Monetary Policy with a Wider Information Set: a Bayesian Model Averaging Approach*

Fabio Milani†
Princeton University

First version: August, 2002
This version: December, 2003

Abstract

Monetary policy has been usually analyzed in the context of small macroeconomic models where central banks are allowed to exploit a limited amount of information. Under these frameworks, researchers typically derive the optimality of aggressive monetary rules, contrasting with the observed policy conservatism and interest rate smoothing.

This paper allows the central bank to exploit a wider information set, while taking into account the associated model uncertainty, by employing Bayesian Model Averaging with Markov Chain Model Composition (MC^3). In this enriched environment, we derive the optimality of smoother and more cautious policy rates, together with clear gains in macroeconomic efficiency.

Keywords: Bayesian model averaging, leading indicators, model uncertainty, optimal monetary policy, interest rate smoothing.

JEL classification: C11, C15, C52, E52, E58.

*I thank Filippo Altissimo, Carlo Favero, Tommaso Monacelli, Giorgio Primiceri, Luca Sala, Lars Svensson, Andrea Tambalotti, Noah Williams, and all participants at the II Workshop in *Dynamic Macroeconomics: Theory and Applications* (Milan, December 2003), for helpful comments and useful discussions. I worked at this paper during my permanence at IGIER, Milan, and I thank them for hospitality. All errors are mine. Summer support from Princeton University is gratefully acknowledged.

†Address for correspondence: Princeton University, Department of Economics, Fisher Hall, Princeton, NJ 08544-1021. E-mail: *fmilani@princeton.edu*. Homepage: http://www.princeton.edu/~fmilani. Monetary Economics Page: http://www.princeton.edu/~fmilani/moneco.html.

1 Introduction

The analysis of optimal monetary policy has been usually carried out in the context of small macroeconomic models, where the central bank is implicitly considered able to exploit only a limited amount of information. For example, simple structural models, like those proposed by Rudebusch and Svensson (2002), and Clarida, Galì and Gertler (2000), generally confine the empirical analysis of monetary policy to frameworks characterized by just three economic variables: inflation, output gap and a short-term real interest rate.

In such a framework, the policy maker's decision problem is, probably, too simplistic and the disconnection with real world central banking appears severe. In fact, in a reality where incomplete information and uncertainty about the state of the economy are the rule, central bankers need to monitor a wide variety of economic data and indicators, at the scope of reaching a more informed and accurate decision.

The monetary policy makers not only focus their attention on current and past values of the target variables, but they also analyze a great number of the so-called 'intermediate targets': the advantage is that they are correlated with the actual target variables, but they are, differently from the latter, more easily and promptly observable. A possible drawback of their use lays in the potential instability between the target and the indicator (this factor has been a cause of the declining importance given to monetary aggregates as intermediate targets).

This attempt to reconsider monetary policy under this novel, more realistic, framework, has encountered recent examples in the literature, see the pioneering work of Bernanke and Boivin (2003).

The present paper inserts itself in this young literature and, focusing on the US case, delivers an attempt to provide a more realistic approximation of the way the Fed behaves. Within this context, this work studies how the expansion of the central bank's information set affects optimal monetary policy, comparing the results with those typically obtained under traditional limited information environments.

In particular, we try to verify whether the incorporation of an enlarged information set, together with the unavoidable associated model uncertainty, can represent a solution to an important unresolved issue in the monetary policy literature: the reconciliation with optimizing behavior of real world central bank's conservatism and interest rate smoothing.

In fact, in the context of small macroeconomic models, it is common to derive the optimality of a much more aggressive and volatile monetary policy rule, at odds with the historically observed one. This puzzle has lead to the development of a rich and active stream of research on the argument.

The potential explanations offered in the literature, regarding the desirability of policy smoothness, are surveyed in Sack and Wieland (2001) and consist of:

1. *Forward-looking expectations.* In the presence of forward-looking market participants, policy rules characterized by partial adjustment will be more effective in stabilizing output and inflation, since a small initial policy move in one direction will be expected to be followed by additional subsequent moves in the same direction. This induces a change in future expectations without requiring a large initial move. An illustration of this reasoning can be found in Woodford (2003a).

2. *Data uncertainty (real-time data).* If macroeconomic variables are measured with error, the central bank moderates its response to initial data releases in order to avoid unnecessary fluctuations in the target variables. The leading example of monetary policy using data in real-time is Orphanides (2001).

3. *Parameter uncertainty.* If there is uncertainty about the parameters of

the model of the economy, an attenuated response to shocks is optimal, as shown in the original paper by Brainard (1967). Several recent papers have reinvestigated this prediction (see, for an example in VAR models, Sack (2000)).

However, none of these characteristics is likely to represent the final word on the issue, as their empirical relevance is contrasted and it has been, typically, found quite limited; we are, therefore, still awaiting a solution of this unexplained issue. We feel that finding an answer about this topic will be strongly connected with improving the modeling of monetary policy-making under uncertainty, another area of fertile and noteworthy research in the last few years.

Therefore, this paper explores whether adding some elements of realism, like a richer information set and the associated model uncertainty, is able to render the optimality of observed gradualism and smoothness. In our environment, the central bank takes into account a variety of other data, in addition to the values assumed by inflation, output gap and the federal funds rate. As we focus on the US case, these further variables being monitored, are identified to be the leading indicators, explicitly published and recognized as important in formulating monetary policy, by the Fed.

As the available information is very large and diverse, the policy maker has to face a considerable degree of model uncertainty and there is the need to recognize which indicators are more successful and reliable predictors of current and future inflation and real activity.

Hence, to take the pervasive model uncertainty, arising in this environment, into account, we employ a technique known as Bayesian Model Averaging (BMA) with Markov Chain Monte Carlo Model Composition (MC^3). This technique has not been extensively used in economic analysis yet (also because of a certain computational effort) and, to my knowledge, has not received application in the optimal monetary policy literature so far[1].

The procedure will be described in detail in the next section, but we can already anticipate that it implies the estimation of all the models coming from every possible combinations of the regressors; the derived coefficients are, then, obtained as averages from their values over the whole set of models, weighted according to the posterior model probabilities.

We believe this technique could be of considerable help in this field, where the consideration of model uncertainty is crucial and has recently generated a growing attention. To date, model uncertainty has been introduced in the policy maker's decision problem, through the use of different techniques. A first attempt has been to add multiplicative (parameter) uncertainty, thus, supposing that the only uncertainty about the model comes from unknown values of the parameters. Another stream of research, among which the well-known work by Onatski and Stock (2002), has applied robust control techniques (minimax approach), assuming that the policy maker plays a game against a malevolent Nature and tries to minimize the maximum possible loss, whereas Nature seeks to maximize this loss. A pitfall of this approach is that it takes into account only an extreme level of uncertainty, ignoring uncertainties about other important aspects of the models. This practice corresponds to the choice of the "least favorable prior" in Bayesian decision theory, but this is just one of the possible assessment of prior model probabilities and, probably, not the most realistic; an attractive discussion of the drawbacks of the robust control literature in monetary policy is Sims (2001). Furthermore, a recent approach to model monetary policy under uncertainty has been the proposal of "thick" modeling, as in Favero and Milani (2001). In their work, they recursively estimate several possible models, generated by different combinations of the included regressors, and

[1] This was true when the first draft of this paper was released. Recently, a couple of fascinating papers, Brock, Durlauf and West (2003) and Cogley and Sargent (2003), have employed BMA to monetary policy issues.

they calculate the associated optimal monetary policies. Under recursive "thin" modeling, the best model, according to some statistical criterion, is chosen in every period and policy is derived. Then, they propose recursive "thick" modeling, as a mean to take into account the information coming from the whole set of models that would be ignored with the previous strategy. Optimal policies for each specification are calculated, and the average (or weighted average, based on some measure of model accuracy) is taken as benchmark monetary policy.

The current paper has two important advantages over their work: first, we examine a much larger degree of model uncertainty, as we consider monetary policy-making under a wider and more realistic information set, allowing the central bank to monitor a variety of data and leading indicators, whereas they only introduce uncertainty about the dynamic structure of the economy (they just consider three variables, inflation, output gap and short-term interest rates, and they contemplate uncertainty about the lags with which these variables affect aggregate supply and demand).

A second important advantage lies in the fact that we use a technique more grounded on statistical theory, like Bayesian Model Averaging. This permits us to obtain a reliable derivation of the more robust variables across different model specifications; with our procedure, we can, also, avoid problems of multicollinearity, which could be harmful using "thick" modeling in the context of a rich information environment like ours.

More generally, this paper aims at contributing to the literature in the following aspects: first, it proposes original estimation techniques and a novel method to incorporate model uncertainty in the empirical analysis of monetary policy. Then, it tries to add more realism in the modeling of central banking, by allowing the exploitation of a wider information set, with the objective of examining if in such an environment a smoother policy instrument path could be optimal.

In its novel framework, this work can help in explaining the optimality of monetary policy conservatism and interest rate smoothing, proposing an original solution that can be added to those suggested in the literature. Our results also stress the importance of taking model uncertainty into consideration in the modeling of optimal monetary policy-making.

The rest of the paper is structured as follows. Section 2 describes the methodology we use. Section 3 offers an analysis of the estimation results, the derivation of optimal monetary policy and comparisons of policy gradualism and macroeconomic efficiency with limited information frameworks. Section 4 concludes and discusses possible extensions of research.

2 Methodology

We suppose that in every period the central bank tries to obtain an estimate of its target variables, basically inflation and the output gap; in order to have an accurate perception of the current state of the economy, the central bank employs a wide variety of variables and indicators, potentially incorporating useful information about future developments of the targeted variables.

Due to the large number of included explanatory variables, we do not consider a unique model with all of them, but, instead, we focus on all the possible combinations obtained with the different regressors. Thus, if the specification contains k potential regressors, we end up with the use of $2^k \times 2$ (as we have two equations) different models: in our case, we have 15 variables per equation, and we consider 4 lags for each of them; hence, we end up dealing with a set of $2^{60} \times 2$ possible models M_j.

We may, therefore, describe the inflation and output equations the policy maker implicitly uses, in the following way:

$$[\text{AS}] \qquad \pi_t = \beta_0^\pi \iota_t + \beta_j^\pi \mathbf{X}_{j,t} + \varepsilon_t^\pi \tag{1}$$

$$\text{[AD]} \qquad y_t = \beta_0^y \iota_t + \beta_j^y \mathbf{X}'_{j,t} + \varepsilon_t^y \tag{2}$$

where ι_t is a t-dimensional vector of ones, β_0^π and β_0^y are constants, β_j^π and β_j^y are vectors of the relevant coefficients for every model j, and the regressors' matrices are represented by $\mathbf{X}_{j,t} = [\pi_{t-1},\ \pi_{t-2},\ \pi_{t-3},\ \pi_{t-4},\ y_{t-1},\ y_{t-2},\ y_{t-3},\ y_{t-4},\ \mathbf{Z}_{t-1},\ \mathbf{Z}_{t-2},\ \mathbf{Z}_{t-3},\ \mathbf{Z}_{t-4}]$, $\mathbf{X}'_{j,t} = [y_{t-1},\ y_{t-2},\ y_{t-3},\ y_{t-4},\ (i_{t-1} - \pi_{t-1}),\ (i_{t-2} - \pi_{t-2}),\ (i_{t-3} - \pi_{t-3}),\ (i_{t-4} - \pi_{t-4}),\ \mathbf{Z}_{t-1},\ \mathbf{Z}_{t-2},\ \mathbf{Z}_{t-3},\ \mathbf{Z}_{t-4}]$, with $\mathbf{Z}_{t-i}$, $i = 1, 2, 3, 4$, including lags of the leading indicators used by the Fed and listed in the appendix. We use quarterly data, from 1969 to 2001. In the estimation, we employ demeaned variables, so we can get rid of the constants.

