Adaptive Learning, Model Uncertainty and Monetary Policy Inertia in a Large Information Environment*

Fabio Milani†
Princeton University

First draft: February 13, 2003
This version: May 27, 2003
(Preliminary and Incomplete.
Comments are Welcome!)

Abstract

This paper studies monetary policy in a large information environment. We allow the central bank to exploit information coming from several indicator variables. Model uncertainty is extremely pervasive in this large information world. We take it into account, by estimating the whole set of possible models through Bayesian Model Averaging, implemented with a Markov Chain Monte Carlo Model Composition (MC^3) strategy. The central bank can learn parameters and models through time. In this setting, we aim at evaluating if the introduction of a large information set can, at least partially, explain the optimality of observed monetary policy inertia.

Keywords: optimal monetary policy, Bayesian Model Averaging, leading indicators, model uncertainty, adaptive learning, interest-rate smoothing, inertia.

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

---

*Preliminary and incomplete. Comments are welcome. Updated versions of the paper will be found at: http://www.princeton.edu/~fmilani/research.html

†*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.

1 Introduction.

The analysis of monetary policy has been typically confined to the use of small macroeconomic models, in which the central bank is implicitly allowed to exploit a very limited amount of information. Widely used frameworks for the empirical analysis of monetary policy are Rudebusch and Svensson (1999, 2002) and Clarida, Gali' and Gertler (2000). Models of this sort typically include just three variables: inflation, a measure of the output gap and the short-term interest rate (policy instrument).

Another large strand of literature has, instead, used VAR models to study monetary policy issues. VARs can be seen as a general unrestricted form to model the economy. We can then apply theoretical restrictions to arrive at small structural macroeconomic models. However, also within this approach, researchers have usually employed small VARs, consisting of few variables: inflation, output gap, federal funds rate and, sometimes, money or commodity prices. They were still characterized by the implicit assumption that the policy maker was able to exploit only a limited amount of information.

This assumption is clearly unrealistic. Central banks monitor a large number of economic data and leading indicators. They need to do it since policy affects target variables with long and uncertain lags. Therefore, they are in the situation of predicting forward looking variables without perfectly knowing the current stance of the economy. The use of several indicators, linked with future developments of the target variables, becomes helpful. A discussion of the reasons behind policy makers' use of many indicators is Kozicki (2001).

Modeling an enlarged information set is probably worthwhile, if we want to provide a more accurate characterization of the central banks' decision environment in the real world.

The pioneering work in this attempt to consider monetary policy in a data-rich environment is Bernanke and Boivin (2001), followed by Bernanke, Boivin and Eliasz (2002). They consider a very large data set and add common factors to a standard monetary VAR to account for the larger degree of information available in real-time to the central bank.

Our paper, although with different techniques, tries to contribute to this young literature. Focusing on the US case, it delivers an attempt to provide a more realistic approximation of how the Fed behaves and to verify how the expansion of the available information set affects optimal monetary policy, comparing the outcomes with limited information counterparts.

In particular, the focus of this paper is on an important unresolved issue in the monetary policy literature: the reconciliation of observed interest-rate smoothing and policy conservatism, with optimizing behavior in the context of theoretical models.

Usually, in fact, dynamic optimization techniques, applied to small macroeconomic models, dictate the optimality of a very aggressive monetary policy rule, leading to an extremely high volatility of the policy instrument. In the reality, instead, a completely different behavior is observed: real-world monetary policy is characterized by strong gradualism and interest-rate smoothing. This puzzle has generated a fertile stream of research on the topic.

In the literature, different explanations, leading to the optimality of policy gradualism, has been identified (an interesting survey is 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, in fact, 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 (1999).

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. An interesting 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).

Rudebusch (2002), instead, takes an alternative view, treating monetary policy inertia as an illusion, due to highly serially correlated unforecastable shocks or measurement error that central banks have to face (e.g., a lasting misperception of the actual level of potential output).

In this paper, we propose and evaluate a different explanation of the observed policy inertia:

4. *Larger Information Set.* We deem that introducing some elements of realism in the analysis, such as the possibility to exploit a larger information set for the central bank (taking into account also the associated model uncertainty, deriving from this data-rich environment), leads to partially solve the interest-rate smoothing puzzle. In practice, extremely aggressive rules are due to a misspecification of the standard

models, which omit many relevant information, indeed used in policy formulation.

In our framework, we allow the central bank to take into account a variety of data. As we focus on the US case, these additional variables are the leading indicators, explicitly rendered public and used by the Federal Reserve in policy formulation.

As the information is quite diverse, the policy maker has to face a considerable degree of model uncertainty and needs to recognize which indicators are more accurate predictors of future inflation and real activity.

To take the pervasive model uncertainty, associated with such a large information environment, into account, we employ Bayesian Model Averaging (BMA) with Markov Chain Monte Carlo Model Composition (MC^3). This technique has recently begun to be used in economics studies (for example, to find robust growth determinants, as in Fernandez, Ley and Steel (2001)); to my knowledge, it has not received application in optimal monetary policy so far.

