

Gerdrup, Karsten R.; Nicolaisen, Jon

Research Report

On the Purpose of Models - the Norges Bank Experience

Staff Memo, No. 06/2011

Provided in Cooperation with:

Norges Bank, Oslo

Suggested Citation: Gerdrup, Karsten R.; Nicolaisen, Jon (2011) : On the Purpose of Models - the Norges Bank Experience, Staff Memo, No. 06/2011, ISBN 978-82-7553-612-7, Norges Bank, Oslo, <https://hdl.handle.net/11250/2507330>

This Version is available at:

<https://hdl.handle.net/10419/210222>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc-nd/4.0/deed.no>

No. 06 | 2011

Staff Memo

On the purpose of models - The Norges Bank experience

Karsten R. Gerdrup and Jon Nicolaisen, Monetary Policy

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

© 2011 Norges Bank

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

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

© 2011 Norges Bank

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

ISSN 1504-2596 (online only)

ISBN 978-82-7553-* (online only)

On the purpose of models - The Norges Bank experience*

Karsten R. Gerdrup[†] Jon Nicolaisen[‡]

July 4, 2011

Abstract

Macroeconomic models are important ingredients in the monetary policy process, and, in the Norwegian case, projecting a forward interest rate path. In this paper we argue that when deciding on a model strategy, it is crucial to consider the purpose of models. If the purpose is to understand basic mechanisms in the economy and implications of economic policy, we need a set of models that highlight these features. If the purpose is to forecast short-term developments, a different set of models may be required. Given the complexity of the real world, we argue that it is better to provide the policymakers with a good characterization of uncertainty instead of only providing point forecasts, i.e. it is better to be "roughly right" than "exactly wrong". A robust strategy for handling uncertainty should be an inherent part of the preferred system of models.

Keywords: Forecasting; Forecast Combination; Monetary policy; Robustness

*We thank Ragna Alstadheim, Kjetil Olsen, Francesco Ravazzolo, and Øistein Røisland for useful comments.

[†] *Corresponding author:* Karsten R. Gerdrup, Norges Bank, Economics Department, Bankplassen 2, 0107 Oslo, Norway. Telephone: +47 22 31 64 40. Fax: +47 22 42 40 62. karsten.gerdrup@norges-bank.no

[‡]Executive Director, Norges Bank Monetary Policy, jon.nicolaisen@norges-bank.no

1 Introduction

In this paper we elaborate on the purpose of models at central banks. We concur with George Box' (1979) statement that *"All models are wrong. But some are useful"*. The overriding evaluation criterion for models (or a combination a models) used at central banks is how useful they prove to be in helping the policymakers conduct monetary policy. We argue that a forecasting model that satisfies a wide range of statistical tests may be of little use if the model does not forecast well out of sample. Similarly, a model constructed for policy purposes is of little use if the model does not, for example, include some representation of forward interest rates.

Forecasting is an important ingredient in the monetary policy process of most central banks. Inflation targeting can indeed be described as targeting forecasts of inflation (Svensson (1997)). Forecasts will always be followed by forecast errors, possibly large ones. We have imperfect knowledge about the state of the economy, and we are uncertain about its structure. Alan Greenspan, the former Chairman of the US Federal Reserve, described this as follows: *"Uncertainty is not just an important feature of the monetary policy landscape; it is the defining characteristic of that landscape."*¹ We argue that it is better to provide the policymakers with a good characterization of that uncertainty, rather than only providing point forecasts. In other words, it is better to "roughly right" than "exactly wrong".²

Norges Bank's criteria for what constitutes a good model depends, broadly speaking, on the time horizon of forecasts and analysis. In the short-run, empirical fit and out-of-sample performance of models are typically the primary concern. In the medium- to long-term perspective, theoretical consistency and the likely interaction between monetary policy and the economy becomes crucial to any model that is designed to analyse

¹In Alan Greenspan's opening remarks to the Jackson Hole symposium in 2003, available at: <http://www.bis.org/review/r030905a.pdf>

²The quote *"It's better to be roughly right than exactly wrong"* is according to Wikiquote attributed to John Maynard Keynes, but the original quote came from Read (1898) where he used 'vague' instead of 'roughly'.

monetary policy. These are widely different purposes. Hence, it may be best to construct different models for each of them.

This paper is organized as follows: In Section 2, we present some thoughts on criteria that we believe models should meet. In Section 3, we elaborate on Norges Bank's experience with models, and based on that, give a brief overview over future work in section 4. Finally, a summary is given in section 5.

2 What is a good model?

2.1 Models versus judgment

The forecasting process at central banks is often iterative. Short-term forecasts may be formed by sector experts' views on economic forces in conjunction with projections by formal models. Detailed knowledge of developments in the different sectors of the economy may be useful for forecasting short-term developments. The sector experts monitor a large amount of data from disparate sources, including information of a more qualitative nature. They also have an understanding of how disaggregated bits of data feed into the preparation of the aggregate numbers that are published with a lag by the statistical agencies. It is thus possible to construct a forecasting system using the detailed knowledge of experts, without extensive use of complex models. Indeed, this has been the strategy of choice for some central banks, at least in terms of centralised modeling efforts.

However, expert judgment is inherently subjective. As soon as you want to provide forecasts of economic developments which could serve as a more objective benchmark or starting point, or if you need to clarify complex interactions between policy and the economy, models are useful to ensure internal consistency. Also, models are prioritising tools than can abstract from unimportant details in policy analysis. Haavelmo's arguments are still valid: *"Theoretical models are necessary tools in our attempts to understand and "explain" events in real life. In fact, even a simple description and classification of real phenomena would probably not be possible or feasible without viewing reality through the*

framework of some scheme conceived a priori", see [Haavelmo \(1944\)](#).

Models can be categorized into stochastic and non-stochastic models. Non-stochastic models involve variables that are deterministic. Non-stochastic models are used in all kinds of qualitative analysis to illustrate the relationship between economic variables (e.g. supply and demand in goods markets), and answer questions like the consequence of a change in an exogenous variable on the set of endogenous variables. Such models may be useful for our understanding and enlightenment, as they tend to be small and pedagogical.

Stochastic models involves random processes described by probability distributions. This implies that even when we know the state of the economy (which we rarely do), the future path of the economy is still uncertain, as it will be hit by stochastic "shocks" in the future. In the medium- to long-term, uncertainty is higher than in the short run because of the accumulation of disturbances.

Econometrics aims to give empirical content to economic models and to the uncertainty surrounding model outcomes. Central banks typically also need to quantify the relationship between variables to foster a better understanding of the structure of the economy. Theory and measurement are interrelated: *"Neither 'theory' nor 'measurement' on their own is sufficient to further our understanding of economic phenomena"* ([Geweke, Horowitz, and Pesaran \(2006\)](#)). Thus, judgment still has a role in assessing the best mix of theory and measurement. Moreover, once a model is operative, the model user may and should combine model contents with detailed knowledge of factors outside the model, in order to reduce the uncertainty of model predictions.

2.2 Criteria for models used in central banks

The overriding evaluation criterion for models (or a combination a models) used at central banks is how useful they prove to be in helping the policymakers conduct monetary policy. A model that satisfies a wide range of statistical tests may be of little use for monetary policy if the model is not able to address the most important issues at hand. Below, we suggest some possible criteria to assess models. The first two criteria are

particularly useful when assessing short-term forecasting models, while criteria 2-6 are particularly useful when assessing models for monetary policy analysis.

