Topics in Macroeconomics

Volume 5, *Issue* 1 2005 *Article* 4

Parameter Instability, Model Uncertainty and the Choice of Monetary Policy

Carlo A. Favero* Fabio Milani†

*IGIER, Bocconi University and CEPR, carlo.favero@uni-bocconi.it

†Princeton University, fmilani@uci.edu

Copyright ©2005 by the authors. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher, bepress, which has been given certain exclusive rights by the author. *Topics in Macroeconomics* is produced by The Berkeley Electronic Press (bepress). http://www.bepress.com/bejm

Parameter Instability, Model Uncertainty and the Choice of Monetary Policy*

Carlo A. Favero and Fabio Milani

Abstract

This paper starts from the observation that parameter instability and model uncertainty are relevant problems for the analysis of monetary policy in small macroeconomic models. We propose to deal with these two problems by implementing a novel "thick recursive modelling" approach. At each point in time we estimate all models generated by the combinations of a base-set of k observable regressors for aggregate demand and supply. We compute optimal monetary policies for all possible models and then consider alternative ways of summarizing their distribution. Our main results show that thick recursive modelling delivers optimal policy rates that track the observed policy rates better than the optimal policy rates obtained under a constant parameter specification with no role for model uncertainty.

KEYWORDS: model uncertainty, optimal monetary policy, interest rate smoothing

*We would like to thank David Romer,the editor, two anonymous referees, Mark Salmon and all the participants at the Bank of Finland/ CEPR Annual Workshop on "Heterogeneous Information and Modeling of Monetary Policy" for comments and useful suggestions. Address for correspondence: Carlo A. Favero, IGIER-Bocconi University, Via Salasco 5, 20136 Milan, Italy. E-mail: carlo.favero@uni-bocconi.it. Homepage: http://www.igier.uni-bocconi.it Fabio Milani, Department of Economics, Fisher Hall, Princeton University, Princeton, NJ 08544, USA. E-mail: fmilani@princeton.edu. Homepage: http://www.princeton.edu/∼fmilani.

Parameter Instability, Model Uncertainty and the Choice of Monetary Policy*

Carlo A. Favero[†] and Fabio Milani[‡]
IGIER, Bocconi University and CEPR; Princeton University

Third *revision* January 2005

Abstract

This paper starts from the observation that parameter instability and model uncertainty are relevant problems for the analysis of monetary policy in small macroeconomic models. We propose to deal with these two problems by implementing a novel "thick recursive modelling" approach. At each point in time we estimate all models generated by the combinations of a base-set of k observable regressors for aggregate demand and supply. We compute optimal monetary policies for all possible models and consider alternative ways of summarizing their distribution. Our main results show that thick recursive modelling delivers optimal policy rates that track the observed policy rates better than the optimal policy rates obtained under a constant parameter specification, with no role for model uncertainty.

Keywords: model uncertainty, parameter instability, optimal monetary policy

JEL classification: E44, E52, F41

*We would like to thank David Romer, the editor, two anonymous referees, Mark Salmon and all the participants at the Bank of Finland/ CEPR Annual Workshop on *"Heterogeneous Information and Modeling of Monetary Policy"* for comments and useful suggestions.

[†] *Address:* Via Salasco 3/5, 20136 Milan (Italy). *E-mail:* carlo.favero@uni-bocconi.it.

[‡] *Address:* Department of Economics, Fisher Hall, Princeton, NJ 08544-1021. *E-mail:* fmilani@princeton.edu; *Homepage:* www.princeton.edu/~fmilani.

1 Introduction

Small macroeconomic models have been extensively, and successfully, used in the recent literature for the analysis of monetary policy.[1]

In this paper, we show that the simple standard specifications adopted in these models feature sizeable parameter instability. Moreover, inference and optimal policy are usually conditioned on a single selected specification, which is taken as the true model of the economy. We show that model uncertainty is also an important feature of the data. We use the standard specification for aggregate demand and supply in a small macroeconomic model but we also consider all possible models generated by different specifications of distributed lags. We find that the distance among alternative models, as measured by a within-sample selection criterion, is small, and that the ranking of models according to some within-sample criterion does not match that obtained by using an out-of-sample forecasting performance criterion.

We propose an approach based on "thick recursive modelling" to deal with both these problems.

We mimic the decision of a monetary policy maker who sets policy rates on the basis of the available data. To this end, at each point in time t we search over a base set of k observable regressors to construct a small structural model of the economy. In each period we estimate a set of regressions spanned by all the possible combinations of the k regressors. We estimate our system equation by equation and we keep the number of regressors k constant for all equations. We estimate our equations by a rolling method, using a fixed window of twenty-two years of quarterly data.[2]

Our econometric procedure delivers 2^k models for aggregate demand and supply at any point in time therefore the choice of monetary policy requires one to take a stand on model, or specification, uncertainty.

A traditional approach taken in the literature is to proceed by specifying a selection criterion to choose the best model in each period. Inference is then conditional on this preferred model. We follow Granger and Jeon (2004) and label this approach 'thin' modelling, in that the optimal monetary policy is described over time by a thin line.

Thin modelling needs to be based on a selection criterion that weights goodness of fit against parsimony of the specification. The literature typically considers Schwarz's Bayesian Information Criterion (BIC), AKAIKE, and adjusted R^2 as selection criteria.

The advantage of this approach is that a potentially non-linear process is modelled by applying recursively a selection procedure among linear models.

[1] See, for example, Rudebusch and Svensson (2002), Sack (2000), and Clarida, Galì and Gertler (2000).

[2] An alternative would be to proceed to a series of recursive regressions, by extending the sample over time. Rolling estimation, however, allows us to better account for potential structural breaks.

The procedure mimics a situation in which the specifications of aggregate demand and supply are chosen in each period from a pool of potentially relevant regressors.

The main limit of thin modelling is that model uncertainty is not considered. In each period the information coming from all the discarded models is ignored for the design of optimal monetary policy. The explicit consideration of estimation risks naturally generates what we define as ‘thick’ modelling, where optimal monetary policy is described by a thick line to take account of the multiplicity of estimated models. The thickness of the line that describes the evolution over time of any estimated parameters across all different models is a direct reflection of the estimation risk. Given the range of all optimal monetary policies, we consider their average (or weighted average) to evaluate comparatively the behavior of policy rates implied by thin and thick modelling. Although the averaging of policies does not arise as a solution to an optimization problem, our approach permits us to account for model uncertainty in a very tractable way. This method constitutes a simple approximation of the optimal solution to the problem of setting policy under model uncertainty. A Bayesian policy maker would in fact rationally choose to weight policies according to the relative posterior model probabilities. Our approach, weighting all policies equally or according to the models' fit, avoids the need of specifying priors over the whole model space. Given the lags with which monetary policy affects the economy, optimal monetary policy must be based on some forecasting model. Forecast combinations have a proven track record and have been often found to produce better forecasts than methods based on the “best” individual model. Moreover, as discussed by Timmermann (2004), simple combinations that ignore correlations between forecast errors often dominate more refined combination schemes aimed at estimating the theoretically optimal combination weights. Timmermann (2004) reviews the possible theoretical factors underlying the empirical success of simple forecast combinations. Some of the cited factors are related to model misspecification, instability and estimation error when the number of models is large relative to the available sample size. Although our paper is not about forecasting, those factors remain relevant to our purposes.

Our empirical results emphasize the importance of uncertainty for policy. The implied optimal monetary policies in a given period show considerable variation over the whole range of potential models of the economy. This evidence highlights the necessity of a method that enables one to take the estimation risk into account when setting policy. Importantly, our exercise should be taken as an exploratory one: we consider a very specific class of models, which are backward-looking, and, as already stated, our approach is not an optimizing one for dealing with model uncertainty.

The paper is structured in four sections. The first section discusses the relevance of parameter instability and model uncertainty in small macroeconomic models of the monetary transmission mechanism. The second section illustrates the differences in the calculation of optimal monetary policy when thin modelling, recursive thin modelling and recursive thick modelling are adopted. The third sections contains the empirical results for the U.S. case.

The last section concludes.

2 Parameter Instability and Model Uncertainty in Small Structural Models

Recent studies of optimal monetary policy in closed economies have adopted a simple two-equation framework. A strand of the literature uses small forward-looking models derived from microfounded behavior of economic agents. Another strand, more interested in the empirical fit of the models, employs simple backward-looking specifications.

A typical model in this latter class is the one estimated by Rudebusch and Svensson (2002), who employ the following representation of aggregate supply and demand of the economy:

$$\begin{aligned} M^{AS} &: \quad \pi_t = \beta_0 + \beta_1 \pi_{t-1} + \boldsymbol{\beta}' X_t^1 + u_t^1, && (1) \\ M^{AD} &: \quad y_t = \gamma_0 + \gamma_1 y_{t-1} + \boldsymbol{\gamma}' X_t^2 + u_t^2, && (2) \\ X_t^{1\prime} &= \begin{bmatrix} \pi_{t-2} & \pi_{t-3} & \pi_{t-4} & y_{t-1} \end{bmatrix}, \\ X_t^{2\prime} &= \begin{bmatrix} y_{t-2} & i_{t-1} - \pi_{t-1} \end{bmatrix}. \end{aligned}$$

The authors estimate the equations using quarterly data over the sample 1961:1 to 1996:4. Inflation, denoted by π_t, is calculated as $100 * (\log(p_t) - \log(p_{t-4}))$ where p_t is the GDP implicit price deflator, the output gap y_t is obtained as $100*(\log(Q_t) - \log(Q_t^*))$ where Q_t denotes actual GDP (in chained 1996 dollars), Q_t^* denotes potential GDP, as measured by the Congress and made publicly available via the Federal Reserve of St. Louis's website, and i_t represents the federal funds rate.