The procedure will be described in detail in the next section. 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, which assumes 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). A recent alternative approach to model monetary policy under uncertainty comes from the proposal of "thick" modeling, as in Favero and Milani (2001). They recursively estimate several possible models, generated by different combinations of the included regressors, and 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.

A recent interesting effort to provide a general treatment of the problem of modeling model uncertainty is Onatski and Williams (2003).

We try to mimic the decision problem of the Fed in real time. In each period, we estimate the economy equations by BMA, taking thus model uncertainty into account, and we obtain the optimal policy rule. In the next period, we repeat the same exercise re-estimating the system and obtaining a new policy rule. We consider adaptive learning between periods, as described for example in Evans and Honkapohja (2001). An alternative could have been the use of Bayesian learning, through Kalman Filter. A more challenging possibility would be to jointly consider estimation and control, and allowing active experimentation by the central bank. Active experimentation is not examined in the present work.

Learning is quite interesting in this setting, since not only it permits to update the coefficients over time, but also, through the use of BMA, it permits to learn the correct model of the economy. This represents a form of learning in misspecified models, which can be expressed in a particularly easy form.

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 aims at explaining, at least partially, 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 additional information and model uncertainty into consideration in the modeling of optimal monetary policy-making.

The rest of the paper is organized as follows. Section 2 introduces the methodology we use, describing adaptive learning and Bayesian Model Averaging estimation. In Section 3, we present some evidence on our estimation results. Section 4 analyzes monetary policy in a large information environment, focusing on policy gradualism (comparing results with limited information settings) and on the excess response of policy to additional information. Section 5 concludes and discusses possible extensions of research.

2 Methodology.

In this section, we describe the techniques we will use in the rest of the paper. In each period, we mimic the decision problem of a central bank, which needs to obtain estimates of the state of the economy and compute optimal policy. We assume that the policy maker monitors several variables and leading indicators. Therefore, she faces a high degree of model uncertainty, since she has to decide what weight to assign to each variable. At every point in time, the central bank re-estimates the equations governing the economy: the parameters are time-varying and also the most likely models of the economy are allowed to vary over time. The policy maker learns both the model coefficients and the model itself by adaptive learning.

2.1 Adaptive Learning.

We seek to mimic the policy maker problem in real-time. In every period, the central bank tries to obtain an estimate of its target variables, inflation and output gap; to do this, a large number of potentially relevant leading indicators are employed.

As we recursively estimate the economy, we need to allow some learning mechanism between periods. We choose adaptive learning, which has already been the focus of a large number of works in the learning literature.

Let's briefly describe it in the simplest case, following Evans and Honkapohja (2001). Suppose that inflation behaves as:

$$\pi_t = \mu + \alpha \pi_t^e + \delta w_{t-1} + \eta_t, \tag{1}$$

where w_{t-1} is an exogenous variable. Suppose that agents in this economy act as econometricians and form expectations based on estimates of a statistical model. Their forecasts are given by:

$$\pi_t^e = a_{t-1} + b_{t-1} w_{t-1}, \tag{2}$$

where the coefficients are updated every period, as new information becomes available. Under Recursive Least Square (RLS) learning, agents run a least squares regression of π_t on z_{t-1}, where $z_t = (1, w_t)$. Then, the parameters' vectors $\phi_t = (a_t, b_t)'$ is recursively calculated as:

$$\phi_t = \phi_{t-1} + t^{-1} R_t^{-1} z_{t-1} (\pi_t - \phi_{t-1}' z_{t-1}), \tag{3}$$

$$R_t = R_{t-1} + t^{-1} (z_{t-1} z_{t-1}' - R_{t-1}), \tag{4}$$

where R_t is the moment matrix. The perceived law of motion (PLM) is given by:

$$\pi_t = (\mu + \alpha a_{t-1}) + (\delta + \alpha b_{t-1}) w_{t-1} + \eta_t, \tag{5}$$

and it is different from (1), the actual law of motion (ALM). Evans and Honkapohja (2001) discusses E-stability and prove that $\phi_t \longrightarrow \overline{\phi}$, as $t \longrightarrow \infty$.

An alternative to adaptive learning would have been the use of Bayesian learning, through the use of Kalman Filter. We believe that our results would not be considerably different. A more radical alternative, would be to allow active experimentation for the central bank. In this case, optimal control and estimation cannot be considered disjointly. The dynamic optimization problem would be given by the minimization of a standard quadratic loss function, subject to the linear equations, describing the evolution of the economy, and the nonlinear updating equations, representing the dynamics of beliefs. In this context, the decision maker can experiment, by changing the policy instrument, to learn the structure of the economy. This experiment is costly, so there will be an optimal level of experimentation. An evaluation of the impact of experimentation on policy can be found in Beck and Wieland (2002), among others. In the current paper, we rule out active experimentation, as our aim is to focus on the potential for a large information set to explain observed policy smoothness. This is not really a paper on learning and we just bound ourselves to use the relative techniques.

