Borradores de ECONOMÍA

Monetary Policy Forecasting in a DSGE Model with Data that is Uncertain,Unbalanced and About the Future.

Por: Andrés González Gómez,Lavan Mahadeva,Diego Rodríguez G.,Luis Eduardo Rojas

Núm. 559

2009

tá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Colombia - Bogotá - Col

BANCO DE LA REPÚBLICA
COLOMBIA

MONETARY POLICY FORECASTING IN A DSGE MODEL WITH DATA THAT IS UNCERTAIN, UNBALANCED AND ABOUT THE FUTURE*.

ANDRÉS GONZÁLEZ GÓMEZ †, LAVAN MAHADEVA ‡, DIEGO RODRÍGUEZ G. §, AND LUIS EDUARDO ROJAS ♣

Abstract. If theory-consistent models can ever hope to forecast well and to be useful for policy, they have to relate to data which though rich in information is uncertain, unbalanced and sometimes forecasts from external sources about the future path of other variables. One example from many is financial market data, which can help but only after smoothing out irrelevant short-term volatility. In this paper we propose combining different types of useful but awkward data set with a linearised forward-looking DSGE model through a Kalman Filter fixed-interval smoother to improve the utility of these models as policy tools. We apply this scheme to a model for Colombia.

Keywords: Monetary Policy, DSGE, Forecast, Kalman Filter

JEL Code: F47, E01, C61.

† Head of the Department of Modelling, Banco de la República de Colombia, Carrera 7 No 14-78, Bogotá D.C. Corresponding author: Email agonzago@banrep.gov.co.

‡ Adviser to the Governor, Banco de la República de Colombia.

§ Deputy Head of the Department of Modelling, Banco de la República de Colombia.

♣ Economist, Department of Modelling, Banco de la República de Colombia.

*This work represents the opinions of the authors alone and not the Board members of the Banco de la República.

Contents

1. Introduction 3
2. The key assumptions, the sequencing of solutions and notation 6
2.1. Crucial assumptions 6
2.2. The sequence of events and the information sets 6
2.3. Notation 7
3. The solved model 7
4. Extending the economic structure to allow for future data 9
5. Allowing for awkward features of the policy forecasting data set 10
5.1. Introducing the data formally 10
5.2. The information set 11
5.3. The solution to the data uncertainty problem 11
6. Philosophies for finding parameter values for the data measurement equation 13
6.1. Tender loving care 13
6.2. Purpose-built data set 14
6.3. Extending the model to fit in data 14
6.4. Data intensive methods 15
6.5. Some general principles to choose the data measurement uncertainty 15
7. Useful outputs from the forecasting system 16
7.1. Data decompositions 17
7.2. Extracting information on the economic content of the model 19
8. Allowing for reporting variables 22
9. The state identification problem 23
10. Some applications 24
10.1. A complete forecast 24
10.2. Experiment to show the importance of imperfectly measured awkward data 26
10.3. Experiments to show the importance of off-model forecasts 26
10.4. Experiment to reveal the identification problem 30
11. Conclusions 32
References 33

1. Introduction

Forecasting is a very different exercise to simulating, especially when those forecasts are used to guide and explain policy. A policy forecast should relate to all the data that features in the public debate, even if that comes in an awkward variety of shapes and forms.

The need to reach out and connect with the relevant data matters to any policy forecasting strategy, but it is especially important when the forecast is based on a dynamic stochastic general equilibrium (henceforth, DSGE) model. DSGE models are distinguished by having a greater theoretical input. Indeed this explains their appeal to policy; with more theory it is easier to use the forecast as a basis for discussion and explanation. Yet more theory often comes at a cost in terms of worse forecast performance because it is harder to match more rigid theoretical concepts to available data. But if any policy model ultimately does not forecast well, it cannot be of much use as a tool for explanations however elegant is the theory.

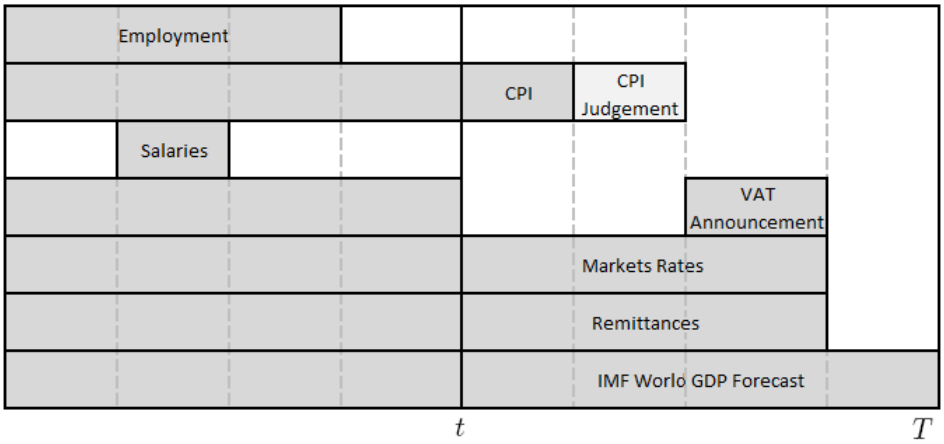
To return to the point we raised in the first paragraph, there are two main types of reasons why we should expect the useful data set for a policy forecast to be awkward:

(1) Real world data is unbalanced.
   (a) Real world data comes with different release lags: we have more up to date information on some series than on other series.
   (b) Real world data comes with different frequencies. Some relevant economic information is only available annually whilst some is in real time.
   (c) There may be useful off-model information on the expected current and future values of exogenous and endogenous variables from other sources which cannot be incorporated into the model system, at least not without making the model cumbersome. Sometimes this information is patchy: we have information on values at some future point. But sometimes the information may be complete for the whole horizon as would be the case if we use the forecasts from other sources for the exogenous variables.
(2) Real world data is subject to time-varying measurement uncertainty. This also has many aspects.
   (a) First real world data is imperfectly measured. For this reason published data may be subject to revision.
   (b) Second more generally, real world data that is available to agents may differ from the economic concept that matters to their decisions.
   (c) Third forecasts from other models too comes with a measurement error.

Figure 1.1 shows us some examples of the kind of awkward data that one can expect to deal with in the real world. Assume we are time t and planning to forecast up until time T.

- First we have series such as employment which comes to us after a longer delay.
- In contrast series such as CPI are very up to date. And then often it is the case that a combination of small monthly models, monthly data on some prices and even information from the institutions which set regulated prices, can give a decent forecast of consumer prices into the next quarter.
- Some data is only available annually. For example in Colombia the only national salary series can be constructed from the income side of annual national accounts.

FIGURE 1.1. A typical real world data set for forecasting



- Then there is information on what can happen quite far ahead. The government can preannounce its VAT plans which will have a first-round effect on inflation.
- Also we have implied forecasts from financial markets data. For example, removing risk premia and other irrelevant effects from a yield curve gives us a data on what the risk-free expectation of future interest rates. Some other important special cases of this type of information would be world interest rates, world commodity price and expected monetary policy rates. The forecast then should be conditioned on the useful information contained in this data.
- Finally there are forecasts from other models. A good example is remittances. Using information on migration trends, exchange rate movements, and some disaggregated capital flows data, a specialist can come up with a good forecast for remittances or at least a forecast that is better than one that a DGSE model can generate internally. Population growth and relative food prices are two other examples of variables that might also be best forecasted separately. Similarly forecasts for the GDP of important trading partners should probably come from forecasters in those countries or from international institutions such as the IMF which have a better capacity to forecast those series.

The contention of this paper is that a model can only provide both decent predictions and useful explanations if it takes account of the real world data set. Obviously, if there is valuable information is this awkward data set then the better forecasts should not ignore it. But more subtly, if the agents whose behaviour we are trying to model might be using this type of information when they form expectations, we would also need to mimic them if we are to expect to pick up their behaviour. For example, if the data is noisy and is infected by short-term movements which have nothing to do with the fundamentals, agents will ignore some of the movements in the data. So then should the forecast.

Through the method described in this paper, a forecast can bring this rich but awkward information to bear in forming a policy forecast with a forward-looking DSGE model. The basic idea is to first solve the model for rational expectations under the assumption that the data up until the end of the forecast horizon is perfectly known. Then in a second stage, the data uncertainty problem is wrapped around

these solutions which are the state equations of a Kalman filter. The innovation is to solve in the first stage assuming that the whole data set up until the end of the forecast could in principle be available but then put in holes where we want the model to forecast. In Kalman filter terminology, the solution is fixed-interval smoothing.

This method has advantages over alternative strategies.

First it allows for measurement error of future data. If that future data are forecasts produced by other models, then this model's forecast should take account of the external models' errors. If that future data were instead announcements or plans made by external institutions, this method allows for the important possibility of imperfect credibility by incorporating that as measurement error. For example, a preannounced tax change may not be fully credible.

Another important special case is financial market data. Financial market data is noisy in the sense that the price that data measures diverges from the economic concept that matters to agents. If agents have a longer holding horizon than financial market participants, they would not react to short-term, reversible, movements in financial data. Not to adjust for this would mean that forecasts bounce around with financial data unrealistically. An influential paper by Lettau and Ludvigson (2004) emphasises that movements in short-term financial market data should be smoothed before we assess their impact on macroeconomic variables.

Policy forecasters commonly smooth financial market data by imposing prior views on the initial values and forecasts of financial variables; these are the controversial constant exchange rate and interest rate assumptions. Economic shocks are then made endogenous to make the model's solutions comply with those priors. Our proposal is an alternative way of getting to the same result, and one which does not involve us rewriting the economic model to make structural shocks endogenous, for now those deviations are classified as data noise terms.

A second advantage of this method is that the model does not need to be rewritten each time the shape of the data of each given series changes. One only has to fill the parts of the data set where there is data and put blanks (NAs!) in where there is nothing available. With this method, one can imagine how the task of building the data set and that of maintaining the forecasting model can become separate and so carried out at different intervals and by different groups of specialists.

Third, we can be much more imaginative with what data we use with this method. In this paper, we show how the measurement equations can be adapted to push the forecast towards where there is interesting information without having to extend the economic part of the model. In this way we can exploit the interesting information in money data without making that part of the endogenous core of the model, for example. We can allow for the peculiarities of the national accounts data, and especially what we know about its revisions policy. We can bring in information from surveys that we know are important in monetary policy decisions.

Another advantage of this approach is that it allows us access to the whole toolkit that comes with the Kalman filter. We show how we can derive decompositions of forecasts according to the contributions of data, and not just according to the contributions of economic shocks as is standard. In fact we can show how estimates of the contributions of the shocks depend on data. Going further, we also show

how we can try and spot where forecasts will not be well identified in terms of shocks. The information context of different series to the forecast can be measured. And then retrospectively, we can use this technology to compare the forecasts across different vintages of data to see if policy mistakes were due to data mismeasurement, as suggested by Orphanides (2001) and Boragan Aruoba (2004). This means that we can present the differences between forecasts across policy rounds in terms of the contribution of news in the data. This incremental way of presenting the forecast to a busy Monetary Policy Committee is at least more efficient. All these outputs help us to integrate the model's forecast into the central bank's policy decision.

In many respects this work is not original but rather a synthesis of a very broad literature. In particular we have borrowed from findings on how to simulate linear rational expectations models for optimal monetary policy under data uncertainty. The key references are Pearlman (1986), Svensson and Woodford (2003), Svensson and Woodford (2004), Swanson (2004) and the toolkit of Gerali and Lippi (2003). In terms of central banks, the pioneers for bringing data uncertainty into the policy forecast problem are the Czech National Bank (Coats, Laxton, and Rose (2003)).

The rest of this paper is organized as follows. Section 2 summarises our key assumptions and clarifies our notation. Section 3 presents the solved economic model in the absence of data uncertainty. In Section 4, that is extended to allow for future data. Section 5 introduced the data into our forecast. Section 6 discusses the different strategies for modelling data uncertainty. Section 7 presents several useful different ways of presenting the policy forecast. Section 8 allows for reporting variables. Section 9 discusses the problem of state identification. Section 10 presents some forecasts using this strategy on Colombian data. Section 11 concludes.

2. The key assumptions, the sequencing of solutions and notation

2.1. **Crucial assumptions.** The input of this paper is the outcome of a micro-founded general equilibrium problem. We assume that in a prior stage, the relevant decisions of agents have been formulated as optimizing problems, that the first and second-order conditions have been derived and that those conditions have then been aggregated to match the data and transformed into a linear dynamic system with all variables only in terms of log deviations from time-invariant steady-state values. Uhlig (1995),Sims (2002) and Klein (2000) explain how this is done.

This paper begins at the next stage where we need to solve to model and match it to available data. By solving the model we mean that we need to model how the monetary policymakers choose their policy instrument and how agents form rational expectations. By forecasting we mean how we want to use this model to fit and predict available data. We want all these decisions to reflect what data is really available. Even then as we shall now see, we make use of a separation theorem to focus on the data part of the solution, assuming that the model has been solved and is given to us in a standard form.

2.2. **The sequence of events and the information sets.** The choices of policymakers and agents happen at a current time $s = t$ which lies in between 1 and the end of the forecast horizon, $T = t + k$. Unlike conventional expositions, it is assumed that potential data set that was available to agents and

policymakers at any time in the past up until now, that is at time u ($u = 1, \ldots, t$), could have include possibly useful off-model information on variables timed from 1 to $u+k$, and also could include information about data uncertainty. Later on we specify exactly what are in the data sets, but here we just mention that they could contain values of future variables which we interpret as forecasts from other models. As we move from the past to the current date, information is never forgotten. The information sets are written as $\mathcal{F}_0, ..., \mathcal{F}_t$ with the property that $\mathcal{F}_i \subset \mathcal{F}_j$ for any $i < j$. The macroeconomic forecaster makes his optimal forecasts based on the same data set as economic agents.