2.2.1 Forecasting models should forecast well

As a starting point, it seems an appropriate criterion for a good forecasting model that it forecasts well out of sample, using the data available at each point in time of forecasts, i.e. real-time data. While this may seem rather obvious, it has not always been the case. For example, complex models of the Cowles-school era³, often did not perform well when being tested out of sample. Due to their complexity, uncertainty bands around forecasts were often very wide, making these models impractical for forecasting purposes.

A solid assessment of both the current economic situation and developments in the next few quarters is typically essential to making sound projections for economic developments over a longer period. This assessment is complicated by the fact that many key statistics are released with a long delay, are subsequently revised and are available at different frequencies. Standing in the beginning of quarter typically means that we lack important information also about the previous quarter. As a consequence, there has been a substantial interest in developing a framework for forecasting the present and recent past. In a seminal paper, [Giannone, Reichlin, and Small \(2008\)](#) provide important amendments to the approximate dynamic factor model, as they adapt the model to account for an unbalanced dataset. As shown by [Maih \(2010\)](#) it may be possible to improve the forecast performance of dynamic stochastic general equilibrium models by conditioning on e.g., financial market information, surveys or short-term forecasts from models that are able to exploit recent data and information from large datasets. Problems with publication lags and subsequent revisions have also prompted the establishment of regional networks by central banks to facilitate the use of more timely information from companies, see [Brekke and Halvorsen \(2009\)](#).

Neither theory nor empirical literature can single out specific types of models or estimation methods that can produce significantly more precise out-of-sample forecasts

³See [Christ \(1994\)](#) for a history of the Cowles Commission.

than others. Identification of a structure or finding causal relationships are of second order importance, since also variables that have not been causally relevant in earlier periods may be so in future. Furthermore, non-causal variables may outperform in forecasting relative to causally relevant variables when the model is misspecified or the economy undergoes structural breaks, see [Clements and Hendry \(2002\)](#). Thus, the state of the art seems to be that a variety of models and methods may produce reasonable forecasts, but none of them seem to outperform the others. This may constitute an argument for combining forecasts of many models, as such combinations may give better forecasts than any of the individual models in the combination.

[Bates and Granger \(1969\)](#) early recognized that combining competing point forecasts can provide more accurate forecasts in terms of lower root-mean-squared forecast. [Bjørnland, Gerdrup, Jore, Smith, and Thorsrud \(2011a\)](#) find support that forecast combination using a large suite of models can beat Norges Bank’s own published forecasts for inflation.

2.2.2 Forecasts should be robust

Academic literature has until quite recently been focusing on developing models that increase forecast accuracy in term of point forecasts, but this is typically only a valid performance criterion if the policymaker’s main concern is to minimize mean-squared-forecast errors. We would argue this is not the case. Rather, what is useful for a policy maker is a forecast that with a high probability is in the vicinity of actual outcomes. Moreover, the implied loss function is probably non-linear. Improving a really bad forecast is more important than an incremental improvement of a good one. In other words, rather than looking for good point forecasts, we should be searching for robust ones.

[Elliott and Timmermann \(2008\)](#) argued that forecasting is one case where one size fits all does not hold, because policy makers may have different preferences and therefore require different optimal point forecasts. It may therefore be better to provide density forecasts instead of point forecasts.

Along these lines, several authors have in recent years provided justification for combining density forecasts ([Jore, Mitchell, and Vahey \(2010\)](#) and [Amisano and Geweke \(2009\)](#)). Many studies find that combining densities is a much better strategy than selecting a particular model *ex ante*, i.e. combination provides insurance against selecting inappropriate models ([Kascha and Ravazzolo \(2010\)](#) and [Bjørnland, Gerdrup, Jore, Smith, and Thorsrud \(2011b\)](#)). Bayesian model averaging (BMA) provides an alternative approach to density combination⁴. See also [Timmermann \(2006\)](#) for a discussion of theoretical and empirical motivations for combining forecasts, and [Wallis \(2011\)](#) for a historical overview.

In our view, the key to a robust forecasting system is that it hedges against uncertain instabilities in models. The obvious way to do this is use many models, both different types of models⁵ and models with different estimation periods, transformation of data, lag lengths, etc. These models should in turn produce density forecasts. In addition to the insurance aspect, we can then answer questions like: "What is the probability that inflation will exceed the inflation target the next four quarters?". Models can be given weights based on a measure of past out of sample forecast performance in terms of their densities, or equal weights could be used if the model suite is first trimmed based on past out of sample forecast performance. The Kullback-Leibler information criterion (KLIC) is an important criterion because it attaches highest weight to the model with highest probability of having "created" the data⁶. Density forecasts based on different weighting schemes should be evaluated on an ongoing basis to judge whether they are well-calibrated.⁷ Even when published density forecasts are not based on formal models, but largely based on judgment, they could still be evaluated.

⁴See [Hoeting, Madigan, Raftery, and Volinsky \(1999\)](#) for an introduction to BMA.

⁵For example vector autoregressive models, factor models, and leading indicator models.

⁶[Mitchell and Hall \(2005\)](#) and [Hall and Mitchell \(2007\)](#) suggest using the KLIC as a unified means of evaluating, comparing and combining competing density forecasts.

⁷Tests can be based on Kullback-Leibler information criteria or *pits*. The pits are the ex ante inverse predictive cumulative distribution evaluated at the ex post actual observations, see [Diebold, Gunther, and Tay \(1998\)](#). A density is correctly specified if the pits are uniform and, for one-step ahead forecasts, independently and identically distributed.

2.2.3 Policy must have a role in policy models

A prerequisite for a model to be used for monetary policy analysis is that monetary policy has a defined role. As a starting point, it is monetary policy that ensures that inflation and other variables in the objective function are brought back to target. This must be reflected in the model.

One way to do this is to specify interest rate rules or loss functions that can help policymakers trade-off the different objectives in monetary policy in a consistent and transparent way, and to react consistently to disturbances. The choice of specification of monetary policy should be suited for being a tool for internal discussions on the appropriate path of forward interest rates, as well as reflecting the preferences of the governing body of the central bank. The two most common general approaches to modeling monetary policy in practice are instrument rules on the one hand (for example a Taylor rule), and solving for the interest rate path that minimizes some central bank loss function subject to a model, referred to as optimal policy⁸. That is, over time, inflation should be determined by policy.

Again, this criterion may seem obvious, but has not always been the case. Often, an assumption of constant interest rates going forward, has been used to forecast inflation. However, such assumptions violate what we know about expectations formation processes and is typically inconsistent with observed data, including observable forward market rates. As a tool for policy analysis, it would thus appear that such forecasts make little sense.

2.2.4 Expectations should be taken seriously

The fundamental role of monetary policy in an inflation targeting regime is to provide the economy with a nominal anchor, i.e. an anchor for inflation and inflation expectations. Low and stable inflation will, in turn, provide a basis for stable developments in output and employment. If monetary policy is credible, future inflation will be expected to

⁸See [Holmsen, Qvigstad, Øistein Røisland, and Solberg-Johansen \(2008\)](#) for a description of Norges Bank's use of optimal policy as a normative benchmark when assessing policy intentions

be equal or close to the inflation target. Thus, monetary policy works mainly through private agents' expectations, and the effectiveness of monetary policy depends on the way central banks communicate their future policy intentions.

