

# The role of model uncertainty and learning in the U.S. postwar policy response to oil prices

Francesca Rondina<sup>†</sup>

Institute for Economic Analysis, CSIC and Barcelona GSE

This draft: June 2010  
comments welcome

## Abstract

This paper studies optimal monetary policy in a framework that explicitly accounts for policymakers' uncertainty about the channels of transmission of oil prices into the economy. More specifically, I examine the robust response to the real price of oil that US monetary authorities would have been recommended to implement in the period 1970 – 2009, had they used the approach proposed by Cogley and Sargent (2005b) to incorporate model uncertainty and learning into policy decisions. In this context, I investigate the extent to which regulators' changing beliefs over different models of the economy play a role in the policy selection process. The main conclusion of this work is that, in the specific environment under analysis, one of the underlying models dominates the optimal interest rate response to oil prices. This result persists even when alternative assumptions on the models' priors change the pattern of the relative posterior probabilities, and can thus be attributed to the presence of model uncertainty itself.

**Keywords:** model uncertainty, learning, robust policy, Bayesian model averaging, oil prices

**JEL Classification:** C52, E43, E58, E65

## 1 Introduction

This paper investigates the role of model uncertainty in an environment in which the monetary authority acknowledges the existence of different competing mechanisms through which oil prices might affect economic performance. More specifically, this work focuses on the policy recommendations originating from an optimization problem that explicitly incorporates policymakers' uncertainty over a finite number of possible model representations of the economy. Using US data for the period 1970 – 2009, this study aims to provide an analysis of the extent to which the optimal response to oil prices is affected by policymakers' changing belief over the model to be adopted for policy decisions. In addition, this paper wants to offer some insights on the dimensions in which the robust policy recommendation relates to the optimal policies suggested by the individual underlying models.

---

\*Contacts: Campus UAB, 08193 Bellaterra, Barcelona, Spain; tel: +34 935 806 612, fax: +34 935 801 452, email: francesca.rondina@iae.csic.es

<sup>†</sup>This paper is a revised version of the second chapter of my Ph.D. dissertation at the University of Wisconsin - Madison. I would like to thank Steven Durlauf for his guidance and advice, and William Brock, Tim Cogley, Morris Davis, Luca Gambetti, Kenneth West and Noah Williams for helpful discussions and comments. All errors remain my own.

A number of contributions have analyzed the response of economic variables to changes in the price of oil.<sup>1</sup> These studies suggest that oil prices can affect economic activity in many different ways. First of all, an increase in the price of oil is likely to increase the general price level, which in turn can reduce employment if wages exhibit some degree of rigidity (Solow, 1980; Pindyck, 1980; Blanchard and Gali, 2007). Second, petroleum derivatives are directly used in a variety of production processes as inputs, and they are employed to provide supporting services as for instance energy and transportation. It follows that changes in their price will directly affect the cost of production for a large number of goods and services (Rasche and Tatom, 1977; Kim and Loungani, 1992). In this context, additional effects might be generated by changes in capacity utilization rates (Finn, 2000), or changes in business markups (Rotemberg and Woodford, 1996). Last, increases in the price of oil might lead to a reallocation of resources between sectors of the economy (Hamilton, 1988), and/or induce a reduction in the purchases of durable and investment goods (Bernanke, 1983).

The large literature studying the effects of oil prices on economic variables has not reached a consensus on which of the described mechanisms should be the main channel through which a change in the price of oil is transmitted in the economy. As a consequence, a number of works have also been debating on the optimal policy response to oil price shocks and, more generally, on the ability of policy to contrast the consequences of oil price variations on economic activity. An example is provided by the argument between Bernanke, Gertler and Watson (1997, 2004) and Hamilton and Herrera (2004) about the role of monetary policy in the economic downturns following the oil price shock episodes of the postwar period.

In this environment, some support can be provided by the literature on model uncertainty in economic policy. In recent years, the two major directions of work that have emerged are the minimax approach (Hansen and Sargent, 2001a; 2001b; 2008) and the Bayesian approach (Brock, Durlauf and West, 2003; 2007). In this second direction, the contribution of Cogley and Sargent (2005b) extends the Bayesian method to incorporate policymakers' learning over the model that best approximates the true data generating process. More specifically, the monetary authority is assumed to base his decisions on a composite framework encompassing a few alternative models of the economy, in which each model is weighted based on its relative posterior probability. These weights, together with the estimated model parameters, are updated in every period as new data becomes available.

In this paper, I consider the problem of a policymaker who is uncertain about three alternative specifications of the mechanism through which oil prices affect the economy, and who wants to use the approach proposed by Cogley and Sargent (2005b) to incorporate this form of uncertainty in the policy decision process. In this context, I study the pattern of the model weights, which can be interpreted as policymakers' beliefs over the true data generating process, and I focus on their relationship with the optimal response to oil prices in the 1970 – 2009 period. I find that in this environment model uncertainty has important effects on the optimal policy recommended by a composite linear quadratic regulator problem that includes the alternative models of the economy. Indeed, I show that uncertainty itself matters more than the actual pattern of the weights attached to each model, as long as these weights are different from zero. In addition, I find that, for a large fraction of the period under analysis, the robust interest rate is lower than the optimal policy recommended by each of the underlying individual models. This result is a consequence of the specific approach used in this paper to account for model uncertainty, in which the policy instrument can generate interactions between the variables of interest in directions that are precluded in the individual models. This finding is consistent with a number of previous works studying monetary policy in contexts characterized by imperfect knowledge of the true data generating process, which also concluded that uncertainty might induce an attenuated policy response to economic variables (see, for instance, Brainard, 1967; Wieland, 2000a; Svensson, 1999; Rudebusch,

---

<sup>1</sup>Mork (1994), Hamilton (2005) and Segal (2007) provide excellent reviews of this literature.