A potentially interesting characteristics of our approach, though, is that it can represent learning in misspecified models, in a very easy way. We will explain this in more detail after having discussed how we account for model uncertainty.

2.2 Model Uncertainty: Bayesian Model Averaging.

Monetary policy-making under uncertainty has been at the center of several studies in recent years. Model uncertainty represents, in fact, an unavoidable characteristics in modeling monetary policy decisions. In particular, in a large information environment like ours, model uncertainty is likely to be very pervasive.

Therefore, 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 the lagged value of each of them and three additional lags for the dependent variable. Hence, we end up dealing with a set of $2^{18} \times 2 = 524,288$ possible models M_j.

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

$$[\text{AS}] \qquad \pi_t = \beta^{\pi}_{0,t}\iota_t + \beta^{\pi}_{j,t}\mathbf{X}_{j,t} + \varepsilon^{\pi}_t \tag{6}$$

$$[\text{AD}] \qquad y_t = \beta^{y}_{0,t}\iota_t + \beta^{y}_{j,t}\mathbf{X}'_{j,t} + \varepsilon^{y}_t \tag{7}$$

where ι_t is a t-dimensional vector of ones, $\beta^{\pi}_{0,t}$ and $\beta^{y}_{0,t}$ are constants, $\beta^{\pi}_{j,t}$ and $\beta^{y}_{j,t}$ are time-varying vectors of the relevant coefficients for every model j, and the regressors' matrices are represented by $\mathbf{X}_{j,t} = [\pi_{t-1},\ \pi_{t-2},\ \pi_{t-3},\ \pi_{t-4},\ y_{t-1},\ \mathbf{Z}_{t-1}]$, $\mathbf{X}'_{j,t} = [y_{t-1},\ y_{t-2},\ y_{t-3},\ y_{t-4},\ (i_{t-1} - \pi_{t-1}),\ \mathbf{Z}_{t-1}]$, with $\mathbf{Z}_{t-1}$ including lagged values of the leading indicators used by the Fed and listed in the appendix. Additional variables $\mathbf{Z}_t = \{z_{t+\tau}\}^{\infty}_{\tau=-\infty}$ are identified as the unavoidable central bank's judgment in Svensson (2003).

We use quarterly data, from 1969 to 2001. In the estimation, we employ demeaned variables, so we can get rid of the constants.

To deal with the considerable model and parameter uncertainty, arising in this wide information environment, we use Bayesian Model Averaging (BMA) estimation with Gibbs sampling, 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 also 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 $\mathbf{M} = \{M_1, ..., M_k\}$, 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:

$$pr\left(\Delta|D\right) = \sum_{k=1}^{K} pr\left(\Delta|M_k, D\right) pr\left(M_k|D\right), \tag{8}$$

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 (8), $pr\left(M_k|D\right)$ is given by:

$$pr\left(M_k|D\right) = \frac{pr\left(D|M_k\right) pr\left(M_k\right)}{\sum_{j=1}^{K} pr\left(D|M_j\right) pr\left(M_j\right)}, \tag{9}$$

where

$$pr\left(D|M_k\right) = \int pr\left(D|\beta_k, M_k\right) pr\left(\beta_k|M_k\right) d\beta_k \tag{10}$$

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 $\mathbf{M}$).

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 (9)). 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:

$$pr\left(M_k\right) = \prod_{j=1}^{p} \pi_j^{\delta_{kj}} \left(1 - \pi_j\right)^{1-\delta_{kj}}, \tag{11}$$

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.

With an enormous number of models, the posterior distributions could be very hard to derive (the number of terms in (8) could be extremely large, and also the integral in (10) could be really hard to compute).

For this reason, we need to approximate the posterior distribution in (8) 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

$$G = \frac{1}{N}\sum_{t=1}^{N} g(M(t)) \ \ \underrightarrow{a.s.} \ \ E(g(M)), \text{as } N \rightarrow \infty, \tag{12}$$

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 (8).

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. In addition, both the estimates of the coefficients and the models themselves are time-varying. As explained in the previous section, coefficients evolve according to:

$$\beta_t^{\mathbf{Y}} = \beta_{t-1}^{\mathbf{Y}} + t^{-1} R_t^{-1} \mathbf{X}_t^{\mathbf{Y}} (\mathbf{Y}_t - \mathbf{X}_t^{\mathbf{Y}\prime} \beta_{t-1}^{\mathbf{Y}}), \quad \mathbf{Y} = \pi, y. \tag{13}$$

where R_t again is the moment matrix. We anticipated in the previous paragraph that learning in our framework can be interpreted as an easy-to-implement example of learning under misspecified models. There is no correct model of the economy, we try to estimate a large number of possible models and assign them a posterior probability to be the 'true' model. Not only the best model is selected, but information coming from the whole set of models is retained and incorporated into the analysis. Our coefficient estimates summarizes the information about which are the most successful predictors of our target variables; they are, therefore, sufficient in accounting for the existing model uncertainty. Re-estimating our equations in every period, we are able, in this way, to take into account also model variation (in the form of variation of the regressors composing single models, or just variation in the probability assigned to the models), besides parameter variation. There is, therefore, learning on all this huge amount of information. The problem remains, however, extremely tractable, thanks to our Bayesian Model Averaging approach. Learning will move towards the true model, through (12).