In this set up, the information set is common, the dynamic model linear, the objective function quadratic and expectations rational. Then the literature on optimal policy under data uncertainty tells us that there is a *separation property* such that the problem can be split into two artificial stages and solved recursively. See the references in the introduction. The first stage, where the rational expectations of agents are solved for, is certainty equivalent; at this point, the second-order properties of the data measurement and economic shocks do not matter. Our scheme allows for future data. For that reason, agents and the forecaster must be allowed to see shocks k periods in advance in this artificial first stage as explained in Schmitt-Grohe and Uribe (2008). In a second stage, this partial solution is combined with a description of the data and the second-order properties of the data measurement and economic shocks to give the final solution.

2.3. **Notation.** The operator E_i means that we are taking expectations of a variable with respect to an information set $\mathcal{F}_i$ of timing i. Matrices and vectors are in bold, with matrices in upper case. $\mathbf{I}_M$ refers to the $(M \times M)$ identity matrix. $\mathbf{0}_{MN}$ is an $(M \times N)$ matrix with zero elements. $Tr\boldsymbol{\Upsilon}$ refers to the trace of matrix $\boldsymbol{\Upsilon}$. $\boldsymbol{\Upsilon}^H$ refers to the transpose of a matrix $\boldsymbol{\Upsilon}$. $\boldsymbol{\Upsilon}^+$ refers to the Moore Penrose inverse of the matrix $\boldsymbol{\Upsilon}$. $\boldsymbol{\Upsilon}(i,j)$ is the ij^{th} element of the matrix $\boldsymbol{\Upsilon}$.

3. The solved model

The starting point of this paper is the outcome of a micro founded general equilibrium problem, that has already been expressed in terms of log-linearised deviations from a steady state:

$$\mathbf{A}E_s\mathbf{x}_{s+1}^0 = \mathbf{B}\mathbf{x}_s^0 + \mathbf{C}\mathbf{z}_{s+1}^0. \tag{3.1}$$

There are N variables in the vector $\mathbf{x}_s^0$. $\mathbf{x}_s^0$ includes all economic variables endogenous for the model. The N_ε variables in $\mathbf{z}_s^0$ are a set of exogenous variables that follow univariate first-order process. We call them economic shocks. Then there is a matrix equation

$$\mathbf{z}_{s+1}^0 = \mathbf{B}_\varepsilon^0\mathbf{z}_s^0 + \mathbf{D}_\varepsilon^0\boldsymbol{\varepsilon}_{s+1} \tag{3.2}$$

where $\mathbf{B}_\varepsilon^0$ and $\mathbf{D}_\varepsilon^0$ are both diagonal matrices, although $\mathbf{B}_\varepsilon^0$ may have a zero entries on its diagonal. Being exogenous, these variables have no expectational error. Any autoregressive component of the shock is captured in the process 3.2 such that $\boldsymbol{\varepsilon}_{s+1}$ is a martingale difference process with respect to the information set at time s. In particular,

$$E_s\boldsymbol{\varepsilon}_{s+1} = \mathbf{0}_{N_\varepsilon,1} \tag{3.3}$$

and, without loss of generality,

$$E_s \boldsymbol{\varepsilon}_{s+1} \boldsymbol{\varepsilon}_{s+1}^H = \mathbf{I}_{N_\varepsilon}. \tag{3.4}$$

The trivial steady-state solution, defined by $\mathbf{x}_s^0 = \mathbf{0}_{N,1}$ and $\boldsymbol{\varepsilon}_s = \mathbf{0}_{N_\varepsilon,1}$ for all time s, exists and is unique. The model begins at that steady state:

$$\mathbf{x}_0^0 = \mathbf{0}_{N,1}. \tag{3.5}$$

Then, departing from the solution Klein (2000) shows, and including in the vector $\mathbf{x}_s^0$ the variables in $\mathbf{z}_s^0$ the solved model can be presented in the form

$$\mathbf{x}_{s+1}^0 = \boldsymbol{\Xi}_1^0 \mathbf{x}_s^0 + \boldsymbol{\Xi}_2^0 \boldsymbol{\varepsilon}_{s+1}. \tag{3.6}$$

The solution in the form of equation 3.6 can be written in partitioned form, with the exogenous variables ($\mathbf{z}_s^0$) recovered in the lower partition:

$$\mathbf{x}_{s+1}^0 = \begin{bmatrix} \boldsymbol{\Xi}_{111}^0 & \boldsymbol{\Xi}_{112}^0 \\ \mathbf{0}_{N_\varepsilon, N-N_\varepsilon} & \mathbf{B}_\varepsilon^0 \end{bmatrix} \mathbf{x}_s^0 + \begin{bmatrix} \boldsymbol{\Xi}_{211}^0 \\ \mathbf{D}_\varepsilon^0 \end{bmatrix} \boldsymbol{\varepsilon}_{s+1}. \tag{3.7}$$

where the matrices $\mathbf{B}_\varepsilon^0$ and $\mathbf{D}_\varepsilon^0$ are defined in equation 3.2. In fact Klein (2000) shows that the upper partition of the equation 3.7 can be partitioned further, were the upper positions of the vector are assigned to the non-predetermined variables, and the rest to the predetermined endogenous variables. A variable is predetermined if its generating process is *backward-looking*, here following the definition of a *backward-looking* process presented in Klein (2000). From now on it is assumed that the vector $\mathbf{x}_s^0$ is arranged in that way and N_p is defined as the number of endogenous predetermined variables.

For us there is another more important property of the solution. It will always be the case that the matrix $\boldsymbol{\Xi}_{111}^0$ is independent of the parameters of the exogenous shock processes (of matrices $\mathbf{B}_\varepsilon^0$ and $\mathbf{D}_\varepsilon^0$). This can be seen in the solutions presented by Klein (2000) pages 1417-1418, equations (5.20) and (5.21).

The $(N - N_\varepsilon) \times N_\varepsilon$ matrices $\boldsymbol{\Xi}_{112}^0$ and $\boldsymbol{\Xi}_{211}^0$ does depend on $\mathbf{B}_\varepsilon^0$ and $\mathbf{D}_\varepsilon^0$ though. $\boldsymbol{\Xi}_{112}^0$ and $\boldsymbol{\Xi}_{211}^0$ are given by

$$\boldsymbol{\Xi}_{112}^0 = \begin{bmatrix} \mathbf{R} \\ \mathbf{L} \end{bmatrix} \qquad \boldsymbol{\Xi}_{211}^0 = \begin{bmatrix} \mathbf{N}\mathbf{D}_\varepsilon^0 \\ \mathbf{0}_{N-N_p-N_\varepsilon, N_\varepsilon} \end{bmatrix}$$

where

$$\begin{aligned} \mathbf{R} &\equiv \mathbf{Z}_{21} \left[\mathbf{Z}_{11}\right]^{-1} \mathbf{L} + \mathbf{N}\mathbf{B}_\varepsilon^0 \\ \mathbf{N} &\equiv \left(\mathbf{Z}_{22} - \mathbf{Z}_{21} \left[\mathbf{Z}_{11}\right]^{-1} \mathbf{Z}_{12}\right) \mathbf{M}, \end{aligned} \tag{3.8}$$

$$\begin{aligned} \mathbf{L} \equiv\ & -\mathbf{Z}_{11} \left[\mathbf{S}_{11}\right]^{-1} \mathbf{T}_{22} \left[\mathbf{Z}_{11}\right]^{-1} \mathbf{Z}_{12}\mathbf{M} \\ & \mathbf{Z}_{11} \left[\mathbf{S}_{11}\right]^{-1} \left[\mathbf{T}_{12}\mathbf{M} - \mathbf{S}_{12}\mathbf{M}\mathbf{B}_\varepsilon^0 + \mathbf{Q}_1\mathbf{C}\right] \\ & \mathbf{Z}_{12}\mathbf{M}\mathbf{B}_\varepsilon^0, \end{aligned} \tag{3.9}$$

and

$$vec\mathbf{M} \equiv \left(\left[\left[\mathbf{B}_{\varepsilon}^{0}\right]^{H} \otimes \mathbf{S}_{22} - \mathbf{I}_{N_{\varepsilon}} \otimes \mathbf{T}_{22} \right]^{-1} \right) vec\left[\mathbf{Q}_2\mathbf{C}\right]. \tag{3.10}$$

N_p is the number of endogenous predetermined variables. Crucially the matrices $\mathbf{S}_{ij},\ \mathbf{T}_{ij},\ \mathbf{Z}_{ij}, \mathbf{Q}_i$ for $i, j = 1, 2$ are only functions of $\mathbf{A}$ and $\mathbf{B}$ and so are independent of $\mathbf{B}_{\varepsilon}^{0}$ and $\mathbf{D}_{\varepsilon}^{0}$. They are determined by a generalized Schur factorization which would be computationally expensive in very large sparse models.

4. Extending the economic structure to allow for future data

To allow for the possibility of future data to affect current behaviour, agents must be able to anticipate shocks up until the forecast horizon. In this section, the model and solution is extended to allow for that possibility along the lines of Schmitt-Grohe and Uribe (2008).

It was assumed that at each time u when a forecast is made, shocks in principle can be anticipated k periods ahead ($k = T - t$). $\boldsymbol{\varrho}_{i,s+1}$ is defined as perfect information on the vector of shocks $\boldsymbol{\varepsilon}_{s+1}$ but known earlier at time $s - i$ for $i = 0, ..., k - 1$. Then the extended vector of exogenous shocks can be written as

$$\mathbf{z}_{s+1} \equiv \left[\mathbf{z}^{\mathbf{0}H}_{s+1}, \boldsymbol{\varrho}^{H}_{0,s+1}, \ldots, \boldsymbol{\varrho}^{H}_{k-1,s+1} \right]^{H}$$

and the extended vector of economic variables becomes

$$\mathbf{x}_{s+1} \equiv \left[\mathbf{x}^{\mathbf{0}H}_{s+1}, \boldsymbol{\varrho}^{H}_{0,s+1}, \ldots, \boldsymbol{\varrho}^{H}_{k-1,s+1} \right]^{H}.$$

The extended system of exogenous variables is now

$$\mathbf{z}_{s+1} = \mathbf{B}_{\varepsilon}\mathbf{z}_{s} + \mathbf{D}_{\varepsilon}\boldsymbol{\varepsilon}_{s+1} \tag{4.1}$$

with

$$\mathbf{B}_{\varepsilon} \equiv \begin{bmatrix} \mathbf{B}_{\varepsilon}^{0} & \mathbf{I}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \cdots & \mathbf{0}_{N_{\varepsilon}} \\ \mathbf{0}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \mathbf{I}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \cdots & \mathbf{0}_{N_{\varepsilon}} \\ \mathbf{0}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \mathbf{I}_{N_{\varepsilon}} & \cdots & \mathbf{0}_{N_{\varepsilon}} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \cdots & \mathbf{I}_{N_{\varepsilon}} \\ \mathbf{0}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \mathbf{0}_{N_{\varepsilon}} & \ldots & \mathbf{0}_{N_{\varepsilon}} \end{bmatrix}$$

and

$$\mathbf{D}_{\varepsilon} \equiv \begin{bmatrix} \mathbf{0}_{N_{\varepsilon},1} \\ \mathbf{0}_{N_{\varepsilon},1} \\ \mathbf{0}_{N_{\varepsilon},1} \\ \vdots \\ \mathbf{0}_{N_{\varepsilon},1} \\ \mathbf{D}_{\varepsilon}^{0} \end{bmatrix}.$$

It immediately follows that the solution to the extended system is of the form

$$\mathbf{x}_{s+1} = \boldsymbol{\Xi}_1 \mathbf{x}_s + \boldsymbol{\Xi}_2 \boldsymbol{\varepsilon}_{s+1} \tag{4.2}$$

with

$$\boldsymbol{\Xi}_1 \equiv \begin{bmatrix} \boldsymbol{\Xi}^0_{111} & \boldsymbol{\Xi}_{112} \\ \mathbf{0}_{N_\varepsilon(k+1),N-N_\varepsilon} & \mathbf{B}_\varepsilon \end{bmatrix},$$

$$\boldsymbol{\Xi}_2 \equiv \begin{bmatrix} \boldsymbol{\Xi}_{211} \\ \mathbf{D}_\varepsilon \end{bmatrix}$$

Comparing the two solutions, the only matrices that needs to be recalculated are $\boldsymbol{\Xi}_{112}$ and $\boldsymbol{\Xi}_{211}$. But those matrices will follow from the same formulae as in the original model, 3.8, 3.9 and 3.10 but when $\mathbf{B}^0_\varepsilon$ is replaced with $\mathbf{B}_\varepsilon$, $\mathbf{D}^0_\varepsilon$ is replaced with $\mathbf{D}_\varepsilon$ and $\mathbf{C}$ is replaced with with $\begin{bmatrix} C & \mathbf{0}_{N-N_\varepsilon,N_\varepsilon k} \end{bmatrix}$.

In this way, extending the model to incorporate future information does not require carrying out a Schur factorization on a system with many more endogenous variables.

A much greater difficulty would only arise if the vectorized formula 3.10 were actually used in the larger extended model. Luckily, as Klein (2000) explains, there is an alternative recursive method for calculating $\mathbf{M}$, that can be applied to the extended system to keep the computational cost manageable. In summary, the extended model solution can piggy back from most of the original solution.

5. Allowing for awkward features of the policy forecasting data set

Up until now, the solution ignores the possibility of uncertain and unbalanced data. In this Section, data is introduced. The idea is essentially to wrap the system 4.2 around an observation system which relates the true values to the observed data values.

5.1. **Introducing the data formally.** Let the vector $\mathbf{y}^u_s$ of size $(N^u_{Ds} \times 1)$ be the observed data that pertains to variables at time s available in information set u . There can be holes in this data: some observations may not be available at the time s in information set of time u. Thus $N^u_{Ds} \leq N_{D\max}$ where $N_{D\max}$ is the maximum number of data series possibly available.