2001; Söderström, 2000, Tetlow and von zur Muehlen, 2001 for the case of structured model uncertainty).<sup>2</sup>

This paper contributes to the existing literature in economics in several aspects. First, my work is related to the large literature studying the impact of oil prices on economic activity. This study does not aim to take a stance in the debate over the different models that have been proposed to explain the effects of oil prices on economic performance. Rather, it wants to point out that an issue of model uncertainty in this area in fact exists. Indeed, I show that the three alternative model specifications considered in this study are characterized by posterior probabilities that are relatively similar in the period under analysis. This implies that the available data does not clearly favour one of these three approximating models over the others. This work is also related to the literature studying the role of monetary policy in contrasting the impact of oil prices on economic variables. In particular, it provides an analysis of the effects of learning and model uncertainty on the optimal policy response to changes in the price of oil. Finally, this paper offers a minor methodological contribution, and a more important empirical contribution, to the literature on model uncertainty and learning in economic policy. From the methodological point of view, this paper provides a simple extension of the approach proposed by Cogley and Sargent (2005b) to bivariate models in which the policy instrument is assumed to be the interest rate instead of the inflation rate. From the empirical point of view, this study proposes an application of the method introduced by Cogley and Sargent (2005b) to an environment in which the lack of consensus over the true model of the economy makes the inclusion of uncertainty in the policy decision process essential to sound policymaking.

The remainder of the paper is organized as follows. Section 2 presents the general setting and explains the policy decision process. Section 3 reports the robust policy results, and compares them with the policy recommendations obtained from the underlying individual models. Section 4 provides an analysis of the sensitivity of the results to changes in some of the assumptions adopted in section 3. Section 5 discusses the issue of private agents' expectations of future inflation in frameworks characterized by model uncertainty. Section 6 concludes.

## 2 General framework

The central assumption of this paper is that the policymaker does not know the true data generating process and chooses to base his decisions on a statistical approximating model. At the same time, he acknowledges that different specifications could be used for this purpose. In particular, I assume that the monetary authority considers a small number of alternative approximating models, each one incorporating a different theory on the channel through which oil prices affect economic variables. Each model specification is composed of three equations, one for the output gap,  $y_t$ , one for the inflation rate,  $\pi_t$  and one for the real price of oil,  $s_t$ . This last is common to all models. In each model, the policy instrument is the interest rate,  $i_t$ , which is selected according to the procedure that will be explained in the last part of this section.

As previously discussed, the literature in economics has proposed many different mechanisms through which oil prices can affect economic performance. In this work, policymakers believe that the true model of the economy can be approximated by one of three alternative econometric specifications.<sup>3</sup> In all the specifications, the real price of oil is assumed to (potentially) have a direct impact on both the output gap, because of the repercussions of the price of petroleum derivatives on the cost of production, and core CPI inflation, because

---

<sup>2</sup>However, other studies focused on environments in which uncertainty resulted in a stronger response of the policy instrument to fluctuations in the variables of interest. See, for instance, Giannoni, 2007; Söderström, 2002; and Tetlow and von zur Muehlen, 2001, for the case of unstructured model uncertainty.

<sup>3</sup>The model specifications that I consider in this study are similar to the ones I examine in a related work (Rondina, 2010). However, while in this last I focus on the evaluation of the performance of alternative policy rules, here I provide an historical analysis of the effect of model uncertainty and learning on the optimal policy response to oil prices. Moreover, Rondina (2010) considers additional forms of model uncertainty that are disregarded here.

of the possible pass-through to the general price level. However, the alternative models differ with respect to the presence, and the nature, of additional effects of oil prices on the variables of interest.

The first approximating model is in the spirit of Solow (1980). In Solow's view, an important channel through which oil prices affect output is nominal wage rigidities. More specifically, Solow argues that variations in the price of oil induce changes in the overall price level, and this in turn affects employment and real variables due to the assumption of nominal wage rigidities. The Solow ( $S$ ) model is represented by the equations:

$$y_t = \alpha_y^S(L) y_{t-1} + \alpha_\pi^S(L) [\pi_{t-1} - E_{t-2}(\pi_{t-1})] + \alpha_s^S(L) s_{t-1} + \omega_{y,t}^S \quad (1)$$

$$\pi_t = \beta_\pi^S(L) \pi_{t-1} + \beta_y^S(L) y_{t-1} + \beta_i^S(L) i_{t-1} + \beta_s^S(L) s_{t-1} + \omega_{\pi,t}^S \quad (2)$$

where  $\omega_{y,t}^S$  and  $\omega_{\pi,t}^S$  are i.i.d.  $N(0, \Sigma_S)$ . In this framework, the term for unanticipated inflation in the output equation aims to capture the additional effect of oil prices on output through nominal wage rigidities.

The second model is inspired by Blanchard and Gali (2007). In this framework, the main channel through which oil prices have an impact on economic activity is real wage and price rigidities. In particular, Blanchard and Gali think that the major consequence of a change in oil prices is a change in the overall price level, which in turn affects employment and real variables due to the assumption of real wage rigidities and Calvo pricing. The Blanchard and Gali ( $BG$ ) specification is described by the following equations:

$$y_t = \alpha_y^{BG}(L) y_{t-1} + \alpha_i^{BG}(L) [i_{t-1} - E_{t-1}(\pi_t)] + \alpha_s^{BG}(L) s_{t-1} + \omega_{y,t}^{BG} \quad (3)$$

$$\Delta\pi_t = \beta_\pi^{BG}(L) \Delta\pi_{t-1} + \beta_y^{BG}(L) y_{t-1} + \beta_s^{BG}(L) s_{t-1} + \omega_{\pi,t}^{BG} \quad (4)$$

where  $\omega_{y,t}^{BG}$  and  $\omega_{\pi,t}^{BG}$  are i.i.d.  $N(0, \Sigma_{BG})$ . This model is similar to a standard new Keynesian framework (see, for instance, Rudebusch and Svensson, 1999), with the main difference being the addition of the real price of oil in both the output and inflation equations. The restriction on the inflation coefficients in (4) is included in order to enforce the long run neutrality of the Phillips curve. This is a common assumption in the empirical frameworks inspired by the new Keynesian Phillips curve (see, for instance, Brock, Durlauf and West, 2007; Cogley and Sargent, 2005b; Primiceri, 2006; Rudebusch, 2001; Rudebusch and Svensson, 1999).