The role of expectations in monetary policy was highlighted by [Woodford \(2005\)](#) when he argued that: *"For not only do expectations about policy matter, but ... very little else matters"*. [Woodford \(2007\)](#) furthermore described inflation forecast targeting as a combination of a decision procedure and a communication policy.

The reason for modeling expectations also has a deeper cause, that is to strive to comply with the Lucas critique ([Lucas \(1976\)](#)). The critique implies that since economic policy is a rule for systematically changing a policy variable in response to conditions, and since changes in policy in this sense must be expected to change the reduced form of macroeconometric models that treat the policy variable as exogenous, the reduced form is not structural, see also [Sims \(1980\)](#). [Robert E. Lucas and Sargent \(1979\)](#) argued forcefully for models with explicit expectations and "deep" parameters on behaviour and technology when they wrote that: *"No general first principle has ever been set down which would imply that, say, the expected rate of inflation should be modeled as a linear function of lagged rates of inflation alone with weights that add up to unity, yet this hypothesis is used as an identifying restriction in almost all existing models. The casual treatment of expectations is not a peripheral problem in these models, for the role expectations is pervasive in them and exerts a massive influence on their dynamic properties (a point Keynes himself insisted on)"*.

We will therefore stress the perhaps obvious point that a monetary policy model must reflect that agents not only take account of today's economic policy, but also form expectations of future policy, and act accordingly. Economic relationships built into a macro model should, however, be based on careful and ongoing empirical analysis for such analysis to be of any value.⁹

⁹The development in the estimation of monetary policy models was in the early 2000s lagging behind the development in monetary policy theories that gave an explicit role to central banks in stabilising prices, see for example [Woodford \(2003\)](#). Many central banks tried to incorporate forward-looking, optimizing agents into already existing large macro models. [Sims \(2002\)](#) gave two reasons for a reappraisal

2.2.5 Policy models should be comprehensible

It is important to have in mind that models are tools, not sources of definitive answers. In order to use successfully a model for monetary policy analysis, it needs to be comprehensible to policy makers. The key advantage of using a well-formulated macroeconomic model is that it imposes structure and discipline on the forecast and policy analysis processes, by revealing and focusing attention on the relevant but perhaps non-obvious implications of what is known or assumed. Fully identified models allow us to have an informed opinion about the effects of policy and estimation of shocks that are driving the economy. To utilize these advantages in actual policy, policy makers must comprehend the implications of the model analysis. This requires models to be relatively simple and centered on main policy issues.

In the past, central banks typically would embark on projects in the tradition of the Cowles Commission, resulting in large and complex models which tried to encompass a set of multiple purposes that varied in scope and detail. Such models were often complex beyond comprehension. Model systems designed for and tailored to a more limited purpose, are, in our opinion, much more efficient.

2.2.6 Policy models should take account of uncertainty

A strategy for handling uncertainty should be an inherent part of the modeling system. Inflation-targeting central banks must constantly deal with pervasive uncertainty regarding both the current situation and the workings of the economy and monetary policy. Yet they must make assumptions and set monetary policy such that inflation

of the probability approach: *"One is scientific...: Economic models will make forecast errors; unless they are probability models, we have no objective way to assess their accuracy, compare them to other models, or improve them based on experience. The other is decision-theoretic: ... policy-makers need to weigh the implications of the model against their own knowledge and judgment; this will necessarily involve weighting model results by their reliability. A probability model provides its own characterization of its reliability."* Many central banks developed Bayesian estimated New-Keynesian DSGE models along the lines of e.g. [Smets and Wouters \(2003\)](#). [Sims \(2007\)](#) could therefore later state that: *"There is a trend,... toward bringing probability modeling and policy modeling back together"*.

is expected to be on target within an appropriate time horizon. Therefore, it is very useful to set out assumptions explicitly in the context of economic models, such that the implications of alternative assumptions, i.e. risks, can be explored and discussed in a systematic way.

There is now a large literature on monetary policy under uncertainty. Generally, the policy implications depend on what the uncertainty relates to, for example, whether there is parameter uncertainty or model uncertainty, and whether the uncertainty is quantifiable or not. If uncertainty is quantifiable, Bayesian model averaging, as suggested by [Brock, Durlauf, and West \(2007\)](#), is a natural approach. However, even if optimal policy in a Bayesian model averaging framework in principle could deal with model uncertainty, it is a very computationally demanding approach, and existing work focus on simple, as opposed to fully optimal, interest rate rules, see e.g., [Cogley, de Paoli, Matthes, Nikolov, and Yates \(2011\)](#) and the references therein. Deriving optimal forecasts based on Bayesian model averaging is, at least at the current stage, a daunting task that would entail that central bank staff must produce multiple model forecasts added with judgments in hectic forecasting rounds. There is thus a practical argument for producing forecasts within one main model, while using other models as cross-checks and inputs to the judgmental adjustments of the forecasts of the core model.

If uncertainty is not quantifiable, i.e., there is Knightian uncertainty, a minimax approach is a common way to deal with such uncertainty. Under minimax, one aims to minimize the loss in a worst case situation, see [Hansen and Sargent \(2008\)](#) for robust control theory. For a central bank with a core model, robust control could be a useful tool for discussing alternative interest rate paths reflecting different preferences on robustness. An advantage is that a robust control exercise is carried out within the core model itself and does thus not require other models. This advantage has, however, also its costs. As argued by [Levin and Williams \(2003\)](#), a robustly optimal policy in one model may give poor results in another model, and may be better suited for dealing with local model uncertainty, i.e., uncertainty within the constrained class of model.

A common approach to deal with uncertainty across models is to use simple interest

rate rules that are specified and calibrated to give reasonably good results in a variety of models. The rationale for simple rules is phrased by [Taylor and Williams \(2010\)](#): *“...optimal policies can be overly fine-tuned to the particular assumptions of the model. If those assumptions prove to be correct, all is well. But, if the assumptions turn out to be false, the costs can be high. In contrast, simple monetary policy rules are designed to take account of only the most basic principle of monetary policy of leaning against the wind of inflation and output movements. Because they are not fine-tuned to specific assumptions, they are more robust to mistaken assumptions.”*

We will argue that a practical solution to the issue of uncertainty is to look for policy strategies that are robust - i.e. gives reasonably good outcomes - over a wide range of assumptions and models. As we show in the next section, we believe it is possible to combine an optimal policy reaction pattern with explicit weight on robust simple rules, so as to provide a hedge against model uncertainty.

3 Norges Bank’s modeling system

Norges Bank’s modeling system is largely built up around the different requirements set up depending on the time horizon of forecasts and analysis, like shown in figure 1. In the short-run, empirical fit and out-of-sample performance of models are typically the primary concern. In the medium- to long-term perspective, theoretical consistency and the likely interaction between monetary policy and the economy becomes crucial to any model that is designed to analyse monetary policy.