This small structure constitutes the set of constraints under which the reaction function of the central bank is derived by minimizing an intertemporal loss function. Optimal setting of interest rates delivers in general a functional specification resembling a Taylor rule. The parameters in the central bank's reaction function are convolutions of the parameters in the structure of the economy and of the parameters describing the preferences of the monetary policy maker. Hence, joint estimation of the simple structure of the economy and the interest rate settings equation allows one to evaluate which set of central bank preferences delivers a path for policy rates closest to that observed in the data.

Under this simple representation of the economy, estimation and policy advice are conditional on a single model being the correct model of the economy. Also the model parameters are assumed constant. We shall refer to this approach as **"thin" modelling**. We use this simple representation to illustrate the importance of parameter instability and model uncertainty for the determination of optimal monetary policy.

2.1 Parameter Instability

We start by replicating Rudebusch and Svensson's results using quarterly data over the period 1961:1-2000:3. Our estimated equations are as follows:[3]

$$\pi_{t+1} = \underset{(0.080)}{0.63}\,\pi_t + \underset{(0.093)}{0.005}\pi_{t-1} + \underset{(0.094)}{0.21}\,\pi_{t-2} + 0.15\pi_{t-3} + \underset{(0.033)}{0.14}\,y_t + \widehat{u}_{1,t+1}, \quad (3)$$

$$y_{t+1} = \underset{(0.075)}{1.24}\,y_t - \underset{(0.075)}{0.31}\,y_{t-1} - \underset{(0.026)}{0.06}\,(i_t - \pi_t) + \widehat{u}_{2,t+1}. \quad (4)$$

To evaluate the potential parameter instability we re-estimate the system by considering two sub-samples. The first sub-sample runs from 1961:1 to 1982:4. Estimation yields:

$$\pi_{t+1} = \underset{(0.107)}{0.70}\,\pi_t - \underset{(0.129)}{0.013}\pi_{t-1} + \underset{(0.13)}{0.18}\pi_{t-2} + 0.14\pi_{t-3} + \underset{(0.045)}{0.16}\,y_t + \widehat{u}_{1,t+1}, \quad (5)$$

$$y_{t+1} = \underset{(0.101)}{1.17}\,y_t - \underset{(0.104)}{0.24}\,y_{t-1} - \underset{(0.038)}{0.106}(i_t - \pi_t) + \widehat{u}_{2,t+1}. \quad (6)$$

By concentrating instead on the second sub-sample 1983:1-2000:2, we obtain:

$$\pi_{t+1} = \underset{(0.12)}{0.35}\pi_t + \underset{(0.12)}{0.05}\pi_{t-1} + \underset{(0.123)}{0.32}\,\pi_{t-2} + 0.27\pi_{t-3} + \underset{(0.044)}{0.11}\,y_t + \widehat{u}_{1,t+1}, \quad (7)$$

$$y_{t+1} = \underset{(0.115)}{1.29}\,y_t - \underset{(0.108)}{0.38}\,y_{t-1} + \underset{(0.03)}{0.027}(i_t - \pi_t) + \widehat{u}_{2,t+1}. \quad (8)$$

We take these results as an indication of parameter instability of economic relevance. Consider inflation persistence and the effect of monetary policy on the output gap, two crucial parameters for the design of optimal monetary policy. Although the sum of the coefficients on the lagged dependent variables in the supply equation is restricted to one in all sub-samples, the weight on the first two lags decreases across periods. Similarly, the effect of real interest rates on the output gap in the aggregate demand equation features an important shift: it is significantly negative in the first sub-sample, with a sizeable long-run effect of about one, and it becomes insignificant (and slightly positive) in the second sub-sample. We perform a Chow test of the null of parameter stability on the two equations to investigate a potential breakpoint at date 1982:4. We reject the hypothesis of no breakpoint at the 5% significance level for the aggregate demand equation; the evidence in against stability of the inflation equation is weaker.

Recently, Pesaran and Timmermann (1995) have proposed recursive modelling as an appropriate approach to deal with parameter instability and non-linearity in the context of small models. Their paper originally proposed recursive modelling for forecasting, but here we explore its application for

[3] All the specifications for the supply equation impose the restriction that the coefficients on the lags of the dependent variable add up to unity.

the analysis of optimal monetary policy. Consider a monetary policy maker who believes that demand and supply equations can be modelled by projecting output and inflation on macroeconomic indicators but does not know the "true" form of the underlying specification and the "true" parameter values. To keep the macro structure simple and comparable to that of Rudebusch and Svensson, we consider a situation in which there is uncertainty only on the specification of the lags with which the relevant variables enter the supply and demand equations. The best option for the policy maker is to search for a suitable model specification among the set of models believed a priori appropriate to describe supply and demand. As time elapses, in the presence of potential parameter instability, such a specification might change, with different variables entering the two equations or the same variables entering with different coefficients. A policy maker with no strong a priori belief on the specification of lags in the demand and supply equations would probably like to update the econometric model to base monetary policy on the best possible representation of the unknown Data Generating Process. Therefore, at each point in time t, the policy maker searches over a base set of k regressors to obtain the best possible specification for output and inflation based on the information available at that time. Recursive modelling mimics such a decision process by assuming that the policy maker estimates at each point in time the entire set of regression models spanned by all the possible permutations of the k regressors and chooses the best model, according to some statistical criteria, to define optimal monetary policy. Hence, in each period, the decision is based on the best specification for inflation and output, out of 2^k models for each variable. Given that the variables and parameters entering the best chosen specification are allowed to vary over time, recursive modelling is capable of accommodating parameter instability and non-linearity in the effect of some factors on output and inflation.

In practice, we implement recursive modelling by considering the following specifications for aggregate demand and supply:

$$M_{i,t}^{AS} \;:\; \pi_t = \beta_0 + \beta_1 \pi_{t-1} + \boldsymbol{\beta}_i' X_{t,i}^1 + u_{t,i}^1, \tag{9}$$

$$M_{i,t}^{AD} \;:\; y_t = \gamma_0 + \gamma_1 y_{t-1} + \boldsymbol{\gamma}_i' X_{t,i}^2 + u_{t,i}^2, \tag{10}$$

where $X_{t,i}^1, X_{t,i}^2$ are $(k_i \times 1)$ vectors of regressors under models $M_{i,t}^{AS}$, $M_{i,t}^{AD}$, obtained as a subset of the base set of regressors X_t^1, X_t^2

$$\begin{aligned} X_t^{1\prime} &= \begin{bmatrix} \pi_{t-2} & \pi_{t-3} & \pi_{t-4} & y_t & y_{t-1} & y_{t-2} & y_{t-3} & y_{t-4} \end{bmatrix} \\ X_t^{2\prime} &= \begin{bmatrix} y_{t-2} & y_{t-3} & y_{t-4} & rr_{t-1} & rr_{t-2} & rr_{t-3} & rr_{t-4} & rr_{t-5} \end{bmatrix} \end{aligned}$$

where $k_i = e' v_i$, e is a $(k \times 1)$ vector of ones, and v_i is a $(k \times 1)$ selection vector, composed of zeros and ones, where a one in its j-th element means that the j-th regressor is included in the model. All variables are defined as above and $rr_t = i_t - \pi_t$. The constant and the lagged dependent variable are always included in all specifications. Uncertainty on the specification of lags implies that the policy maker searches over $2^8 = 256$ specifications to select in each period the relevant demand and supply equations. The selection is based on traditional criteria such as adjusted R^2, Akaike Information Criterion, or Schwarz's Bayesian Information Criterion. By considering uncertainty only about the dynamic structure of the economy, we are certainly underestimating the true level of uncertainty faced by policy makers. Nevertheless, our choice still permits one to account for some important uncertainty in the policy maker's problem, such as the number of lags with which policy affects the economy. Since policy affects the economy only with some delay, the policy maker needs to forecast macroeconomic conditions in the future. The extent of such delayed effects on the economy determines how forward-looking optimal monetary policy has to be.

Other important sources of model uncertainty are omitted from our analysis. For example, it would be interesting to consider the uncertainty about the backward or forward-looking nature of the model. In fact this is a matter of current debate. The difficulty one identifying these alternative specifications and the complications generated by the explicit consideration of a forward-looking structure in our framework, however, lead us to refrain from considering such additional sources of uncertainty.

Following Granger and Jeon (2004), we label **recursive "thin" modelling** the approach that, after estimating all possible models, selects a single model in each period to design optimal policy. In this case, optimal monetary policy will be described over time by a thin line. We shall instead label as **recursive "thick" modelling** the approach that retains and exploits the information coming from the whole set of estimated models.

2.2 Model Uncertainty

The main limitation of thin modelling is that model, or specification, uncertainty is not considered. In each period, both inference and optimal policy remain conditional on the preferred model being the true model of the economy. The information coming from the discarded $\left(2^k - 1\right) * 2$ models for aggregate demand and supply is ignored for the determination of optimal monetary policy.

This is a problematic simplification of the 'thin' modelling approach.