The last approximating model is in the spirit of Kim and Loungani (1992) and Hamilton (1988, 1996, 2005). In an economy characterized by perfect competition and flexible prices and wages, changes in the price of oil affect employment and output directly through a change in the wage level and productivity. In this framework, monetary policy is not able to contrast the consequences of a change in oil prices on output, but it might have a significant role in controlling its effects on the inflation rate. Therefore, the Hamilton ( $H$ ) model is characterized by the equations:

$$y_t = \alpha_y^H(L) y_{t-1} + \alpha_s^H(L) s_{t-1} + \omega_{y,t}^H \quad (5)$$

$$\pi_t = \beta_y^H(L) y_{t-1} + \beta_\pi^H(L) \pi_{t-1} + \beta_i^H(L) i_{t-1} + \beta_s^H(L) s_{t-1} + \omega_{\pi,t}^H \quad (6)$$

where  $\omega_{y,t}^H$  and  $\omega_{\pi,t}^H$  are i.i.d.  $N(0, \Sigma_H)$ .

## 2.1 The process for the real price of oil

The three approximating models that the policymaker wants to consider for policy decisions are completed by a description of the process for the real price of oil:

$$s_t = \delta_t + \rho s_{t-1} + o_t \quad (7)$$

This  $AR(1)$  process is characterized by a time-varying intercept  $\delta_t$ , and by a constant autoregressive coefficient  $\rho$ . The innovation  $o_t$  is a transitory oil price shock, which is assumed to be uncorrelated over time, and to have mean zero and constant variance  $\sigma_o^2$ . The process described by (7) aims to capture the features of nonlinearity and nonstationarity that seem to characterize the behavior of the real price of oil,<sup>4</sup> and is a generalization of the representation used in Blanchard and Gali (2007), which simply set  $\delta_t = 0$  at each time  $t$ .<sup>5</sup>

The process in (7) can be written in matrix form as:

$$s_t = X'_{s,t} \theta_{s,t} + o_t \quad (8)$$

where  $\theta_{s,t} = [\delta_t \ \rho]'$  and  $X_{s,t} = [1 \ s_{t-1}]'$ . The vector of coefficients  $\theta_{s,t}$  in (8) is time dependent because of the time-varying intercept  $\delta_t$ . I assume that:

$$\theta_{s,t} = \theta_{s,t-1} + \varepsilon_{s,t} \quad (9)$$

where  $\varepsilon_{s,t} = [\varepsilon_t \ 0]'$  and  $\varepsilon_t$  is i.i.d. Gaussian, with mean zero and variance  $\sigma_\varepsilon^2$ . The variance of the vector  $\varepsilon_{s,t}$ , denoted by  $Q$ , is then a  $2 \times 2$  matrix with zeros everywhere except for the element in the first row, first column, which is  $\sigma_\varepsilon^2$ . Thus, the process for the real price of oil is affected by two types of shocks: one,  $o_t$ , which is transitory (even if, given the estimated value of  $\rho$ , it turns out to be very persistent) and one,  $\varepsilon_t$ , which represents the drift of the intercept  $\delta_t$ , and is permanent instead. I make the simplifying assumption that:

$$E \begin{pmatrix} o_t \\ \varepsilon_{s,t} \end{pmatrix} \begin{pmatrix} o_t & \varepsilon_{s,t} \end{pmatrix} = V = \begin{bmatrix} \sigma_o^2 & 0 \\ 0 & Q \end{bmatrix} \quad (10)$$

which implies that the innovation to  $\delta_t$  is uncorrelated to the oil price shock  $o_t$ .<sup>6</sup> Finally, both shocks are assumed to be uncorrelated to the innovations in the output and inflation equations in each model.

The three approximating models described by (1) – (6) and (7) have the common feature of being backward looking specifications. This decision can be interpreted as the response of policymakers to an additional concern for robustness. More specifically, I assume that faced with uncertainty on the direction of the contemporaneous relationship between the variables of interest, the monetary authority simply decides to use approximating specifications that do not explicitly model this feature. This choice is in the spirit of limiting the restrictions that are imposed to the models used for policy purposes, so that the obtained policy recommendations can be valid in a more general set of possible data generating processes.

In addition to the assumption of backward looking specifications, equations (7) – (10) characterize a process for the real price of oil that is exogenous and not affected by the behavior of output and core CPI inflation. This choice is consistent with previous studies considering oil prices as determined before domestic variables (for instance, Hamilton, 1983 and 1996; Bernanke, Gertler and Watson, 1997; Blanchard and Gali, 2007). However, there is also a large literature suggesting that oil prices might have been affected by world demand, and thus US domestic output, especially in the last two/three decades (among the others, Rotemberg and Woodford, 1996; Baumeister and Peersman, 2008; Kilian, 2008; Lippi and Nobili, 2008; Herrera and Pesavento, 2009).

<sup>4</sup>Pindyck (1999) shows that the long-run behavior of oil prices is best forecasted by models of reversion to stochastically fluctuating trend lines. Blanchard and Gali (2007) use a stationary  $AR(1)$  process for the real price of oil, but they suggest that this variable would be better described as non-stationary.

<sup>5</sup>While the process described by (7) will be used in the baseline scenario, section 4 studies the optimal policy recommendation in the case the real price of oil is defined as in Blanchard and Gali (2007).

<sup>6</sup>This assumption is motivated by a need to reduce the computational burden in the estimation of the parameters of the process for  $s_t$ .

The assumption of an exogenous process for the real price of oil was motivated by a necessity to simplify the environment under analysis, mostly for computational issues. This assumption might seem quite restrictive; however, there are two reasons why I believe that, in this specific framework, the consequences of this choice are not too severe. First, I considered the possibility of allowing the oil price shock  $o_t$  to be correlated with the innovations in the output and inflation equations in each model. This extension would have been able to account for a possible contemporaneous relationship between the real price of oil and domestic variables, without imposing any restrictions on the direction of this relationship. The estimation of the parameters of each approximating model in this more general framework was relatively straightforward. However, I found that the covariances between the innovation  $o_t$  and the innovations in the output and inflation equation were so small that these parameters did not change much relative to the baseline case, and the estimated process for the real price of oil was very similar in all the models. At the same time, this extension, added a lot of complexity to the procedure used for the computation of the posterior probabilities for each model, which were needed for the policy decision process.<sup>7</sup> Thus, my conclusion was that allowing for a contemporaneous relationship between the real price of oil and domestic variables did not add much information relative to the assumption of an exogenous process for the real price of oil, while at the same time increasing the computation time significantly. A second reason why the choice of an exogenous process for  $s_t$  has smaller repercussions in this framework is related to the specific characterization in (7) – (10), and in particular to the assumption of a drift in the intercept  $\delta_t$ . Indeed, this assumption allows for some additional flexibility, which partially reduces the limitations imposed by the exogeneity of  $s_t$ . For instance, it implies that the covariances between the variables of interest evolve over time, which might encompass some of the postulated changes in the relationship between the real price of oil and output.