We notice, therefore, that inflation depends on lags of itself and of a real activity measure (the output gap, in our case), and the output gap depends on its past values and on lagged real short term interest rate. This is a very common specification in monetary policy analysis. In addition, in our framework, we have allowed for the consideration of several other potential predictors, which enters the vector $\mathbf{Z}_t$. As said, we do not, obviously, consider a single model with all those variables, but we aim at estimating all the possible models obtained with the different combinations of the regressors, by using a procedure, which allows us to retain the potentially very important information coming from each model.

To deal with the considerable model and parameter uncertainty, characterizing this wide information environment, we use Bayesian Model Averaging (BMA) estimation, which allows inference averaged over all models. To solve the computational burden, we employ a technique known as Markov Chain Monte Carlo Model Composition (MC^3), which derives from the work of Madigan and York (1995). The use of BMA is necessary, as we have included a too large number of regressors to be estimated within one single model; this technique enables us to account for model uncertainty, identifying the most robust predictors across all the possible specifications.

Our estimation procedure enables us to deal also with heteroskedasticity and the presence of outliers; by applying the Markov Chain Monte Carlo method, instead, we are able to derive the posterior distribution of a quantity of interest through the generation of a process which moves through model space. Our choice of prior distribution reflects what is done in Raftery, Madigan and Hoeting (1997), i.e. the use of data-dependent "weakly informative" priors. In particular, following them, we use the standard normal-gamma conjugate class of priors:

$$\beta \sim N\left(\mu, \sigma^2 V\right),$$

$$\frac{\upsilon\lambda}{\sigma^2} \sim \chi^2_\upsilon,$$

where υ, λ, the matrix V and the vector μ are hyperparameters to be chosen. The distribution of β is centered on 0, we use $\mu = (0, 0, ..., 0)$, as our variables are demeaned and, thus, we can avoid the constant term. The covariance matrix V equals σ^2 times a diagonal matrix with entries given by $\left(var(Y), \frac{\phi^2}{var(X_1)}, \frac{\phi^2}{var(X_2)}, ..., \frac{\phi^2}{var(X_k)}\right)$, where here Y stands for the dependent variable, X_j, $j = 1, ..., k$, for the j-th regressor, and ϕ is a hyperparameter to be chosen. In our estimation, we select $\upsilon = 4$, $\lambda = 0.25$, $\phi = 3$ (we have experimented different values, but the results are substantially unchanged).

By means of Bayesian estimation, the parameters are averaged over all possible models using the corresponding posterior model probabilities as weights; in accordance with the literature, exclusion of a regressor means that the corresponding coefficient is zero.

This procedure is clearly better than just considering a single best model M^*, and acting as if it was the 'true' model, since this procedure would ignore the, potentially considerable, degree of model uncertainty and would lead to underestimation of uncertainty about the quantities of interest.

The Bayesian solution to this problem is the following: define $\mathcal{M} = \{M_1, ..., M_k\}$, the set of all possible models, and assume Δ is a quantity of interest. Then, the posterior distribution of Δ given the observed data D is:

$$pr\left(\Delta|D\right) = \sum_{k=1}^{K} pr\left(\Delta|M_k, D\right) pr\left(M_k|D\right), \tag{3}$$

which is an average of the posterior distributions under each model, weighted by the respective posterior model probabilities. This is exactly what is known as Bayesian Model Averaging (BMA). From (3), $pr\left(M_k|D\right)$ is given by:

$$pr\left(M_k|D\right) = \frac{pr\left(D|M_k\right) pr\left(M_k\right)}{\sum_{j=1}^{K} pr\left(D|M_j\right) pr\left(M_j\right)}, \tag{4}$$

where

$$pr\left(D|M_k\right) = \int pr\left(D|\beta_k, M_k\right) pr\left(\beta_k|M_k\right) d\beta_k \tag{5}$$

represents the marginal likelihood of model M_k, obtained as the product of the likelihood $pr\left(D|\beta_k, M_k\right)$ and the prior density of β_k under model M_k, $pr\left(\beta_k|M_k\right)$; β_k is the vector of parameters of model M_k, and $pr\left(M_k\right)$ is the prior probability of M_k being the 'true' model (note that all the probabilities are implicitly conditional on the set of all possible models $\mathcal{M}$).

Before implementing any method of estimation, we need to specify a prior distribution over the competing models M_k (i.e., we need to assign a value to $pr\left(M_k\right)$ in expression (4)). The obvious neutral choice, when there is no a priori belief, would be to consider all models as equally likely. Otherwise, when we have prior information about the importance of a regressor, we can use a prior probability for model M_k:

$$pr\left(M_k\right) = \prod_{j=1}^{p} \pi_j^{\delta_{kj}} \left(1 - \pi_j\right)^{1-\delta_{kj}}, \tag{6}$$

with $\pi_j \in [0, 1]$ representing the prior probability of $\beta_j \neq 0$ and δ_{kj} is a variable assuming value 1 if the variable j is included in model M_k, and value 0 if it is not. Here, we consider $\pi_j = 0.5$, which corresponds to a Uniform distribution across model space. In this case, the prior probability of including each regressor is 1/2, independently of which other predictors are already included in the model.

Obviously, different choices would be possible, for example, an interesting case would be considering $\pi_j < 0.5 \, \forall j$, which corresponds to imposing a penalty on larger models. It is probably worthwhile to mention that even if our prior choice does not excessively penalize larger models, our estimation results will favor very parsimonious specifications as our best models.

With an enormous number of models, the posterior distributions could be very hard to derive (the number of terms in (3) could be extremely large, and also the integral in (5) could be really hard to compute).

For this reason, we need to approximate the posterior distribution in (3) using a Markov Chain Monte Carlo approach, which generates a stochastic process which moves through model space. An alternative approach, not implemented here, would have been the use of Occam's Window, which implies averaging over a subset of models supported by the data.

Our method works as follows: we construct a Markov Chain $\{M_t, t = 1, 2, 3, ...\}$ with state space M and equilibrium distribution $pr\,(M_j|D)$, then we simulate this Markov Chain for $t = 1, ..., N$, with N the number of draws. Under certain regularity conditions, it is possible to prove that, for any function $g(M_j)$ defined on M, the average

$$G = \frac{1}{N}\sum_{t=1}^{N} g(M(t)) \;\; \xrightarrow{a.s.} \;\; E(g(M)), \text{as } N \to \infty, \tag{7}$$

i.e. it converges almost surely to the population moment (for a proof, see Smith and Roberts, 1993). Setting $g(M) = pr\,(\Delta|M, D)$, we can calculate the quantity in (3).

In the implementation, given that the chain is currently at model M_s, a new model, say M_i,which belongs to the space of all models with either one regressor more or one regressor less than M_s, is considered randomly through a Uniform distribution. The chain moves to the newly proposed model M_i with probability $p = \min\left\{1, \frac{pr(M_i|D)}{pr(M_s|D)}\right\}$, and stays in state M_s with probability $1-p$.

The goal of the procedure is to identify the models with highest posterior probabilities: only a limited subset of the models is thus effectively used in the estimation, but, in any case, a subset representing an important mass of probability.

In the estimation, all the regressors are employed and the coefficients' values result from the averaging over all possible models, using, as weights, the posterior model probabilities, which, in turn, are based on the number of visits of the chain. As previously mentioned, when a regressor is not included in a specification its coefficient is zero. If a regressor is not a significant predictor for the dependent variable, it is assigned a coefficient close to zero with a high p-value.

After accounting for model uncertainty, the posterior mean and variances of the coefficients will be:

$$E(\beta \mid D) = \sum_{M_j \in \mathcal{M}} pr(M_j \mid D) E(\beta \mid D, M_j) \tag{8}$$

$$var(\beta \mid D) = \sum_{M_j \in \mathcal{M}} pr(M_j \mid D) var(\beta \mid D, M_j) + \sum_{M_j \in \mathcal{M}} pr(M_j \mid D)(E(\beta \mid D, M_j) - E(\beta \mid D))^2 \tag{9}$$

As already explained, the posterior mean will be the weighted average of the posterior means across each model, weighted by the model posterior probabilities. The posterior variance will be, instead, the sum of two terms: again a weighted average of all the variances across models plus a novel term, $\sum_{M_j \in \mathcal{M}} pr(M_j \mid D)(E(\beta \mid D, M_j) - E(\beta \mid D))^2$, which reflects the variance across models of the expected β. This term accounts for the variance explicitly due to model uncertainty and it is new. Even being the variance constant across models, we would have $var(\beta \mid D, M_j) < var(\beta \mid D)$ as long as there is variation in $E(\beta)$ across models. This shows that not accounting for model uncertainty leads to the underestimation of uncertainty about the quantities of interest.

2.1 How to *Model* Model Uncertainty: why BMA?

In this section, we discuss some advantages and the intuition behind our decision to *model* model uncertainty (to borrow the expression from Onatski and Williams (2003)) by BMA. In particular, we would like to point out some potential favorable characteristics of BMA compared to the use of robust control, currently one of the most used methods to insert model uncertainty in monetary policy settings.

With respect to robust control, BMA can allow the consideration of a bigger amount of uncertainty, since it permits to deal with a huge set of models, where each single model can be false. Under robust control, on the other hand, there is the need to start from a reference 'true' model and just consider perturbations around it. As said, the minimax approach corresponds to the choice of the "least favorable prior" in Bayesian decision theory: it is not certain that this is the most realistic choice. Another potential problem, pointed out by Brock, Durlauf and West (2003), is that the minimax assumption leads to ignore posterior model probabilities. A model, which represents the "worst case" in terms of possible losses will be the only specification used to shape policy, even if it has very low posterior probabilities. This is certainly useful in some applications (e.g., for global climate change discussions) but the appropriateness for monetary policy is not so clear.

In addition, BMA easily allows to compare both *nested* and *non-nested* competing models, a task that is not always straightforward from a classical perspective.

From a more practical point of view, with BMA we aim to model the behavior of a central banker that has in mind several models of the economy, without knowing the 'true' model; however, the central banker is able to assign, at least implicitly, some probabilities to the different specifications. She has in his mind a sort of weighted average of all the possibilities, where the weights are the model probabilities. Each model has different implications for policy. Because of this uncertainty, policy decisions will not be based on a single model, but they will be affected by all the possible alternatives.

This way of formulating policy seems quite realistic. Indeed, it emerged from interviews with Fed's staff economists, as reported in Sims (2002). Interestingly, they refer to this method of weighting probabilities as "unscientific" (where they believed that "scientific" would have been to test hypotheses of interest at the 5% significance level), whereas they are pure applications of Bayesian decision theory to their problem.