First, the distance among models, as measured by the chosen selection criterion, is small. Moreover, the ranking of models according to a within-sample performance criterion does not match that obtained by using an out-of-sample forecasting performance criterion. Figures 1 and 2 make this point by showing the cross-plot of the respective model ranking according to the adjusted R^2 and to the Theil's U for the 256 models of aggregate demand and supply using estimation based on the last window of observations in our sample.[4]

[4] For brevity, we report the plot using only the last sample period. The picture emerging from all other sample points is very similar.

Clearly, the ranking of models according to the adjusted R^2 is not only different but also little correlated with the ranking of models based on the Theil's U. For example, the top ranked model using the in-sample criterion would rank according to the out-of-sample criterion around 100th for the demand equation and close to last for the supply equation. Considering the lags with which the policy instrument affects output gap and inflation, optimal monetary policy has to be based on forecasts for the relevant variables: it is therefore not clear at all that the best model selected by the adjusted R^2 is the most appropriate for the design of monetary policy. The first two figures show how hard it is to decide among different models of demand and supply. To evaluate the importance of this choice we need to measure the potential relevance of model uncertainty. To this aim, we consider the distribution of some key parameters in our small structural model across the different models of aggregate demand and supply.

Figure 3 reports the distribution across models and time of the two crucial parameters that determine the effect of monetary policy on inflation in our model economy: the effect of the output gap on inflation (top) and the effect of real interest rates on the output gap (bottom).

We consider long-run effects. The figure shows in the panel on the left the variation of the coefficients across models (for the last window of observations) and in the panel on the right the variation across time (considering in each period just the best model). The plots show substantial variation over both dimensions. Notice that a policy maker who adopts the recursive thin modelling strategy using the last window of observations would measure the impact of an interest rate move on real activity as equal to -0.113 and the impact of output gap on inflation as equal to 0.054.[5] Suppose instead that the policy maker is concerned about model uncertainty. By considering the whole set of possible models, the policy maker would obtain estimates of the real interest rate coefficient ranging from -0.164 to -0.015 (excluding the cases where the effect is 0). Similarly, the estimated slope coefficient in the supply equation would range from 0.01 to 0.09. Such differences across the model spectrum might imply important consequences for policy.

A natural way to interpret model uncertainty is to refrain from the assumption of the existence of a "true" model and attach instead probabilities to different possible models. This approach has been labelled 'Bayesian Model Averaging'; see for example Hoeting et al. (1999), and Raftery et al. (1997).

The main difficulty with the application of Bayesian Model Averaging to problems like ours lies with the specification of prior distributions for all the $2 * 2^k$ parameters in the equations. The results reported in Figures 1 and 2 show clearly that the ranking of models in terms of their within sample performance does not match the ranking of models in terms of their out-of-

[5] These point estimates correspond to the coefficient for model 1 from the left-hand side plots.

sample forecasting performance. In face of the risk involved in choosing a weighting scheme, we opt for the selection method proposed by Granger and Jeon (2004) of using a '... *procedure [which] emphasizes the purpose of the task at hand rather than just using a simple statistical pooling...'*. Therefore, we derive the optimal monetary policy associated with each specification for the simple aggregate demand-supply system and we then consider the average monetary policy obtained by giving equal weight to each alternative monetary policy. As an alternative, we also weight policies according to the model probabilities, calculated as a function of the likelihood corrected for degrees of freedom. This latter choice comes closer to Bayesian alternatives but it does not require the specification of priors for all the different models.

3 Optimal Monetary Policy

To assess the impact of recursive thick modelling, we calculate the optimal federal funds rate paths based on the following model choices:

- **Thin modelling**: the Rudebusch and Svensson model.
- **Recursive thin modelling**: the model with the best adjusted R^2.
- **Recursive thin modelling**: the best forecasting model (lowest Theil U).
- **Recursive thick modelling**: the average (simple or weighted) optimal monetary policy.

The central bank minimizes an intertemporal loss function of the form:

$$E_t \left\{ \sum_{\tau=0}^{\infty} \phi^{\tau} \left[\lambda_\pi \pi^2_{t+\tau} + \lambda_y y^2_{t+\tau} + \lambda_R (i_{t+\tau} - i_{t+\tau-1})^2 \right] \mid \mathcal{M} \right\}, \qquad (11)$$

where ϕ is the discount factor, E_t is the usual expectation operator, and $\mathcal{M}$ represents the set of possible models; the period loss function is quadratic in the deviations of output and inflation from their targets, and it includes a penalty for the policy instrument's volatility. The parameters λ_π, λ_y, and λ_R represent the relative weights of inflation stabilization, output gap stabilization, and interest rate smoothing, and are assumed to sum to 1.

We shall solve the optimization problem under different assumptions over preferences to evaluate which weighting scheme delivers the best performance in replicating the observed data.

More in detail, we calculate the optimal monetary policy rules implied by the four described approaches, under five alternative specifications for preferences:

- CASE 1. **Pure (strict) inflation targeting**: $\lambda_\pi = 1$, $\lambda_y = 0$, $\lambda_r = 0$.
- CASE 2. **Pure inflation targeting with interest rate smoothing (strong)**: $\lambda_\pi = 0.7$, $\lambda_y = 0$, $\lambda_r = 0.3$.
- CASE 3. **Flexible inflation targeting**: $\lambda_\pi = 0.5$, $\lambda_y = 0.5$, $\lambda_r = 0$.
- CASE 4. **Flexible inflation targeting with interest rate smoothing**: $\lambda_\pi = 0.4$, $\lambda_y = 0.4$, $\lambda_r = 0.2$.
- CASE 5. **Pure inflation targeting with interest rate smoothing (weak)**: $\lambda_\pi = 0.9$, $\lambda_y = 0$, $\lambda_r = 0.1$.

Before going into the details of each case, two problems relevant to the implementation of recursive modelling are worth some discussion. First, there are specifications in which the question of optimal monetary policy is not worth addressing because monetary policy has no effect on target variables. We have consequently excluded all specifications that feature a zero effect of interest rates on the output gap and/or a zero effect of the output gap on inflation. Second, thick modelling delivers 256 specifications for aggregate demand and 256 specifications for aggregate supply. When demand and supply are combined in a model the curse of dimensionality is relevant and the total number of possible models becomes $256^2 = 65\,536$. To keep the number of models limited we ordered specifications for aggregate demand and supply in terms of their performance to build models by considering aggregate demand and supply equations with the same position in their respective ranking. This strategy led us to consider 256 models. Although admittedly ad hoc, this choice still permits us to show the large amount of uncertainty associated with the policy maker's decision problem.

3.1 Thin Modelling

Under thin modelling the optimization problem is solved subject to the dynamic structure of the economy, which is given by the constant parameter specifications of aggregate demand and supply adopted by Rudebusch and Svensson, as in (1) and (2). The parameters are estimated on the whole available sample.

As shown in the Appendix, the optimal policy rule computed by re-writing the model in state-space form and solving the relevant optimal control problem is given by

$$i_t = f \left[\begin{array}{cccc} \pi_{t-1} & y_{t-1} & X_t^{1\prime} & X_t^{2\prime} \end{array} \right],$$

where f is the optimal feedback vector, which depends both on the parameters describing the preferences of the central bank and on the parameters describing the stochastic difference equations for aggregate demand and supply.

The only uncertainty affecting the economy consists of additive disturbances entering the model equations. The certainty-equivalence principle holds: additive uncertainty has no effect on the optimal rule.

The policy maker knows the "true" model characterizing the economy. Neither parameter uncertainty nor model uncertainty are relevant within this framework.

3.2 Recursive Thin Modelling

Recursive thin modelling implies that the policy maker investigates more deeply the constraints under which optimal policy is designed. At any point in time, all possible models are estimated and the best model, according to some criterion, is chosen. The process is iterated, allowing for modifications in the specification of demand and supply as new information accrues. In practice, estimation is based on a rolling window of fixed length.

Rolling regressions might induce bias in the coefficients because of structural breaks. However, we have chosen to implement this procedure as it permits us to mimic the situation of a policy maker who obtains data in real time and learns slowly about the break.

Recursive modelling is implemented by considering specifications for aggregate demand and supply given by (9) and (10).

All estimated models are then ranked in accordance to a selection criterion and the best model is chosen. In light of the evidence on the differences in ranking of models when within-sample or out-of-sample performances are considered, we rank models using both the adjusted R^2 and Theil's U as selection criteria.

Table 1 reports the percentages of inclusion of the different variables in the evolving best model of the economy, ranked according to the adjusted R^2 and Theil U.

From the table it is apparent that the variables selected to belong to the best model vary over time. Again, important differences do exist among the best specifications according to the in-sample and out-of-sample criteria.

The optimal policy is then derived conditional on the chosen model and it takes the form

$$i_t = f_t \left[\begin{array}{cccc} \pi_{t-1} & y_{t-1} & X_{t,i}^{1\prime} & X_{t,i}^{2\prime} \end{array} \right].$$

The optimal rule is time-varying along two dimensions: the size of the coefficients and the set of variables to which monetary policy responds.

Note that optimization is performed in every period as if the parameters would remain constant in the future, whereas they will be instead updated in the following periods. In this respect, the behavior of the policy maker is sub-optimal. Optimality would require in fact active experimentation by the policy maker to learn faster the structure of the economy. Active experimentation is not explored in the present paper.

3.3 Recursive Thick Modelling

So far optimal monetary policy has been designed at each sample point by estimating all the possible models but by optimizing just once, taking the best model as the relevant constraint. As we have discussed, this procedure does not retain information from all non-selected models.