I believe that it would be very interesting to investigate alternative ways of accounting for the possible endogeneity of the process for the real price of oil and, more generally, for the uncertainty on the direction of the contemporaneous relationships between the variables in (1) – (7). However, I decided to start with a simpler setting and to leave this extension for further research.

## 2.2 Learning

Following the recent literature on policy design and evaluation in uncertain environments, the policymaker wants to adopt an approach that incorporates the three approximating models described by (1) – (6) and (7) in the policy decision process. As I already mentioned, the monetary authority chooses to follow the procedure proposed by Cogley and Sargent (2005b), which is based on the Bayesian model averaging techniques introduced by Brock, Durlauf and West (2003, 2007). This procedure accounts for the fact that each model specification has a different probability of being the true data generating process, and that this probability might change over time as new information becomes available. More specifically, in each period the policymaker learns from the new data on two different levels. First, he reviews the estimates of the parameters of each model. Second, he updates the relative posterior probability attached to each one of the three approximating specifications.

### 2.2.1 Learning on the process for the real price of oil

The process for the real price of oil is characterized by a vector of coefficients  $\theta_{s,t}$ , which drifts over time, and by a constant covariance matrix  $V$ . These variables are unknown to the policymaker, and they need to be estimated and updated in each period as new data becomes available.

---

<sup>7</sup>More specifically, this extension required the posterior probabilities to be computed numerically, while in the baseline framework used in this paper they can be computed using simple analytical formulas.

The initial state for the vector  $\theta_s$ , denoted as  $\theta_{s,0}$ , and the hyperparameters in the covariance matrix  $V$  are assumed to be independent of each other and across blocks, so that we have:

$$p(\theta_{s,0}, V) = p(\theta_{s,0}) p(\sigma_o^2) p(Q) \quad (11)$$

In addition, the process assumed in (9) implies a very specific form for the matrix  $Q$ , which has  $\sigma_\varepsilon^2$  as the element in the first row, first column, and zero elsewhere. Thus:  $p(Q) = p(\sigma_\varepsilon^2)$ . The priors for the hyperparameters  $\sigma_o^2$  and  $\sigma_\varepsilon^2$  are assumed to be inverse Wishart distributions, so that the prior for the covariance matrix  $V$  can be written as:

$$\begin{aligned} p(V) &= p(\sigma_o^2) p(Q) = p(\sigma_o^2) p(\sigma_\varepsilon^2) \\ &= IW(B_{o,0|0}, v_{o,0}) IW(B_{\varepsilon,0|0}, v_{\varepsilon,0}) \end{aligned} \quad (12)$$

where  $B_{o,0|0}$  and  $B_{\varepsilon,0|0}$  are initial inverse scale parameters and  $v_{o,0}$  and  $v_{\varepsilon,0}$  are initial degrees of freedom. On the other hand, the prior on the vector of coefficients,  $p(\theta_{s,0})$ , is a multivariate normal distribution:

$$p(\theta_{s,0}) = N(\theta_{s,0}, P_{s,0}) \quad (13)$$

One of the central assumptions of this paper is that policymakers learn from the new data that becomes available in each period. This process of learning will be reflected in (11), and in particular in the prior for the covariance matrix  $V$ . The vector of coefficients  $\theta_s$  drifts over time, and it seems plausible that new data will not motivate policymakers to revise their prior on the value of the coefficients at time zero. However, things are different for  $V$ . Indeed, this covariance matrix is assumed to be time invariant, so that it is reasonable to think that the monetary authority will want to account for the new information that becomes available by updating the prior for  $V$  over time. Thus, in the learning process over the history of  $s_t$ , I will allow (12) to change. The empirical section and Appendix 2 describe how I computed the parameters in (12) and (13), and how the prior  $p(V)$  is updated in each period.

Now, let  $s^t$  summarize the history of  $s_t$  up to time  $t$ . In each period, the policymaker is interested in the joint posterior  $p(\theta_s^t, V | s^t)$ .<sup>8</sup> This posterior distribution is unknown, and a MCMC algorithm, in the form of a Gibbs sampler, is used to simulate draws from it. In particular, the procedure that I employ is very similar to that described in Cogley and Sargent (2001). Further details on the implementation of the Gibbs sampler in this framework are given in Appendix 2; for a more technical description of this approach, see Cogley and Sargent (2001).

### 2.2.2 Learning on the parameters of the approximating models

The process of learning on the models' estimated parameters follows Cogley and Sargent (2005b). The framework that I adopt in this paper requires a few technical adjustments relative to Cogley and Sargent's work, which pertain to the distribution of the parameters and their updating formulas. These changes are minor, and are only due to the fact that here each approximating model is characterized by a bivariate distribution (over output gap and core CPI inflation) instead of a univariate distribution.

---

<sup>8</sup>Actually, at each  $t$  the policymaker is ultimately interested in estimating the vector of parameters  $\theta_{s,t}$ , which is the information he needs for policy decisions. If  $V$  was known, we could simply estimate  $\theta_{s,t}$  using the Kalman filter. However, since the covariance matrix  $V$  is assumed to be unknown, the entire history  $\theta_s^t$  and  $V$  need to be jointly estimated.

The output and inflation equations in (1) – (6) can be written in more general terms as:

$$y_t = X'_{y,t} \theta_y + \omega_{y,t} \quad (14)$$

$$\pi_t = X'_{\pi,t} \theta_\pi + \omega_{\pi,t} \quad (15)$$

where  $X_{y,t}$  and  $X_{\pi,t}$  are vectors of right hand side variables, and  $\theta_y$  and  $\theta_\pi$  are vectors of coefficients. The dimension of  $X_{y,t}$  and  $\theta_y$  is  $k_y$ , while the dimension of  $X_{\pi,t}$  and  $\theta_\pi$  is  $k_\pi$ . Each approximating model can also be rewritten in the general form:

$$Y_t = X'_t \theta + \omega_t \quad (16)$$

where  $Y_t = [y_t \ \pi_t]'$  is a  $2 \times 1$  vector of left-hand side variables,  $X_t = X_{y,t} \oplus X_{\pi,t} = \begin{bmatrix} X_{y,t} & 0 \\ 0 & X_{\pi,t} \end{bmatrix}$  is a  $k \times 2$  matrix of regressors, with  $k = k_y + k_\pi$ ,  $\theta_t = [\theta'_y \ \theta'_\pi]'$  is a  $k \times 1$  vector of coefficients and  $\omega_t = [\omega_{y,t} \ \omega_{\pi,t}]'$  is the  $2 \times 1$  vector of innovations, with covariance matrix  $\Sigma$ .