3 Optimal Monetary Policy with a Wider Information Set

3.1 Bayesian Estimation

As a first step, we estimate the two equations (1), (2), used by the central bank to represent the dynamics of the target variables. As we assume that, in our framework, the policy maker deals with a great amount of information in addition to the usual data, it is important to identify which variables are important indicators of the developments of the economy. BMA is ideal to derive such an information, as it estimates several possible models and assigns them probabilities to be the "true" model; then, it permits to detect which regressors are indeed successful explanatory variables and which are not worth giving much attention.

One thing to note is that the leading indicators we use are likely to include some variables, which can be highly collinear: the Markov Chain Monte Carlo procedure, working in the way described above, will avoid models with collinear regressors, assigning, on the contrary, high posterior probabilities to combinations of regressors not characterized by this problem. In fact, when the Chain is at model M_s and a new model with a further regressor, suppose collinear with one of those already included, is proposed through the Uniform Distribution, it is likely that the Chain will not move to the new model, as the additional variable does not convey more useful information[2].

In our approach, we estimate the model over the whole sample. This implies that we believe that the central bank makes use of all the possible information and estimates all the possible models, identifying, in this way, those which carry a higher probability of being the true model. But our assumption is that these 'better' models are so during the whole sample. The leading indicators, which are more successful predictors of the target variables, remain the same during the entire period of our consideration; their relation with inflation and output is, therefore, stable.

Thus, the policy makers try every possible model, but when they discover which models are closer to the 'truth', they think these are stable through time.

An obvious alternative would be to consider all the competing models and also assume that the best models change from period to period. In this framework, the policy makers update their more successful models in every period, giving more weights to different leading indicators each time. The models and the parameters are, therefore, time-varying.

A potential problem with a time-varying framework, like the one described, would be that, when we turn to derive the optimal monetary policy rule, we would encounter time-inconsistency. In fact, in every period we would have

[2] Bernanke and Boivin (2003), in their analysis of monetary policy in a data-rich environment, have dealt with this problem using dynamic common factors.

a different model of the economy, with different regressors and different coefficients, and we would have to calculate the central bank's reaction function under the period's framework. But, if we derive it by dynamic programming in every period, we would act as if the rule would remain stable all over the future, which is clearly not the case here, because it will be updated already in the next period. Therefore, in such a time-varying world, at the scope of optimal monetary policy, it would be necessary to consider some learning mechanisms from one period to the other, in order to face the described inconsistency. This issue is taken up in Milani (2003), which introduces adaptive learning in a similar framework.

In the present paper, however, we consider the case of stable models across the whole sample period.

We thus estimate the equations (1) and (2) with Bayesian Model Averaging. Our Markov Chain Monte Carlo simulation is based on 51,000 draws, with the initial 1,000 draws omitted for burn-in. The results are reported in Table 1-2 and Table 3-4 for the supply and demand specifications, respectively.

Insert Table 1-2 here.

In Table 1, we see that the chain has visited 31,859 models; among these, we report the 10 models, which have been assigned the highest posterior probabilities. In the table, 1 stands for inclusion of a regressor, 0 for its exclusion.

We notice that among all the models, the most supported by the data is characterized by a little bit more than a 1% probability to be the "true" model. This clearly indicates that there is enormous uncertainty about the correct model of the economy. From our estimation results, we can, thus, understand the superiority of a method capable of taking model uncertainty into account versus choosing a single 'best' model, since the posterior probability is spread among several models.

In Table 2, we report the posterior estimates of the coefficients, which are obtained by averaging over the whole set of models, with weights equal to the respective posterior model probabilities, together with the associated t-statistics and p-values. As already explained, a regressor which is not usually included in the selected models is assigned a near zero coefficient with a high p-value.

From the results, we can, therefore, notice that, besides lagged inflation, other variables are useful determinants of future inflation; these are CPI inflation, new orders and the output gap. CPI inflation, apparently, is able to take into account a fraction of price variation not captured by the standard measure of inflation. The most useful indicators of the situation of real activity are observed to be the value of new orders in the previous period, and the, commonly used, output gap. This latter, however, has been found to have an impact on inflation only after four periods (this is consistent with standard models of monetary policy, which assumes one year is needed for changes in real activity to affect inflation).

The same reasonings apply for the demand estimation results, where the most likely model accounts for 2.4% of the total mass probability.

Insert Table 3-4 here.

In this case, successful determinants of the output gap are its lagged value, the real interest rate, the indicator of consumer confidence and housing starts (the latter seem able to capture effects of the construction sector, not taken into account by our output gap measure). Variations in the real interest rate have an effect over real activity after two quarters.

The models visited by the chain have been individually estimated by OLS; in this situation, where the regressors are not the same across the two equations, and the residuals can be correlated, OLS is not the most efficient estimator. Therefore, we deem it necessary to evaluate whether a joint estimation of our equations substantially changes the results. Our specifications are simultaneously estimated by the Seemingly Unrelated Regression (SUR) method, the efficient estimator in this case; the coefficients are, again, obtained as weighted averages, across the whole set of models visited by the MCMC procedure, with the posterior model probabilities as weights. The results are shown in Table 5.

Insert Table 5 here.

We can, easily, notice that the estimates are totally similar to the equation-by-equation results.

In order to evaluate the convergence of the sampler, we have performed the simulation starting from different initial conditions: the results are unchanged. We have also experimented different lag structures, to verify that our findings are robust across different specifications. Again, the significant variables in the estimation and the monetary policy outcome, which will be described in the next section, are absolutely similar.

3.2 Optimal Monetary Policy: Conservatism and Interest Rate Smoothing

After having estimated the equations, we want to derive the optimal monetary policy the central bank would have followed under this framework. It is our intention, in particular, to examine how the amplification of the policy maker's information set (together with the existing model uncertainty) affects the optimal reaction function and how this compares with those obtained under more traditional macroeconomic models.

In particular, we consider the unresolved issue of the strong divergence between optimal monetary policy as derived from theory, which indicates the optimality of much more variable and aggressive interest rate paths, and central banks' behavior observed in practice, instead, characterized by pronounced policy "conservatism" (attenuation of the feedback coefficients regarding inflation and output) and "interest rate smoothing" (partial adjustment to the evolution of the economy, reflected in small deviations from previous period interest rate value).

We verify whether the allowance of a wider information set determines significant changes over optimal monetary policy decisions.

To do this, we have to solve the stochastic dynamic optimization problem of a central bank, which wants to minimize an intertemporal loss function, quadratic in the deviation of inflation and output gap from their respective targets and with a further term denoting a penalty over interest rate excessive variability. The period loss function is, therefore, given by:

$$L_t = \lambda_\pi \pi^2 + \lambda_y y^2 + \lambda_{\Delta i}(i_t - i_{t-1})^2. \tag{10}$$

The optimization is performed under the constraints given by the dynamics of the economy. Were we considering just equations (1) and (2) as our constraints, it would correspond to consider our information variables in Z_t as purely exogenous with respect to policy. However, variables like CPI inflation, new orders and others are certainly affected by policy and, therefore, we have to take their endogeneity into account if we want our policy to be optimal. To solve this issue, we choose to treat all our additional variables as endogenous. To this scope, we re-estimate the whole system by BMA (exploiting the simultaneity by using SUR) where we have one equation for each of the 15 variables (the regressors are the same of those in the output and inflation equations). The estimated parameters for the aggregate supply and demand equations are very similar to those previously obtained and we do not report them.

In standard macro models, the optimal rule is usually given by:

$$i_t^* = fX_t, \tag{11}$$

i.e., the policy instrument is fixed in every period in response to the evolution of the state variables[3]. The rule generally resembles a traditional Taylor rule, where the federal funds rate responds to deviations of inflation and output gap, or, also, a Taylor rule with partial adjustment.

A Taylor rule expressed as a linear function of inflation and the output gap will be optimal only if these are sufficient statistics of the state of the economy and they are perfectly observed. These conditions are probably not met in reality.

The introduction of a larger information set can be approached by assuming that the central bank makes use of all the available data to produce forecasts of inflation and output and then calculates an optimal rule with only these target variables' forecasts as arguments.

Our approach, instead, consists on letting the central bank directly respond to all the available series and leading indicators, thus allowing a more complete optimal feedback rule. In fact, when taking a decision, the monetary policy maker responds to the developments of the indicators she is monitoring: it seems correct to evaluate which variables are more successful in predicting inflation and real activity state (we do this by means of BMA) and then calculate those variables' optimal feedback coefficients, which are thus based on the extent they are indeed useful predictors of the movements of the target variables.

The optimal monetary policy rule we obtain is:

$$i_t^* = f\left[\pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, y_t, y_{t-1}, y_{t-2}, y_{t-3}, \mathbf{Z}_t, \mathbf{Z}_{t-1}, \mathbf{Z}_{t-2}, \mathbf{Z}_{t-3}, i_{t-1}, i_{t-2}, i_{t-3}\right], \tag{12}$$

where, in addition to the usual Taylor rule terms, the optimal interest rate, each period, is adjusted in relation to the situation of several economic indicators through the feedback coefficients found in the 1×63 vector f. This seems to better represent the real world case, where the Fed responds to a variety of

[3]The derivation is, by now, standard and we omit it. The interested reader can find a thorough derivation in the appendix in Rudebusch and Svensson (2002), among others.

different information.

In our setting, the feedback coefficients in f will be convolutions of the policy weights, the structural parameters governing the dynamics of the economy and the relative model probabilities (through the $pr(M_j \mid D)$ term). This latter term makes clear the dependence of policy on the uncertain environment.

From a signal extraction point of view, if the predictive accuracy of a variable increases, policy will respond more to changes in this variable (higher feedback coefficient).

To evaluate the effects of the widening of the information set, we compare the optimal reaction functions and the implied optimal federal funds target rate paths (calculated by applying the optimal feedback coefficients to the actual state of the economy, in every period), obtained under traditional backward-looking representations of the economy, as the Rudebusch and Svensson (2002)'s model (RS), which takes into consideration only three variables (inflation, output and short term interest rate), and in the context of our higher dimension framework.

We recall that Rudebusch and Svensson's specification is given by two simple equations of the following form:

$$\pi_{t+1} = \beta_1 \pi_t + \beta_2 \pi_{t-1} + \beta_3 \pi_{t-2} + \beta_4 \pi_{t-3} + \beta_5 y_t + \varepsilon_{t+1}, \tag{13}$$

$$y_{t+1} = \gamma_1 y_t + \gamma_2 y_{t-1} + \gamma_3 (i_t - \pi_t) + \eta_{t+1}. \tag{14}$$

We would like to be certain that our results do not depend on the particular choice of the RS model as a benchmark. Therefore, we also compare our results to a richer backward-looking specification, more similar to our setting, represented by a monetary VAR, augmented by some of our information variables. We choose the following:

$$\mathbb{X}_t = \Phi(L) \mathbb{X}_{t-1} + \mathbf{e}_t, \tag{15}$$

where $\mathbb{X}_t = [\pi_t, y_t, i_t, Z_t]$, and Z_t contains some of the leading indicators we have found more important in estimation ($Z_t =${*CPI infl, housing starts, new orders, consumer confidence*}).