3.1 Short-term forecasts

Norges Bank’s system for model-based short-term forecasts serves as an objective benchmark, or starting-point, for short-term forecasts, see figure 2 for an overview over the system¹⁰. In the modeling set-up, we follow the criteria for a good forecasting model

¹⁰See [Bache, Brubakk, and Maih \(2010\)](#) for a thorough description of the forecasting and analysis system at Norges Bank

Figure 1. Short-term, medium-term or long-run - different methods and models

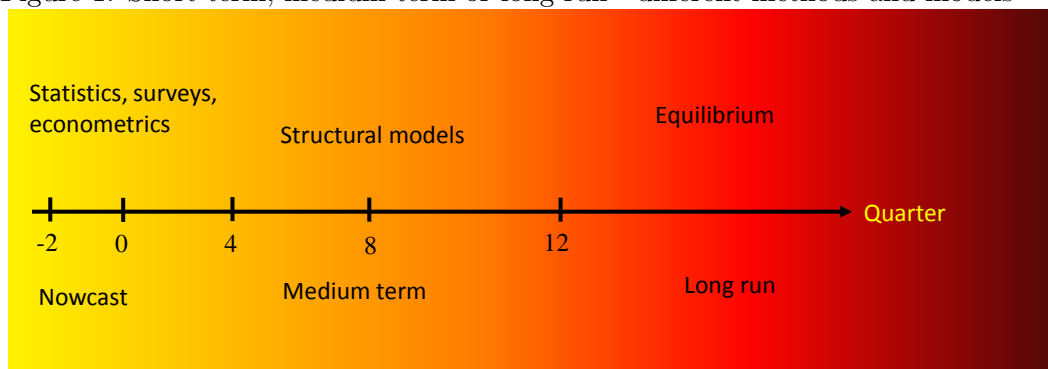
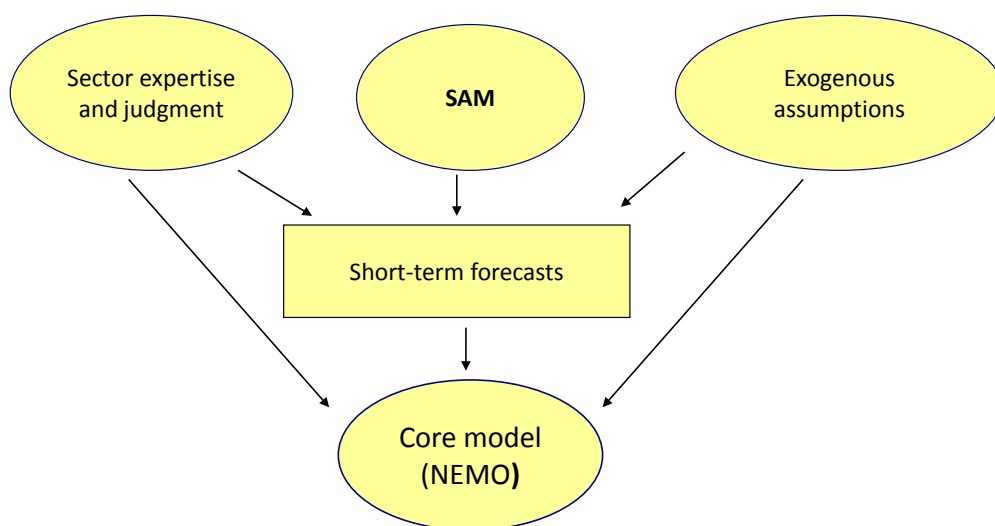


Figure 2. The forecasting and monetary policy analysis system at Norges Bank



and robust forecasts that we outlined in section 2.

The system is called SAM - System for Averaging Models - and it produces currently short-term density forecasts (up to five quarters) for Mainland GDP and CPI-ATE (consumer prices adjusted for taxes and without energy).¹¹ Models that historically have proven successful in empirical literature and in central banks, are included in the suite. A total of 221 models are used to forecast GDP and 171 models are used to forecast CPI-ATE, see table 1.

Vector autoregressive models (VARs) without restrictions stand for a major part of the models, see ‘VAR’-ensemble in table 1. Following Sims (1980), VARs have been utilized extensively in macroeconomic forecasting since they allow us to estimate simultaneous macromodels as unrestricted reduced forms, treating all variables as endogenous¹². Although VARs are useful in forecasting, recent work by Clark and McCracken (2010) suggest that such models may be prone to instabilities. Thus, we use different estimation windows, transformation of data and lag lengths in the model suite in SAM. Bayesian techniques have become a useful supplement to classical econometrics. We have also included some small Bayesian VARs in the suite, but our aim is to introduce also Bayesian VARs with many variables.

There is a large amount of studies showing that leading indicators can be useful for economic forecasting, see among others Banerjee, Marcellino, and Masten (2005), Banerjee and Marcellino (2006) and Marcellino (2006), see ‘Indicator’-ensemble in table 1. We have included many simple indicator models in SAM formulated as a bivariate VAR between the variable in interest (inflation or Mainland GDP) and the indicator. The challenge with ragged edge data when forecasting is solved by utilizing the Kalman-

¹¹See Aastveit, Gerdrup, and Jore (2011) for a more thorough description of the system.

¹²Sims was highly skeptical towards the large-scale macroeconomic models of the time. While admitting that estimation techniques used for VARs could be improved, he stated that *“the opportunity it offers to drop the discouraging baggage of standard, but incredible, assumptions macroeconometricians have been used to carrying may make the road attractive.”*

Table 1. Models and ensembles in SAM

Ensembles		Number of models for CPIATE	Number of models for GDP
Factor	Factor models, monthly and quarterly	5	5
VAR	(V)AR models, bayesian and classical incl. a DSGE and VEqCM	161	149
Indicator	Indicator models, monthly and quarterly	5	67
Sum		171	221

filter. Indicators include mainly financial information, surveys¹³, industrial production, order statistics, labor market developments, money and credit. The indicators can have different frequencies, and this is solved by using some sort of bridge equation to transform for example monthly data to quarterly indicators.

Dynamic factor models is another important model type implemented in SAM, see ‘Factor’-ensemble in table 1. The objective of factor models is to summarize the information contained in large datasets, while at the same time reducing their dimension. In other words, to reduce the parameter space. These types of models have been increasingly popular at central banks as they tend to have good forecasting properties. The factor model are based on either monthly or quarterly information, and for each of these, models with different number of factors are included.

Finally, a New-Keynesian DSGE model and a macroeconometric vector equilibrium model (VEqCM), are included (in ‘VAR’-ensemble in table 1). Even though these are not constructed for forecasting well in the short-term, they may have good forecasting properties a couple of quarters ahead.

Each model in SAM provide density forecasts, see upper panel in figure 3. We give each model a weight based on past forecasting performance. This is done in two steps.

¹³Examples of other surveys are Statistics Norway’s business tendency survey and TNS Gallup’s Consumer Confidence Index.

In the first step we group models that loosely share the same information set and/or model structure (see also [Gerdrup, Jore, Smith, and Thorsrud \(2009\)](#)). The reason for this is that we do not want the combined density forecast to be dominated by a certain model type just because the number of this model type is large due to uncertainties regarding lag lengths, transformation of data etc. Furthermore, by grouping models we become more certain that models with more timely information - like the factor models with monthly information - have a chance of receiving a reasonably high weight. In the current version of SAM we have divided the models into three groups. In the next step we combine the predictive densities from the groups into a combined density forecast. The ensemble densities and the combined density are shown the lower panel in figure 3.