For each model, the vector of coefficients  $\theta$  in (16) and the matrix  $\Sigma$  are unknown, and they need to be estimated. Given the assumptions on the innovations in (1) – (6), this can be done using standard techniques. As for the process for the real price of oil, the estimates of these parameters will reflect all the information available in each period, and will be updated with the arrival of new data. However, the vector  $\theta$  is not assumed to drift over time. This makes the estimation and updating procedure for  $\theta$  and  $\Sigma$  a lot simpler than the approach used for  $\theta_s^t$  and  $V$ .

I assume that the policymaker adopts a normal-inverse Wishart prior on the joint distribution  $p(\theta_0, \Sigma)$ :

$$p(\theta_0, \Sigma) = p(\theta_0 | \Sigma) p(\Sigma) \quad (17)$$

with:

$$p(\Sigma) = IW(B_0, v_0)$$

$$p(\theta_0 | \Sigma) = N(\theta_0, P_0)$$

The empirical section describes how I computed the parameters of these distributions. The prior (17) at time zero will be used to obtain the posterior for the first period. Then, this posterior will be used as the prior for the following period, and so on until the end of the sample. For each model, the representation in (16) is linear, given the vector  $\theta$  and the covariance matrix  $\Sigma$ , and the innovations are i.i.d. and conditionally normal given the right hand variables. Therefore, each approximating model is characterized by a multivariate normal conditional likelihood. The normal-inverse Wishart distribution is a conjugate prior for the multivariate normal likelihood, which implies that the parameters' posterior probabilities follow a normal-inverse Wishart distribution as well.<sup>9</sup> As a consequence, the mean values of the joint posterior for  $\theta$  and  $\Sigma$  can be computed in each period using simple formulas. Appendix 2 describes this procedure in more detail, and provides the updating formulas for the parameters of the joint posterior distributions for the vector of coefficients  $\theta$  and covariance matrix  $\Sigma$ .

---

<sup>9</sup>For more details on the properties of conjugate distributions, see Gelman, Carlin, Stern and Rubin (2003).



### 2.2.3 Learning on the model weights

As for the models' parameters, policymakers' learning on the models' relative weights also follows Cogley and Sargent (2005b). In the policy decision process, each model is weighted based on its posterior probability, which is computed using the available information. Let  $Z^t = (Y^t, X^t)$  summarize the history of the left hand and right hand variables, up to time  $t$ . The posterior probability of model  $i$  given data up to time  $t$  is defined according to Bayes's theorem as:

$$p(M_i | Z^t) \propto m_{it} \cdot p(M_i) \equiv w_{it} \quad (18)$$

where  $p(M_i)$  and  $p(M_i | Z^t)$  are the prior and posterior probabilities of model  $i$  respectively, while  $m_{it}$  is the marginalized likelihood function for model  $i$  at date  $t$ .

In this framework, the policymaker considers the  $S$ ,  $BG$  and  $H$  specifications as the only possible alternative approximating models in his policy decision process. Therefore, he will normalize the model weights so that in each period they add up to one. For any model  $i$ ,  $i = \{S, BG, H\}$ , the normalized weight for time  $t$  can thus be obtained as:

$$\alpha_{it} = \frac{w_{it}}{w_{S,t} + w_{BG,t} + w_{H,t}} \quad (19)$$

Appendix 2 provides a definition of the marginalized likelihood function in this framework, and describes the procedure used to compute and update the model weights.

## 2.3 Interest rate setting

The monetary authority is assumed to control the policy instrument  $x_t$ , which is related to the interest rate  $i_t$  according to the following rule:

$$i_t = x_{t|t-1} + \xi_t \quad (20)$$

where  $\xi_t$  is an i.i.d. normal shock with mean zero and variance  $\sigma_\xi^2$ . Here,  $x_{t|t-1}$  denotes the value of the instrument for time  $t$ , computed using information up to time  $t - 1$ . In each period, the value of  $x_{t|t-1}$  is obtained as the solution to a linear quadratic dynamic programming problem which will be described in the following of this section.<sup>10</sup>

A central assumption in the policy decision process is that, in each period, the monetary authority considers the estimated coefficients as if they were constant and not (possibly) changing over time. This is true for the parameters in the output and inflation equations, as well as for those in the process for the real price of oil. This assumption follows Cogley and Sargent (2005b) and, as already observed in their work, has important implications for the policy decision process. First of all, this assumption allows to bypass the issue of parameter uncertainty highlighted by Brainard (1967), since the decision rule will not depend on posterior variances or higher moments of the parameters' distribution. For  $s_t$  in particular, this assumption greatly simplifies the optimization process since the uncertainty about the estimated value of  $V$  would not only affect the past history  $\theta_s^t$ , but also the probability distribution of future trajectories of the vector  $\theta_s$ . Second, the assumption of constant parameters originates a learning process that has been defined "passive", since the policymaker takes into account the new information that becomes available but ignores the effects of current decisions on the future value of the models' parameters and posterior probabilities. This approach to learning eliminates the incentives for experimentation that might arise from exploiting the connection between policymakers' current

<sup>10</sup>The relationships between  $x_{t|t-1}$  and  $i_t$  described by (20) is only relevant in the regulator problem, since it establishes how the policy instruments  $x_{t|t-1}$  affects the outcomes of interest. However, it has no effect on the estimation of the models' parameters and posterior probabilities. Indeed, these are computed and updated in each period using the actual value of  $i_t$  obtained from the data, not the value of policy instrument.

decisions and future beliefs.<sup>11</sup>

The assumption that the estimated parameters do not change over time produces a policy recommendation that is suboptimal, because it does not include all the information available to the policymaker at the time the policy is set. However, incorporating parameter uncertainty, and accounting for the relationship between policy decisions and future information would make the problem significantly more difficult to solve. In this specific framework, this might not be worthwhile, since the optimal policy is only implemented for one period, and then revised in the following period given the new estimates of the models' parameters and weights. In other words, while the optimization process will suggest a pattern of policies for every future  $t$ , the policymaker is essentially only interested in the policy recommendation for the next period, since this is the only one that will actually be implemented given the estimated set of parameters and posterior probabilities. For this reason, the approach adopted in Cogley and Sargent (2005b) and followed in this work seems to be an efficient way of defining the regulator problem in an environment characterized by model uncertainty and learning.

