previous recursion. Figure 2 compares the univariate RMSEs from the DSGE model with optimal policy to those obtained using a DSGE-VAR approach and the BVAR. For most variables, relaxing the cross-equation restrictions in the DSGE model towards an unrestricted VAR improves the forecasting performance. However, in terms of forecasting performance the DSGE-VAR models are inferior to the BVAR with a Minnesota prior. This findings is confirmed in figure 3 which reports a multivariate measure of forecast accuracy. The fact that the BVAR outperforms the DSGEs is perhaps not surprising. The BVAR prior tilts the unrestricted VAR towards univariate random walks. Given that inflation and interest rates are only borderline stationary in our sample, this seems like a very reasonable prior.

From figure 4, we note that the interest rate forecasts from the DSGE-VAR are quite close to the official forecasts. Hence, accounting for misspecification brings the model interest rate projections much more in line with the published forecasts. The same holds for inflation (see figure 5): the DSGE-VAR forecasts are closer to both actual inflation and the official projections than the DSGE model forecasts.

One tentative conclusion one could draw from this exercise is that the DSGE-VAR model mimics the combination of pure model forecasts and judgment inherent in the official Norges Bank forecasts. As noted above, the DSGE model employed in this paper is broadly similar to the core model used for policy projections at Norges Bank. However, arriving at the final official projections is a complex process, involving input from other forecasting models, add factors and off-model considerations. Our results indicate that the iterative forecasting process used by the Norges Bank can be well represented by a DSGE-VAR model, where the restrictions from the core DSGE model can be interpreted as a prior on the VAR parameters.

A notable feature of the interest rate and inflation projections from the DSGE model with optimal policy is that they 'overshoot' the long-run level in the medium run. This feature of optimal policy under commitment is less pronounced in Norges Bank's projections since 2005, and is not a feature of the DSGE-VAR forecasts. This is a sign that the model with optimal monetary policy is misspecified. One interpretation is that the Norges Bank does not fully exploit the expectations channel when setting policy, or alternatively, that it perceives the gains from commitment in the current specification of the DSGE model to

[18] Similar findings are again reported by Del Negro & Schorfheide (2008b).

be too large (e.g., that the price-setters in the model are in a sense too forward-looking).

4 Concluding remarks

The results in this paper suggest that the empirical merits of the DSGE model estimated with optimal monetary policy is comparable to the performance of a model with a simple Taylor-type rule – both with regards to in-sample and out-of-sample measures of fit. This conclusion also holds when taking account of misspecification. One way of interpreting the DSGE-VAR results, is that introducing optimal monetary policy reduces the degree of misspecification. Interestingly, in contrast to the model using a simple rule, the policy parameters in the optimal policy model appear to be quite robust to misspecification. Hence, based on our empirical findings and given the superior theoretical (and intuitive) appeal of the optimizing approach to monetary policy, we argue that the optimal policy framework should be a natural ingredient in any DSGE model describing central bank behaviour.

However, as is evidenced both by the optimal value of the DSGE-VAR hyperparameter and the forecasting performance of the different models, model misspecification remains a serious concern for the use of DSGE models in practical policy analysis. Hence, one tentative conclusion that could be drawn from our analysis is that the empirical gains from assuming optimal monetary policy is of second order relative to improving the modelling of the transmission mechanism itself. In this respect, it is interesting to note that off-model considerations appear to bring the official projections closer to the DSGE-VAR model, and, as a result, also reduce the forecast errors.

References

Adjemian, S., Pariès, M. D., & Moyen, S. (2008). *Towards a Monetary Policy Evaluation Framework*. Working Paper 942, ECB.

Adolfson, M., Laséen, S., Lindé, J., & Svensson, L. E. (2009). Optimal monetary policy in an operational medium-sized DSGE model. Mimeo June 2009, available at Lars E. O. Svensson's research page at www.iies.su.se. A previous version was published as NBER Working Paper 14092 in 2008.

Adolfson, M., Laseén, S., Lindé, J., & Villani, M. (2007a). Bayesian estimation of an open economy DSGE model with incomplete pass-through. *Journal of International Economics*, 72, 481–511.

Adolfson, M., Laséen, S., Lindé, J., & Villani, M. (2007b). Evaluating an estimated New Keynesian small open economy model. *Journal of Economic Dynamics and Control*. forthcoming.