To implement thick modelling, we consider a situation in which the central banker not only estimates all possible models but also derives all the associated optimal monetary policies. Then the adopted monetary policy explicitly considers the distribution of all optimal policies associated with each different model of the economy.

In order to interpret the observed monetary policy rates as the outcome of thick modelling we consider three alternative methods for mapping in each period the distribution of optimal policies into a single rate. Our first choice is a simple average of the different optimal rates obtained under each different model :

$$\begin{aligned} i_t^* &= \frac{1}{n}\sum_{j=1}^{n} i_t^j, \\ i_t^j &= f_t^j \left[\begin{array}{cccc} \pi_{t-1} & y_{t-1} & X_{t,j}^{1\prime} & X_{t,j}^{2\prime} \end{array} \right]. \end{aligned}$$

In our framework uncertainty concerns only the dynamic structure of the economy, i.e. which lags of the considered variables are important. This implies that the different models are highly similar with each other and also the capacity of the policy maker to assign different weights to them is limited. Therefore, the assumption that all models and all policies are a priori equally likely seems realistic. Even if the models are apparently similar, they might still lead to important differences in the implied policies (and they actually will, as we show in the next section).

We explore, however, two alternative solutions that correspond to weighting policies according to some measures of accuracy of the underlying models.

The first alternative consists of weighting optimal policies according to the value of Schwarz's Bayesian Information Criterion. The derived optimal policy is a weighted average of all policies, with a higher weight attached to rules derived from models with a lower BIC.

The second alternative weights policy rules according to the approach proposed by Doppelhofer, Miller, and Sala-i-Martin (2004) and called Bayesian Averaging of Classical Estimates (BACE). This approach combines a Bayesian concept, the averaging across different models, with classical OLS estimation; this method can be derived as a limiting case of standard Bayesian estimation when prior information becomes dominated by the data.

Within this approach, the alternative policy rules are weighted according to the following quantity, derived for all models:

$$P\left(M_j|D\right) = \frac{P(M_j)T^{-k_j/2}SSE_j^{-T/2}}{\sum_{j=1}^{2^k} P(M_j)T^{-k_j/2}SSE_j^{-T/2}}, \tag{12}$$

where $P\left(M_j|D\right)$ represents the posterior probability of the model M_j given the data D, $P(M_j)$ is the prior probability of model M_j, which is set to $1/\Sigma M_j$, T is number of observations, k_j is the number of included regressors in model M_j, and SSE_j is the corresponding sum of squared errors from classical OLS estimation.

This latter alternative comes closer to the Bayesian optimal decision problem. The Bayesian policy maker would in fact minimize the expected loss function weighting policies according to their posterior model probabilities. The difference here is that priors about model parameters are not used.

Our effort to account for model uncertainty differs from two common solutions adopted in the literature: the insertion of multiplicative (parameter) uncertainty and the use of robust control techniques.

The existence of multiplicative uncertainty, as first shown by Brainard (1967), implies that the optimal policy rule is also affected by the variances of the estimated parameters, not only by their first moments. Uncertainty about the model parameters typically causes attenuation of the central bank's optimal response.[6]

Robust control (see for example Onatski and Stock 2002) assumes instead that the policy maker plays a game against a malevolent Nature and tries to minimize the maximum possible loss (minimize the loss in the worst-case scenario), whereas his opponent, Nature, tries to maximize his loss.

Onatski and Stock assume that there exists a reference model M, which is known to be an approximation of the true model of the economy, with an unknown deviation Δ, belonging to the set of perturbations D, from the true model. Being K the set of policy rules and $R(K, M + \Delta)$ the risk of policy K when the real model is $M + \Delta$, the robust control problem is given by:

$$\min_{\{K\}} \sup_{\Delta \in D} R\left(K, M + \Delta\right).$$

The robust solution to this problem is very different from the solution under multiplicative uncertainty. Here uncertainty might induce the policy maker to favor a more aggressive policy than in the certainty-equivalent state, in order to minimize the welfare loss in the worst case alternative.

We focus on model uncertainty in its simplest possible form, i.e.uncertainty about the specification of the relevant dynamics. Admittedly our treatment of model uncertainty does not arise as a solution of an optimization problem under uncertainty. Nonetheless, our exploratory exercise might offer some useful indication concerning the choice of monetary policy.

[6]For recent revisitations of this result, see Sack (2000) and Söderström (1999, 2002).

4 Empirical Results

Our empirical results are summarized in Table 2. As described above, we consider five possible parameterization of the loss function and four modelling strategies: thin modelling (adopting the parameterization in Rudebusch and Svensson), recursive thin modelling using a within-sample performance selection criterion (best adjusted R^2), recursive thin modeling using an out-of-sample performance selection criterion (Theil's U), and recursive thick modeling. In the latter case, we report results for three different alternatives: the simple average optimal policy across all possible models, and the two weighted averages with weights calculated according to models' BIC and BACE values.

We end up with 30 optimal federal funds rate series to be compared with the observed one. Table 2 reports the first two moments of the optimal and actual series.

The results clearly show that observed monetary policy is nowhere near the optimal monetary policy when no positive weight is attached to interest rate smoothing in the loss function of the policy maker.

When we allow a strictly positive weight to interest rate smoothing, some optimal policy rate series feature first two moments comparable to those displayed by the actual policy rates.

In particular, with preference weights $\lambda_\pi = 0.7, \lambda_y = 0, \lambda_r = 0.3$, the optimal interest rates obtained under recursive thin modelling and recursive thick modelling have means and standard deviations quite close to those of the actual series.

Also in the case $\lambda_\pi = 0.4, \lambda_y = 0.4, \lambda_r = 0.2$, we find at least two series that can replicate quite well the first two moments of the actual policy rates: the one obtained with recursive thin modelling (and adjusted R^2 as selection criterion), and the one derived under recursive thick modelling, weighting policies according to the BACE approach (the other policies under recursive modelling are also not very far).

When the weights are $\lambda_\pi = 0.9, \lambda_y = 0, \lambda_r = 0.1$, i.e. the weak interest rate smoothing case, no series come very close to the observed one.

Interestingly, the optimal policy series obtained in the context of thin modelling (à la Rudebusch-Svensson) are never similar to the actual policy series. In that framework, a considerably higher preference for interest rate smoothing in the loss function would be necessary.

We see that the series generated by recursive thin and thick modelling always feature a lower volatility than that of the series generated by thin modelling.

An important characteristic to analyze is also the "thickness" of our derived optimal policy choices: the "thickness" of the line gives a clear indication of the amount of model uncertainty surrounding the policy maker's decision problem.

In order to measure this uncertainty, we focus on the width of optimal interest rates set in every period for the different models. We are also interested in examining whether this width is constant or time-varying, in order to understand the evolution of uncertainty over time.

An evaluation of these aspects can be based on Figures 4 and 5. Figure 4 shows the thickness of monetary policy choices by plotting all the optimal interest rate series obtained under each model together with the selected average policy (the darker line). The results are reported for the five different cases for preferences.

First, we see that cases 1 and 3, which represent the alternatives without smoothing, imply unrealistically high and volatile optimal series.

Concentrating instead on the other more realistic preferences, the figures show a large amount of model uncertainty. As an example, in a representative period when the average policy would imply a rate around 12%, the thickness of the line is such that different models would suggest policy rates as high as 18% or as low as 5% (see the uncertainty at the end of the sample of these three cases, for example).

This level of uncertainty is likely to underestimate the true uncertainty, as we are considering only uncertainty about the dynamic structure of the economy, within a very specific class of models.

Notice that the alternative use of recursive (instead of rolling) regressions would reduce the variability of models and parameters over time. As a result, the thickness of the set of monetary policies in the second half of the sample would probably decrease. However, by its nature, recursive estimation would underestimate parameter instability in the presence of structural breaks toward the end of the sample.

In the optimization we do not include a zero lower bound on interest rates. The inclusion of the zero bound would also reduce monetary policy thickness under some central bank preferences. The difference is likely to be small, however, since the zero bound constraint is found to be rarely binding in our optimizations.

It is also apparent that the width of the band of optimal interest rates varies over time. The time variation can be more easily assessed from figure 5, where we have plotted the band width dynamics over the sample. The dynamics, for the three more realistic cases with smoothing (cases 2, 4 and 5), indicate an almost U-shaped pattern. In fact, the width seems very large during the beginning and the end of the sample, while it is lower (indicating less uncertainty) in the central part of our sample.

It might also be useful to understand if the alternative policies are uniformly distributed towards the average or are skewed in some ways. Some skewness towards high values is evident for all cases without interest rates smoothing. Focusing instead on the cases with smoothing, we notice a strong skewness toward low values: the plots show that very low values (much lower than the average) are more likely than very high ones. The average policy is in fact closer to the maximum than to the minimum possible policy.

Finally, we turn to another important characteristic of interest rates: "smoothness", i.e. the persistence of derived policy rates. As we have seen, modelling actual policy rates seems to require the direct inclusion of interest rate smoothing in central banks' preferences as an ad hoc solution. In a recent survey of the empirical literature, Sack and Wieland (2000) have discussed three main motives for interest rate smoothing that do not require the direct inclusion of volatility of interest rates in the loss functions of the monetary policy makers.

The first motive is forward-looking expectations. In models with forward-looking expectations, estimated policy rules with inertia are more effective in stabilizing output and inflation for a given level of volatility of the policy instrument. If policy features a high degree of partial adjustment, then forward-looking market participants will expect an initial policy move to be followed by additional moves in the same direction. Such an effect on expectations increases the impact of policy on output and inflation. Smoothing is then induced by the structure of the economy and there is no need to include some cost in the preferences of central banks to generate the observed behavior of interest rates.