Under the assumption of passive learning, and because the policymaker will refrain from including parameter uncertainty into policy decision-making, the optimization problem becomes standard and can be solved in each period given the estimated value of the models' parameters and weights. The setup of the regulator problem follows Cogley and Sargent (2005b). Each of the models described by (1) – (7) has a state-space representation in the form:<sup>12</sup>

$$S_{i,t} = A_i S_{i,t-1} + B_i x_{t|t-1} + C_i \omega_{i,t} \quad (21)$$

The policymaker considers a composite problem in which each model specification is weighted based on its relative posterior probability, defined by (19). The loss function for the composite problem is in the standard linear quadratic form:

$$\begin{aligned} \mathcal{L}_E &= E_{t-1} \sum_{j=0}^{\infty} \beta^j \left[ y_{t+j}^2 + \lambda_{\pi} (\pi_{t+j} - \pi^*)^2 + \lambda_i (i_{t+j} - i_{t+j-1})^2 \right] \\ &= E_{t-1} \sum_{j=0}^{\infty} \beta^j (S'_{E,t+j} Q_{E,t} S_{E,t+j}) \end{aligned} \quad (22)$$

where  $\pi^*$  is the target inflation rate,  $\lambda_{\pi}$  is the weight attached to inflation relative to the output gap, and  $\lambda_i$  is a cost that policymakers incur for changing the value of the policy instrument.

The matrix  $S_{E,t}$  is the state for the composite problem, defined as:

$$S_{E,t} = A_E S_{E,t-1} + B_E x_{t|t-1} + C_E \omega_{E,t} \quad (23)$$

with  $S_{E,t} = [S'_{S,t} \ S'_{BG,t} \ S'_{H,t}]'$ ,  $\omega_{E,t} = [\omega'_{S,t} \ \omega'_{BG,t} \ \omega'_{H,t}]'$  and

$$A_E = \begin{bmatrix} A_S & 0 & 0 \\ 0 & A_{BG} & 0 \\ 0 & 0 & A_H \end{bmatrix} \quad (24)$$

$$B_E = [B'_S \ B'_{BG} \ B'_H] \quad (25)$$

<sup>11</sup>These incentives for experimentation are investigated in the works of Wieland (2000a, 2000b), Beck and Wieland (2002), and Cogley, Colacito and Sargent (2007).

<sup>12</sup>Additional details on the state vector and matrices of coefficients for each model are given in Appendix 3.

$$C_E = \begin{bmatrix} C_S & 0 & 0 \\ 0 & C_{BG} & 0 \\ 0 & 0 & C_H \end{bmatrix} \quad (26)$$

The matrix  $Q_E$  is a weighting matrix counting each model in proportion to its posterior probability. In more detail:

$$Q_E = \begin{bmatrix} \alpha_S M'_S Q M_S & 0 & 0 \\ 0 & \alpha_{BG} M'_{BG} Q M_{BG} & 0 \\ 0 & 0 & \alpha_H M'_H Q M_H \end{bmatrix} \quad (27)$$

where  $\alpha_S$ ,  $\alpha_{BG}$  and  $\alpha_H$  are the model weights and  $M_i$  is a selection matrix that has the function to pick the target variables from the state vector  $S_{i,t}$ . The matrix  $Q$  takes the form:

$$Q = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \lambda_\pi & 0 \\ 0 & 0 & \lambda_i \end{bmatrix}$$

The weights  $\alpha_S$ ,  $\alpha_{BG}$  and  $\alpha_H$  and the matrices  $A_E$ ,  $B_E$ ,  $C_E$  and  $Q_E$  change from period to period as the policymaker updates his beliefs in light of the new information that becomes available.

In each period, this problem delivers an optimal policy rule that is a linear function of the composite state:  $x_{t|t-1} = -F_E S_{E,t-1}$ , where  $F_E$  is a row vector of policy coefficients. Given the optimal policy, the state space representation can be rewritten as:

$$S_{E,t} = (A_E - B_E F_E) S_{E,t-1} + C_E \omega_{E,t} \quad (28)$$

with  $(A_E - B_E F_E)$  representing the deterministic portion of the transition law for the composite state vector.

### 3 Empirical results

The approximating models were estimated using the number of lags reported in table 1. For all the variables, the exclusion from one or more of the equations was dictated by the structure of each individual model, as described in (1)–(6). The lags for the included variables were chosen using the Bayesian information criterion.<sup>13</sup> Finally, constant terms were also incorporated in all estimations.

Table 1 - Lags used in the estimation procedure

	<i>l.h.v</i>	<i>y</i>	$\pi$	<i>i</i>	<i>s</i>
<i>S</i>	<i>y</i>	3	2	0	1
	$\pi$	1	3	2	1
<i>BG</i>	<i>y</i>	3	0	1	1
	$\pi$	2	3	0	1
<i>H</i>	<i>y</i>	3	0	0	1
	$\pi$	1	3	2	1

Note: number of lags of the output gap (*y*), inflation ( $\pi$ ), interest rate (*i*) and real price of oil (*s*) included in each equation (with left hand variable *y* or  $\pi$ ) for each approximating model: Solow (S), Blanchard-Gali (BG) and Hamilton (H).

<sup>13</sup>This work does not account for policymakers' uncertainty about the number of lags to be used in the estimation of each approximating model. This form of uncertainty is considered in Rondina (2010).