Adolfson, M., Lindé, J., & Villani, M. (2007c). Forecasting performance of an open economy DSGE model. *Econometric Reviews*, 26, 289–328.

Brubakk, L., Husebø, T. A., Maih, J., Olsen, K., & Østnor, M. (2006). *Finding NEMO: Documentation of the Norwegian Economy Model*. Norges Bank Staff Memo 2006/6.

Christiano, L. J., Eichenbaum, M., & Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113, 1–45.

Del Negro, M. & Schorfheide, F. (2004). Priors from general equilibrium models for VARs. *International Economic Review*, 45, 643–673.

Del Negro, M. & Schorfheide, F. (2008a). Forming priors for DSGE models (and how it affects the assessment of nominal rigidities). *Journal of Monetary Economics*, 55, 1191–1208.

Del Negro, M. & Schorfheide, F. (2008b). *Inflation Dynamics in a Small Open-Economy Model under Inflation Targeting: Some Evidence for Chile*. Staff Reports 329, Federal Reserve Bank of New York.

Del Negro, M. & Schorfheide, F. (2009). Monetary policy analysis with potentially misspecified models. *American Economic Review*, 99, 1415–1450.

Del Negro, M., Schorfheide, F., Smets, F., & Wouters, R. (2007a). On the fit of New Keynesian models. *Journal of Business Economic Statistics*, 25, 123–143.

Del Negro, M., Schorfheide, F., Smets, F., & Wouters, R. (2007b). Rejoinder. *Journal of Business and Economic Statistics*, 25, 159–162.

Dennis, R. (2004). Inferring policy objectives from economic outcomes. *Oxford Bulletin of Economics and Statistics*, 66 (Supplement), 735–764.

Edge, R. M., Kiley, M. T., & Laforte, J.-P. (2010). A comparison of forecast performance between Federal Reserve staff forecasts, simple reduced-form models, and a DSGE model. *Journal of Applied Econometrics*. Forthcoming.

Fernández-Villaverde, J., Rubio-Ramírez, J. F., Sargent, T., & Watson, M. W. (2007). ABCs (and Ds) of understanding VARs. *American Economic Review*, 97, 1021–1026.

Geweke, J. (1999). Using simulation methods for bayesian econometric models: Inference, development and communication. *Econometric Reviews*, 18, 1–126.

Holmsen, A., Qvigstad, J. F., & Røisland, Ø. (2007). *Implementing and Communicating Optimal Monetary Policy*. Staff Memo 2007/3, Norges Bank.

Ilbas, P. (2008a). *Estimation of Monetary Policy Preferences in a Forward-Looking Model: A Bayesian Approach*. Working Paper Research 129, National Bank of Belgium. Forthcoming International Journal of Central Banking.

Ilbas, P. (2008b). *Revealing the Preferences of the US Federal Reserve*. Working Paper 2008/21, Norges Bank.

Juillard, M. & Pelgrin, F. (2007). *Computing Optimal Policy in a Timeless-Perspective: An Application to a Small Open Economy*. Working Paper 2007-32, Bank of Canada.

Lane, P. R. (2001). The new open economy macroeconomics: A survey. *Journal of International Economics*, 54, 235–266.

Laxton, D. & Pesenti, P. (2003). Monetary rules for small, open, emerging economies. *Journal of Monetary Economics*, 50, 1109–1146.

Ravenna, F. (2007). Vector autoregressions and reduced form representations of DSGE models. *Journal of Monetary Economics*, 54, 2048–2064.

Sims, C. A. (2003). Probability models for monetary policy decisions. Unpublished manuscript available Sims' home page at Princeton University.

Smets, F. & Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the Euro area. *Journal of the European Economic Association*, 1, 1123–1175.

Smets, F. & Wouters, R. (2004). Forecasting with a Bayesian DSGE model: An application to the Euro Area. *Journal of Common Market Studies*, 42, 841–867.

Svensson, L. E. O. (2003). What is wrong with Taylor rules? Using judgement in monetary policy through targeting rules. *Journal of Economic Literature*, 41, 426–477.

Woodford, M. (1999). *Optimal Monetary Policy Inertia*. NBER Working Paper 7261.