This measurement system is written as

$$\begin{aligned} \underset{N^u_{Ds}\times 1}{\mathbf{y}^u_s} &= \underset{N^u_{Ds}\times N_{D\max}}{\mathbf{R}^u_s}\ \underset{N_{D\max}\times N}{\mathbf{H}}\ \underset{N\times 1}{\mathbf{x}_t} \\ &+ \underset{N^u_{Ds}\times N_{D\max}}{\mathbf{R}^u_s}\ \underset{N_{D\max}\times N_\delta}{\mathbf{H}_\delta}\ \underset{N_\delta\times 1}{\boldsymbol{\delta}^u_s} + \underset{N^u_{Ds}\times N_{D\max}}{\mathbf{R}^u_s}\ \underset{N_{D\max}\times 1}{\mathbf{v}^u_s} \end{aligned} \tag{5.1}$$

with

$$\mathbf{v}^u_s \sim N\left(\mathbf{0}, \underset{N_{D\max}\times N_{D\max}}{\boldsymbol{\Sigma}^u_s}\right). \tag{5.2}$$

δ^u_s is vector of deterministic variables such as dummies, constants or trends which can affect data measurement. (Economic adjustments to the model are already incorporated in the states). $\mathbf{v}^u_s$ is the normally

distributed stochastic component of the data error which have variance-covariance matrices $\mathbf{\Sigma}_s^u$ that vary both with the time of the data and also with the information set.

Note that the data errors are assumed to be independent of the economic shocks. This assumption would seem to be unrealistic: in many cases, one would expect data measurement problems to be related to the economic cycle or the entry or exit of members of the sample during booms and recessions. It can easily be relaxed in the Kalman filter algorithms. However the greater generality makes it harder to identify shocks. And so the possibly strong restriction that the economic and data uncertainty shocks are independent is imposed in order to avoid getting embroiled in identification problems at this point. Problems of identification are discussed later on.

$\mathbf{R}_s^u$ is a *selector matrix* which alters the number of rows to suit the number of series with time s data observations within the time u information set. $s = 1, \ldots, T$ and $u = 1, \ldots, t$. It is assumed that there is data available in principle from time 0 to time T , where T is the end of the forecast, but this is quite general as $\mathbf{R}_s^u$ can fill in the holes.

Then the selector matrices $\mathbf{R}_s^u$ are formed by taking the $N_{D\max} \times N_{D\max}$ identity matrix and deleting rows corresponding to when there are no data series available at time s within information set u.

Naturally then

$$\mathbf{y}_s^u = \mathbf{R}_s^u \mathbf{y}_s^{u*} \tag{5.3}$$

with $\mathbf{y}_s^{u*}$ being a mythical data set where the maximum number of data series is always available (although with measurement error).

That the variance-covariance matrices of the data error $\mathbf{\Sigma}_s^u$ are both time and information set heteroscedastic means that available data can be weighted according to how reliable that data is thought to be, but also in real time; that is relative to when decisions are made and expectations formed.

5.2. **The information set.** The information sets available to agents, policymakers and the forecasters alike at time u is given by

$$\mathcal{F}_u = \left\{ \mathbf{\Xi}_1, \mathbf{\Xi}_2, \mathbf{H}, [\mathbf{R}_s^u]_{s=1}^{u+k}, [\mathbf{\Sigma}_s^u]_{s=1}^{u+k}, [\boldsymbol{\delta}_s^u]_{s=1}^{u+k}, [\mathbf{y}_s^u]_{s=1}^{u+k}, \mathcal{F}_{u-1} \right\} \tag{5.4}$$

for $u = 1, \ldots, t$, $t + k = T$ and $\mathcal{F}_0 = \{\mathbf{x}_0, \mathbf{P}_0, \mathbf{\Sigma}_0\}$. This structure satisfies the descriptions given in section 2.2. As shall be seen shortly, the information set up allows for a data set with the particular characteristics mentioned in the introduction.

5.3. **The solution to the data uncertainty problem.** The idea is to derive the expectations and forecasts of the economic variables in the model as the fixed-interval smoothed state estimates of the economic system 3.6 conditional on the measurement system 5.1 and 5.2 and the information structure 5.4.

Although the exposition of the Kalman filter is standard, see Harvey (1991); Durbin and Koopman (2001) for example, it is worth repeating here so that we can clarify our particular notation. Given assumptions that the residuals are Gaussian, the Kalman filter produces the minimum mean squared linear estimator of the state vector $\mathbf{x}_{t+1}$ using the set of observations for time t in information set $\mathcal{F}_u$. Let us call that estimate $\widehat{\mathbf{x}}_{t+1|t}^u$ and its associated covariance $\widehat{\mathbf{P}}_{t+1|t}^u$. The series of these estimates are a

building block towards what we are really interested in, the fixed interval smoothed estimates. We carry out this calculation for each information set, although typically we will only be interested in the forecasts from the last (the current) time u = t information set.

Begin with the initial values contained in information set $\mathcal{F}_1$. Then the recursion runs from $s = 1$ to $s = T - 1$ over

$$\widehat{\mathbf{x}}^u_{s+1|s} = \mathbf{\Xi}_1 \widehat{\mathbf{x}}^u_{s|s-1} + \mathbf{K}^u_s w_{\mathbf{s}} \tag{5.5}$$

where

$$w^{\mathbf{u}}_{\mathbf{s}} = \mathbf{y}^{\mathbf{u}}_{\mathbf{s}} - \mathbf{R}^{\mathbf{u}}_{\mathbf{s}} \mathbf{H} \widehat{\mathbf{x}}^u_{s|s-1} - \mathbf{R}^{\mathbf{u}}_{\mathbf{s}} \mathbf{H}_\delta \delta^u_s$$

and the $N \times N^u_{Ds}$ gain matrices are

$$\mathbf{K}^u_s = \mathbf{\Xi_1} \mathbf{P}^{\mathbf{u}}_{\mathbf{s|s-1}} \left[\mathbf{R}^{\mathbf{u}}_{\mathbf{s}} \mathbf{H}\right]^{\mathbf{H}} \left[\mathbf{L}^{\mathbf{u}}_{\mathbf{s}}\right]^{-\mathbf{1}}, \tag{5.6}$$

the $N \times N$ covariance matrices of one-step estimation error are given by the Riccati equation,

$$\hat{\mathbf{P}}^u_{s+1|s} = \mathbf{\Xi}_1 \hat{\mathbf{P}}^u_{s|s-1} [\mathbf{L}^{\mathbf{u}}_{\mathbf{s}}]^{\mathbf{H}} + \mathbf{\Xi}_2 \left[\mathbf{\Xi}_2\right]^H,$$

and the $N^u_{Ds} \times N^u_{Ds}$ covariance matrices of the one-step-ahead prediction errors in the observation data $\mathbf{L}^u_s$ are defined as

$$\mathbf{L}^u_s \equiv \mathbf{R}^u_s \mathbf{H} \hat{\mathbf{P}}^u_{s|s-1} \left[\mathbf{R}^u_s \mathbf{H}\right]^H + \mathbf{R}^u_s \mathbf{\Sigma}^u_s \left[\mathbf{R}^u_s\right]^H.$$

The initial values are given as

$$\widehat{\mathbf{x}}^u_{1|0} = \mathbf{\Xi}_1 \mathbf{x}_0$$

and

$$\hat{\mathbf{P}}^u_{1|0} = \mathbf{\Xi}_1 \mathbf{P}_0 \left[\mathbf{\Xi}_1\right]^H + \mathbf{\Xi}_2 \left[\mathbf{\Xi}_2\right]^H.$$

We also need an updated estimate at least at time T :

$$\widehat{\mathbf{x}}^u_{T|T} = \widehat{\mathbf{x}}^u_{T|T-1} + \mathbf{P}^u_{T|T-1} \left[\mathbf{R}^u_T \mathbf{H}\right]^H \left[\mathbf{F}^u_T\right]^{-1} w^u_s$$

with a series of variance-covariance matrices:

$$\hat{\mathbf{P}}^u_{s|s} = \hat{\mathbf{P}}^u_{s|s-1} - \hat{\mathbf{P}}^u_{s|s-1} \left[\mathbf{R}^u_s \mathbf{H}\right]^H \left[\mathbf{F}^u_s\right]^{-1} \mathbf{R}^u_s \mathbf{H} [\hat{\mathbf{P}}^u_{s|s-1}]^{\mathbf{H}}.$$

where

$$\mathbf{F}^u_s = \mathbf{R}^u_s \mathbf{H} [\hat{\mathbf{P}}^{\mathbf{u}}_{\mathbf{s|s-1}}] \left[\mathbf{R}^{\mathbf{u}}_{\mathbf{s}} \mathbf{H}\right]^{\mathbf{H}} + \mathbf{R}^{\mathbf{u}}_{\mathbf{s}} \mathbf{\Sigma}^{\mathbf{u}}_{\mathbf{s}} [\mathbf{R}^{\mathbf{u}}_{\mathbf{s}}]^{\mathbf{H}}$$

The fixed-interval smoothed estimates are instead given by working backwards from $\widehat{\mathbf{x}}^u_{T|T}$ with the following recursion

$$\widehat{\mathbf{x}}^u_{s|T} = \widehat{\mathbf{x}}^u_{s|s-1} + \hat{\mathbf{P}}^{\mathbf{u}}_{\mathbf{s|s-1}} \mathbf{r}^{\mathbf{u}}_{\mathbf{s-1}}, \tag{5.7}$$

$$\mathbf{r}^u_{s-1} = \mathbf{R}^u_s \mathbf{H} [\mathbf{F}^{\mathbf{u}}_{\mathbf{s}}]^{-\mathbf{1}} \mathbf{w}^{\mathbf{u}}_{\mathbf{s}} + [\mathbf{L}^{\mathbf{u}}_{\mathbf{s}}]^{\mathbf{H}} \mathbf{r}^{\mathbf{u}}_{\mathbf{s}}$$

with $r_T^u = 0$ for $s = T-1,\ T-2, ..., 1$ and the associated variance-covariance matrices of the smoothed prediction error given by

$$\hat{\mathbf{P}}^u_{s|T} = \hat{\mathbf{P}}^u_{s|s-1} + \hat{\mathbf{P}}^u_{s|s-1}\mathbf{N}^u_{s-1}\hat{\mathbf{P}}^u_{s|s-1}, \tag{5.8}$$

$$\mathbf{N^u_{s-1}} = \mathbf{R}^u_s\mathbf{H^H F^{u-1}_s R^u_s H} + [\mathbf{L^u_s}]^{\mathbf{H}}\mathrm{N}^{\mathbf{u}}_{\mathbf{s}}\mathbf{L^u_s}$$

The one-step ahead predictions of the data are given by

$$\widehat{\mathbf{y}}^u_{s+1|s} = \mathbf{R}^u_s\widehat{\mathbf{y}}^{u*}_{s+1|s} = \mathbf{R}^u_s\mathbf{H}\widehat{\mathbf{x}}^u_{s+1|s} + \mathbf{R}^u_s\mathbf{H}_\delta\boldsymbol{\delta}^u_s \text{ for } s = 1, \ldots, T-1 \tag{5.9}$$

and the smoothed predictions follow a process

$$\widehat{\mathbf{y}}^u_{s|T} = \mathbf{R}^u_s\widehat{\mathbf{y}}^{u*}_{s|T} = \mathbf{R}^u_s\mathbf{H}\widehat{\mathbf{x}}^u_{s|T} + \mathbf{R}^u_s\mathbf{H}_\delta\boldsymbol{\delta}^u_s.$$

The $N^u_{Ds}\times N^u_{Ds}$ variance-covariance matrix of the forecast error in the smoothed predictions of the fitted data at time s conditional on the information set u, $\mathbf{G}^u_s$, is given by

$$\begin{aligned}\mathbf{G}^u_s &\equiv E\left[\mathbf{y}^u_s - \widehat{\mathbf{y}}^u_{s|T}\right]\left[\mathbf{y}^u_s - \widehat{\mathbf{y}}^u_{s|T}\right]^H \\ &= \mathbf{R}^u_s\mathbf{H}\mathbf{P}^u_{s|T}\left[\mathbf{R}^u_s\mathbf{H}\right]^H + \mathbf{R}^u_s\boldsymbol{\Sigma}^u_s\left[\mathbf{R}^u_s\right]^H.\end{aligned}$$

The forecasts and fitted values of our model, which are identical to the expectations of agents, are given as

$$E_u\mathbf{x}_s = \widehat{\mathbf{x}}^u_{s|T} \text{ for } s = 1, \ldots T,$$

Although the Kalman filter algorithm also gives us other useful statistics which we can use in analyzing and presenting the forecast, as we shall see below.

6. Philosophies for finding parameter values for the data measurement equation

To make this idea operational the schemes by which values for the matrices $\mathbf{H}$, $[\boldsymbol{\delta}^u_s]^T_{s=1}$and $[\boldsymbol{\Sigma}^u_s]^T_{s=1}$ that describe the data measurement equation can be found need to be described.

6.1. **Tender loving care.** Priors for each of the elements of these matrices could be justified and imposed on a case by case basis, and the rest estimated. Providing the estimations are identified, a large degree of generality could be allowed for here. The model could be estimated in separate parts (ignoring some of the interrelations) or it could be estimated together, or in blocks. In any case an appropriate estimation method, probably Bayesian, which balances prior information against data information could be applied to the problem. As the priors are not automatic, and both the state-space and the data set could be very large, this could be costly to carry out and possibly even to maintain. But this is probably the most accurate method for building the data uncertainty part of the model, and may be what's needed if the demands of policy are such that the forecasts have to be consistent with all monitored indicators.

6.2. **Purpose-built data set.** The most common method for construct the data set for a forecasting model is by finding one data series to match each important state variable. Then there will be some series, the Lagrange multipliers or the flexible price variables for which no good data series is available at all. They could be left to be solved within the model.