The second motive is data-uncertainty. According to this motive a moderate responsiveness of interest rates to initial data releases is optimal when the data are measured with errors. In fact, an aggressive policy response would induce unnecessary fluctuations in policy rates resulting in unintended fluctuations in output and inflation.

The third motive is uncertainty about the parameters. This is a revisiting of the classical argument offered by Brainard (1967). When policy makers are uncertain about the key parameters that determine the transmission of monetary policy in the adopted structural model of the economy, aggressive policy moves would be more likely to have unpredictable consequences on output and inflation. Gradual policy is then optimal to minimize fluctuations of output and inflation around their targets.

Rudebusch (2002) pushes the argument even further to label monetary policy inertia as an illusion, reflecting the episodic unforecastable persistent shocks that central banks face. His views are supported by the empirical evidence from the term structure of interest rates, which does not indicate the large amount of forecastable variation in interest rates at horizons of more than three months that monetary policy inertia would imply.

It is then interesting to evaluate the role of the omitted consideration of model uncertainty and parameter instability as the potential sources of the observed persistence in interest rates. We have reported the relevant results in Table 3, which shows the estimated AR(1) coefficient in a regression of the policy rate on a constant and its lagged value.

The considerable observed persistence of the actual series is well replicated only by optimal policy rates derived allowing explicitly for a smoothing motive in the loss function. Moving from thin modelling to recursive thin and thick modelling does not generate a systematic change in the persistence parameter.

5 Conclusions

This paper starts from the observation that parameter instability and model uncertainty are very relevant in the specification of the constraints under which the monetary policy maker operates. We have implemented "thick recursive modelling" to simultaneously deal with the two problems.

At any point in time we mimicked the decision of a monetary policy maker who sets policy rates on the basis of the available data. We have assumed that in each period t the policy maker searches over a set of observable k regressors to construct a small structural model of the economy. The policy maker then estimates the whole set of regressions spanned by all the possible permutations of the k regressors.

This econometric procedure delivers 2^k models describing aggregate demand and supply at any point in time, making clear the need to account for model uncertainty. To incorporate model uncertainty in the optimal policy decision, we proceed by computing all the optimal monetary policies implied by each single model, and by then taking their average or weighted average as our benchmark optimal monetary policy.

We then compare the observed policy rates with the optimal rates generated by the traditional "thin modelling" approach and by our proposed "thick modelling" approach.

Our results confirm the difficulty of recovering the deep parameters describing the preferences of the monetary policy makers from their observed behavior. This is because optimal policy depends both on the parameters describing the preferences of the policy maker and on those defining the structure of the economy. Model uncertainty and parameter instability imply very low precision in the estimation of the structure of the economy and therefore the observational equivalence of optimal policy rates generated by different preference parameters.

However, we also find that thin and thick recursive modelling are capable of rationalizing the observed interest rate fluctuations much better than the simple constant parameters specification adopted by Rudebusch and Svensson.

Finally, we have evaluated the role of the omitted consideration of model uncertainty and parameter instability as the potential sources of the observed persistence in interest rates. We found that thick modelling alone does not help in replicating the smoothness of U.S. policy rates. The considerable observed persistence of the actual U.S. policy rates is well replicated only by optimal policy rates derived allowing explicitly for a smoothing motive in the loss function.

A Appendix: The optimal control problem

In this appendix we illustrate the derivation of the solution of the central bank's optimization problem under all the different modelling strategies adopted in the paper.

A.1 Thin modelling

Assume that the central bank minimizes an intertemporal loss function of the form:

$$E_t \sum_{\tau=0}^{\infty} \phi^{\tau} L_{t+\tau}, \tag{13}$$

where ϕ is the discount factor and E_t is the usual expectations' operator. The central bank, thus, minimizes the expected discounted sum of future values of a loss function, L_t, given in each period by:

$$L_t = \lambda_{\pi} \pi_t^2 + \lambda_y y_t^2 + \lambda_R (i_t - i_{t-1})^2. \tag{14}$$

When the discount factor ϕ approaches unity, the intertemporal loss function approaches the unconditional mean of the period loss function

$$E\left[L_t\right] = \lambda_{\pi} Var\left[\pi_t\right] + \lambda_y Var\left[y_t\right] + \lambda_R Var\left[i_t - i_{t-1}\right]. \tag{15}$$

The discussed optimization problem is then solved subject to the dynamics of the economy, which is usually given by stochastic difference equations. We first make use of a standard representation of the economy like the one employed by Rudebusch and Svensson (2002) and consisting of two simple empirical relations for inflation and output gap:

$$\pi_{t+1} = \beta_0 + \beta_1 \pi_t + \boldsymbol{\beta}' X_{t+1}^1 + u_{t+1}^1, \tag{16}$$

$$y_{t+1} = \gamma_0 + \gamma_1 y_t + \boldsymbol{\gamma}' X_{t+1}^2 + u_{t+1}^2, \tag{17}$$

where π, y stand for the inflation rate and the output gap, respectively, and X_{t+1}^1, X_{t+1}^2 correspond to the following regressors

$$X_{t+1}^{1'} = \left[\pi_{t-1}\ \pi_{t-2}\ \pi_{t-3}\ y_t\right], \tag{18}$$

$$X_{t+1}^{2'} = \left[y_{t-1}\ i_t - \pi_t\right]; \tag{19}$$

β', γ' are vectors of parameters which we can express as

$$\boldsymbol{\beta}' = \left[\beta_2\ \ \beta_3\ \ \beta_4\ \ \beta_5\right], \tag{20}$$

$$\boldsymbol{\gamma}' = [\gamma_2 \quad \gamma_3]\,. \tag{21}$$

Finally, u^1_{t+1}, u^2_{t+1} are *iid* shocks with variances $\sigma^2_{u^1_{t+1}}$, $\sigma^2_{u^2_{t+1}}$.

In order to calculate the optimal policy rule, it is convenient to rewrite the model in state-space form as

$$\mathbf{X}_{t+1} = A\mathbf{X}_t + Bi_t + \varepsilon_{t+1}. \tag{22}$$

X_t is the vector of state variables $[\pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, y_t, y_{t-1}]$, i_t is the policy instrument (the federal funds rate), and ε_{t+1} is a vector of structural shocks. A and B are parameters' matrices, with dimensions 6×6 and 6×1 respectively, given by:

$$A = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 & \beta_5 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \gamma_3 & 0 & 0 & 0 & \gamma_1 & \gamma_2 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ -\gamma_3 \\ 0 \end{bmatrix}. \tag{23}$$

The loss function can now be rewritten as:

$$L_t = \mathbf{X}'_t Q \mathbf{X}_t, \tag{24}$$

where Q is the 6×6 weights matrix, with λ_π, λ_y as elements $(1, 1)$ and $(5, 5)$, respectively, and zeros elsewhere. The central bank solves the optimal control problem

$$J(\mathbf{X}_t) = \min_{i_t} \left\{ \mathbf{X}'_t Q \mathbf{X}_t + \phi E_t J(\mathbf{X}_{t+1}) \right\}, \tag{25}$$

subject to the laws of evolution of the economy (16) and (17). After deriving the first-order condition for the minimization problem, we have that the solution for the optimal interest rate is

$$i_t = f\mathbf{X}_t, \tag{26}$$

where f is the optimal feedback vector given by

$$f = -(R + \phi B'VB)^{-1} \phi B'VA, \tag{27}$$

and the matrix V is obtained as the solution of the following Riccati equation:

$$V = Q + \phi(A + Bf)'V(A + Bf) + f'Rf, \tag{28}$$

where R incorporates the interest rate smoothing objective. We obtain that the central bank sets the optimal policy instrument value in every period as a function of the current and lagged values of the state variables as well as lagged values of the instrument itself.

Given this optimal policy rule, the dynamics of the relevant variables is defined as follows:

$$\mathbf{X}_{t+1} = M\mathbf{X}_t + \varepsilon_{t+1}, \tag{29}$$

with the matrix M given by

$$M = A + Bf. \tag{30}$$

A.2 Recursive thin modelling

We consider now the following representation for aggregate supply and demand equations:

$$\pi_{t+1} = \beta_0 + \beta_1 \pi_t + \boldsymbol{\beta}_i' X_{t+1,i}^1 + u_{t+1,i}^1, \tag{31}$$

$$y_{t+1} = \gamma_0 + \gamma_1 y_t + \boldsymbol{\gamma}_i' X_{t+1,i}^2 + u_{t+1,i}^2, \tag{32}$$

where $X_{t+1}^1 = [\pi_{t-1}, \pi_{t-2}, \pi_{t-3,} y_{t+1}, y_t, y_{t-1}, y_{t-2}, y_{t-3}]$,
$X_t^2 = [y_{t-1}, y_{t-2}, y_{t-3}, rr_t, rr_{t-1}, rr_{t-2}, rr_{t-3}, rr_{t-4}]$.