We compare the optimal federal funds rate paths obtained under our large information setting (BMA) and under the two competing frameworks (RS, VAR), with the actual series, historically implemented by the Federal Reserve. Moreover, we consider four different cases concerning various aversions to interest rate variability ($\lambda_{\Delta i} = 0, 0.07, 0.2, 0.5$), and three possible preference alternatives: strict inflation targeting, strict output targeting, and equal weight to inflation and output stabilization. This is necessary, due to the difficulty in identifying central bank preferences.

The results under all these cases and in the various models are reported in Table 6.

Insert Table 6 here.

If we allow the central bank to deal with an increased and more realistic amount of information and we take into account the existing model uncertainty, we can obtain optimal federal funds rate paths quite similar to the actual one, by considering a small 0.07 penalty on interest rate variability in the loss function (against a relative weight of 0.93 given to inflation).

We see from the table that just with a very small penalty on policy instrument volatility (assuming "strict" inflation targeting and $\lambda_{\Delta i} = 0.07$), we are able to obtain optimal federal funds series close to the historically realized one; over the sample, it is characterized by mean and standard deviation not far from those of the actual funds rate (mean and standard deviation equal to 7.41 and 3.40, compared with the actual 7.34 and 3.15, respectively). It is also evident the improvement over the consideration of optimal monetary policy under alternative backward-looking specifications, where we end up with less realistic interest rate series (too aggressive and volatile, std.= 9.77 for RS and std.= 14.97 for VAR).

Even allowing for a sensibly stronger preference for smoothing in the objective function (say $\lambda_{\Delta i} = 0.2$ or also $\lambda_{\Delta i} = 0.5$), the optimal rules we obtain do not lead to funds rate's paths characterized by standard deviations compatible with the actual one. In fact, in these cases, the interest rate series derived from the RS and VAR frameworks continue to be excessively volatile (an exception is found under strict output targeting with a very high preference for smoothing and RS model, $\lambda_{\Delta i} = 0.5$); the series obtained under wider information sets do not feature enough variability, compared to the actual one, as the care for smoothing is too high.

When no weight at all is assigned to interest rate smoothing in the loss function (i.e. $\lambda_{\Delta i} = 0$), the optimal federal funds rate series never come close to the actual ones.

However, there seems to be a clear indication: when a wider information set and model uncertainty are accounted for, optimal policy rates compatible with the observed ones are obtained with the presence of a small penalty on interest rate volatility ($\lambda_{\Delta i} = 0.07$), whereas, when we are in the context of standard alternatives, RS or VAR augmented with some informational variables, even unrealistically high penalties ($\lambda_{\Delta i} = 0.5$) are not usually able to deliver optimal series comparable to the actual one.

This is clearly important, as we can much more easily justify a very low penalty, with the desire to avoid extreme and unrealistic jumps in the federal funds rate, while we do not think a much bigger weight to be justifiable.

Another characteristic about interest rates it is worth exploring, besides volatility, is persistence. In the table, we show the estimated AR(1) coefficient of a regression of the series on a constant and its one-period lagged value. We notice that all the optimal series are able to replicate the strong persistence, which characterizes policy rates in reality (optimal rates under wider information seem to be, generally, a little more persistent than traditional model's counterparts).

Figure 1 shows the path of our optimal policy rate (under $\lambda_\pi = 0.93, \lambda_{\Delta i} = 0.07$), together with the actual series. The tracking of the actual variable is certainly far from being perfect, but the sizes of the funds rates are comparable. The corresponding graph for the optimal series under RS and VAR, on the other hand, is absolutely unrealistic and too volatile (with some values higher than one hundred). Those are therefore not very indicative and not reported.

Insert Figure 1 here.

Now, in Table 7 and 8, we report the optimal feedback coefficients in the calculated reaction functions, for the case $\lambda_\pi = 0.93$, $\lambda_y = 0$, $\lambda_{\Delta i} = 0.07$, for our model and the RS model, respectively.

Insert Table 7-8 here.

It is immediate to notice a strong attenuation of the feedback coefficients in the wider information framework, compared to the traditional one, where the reaction function indicates a far too aggressive response to the evolution of the economy. In fact, the sum of the feedback coefficients to inflation (both GDP deflator and CPI inflation) and output gap amount to 1.17 and 0.42, respectively, in the wider information case, against 3.89 and 1.97 in the alternative framework. It seems that a bigger information set and the consideration of the associated model uncertainty lead to a much more 'conservative' monetary policy, thus, proposing themselves as candidate explanations of the pronounced policy conservatism, observed in practice.

3.2.1 Large Information and Interest-Rate Smoothing

A not so expected result is that we do not find any evidence of interest-rate smoothing. In fact, the optimal feedback coefficients on lagged federal funds rates are close to 0, contrasting with commonly estimated values around 0.9. Our optimal policy rate series is smooth, but the smoothness does not come from the optimality of small deviations from past rates. Our results points towards the illusion version of interest-rate smoothing, due to Rudebusch (2002). In fact, we have seen that the optimal feedback coefficients to past policy rates is almost zero. However, if we regress the optimal funds rate on its lagged value and on current output gap and inflation (demeaned variables), estimating the following standard Taylor rule with partial-adjustment:

$$i_t^* = \rho i_{t-1}^* + (1-\rho)(g_\pi \pi_t + g_y y_t) + \nu_t, \tag{16}$$

we obtain the following results:

$$i_t^* = \underset{(0.052)}{0.571} i_{t-1}^* + \underset{(0.052)}{0.429} (\underset{(0.048)}{1.562} \pi_t + \underset{(0.060)}{0.183} y_t) + \nu_t, \tag{17}$$

The interest-rate smoothing term is not around 0.9, but still from estimation there is the perception of a considerable degree of partial adjustment (0.57), when in fact there was none in the optimal rule. Our findings are therefore consistent with the illusion argument of Rudebusch (2002). This kind of attenuation, due to the fact that the Fed responds to many different shocks and has large information, was also documented in Tetlow (2000), who used the much larger FRB/US macro model. Like us, he was also able to obtain the empirical equivalence of a non-inertial optimal rule and a much more inertial empirical counterpart.

The results do not simply reflect the serial correlation of the information variables. The leading indicators have been rendered stationary (considering their growth rates) and usually have small autoregressive coefficients. An indication that our smooth policy rate series is not just driven from the serial correlation of the leading indicators can derive from our VAR findings. The optimal policy rate coming from the VAR, augmented with informational variables, is still extremely volatile. If serial correlation of the leading indicators were driving the results, we would expect policy from the VAR to be smooth as well. As seen, this does not happen.

In conclusion, we can affirm that the explicit consideration of leading indicators leads to a substantial attenuation of optimal policy rules and to a smoother interest rate path, thus partially helping to reconcile macroeconomic theory results with reality.

The intuition behind our finding of a smoother instrument path is the following: in our framework, the central bank, constantly, monitors many economic variables and, as a consequence, when taking a policy decision, it responds, in different measures, to all their deviations from equilibrium. However, in this environment, the several and diverse indicators are likely to provide mixed indications of the current stance of the economy and, therefore, contrasting implications for policy. As a consequence of this need to respond to this heterogeneous information, the central bank's reaction seems attenuated (in relation to changes in the target variables) and deviations from previous period values of the instrument are small, inducing a smooth path of policy. On the other hand, in the context of standard macroeconomic models, the policy maker is assumed to have perfect information about the target variables, inflation and output gap. When these series depart from their targets, theory implies the optimality of a sudden and aggressive adjustment back towards equilibrium (as there is no uncertainty about the actual situation of the economy). This, clearly, means that extremely pronounced and volatile instrument changes are optimal in such a framework.

This could represent a new and original explanation of realized monetary policy gradualism, in addition to the traditional ones suggested in the literature and consisting of central bank's preference for smoothing (due to the care for financial markets stability, etc.), uncertainty (parameter, data and model uncertainty) and forward-looking expectations.

Following this view, therefore, the common inability to obtain the optimality of a gradual monetary policy and of interest rate smoothing could be due to a misspecification of traditional simple macro models (which do not take into account the fact that central banks dispose of a much richer information set) and to the failure to account for the considerable degree of model uncertainty that permeates this large information environment. Inserting these two characteristics leads, in fact, to the optimality of a policy behavior comparable with what happens in reality, without the need of placing a strong preference for smoothing in the loss function.

3.3 Monetary Policy Efficiency

We want now to analyze whether the consideration of an expanded information set leads to an improvement in efficiency for a central bank, taking decisions in the way we have described (thus, using all the available information and taking model uncertainty into account).

We suppose that efficiency is measured by a loss function of the common form:

$$LOSS = \lambda_{\pi} Var(\pi_t) + \lambda_y Var(y_t) + \lambda_{\Delta i} Var\left(i_t^* - i_{t-1}^*\right) \tag{18}$$

where λ_{π} represents the weight of inflation stabilization, $\lambda_y = (1 - \lambda_{\Delta i} - \lambda_{\pi})$ is the weight assigned to output stabilization and $\lambda_{\Delta i}$ is the interest rate variability penalty. The three weights sum to 1.

By varying λ_{π} from 0 to $1 - \lambda_{\Delta i}$, we can easily derive the efficiency frontier. We derive the loss at various points of the efficiency frontier, for given interest rate smoothing penalty preferences ($\lambda_{\Delta i}$). Our focus is on comparison between loss (efficiency) under our wide-information policy-making and those realized in the context of the RS and VAR specifications.

Insert Figure 2-3 here.

We report loss comparisons for four cases: $\lambda_{\Delta i} = 0$, $\lambda_{\Delta i} = 0.07$, $\lambda_{\Delta i} = 0.2$ and $\lambda_{\Delta i} = 0.5$. The red triangles represent the losses caused by the application of optimal policy rules under the RS and VAR models, respectively, the blue rounds regard our wider information case.

It seems evident that the losses under the traditional cases are always larger; thus, the exploitation of a wider information set, accounting for model uncertainty, leads to clear gains in macroeconomic efficiency.

3.4 Extension: Forward-looking Models

In our analysis, we have considered a backward-looking specification and we chose to compare our results to other backward-looking benchmark alternatives, to maintain the interpretation consistent. Our BMA methodology can be extended to forward-looking models as well. However a careful study of BMA with forward-looking models and large information goes beyond the scope of this paper. In this section, we try to infer if large information can still provide some attenuation in the context of a complete forward-looking DSGE model, such as the New Keynesian model, described for example in Woodford (2003b). Let's consider the following hybrid New Keynesian model:

$$y_t = (1-\delta)y_{t-1} + \delta E_t y_{t+1} - \widetilde{\sigma}^{-1}(i_t - E_t \pi_{t+1}) + g_t \tag{19}$$

$$\pi_t = \gamma_b \pi_{t-1} + \gamma_f E_t \pi_{t+1} + \kappa y_t + h_t \tag{20}$$

$$g_t = \mu g_{t-1} + \widetilde{g}_t$$

$$h_t = \rho h_{t-1} + \widetilde{h}_t$$

where $\delta = \frac{1}{1+\omega_c}$, $\widetilde{\sigma} = \frac{1+\omega_c}{1-\omega_c}\sigma$, $\kappa = \frac{(1-\omega_p)(1-\beta\theta)(1-\theta)(\varphi+\sigma)}{\theta+\omega_p(1-\theta(1-\beta))}$, $\gamma_b = \frac{\omega_p}{\theta+\omega_p(1-\theta(1-\beta))}$, $\gamma_f = \frac{\beta\theta}{\theta+\omega_p(1-\theta(1-\beta))}$, and ω_c is the fraction of rule-of-thumb agents in consumption, σ is the elasticity of intertemporal substitution, ω_p is the fraction of rule-of-thumb firms, β is the discount factor, θ is the Calvo parameter and φ

the elasticity of disutility of work. This model can be explicitly derived from microfoundations and has been developed by Amato and Laubach (2003). We calibrate the parameters to the following values: $\beta = 0.99$, $\mu = \rho = 0$, $\theta = 0.75$, $\varphi = 1$, $\sigma = 1$, $\omega_c = \omega_p = 0.5$, $g_t, h_t \sim N(0, 0.1)$.