For those variables for which there is a series in the data set, the choice of $\mathbf{H}$ is straightforward: where the model concept has a series to match a state variable, the row and column of $\mathbf{H}$ will be as if it were an identity matrix. For example if the national accounts GDP data corresponds to the value-added output in the model and GDP is the nth data series and value-added output the mth state variable, $\mathbf{H}$ will be have a 1 in entry (m, n) and zero elsewhere in the mth row and the nth column. Even if it can be assumed that data series might on average rise and fall alongside the model concept, there is less reason to argue that it will be unbiased or without some noise. On these grounds, some systematic bias, $\boldsymbol{\delta}_s^u(m) > 0$, and some data measurement error, $\boldsymbol{\Sigma}_s^u(m, m) > 0$, might be allowed interfere in these relations.

Where the model concept has no data, the consistent and transparent solution would be to eliminate that row in $\mathbf{H}$ and let the model solve for these series, by combining the available data with knowledge of model structure and parameter estimates.

This method is less costly, as data series are chosen essentially based on only what state variables are important in the model. But the forecasts may be much worse than an approach that also considers what useful data is available in designing the match. The purpose-built data set ignores the useful technology that is in this paper for bringing in other useful data.

6.3. **Extending the model to fit in data.** But there may be variables which are not needed in the core solution of the model but for which there is useful data. It can be very valuable to extend the model to include data from these variables for two reasons. First if that data is informative it should directly improve the forecast. Second if that data is what agents and policymakers use, the model will need to incorporate that data if it is to predict their behaviour well. One example is money aggregates which on occasion provide very useful and timely information for monetary policy but which do not play a critical part in the solution of many DSGE models. Another example is when there is only annual data available for a quarterly model.

It is assumed that there are N_M of these variables, called $\mathbf{m}_s$ and that can be written as a static function of the variables in the model:

$$\underset{N_M \times N_M}{\boldsymbol{\Xi}_4} \underset{N_M \times 1}{\mathbf{m}_s} = \underset{N_M \times N}{\boldsymbol{\Xi}_3} \underset{N \times 1}{\mathbf{x}_s} + \underset{N_M \times N_M}{\boldsymbol{\Xi}_5} \underset{N_M \times 1}{\boldsymbol{\varepsilon}_{m,s}} \tag{6.1}$$

with the $\boldsymbol{\varepsilon}_{m,s}$ and the $\boldsymbol{\varepsilon}_s$ uncorrelated with each other and $\boldsymbol{\Xi}_4$ non-singular. The key assumption here is that to the solution for these variables can take place in a second stage after the main model has been solved. For example assume that annual data only is available on a flow series, such as GDP

$$y_{a,s} \simeq \sum_{i=0}^{3} 0.25 z_{q,s-i}$$

where $y_{a,s}$ is the data on the annual series and $z_{q,s-i}$ is the unobserved (seasonally adjusted) quarterly state and where both the annual and quarterly series are expressed in terms of log deviations of steady state. This could be introduced into the model in the form of equation 6.1.

Let us assume that these variables are related to the data according to the equation

$$\underset{N^u_{Ds}\times 1}{\mathbf{y}^u_s} = \underset{N^u_{Ds}\times N_{D\max}}{\mathbf{R}^u_s}\ \underset{N_{D\max}\times N_M}{\mathbf{H}_M}\ \underset{N_M\times 1}{\mathbf{m}_s}$$
$$+\mathbf{R}^u_s \underset{N_{D\max}\times N_{D\max}}{\boldsymbol{\Xi}_6}\ \underset{N_{D\max}\times 1}{\boldsymbol{\delta}^u_s} + \mathbf{R}^u_s \underset{N_{D\max}\times N_{D\max}}{\boldsymbol{\Xi}_7}\ \underset{N_{D\max}\times 1}{\mathbf{v}^u_s}.$$

Then this extension of the model can be built in by adapting the measurement equation according to

$$\mathbf{y}^u_s = \mathbf{R}^u_s\left(\mathbf{H}_M\left[\boldsymbol{\Xi}_4\right]^{-1}\boldsymbol{\Xi}_3+\mathbf{H}\right)\mathbf{x}_t + \mathbf{R}^u_s\left(\mathbf{H}_\delta+\boldsymbol{\Xi}_6\right)\boldsymbol{\delta}^u_s$$
$$+\mathbf{R}^u_s\left(\mathbf{I}_{N^u_{Ds}}+\boldsymbol{\Xi}_7\right)\mathbf{v}^u_s + \mathbf{R}^u_s\mathbf{H}_M\left[\boldsymbol{\Xi}_4\right]^{-1}\boldsymbol{\Xi}_5\boldsymbol{\varepsilon}_{m,s}.$$

The rest of the results of the paper would follow if the following replacements are made in the measurement equation

$$\begin{aligned}
\mathbf{H} &\text{ with } \left(\mathbf{H}_M\left[\boldsymbol{\Xi}_4\right]^{-1}\boldsymbol{\Xi}_3+\mathbf{H}\right);\\
\mathbf{H}_\delta &\text{ with } \left(\mathbf{H}_\delta+\boldsymbol{\Xi}_6\right)\\
\mathbf{v}^u_s &\text{ with } \begin{bmatrix}\left(\mathbf{I}_{N^u_{Ds}}+\boldsymbol{\Xi}_7\right)\\ \mathbf{H}_M\left[\boldsymbol{\Xi}_4\right]^{-1}\boldsymbol{\Xi}_5\end{bmatrix}\begin{bmatrix}\mathbf{v}^u_s\\ \varepsilon_{m,s}\end{bmatrix}^H\\
\text{and } \boldsymbol{\Sigma}^u_s &\text{ with } \begin{bmatrix}\left(\mathbf{I}_{N^u_{Ds}}+\boldsymbol{\Xi}_7\right)\\ \mathbf{H}_M\left[\boldsymbol{\Xi}_4\right]^{-1}\boldsymbol{\Xi}_5\end{bmatrix}\times\begin{bmatrix}\boldsymbol{\Sigma}^u_s & \mathbf{0}_{N_{D\max},N_M}\\ \mathbf{0}_{N_{D\max},N_M} & \mathbf{I}_{N_M}\end{bmatrix}^H\times\begin{bmatrix}\left(\mathbf{I}_{N^u_{Ds}}+\boldsymbol{\Xi}_7\right)\\ \mathbf{H}_M\left[\boldsymbol{\Xi}_4\right]^{-1}\boldsymbol{\Xi}_5\end{bmatrix}^H.
\end{aligned}$$

If the form 6.1 is not justified, then the original system would have to be enlarged to incorporate this information.

6.4. **Data intensive methods.** At the other extreme, very many data series, many more than there are states, could be included as in Boivin and Giannoni (2006). With very many data series, it will be infeasible to separately impose priors on the elements of the matrices linking each data series to each unobserved state variables.

What might be possible is to group the data series in terms of which state variable they contain some useful information for, with the groups not necessarily being exclusive. There will be at most N groups of data.

One possible identifying assumption is that the common information that this group contains is proportional to the state variable that indexes that group. Standard dynamic factor analysis methods can be used to estimate and forecast each state variable, with the added twist here that the unobserved behaviour of the state variables is restricted to be consistent with the economic model.

6.5. **Some general principles to choose the data measurement uncertainty.** A more difficult decision is the choice of $\boldsymbol{\Sigma}^u_s$. Imposing case-by-case priors and estimating will almost certainly be infeasible, and also unidentifiable. It would be better to impose some structure to reduce the degrees of freedom.

One solution could be to assume some process for the variance to increase slowly over the sample, starting from time 0 up until the end of the forecast, but with that rate of increase slowing down. The idea is then to estimate or calibrate that process taking each series at a time or much more ambitiously to estimate the series of the matrices as a whole. But there may be some uneven heteroskedasticity around the current period because of the calender of national accounts release. This could be dealt with by spliting the data set into three periods and put in slope dummies into the variance process to adjust for the regime shift.

- The first period goes from time 0 to time $T1$, and takes us up to two quarters before the current quarter for which national accounts data is available.
- The second period is from time $T1$ to time $T2$ which is two quarters after the current quarter. Within this year, centred around the current time, judgement plays a great part.
- The final period is from time $T2$ up until the end of the forecast at time T.

$$\begin{gathered}
\boldsymbol{\Sigma}_s(i,i) = \boldsymbol{\Sigma}_{end}(i,i)^{(1-\alpha_i\delta_{\Sigma i1s}\delta_{\Sigma i2s})}\boldsymbol{\Sigma}_{s-1}(i,i)^{\alpha_i\delta_{\Sigma i1s}\delta_{\Sigma i2s}} \\
\text{for } s = 1, ..., T \\
\text{with } \boldsymbol{\Sigma}_0(i,i) < \boldsymbol{\Sigma}_{end}(i,i); \\
\boldsymbol{\delta}_{\Sigma i1s} = \begin{cases} 1 & \text{for } s < t-3 \text{ qrts} \\ \boldsymbol{\delta}_{\Sigma i1} > 1 & s \geq t-3 \text{ qrts} \end{cases} \\
\boldsymbol{\delta}_{\Sigma i2s} = \begin{cases} 1 & \text{for } s < t+1 \\ \boldsymbol{\delta}_{\Sigma i2} > 1 & s \geq t+1 \text{ qrts} \end{cases} \\
\boldsymbol{\Sigma}_s^u(i,i) = \boldsymbol{\Sigma}_{s-u+t-1}(i,i) \text{ for } s = 1, \ldots, T \text{ and } u = 1, \ldots, t \\
\text{and } \boldsymbol{\Sigma}_s^u(i,j) = 0 \text{ for } i \neq j.
\end{gathered}$$

For each data series, the $T \times t$ data measurement variances across both time and information sets are summarised by five parameters: a starting value $\boldsymbol{\Sigma}_0(i,i)$, a final value $\boldsymbol{\Sigma}_{end}(i,i)$, a rate of improvement α, and two slope dummies to separate out the window surrounding the current time, $\boldsymbol{\delta}_{\Sigma i1}$ and $\boldsymbol{\delta}_{\Sigma i2}$. If a real time data base was available, that could be used that to get estimates of the scale of data uncertainty at a particular lag, for example at $t-3$, and through that route estimate $\boldsymbol{\Sigma}_{t-3}(i,i)$. These kind of estimates should help pin down values of the parameters here.

7. Useful outputs from the forecasting system

A policy forecast is judged not just on its predictability, but also on how well it tells stories. Indeed, often the job of the policy forecaster is to explain what went wrong! In order to do this, it is important to be able to decompose the forecast in two dimensions: in terms of the data, and second in terms of the economic and data measurement shocks. The contributions on either dimension are myriad, and so it is also important to think of interesting ways of summarising this information. Finally it would be useful to have a metric for assessing the information worth of bits of data or whole series.

7.1. **Data decompositions.** Consider the $N \times N_{Ds}^u$ multiplier matrices $\left(\mathbf{w}_j^u\left(\widehat{\mathbf{x}}_{s|T}^u\right) \text{ for } j=0,\ldots,T\right)$ which when multiplied by the selector matrix $\mathbf{R}_j^u$ determine how much each piece of information (data plus adjustments) from the mythical full data set at time s based on the information set u $\left(\mathbf{y}_j^{u*}\right)$ affects the smoothed estimate at time s. These multipliers are defined by the relationship

$$\widehat{\mathbf{x}}_{s|T}^u = \sum_{j=0}^{T} \mathbf{w}_j^u\left(\widehat{\mathbf{x}}_{s|T}^u\right)\mathbf{R}_j^u\left[\mathbf{y}_j^{u*} - \mathbf{H}_\delta \delta_j^u\right]. \tag{7.1}$$

Define the $N \times N_{Ds}^u$ matrices $\mathbf{\Psi}_s^u$,

$$\begin{aligned}\mathbf{\Psi}_s^u \equiv &\left[\mathbf{R}_{s-1}^u\mathbf{H}\right]^H\left[\left[\mathbf{L}_s^u\right]^{-1} + \left[\mathbf{K}_s^u\right]^H\left[\mathbf{R}_s^u\mathbf{H}\right]^H\left[\mathbf{L}_s^u\right]^{-1}\mathbf{R}_s^u\mathbf{H}\mathbf{K}_s^u\right]\\ &-\left[\mathbf{\Xi}_1\right]^H\left[\mathbf{R}_s^u\mathbf{H}\right]^H\left[\mathbf{L}_s^u\right]^{-1}\mathbf{R}_s^u\mathbf{H}\mathbf{K}_s^u \text{ for } s=0,\ldots,T.\end{aligned}$$

Then the multiplier matrices are calculated in part backwards and in part forwards from time s. The recursion formulas from Koopman and Harvey (2003) are

$$\begin{aligned}\mathbf{w}_j^u\left(\widehat{\mathbf{x}}_{s|T}^u\right) &= \left(\mathbf{I}_{N\times N} - \mathbf{P}_{s|s-1}^u\left[\mathbf{R}_{s-1}^u\mathbf{H}\right]^H\left[\mathbf{L}_{s-1}^u\right]^{-1}\mathbf{R}_{s-1}^u\mathbf{H}\right)\mathbf{w}_j^u\left(\widehat{\mathbf{x}}_{s|s-1}^u\right);\\ \mathbf{w}_j^u\left(\widehat{\mathbf{x}}_{s|s-1}^u\right) &= \mathbf{\Theta}_{s,j-1}^u\mathbf{K}_s^u \text{ and } \mathbf{\Theta}_{s,j-1}^u = \mathbf{\Theta}_{s,j}^u\mathbf{\Xi}_1 - \mathbf{w}_j^u\left(\widehat{\mathbf{x}}_{s|s-1}^u\right)\mathbf{R}_{s-1}^u\mathbf{H}\\ \text{with } \mathbf{\Theta}_{s,s-1}^u &= \mathbf{I}_{N\times N} - \mathbf{P}_{s|s-1}^u\left[\mathbf{R}_{s-1}^u\mathbf{H}\right]^H\left[\mathbf{L}_{s-1}^u\right]^{-1}\mathbf{R}_{s-1}^u\mathbf{H}\\ &\text{for } j=s-1,\ldots,0.\end{aligned}$$

and

$$\begin{aligned}\mathbf{w}_j^u\left(\widehat{\mathbf{x}}_{s|T}^u\right) &= \mathbf{\Theta}_{s,j}^{*u}\mathbf{\Psi}_s^u\\ \mathbf{\Theta}_{s,j+1}^{*u} &= \mathbf{\Theta}_{s,j}^{*u}\left[\mathbf{\Xi}_1 - \mathbf{K}_s^u\mathbf{R}_s^u\mathbf{H}\right]^H\\ \text{with } \mathbf{\Theta}_{s,s}^{*u} &= \mathbf{P}_{s|s-1}^u\\ &\text{for } j=s,\ldots,T.\end{aligned}$$

The *contributions* of each piece of data to these estimates are then simply that piece of data multiplied by its multiplier, only that here they are also indexed by the information sets through u.