The individual models (in the first step) and the ensembles (in the second step) are combined using a linear opinion pool¹⁴:

$$p(y_{\tau,h}) = \sum_{i=1}^N w_{i,\tau,h} g(y_{\tau,h}|I_{i,\tau}), \quad \tau = \underline{\tau}, \dots, \bar{\tau} \quad (1)$$

where $I_{i,\tau}$ is the information set used by model i to produce the density forecast $g(y_{\tau,h}|I_{i,\tau})$ for variable y at forecasting horizon h . $\underline{\tau}$ and $\bar{\tau}$ are the period over which the individual forecasters's densities are evaluated, and finally $w_{i,\tau,h}$ are a set of non-negative weights that sum to unity. The evaluation period is from 2001 to present.

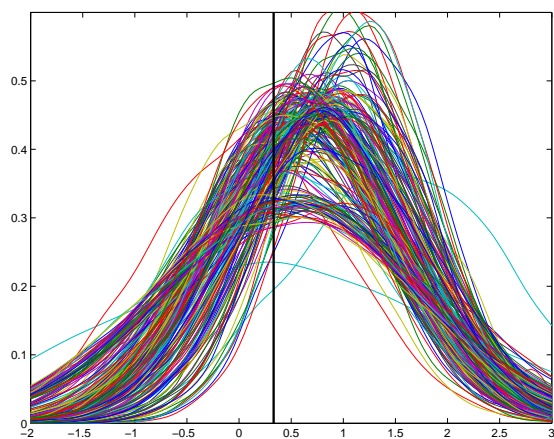
Combining the N density forecasts according to equation 1 can potentially produce a combined density forecasts with characteristics quite different from those of the individual forecasters. Since the combined density is a linear combination of all the individual forecasters' densities, the variance of the combined density forecast will in general be higher than that of individual models. However, this is not necessarily deleterious, as the combined density may perform better than the individual density forecasts when evaluated.

Many different weighting schemes have been proposed in the literature to define $w_{i,\tau,h}$. In SAM we apply logarithmic score (log score) weights¹⁵ for GDP and inverse mean

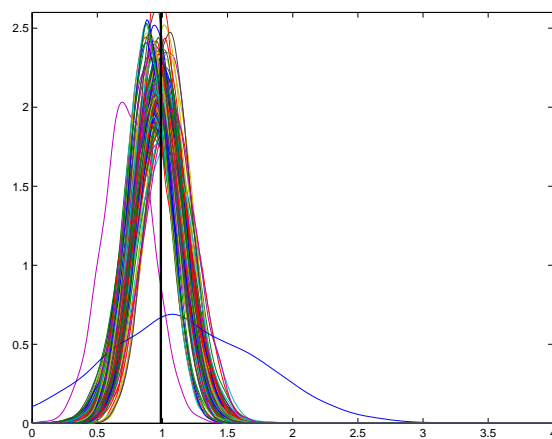
¹⁴The logarithmic pool is an alternative method, see [Kascha and Ravazzolo \(2010\)](#) for an analysis of the properties of the two different pooling methods.

¹⁵See [Jore, Mitchell, and Vahey \(2010\)](#) for a description.

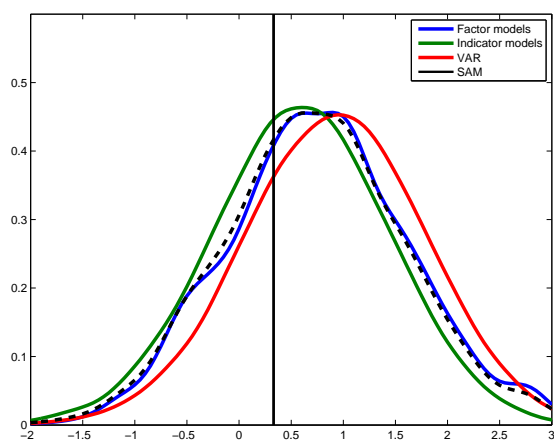
Figure 3. Density forecasts for individual models, ensembles, and SAM for 2010Q4 compared with outturn. Mainland GDP (qoq-growth) and CPI-ATE (yoy-growth). Per cent



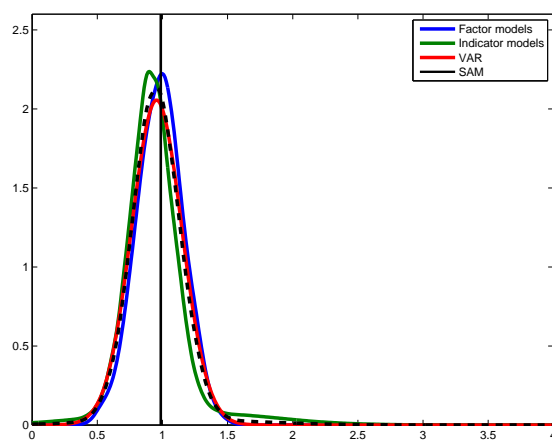
(a) Individual models, GDP



(b) Individual models, CPI-ATE



(c) Ensembles and SAM, GDP



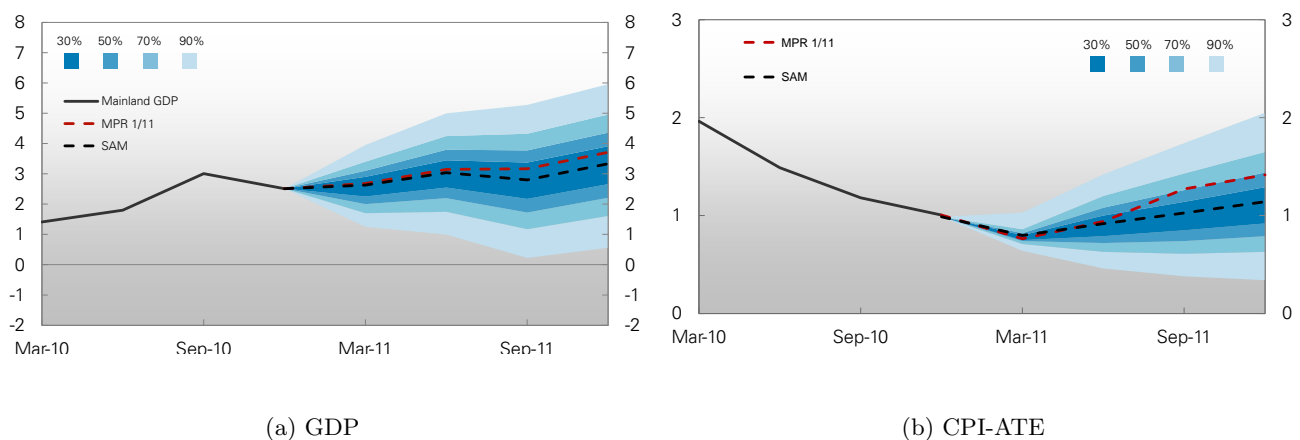
(d) Ensembles and SAM, CPI-ATE

Note: The forecasts were made with data available in mid-October, which was the cut-off date in MPR 3/10. GDP for 2010Q3 was not first available in late November.

Source: Norges Bank

squared errors (MSE) weights for CPI-ATE because these weighting schemes perform well (out-of-sample) in terms of density forecasting (maximize log score)¹⁶. The weights are derived based on out-of-sample performance and horizon specific. Furthermore, the weights are recursively updated. Inverse MSE weights are close to equal weights when we have many models.