We calculate the optimal policy under the case $\lambda_\pi = 0.93$, $\lambda_y = 0$, $\lambda_{\Delta i} = 0.07$, where we obtained realistic results under our large information model. Here, we end up again with an unrealistic series (std=12.83) for the optimal policy rates. The results are not due to the choice of calibrated values or the addition of backward-looking terms besides forward-looking expectations. Weights for $\lambda_{\Delta i}$ around 0.5 are needed to have volatility comparable to the empirical one. There is therefore room for improvement by adding larger information sets. We do not expect the direction of change to be different compared to the backward-looking results. Still the central bank will respond to many more variables and this is likely to lead to a smoother path. If the central bank was just responding to the target variables, it would be intuitively optimal to act aggressively in order to eliminate their deviations from steady state (leading to an unrealistically volatile optimal policy path).

4 Conclusions

In the literature, optimal monetary policy rules, derived in the framework of small macroeconomic models relating inflation, output gap and a short term interest rate, are much more aggressive and imply higher volatility of the policy instrument than what is observed in practice. In order to reconcile these rules with reality, it is necessary to assign a considerable weight to an interest smoothing objective in the central bank's loss function.

We have considered that the policy maker is not able, in practice, to perfectly observe the value of her target variables and, thus, she employs several leading indicators to get a more accurate feeling of the state of the economy. Therefore, we have allowed the central bank to use a wider information set, taking also into account the associated model uncertainty, implicit in the use of many additional variables.

The consideration of this larger and more realistic information set has been found to have an important impact on the optimal monetary policy. In more detail, the results indicate that it leads to a substantial attenuation of the optimal response of the instrument to disequilibrium and to a smoother interest rate path, if compared with the optimal solutions obtained under standard models of monetary transmission.

It remains to be understood if the obtained attenuation has, mainly, to be ascribed to the wider information set or to the consideration of model uncertainty.

If we consider the same technique we used, BMA, to take model uncertainty into account, but assuming that the uncertainty regards only the dynamic structure of the economy (without introducing leading indicators in the analysis), we find that policy attenuation is much smaller.

This seems to signal that the explicit introduction of a wider information set is indeed important and it is itself a cause of interest rate smoothness. We are also aware of the fact that model uncertainty is undoubtedly more important in such an environment, where the policy maker has to deal with a huge amount of diverse information (model uncertainty is likely to have less impact if the only uncertainty is about dynamic structure).

We think that the consideration of a larger information set can be a further explanation, at least a partial one, of real world monetary policy conservatism and interest rate smoothing, in addition to other proposed in the literature and which comprises uncertainty, in the form of parameter, model, or data uncertainty, and forward-looking behavior (besides the explicit introduction of a preference for smoothing in the central bank's loss function).

Also the presence of a penalty on interest rate variability is not worrying, as we can imagine how a very low penalty can be justifiable to avoid extreme and unrealistic jumps in the federal funds rate, while we do not think a much bigger weight to be reasonable.

It remains to be seen if the contemporaneous insertion of the different proposed explanations of policy conservatism and smoothing are able to explain the behavior of the funds rate without any further assumption in the central bank's loss function at all.

In addition, we have seen that allowing the central bank to exploit information coming from different indicators and economic variables, besides the traditional target variables, leads to clear gains in macroeconomic efficiency, no matter what the preference weights are.

In future research, it would be worth employing other techniques to account for this large information, like dynamic common factors, to check if our results are robust to different modeling choices, or using the Kalman Filter to assign optimal weights to the various indicators, as described in Svensson and Woodford (2003). Another possible result, that could change with the use of a richer information set, regards the common 'price puzzle' in VAR models, which can probably be eliminated in our new framework.

Finally, we hope that the use of BMA, in order to account for the information coming from different plausible models of the economy and to explicitly consider model uncertainty (which is, without any doubt, an extremely relevant and realistic characteristic in monetary policy-making), could begin to be used in the monetary policy literature. An important extension would be to use BMA with more structural models and incorporating forward-looking expectations as well[4].

[4] A first step in this direction is the recent work by Cogley and Sargent (2003).

References

[1] Amato, J.D., and T. Laubach, (2003). "Rule-of-thumb behaviour and monetary policy", *European Economic Review* 47-5, 791-831.

[2] Bernanke, B., and J. Boivin, (2003). "Monetary Policy in a Data-Rich Environment", *Journal of Monetary Economics* 50-3, 525-546.

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

[4] Brock, W.A., Durlauf, S.N., and K.D. West, (2003). "Policy Evaluation in Uncertain Economic Environments", *Brooking Papers on Economic Activity* 1.

[5] 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.

[6] Cogley, T., and T.J. Sargent (2003). "The Conquest of U.S. Inflation: Learning and Robustness to Model Uncertainty", mimeo, UC Davis and New York University.

[7] Favero, C.A., and F. Milani, (2001). "Parameter Instability, Model Uncertainty and the Choice of Monetary Policy", IGIER Working Paper No. 196, June.

[8] Fernández, C., E. Ley and M.F.J. Steel, (2001). "Model Uncertainty in Cross-Country Growth Regression", *Journal of Applied Econometrics*, Vol. 16, No. 5, pp. 563-576.

[9] Madigan, D., and J. York, (1995). "Bayesian Graphical Models for Discrete Data", *International Statistical Review* 63, 215-232.

[10] Milani, F., (2003). "Adaptive Learning, Model Uncertainty and Monetary Policy Inertia in a Large Information Environment", mimeo, Princeton University.

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

[12] Onatski, A., and N. Williams, (2003). "Modeling Model Uncertainty", *Journal of the European Economic Association* Vol. 1, Issue 5, 1087-1122.

[13] Orphanides, A., (2001). "Monetary Policy Rules Based on Real-Time data", *American Economic Review* 91, 964-985.

[14] Raftery, A.E., D. Madigan and J.A. Hoeting, (1997). "Bayesian Model Averaging for Linear Regression Models", *Journal of the American Statistical Association* 92, 179-191.

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

[16] Rudebusch, G.D., and L.E.O. Svensson, (1999). "Policy Rules for Inflation Targeting", in J. Taylor (Ed.), *Monetary Policy Rules*, pp. 203-246, Chicago: University of Chicago Press.

[17] Rudebusch, G.D., and L.E.O. Svensson, (2002). "Eurosystem Monetary Targeting: Lessons from US Data", *European Economic Review* 46 (3), pp. 417-442.

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

[19] Sack, B., and V. Wieland, (2000). "Interest-rate Smoothing and Optimal Monetary Policy. A Review of Recent Empirical Evidence", *Journal of Economics and Business* (52) 1-2, pp. 205-228.

[20] Sims, C., (2001). "Pitfalls of a Minimax Approach to Model Uncertainty", *American Economic Review* 91 (2), pp. 51-54.

[21] Sims, C., (2002). "The Role of Models and Probabilities in the Monetary Policy Process", *Brookings Papers on Economic Activity* 2, 1-62.

[22] Smith, A.F.M., and G.O. Roberts, (1993). "Bayesian Computation via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods (with discussion)", *Journal of the Royal Statistics Society* Series B, 55, 3-23.

[23] Svensson, L.E.O., and M. Woodford, (2003). "Indicator Variables for Optimal Policy", *Journal of Monetary Economics* 50 (3), 691-720.

[24] Tetlow, R., (2000). "The Fed Is Not As Ignorant As You Think", unpublished manuscript.

[25] Woodford, M., (2003a). "Optimal Interest-Rate Smoothing", *Review of Economic Studies*, Vol. 70 Issue 4, page 861.

[26] Woodford, M., (2003b). *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.

TABLE 1

Bayesian Model Averaging Estimates
Dependent Variable = inflation
R-squared = 0.9900
sigma^2 = 0.0688
Nobs, Nvars = 124, 60
ndraws = 51000
nu,lam,phi = 4.000, 0.250, 3.000
# of models = 31859
time(seconds) = 40602.7240
****

Model averaging information

Model	infl1	infl2	infl3	infl4	consconf1	consconf2	consconf3	consconf4	cpiinfl1	cpiinfl2	cpiinfl3	cpiinfl4	empl1	empl2	empl3	empl4
model 1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
model 2	1	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0
model 3	1	0	1	0	1	0	0	0	0	0	0	0	0	0	1	0
model 4	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
model 5	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
model 6	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
model 7	1	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
model 8	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
model 9	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
model 10	1	0	0	0	1	0	0	0	0	0	0	0	0	0	1	0

Model	housing1	housing2	housing3	housing4	invsales1	invsales2	invsales3	invsales4	m21	m22	m23	m24	napm1	napm2	napm3	napm4
model 1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 5	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 6	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 7	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Model	neword1	neword2	neword3	neword4	outgap1	outgap2	outgap3	outgap4	retail1	retail2	retail3	retail4	shipments1	shipments2	shipments3	shipments4
model 1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
model 2	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0
model 3	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0
model 4	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0
model 5	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
model 6	1	0	0	0	1	0	0	1	0	0	0	0	0	1	0	0
model 7	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
model 8	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0
model 9	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0
model 10	1	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0

Model	stock1	stock2	stock3	stock4	unford1	unford2	unford3	unford4	vehicles1	vehicles2	vehicles3	vehicles4	Prob	Visit
model 1	0	0	0	0	0	1	0	0	0	0	0	0	0.314	4
model 2	0	0	0	0	0	0	0	1	0	0	0	0	0.315	5
model 3	0	0	0	0	0	0	0	1	0	0	0	0	0.316	2
model 4	0	0	0	0	0	0	0	1	0	0	0	0	0.374	2
model 5	0	0	0	0	1	0	0	0	0	0	0	0	0.387	20
model 6	0	0	0	0	0	1	0	0	0	0	0	0	0.447	2
model 7	0	0	0	0	1	0	0	0	0	0	0	0	0.648	8
model 8	0	0	0	0	0	0	0	1	0	0	0	0	0.651	1
model 9	0	0	0	0	0	0	0	1	0	0	0	0	0.716	11
model 10	0	0	0	0	0	0	0	1	0	0	0	0	1.137	3