Publishing forecasts is about communications as much as anything else. But if the central bank wants to communicate to the public it should always refer to data that is publicly and objectively available. For that reason it is important also to present the contributions of data to the smoothed estimates of the data series themselves. These are given by

$$\mathbf{R}_s^u\widehat{\mathbf{y}}_{s|T}^{u*} = \mathbf{R}_s^u\mathbf{H}\sum_{j=0}^{T}\mathbf{w}_j^u\left(\widehat{\mathbf{x}}_{s+1|T}^u\right)\mathbf{R}_j^u\left(\mathbf{y}_j^{u*} - \mathbf{H}_\delta\boldsymbol{\delta}_j^u\right) + \mathbf{R}_s^u\mathbf{H}_\delta\boldsymbol{\delta}_s^u. \tag{7.2}$$

A popular maxim in policy forecasting is that it is the source of the shock that matters, meaning that the explanations of the forecast depend very much on what shocks are believed driving the forecast. It would be therefore useful to know how we determine these shocks, in terms of which bits data tells us that

there is a demand versus a supply side shock for example. The economic shocks are given as the last kN_ε members of the state vector, and so equation 7.1 can be used to generate this interesting decomposition.

7.1.1. *Summarizing the information given by the data decompositions.* The decompositions in this section provide a myriad of information. It is usually convenient to summarize that information along some relevant dimension. There are many interesting possibilities but three standard aggregations spring to mind:

(1) *The contribution of each individual data series* in a given set is given by summing the contributions of observations on each series over time.
(2) *The impact of the news in each data series on each forecasted variable* can be calculated by subtracting the contribution of each data series as it was in the previous information set from the contribution of that series as it is in the current data set.
(3) The data set into the part that reflects off-model judgement and the rest. *The role of judgements* is then the sum of the contributions of those pieces of data which are designated to be judgement.
(4) *T*hese calculations can be only in terms of *the contribution only to key variables*, for example inflation and GDP, and then only *over the more interesting periods*, such as the forecast.

7.1.2. *The information contribution of each piece of data to the forecast.* The expected contributions of particular pieces of data to the forecasted variable were derived previously in this section. It is interesting to complement that with some measure of how much useful information each data series brings to the forecast, which can be thought of as a forecast error variance decomposition in terms of the data series.

Tinsley, Spindt, and Friar (1980) and later Coenen, Levin, and Wieland (2004) describe how the information contribution of a piece of data to a forecast is related to the reduction in uncertainty that using that data brings to the forecast. The expected uncertainty in two sets of variables x and y is defined as

$$H(y,x) \equiv -E\left[\ln\left(f(y,x)\right)\right]$$

where $f(y,x)$ is their joint density. Consistently the uncertainty in just y is

$$H(y) = -E\left[\ln\left(f(y)\right)\right]$$

where $f(y)$ is the marginal density of y, and the conditional uncertainty of y given x is

$$H(y \mid x) = -E\left[\ln\left(f(y \mid x)\right)\right]$$

where $f(y \mid x)$ is the conditional density of y given x. The mutual information content of x and y is then defined by the reduction in expected uncertainty when x is used to predict y:

$$I(y \mid x) \equiv H(y) - H(y \mid x)$$

and is always positive definite. If y and x are jointly normally distributed, Tinsley, Spindt, and Friar (1980) show that the mutual information content is then

$$I(y \mid x) = 0.5 \ln\left(\frac{\mathbf{\Omega}_{yy}}{\left|\mathbf{\Omega}_{yy} - \mathbf{\Omega}_{yx}\left[\mathbf{\Omega}_{xx}\right]^{-1}\mathbf{\Omega}_{xy}\right|}\right)$$

where $\mathbf{\Omega}_{z_1 z_2}$ is the covariance matrix between two vectors z_1 and z_2. If x is partioned into two sets of information, the extra gain in using x_1 alongside x_2 over just using x_1 is $G(y \mid x_1, x_2)$ where

$$\begin{aligned} G(y \mid x_1, x_2) &\equiv I(y \mid x) - I(y \mid x_1) \\ &= 0.5 \ln \left(\frac{\left| \mathbf{\Omega}_{yy} - \mathbf{\Omega}_{yx_1} \left[\mathbf{\Omega}_{x_1 x_1} \right]^{-1} \mathbf{\Omega}_{x_1 y} \right|}{\left| \mathbf{\Omega}_{yy} - \mathbf{\Omega}_{yx} \left[\mathbf{\Omega}_{xx} \right]^{-1} \mathbf{\Omega}_{xy} \right|} \right). \end{aligned} \tag{7.3}$$

This is the percentage difference between the square root of the determinant of the conditional variance covariance matrix of y just using x_1 and the square root of the determinant of the conditional variance covariance matrix using both x_1 and x_2. As such it seems to be a good measure of the marginal information content of x_2 over and above x_1. Note that it is not the same as mutual information content of y and x_2, which is $I(y \mid x_2)$.

This measures the information value to estimating all the states in the forecast over the whole sample. But what is more likely to be of interest is to assess the information content of a data series in terms of predicting just one state variable, and even then over a particular time window. For example the relevant question might be, how useful is a particular data series in predicting inflation over the two-year forecast horizon? An appropriate answer would then be given by the following statistic based on a scalar version of 7.3 above:

Proposition. *The extra information content of the jth data series in predicting the ith variable on average over the two year forecast horizon could then be measured by*

$$G^{u}_{s=t,t+7}(i,j) \equiv \left(0.5 \sum_{\mathrm{s=t}}^{\mathrm{t+7}} \ln \frac{\left(\mathbf{P}^{u}_{s|T}(i,i) \right)}{\left(\mathbf{P}^{u}_{s|T^{-j}}(i,i) \right)} \right) \tag{7.4}$$

where the ith variable is of interest, t is the current time and the frequency is quarterly, and T^{-j} indicates that all but the jth data series whose information content we want to measure to calculate the covariance of the states is being used.

7.2. **Extracting information on the economic content of the model.** It is also interesting to see to what extent the forecasted values are driven by economic shocks. The economic content of the model can also be presented as impulse responses or variance decompositions as in Gerali and Lippi (2003).

7.2.1. *Economic decompositions.* The last kN_ε state variables describe the exogenous economic shocks, also in terms of when they were first spotted. $\widehat{\mathbf{z}}^{\varepsilon u}_{s|T}$ refers to the vector of state estimates of those state variables at time s using information set u. The lower partition described in equation 4.2 and these smoothed state estimates to recover smoothed estimates of the white noise components of these shocks as

$$\widehat{\varepsilon}^{u}_{s|T} = \left[\mathbf{D}_\varepsilon \right]^{-1} \left[\widehat{\mathbf{z}}^{\varepsilon u}_{s|T} - \mathbf{B}_\varepsilon \widehat{\mathbf{z}}^{\varepsilon u}_{s-1|T} \right]. \tag{7.5}$$

Then the contribution of the ith shock to the smoothed estimate at time s on information set u is

$$\widehat{\mathbf{x}}_{s|T}^{\varepsilon iu} = \sum_{j=0}^{s} [\boldsymbol{\Xi}_1]^{-j} \boldsymbol{\Xi}_2 \mathbf{R}^{\varepsilon i} \widehat{\boldsymbol{\varepsilon}}_{j|T}^{u}, \tag{7.6}$$

where $\mathbf{R}^{\varepsilon i}$ is a $k(N - N_\varepsilon) \times k(N - N_\varepsilon)$ matrix formed by taking an identity matrix and putting zeros in all but the ith diagonal element. The contribution to the smoothed fitted values of the data series will be

$$\mathbf{R}_s^u \widehat{\mathbf{y}}_{s|T}^{\varepsilon iu*} = \mathbf{R}_s^u \mathbf{H} \widehat{\mathbf{x}}_{s|T}^{\varepsilon iu}.$$

Here it should be remembered that the sum of the contributions across all shock will equal the smoothed estimate of the data series minus the mean bias terms ($\mathbf{R}_s^u \mathbf{H}_\delta \boldsymbol{\delta}_s^u$).

7.2.2. *Impulse responses.* To obtain the economic decomposition of the states, the decomposition is first written as

$$\begin{aligned} \widehat{\mathbf{x}}_{s|T}^{u} &= \sum_{j=0}^{T} \mathbf{W}_{s,j|T}^{u} \left(\mathbf{y}_j^{u*} - \mathbf{H}_\delta \delta_j^u \right) \\ &= \sum_{j=0}^{T} \mathbf{W}_{s,j|T}^{u} \left[\mathbf{H}\mathbf{x}_j + \mathbf{v}_j^u \right] \end{aligned}$$

where the $N \times N_{D\max}$ matrices $\mathbf{W}_{s,j|T}^{u}$ are given by $\mathbf{W}_{s,j|T}^{u} \equiv \mathbf{w}_j^u \left(\widehat{\mathbf{x}}_{s|T}^{u} \right) \mathbf{R}_j^u$.

Substituting out for the true states gives

$$\widehat{\mathbf{x}}_{s|T}^{u} = \sum_{j=0}^{T} \mathbf{W}_{s,j|T}^{u} \left[\mathbf{H} \sum_{k=0}^{j} (\boldsymbol{\Xi}_1)^{j-k} \boldsymbol{\Xi}_2 \boldsymbol{\varepsilon}_k + \mathbf{v}_j^u \right] \tag{7.7}$$

which decomposes the state estimate into the contribution of each economic shock and each data noise shock. Impulse responses to economic shocks or data errors can now be derived from expression 7.7.

Similarly the smoothed estimates of the data and the policy interest rate can be decomposed into shocks with a view to obtaining the impulse responses to the fitted data values:

$$\mathbf{R}_s^u \widehat{\mathbf{y}}_{s|T}^{u*} = \mathbf{R}_s^u \mathbf{H} \sum_{j=0}^{T} \mathbf{W}_{s,j|T}^{u} \left[\mathbf{H} \sum_{k=0}^{j} (\boldsymbol{\Xi}_1)^{j-k} \boldsymbol{\Xi}_2 \boldsymbol{\varepsilon}_k + \mathbf{v}_j^u \right] + \mathbf{R}_s^u \mathbf{H}_\delta \delta_s^u.$$

7.2.3. *Variance decompositions.* The variance decomposition of a state estimate is the proportion of variance explained by the variance of each type of shock at each horizon. The unconditional variance of the

state estimates is first derived as

$$
\begin{aligned}
& E\left[\widehat{\mathbf{x}}_{s|T}^{u}\left[\widehat{\mathbf{x}}_{s|T}^{u}\right]^{H}\right] \\
&= \sum_{k=0}^{T}\left[\sum_{j=k}^{T}\mathbf{W}_{s,j|T}^{u}\mathbf{H}\left(\boldsymbol{\Xi}_{1}\right)^{j-k}\boldsymbol{\Xi}_{2}\right]E\left[\boldsymbol{\varepsilon}_{k}\left[\boldsymbol{\varepsilon}_{k}\right]^{H}\right]\left[\sum_{j=k}^{T}\mathbf{W}_{s,j|T}^{u}\mathbf{H}\left(\boldsymbol{\Xi}_{1}\right)^{j-k}\boldsymbol{\Xi}_{2}\right]^{H} \\
&+ \sum_{j=0}^{T}\mathbf{W}_{s,j|T}^{u}E\left[\mathbf{v}_{j}\left[\mathbf{v}_{j}\right]^{H}\right]\left[\mathbf{W}_{s,j|T}^{u}\right]^{H} \\
&= \sum_{k=0}^{T}\left[\sum_{j=k}^{T}\mathbf{W}_{s,j|T}^{u}\mathbf{H}\left(\boldsymbol{\Xi}_{1}\right)^{j-k}\boldsymbol{\Xi}_{2}\right]E\left[\boldsymbol{\varepsilon}_{k}\left[\boldsymbol{\varepsilon}_{k}\right]^{H}\right]\left[\sum_{j=k}^{T}\mathbf{W}_{s,j|T}^{u}\mathbf{H}\left(\boldsymbol{\Xi}_{1}\right)^{j-k}\boldsymbol{\Xi}_{2}\right]^{H} \\
&+ \sum_{j=0}^{T}\mathbf{W}_{s,j|T}^{u}\boldsymbol{\Sigma}_{j}^{u}\left[\mathbf{W}_{s,j|T}^{u}\right]^{H}. \qquad (7.8)
\end{aligned}
$$

Given a suitable calibration about the variance covariance matrix of economic shocks $\left(E\left[\boldsymbol{\varepsilon}_{k}\left[\boldsymbol{\varepsilon}_{k}\right]^{H}\right]\right)$ which could be an identity matrix, expression 7.8 permits the decomposition.