	Marginal data density	Weight on DSGE λ
DSGE optimal policy	3223.0	–
DSGE simple rule	3157.3	–
DSGE-VAR optimal policy	3247.8	1.1381
DSGE-VAR simple rule	3223.1	0.8929

Table 1: The fit of the DSGE and DSGE-VAR models under different assumptions about monetary policy

	Prior mean	Posterior mean	
		DSGE	DSGE-VAR
Optimal policy			
Weight on output gap ω_y	0.5	0.2508	0.2285
Weight on interest rate $\omega_{\Delta r}$	0.2	0.4400	0.4688
Simple instrument rule			
Weight on inflation ω_π	2.0	1.5031	1.7986
Weight on output gap ω_y	0.2	0.4552	0.3830
Weight on interest rate ω_r	0.8	0.6720	0.7021
Weight on real exchange rate ω_{rer}	0.0	0.0202	0.0033

Table 2: Estimates of monetary policy preferences in DSGE and DSGE-VAR with optimal monetary policy. The weight on the inflation term in the loss function is normalised to unity.

Parameter	Prior		DSGE opt	DSGE-VAR opt	DSGE simp	DSGE-VAR simp
	Type	Mean (std)	Mean (std)	Mean (std)	Mean (std)	Mean (std)
α	Beta	0.30 (0.020)	0.2973 (0.0196)	0.2908 (0.0217)	0.2968 (0.0201)	0.2844 (0.0202)
ψ	Inv gam	5.50 (0.500)	4.9056 (0.5624)	5.1585 (0.5095)	4.8298 (0.4838)	5.2486 (0.5409)
ζ	Inv gam	3.00 (0.200)	2.7867 (0.2807)	2.8721 (0.2940)	2.8622 (0.2771)	2.8984 (0.2878)
μ	Inv gam	1.10 (0.200)	1.1472 (0.0494)	1.2116 (0.0621)	1.2497 (0.0557)	1.3252 (0.0571)
μ^*	Inv gam	1.10 (0.200)	1.1864 (0.1387)	1.0665 (0.2220)	1.3298 (0.2263)	1.0856 (0.2988)
b_C	Beta	0.75 (0.050)	0.7879 (0.0450)	0.7202 (0.0395)	0.7989 (0.0348)	0.7411 (0.0333)
ϕ^M	Inv gam	1.00 (1.000)	1.7113 (0.1740)	1.8087 (0.3929)	1.7180 (0.2703)	1.8344 (0.3232)
ϕ^Q	Inv gam	1.00 (1.000)	1.9145 (0.1229)	1.8168 (0.0.4621)	2.4521 (0.4844)	2.1027 (0.9027)
ϕ^W	Inv gam	1.00 (1.000)	2.7469 (0.4183)	2.2420 (0.5950)	2.8860 (0.4940)	2.1393 (0.7644)
ϕ_2^I	Gam	10.00 (5.000)	17.7661 (2.7254)	12.2380 (3.5670)	19.180 (3.9843)	14.0437 (4.0315)
ϕ_2^B	Inv gam	0.02 (0.005)	0.0173 (0.0037)	0.0167 (0.0033)	0.0166 (0.0027)	0.0199 (0.0031)