****

TABLE 2

BMA Posterior Estimates

Variable	Coefficient	t-statistic	t-probability
infl(-1)	0.90072	25.284134	0
infl(-2)	-0.028462	-0.340188	0.734291
infl(-3)	-0.01729	-0.349441	0.727351
infl(-4)	-0.003371	-0.094077	0.9252
consconf(-1)	-0.256817	-1.747004	0.083112
consconf(-2)	-0.022229	-0.145921	0.884221
consconf(-3)	-0.000188	-0.001445	0.99885
consconf(-4)	-0.007333	-0.04296	0.965802
cpiinfl(-1)	0.073531	2.233223	0.027328
cpiinfl(-2)	0.012653	0.364784	0.715894
cpiinfl(-3)	0.000674	0.020192	0.983922
cpiinfl(-4)	-0.000282	-0.00957	0.99238
empl(-1)	-0.001181	-0.024751	0.980293
empl(-2)	0.000982	0.025117	0.980002
empl(-3)	0.004174	0.158311	0.874469
empl(-4)	0.001432	0.061294	0.951223
housing(-1)	0.228758	1.291613	0.198893
housing(-2)	-0.001389	0.001355	0.998921
housing(-3)	0.013088	0.062499	0.950266
housing(-4)	-0.003882	-0.016008	0.987254
invsales(-1)	0.000022	0.00571	0.995453
invsales(-2)	0.095952	0.124392	0.901206
invsales(-3)	0.085339	0.121116	0.903795
invsales(-4)	0.285341	0.409081	0.683186
m2(-1)	0.001216	0.056305	0.955189
m2(-2)	-0.000754	-0.030573	0.975659
m2(-3)	-0.00002	-0.001843	0.998532
m2(-4)	0.000008	0.000754	0.9994
napm(-1)	0.001069	0.003022	0.997594
napm(-2)	0.00613	0.017596	0.985989
napm(-3)	0.010168	0.031783	0.974696
napm(-4)	0.00631	0.020248	0.983878
neword(-1)	0.030139	6.350585	0
neword(-2)	0.007328	0.764999	0.445726
neword(-3)	-0.000321	-0.070042	0.944273
neword(-4)	0.000682	0.151985	0.879446
outgap(-1)	-0.036333	-1.351101	0.179123
outgap(-2)	-0.001817	-0.060566	0.951803
outgap(-3)	0.000711	0.033042	0.973694
outgap(-4)	0.097277	5.054008	0.000002
retail(-1)	0.0001	0.006299	0.994984
retail(-2)	-0.002081	-0.155212	0.876907
retail(-3)	-0.001331	-0.108722	0.913598
retail(-4)	0.000364	0.031169	0.975185
shipments(-1)	-0.000151	-0.007256	0.994222
shipments(-2)	-0.014203	-1.323847	0.187989
shipments(-3)	-0.000039	-0.007378	0.994125
shipments(-4)	-0.000056	-0.004794	0.996182
stock(-1)	-0.000099	-0.045489	0.963791
stock(-2)	-0.000234	-0.104387	0.917031
stock(-3)	-0.000007	-0.003296	0.997375
stock(-4)	0.000005	0.002234	0.998221
unford(-1)	0.00531	1.006805	0.315989
unford(-2)	0.002907	0.572049	0.568324
unford(-3)	-0.00127	-0.100061	0.920458
unford(-4)	0.003913	0.740212	0.460571
vehicles(-1)	-0.000061	-0.019881	0.98417
vehicles(-2)	-0.000181	-0.06727	0.946475
vehicles(-3)	-0.000248	-0.098569	0.92164
vehicles(-4)	0.000183	0.075855	0.939656

Notes: The posterior estimates of the coefficients are obtained by averaging over the whole set of models, with weights given by the respective posterior model probabilities. A regressor which is not usually included in the selected models is assigned a near zero coefficient with high p-value.

TABLE 3

Bayesian Model Averaging Estimates

Dependent Variable = output gap
R-squared = 0.9333
sigma^2 = 0.3980
Nobs, Nvars = 124, 60
ndraws = 51000
nu,lam,phi = 4.000, 0.250, 3.000
# of models = 15870
time(seconds) = 23946.8310
****

Model averaging information

Model	outgap1	outgap2	outgap3	outgap4	consconf1	consconf2	consconf3	consconf4	cpiinfl1	cpiinfl2	cpiinfl3	cpiinfl4	empl1	empl2	empl3	empl4
model 1	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
model 2	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
model 3	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
model 4	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
model 5	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0
model 6	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
model 7	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
model 8	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
model 9	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
model 10	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0

Model	housing1	housing2	housing3	housing4	invsales1	invsales2	invsales3	invsales4	m21	m22	m23	m24	napm1	napm2	napm3	napm4
model 1	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
model 2	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
model 3	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
model 4	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
model 5	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
model 6	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
model 7	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
model 8	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
model 9	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
model 10	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0

Model	neword1	neword2	neword3	neword4	retail1	retail2	retail3	retail4	shipments1	shipments2	shipments3	shipments4	stock1	stock2	stock3	stock4
model 1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 2	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
model 3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 5	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
model 6	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
model 7	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
model 8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
model 9	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
model 10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

Model	unford1	unford2	unford3	unford4	vehicles1	vehicles2	vehicles3	vehicles4	reals1	reals2	reals3	reals4	Prob	Visit
model 1	0	0	0	0	0	0	0	0	0	1	1	1	0.756	1
model 2	0	0	0	0	0	0	0	0	0	1	1	0	0.81	1
model 3	0	0	0	0	0	0	0	0	1	0	0	0	0.89	71
model 4	0	0	0	0	0	0	0	0	0	1	1	0	1.108	25
model 5	0	0	0	0	0	0	0	0	0	1	0	0	1.129	6
model 6	0	0	0	0	0	0	0	0	1	0	0	0	1.222	25
model 7	0	0	0	0	0	0	0	0	0	1	1	0	1.281	2
model 8	0	0	0	0	0	0	0	0	0	1	0	0	1.396	92
model 9	0	0	0	0	0	0	0	0	0	1	0	0	1.667	24
model 10	0	0	0	0	0	0	0	0	0	1	0	0	2.399	105

****

TABLE 4

BMA Posterior Estimates

Variable	Coefficient	t-statistic	t-probability
outgap(-1)	0.660132	14.391195	0
outgap(-2)	0.010224	0.126782	0.899319
outgap(-3)	0.001629	0.029052	0.97687
outgap(-4)	0.000863	0.019407	0.984547
consconf(-1)	1.766744	4.560447	0.000012
consconf(-2)	0.007527	0.015301	0.987816
consconf(-3)	0.006621	0.014715	0.988283
consconf(-4)	0.003451	0.008739	0.993041
cpiinfl(-1)	-0.002819	-0.100191	0.920354
cpiinfl(-2)	-0.000951	-0.041875	0.966666
cpiinfl(-3)	-0.000592	-0.02999	0.976123
cpiinfl(-4)	-0.00133	-0.050583	0.959739
empl(-1)	-0.001872	-0.034859	0.972248
empl(-2)	-0.002375	-0.046852	0.962707
empl(-3)	-0.001212	-0.025841	0.979426
empl(-4)	-0.000846	-0.019791	0.984242
housing(-1)	2.423494	5.683104	0
housing(-2)	-0.046673	-0.056749	0.954837
housing(-3)	-0.011626	-0.01971	0.984306
housing(-4)	-0.000101	-0.000587	0.999533
invsales(-1)	-0.190449	-0.106462	0.915388
invsales(-2)	-0.033334	-0.024089	0.98082
invsales(-3)	0.004499	-0.002561	0.99796
invsales(-4)	-0.01603	-0.010195	0.991882
m2(-1)	-0.038681	-1.612485	0.109399
m2(-2)	-0.017002	-0.665283	0.507105
m2(-3)	-0.000298	-0.048419	0.96146
m2(-4)	0.000564	0.010192	0.991884
napm(-1)	0.011442	0.016878	0.986561
napm(-2)	-0.014369	-0.023563	0.981239
napm(-3)	-0.009396	-0.016689	0.986711
napm(-4)	-0.003664	-0.007216	0.994254
neword(-1)	-0.000066	-0.007338	0.994157
neword(-2)	-0.000134	-0.016149	0.987141
neword(-3)	-0.000092	-0.011525	0.990823
neword(-4)	-0.000009	-0.001566	0.998753
retail(-1)	-0.000578	-0.020487	0.983688
retail(-2)	-0.000207	-0.007904	0.993707
retail(-3)	-0.000361	-0.014407	0.988529
retail(-4)	-0.00066	-0.02545	0.979737
shipments(-1)	-0.000166	-0.013949	0.988893
shipments(-2)	-0.000125	-0.011315	0.99099
shipments(-3)	-0.000038	-0.004012	0.996805
shipments(-4)	-0.000016	-0.001879	0.998503
stock(-1)	0.003072	0.667899	0.50544
stock(-2)	0.000617	0.135188	0.892682
stock(-3)	0.000155	0.034039	0.972901
stock(-4)	0.000111	0.024429	0.980549
unford(-1)	-0.000083	-0.010245	0.991842
unford(-2)	-0.000124	-0.015038	0.988026
unford(-3)	-0.000058	-0.007699	0.99387
unford(-4)	-0.000006	-0.001497	0.998808
vehicles(-1)	-0.000026	-0.004229	0.996633
vehicles(-2)	-0.00001	-0.001786	0.998578
vehicles(-3)	-0.000009	-0.001543	0.998771
vehicles(-4)	0.000147	0.023446	0.981332
reals(-1)	-0.018245	-0.693037	0.489582
reals(-2)	-0.078001	-2.416551	0.017125
reals(-3)	0.018494	0.311226	0.756152
reals(-4)	-0.006654	-0.202006	0.840243

Notes: The posterior estimates of the coefficients are obtained by averaging over the whole set of models, with weights given by the respective posterior model probabilities. A regressor which is not usually included in the selected models is assigned a near zero coefficient with high p-value.