In each period only a subset of regressors is selected. The parameter vectors are given by

$$\boldsymbol{\beta}_i' = [\beta_2, \beta_3, \beta_4, \beta_5, \beta_6, \beta_7, \beta_8, \beta_9], \tag{33}$$

$$\boldsymbol{\gamma}_i' = [\gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6, \gamma_7, \gamma_8, \gamma_9]. \tag{34}$$

Re-writing the system in state-space form, we have:

$$\Gamma_{t+1}^1 \mathbf{X}_{t+1} = M_{t+1}^1 \mathbf{X}_t + W_{t+1}^1 i_t + \varepsilon_{t+1}, \tag{35}$$

where

$$\mathbf{X}_t = [1, \pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, \pi_{t-4}, y_t, y_{t-1}, y_{t-2}, y_{t-3}, i_{t-1}, i_{t-2}, i_{t-3}, i_{t-4}], \tag{36}$$

X_t is the 14×1 vector of state variables including a constant, current and lagged values of inflation, current and lagged values of the output gap and lagged values of the nominal interest rate (the federal funds rate). The central bank's policy instrument is denoted by i_t, whereas ε_{t+1} is the vector of shocks.

Here, the matrices M and W are, not invariant over time. They are, in fact, characterized by the subscript t, $t = 1, ..., 70$, which indicates the period to which they refer. The superscript 1 stands for the ranking of the selected model. In each period models are ranked in accordance to some selection criterion and the best model is selected.

As the economy is recursively estimated, the parameter matrix M_t^1, with dimension 14×14, contains the coefficients obtained for the corresponding period t. This matrix has the second and the seventh rows in period t, $t = 1, ..., 70$, given by:

$$\left[\beta_0^{t,1}, \beta_1^{t,1}, \beta_2^{t,1}, \beta_3^{t,1}, \beta_4^{t,1}, 0, \beta_6^{t,1}, \beta_7^{t,1}, \beta_8^{t,1}, \beta_9^{t,1}, 0, 0, 0, 0\right], \tag{37}$$

$$\left[\gamma_0^{t,1}, 0, 0, 0, 0, 0, \gamma_1^{t,1}, \gamma_2^{t,1}, \gamma_3^{t,1}, \gamma_4^{t,1}, \gamma_6^{t,1}, \gamma_7^{t,1}, \gamma_8^{t,1}, \gamma_9^{t,1}\right], \tag{38}$$

with zeros and occasional ones in the other places; the βs represent the parameters of the inflation equation, whereas the γs are those in the output gap relation. W_t^1 is a 14×1 parameter vector with elements:

$$\left[0, 0, 0, 0, 0, 0, \gamma_5^{t,1}, 0, 0, 0, 1, 0, 0, 0\right]. \tag{39}$$

The matrix Γ_{t+1}^1 is inserted to account for the simultaneity between output gap and inflation, it has ones on the diagonal and zeros in every place other than position $(2, 7)$ where we have the parameter $-\beta_5^{t,1}$.

Then, we find $A_t^1 = (\Gamma_t^1)^{-1} M_t^1$ and $B_t^1 = (\Gamma_t^1)^{-1} W_t^1$ obtaining the usual representation:

$$\mathbf{X}_{t+1} = A_{t+1}^1 \mathbf{X}_t + B_{t+1}^1 i_t + \varepsilon_{t+1}. \tag{40}$$

The parameters are allowed to change over time and therefore also the derived optimal rule has varying optimal coefficients over time.

We end up with an optimal monetary policy rule of the form:

$$i_t = f_t^1 \mathbf{X}_t, \tag{41}$$

with the superscript 1 as we are considering the best model, $t = 1, ..., 70$, and the feedback vector f expressed as

$$f_t^1 = -(R + \phi B_t^{1\prime} V B_t^1)^{-1} \phi B_t^{1\prime} V A_t^1, \tag{42}$$

which is now a 70×14 matrix since the 14 optimal coefficients are recalculated in every period.

A.3 Recursive thick modelling

Here we derive the optimal policy rule characterized by recursive optimal coefficients for each possible model.

The minimization problem is subject to the constraint given by the dynamics of the economy

$$\mathbf{X}_{t+1} = A_{t+1}^j \mathbf{X}_t + B_{t+1}^j i_t + \varepsilon_{t+1}, \tag{43}$$

with t indicating the observations from 1983:01 to 2000:02 and where j is the superscript relative to the model employed. We estimate 255 models coming from every possible combination of the different regressors; however, we exclude from this set of models those not incorporating an effect of monetary policy on output gap and inflation. We end up with a set of 241 relevant models; thus we are considering $j = 1, ..., 241$. The matrices A_{t+1}^j and B_{t+1}^j are calculated as $A_{t+1}^j = (\Gamma_{t+1}^j)^{-1} M_{t+1}^j$ and $B_{t+1}^j = (\Gamma_{t+1}^j)^{-1} W_{t+1}^j$.

The matrix M_t^j has the second and the seventh rows in period t, $t = 1, ..., 70$, and for every estimated model j, $j = 1, ..., 241$, given by:

$$\left[\beta_0^{t,j}, \beta_1^{t,j}, \beta_2^{t,j}, \beta_3^{t,j}, \beta_4^{t,j}, 0, \beta_6^{t,j}, \beta_7^{t,j}, \beta_8^{t,j}, \beta_9^{t,j}, 0, 0, 0, 0\right], \tag{44}$$

$$\left[\gamma_0^{t,j}, 0, 0, 0, 0, 0, \gamma_1^{t,j}, \gamma_2^{t,j}, \gamma_3^{t,j}, \gamma_4^{t,j}, \gamma_6^{t,j}, \gamma_7^{t,j}, \gamma_8^{t,j}, \gamma_9^{t,j}\right], \tag{45}$$

with zeros and occasional ones in the other places; the βs represent the parameters of the inflation equation, whereas the γs are those in the output gap relation. W_t^j is a 14×1 parameter vector with elements:

$$\left[0, 0, 0, 0, 0, 0, \gamma_5^{t,j}, 0, 0, 0, 1, 0, 0, 0\right]. \tag{46}$$

The matrix Γ_{t+1}^j accounts for the simultaneity between output gap and inflation and has parameter $-\beta_5^{t,j}$ in position $(2, 7)$.

The optimal policy rule is:

$$i_t^j = f_t^j \mathbf{X}_t, \tag{47}$$

where f_t^j is now a $70 \times 14 \times 241$ matrix, as it reports parameters resulting from every specification.

We implement thick modelling by calculating the average optimal monetary policy:

$$i_t^* = \sum_{j=1}^{241} \omega_j i_t^j, \tag{48}$$

where ω_j equals $\frac{1}{241}$ in the simple average case, or it is a weight based on the model's BIC or on the BACE quantity in (12).

Thus the optimal federal funds rate selected in every period by the central bank is the average (or a weighted average) of all the possible optimal decisions, which would have been taken under the several possible models of the economy.

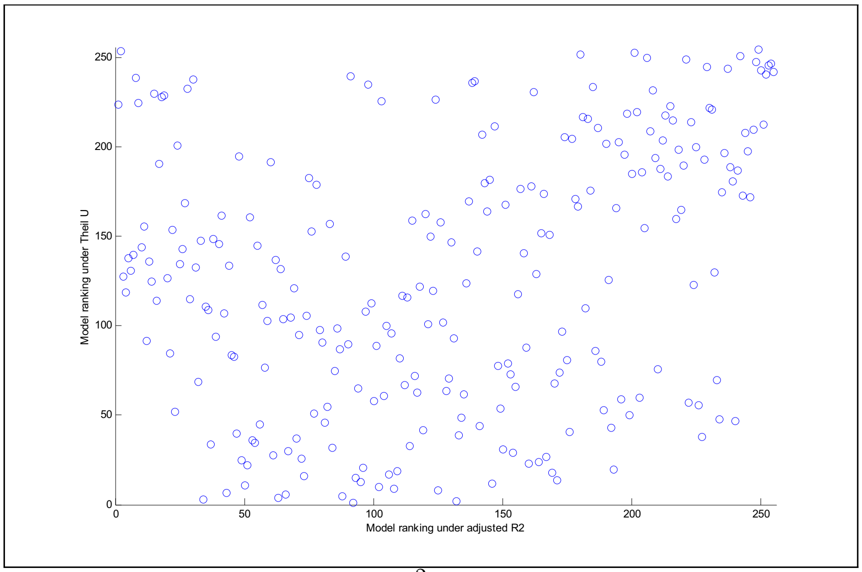

Figure 1 - Theil U vs. adjusted R^2: cross-plot of the respective model rankings for the 255 possible models of aggregate supply at last sample period.

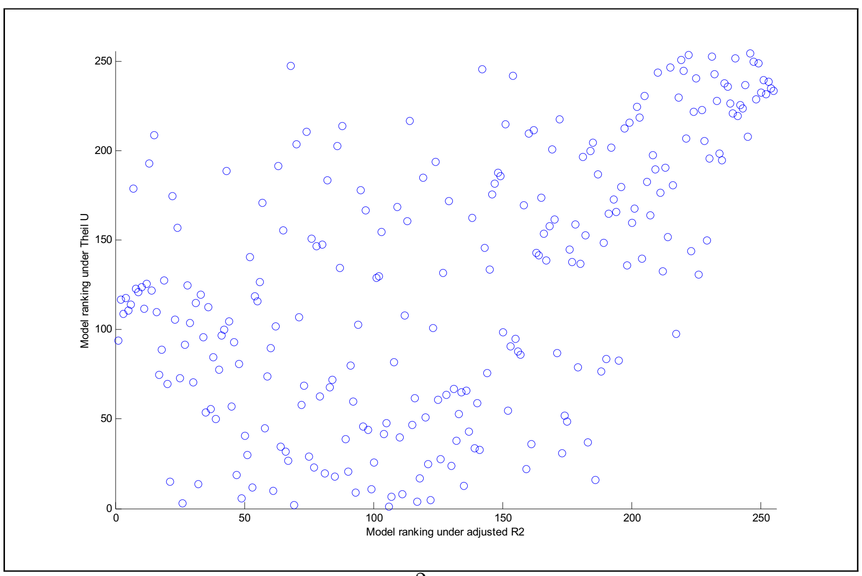

Figure 2 - Theil U vs. adjusted R^2: cross-plot of the respective model rankings for the 255 possible models of aggregate demand at last sample period.