3 Estimation.

We can proceed to recursively estimate the two equations (6), (7) by Bayesian Model Averaging.

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 help in avoiding 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[1].

Our Markov Chain Monte Carlo simulation (Gibbs sampling) is based on 51,000 draws, with the first 1,000 draws omitted for initial burn-in (to attenuate the effect of initial conditions on final estimates). We are increasing the number of draws for future versions of the paper. Coefficients' estimates however seem to have reached convergence and do not significantly change

[1] Bernanke and Boivin (2001), in their analysis of monetary policy in a data-rich environment, have dealt with this problem using dynamic common factors.

with longer simulations.

The first estimation is carried out using data from the beginning of the sample until 1987:03. Then, each period, we redo the exercise, until 2000:04. We focus on this interval because we want to be in a single policy regime, when we compare our optimal policy rule with the empirical one. The selected sample corresponds to the Greenspan period as Chairman of the Fed.

As an example of our estimation results, we report our findings for the last period of the sample in Table 1 and 2, for the inflation and output specifications, respectively.

Insert Table 1-2 about here.

In Table 1, we see that the chain has visited 8,430 models; among these, we report the models, which were characterized by a higher posterior probability greater than 2%. 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 less than 5% probability to be the "true" model. This 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.

We also 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.

Important predictors for inflation, besides its lagged value and CPI inflation, seem to be housing start and new and unfilled orders. Probably, this leading indicator provide a more accurate indication of the current state of real activity than the commonly used output gap, and give important information about future price pressures. CPI inflation accounts for the additional effect on inflation of prices of imported final goods[2].

Table 2 reports the corresponding results for the demand equation: here the most favored model accounts for more than 15% of the total probabilty mass. Significant 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). Also money supply enters significantly.

To evaluate the importance of time variation in the coefficients and, indirectly, in the models, we plot the dynamics of the relative estimates over time.

Insert Figure 1-2 about here.

There seems to be a not irrelevant variation in the effects of the leading indicators over time. For example, from figure 2, we can see an increase in the effect of financial wealth on output. Particularly evident is the increase in monetary policy strength, moving across the sample, as indicated by the real interest rate coefficient on output gap.

The models visited by the chain have been, individually, estimated by OLS; in a 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. The efficient estimator would be to simultaneously estimate the

[2]In fact, being ω the share of imported goods in the CPI, π_t domestic inflation, π_t^f foreign inflation, CPI inflation π_t^c can be expressed as: $\pi_t^c = (1-\omega)\pi_t + \omega\pi_t^f$.

two specifications by the Seemingly Unrelated Regression method (SUR). We find that estimates under this joint estimation are totally similar to the equation-by-equation results, so we remain confident on our OLS.

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.

4 Monetary Policy in a Large Information Environment.

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 that obtained under more traditional small 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:

$$L_t = \lambda_\pi \pi^2 + \lambda_y y^2 + \lambda_{\Delta i}(i_t - i_{t-1})^2. \tag{14}$$

The optimization is performed under the constraints given by the dynamics of the economy.

The optimal rule is usually given by the following specification:

$$i_t^* = f X_t, \tag{15}$$

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.

However, a Taylor rule expressed as a linear function of inflation and the output gap will not be optimal, unless these are sufficient statistics of the state of the economy and they are perfectly observed. We do not think this to be the case.

Our approach consists on letting the central bank directly respond to all the available series and leading indicators. When taking a decision, the monetary policy maker evaluates which variables are more successful in predicting inflation and real activity state (we approximate this by means of Bayesian

[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.

Model Averaging) and then calculates the optimal feedback coefficients to those variables.

The optimal monetary policy rule we consider is:

$$i_t^* = f\left[\pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, y_t, y_{t-1}, y_{t-2}, y_{t-3}, \mathbf{Z}_t\right], \tag{16}$$

where, in addition to current and lagged inflation and output gap (the usual Taylor rule terms), the optimal interest rate is adjusted, each period, in relation to the situation of several economic indicators, through the feedback coefficients found in the 1×21 vector f. This seems to better represent the real world case, where the Fed responds to a variety of different information.

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 a traditional representation of the economy, as the one used by Rudebusch and Svensson (2002), which takes into consideration only three variables (inflation, output and short term interest rate), and in the context of our higher dimension framework.

Both are then compared to the actual federal funds rate path, historically implemented by the Federal Reserve. We focus on the period 1987:03-2000:04, to track the evolution of monetary policy decisions from the start of the Greenspan era to the end of our sample.