TABLE 5 - Seemingly Unrelated Regression (SUR)

BMA Posterior Estimates (SUR)

Aggregate Supply equation

Variable	Coefficient	t-statistic	t-probability
infl(-1)	0.901393	26.24134	0
infl(-2)	-0.027743	-0.344465	0.73108
infl(-3)	-0.017191	-0.360359	0.719192
infl(-4)	-0.00328	-0.095327	0.924209
consconf(-1)	-0.258598	-1.822802	0.070743
consconf(-2)	-0.022566	-0.153205	0.878486
consconf(-3)	-0.000232	-0.001753	0.998604
consconf(-4)	-0.007517	-0.045714	0.963612
cpiinfl(-1)	0.07191	2.268482	0.025029
cpiinfl(-2)	0.012625	0.377492	0.706453
cpiinfl(-3)	0.000722	0.022316	0.982232
cpiinfl(-4)	-0.000277	-0.009707	0.992271
empl(-1)	-0.001183	-0.026087	0.97923
empl(-2)	0.000932	0.024491	0.9805
empl(-3)	0.004044	0.158947	0.873969
empl(-4)	0.001364	0.060511	0.951846
housing(-1)	0.226209	1.324441	0.187793
housing(-2)	-0.001766	-0.000071	0.999943
housing(-3)	0.012649	0.062646	0.95015
housing(-4)	-0.003867	-0.016647	0.986745
invsales(-1)	-0.001798	0.004045	0.996779
invsales(-2)	0.093031	0.124667	0.900989
invsales(-3)	0.084389	0.12384	0.901643
invsales(-4)	0.28404	0.420986	0.674494
m2(-1)	0.001197	0.057813	0.95399
m2(-2)	-0.000743	-0.031439	0.97497
m2(-3)	-0.000018	-0.001709	0.998639
m2(-4)	0.000011	0.001071	0.999147
napm(-1)	0.001463	0.004333	0.996549
napm(-2)	0.005961	0.017782	0.985841
napm(-3)	0.009853	0.032035	0.974496
napm(-4)	0.006146	0.020433	0.983731
neword(-1)	0.030387	6.643915	0
neword(-2)	0.007094	0.769392	0.443124
neword(-3)	-0.000325	-0.073558	0.941481
neword(-4)	0.00068	0.157702	0.874948
outgap(-1)	-0.036455	-1.40683	0.161979
outgap(-2)	-0.001779	-0.06155	0.95102
outgap(-3)	0.000729	0.035037	0.972106
outgap(-4)	0.098454	5.307887	0
retail(-1)	0.000104	0.006825	0.994566
retail(-2)	-0.002038	-0.157878	0.87481
retail(-3)	-0.00132	-0.112151	0.910885
retail(-4)	0.000355	0.031647	0.974804
shipments(-1)	-0.000146	-0.007281	0.994202
shipments(-2)	-0.013912	-1.348229	0.180042
shipments(-3)	-0.000036	-0.007125	0.994327
shipments(-4)	-0.000055	-0.004852	0.996137
stock(-1)	-0.000094	-0.044878	0.964276
stock(-2)	-0.000227	-0.105099	0.916467
stock(-3)	-0.000007	-0.003155	0.997488
stock(-4)	0.000005	0.002498	0.998011
unford(-1)	0.005205	1.021683	0.30892
unford(-2)	0.002808	0.576278	0.565472
unford(-3)	-0.00129	-0.105703	0.915989
unford(-4)	0.00393	0.769187	0.443245
vehicles(-1)	-0.000062	-0.020831	0.983414
vehicles(-2)	-0.000181	-0.069955	0.944342
vehicles(-3)	-0.000245	-0.101441	0.919364
vehicles(-4)	0.000189	0.081587	0.935106

Aggregate Demand equation

Variable	Coefficient	t-statistic	t-probability
outgap(-1)	0.658511	14.741847	0
outgap(-2)	0.010361	0.132562	0.894755
outgap(-3)	0.001703	0.03127	0.975104
outgap(-4)	0.000901	0.020861	0.98339
consconf(-1)	1.770315	4.692429	0.000007
consconf(-2)	0.007347	0.015427	0.987716
consconf(-3)	0.006794	0.015541	0.987626
consconf(-4)	0.003438	0.008971	0.992857
cpiinfl(-1)	-0.002783	-0.101448	0.919358
cpiinfl(-2)	-0.000931	-0.042211	0.966398
cpiinfl(-3)	-0.000587	-0.030436	0.975768
cpiinfl(-4)	-0.001329	-0.051731	0.958826
empl(-1)	-0.001928	-0.036891	0.970631
empl(-2)	-0.002389	-0.048436	0.961447
empl(-3)	-0.001204	-0.02639	0.978989
empl(-4)	-0.000839	-0.020175	0.983936
housing(-1)	2.442067	5.88123	0
housing(-2)	-0.048643	-0.060759	0.951649
housing(-3)	-0.011578	-0.020137	0.983967
housing(-4)	-0.000211	-0.000815	0.999351
invsales(-1)	-0.195799	-0.112525	0.910589
invsales(-2)	-0.035049	-0.025833	0.979432
invsales(-3)	0.00428	-0.002871	0.997714
invsales(-4)	-0.016488	-0.010774	0.991421
m2(-1)	-0.038699	-1.65984	0.099474
m2(-2)	-0.017228	-0.691801	0.490356
m2(-3)	-0.000335	-0.050817	0.959553
m2(-4)	0.000565	0.010533	0.991613
napm(-1)	0.011241	0.017019	0.986449
napm(-2)	-0.014227	-0.023993	0.980897
napm(-3)	-0.009234	-0.016869	0.986568
napm(-4)	-0.003582	-0.00726	0.994219
neword(-1)	-0.000069	-0.007777	0.993807
neword(-2)	-0.000134	-0.016688	0.986712
neword(-3)	-0.000092	-0.011882	0.990539
neword(-4)	-0.000008	-0.001443	0.998851
retail(-1)	-0.000571	-0.020752	0.983477
retail(-2)	-0.000218	-0.008521	0.993215
retail(-3)	-0.00037	-0.015129	0.987954
retail(-4)	-0.000652	-0.025796	0.979461
shipments(-1)	-0.000171	-0.014748	0.988257
shipments(-2)	-0.000125	-0.011631	0.990739
shipments(-3)	-0.000038	-0.004146	0.996699
shipments(-4)	-0.000017	-0.001988	0.998417
stock(-1)	0.00304	0.680234	0.497624
stock(-2)	0.000609	0.137356	0.890972
stock(-3)	0.000158	0.035648	0.971621
stock(-4)	0.000119	0.027029	0.97848
unford(-1)	-0.000083	-0.010401	0.991718
unford(-2)	-0.000123	-0.015293	0.987823
unford(-3)	-0.000056	-0.007727	0.993847
unford(-4)	-0.000005	-0.001391	0.998892
vehicles(-1)	-0.000029	-0.004877	0.996116
vehicles(-2)	-0.000013	-0.002335	0.998141
vehicles(-3)	-0.000011	-0.001946	0.998451
vehicles(-4)	0.00015	0.024699	0.980335
reals(-1)	-0.018252	-0.712239	0.477655
reals(-2)	-0.078419	-2.495721	0.013884
reals(-3)	0.018652	0.323667	0.746736
reals(-4)	-0.006592	-0.206002	0.837127

	FEDERAL FUNDS RATE PATHS			
$\lambda_\pi = 0.93, \lambda_y = 0, \lambda_{\Delta i} = 0.07$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	7.41	7.13	6.81	7.34
STD.	3.40	9.77	14.97	3.15
PERSISTENCE	0.988	0.944	0.935	0.934
$\lambda_\pi = 0.465, \lambda_y = 0.465, \lambda_{\Delta i} = 0.07$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	7.39	7.11	6.73	7.34
STD.	2.29	8.05	12.39	3.15
PERSISTENCE	0.981	0.92	0.91	0.934
$\lambda_\pi = 0, \lambda_y = 0.93, \lambda_{\Delta i} = 0.07$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	7.37	7.13	6.71	7.34
STD.	0.97	6.10	11.24	3.15
PERSISTENCE	0.969	0.882	0.897	0.934
$\lambda_\pi = 0.8, \lambda_y = 0, \lambda_{\Delta i} = 0.2$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	7.39	7.21	6.99	7.34
STD.	1.30	6.59	9,19	3.15
PERSISTENCE	0.98	0.94	0.928	0.934
$\lambda_\pi = 0.4, \lambda_y = 0.4, \lambda_{\Delta i} = 0.2$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	7.39	7.21	6.92	7.34
STD.	0.83	5.64	8.43	3.15
PERSISTENCE	0.981	0.927	0.907	0.934
$\lambda_\pi = 0, \lambda_y = 0.8, \lambda_{\Delta i} = 0.2$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	7.37	7.22	6.89	7.34
STD.	0.32	4.39	8.04	3.15
PERSISTENCE	0.97	0.90	0.893	0.934
$\lambda_\pi = 0.5, \lambda_y = 0, \lambda_{\Delta i} = 0.5$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	7.39	7.25	7.12	7.34
STD.	0.37	4.76	5.8	3.15
PERSISTENCE	0.981	0.937	0.928	0.934
$\lambda_\pi = 0.25, \lambda_y = 0.25, \lambda_{\Delta i} = 0.5$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	7.39	7.26	7.08	7.34
STD.	0.23	4.35	5.51	3.15
PERSISTENCE	0.981	0.932	0.907	0.934
$\lambda_\pi = 0, \lambda_y = 0.5, \lambda_{\Delta i} = 0.5$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	7.38	7.26	7.05	7.34
STD.	0.08	3.86	5.32	3.15
PERSISTENCE	0.969	0.925	0.894	0.934
$\lambda_\pi = 1, \lambda_y = 0, \lambda_{\Delta i} = 0$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	12.03	0.36	5.82	7.34
STD.	469.8	649	63.1	3.15
PERSISTENCE	0.983	0.969	0.952	0.934
$\lambda_\pi = 0.5, \lambda_y = 0.5, \lambda_{\Delta i} = 0$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	7.62	5.75	6.33	7.34
STD.	41.32	41.71	21.5	3.15
PERSISTENCE	0.985	0.912	0.917	0.934
$\lambda_\pi = 0, \lambda_y = 1, \lambda_{\Delta i} = 0$	Optimal FF (BMA)	Optimal FF (RS)	Optimal FF (VAR)	Actual FF
MEAN	7.22	6.05	6.43	7.34
STD.	16.68	31.42	16.48	3.15
PERSISTENCE	0.971	0.893	0.901	0.934

Table 6 - Optimal and actual federal funds rate paths (BMA, RS, information-augmented VAR).

	Infl.	*Infl.(-1)*	*Infl.(-2)*	*Infl.(-3)*	*Cons.Conf.*	*Cons.Conf.(-1)*	*Cons.Conf.(-2)*	*Cons.Conf.(-3)*
FEEDBACK COEFF.	0.429	0.021	0.025	-0.002	-1.65	-0.847	-0.573	-0.013

CPI infl.	*CPI infl.(-1)*	*CPI infl.(-2)*	*CPI infl.(-3)*	*Empl.*	*Empl.(-1)*	*Empl.(-2)*	*Empl.(-3)*
0.866	-0.107	-0.036	-0.025	0.196	-0.055	0.011	0.002

Hous.	*Hous.(-1)*	*Hous.(-2)*	*Hous.(-3)*	*Inv/Sales*	*Inv/Sales(-1)*	*Inv/Sales(-2)*	*Inv/Sales(-3)*
0.089	-0.007	-0.0002	0.0005	0.486	0.002	0.001	-0.096

M2	*M2(-1)*	*M2(-2)*	*M2(-3)*	*NAPM*	*NAPM(-1)*	*NAPM(-2)*	*NAPM(-3)*
-0.056	0.004	-0.006	-0.001	0.069	0.021	0.021	-0.014

New Ord.	*New Ord.(-1)*	*New Ord.(-2)*	*New Ord.(-3)*	*Outgap*	*Outgap(-1)*	*Outgap(-2)*	*Outgap(-3)*
0.060	0.044	0.044	0.050	0.272	0.069	0.063	0.019

Retail	*Retail(-1)*	*Retail(-2)*	*Retail(-3)*	*Shipm.*	*Shipm.(-1)*	*Shipm.(-2)*	*Shipm.(-3)*
0.021	0.004	0.006	0.007	-0.050	-0.068	-0.045	-0.049

Stock	*Stock(-1)*	*Stock(-2)*	*Stock(-3)*	*Unf.Ord.*	*Unf.Ord.(-1)*	*Unf.Ord.*(-2)	*Unf.Ord.(-3)*
-0.006	-0.001	-0.001	0.0002	-0.169	-0.004	0.022	0.032