It would be useful to draw bands of uncertainty around the forecasts. The forecast error of the estimated state variable is

$$
\begin{aligned}
\mathbf{x}_{s}-\widehat{\mathbf{x}}_{s|T}^{u} &= \mathbf{x}_{s}-\sum_{j=0}^{T}\mathbf{W}_{s+1,j|T}^{u}\left[\mathbf{H}\sum_{k=0}^{j}\left(\boldsymbol{\Xi}_{1}\right)^{j-k}\boldsymbol{\Xi}_{2}\boldsymbol{\varepsilon}_{k}+\mathbf{v}_{j}^{u}\right] \\
&= \sum_{k=0}^{s}\left(\boldsymbol{\Xi}_{1}\right)^{s-k}\boldsymbol{\Xi}_{2}\boldsymbol{\varepsilon}_{k}+\mathbf{v}_{s}^{u}-\sum_{j=0}^{T}\mathbf{W}_{s,j|T}^{u}\left[\mathbf{H}\sum_{k=0}^{j}\left(\boldsymbol{\Xi}_{1}\right)^{j-k}\boldsymbol{\Xi}_{2}\boldsymbol{\varepsilon}_{k}+\mathbf{v}_{j}^{u}\right].
\end{aligned}
$$

So if by defining the matrices $\mathbf{W}_{s,j|T}^{*u}$ such that

$$
\mathbf{W}_{s,j|T}^{*u}=\begin{cases}\mathbf{W}_{s,j|T}^{u} \text{ for } j\neq s \\ \mathbf{I}_{N}-\mathbf{W}_{s,j|T}^{u} \text{ for } j=s\end{cases}
$$

then the forecast error covariance of the state estimates can be written as

$$
\begin{aligned}
\mathbf{G}_{s}^{u} &= E\left[\left[\mathbf{x}_{s}-\widehat{\mathbf{x}}_{s|T}^{u}\right]\left[\mathbf{x}_{s}-\widehat{\mathbf{x}}_{s|T}^{u}\right]^{H}\right] \\
&= \sum_{k=0}^{T}\left[\sum_{j=k}^{T}\mathbf{W}_{s,j|T}^{*u}\mathbf{H}\left(\boldsymbol{\Xi}_{1}\right)^{j-k}\boldsymbol{\Xi}_{2}\right]E\left[\boldsymbol{\varepsilon}_{k}\left[\boldsymbol{\varepsilon}_{k}\right]^{H}\right]\left[\sum_{j=k}^{T}\mathbf{W}_{s,j|T}^{*u}\mathbf{H}\left(\boldsymbol{\Xi}_{1}\right)^{j-k}\boldsymbol{\Xi}_{2}\right]^{H} \\
&+ \sum_{j=0}^{T}\mathbf{W}_{s,j|T}^{u}\sum_{j=0}^{T}E\left[\mathbf{v}_{j}^{u}\left[\mathbf{v}_{j}^{u}\right]^{H}\right]\left[\mathbf{W}_{s,j|T}^{u}\right]^{H} \\
&+ E\left[\mathbf{R}_{s}^{u}\mathbf{v}_{s}^{u}\left[\mathbf{R}_{s}^{u}\mathbf{v}_{s}^{u}\right]^{H}\right]
\end{aligned}
$$

or

$$\begin{aligned}
&\mathbf{G}_s^u \\
&= \sum_{k=0}^{T} \left[\sum_{j=k}^{T} \mathbf{W}_{s,j|T}^{*u} \mathbf{H} \left(\mathbf{\Xi}_1 \right)^{j-k} \mathbf{\Xi}_2 \right] E \left[\boldsymbol{\varepsilon}_k \left[\boldsymbol{\varepsilon}_k \right]^H \right] \left[\sum_{j=k}^{T} \mathbf{W}_{s,j|T}^{*u} \mathbf{H} \left(\mathbf{\Xi}_1 \right)^{j-k} \mathbf{\Xi}_2 \right]^H \\
&+ \sum_{j=0}^{T} \mathbf{W}_{s,j|T}^{u} \mathbf{\Sigma}_j^u \left[\mathbf{W}_{s,j|T}^{u} \right]^H \\
&+ \mathbf{R}_s^u \mathbf{\Sigma}_s^u \left[\mathbf{R}_s^u \right]^H .
\end{aligned}$$

Similarly the covariance of the estimates of the fitted variables is

$$E \left[\mathbf{R}_s^u \widehat{\mathbf{y}}_{s|T}^{u*} \left[\mathbf{R}_s^u \widehat{\mathbf{y}}_{s|T}^{u*} \right]^H \right] = \mathbf{R}_s^u \mathbf{H} E \left[\widehat{\mathbf{x}}_{s|T}^u \left[\widehat{\mathbf{x}}_{s|T}^u \right]^H \right] \left[\mathbf{R}_s^u \mathbf{H} \right]^H$$

with the covariance of the forecast error of fitted data in terms of shocks as

$$E \left[\left[\mathbf{y}_{s|T}^u - \mathbf{R}_s^u \widehat{\mathbf{y}}_{s|T}^{u*} \right] \left[\mathbf{y}_{s|T}^u - \mathbf{R}_s^u \widehat{\mathbf{y}}_{s|T}^{u*} \right]^H \right] = \mathbf{R}_s^u \mathbf{H} \mathbf{G}_s^u \left[\mathbf{R}_s^u \mathbf{H} \right]^H + \mathbf{R}_s^u \mathbf{\Sigma}_s^u \left[\mathbf{R}_s^u \right]^H .$$

It should be noted however that these expressions do not take account of estimation uncertainty. That would seem to be more obviously deficient in a set up where the data are formally separated from the model variables. In our opinion, then these expressions should be superceded by a formal combined treatment of uncertainty and estimation, and one which is designed for policy. Such a scheme is outlined in Sims (2002) and applied by Adolfson, Andersson, Lindé, Villani, and Vredin (2005).

8. Allowing for reporting variables

One way to better exploit this wealth of information, is to introduce variables into the economic model which are there just for reporting. This is a very common practice in policy forecasting as round by round the way the forecast is presented changes depending on particular issues.

Formally, the reporting variables could follow the process

$$\mathbf{rp}_t = \mathbf{H}_{rp} \mathbf{x}_s + \epsilon_{rp,s}$$

where $\epsilon_{rp,s}$ is a vector of white noise residuals, allowing for the fact that reporting is not exact. Schorfheide, Sill, and Kryshko (2008) discuss the importance of allowing for these non-modelled variables and suggest how such auxiliary equations can be estimated. Here more simply, the variance of these residuals is assumed to be given. In keeping with the spirit of what a reporting variable is it is also assumed that these residuals are independent of the shocks and data noise terms in the rest of the model. Otherwise these variables would have to be incorporated into the model.

The smoothed estimates of the reporting variables are

$$\widehat{\mathbf{rp}}_{s|T}^u = \mathbf{H}_{rp} \widehat{\mathbf{x}}_{s|T}^u + \epsilon_{rp,s}.$$

Then the forecast of these variables can be decomposed into the contributions of economic shocks with

$$\widehat{\mathbf{rp}}_{s|T}^{\varepsilon iu} = \mathbf{H}_{rp}\widehat{\mathbf{x}}_{s|T}^{\varepsilon iu}$$

from equation 7.6.

The impulse responses are

$$\widehat{\mathbf{rp}}_{s|T}^{u} = \mathbf{H}_{rp}\left[\sum_{j=0}^{T}\mathbf{W}_{s+1,j|T}^{u}\left[\mathbf{H}\boldsymbol{\Xi}_2\sum_{k=0}^{j}\left(\boldsymbol{\Xi}_1\right)^{j-k}\boldsymbol{\Xi}_2\boldsymbol{\varepsilon}_k + \mathbf{v}_j^u\right]\right] + \epsilon_{rps};$$

and the variance decomposition follows

$$E\left[\widehat{\mathbf{rp}}_{s|T}^{u}\left[\widehat{\mathbf{rp}}_{s|T}^{u}\right]^H\right] = \mathbf{H}_{rp}E\left[\widehat{\mathbf{x}}_{s|T}^{u}\left[\widehat{\mathbf{x}}_{s|T}^{u}\right]^H\right]\left[\mathbf{H}_{rp}\right]^H.$$

The variance-covariance matrix of forecast errors is

$$E\left[\left[\mathbf{rp}_{s|T}^{u} - \widehat{\mathbf{rp}}_{s|T}^{u}\right]\left[\mathbf{rp}_{s|T}^{u} - \widehat{\mathbf{rp}}_{s|T}^{u}\right]^H\right] = \mathbf{H}_{rp}\mathbf{G}_s^u\left[\mathbf{H}_{rp}\right]^H.$$

9. The state identification problem

An identification problem arises when the values of parameters cannot be separately estimated given the combination of a theoretical model and a data set. This type of identification problems bedevil the estimation of DSGE models, as Canova and Sala (2006), Fukac and Pagan (2006) and Guerron-Quintana (2007) explain. Here it is assumed that the parameter values are known, and so that particular problem has been set aside. However it is still very possible that there is an identification problem in the estimates of state variables: the data and model together might fail to separately identify some states. Indeed as the solution for states is a first step in the recursions used to estimate parameters, for example as in maximum likelihood estimation, state identification could be seen a necessary condition for parameter identification. These problems are not generic to the proposal in this paper but this procedure will certainly bring these problems to the fore. Therefore it is worthwhile to spend some time discussing the problem of state identification.

Weak identification can be expected to especially affect our estimates of the economic shocks. Remember that economic shocks have no direct data counterpart to tie their value down to. Then it becomes more likely that the forecast of an variable even on which good data is available is based on very unreliable estimates of what shocks cause those movements.

There are many different ways in whch these problems can show themselves, if one digs deep enough. One classic symptom would be if the estimates of two shocks that affect one variable for which there is a data counterpart are estimated to have large but offsetting contributions. Another would be if the estimate of a shock fails to update from its initial prior value. A different way of looking at this problem is to note that the data series are not separately informative about state estimates, and are in this sense multicollinear. This then points to yet another sympton which would be if the estimates of particular states were too sensitive to changes in the data set, as explained in Watson (1983).

One subcategory of this shock identification problem are when data measurement errors cannot be separately estimated from economic shocks. This could reveal itself when the estimate of a data measurement error is negatively correlated with the economic shock. Remember that the covariances between data measurement and economic shocks are assumed to be zero. Relaxing that assumption would be more realistic but also open us up to more identification errors of this type.

It is difficult to offer general solutions to state identification problems. It is important to remember that a failure of identification is to do with the flawed combination of model and data and neither individually. Thus there may be some shocks that can be identified with some data but not with others. And there may be some shocks which are not identifiable by any data that are available! The solution to identification may lie either in changing the model; or in looking for new data; and will probably involve some compromises in either or both directions. Solutions would seem to be case by case.

But it is possible to formulate general tests to detect identification problems. Burmeister and Wall (1982) offer one test for the identification problems for parameter estimates in state space models. Their suggestion is to look out for very large correlations among parameter estimates. Analogously here high correlations among the estimates of the economic shocks would reveal poor identification.

The variance-covariance matrix of the impact of the economic shocks (which are included in the state vector) is given in equation 7.8. Then the following test can be constructed:

Proposition. *Given an information set u, the estimate of the economic shock impacts i and j at time s are not likely to be well identified separately if*

$$\left| \frac{\mathbf{P}^{u}_{s|T}(i,j)}{\left(\mathbf{P}^{u}_{s|T}(i,i)\,\mathbf{P}^{u}_{s|T}(j,j)\right)^{0.5}} \right| > 0.95. \tag{9.1}$$

The identification problem can also be assessed from the point of view of whether adding a particular series to a given data set brings in more information in identifying a group of variables, using the test 7.4.

10. Some applications

In this section some experiments from applying this approach to a quarterly DSGE model for Colombia are reported. The model (PATACON) is described in González, Mahadeva, Prada, and Rodríguez (2008). It is a open economy model which and has been calibrated to fit Colombian data for the period 2003-2006. The database, of a quarterly frequency, runs from 1994Q1 to 2006Q4 is described in Mahadeva and Parra (2008).

10.1. **A complete forecast.** To begin with, the first aim is to show that this model and its database is at least capable of generating a decent forecast. One can in principle generate a forecast with very little data, but that forecast could be quite poor. Adding more data can in principle improve that forecast.

The next two figures show the forecasts of the model for quarterly real value-added income (nominal GDP divided by the CPI) growth. 12 series are included in the data base up until period 40 (10 years of data), and it is assumed that they are all perfectly measured. In this experiment, the model allows for 12 structural shocks all of which seem to reasonably well identified.