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

$$\pi_{t+1} = \beta_1 \pi_t + \beta_2 \pi_{t-1} + \beta_3 \pi_{t-2} + \beta_4 \pi_{t-3} + \beta_5 y_t + \varepsilon_{t+1}, \tag{17}$$

$$y_{t+1} = \gamma_1 y_t + \gamma_2 y_{t-1} + \gamma_3 (i_t - \pi_t) + \eta_{t+1}. \tag{18}$$

4.1 Monetary Policy Inertia.

In the context of a traditional information environment, in order to obtain funds rate series compatible with the observed one, it is necessary to assign a considerable weight to an interest-rate smoothing objective in the central bank's preference function.

On the other hand, if we allow the central bank to deal with an increased and more realistic amount of information, and taking into account the existing model uncertainty, we can obtain optimal federal funds rate paths very similar to the actual one, just by considering a negligible 0.0005 penalty on interest rate variability in the loss function (against a relative weight of 0.9995 given to inflation).

Insert Table 3 here.

We see from the table that just with a very small penalty on policy instrument volatility (assuming "strict" inflation targeting and $\lambda_{\Delta i} = 0.0005$), we are able to obtain optimal federal funds series quite close to the historically realized one; over the sample, it is characterized by mean and standard deviation not too far from the actual funds rate (mean and standard deviation equal to 5.48 and 1.18, compared with the actual 5.78 and 1.72, respectively). It is also evident the improvement over the consideration of optimal monetary policy under a limited information context, where we end up with extremely unrealistic interest rate series (far too aggressive and volatile, std.= 27.9).

In this latter case, even allowing for a sensibly stronger preference for smoothing in the objective function (say $\lambda_{\Delta i} = 0.2$), the optimal rules we obtain do not lead to funds rate's paths characterized by a standard deviation compatible with the actual one. In fact, the interest rate series derived from the Rudebusch-Svensson framework, remain excessively volatile; 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.

4.2 Excess Policy Response.

Let's start from the Rudebusch-Svensson specification. We want to evaluate if considering additional information can help to better explain inflation and output gap. We consider the following equations, where inflation and output gap are regressed on the set of leading indicators:

$$\begin{bmatrix} \pi_t \\ y_t \end{bmatrix} = \begin{bmatrix} \gamma^\pi \\ \gamma^y \end{bmatrix} \mathbf{Z}_{t-1} + \begin{bmatrix} \varepsilon_t^\pi \\ \varepsilon_t^y \end{bmatrix}, \tag{19}$$

where $\mathbf{Z}_{t-1}$ consists of all the leading indicators. We summarize the information about our target variables, coming from this additional information in $\widehat{\pi}_t$ and $\widehat{y}_t$, the fitted values obtained by using just the leading indicators.

A first experiment consists in evaluating if this additional amount of information is indeed useful in better characterizing the behavior of output and inflation. To see this, let's consider the two Rudebusch-Svensson equations (17) and (18) and let's augment them with an additional term, given by $\widehat{\pi}_t$ and $\widehat{y}_t$, respectively. If they enter with a significant coefficient, this would mean that more information should be taken into account in modeling the economy. This would signal the existence of relevant omitted variables. We estimate the following:

$$\pi_{t+1} = \beta_1 \pi_t + \beta_2 \pi_{t-1} + \beta_3 \pi_{t-2} + \beta_4 \pi_{t-3} + \beta_5 y_t + \xi^\pi \widehat{\pi}_t + \varepsilon_{t+1}, \tag{20}$$

$$y_{t+1} = \gamma_1 y_t + \gamma_2 y_{t-1} + \gamma_3 (i_t - \pi_t) + \xi^y \widehat{y}_t + \eta_{t+1}. \tag{21}$$

We obtain estimates for the ξ^π and ξ^y coefficients equal to 0.33 (standard error 0.057) and 0.46 (s.e 0.06), respectively. Additional information enters, therefore, with a highly significant coefficient. The relevance of a larger information set seems, thus, unquestionable.

Let's turn now to the policy rule. The central bank's reaction function is usually expressed in the form of a Taylor rule:

$$i_t = f_\pi \pi_{t-1} + f_y y_{t-1} + \nu_t. \tag{22}$$

We aim at evaluating whether the central bank actually exploits a larger information set when setting policy. To this scope, we derive:

$$i_t = \mathbf{f}\mathbf{Z}_{t-1} + \nu_t, \tag{23}$$

where the policy instrument is changed according to the state of the economy, expressed by the set of leading indicators $\mathbf{Z}_{t-1}$, through the vector of feedback coefficients $\mathbf{f}$. From (23), we can derive $\widehat{i_t}$ and evaluate if there exists an excess policy response, by inserting this fitted value in the standard Taylor rule (22):

$$i_t = f_\pi \pi_{t-1} + f_y y_{t-1} + \xi \widehat{i_t} + \nu_t. \tag{24}$$

We obtain $\xi = 1.05$ (with standard error 0.12), a large and highly significant excess policy reaction. This signals that a reaction to additional information is indeed present. This indicates that the Fed makes use of more information in taking policy decisions, than what is usually assumed in macroeconomic modeling.