There is some debate in the literature about the number of lags of the oil variable that should be included in empirical analysis (see Hamilton and Herrera, 2004). Here, the Bayesian criterion favors models that only include one lag of the real price of oil. This might appear to limit the effects of oil prices on the other variables of interest and, consequently, to induce a reduced policy response to their variations. However, the specific process for  $s_t$  that I assume in this framework implies that shocks to the real price of oil are either very persistent or permanent, which suggests that policymakers will still want to respond to changes in this variable.<sup>14</sup>

I assume that private agents forms expectations using a backward-looking approach, so that:  $E_{t-1}(\pi_t) = \frac{1}{4} \sum_{j=1}^4 \pi_{t-j}$ . This assumption precludes all types of forward looking behavior on the side of the public, thus eliminating any issues related to the role of model uncertainty in the expectation formation process. Section 5 provides a more extensive discussion about the definition of expectations of future variables in a model uncertainty environment.

In the process for the real price of oil, the parameters in (11) need to be selected. For the inverse Wishart distributions,  $B_{o,0|0}$  was set as the sum of squared residuals from the OLS estimates of a stationary process for (7) using an initial training sample, while  $B_{\varepsilon,0|0}$  was set as the variance of the OLS estimate for  $\delta$  in the same stationary process multiplied by the degrees of freedom. Both  $v_{o,0}$  and  $v_{\varepsilon,0}$  were set equal to the degrees of freedom in the initial training sample. For the vector of coefficients,  $\theta_{s,0}$  was set as the OLS point estimate in the stationary process mentioned above, and  $P_{s,0}$  as the corresponding variance. These choices for the prior distributions are quite standard, even if for the covariance matrix  $V$  the prior is more informative than those usually adopted in the literature (see, for instance, Cogley and Sargent, 2001 and 2005a or Primiceri, 2005). This choice was motivated by the fact that at the beginning of the sample period the policymaker will need to estimate the posterior distribution  $p(\theta_s^t, V | s^t)$  from a very small number of observations. In addition, just at the beginning of the sample, the process for the real price of oil is subject to large changes due to the first oil price shock of 1973 – 74. In these circumstances, a more disperse prior on the covariance matrix  $V$  would not deliver a good convergence of the Gibbs sampler within a reasonable number of iterations.<sup>15</sup> Moreover, this framework is characterized by learning, which also involves the prior  $p(V)$ . Thus, even if the prior on the covariance matrix  $V$  is tighter than usually assumed in the literature, the addition of learning still enables the data to have a central role in determining the posterior distribution  $p(\theta_s^t, V | s^t)$ .<sup>16</sup> Section 3 examines the sensitivity of the results to alternative prior specifications.

At each time  $t$ , I simulated draws from the posterior distribution  $p(\theta_s^t, V | s^t)$  by performing 10,000 iterations of the Gibbs sampler and discarding the first 2,000 for convergence.<sup>17</sup> The history  $\theta_s^t$  and the variances  $\sigma_{o,t}^2$  and  $\sigma_{\varepsilon,t}^2$  were then estimated by computing the averages in the retained draws. In the regulator problem, the policymaker will use the estimated value of  $\rho$  from  $\theta_s^t$  as the autoregressive coefficient in the process for the real price of oil<sup>18</sup>, while the value of  $\delta_t$  in  $\theta_{s,t}$  will be the starting point of future trajectories for the intercept of this process. Finally, the policymaker will use the estimated variances  $\sigma_{o,t}^2$  and  $\sigma_{\varepsilon,t}^2$  to update  $p(V)$  in (11).

<sup>14</sup>As a robustness check, I looked at the number of lags recommended by the Akaike criterion. This criterion would favor models with a higher number of lags of  $y$ ,  $\pi$  and  $i$ , in particular in the inflation equations. However, for the real price of oil the number of selected lags was one, as in the Bayesian criterion.

<sup>15</sup>Here, I need to perform a number of iterations of the sampler in each period. In addition, the history to be estimated increases in length in each period. For this reason, it is necessary to keep the iterations performed by the Gibbs sampler to a limited number in order to make the computation manageable in terms of time.

<sup>16</sup>As a robustness check, I estimated the joint posterior  $p(\theta_s^T, V | s^T)$ ,  $T = 2009 : III$ , using the entire sample and a very disperse prior for both  $V$  and  $\theta_{s,0}$ . I verified that the estimates obtained in the paper with the described priors and the learning process on  $p(V)$  eventually converge to values that are consistent with those obtained from this exercise.

<sup>17</sup>I assessed the convergence of the Gibbs sampler by looking at the autocorrelation functions of the draws. The 20<sup>th</sup> order autocorrelations were all very small, well below 0.1, which suggests efficiency of the algorithm in converging to the distribution of interest.

<sup>18</sup>In the process described by (7) - (10), the value of  $\rho$  is assumed to be time invariant. The algorithm used to estimate  $\theta_s^t$  will actually deliver this result; see Appendix 2 for a discussion.

This new prior will be denoted as:  $p(V) = IW(B_{o,0|t}, v_{o,0}) IW(B_{\varepsilon,0|t}, v_{\varepsilon,0})$ . Given this new prior and the new observation  $s_{t+1}$ , the Gibbs sampling algorithm can be repeated to obtain draws from the posterior distribution  $p(\theta_s^{t+1}, V | s^{t+1})$ , and so on until the end of the sample. Notice that the history  $\theta_s^t$  to be estimated increases in length as time passes, thus making the problem computationally longer in each period.

The estimation procedure for  $p(\theta, \Sigma | Z^t)$  is simpler than the one required for the joint posterior  $p(\theta_s^t, V | s^t)$ . For each model, the parameters in (17) were computed from the OLS estimates obtained using the initial training sample. More specifically, in the inverse Wishart distribution  $B_0$  was set as the estimated sum of squared residuals, and  $v_0$  as the degrees of freedom in the training sample, while in the Gaussian distribution,  $\theta_0$  was set as the point estimate from the OLS regression, and  $P_0$  as its covariance matrix. As I previously mentioned, the normal-inverse Wishart distribution is a conjugate prior for the multivariate normal likelihood, which implies that the posterior on the parameters of each approximating model is also a normal-inverse Wishart distribution. In particular, at any time  $t$  the posterior will be:

$$p(\theta, \Sigma | Z^t) = p(\theta | \Sigma, Z^t) p(\Sigma | Z^t) \quad (29)$$

with:

$$p(\Sigma | Z^t) = IW(B_t, v_t)$$

$$p(\theta | \Sigma, Z^t) = N(\theta_t, P_t)$$

where  $v_t = v_{t-1} + 1$ . The formulas used to compute  $B_t$ ,  $\theta_t$  and  $P_t$  are provided in Appendix 2. The posterior for time  $t$  can then be used as the prior for time  $t + 1$ , and so on until the end of the sample. In each period, the policymaker will use the estimated mean of  $p(\theta | \Sigma, Z^t)$  for each model as the value of the coefficients in the policy decision process.