FIGURE 10.1. Forecast of Real GDP growth with data up to period 40

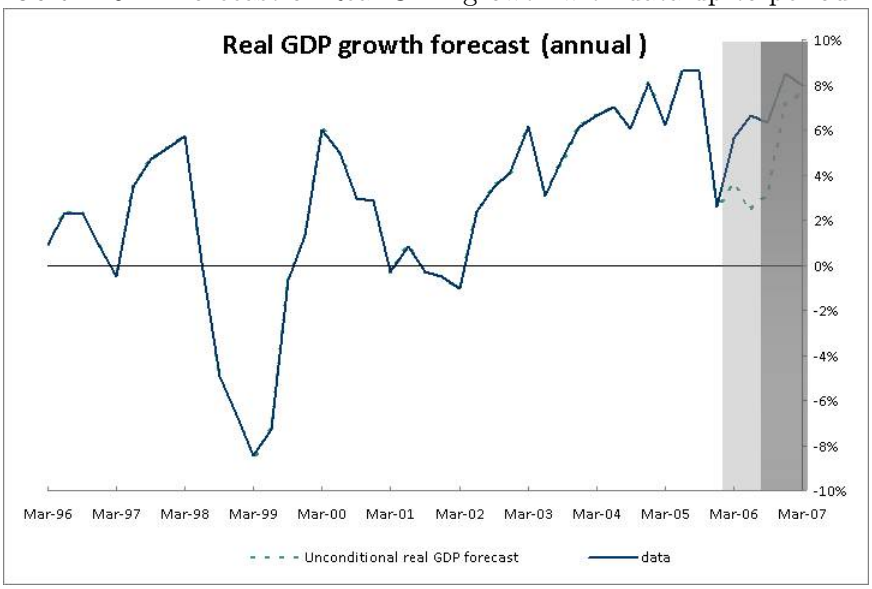
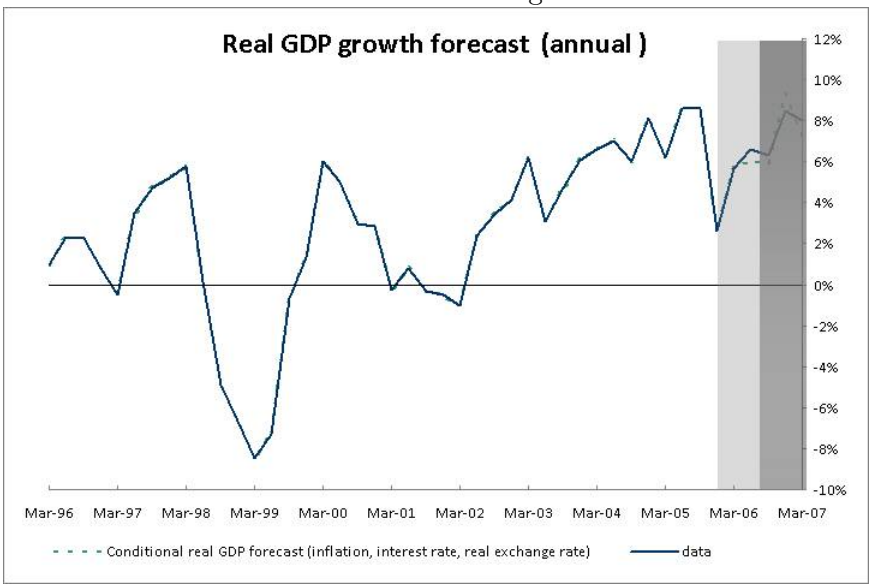


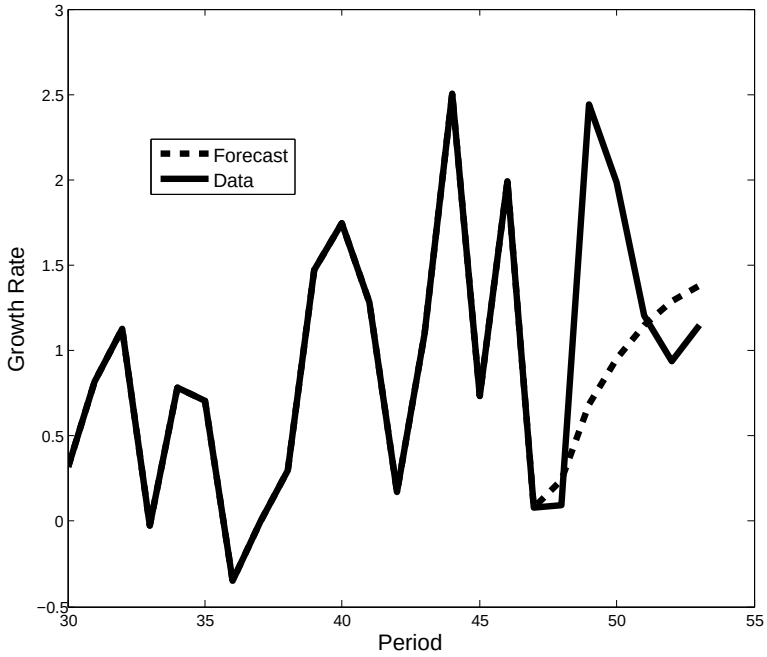
FIGURE 10.2. Forecast of Real GDP growth with extra data



The first figure presents the forecast with no other data beyond the 40 quarters cutoff point. The forecast in the second figure is based on an extra two quarters of inflation, interest rate and real exchange rate data.

Comparing the two forecasts makes the simple point that this model is capable of making a decent forecast, but only if the information on recent data is included. This is true even if that information is from other series, given that the theoretical part of the model is able to interpret the linkages between these series.

FIGURE 10.3. Forecast of consumption growth without extra CPI data



10.2. **Experiment to show the importance of imperfectly measured awkward data.** This next set of experiments is about the advantage of allowing for measurement error in unbalanced data. The first in this next set of experiments uses data on consumption and inflation only until period 47 and no more and assumes that both series are well measured. The results are in Figure 10.3. Then three extra quarters of data on CPI inflation is introduced, while still assuming that both those series are well measured. This is a very similar exercise to that of the previous section. Comparing Figure 10.3 with Figure 10.4, it seems that the extra data on CPI ensures that the boom in consumption in between periods 47-52 was picked up.

However we should be careful about leaping to the conclusion that more data always improves a forecast. All that has been shown is that the extra data on prices helps the forecast track consumption better. If those extra two data points in the CPI were badly measured, then what we have produced is an even worse forecast. This is a very real possibility in Colombia because the statistical authority only updates its aggregate CPI weights once every five years, and then even near term forecasts are subject to measurement error.

To bring out this point, the next experiment compare the previous two forecasts with one in which we allow for data mismeasurement in the extra three CPI data points, in Figure 10.5 This forecast is now a compromise between the two others and perhaps reflects a better balance between not ignoring important data and not chasing it too closely.

10.3. **Experiments to show the importance of off-model forecasts.** The next set of figures illustrate the use of the data decompositions. In what follows there are only two economic shocks, a demand and a supply shock, and it is assumed that the supply shock dominates. In the first experiment there is only

FIGURE 10.4. Forecast of consumption growth with CPI data

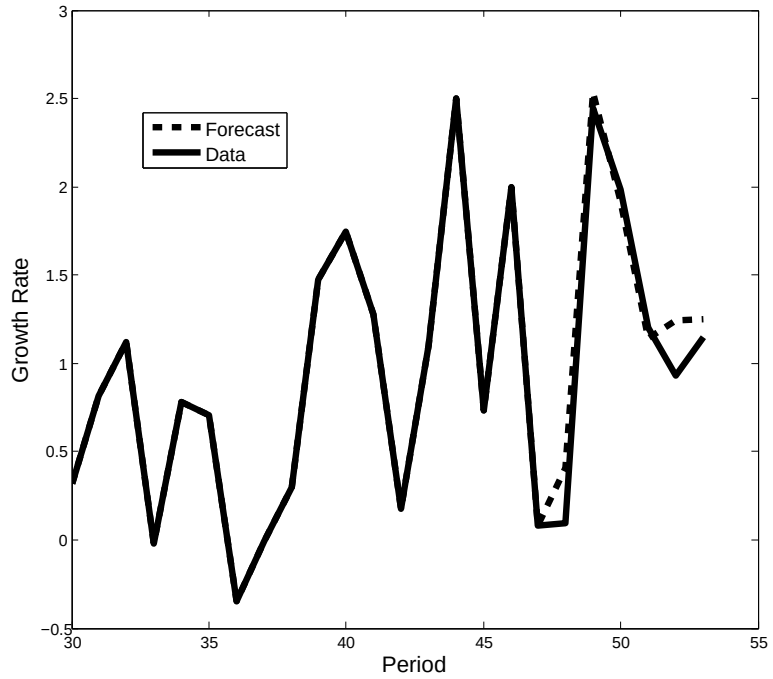


FIGURE 10.5. Forecast of consumption growth allowing for error in CPI data

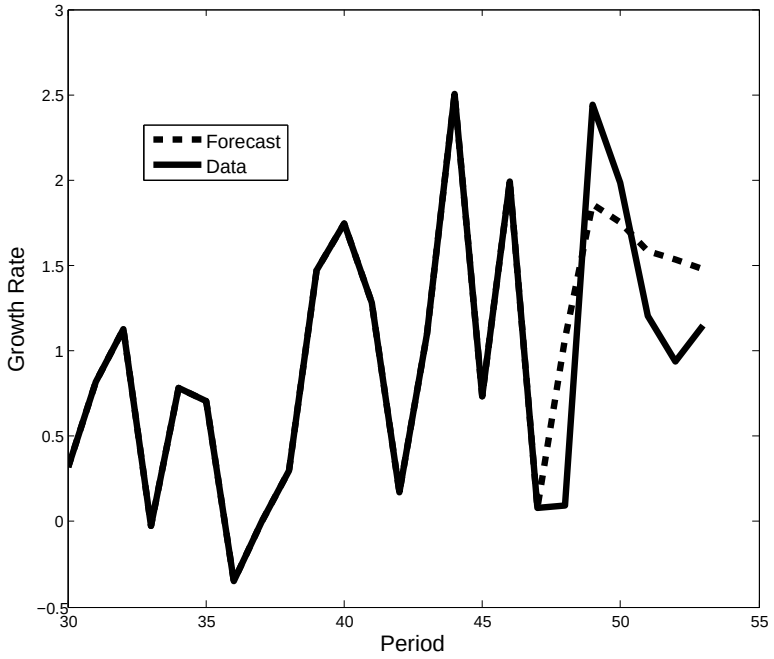
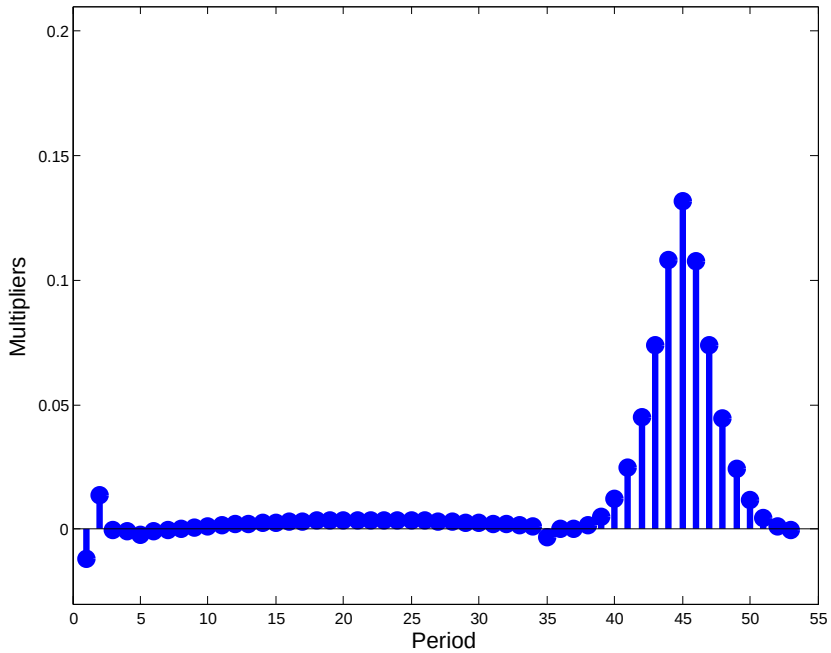


FIGURE 10.6. Kalman Smoother Multipliers on Consumption Data



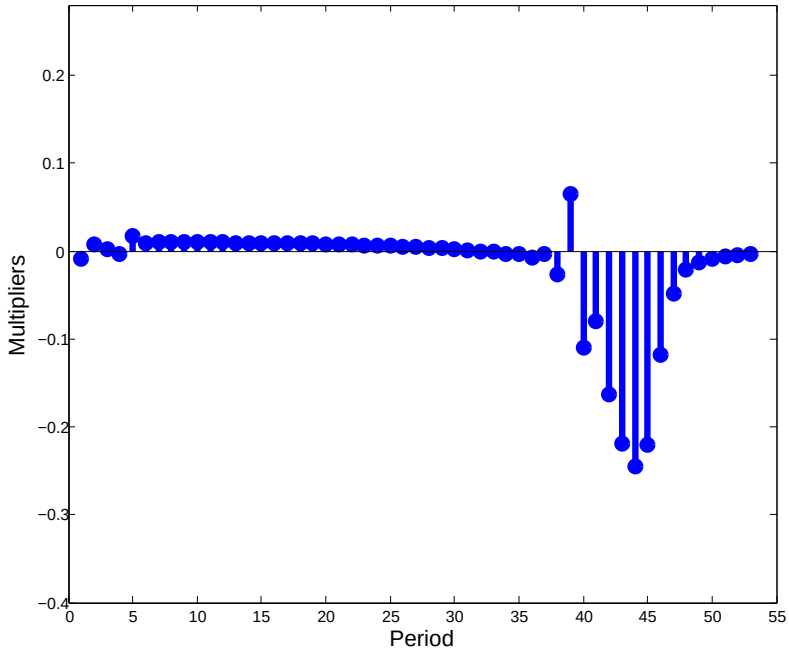
Included Series: Consumption.
Multipliers corresponding to time series data of: Consumption.
Series with measurement error: Consumption.
Last period of consumption available data: 53.

noisy data on consumption. The task is to predict all variables in the model including true consumption itself. Figure 10.6 shows the multipliers of consumption data in predicting true consumption in period 45 in this experiment. Notice that the distribution is symmetric about the current period. Given that there are measurement errors, surrounding data helps to estimate the unobserved state at the current period, only that here that pattern reflects in part the dynamics of the DSGE model.

The next experiment looks into exactly how the information from off-model forecast of inflation in the near term helps in predicting consumption now. The experiment assumes perfectly measured data on consumption and inflation only until period 40, and then inflation data only for the next two years. These extra two years of inflation data are assumed to have measurement error, just as if they were forecasts from a separate inflation model.

The multipliers on estimating consumption in period 45, shown in Figure 10.7, describe how inflation data even a year ahead plays some part in helping us understand what is happening to consumption in the absence of timely consumption data. The multipliers are negative because higher prices imply a lower level of consumption, conditional on the supply shock being important. Shown this figure, a policymaker would understand how his off-model inflation forecast is consistent with his forecast for consumption and so ultimately GDP.