5 Conclusions.

In the paper, we provided an attempt to incorporate a larger information set in the standard central bank's decision problem. Since the policy maker does not have perfect knowledge about the stance of the economy, she needs to monitor several variables, the 'leading indicators', to better forecast the evolution of the target variables.

In this framework, model uncertainty becomes an extremely relevant issue. We take into account by proposing estimation through Bayesian Model Averaging. This procedure, in fact, allows us to estimate the whole set of believed possible models of the economy. The final estimates give us an indication of the existing uncertainty.

In this uncertain environment, we have allowed the central bank to learn over time the value of the model coefficients. Those coefficients are Bayesian estimates obtained as weighted average of the values over all the models, weighted by the respective posterior probabilities. Therefore, the form of adaptive learning we have used can be seen as a way to also learn the best models through time.

Our effort was to insert some elements of bigger realism in the modeling of monetary policy-making. In our framework, our aim was to evaluate if the allowance of larger information set could be able to partially solve the interest-rate smoothing puzzle in optimal monetary policy. This consists of a much more gradual and conservative real world monetary policy, compared to the optimal policy suggested by dynamic optimization results in macroeconomic models. We have shown that in a large information environment, a more gradual monetary policy can be justified. At least to some extent, large information sets can be useful to attenuate this puzzle, and probably warrant further investigation.

We should examine if these results are robust to different modeling choices. For example, much bigger data sets than ours could be incorporated into the policy maker's decision problem and information could be summarized by dynamic common factors. The optimality of smoothing in a model with factors is currently the subject of a separate work. A 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. Other extensions could include more sophisticated methods to introduce learning. In particular, allowing for active experimentation could change the results in both directions (i.e., leading to either a more aggressive or more cautious policy rule).

Finally, we hope that the use of Bayesian Model Averaging, 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.

References

[1] Beck, G.W., and V. Wieland, (2002). "Learning and Control in a Changing Economic Environment", *Journal of Economic Dynamics and Control* 26, 1359-1377.

[2] Bernanke, B., and J. Boivin, (2001). "Monetary Policy in a Data-Rich Environment", forthcoming in *Journal of Monetary Economics.*

[3] Bernanke, B., J. Boivin, and P. Eliasz, (2002). "Measuring the Effects of Monetary Policy: a Factor-Augmented Vector Autoregressive (FAVAR) Approach", mimeo, Princeton University.

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

[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] Evans, G.W., and S. Honkapohja, (2001). *Learning and Expectations in Macroeconomics*, Princeton University Press, Princeton, New Jersey.

[7] Favero, C.A., and F. Milani, (2001). "Parameters' Instability, Model Uncertainty and Optimal 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] Kozicki, S., (2001). "Why do Central Banks Monitor so Many Inflation Indicators?", *Federal Reserve Bank of Kansas City Economic Review Articles*, Third Quarter.

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

[11] Milani, F., (2002). "Monetary Policy with a Wider Information Set: a Bayesian Model Averaging Approach", mimeo, Princeton University.

[12] 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.

[13] Onatski, A., and N. Williams, (2003). "Modeling Model Uncertainty", *Journal of the European Economic Association*, forthcoming.

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

[15] 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.

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

[17] 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.

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

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

[20] 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.

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

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

[23] 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.

[24] Svensson, L.E.O., (2003). "What is Wrong with Taylor Rules? Using Judgment in Monetary Policy through Targeting Rules", *Journal of Economic Literature*, forthcoming.

[25] Svensson, L.E.O., and M. Woodford, (2002). "Indicator Variables for Optimal Policy", working paper, Princeton University, January.

[26] Woodford, M., (1999). "Optimal Monetary Policy Inertia", NBER Working Paper No. 7261, July.

Bayesian Model Averaging Estimates
Dependent Variable = Infl
R-squared = 0.9882
sigma^2 = 0.0791
Nobs, Nvars = 123, 18
ndraws =51000
nu,lam,phi = 4.000, 0.250, 3
# of models = 8430
time(seconds) = 9904.759

****

Model averaging information

Model	infl1	infl2	infl3	infl4	consconf1	cpiinfl1	empl1	housing1	invsales1	m21
model 1	1	0	0	1	0	1	1	1	0	0
model 2	1	0	1	0	0	1	1	1	0	0
model 3	1	0	0	0	0	1	1	1	0	0
model 4	1	1	0	0	0	1	1	1	0	0

Model	napm1	neword1	outgap1	retail1	shipm1	stock1	unford1	vehicles1	Prob	Visit
model 1	0	1	0	0	0	1	1	0	2.06	43
model 2	0	1	0	0	0	1	1	0	2.883	59
model 3	0	1	0	0	0	1	1	0	3.288	85
model 4	0	1	0	0	0	1	1	0	4.694	67