Figure 4. Density forecasts for Mainland GDP and CPI-ATE from SAM and point forecasts from MPR 1/11. Four-quarter growth. Per cent



Source: Norges Bank

SAM provides model-based forecasts, but the final short-term forecasts are in general subject to judgment. The final short-term forecasts are used as starting values and conditional assumptions in NEMO. SAM forecasts are updated regularly and are published on the Bank’s website in conjunction with each monetary policy meeting in Norges Bank’s Executive Board. Figure 4 depict the fan charts for CPIATE and GDP from SAM published in MPR 1/11 (March 2011). Inflation and GDP were judged to in-

¹⁶Note that maximizing the log score is the same as minimizing the Kullback-Leibler distance between the models and the true but unknown density. [Mitchell and Wallis \(2010\)](#) show that the difference in log scores between an “ideal” density and a forecast density, that is the Kullback-Leibler information criterion, can be interpreted as a mean error in a similar manner to the use of the mean error or bias in point forecast evaluation.

crease somewhat more than the mean of the SAM densities. The current version of SAM only produces forecasts for two variables, but our aim is to produce density forecasts for most of the observable endogenous variables in the Bank's core macroeconomic model, NEMO (Norwegian Economy Model).

3.2 Monetary policy analysis and interest rate forecasts

The forecasting system is organized around NEMO, which is a medium-scale, small, open economy DSGE model similar in size and structure to the New-Keynesian DSGE models developed in recent years by many other central banks, see [Alstadheim, Bache, Holmsen, Maih, and Øistein Røisland \(2010\)](#) for a more detailed description. NEMO has been estimated using Bayesian techniques on quarterly data for mainland Norwegian economy, see [Bache, Brubakk, and Maih \(2010\)](#).

When producing policy analysis, the forecast produced by the short-term forecasting system is used as input into NEMO (see figure 2). Thus, the policy analysis in NEMO is conditioned on the set of short-term forecasts produced by other models and judgment. This also implies that NEMO is used to give our short-term forecast an interpretation consistent with the view of the economy inherent in the model. This interpretation takes the form of initial shocks to the model. The role of monetary policy is then to bring target variables back on track, given the initial conditions at the starting point.

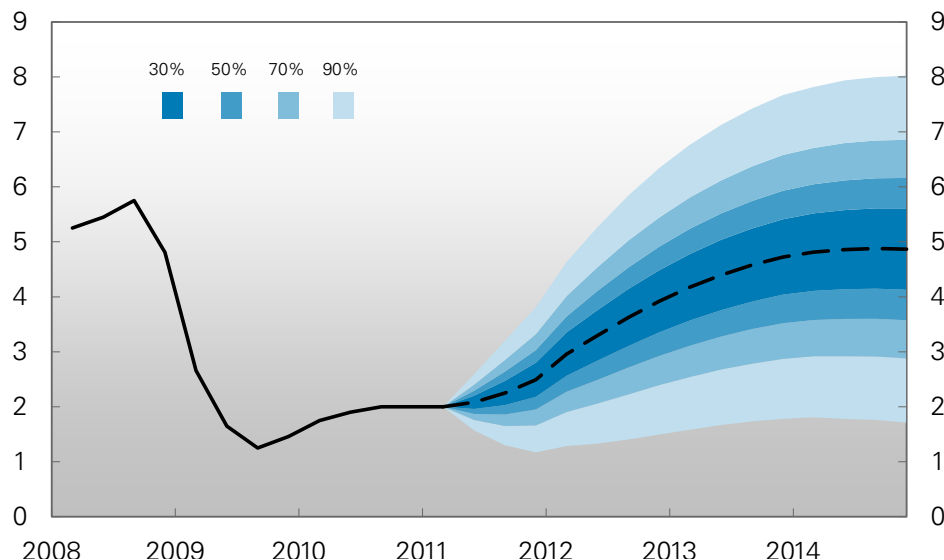
Norges Bank has since 2005 published its own interest rate path, see 5. The uncertainty in the forecasts of the main policy variables¹⁷ are shown as fan charts. The fan charts are in turn based in part on the model's density forecast, and in part on judgment. For example, the output gap density from NEMO is typically deemed too wide to be well-calibrated.

Forecasts have in the last couple of years been based on optimal policy¹⁸ under

¹⁷The main policy variables are: key policy rate, output gap, consumer price inflation, and consumer prices adjusted for tax changes and excluding temporary changes in energy prices

¹⁸See [Holmsen, Qvigstad, and Øistein Røisland \(2007\)](#) for an overview of the Bank's communication with focus on the implementation and communication of the optimal monetary policy

Figure 5. Projected key policy rate in the baseline scenario with fan chart. Per cent. 2008Q1-2014Q4



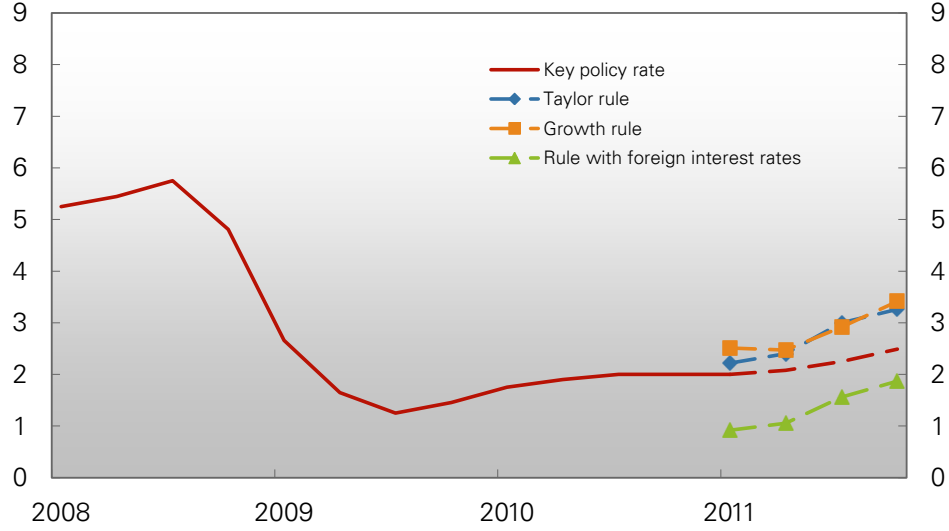
Source: Norges Bank (MPR 1/11)

varying degrees of commitment given the structure of the economy¹⁹. However, we believe there are major first-order uncertainties connected to the choice of models that need to be handled. More specifically, we have taken account of model uncertainty by extending the operational loss function in NEMO. The way we do this is to weigh the primary trade-off in monetary policy against a weight on a simple interest rate rule.

The basic argument for this procedure is that if there is uncertainty about the "true" model, then we need to make sure that the derived optimal policy path in our model works reasonably well also under alternative assumptions about the underlying transmission mechanism and structure of the economy. As earlier mentioned, simple rules can be specified and tailored to give reasonably good results in a variety of models (see e.g. [Taylor and Williams \(2010\)](#)). By putting some weight on the distance to such a

¹⁹The empirical fit of the model with optimal policy is found to be as good as the model with a simple rule, see [Bache, Brubakk, and Maih \(2010\)](#). This result is robust to allowing for misspecification following the DSGE-VAR approach proposed by [Negro and Schorfheide \(2004\)](#). The unconditional interest rate forecasts from the DSGE-VARs are close to Norges Bank's official forecasts since 2005.