FIGURE 10.7. Kalman Smoother Multipliers on Inflation Data

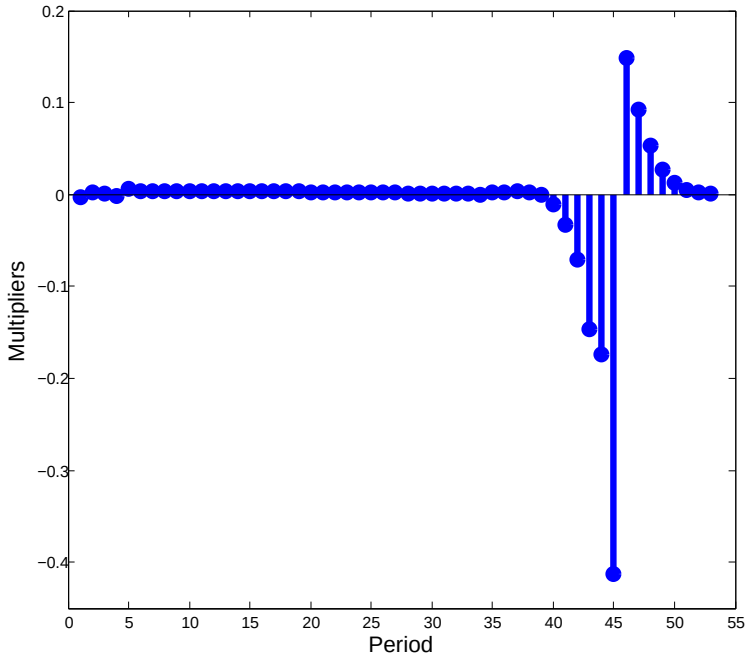


Included Series: Consumption, Inflation.
Multipliers corresponding to time series data of: Inflation.
Series with measurement error: Inflation (t>40).
Last period of consumption available data: 40.

This can be compared to a forecast with data on both consumption and inflation up until period 53. Consumption is imperfectly measured but the inflation data is assumed to be without error. Figure 10.8 plots the multipliers on inflation data in predicting consumption at period 45. It is interesting to see that the multipliers on future inflation data are now positive in predicting current consumption. This is to do with the economic dynamics of the model. Given that the supply shock is important, higher prices now mean lower consumption level now, and so higher prices in the future mean lower consumption expected in the future. Intertemporal consumption dictates that lower expected consumption in the future would raise consumption now and so the multiplier on future inflation data is positive.

Data about the future in the form of forecasts from external sources can be useful but is uncertain. To bring this out, the next experiment examines the role of imperfectly credible announcements of future data outturns. In this experiment, the forecaster is assumed to be in period 40 and trying to forecast consumption a year ahead. Consumption and inflation data only up until period 40 are available and both those series are perfectly measured. Now, information is provided to agents and forecaster alike on what remittances are going to be for the next two years (from periods 41 to 49). Those announcements are not perfectly credible though; there is measurement error. The multipliers reveal how that future imperfect information on an important exogenous variable matter for the forecast of consumption in period 45. This could also be presented by picking out the contribution to consumption growth in period 45 of the remittance data. In more sophisticated experiments, this technology can be used to explain

FIGURE 10.8. Kalman Smoother Multipliers on Inflation Data

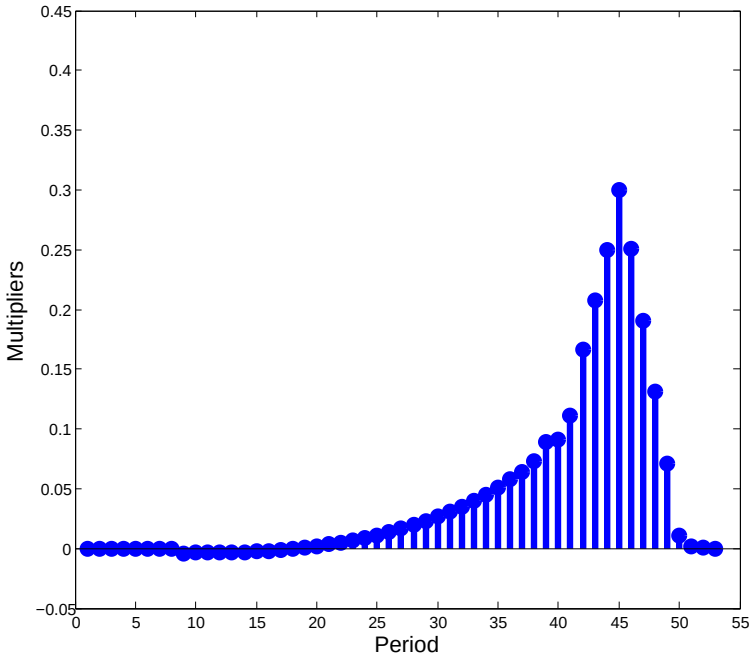


Included Series: Consumption, Inflation.
Multipliers corresponding to time series data of: Inflation.
Series with measurement error: Consumption.
Last period of consumption available data: 53.

the importance of financial market data on the expected future values of asset prices on forecasts of macroeconomic variables.

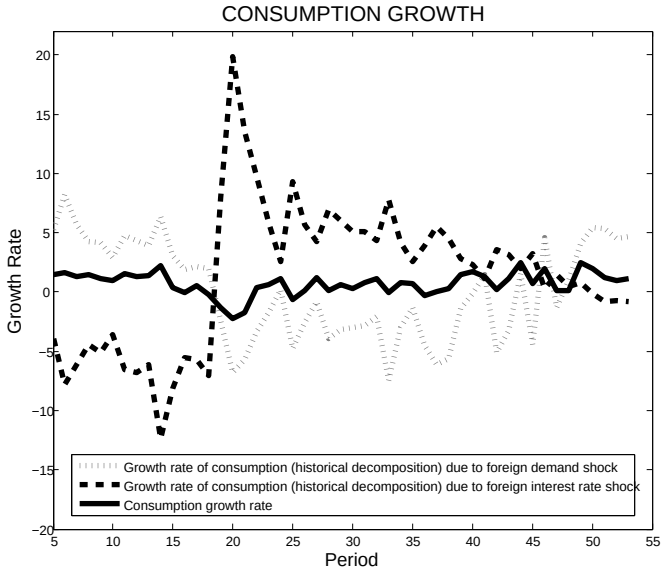
10.4. **Experiment to reveal the identification problem.** A final figure illustrates the identification problems that we discussed in section 9. The figure shows the contribution of two main shocks to consumption growth when we do not allow for measurement error in the CPI and it is assumed that there are only two economic shocks. The model has picked up what it identifies to be a very large offsetting contribution of the two shocks. It does not seem likely that these really have had offsetting effects on consumption, but rather that the model and data cannot separately identify the contribution of each, only their linear combination.

FIGURE 10.9. Kalman Smoother Multipliers on Remittances Data



Included Series: Consumption, Inflation, Remittances.
Multipliers corresponding to time series data of: Remittances.
Remittances 0 to 4-period anticipated shocks
Series with measurement error: Remittances (t>40)
Last period of consumption and Inflation available data: 40.Included

FIGURE 10.10. Economic decomposition of consumption growth



11. Conclusions

The data that is informative for making monetary policy decisions comes in many shapes and sizes and is uncertain. In this paper we propose one possible way of putting that awkward but still useful data set to work in forecasting from a linear dynamic forward-looking model.

There are some important practical advantages to this approach. First it allows theoretical models access to the rich set of real world data that is actually in use. The method also stimulates different ways of presenting and explaining the forecasting. For example we can present a forecast in terms of what data explains the decisions over key variables, and not just what set of shocks causes that forecast. The difference is that as the data is observed, this makes the forecast more transparent. We also show how this method can cope with less than perfectly credible announcements of future information, such as that contained in financial market data. Last but not least, this method has the advantage of separating the data preparation process from the model formulation and solution process. This would better suit how central banks carry out their policy forecasts in practice, given that these two activities are quite specialized.

That said, although we consider this approach to be a step forward, there remain though some important aspects of the policy forecasting problem which have not been taken account of.

First there is evidence, at least anecdotal, that a great part of the forecast error associated with DSGE models is down to drift in economic relations which are assumed to be fixed in the steady state of these models[1]. For example it is often observed that the imports to GDP and exports to GDP ratios lift persistently with greater openness. This data uncertainty method is not designed to deal with such trending economic shocks.

Second we do not discuss how the parameters of the economic model are estimated or calibrated. We discuss alternative procedures for estimating the measurement system that links the data to the unobserved economic model, but we do not do that in much depth. As such we do not allow for parameter estimation effects in this decomposition of the contributions of different data series. Neither have we yet tackled the problem of estimating the whole distribution of these impulse responses and decompositions of the forecasts and not just the mode value, as this would need to take account of parameter estimation uncertainty. We do provide some expressions for the asymptotic conditional uncertainty of the forecast, but without formal treatment of estimation uncertainty we cannot claim that this is a serious presentation of the forecast distribution. Nor do we allow for second-order approximations of the form popularised by Schmitt-Grohe and Uribe (2004). Neither do we allow for asymmetric risks in the forecast which has become common central bank practice following Britton, Fisher, and Whitley (1998). Finally as the set up here is based on a linear model, or more precisely a linearised version of a nonlinear model, this strategy does not apply to nonlinear solution methods, such as that presented in Laxton and Juillard (1996). We hope that ourselves and others will be able to apply some common solutions to these problems within our apparatus.

[1]These errors were still there in the earlier generations of models, but were more easily soaked up by ad hoc residual adjustment, trends and constants that were permitted by those models being looser with theory.

References

Adolfson, M., M. Andersson, J. Lindé, M. Villani, and A. Vredin (2005): "Modern Forecasting Models in Action: Improving Macroeconomic Analyses at Central Banks," Working Paper Series 188, Sveriges Riksbank (Central Bank of Sweden).

Boivin, J., and M. Giannoni (2006): "DSGE Models in a Data-Rich Environment.," Working Paper T0332, NBER.

Boragan Aruoba, S. (2004): "Data Uncertainty in General Equilibrium," Computing in Economics and Finance 2004 131, Society for Computational Economics.

Britton, E., P. Fisher, and J. Whitley (1998): "The Inflation Report Projections: Understanding the Fanchart," Quarterly Bulletin 1, Bank of England.

Burmeister, E., and K. Wall (1982): "Kalman filtering estimation of unobserved rational expectations with an application to the German hyperinflation," *Journal of Econometrics*, 20(2), 255–284.

Canova, F., and L. Sala (2006): "Back to square one: identification issues in DSGE models," Working Paper Series 583, European Central Bank.

Coats, W., D. Laxton, and D. Rose (eds.) (2003): *The Czech National Bank's forecasting and policy analysis system*. The Czech National Bank.

Coenen, G., A. Levin, and V. Wieland (2004): "Data Uncertainty and the Role of Money as an Information Variable for Monetary Policy.," *European Economic Review*, 49(4), 975–1006.

Durbin, J., and S. Koopman (2001): *Time series analysis by state space methods (Oxford Statistical Series)*, vol. 24. Oxford University Press.

Fukac, M., and A. Pagan (2006): "Issues in Adopting DSGE Models for Use in the Policy Process," Working Papers 2006/6, Czech National Bank, Research Department.

Gerali, A., and F. Lippi (2003): "Optimal Control and Filtering in Forward-Looking Economies.," Working Paper 3706, Centre of Economic Policy Research.

González, A., L. Mahadeva, J. Prada, and D. Rodríguez (2008): "Policy Analysis Tool Applied to Colombian Needs," Banco de la República.

Guerron-Quintana, P. (2007): "What You Match Does Matter: The Effects of Data on DSGE Model Estimation.," Working Paper ?, North Carolina State University.

Harvey, A. (1991): *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, Cambridge.

Klein, P. (2000): "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model.," *Journal of Economic Dynamics and Control*, 24(10), 1405–1423.

Koopman, S., and A. Harvey (2003): "Computing Observation Weights for Signal Extraction and Filtering.," *Journal of Economic Dynamics and Control*, 27(7), 1317–1333.

Laxton, D., and M. Juillard (1996): "A Robust and Efficient Method for Solving Nonlinear Rational Expectations Models," IMF Working Papers 96/106, International Monetary Fund.

Lettau, M., and S. Ludvigson (2004): "Understanding Trend and Cycle in Asset Values: Reevaluating the Wealth Effect on Consumption," *American Economic Review*, 94(1), 276–299.

Mahadeva, L., and J. Parra (2008): "Testing a DSGE Model and its Partner Database.," BORRADORES DE ECONOMIA 004507, BANCO DE LA REPÚBLICA.

Orphanides, A. (2001): "Monetary Policy Rules Based on Real-Time Data.," *American Economic Review*, 91(4), 964–984.

Pearlman, J. (1986): "Diverse Information and Rational Expectations Models," *Journal of Economic Dynamics and Control*, 10(1-2), 333–338.

Schmitt-Grohe, S., and M. Uribe (2004): "Solving dynamic general equilibrium models using a second-order approximation to the policy function," *Journal of Economic Dynamics and Control*, 28(4), 755–775.

——— (2008): "What's News in Business Cycles," Working Paper 14215, NBER.

Schorfheide, F., K. Sill, and M. Kryshko (2008): "DSGE model-based forecasting of non-modelled variables," Discussion paper.

Sims, C. A. (2002): "Solving Linear Rational Expectations Models.," *Computational Economics*, 20(1-2), 1–20.

Svensson, L., and M. Woodford (2003): "Indicator Variables for Optimal Policy.," *Journal of Monetary Economics*, 50, 691–720.

——— (2004): "Indicator Variables for Optimal Policy under Asymmetric Information," *Journal of Economic Dynamics and Control*, 28, 661–690.

Swanson, E. (2004): "On Signal Extraction and Non-Certainty Equivalence in Optimal Monetary Policy Rules.," *Macroeconomic Dynamics*, 8(1), 27–50.

Tinsley, P., P. Spindt, and M. Friar (1980): "Indicator and filter attributes of monetary aggregates : A nit-picking case for disaggregation," *Journal of Econometrics*, 14(1), 61–91.

Uhlig, H. (1995): "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily.," Discussion Paper 9597, Tilburg University CentER.

Watson, P. (1983): "Kalman Filtering as an Alternative to Ordinary Least Squares-Some Theoretical Considerations and Empirical Results," *Empirical Economics*, 8(2), 71–85.