****

	BMA Posterior Estimates		
Variable	Coefficient	t-statistic	t-probability
infl1	0.989508	13.5742	0
infl2	-0.107356	-1.16659	0.245631
infl3	-0.047529	-0.88069	0.380201
infl4	-0.020755	-0.55718	0.578416
consconf1	-0.044897	-0.28719	0.774452
cpiinfl1	0.124894	3.60133	0.000458
empl1	-0.035393	-1.23553	0.21899
housing1	0.487761	2.74109	0.007037
invsales1	0.025956	0.03329	0.973501
m21	0.001433	0.11204	0.910977
napm1	0.008723	0.03807	0.969693
neword1	0.015809	2.98109	0.003463
outgap1	0.001163	0.05984	0.952381
retail1	-0.000906	-0.0536	0.957345
shipments1	-0.002388	-0.16711	0.867561
stock1	-0.004032	-1.76043	0.08082
unford1	0.011516	2.32201	0.021877
vehicles1	-0.00115	-0.35431	0.723715

Table 1 - BMA Estimates (infl. eq.).

Bayesian Model Averaging Estimates
Dependent Variable = Output Gap
R-squared = 0.9321
sigma^2 = 0.4039
Nobs, Nvars = 123, 18
ndraws = 51000
nu,lam,phi = 4.000, 0.250, 3
# of models = 4158
time(seconds) = 5498.246

****

Model averaging information

Model	outgap1	outgap2	outgap3	outgap4	consconf1	cpiinfl1	empl1	housing1	invsales1	m21
model 1	1	0	1	0	1	0	0	1	0	1
model 2	1	0	0	0	1	0	0	1	0	0
model 3	1	1	0	0	1	0	0	1	0	1
model 4	1	0	0	0	1	0	0	1	0	1
model 5	1	0	0	0	1	0	0	1	0	1

Model	napm1	neword1	retail1	shipm1	stock1	unford1	ehicles	reals1	Prob	Visit
model 1	0	0	0	0	1	0	0	1	2.099	261
model 2	0	0	0	0	1	0	0	1	2.428	667
model 3	0	0	0	0	1	0	0	1	4.542	262
model 4	0	0	0	0	0	0	0	1	11.785	721
model 5	0	0	0	0	1	0	0	1	15.292	497

****

BMA Posterior Estimates

Variable	Coefficient	t-statistic	t-probability
outgap1	0.646787	13.13242	0
outgap2	0.015864	0.189623	0.849917
outgap3	0.004876	0.079459	0.936797
outgap4	0.002226	0.045822	0.963526
consconf1	1.762535	4.434032	0.00002
cpiinfl1	-0.004006	-0.16523	0.869034
empl1	-0.005027	-0.092235	0.926661
housing1	2.41061	5.751456	0
invsales1	-0.620398	-0.386445	0.699835
m21	-0.052378	-2.258083	0.025703
napm1	0.141811	0.194781	0.845886
neword1	-0.000445	-0.0414	0.967044
retail1	-0.000972	-0.034964	0.972165
shipments´	-0.001024	-0.075037	0.940307
stock1	0.005927	1.284831	0.201266
unford1	-0.000458	-0.056068	0.955379
vehicles1	-0.000067	-0.010979	0.991258
reals1	-0.077983	-3.03921	0.002898

Table 2 - BMA Estimates (output eq.).

	FEDERAL FUNDS RATE PATHS		
λ_π= 0.9995, λ_y= 0, $\lambda_{\Delta i}$= 0.0005	Optimal FF (large information)	Optimal FF (R-S)	Actual FF
MEAN	5.48	-44.95	5.78
STD.	1.18	27.95	1.72
PERSISTENCE	0.808	0.912	0.969
λ_π= 0.8, λ_y= 0, $\lambda_{\Delta i}$= 0.2	Optimal FF (large information)	Optimal FF (R-S)	Actual FF
MEAN	7.38	3.11	5.78
STD.	0.002	3.11	1.72
PERSISTENCE	0.807	0.927	0.969

Table 3 - Optimal and actual federal funds rate paths.