Figure 6. Key policy rate and calculations based on simple monetary policy rules. Per cent. 2008Q1-2011Q4



Note: The calculations are based on Norges Bank's projections for the output gap, consumer prices adjusted for tax changes and excluding temporary changes in energy prices (CPIXE) and 3-month money market rates. To ensure comparability with the key policy rate, the simple rules are adjusted for risk premiums in 3-month money market rates.

Source: Norges Bank (MPR 1/11)

simple rule, we thus take out some "insurance" against model misspecification. That is, by weighing in such rules, we formulate a more robust policy path.

The minimization problem is formulated as follows:

$$E \sum_{t=0}^{\infty} \beta^t [(\pi_t - \pi^*)^2 + \lambda y_t^2 + \gamma (i_t - i_{t-1})^2 + \eta (i_t - i_t^T)], \quad (2)$$

where $i_t^T = r_t^* + \pi^* + 1.5(\pi_t - \pi^*) + 0.5y_t$.

The weight β is the discount factor, π is inflation, i is the interest rate, y is the output gap. η determines how much the central bank aims to guard against bad results due to model uncertainty, while λ and γ determines the weight on output and interest rate smoothing, respectively. By specifying η , the use of simple rules as cross-checks and guidelines can be modeled in a precise way.

In the representation above, the two first terms (inflation and output gap) may be

interpreted as the traditional trade-off involved in policy formulation, given the model. The two last terms (interest smoothing and the distance to a simple rule) modify this optimal path to take account of uncertainty and produce a more robust policy forecast.

We consider also alternative formulations of robust rules, see 6 for alternative simple rules in addition to the key policy rate in the baseline scenario. First, a version where the output gap is replaced by the output *growth* gap, i.e., $i_t = r_t^* + \pi^* + 1.5(\pi_t - \pi^*) + 0.5(y_t - y_{t-1})$. This version is meant to be more robust against uncertainties in estimating potential output and is inspired by work by [Orphanides and van Norden \(2002\)](#) and [Rudebusch \(2002\)](#). Second, we use a version of the Taylor rule where the foreign real interest rate is added to the rule, i.e., $i_t = r_t^* + \pi^* + 1.5(\pi_t - \pi^*) + 0.5y_t + 0.5r_t^f$. This rule takes into account that changes in the interest rate among Norway’s trading partners may result in changes in the exchange rate and thereby influence the inflation outlook. The choice of coefficient values for these rules are based on the original Taylor rule and have not been subject to any optimization.

4 Future work

Taking model uncertainty seriously in monetary policy analysis is currently a high priority area at Norges Bank. First, research is directed towards bringing financial stability into monetary policy analysis by integrating financial variables into the macro models that we use for analysis and forecasting. Second, research is directed towards extending the Bank’s analysis on robust policy. These two areas are related because the first area may bring about more relevant models that can be used in different analysis of robust policy.

Research on robust monetary policy will be concentrated, first, on estimating robust simple rules that can be used within the optimal policy framework used to close NEMO. The simple rules will be evaluated in a variety of different models, including different versions of NEMO (for example with and without financial frictions), Vector Equilibrium models, and other models. Second, Bayesian or minimax approaches to analyse

robust monetary policy will be explored. While implementing better robust simple rules imply small changes in how we currently work when analysing policy, implementation of Bayesian or minimax approaches may in practice be more challenging. First, producing multiple model forecasts with optimal policy added with judgment is a non-trivial task, both conceptually and computationally. Second, and perhaps more importantly, robust optimal policy in one model may give poor results in another model.

We will continue to pursue a strategy which combines optimal and robust policy in terms of a specified loss function. One aspect of this, is the possibility of taking direct account of concerns about financial stability in setting interest rates, i.e. "lean against the wind". However, this is still largely uncharted territory.

5 Summary

In this paper we have argued that when deciding on a model strategy, it is crucial to consider the purpose of the models. The overriding evaluation criterion for models used at central banks is how useful they prove to be in helping the policymakers conduct monetary policy.

Norges Bank's criteria for what constitutes a good model depends, broadly speaking, on the time horizon of forecasts and analysis. In the short-run, empirical fit and out of sample performance of models are typically the primary concern. In the medium- to long-term perspective, theoretical consistency and the likely interaction between monetary policy and the economy becomes crucial to any model that is designed to analyse monetary policy.

As a starting point, it seems an appropriate criterion for a good forecasting model that it forecasts well out-of-sample, using real-time data. This criterion is particularly useful when assessing short-term forecasting models. Second, we believe that forecasts should be robust. In our view, the key to a robust forecasting system is that it hedges against uncertain instabilities in models. Furthermore, we believe that it is important to give policy makers a full characterisation of the uncertainty in the forecasts, i.e. it is

better to be "roughly right" than "exactly wrong". Third, policy must be defined in policy models. This means that monetary policy should have a defined role (formulated as simple rules or optimal policy) in bringing the policy variables back to target. Fourth, expectations should be taken seriously. Fifth, policy models should be comprehensible to policy makers in order to be used successfully. Finally, a strategy for handling uncertainty should be an inherent part of the modeling system.

While Norges Bank is trying to comply with these criteria, there is much more left to be done on robust policy formulation. This could be done, perhaps, by estimating robust simple rules and combining these with policy optimisation.

References

- AASTVEIT, K. A., K. GERDRUP, AND A. S. JORE (2011): “Short-term forecasting of GDP and Inflation in Real-Time: Norges Bank’s system for averaging models,” Staff Memo forthcoming, Norges Bank.
- ALSTADHEIM, R., I. W. BACHE, A. HOLMSEN, J. MAIH, AND ØISTEIN RØISLAND (2010): “Monetary policy analysis in practice,” Staff Memo 2010/11, Norges Bank.
- AMISANO, G., AND J. GEWEKE (2009): “Optimal Prediction Pools,” Working Paper Series 1017, European Central Bank.
- BACHE, I. W., L. BRUBAKK, AND J. MAIH (2010): “Simple rules versus optimal policy: what fits?,” Working Paper 2010/03, Norges Bank.
- BANERJEE, A., AND M. MARCELLINO (2006): “Are there any reliable leading indicators for US inflation and GDP growth?,” *International Journal of Forecasting*, 22(1), 137–151.
- BANERJEE, A., M. MARCELLINO, AND I. MASTEN (2005): “Leading Indicators for Euro-area Inflation and GDP Growth,” *Oxford Bulletin of Economics and Statistics*, 67(s1), 785–813.
- BATES, J., AND C. GRANGER (1969): “The combination of forecasts,” *Operations Research Quarterly*, 20(4), 451–468.
- BJØRNLAND, H. C., K. GERDRUP, A. S. JORE, C. SMITH, AND L. A. THORSRUD (2011a): “Does forecast combination improve Norges Bank inflation forecasts?,” *Oxford Bulletin of Economics and Statistics*, 73(4).
- (2011b): “Weights and pools for a Norwegian density combination,” *The North American Journal of Economics and Finance*, 22(1), 61–76.
- BREKKE, H., AND K. W. HALVORSEN (2009): “Norges Bank’s regional network: fresh and useful information,” Economic Bulletin 2, Norges Bank.