Vehicles	*Vehicles(-1)*	*Vehicles(-2)*	*Vehicles(-3)*	*Fed Funds(-1)*	*Fed Funds(-2)*	*Fed Funds(-3)*
0.005	-0.003	-0.004	-0.004	-0.007	0.003	-0.0008

Table 7 - Optimal Reaction Function: wider information set (case $\lambda_\pi = 0.93, \lambda_y = 0, \lambda_{\Delta i} = 0.07$)

	Infl	*Infl(-1)*	*Infl(-2)*	*Infl(-3)*	*Outgap*	*Outgap(-1)*
FEEDBACK COEFF.	7.02	-2.95	-0.63	0.45	2.70	-0.73

Table 8 - Optimal Reaction Function: RS model (case $\lambda_\pi = 0.93, \lambda_y = 0, \lambda_{\Delta i} = 0.07$)

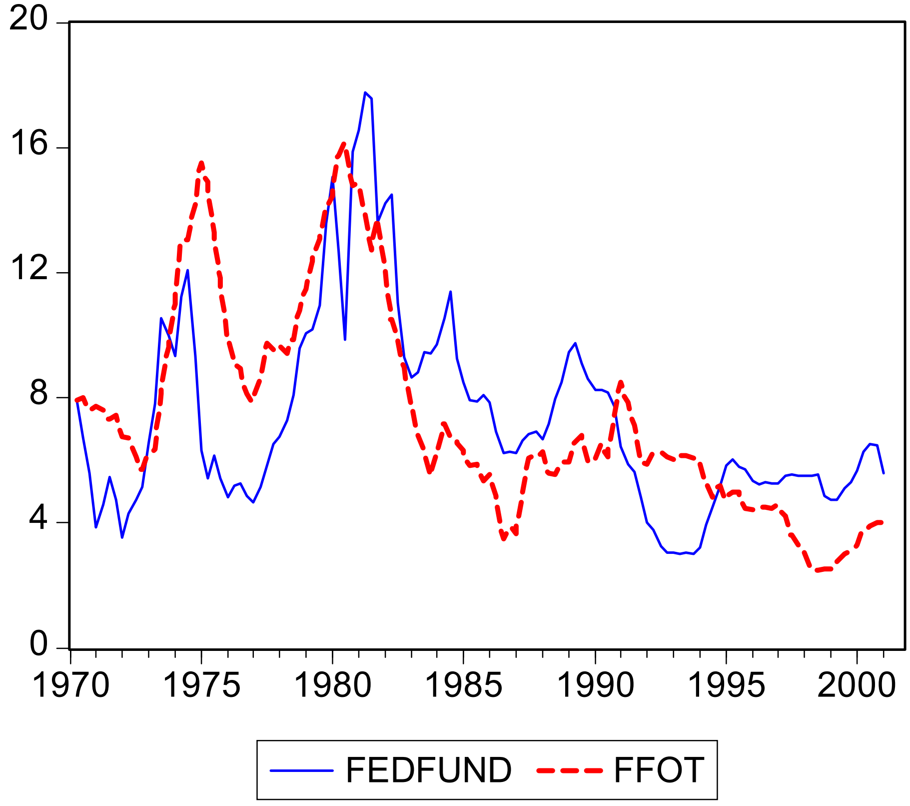

Figure 1 - Actual (FEDFUND) and optimal (FFOT) federal funds rate paths.

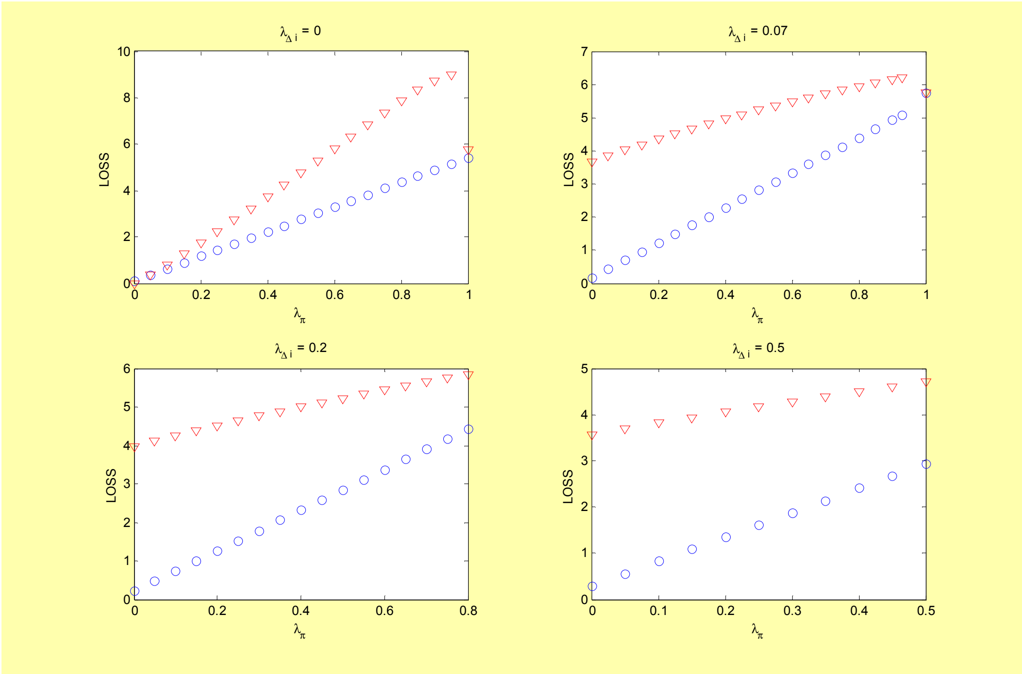

Figure 2 - Monetary Policy Efficiency (BMA vs. RS)

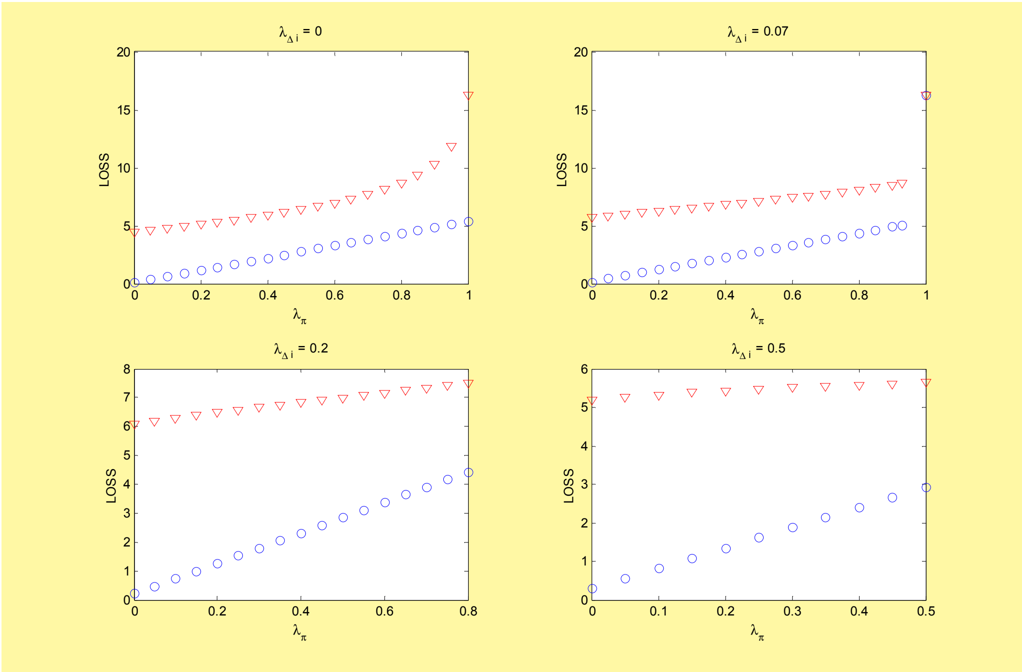

Figure 3 - Monetary Policy Efficiency (BMA vs. VAR)

A Data Appendix

The leading indicators we have incorporated in the central bank's information set (in addition to inflation, output gap and federal funds rate) are:

- Consumer Price Index
- Employment
- Housing Starts
- Inventory/Sales ratio
- Money Supply (M2)
- Consumer Confidence
- NAPM (National Association of Purchasing Managers) survey
- New Orders of Durable Goods
- Retail Sales
- Shipments of Durable Goods
- Stock Market
- Unfilled Orders of Durable Goods
- Vehicles' Sales

All the data are quarterly, from 1969:02 to 2001:01, and taken from FRED, the database of the Federal Reserve Bank of Saint Louis, or DATASTREAM..

Variables	Code	Description	Source
INFL.	GDPDEF	GDP: IMPLICIT PRICE DEFLATOR 1996=100, SA	FRED
OUTGAP	GDPC1	REAL GDP BILLIONS OF CHAINED 1996 DOLLARS, SA	FRED
	GDPPOT	REAL POTENTIAL GDP BILLIONS OF CHAINED 1996 DOLLARS	FRED
CONS.CONF.	USCNFCONQ	US CONSUMER CONFIDENCE: THE CONFERENCE BOARD'S INDEX FOR US SADJ	DATASTREAM
CPI INFL.	USCP....F	US CPI, ALL URBAN SAMPLE: ALL ITEMS NADJ	DATASTREAM
EMPL.	USEMPNAGE	US EMPLOYED - NONFARM INDUSTRIES TOTAL (PAYROLL SURVEY) VOLA	DATASTREAM
HOUSING	USPVHOUSE	US NEW PRIVATE HOUSING UNITS STARTED (ANNUAL RATE) VOLA	DATASTREAM
INV/SALES	USBSINVLB	US TOTAL BUSINESS INVENTORIES (END PERIOD LEVEL) CURA	DATASTREAM
	USBSSALEB	US TOTAL BUSINESS SALES CURA	DATASTREAM
M2	USM2....B	US MONEY SUPPLY M2 CURA	DATASTREAM
NAPM	USCNFBUSQ	US NATIONAL ASSN OF PURCHASING MANAGEMENT INDEX(MFG SURVEY) SADJ	DATASTREAM
NEW ORD.	USNODURBB	US NEW ORDERS FOR DURABLE GOODS INDUSTRIES(DISC.) CURA	DATASTREAM
RETAIL SAL.	USRETTOTB	US TOTAL VALUE OF RETAIL SALES CURA	DATASTREAM
SHIPMENTS	USSHDURGB	US SHIPMENTS OF DURABLE GOODS(DISC.) CURA	DATASTREAM
STOCK IND.	US500STK	US STANDARD & POOR'S INDEX OF 500 COMMON STOCKS(MONTHLY AVE)	DATASTREAM
UNF. ORD.	USUODURBB	US UNFILLED ORDERS FOR DURABLE GOODS(DISC.) CURA	DATASTREAM
VEHICLES	USPCARRSF	US NEW PASSENGER CARS-RETAIL SALES: TOTAL VEHICLES NADJ	DATASTREAM
FED. FUNDS	USFEDFUN	US FEDERAL FUNDS RATE	DATASTREAM

Inflation has been calculated as $(\log(p_t) - \log(p_{t-4})) * 100$, output gap as $(\log(y_t) - \log(y^*)) * 100$. For all the non-stationary series, we have considered their annual growth rates.