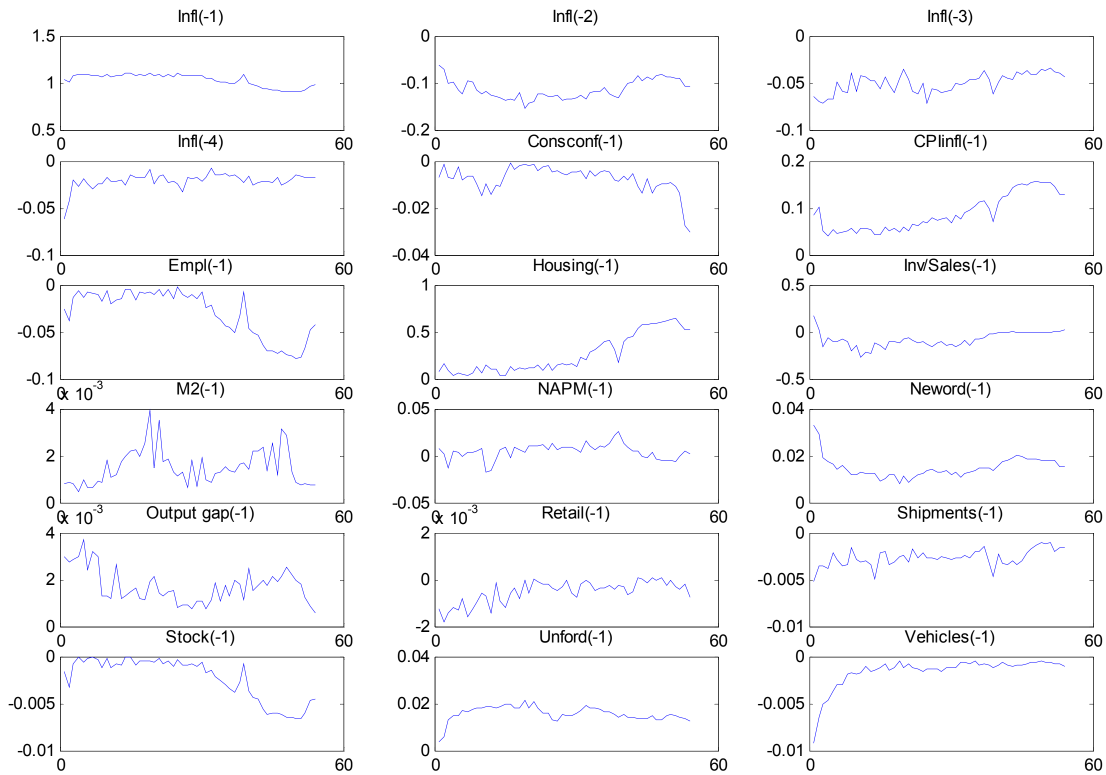

Figure 1 - BMA time-varying coefficients (inflation eq.).

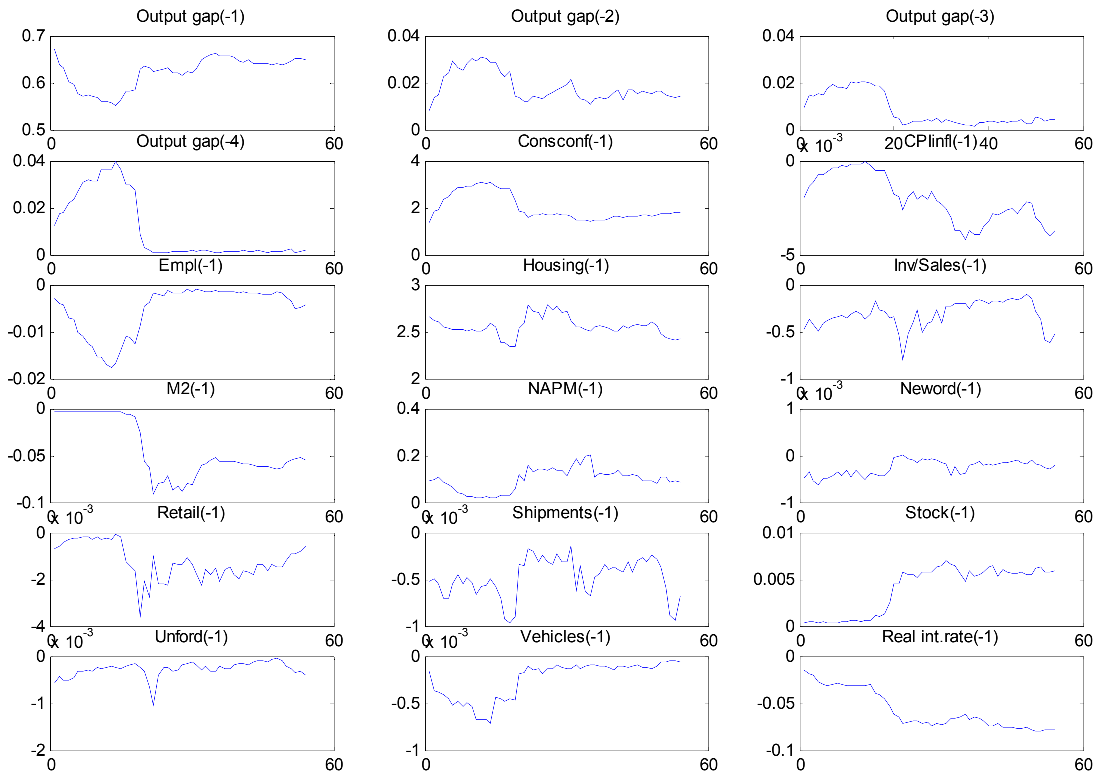

Figure 2 - BMA time-varying coefficients (output eq.).

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.