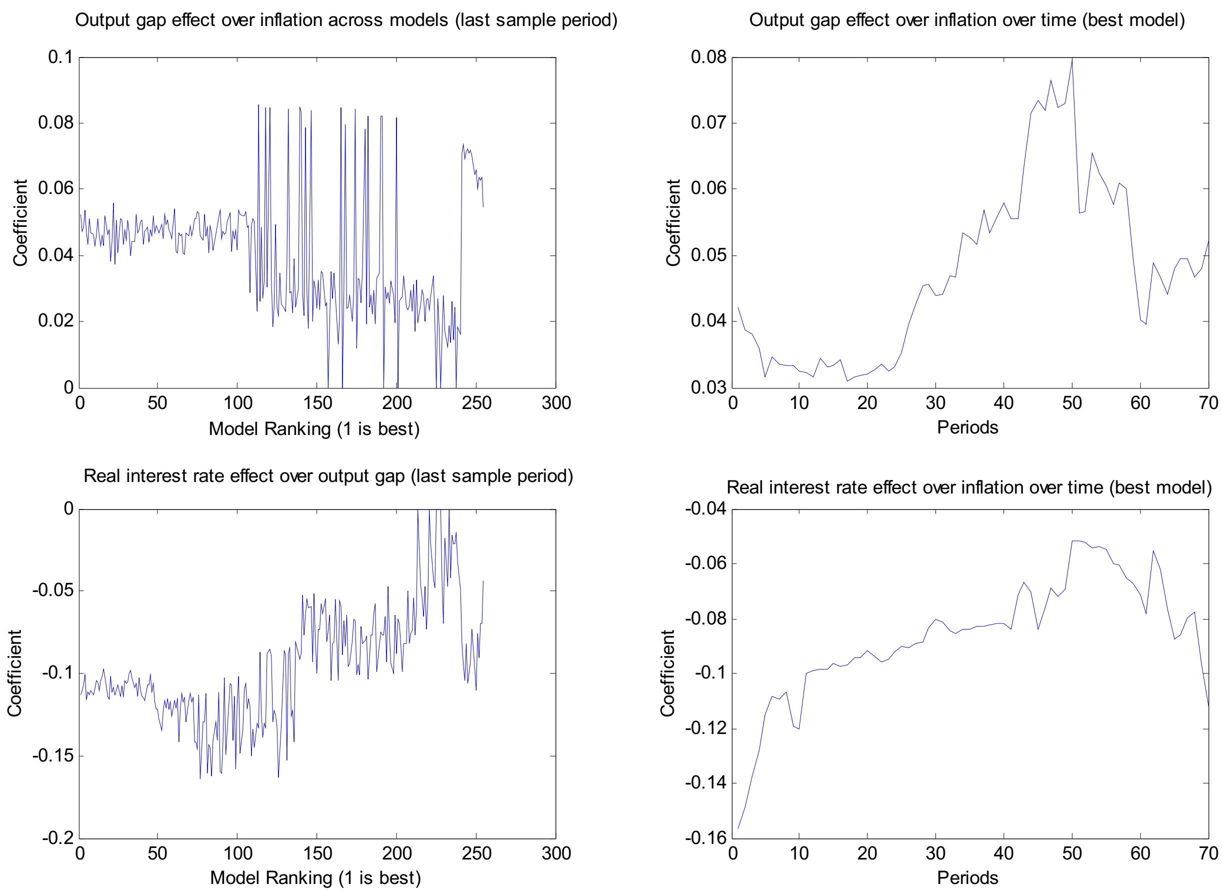

Figure 3 - Variation of the coefficients across models and time.

Note: Top graphs refer to the effect of the output gap over inflation, bottom graphs to the effect of the real interest rate over output gap. Graphs in the first column display variation across models, graphs in the second column display variation over time.

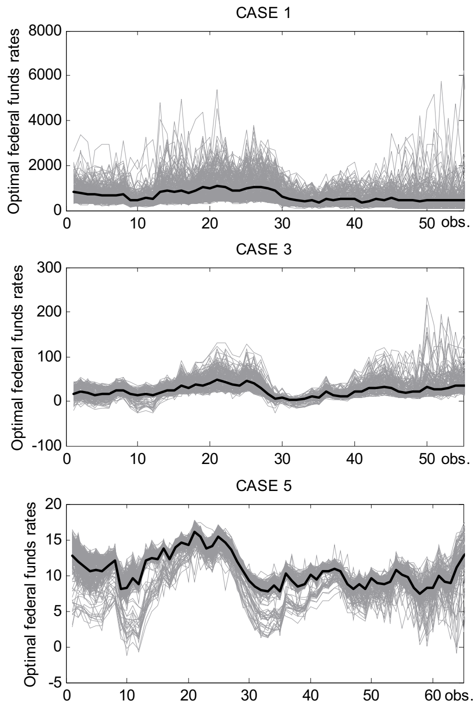

Figure 4 - Monetary policy thickness.

Each subfigure reports the distribution across models of optimal policy rates and their average (darker line).

Given that central bank minimizes an intertemporal loss function of the form:

$$E_t \left\{ \sum_{\tau=0}^{\infty} \phi^{\tau} \left[\lambda_\pi \pi_{t+\tau}^2 + \lambda_y y_{t+\tau}^2 + \lambda_R (i_{t+\tau} - i_{t+\tau-1})^2 \right] \mid \mathcal{M} \right\}$$

The five reported cases are obtained as follows:

- CASE 1. Pure (strict) inflation targeting: $\lambda_\pi = 1$, $\lambda_y = 0$, $\lambda_r = 0$.
- CASE 2. Pure inflation targeting with interest rate smoothing (strong): $\lambda_\pi = 0.7$, $\lambda_y = 0$, $\lambda_r = 0.3$.
- CASE 3. Flexible inflation targeting: $\lambda_\pi = 0.5$, $\lambda_y = 0.5$, $\lambda_r = 0$.
- CASE 4. Flexible inflation targeting with interest rate smoothing: $\lambda_\pi = 0.4$, $\lambda_y = 0.4$, $\lambda_r = 0.2$.
- CASE 5. Pure inflation targeting with interest rate smoothing (weak): $\lambda_\pi = 0.9$, $\lambda_y = 0$, $\lambda_r = 0.1$.

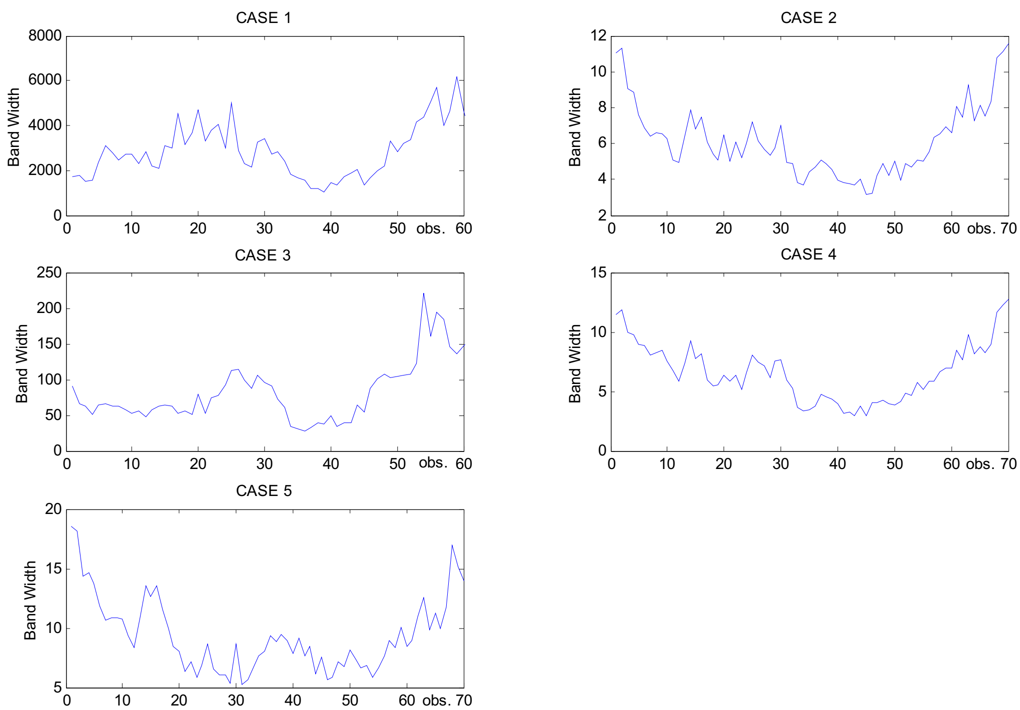

Figure 5 - Optimal policies' band width.