Table 3: Posterior mean of DSGE model parameters I

Parameter	Prior		DSGE opt	DSGE-VAR opt	DSGE simp	DSGE-VAR simp
	Type	Mean (std)	Mean (std)	Mean (std)	Mean (std)	Mean (std)
σ_z	Inv gam	0.0050 (Inf)	0.0115 (0.0013)	0.0060 (0.0011)	0.0140 (0.0011)	0.0062 (0.0013)
σ_r	Inv gam	0.0025 (Inf)	0.0015 (0.0002)	0.0011 (0.0001)	0.0017 (0.0001)	0.0010 (0.0001)
σ_ψ	Inv gam	1.0000 (Inf)	0.9140 (0.2168)	0.7140 (0.1640)	1.2070 (0.2393)	0.6606 (0.2587)
σ_θ	Inv gam	1.0000 (Inf)	0.5644 (0.0617)	0.5677 (0.1292)	0.5083 (0.0967)	0.5316 (0.1316)
σ_I	Inv gam	1.0000 (Inf)	1.5296 (0.2212)	0.7492 (0.2687)	1.5836 (0.2767)	0.7876 (0.3004)
σ_L	Inv gam	0.0050 (Inf)	0.0112 (0.0013)	0.0068 (0.0010)	0.0147 (0.0012)	0.0071 (0.0013)
σ_B	Inv gam	0.0100 (Inf)	0.0028 (0.0001)	0.0039 (0.0005)	0.0031 (0.0004)	0.0038 (0.0005)
σ_{mc^*}	Inv gam	0.0100 (Inf)	0.2689 (0.0209)	0.2465 (0.0722)	0.2903 (0.0629)	0.2083 (0.0710)
σ_{y^*}	Inv gam	0.0100 (Inf)	0.0348 (0.0036)	0.0219 (0.0028)	0.0342 (0.0026)	0.0214 (0.0031)
σ_g	Inv gam	0.0100 (Inf)	0.0541 (0.0058)	0.0315 (0.0046)	0.0615 (0.0046)	0.0322 (0.0048)
σ_u	Inv gam	0.0100 (Inf)	0.0527 (0.0068)	0.0276 (0.009)	0.0593 (0.0090)	0.0293 (0.0119)
λ_z	Beta	0.8500 (0.1)	0.1316 (0.0730)	0.1224 (0.0712)	0.0762 (0.0389)	0.0957 (0.0469)
λ_ψ	Beta	0.8500 (0.1)	0.7396 (0.0757)	0.4981 (0.0405)	0.6681 (0.0495)	0.4945 (0.0426)
λ_θ	Beta	0.8500 (0.1)	0.6767 (0.1232)	0.3310 (0.0725)	0.7495 (0.0605)	0.3830 (0.0521)
λ_I	Beta	0.8500 (0.1)	0.1538 (0.0760)	0.1585 (0.0711)	0.1166 (0.0589)	0.1591 (0.0631)
λ_L	Beta	0.8500 (0.1)	0.8538 (0.0400)	0.7429 (0.0310)	0.8632 (0.0284)	0.8215 (0.0353)
λ_B	Beta	0.8500 (0.1)	0.9386 (0.0486)	0.8194 (0.0213)	0.9252 (0.0196)	0.7862 (0.0195)
λ_{mc^*}	Beta	0.8500 (0.1)	0.5617 (0.1283)	0.4181 (0.1009)	0.5060 (0.1007)	0.4479 (0.0987)
λ_{y^*}	Beta	0.8500 (0.1)	0.6125 (0.0872)	0.2548 (0.0544)	0.6773 (0.0522)	0.3879 (0.0508)
λ_g	Beta	0.8500 (0.1)	0.8840 (0.0777)	0.6830 (0.0223)	0.8497 (0.0331)	0.6489 (0.0330)
λ_r	Beta	0.8500 (0.1)	0.7844 (0.0451)	0.5317 (0.0533)	0.5251 (0.0195)	0.5017 (0.0200)
λ_u	Beta	0.8500 (0.1)	0.3143 (0.0678)	0.3902 (0.0935)	0.2913 (0.0663)	0.3296 (0.0664)