- BROCK, W. A., S. N. DURLAUF, AND K. D. WEST (2007): “Model uncertainty and policy evaluation: Some theory and empirics,” *Journal of Econometrics*, 136(2), 629–664.
- CHRIST, C. F. (1994): “The Cowles Commission’s Contributions to Econometrics at Chicago, 1939-1955,” *American Economic Association*, 32(1), 30–59.
- CLARK, T. E., AND M. W. MCCracken (2010): “Averaging forecasts from VARs with uncertain instabilities,” *Journal of Applied Econometrics*, 25(1), 5–29.
- CLEMENTS, M. P., AND D. F. HENDRY (2002): “An Overview of Economic Forecasting,” in *A Companion to Economic Forecasting*, ed. by M. Clements, and D. Hendry, pp. 1–18. Blackwell Publishing, Oxford.
- COGLEY, T., B. DE PAOLI, C. MATTHES, K. NIKOLOV, AND T. YATES (2011): “A Bayesian approach to optimal monetary policy with parameter and model uncertainty,” Working Paper 414, Bank of England.
- DIEBOLD, F. X., T. A. GUNTHER, AND A. S. TAY (1998): “Evaluating Density Forecasts with Applications to Financial Risk Management,” *International Economic Review*, 39(4), 863–83.
- ELLIOTT, G., AND A. TIMMERMAN (2008): “Economic Forecasting,” *Journal of Economic Literature*, 46(1), 3–56.
- GERDRUP, K. R., A. S. JORE, C. SMITH, AND L. A. THORSRUD (2009): “Evaluating ensemble density combination - forecasting GDP and inflation,” Working Paper 2009/19, Norges Bank.
- GEWEKE, J., J. HOROWITZ, AND M. PESARAN (2006): “Econometrics: A Bird’s Eye View,” Cambridge Working Papers in Economics 0655, Faculty of Economics, University of Cambridge.

- GIANNONE, D., L. REICHLIN, AND D. SMALL (2008): “Nowcasting: The real-time informational content of macroeconomic data,” *Journal of Monetary Economics*, 55(4), 665–676.
- HAAVELMO, T. (1944): “The Probability Approach in Econometrics,” *Econometrica*, 12(Supplement), 1–115.
- HALL, S. G., AND J. MITCHELL (2007): “Combining density forecasts,” *International Journal of Forecasting*, 23(1), 1–13.
- HANSEN, L. P., AND T. J. SARGENT (2008): *Robustness*. Princeton University Press.
- HOETING, J. A., D. MADIGAN, A. RAFTERY, AND C. VOLINSKY (1999): “Bayesian model averaging: A tutorial,” *Statistical Science*, (14), 382–417.
- HOLMSEN, A., J. F. QVIGSTAD, AND ØISTEIN RØISLAND (2007): “Implementing and communicating optimal monetary policy,” Staff Memo 2007/3, Norges Bank.
- HOLMSEN, A., J. F. QVIGSTAD, ØISTEIN RØISLAND, AND K. SOLBERG-JOHANSEN (2008): “Communicating monetary policy intentions: The case of Norges Bank,” Working Paper 2008/20, Norges Bank.
- JORE, A. S., J. MITCHELL, AND S. P. VAHEY (2010): “Combining forecast densities from VARs with uncertain instabilities,” *Journal of Applied Econometrics*, 25(4), 621–634.
- KASCHA, C., AND F. RAVAZZOLO (2010): “Combining inflation density forecasts,” *Journal of Forecasting*, 29(1-2), 231–250.
- LEVIN, A. T., AND J. C. WILLIAMS (2003): “Robust Monetary Policy with Competing Reference Models,” *Journal of Monetary Economics*, 50, 945–975.
- LUCAS, R. J. (1976): “Econometric policy evaluation: A critique,” *Carnegie-Rochester Conference Series on Public Policy*, 1(1), 19–46.

- MAIH, J. (2010): “Conditional forecasts in DSGE models,” Working Paper 2010/07, Norges Bank.
- MARCELLINO, M. (2006): “Leading Indicators,” in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. W. J. Granger, and A. Timmermann, vol. 1, pp. 879–960. Elsevier, Amsterdam.
- MITCHELL, J., AND S. G. HALL (2005): “Evaluating, Comparing and Combining Density Forecasts Using the KLIC with an Application to the Bank of England and NIESR ‘Fan’ Charts of Inflation,” *Oxford Bulletin of Economics and Statistics*, 67(s1), 995–1033.
- MITCHELL, J., AND K. WALLIS (2010): “Evaluating Density Forecasts: Forecast Combinations, Model Mixtures, Calibration and Sharpness,” *Journal of Applied Econometrics*, forthcoming.
- NEGRO, M. D., AND F. SCHORFHEIDE (2004): “A DSGE-VAR for the Euro Area,” Computing in Economics and Finance 2004 79, Society for Computational Economics.
- ORPHANIDES, A., AND S. VAN NORDEN (2002): “The Unreliability of Output-Gap Estimates in Real Time,” *The Review of Economics and Statistics*, 84(4), 569–583.
- READ, C. (1898): *Logic, Deductive and Inductive*. Grant Richards, London.
- ROBERT E. LUCAS, J., AND T. J. SARGENT (1979): “After Keynesian macroeconomics,” *Quarterly Review*, (Spr).
- RUDEBUSCH, G. D. (2002): “Assessing Nominal Income Rules for Monetary Policy with Model and Data Uncertainty,” *Economic Journal*, 112(479), 402–432.
- SIMS, C. A. (1980): “Macroeconomics and Reality,” *Econometrica*, 48(1), 1–48.
- SIMS, C. A. (2002): “The Role of Models and Probabilities in the Monetary Policy Process,” *Brookings Papers on Economic Activity*, 33(2002-2), 1–62.

- (2007): “Monetary Policy Models,” *Brookings Papers on Economic Activity*, 38(2007-2), 75–90.
- SMETS, F., AND R. WOUTERS (2003): “An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area,” *Journal of the European Economic Association*, 1(5), 1123–1175.
- SVENSSON, L. E. O. (1997): “Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets,” NBER Working Papers 5797, National Bureau of Economic Research, Inc.
- TAYLOR, J. B., AND J. C. WILLIAMS (2010): “Simple and Robust Rules for Monetary Policy,” in *Handbook of Monetary Economics*, ed. by B. M. Friedman, and M. Woodford, vol. 3, chap. 15, pp. 829–859. Elsevier, 1 edn.
- TIMMERMANN, A. (2006): “Forecast Combinations,” in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. W. J. Granger, and A. Timmermann, vol. 1, pp. 136–96. Elsevier, Amsterdam.
- WALLIS, K. (2011): “Combining forecasts - forty years later,” *Applied Financial Economics*, 21(1-2), 33–41.
- WOODFORD, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press, Princeton.
- (2005): “Central Bank Communication and Policy Effectiveness,” NBER Working Papers 11898, National Bureau of Economic Research, Inc.
- (2007): “Forecast Targeting as a Monetary Policy Strategy: Policy Rules in Practice,” NBER Working Papers 13716, National Bureau of Economic Research, Inc.