Given that central bank minimizes an intertemporal loss function of the form:

$$E_t \left\{ \sum_{\tau=0}^{\infty} \phi^{\tau} \left[\lambda_{\pi} \pi_{t+\tau}^2 + \lambda_y y_{t+\tau}^2 + \lambda_R (i_{t+\tau} - i_{t+\tau-1})^2 \right] \mid \mathcal{M} \right\}$$

The five reported cases are obtained as follows:

- CASE 1. Pure (strict) inflation targeting: $\lambda_{\pi} = 1$, $\lambda_y = 0$, $\lambda_r = 0$.
- CASE 2. Pure inflation targeting with interest rate smoothing (strong): $\lambda_{\pi} = 0.7$, $\lambda_y = 0$, $\lambda_r = 0.3$.
- CASE 3. Flexible inflation targeting: $\lambda_{\pi} = 0.5$, $\lambda_y = 0.5$, $\lambda_r = 0$.
- CASE 4. Flexible inflation targeting with interest rate smoothing: $\lambda_{\pi} = 0.4$, $\lambda_y = 0.4$, $\lambda_r = 0.2$.
- CASE 5. Pure inflation targeting with interest rate smoothing (weak): $\lambda_{\pi} = 0.9$, $\lambda_y = 0$, $\lambda_r = 0.1$.

	Percentage of appearances in the best model			
	Supply Equation		Demand Equation	
	Adjusted R^2	Theil U	Adjusted R^2	Theil U
π_{t-1}	100%	100%	–	–
π_{t-2}	100%	0%	–	–
π_{t-3}	4.29%	0%	–	–
π_{t-4}	18.57%	0%	–	–
y_t	41.43%	18.57%	–	–
y_{t-1}	30%	31.43%	100%	100%
y_{t-2}	58.57%	32.86%	1.43%	32.86%
y_{t-3}	40%	38.57%	25.71%	28.57%
y_{t-4}	28.57%	62.86%	81.43%	22.86%
rr_{t-1}	–	–	0%	21.43%
rr_{t-2}	–	–	100%	35.71%
rr_{t-3}	–	–	74.29%	32.86%
rr_{t-4}	–	–	85.71%	67.14%
rr_{t-5}	–	–	5.71%	31.43%

Table 1 - Percentage of appearance of the regressors in the highest ranked model over sample.

Table 2 - Optimal and actual federal funds rate paths: DESCRIPTIVE STATISTICS.							
Loss Function	Thin	Best R^2	Best U	Thick			FF
				EW	BIC	BACE	
$L_t = \lambda_\pi \pi^2 + \lambda_y y^2 + \lambda_R (i_t - i_{t-1})^2$	*Mean* *Std*	*Mean* *Std*	*Mean* *Std*	*Mean* *Std*	*Mean* *Std*	*Mean* *Std*	*Mean* *Std*
$\lambda_\pi = 1, \lambda_y = 0, \lambda_r = 0$	340.90 101.57	281.63 178.62	658.26 691.41	655.52 268.34	661.17 270.23	288.03 170.41	6.26 1.98
$\lambda_\pi = 0.7, \lambda_y = 0, \lambda_r = 0.3$	10.54 2.52	6.62 1.00	5.28 2.39	6.09 1.05	6.14 1.06	6.70 1.05	6.26 1.98
$\lambda_\pi = 0.5, \lambda_y = 0.5, \lambda_r = 0$	47.27 30.61	15.94 6.00	53.57 101.25	26.07 12.80	26.29 12.87	16.32 6.09	6.26 1.98
$\lambda_\pi = 0.4, \lambda_y = 0.4, \lambda_r = 0.2$	10.25 2.76	5.84 1.74	4.27 2.70	5.10 1.38	5.15 1.40	5.96 1.68	6.26 1.98
$\lambda_\pi = 0.9, \lambda_y = 0, \lambda_r = 0.1$	13.08 3.56	11.43 2.49	8.80 3.48	10.67 2.27	10.76 2.30	11.35 2.44	6.26 1.98

The table reports the first two moments of observed interest and optimal policy rates derived by implementing four alternative modelling strategies: thin modelling (thin), thin modelling by implementing a within-sample performance as a selection criterion (Best R^2), recursive thin modelling by implementing an out-of-sample performance as a selection criterion (Best U), and thick modelling by taking the average optimal rate across all possible models. As regards thick modelling, we consider, for sensitivity analysis, three different possible weighting schemes: equal weights (EW) for all models, weights according to the model's BIC, and according to Sala-i-Martin's BACE criterion. The last column reports the first two momnts of the actual Federal Funds rate.

Table 3 - Optimal and actual federal funds rate paths: PERSISTENCE.						
Loss Function: $L_t = \lambda_\pi \pi^2 + \lambda_y y^2 + \lambda_R (i_t - i_{t-1})^2$						
Loss Function	Thin	Best R^2	Best U	Thick		
				E.W.	BIC	BACE
	Persistence (AR(1) coeff.)					
$\lambda_\pi = 1, \lambda_y = 0, \lambda_r = 0$	0.964	0.897	0.404	0.635	0.637	0.917
$\lambda_\pi = 0.7, \lambda_y = 0, \lambda_r = 0.3$	0.915	0.732	0.681	0.795	0.797	0.769
$\lambda_\pi = 0.5, \lambda_y = 0.5, \lambda_r = 0$	0.883	0.811	0.637	0.830	0.830	0.815
$\lambda_\pi = 0.4, \lambda_y = 0.4, \lambda_r = 0.2$	0.897	0.781	0.677	0.764	0.764	0.784
$\lambda_\pi = 0.9, \lambda_y = 0, \lambda_r = 0.1$	0.920	0.851	0.727	0.833	0.834	0.833

The table reports the interest rate persistence (AR(1) coefficient, in a regression with constant and lagged value) derived by implementing four alternative modelling strategies: thin modelling (thin), thin modelling by implementing a within-sample performance as a selection criterion (Best R^2), recursive thin modelling by implementing an out-of-sample performance as a selection criterion (Best U),and thick modelling by taking the average optimal rate across all possible models. As regards thick modelling, we consider, for sensitivity analysis, three different possible weighting schemes: equal weights(E.W.) for all models, weights according to the model's BIC, and according to Sala-i-Martin's BACE criterion.

References

[1] Anderson, E.W., L.P. Hansen, E.R. McGrattan, and T.J. Sargent, (1997). "Mechanics of forming and estimating dynamic linear economies" in Amman, H.M., Kendrick D.A., Rust J., eds. *Handbook of Computational Economics*, Vol. 1.

[2] Brainard, W., (1967). "Uncertainty and the effectiveness of policy", *American Economic Review*, Papers and Proceedings 57, 411-425.

[3] Clarida, R., J. Galì, and M. Gertler, (2000). "Monetary policy rules and macroeconomic stability: evidence and some theory.", *Quarterly Journal of Economics* 115, 147-180.

[4] Doppelhofer, G., R.I. Miller, and X. Sala-i-Martin, (2004). "Determinants of long-term growth: a Bayesian averaging of classical estimates (BACE) approach", *American Economic Review*, Vol. 94, No. 4, September.

[5] Granger, C.W.J., and Y. Jeon, (2004). "Thick modeling", *Economic Modelling*, Vol. 21, Issue 2, 191-385.

[6] Hansen, L.P., and T.J.Sargent, (2003). "Robust control of forward-looking models", *Journal of Monetary Economics*, Vol. 50, Issue 3, 581-604.

[7] Hoeting, J.A., D. Madigan, A.E. Raftery, and C.T. Volinsky, (1999). "Bayesian model averaging: a tutorial", *Statistical Science* 14:4, 382-417.

[8] Onatski, A., and J.H. Stock, (2002). "Robust monetary policy under model uncertainty in a small model of the US economy", *Macroeconomic Dynamics* Vol. 6, 85-110.

[9] Pesaran, M.H., and A. Timmermann, (1995). "Predictability of stock returns: robustness and economic significance", *Journal of Finance*, Vol. 50, Issue 4, 1201-1228.

[10] ——, and ——, (2000). "A recursive modelling approach to predicting UK stock returns" *The Economic Journal*, Vol. 110, 159-191.

[11] Raftery, A.E., D. Madigan, and J.A. Hoeting, (1997). "Bayesian model averaging for linear regression models", *Journal of the American Statistical Association,* Vol. 92, 179-191.

[12] Rudebusch, G.D., (2002). "Term structure evidence on interest rate smoothing and monetary policy inertia", *Journal of Monetary Economics,* Vol. 49, Issue 6, 1161-1187.

[13] ———, and L.E.O. Svensson, (2002). "Eurosystem monetary targeting: Lessons from U.S. data", *European Economic Review*, Vol. 46, Issue 3, 417-442.

[14] Sack, B., (2000). "Does the Fed act gradually? A VAR analysis", *Journal of Monetary Economics,* Vol. 46, Issue 1, 229-256.

[15] ———, and V. Wieland, (2000). "Interest-rate smoothing and optimal monetary policy. A review of recent empirical evidence.", *Journal of Economics and Business,* Vol. 52, Issues 1-2, 205-228.

[16] Söderström, U., (1999). "Should central banks be more aggressive?", Working Paper in Economics and Finance, No. 309, Stockholm School of Economics, March.

[17] ——— (2002). "Monetary policy with uncertain parameters", *Scandinavian Journal of Economics,* Vol. 104, Issue 1, 125-145.

[18] Tetlow, R.J., and P. von zur Muehlen, (2001). "Robust monetary policy with misspecified models: Does model uncertainty always call for attenuated policy?", *Journal of Economic Dynamics and Control,* Vol. 25, Issues 6-7, 911-949.

[19] Timmermann, A., (2004). "Forecast combination", mimeo, University of California at San Diego.