Table 4: Posterior mean of DSGE model parameters II

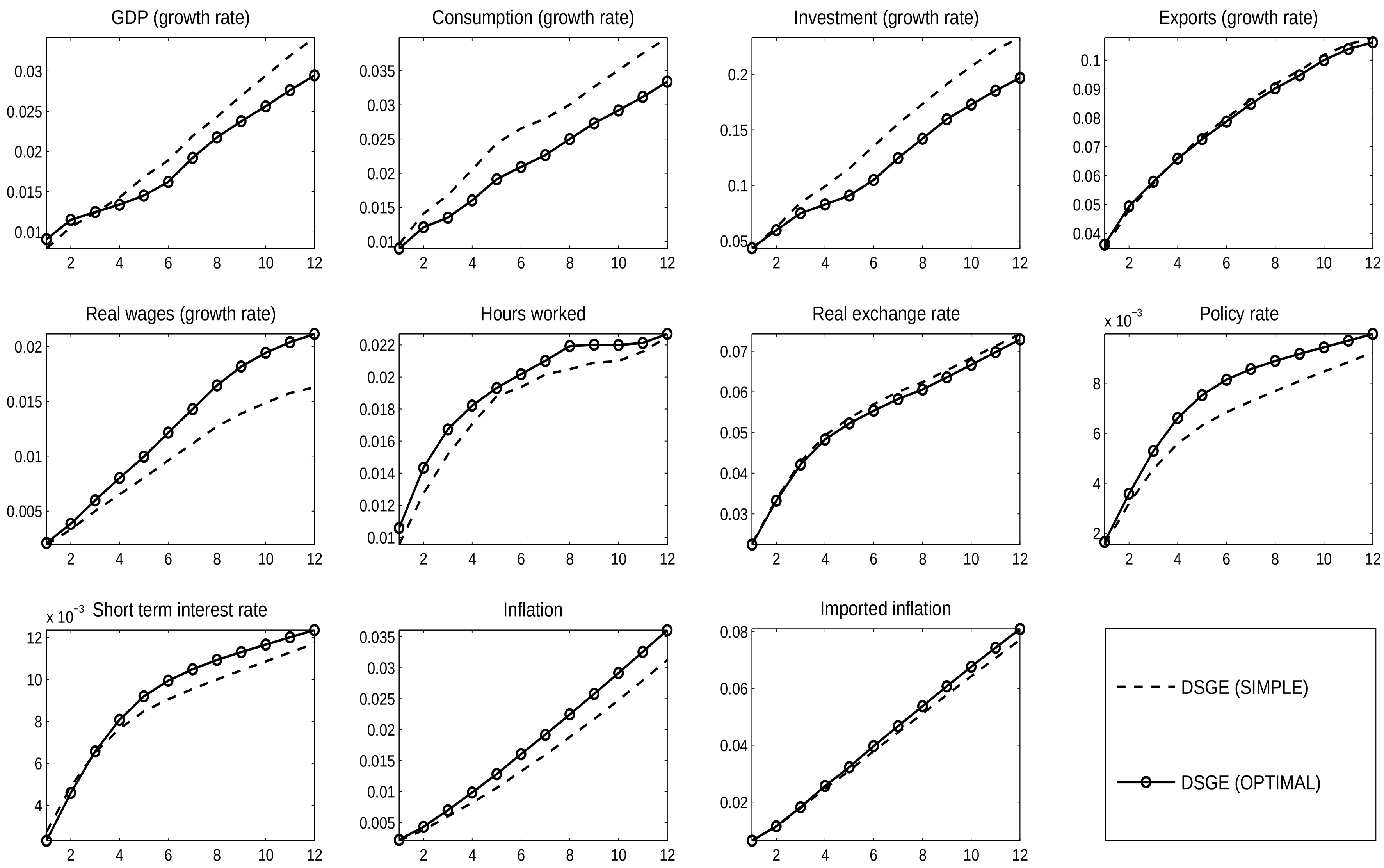

Figure 1: Univariate RMSEs for DSGE model with optimal policy and DSGE with simple instrument rule

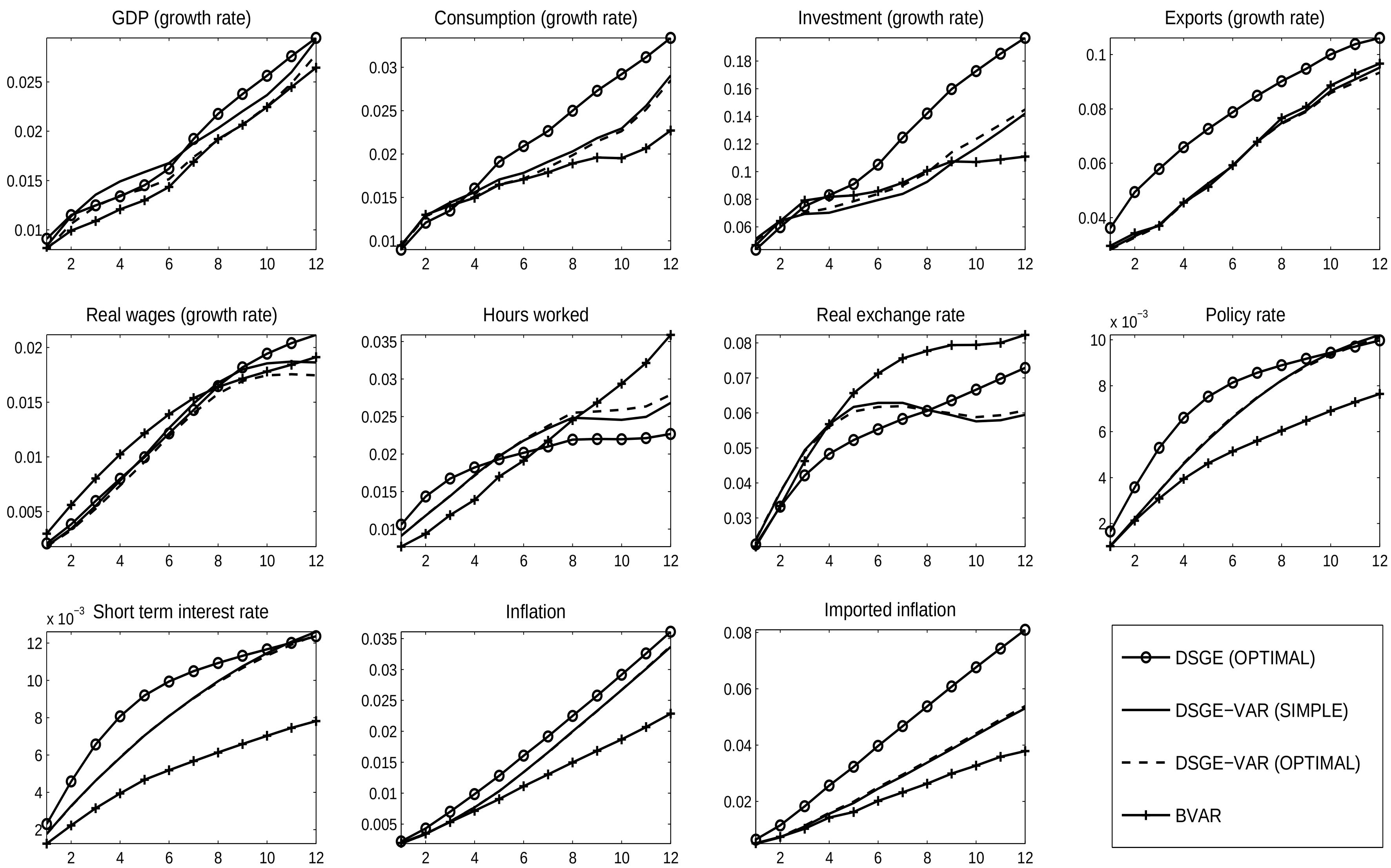

Figure 2: Univariate RMSEs for DSGE model with optimal policy, DSGE-VAR with simple instrument rule, DSGE-VAR with optimal policy and BVAR

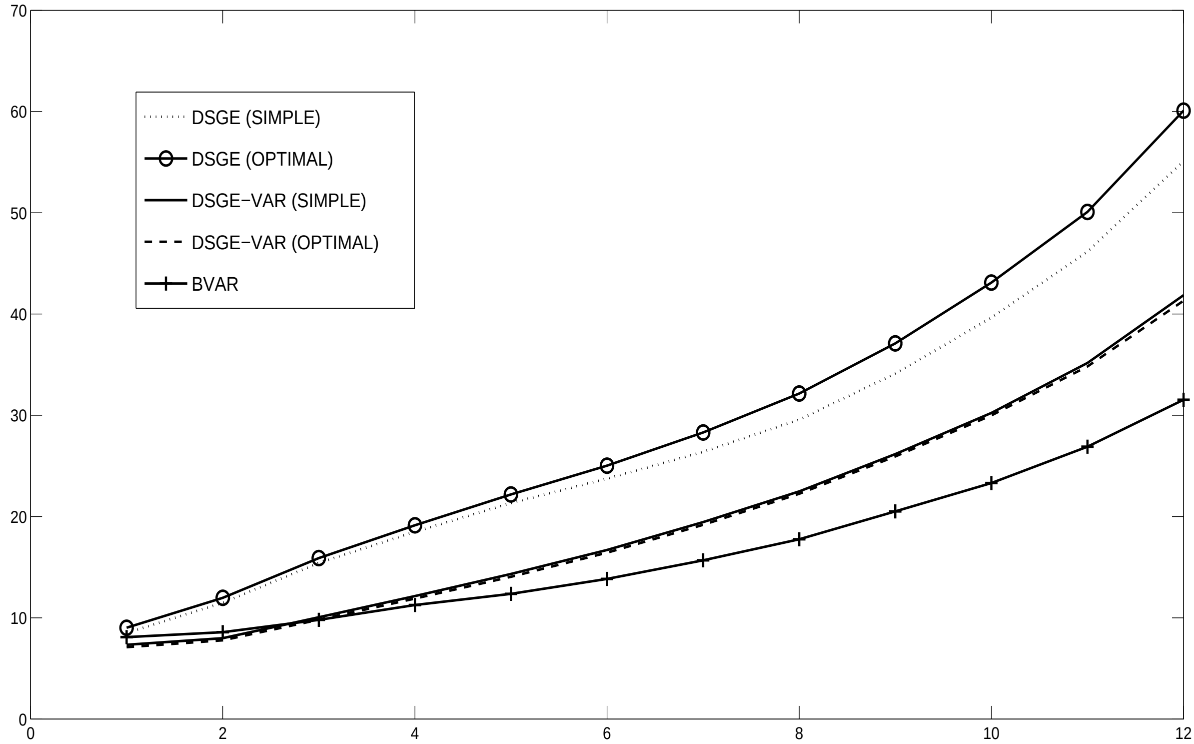

Figure 3: Multivariate trace statistic for DSGE model with optimal policy, DSGE model with simple instrument rule, DSGE-VAR with optimal policy, DSGE-VAR with simple instrument rule and BVAR

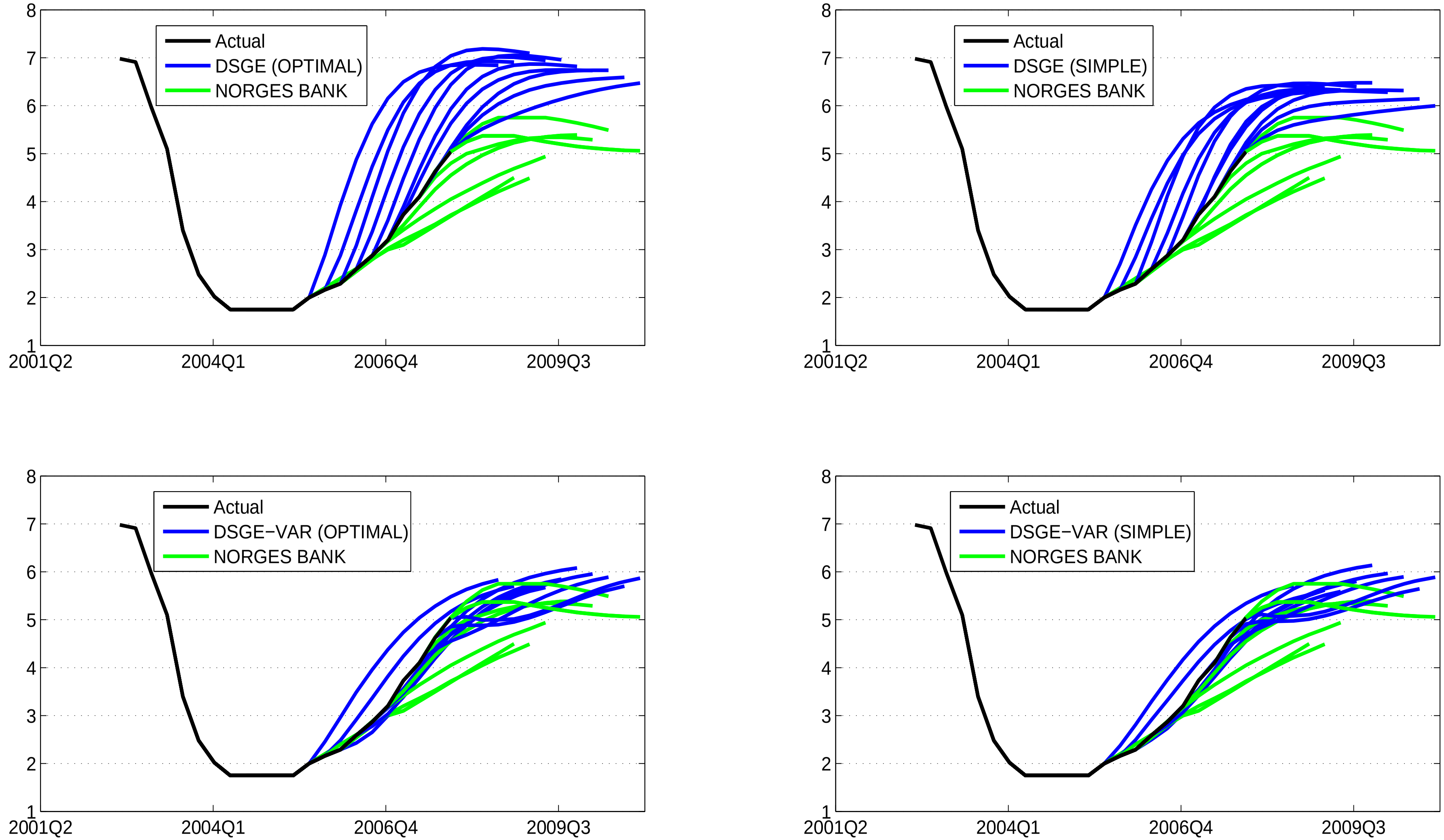

Figure 4: Actual policy rate, Norges Bank's official forecasts and model forecasts

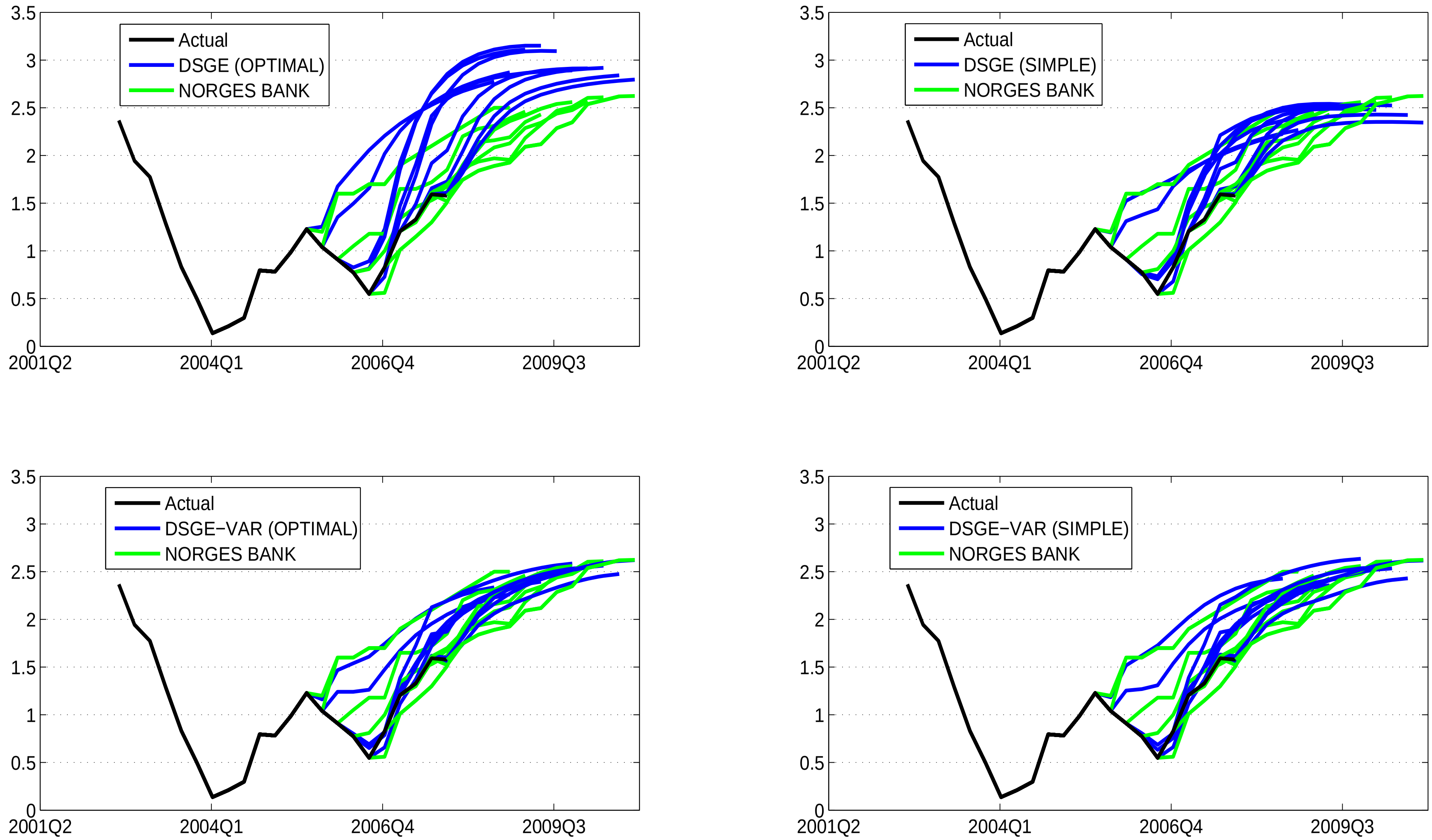

Figure 5: Actual four quarter inflation, Norges Bank's official forecasts and model forecasts