Finally, standard values are assumed for the parameters in the regulator's loss function. The discount factor  $\beta$  is set so that the annual discount rate is 4%. In the matrix  $Q$ ,  $\lambda_\pi$  is set to one, reflecting an equal weight of inflation and unemployment (as in Cogley and Sargent, 2005b, and Brock, Durlauf and West, 2007), while the cost of changing the policy instrument is  $\lambda_i = 0.1$  (as in Brock, Durlauf and West, 2007). This choice of weights is also in the range considered by Levin and Williams (2003). The value of the target inflation rate is set to 2%, as in Primiceri (2006). In the baseline case the initial model weights are  $\alpha_{S,0} = \alpha_{BG,0} = \alpha_{H,0} = 1/3$ , reflecting an uniform prior across the three approximating models.<sup>19</sup>

The data used in the estimation procedure is described in Appendix 1. The recursion starts in the first quarter of 1970.

### 3.1 The process for the real price of oil

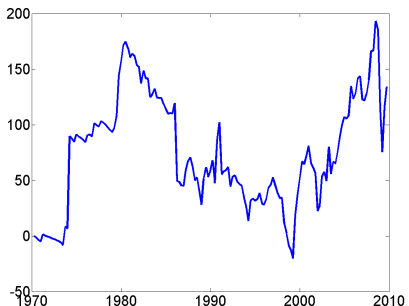
Figure 1 shows the pattern of the real price of oil in the period 1970 : I – 2009 : III. Figure 2 reports the estimated parameters for the process described by (7) – (10). The top panel illustrates the behavior of the coefficients  $\delta_t$  and  $\rho$  in (7) and the bottom panel the pattern of the variances  $\sigma_\varepsilon^2$  and  $\sigma_o^2$  in  $V$ . These values are the averages of the retained draws from the posterior  $p(\theta_s^t, V | s^t)$  generated using the Gibbs sampler.

Figure 2 shows that policymakers' estimates of  $\sigma_o^2$  drastically changed right after the first large increase in oil prices in 1973 – 74, while the value of  $\sigma_\varepsilon^2$  remained very low in the same period. If we think about these values as policymakers' beliefs over the true variances  $\sigma_\varepsilon^2$  and  $\sigma_o^2$ , then this pattern suggests that the regulator initially interpreted the first oil price shock as being due to a larger than expected value of  $o_t$ , rather than

<sup>19</sup>Section 3 studies the sensitivity of the results to changes in the initial model weights and in some of the parameters in the regulator's loss function.

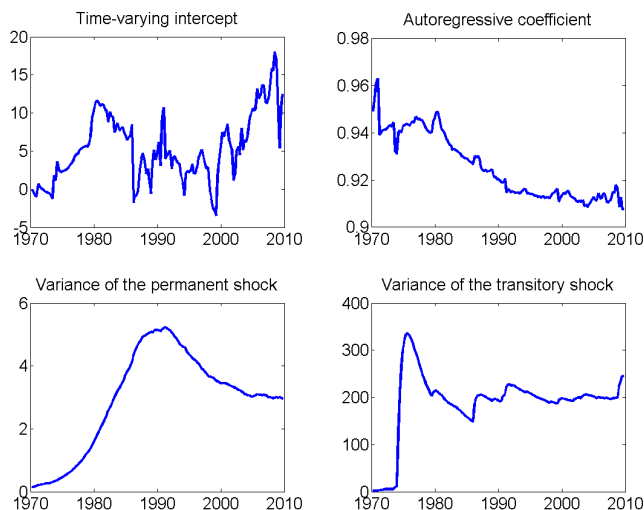
to a permanent shift in the mean of the process for  $s_t$ . However, after this first adjustment, the estimated value of  $\sigma_o^2$  decreased rapidly, while  $\sigma_\varepsilon^2$  started increasing, particularly since the early 1980s. Thus, over time policymakers adjusted their beliefs in the light of the new data, and became persuaded of the fact that, in addition to the transitory shock  $o_t$ , the process for the real price of oil was also affected by relevant permanent changes. Eventually, the estimated  $\sigma_o^2$  reached a relatively stable value in the late 1980s. On the other hand, the learning process on  $\sigma_\varepsilon^2$  led to an initial overestimate of its value, which was then reviewed during the 1990s and reached a nearly stable value after the year 2000.

Figure 1 - The real price of oil



Note: The real price of oil in the period 1970 : I – 2009 : III. See Appendix 1 for a description of the data.

Figure 2 - Parameters in the process for the real price of oil

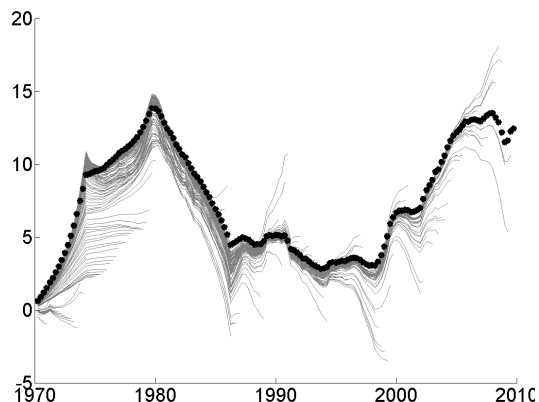


Note: Pattern of the parameters in the process for the real price of oil. The top panel reports the time-varying intercept  $\delta_t$  (left) and autoregressive coefficient  $\rho$  (right) in (7). The bottom panel reports the variances  $\sigma_\varepsilon^2$  (left) and  $\sigma_o^2$  (right). These parameters were estimated using the Gibbs sampler procedure described in Appendix 2.

The pattern of the variances  $\sigma_\varepsilon^2$  and  $\sigma_o^2$  is reflected in the estimated values of  $\delta_t$  and  $\rho$ . In the 1970s and 1980s, the value of  $\sigma_\varepsilon^2$  was still quite small, even if increasing. For this reason, the time-varying intercept displays a limited reaction to the two large oil price increases of 1973 – 74 and 1979 – 80, while starting from the early 1990s this variable captures a larger fraction of the variation in  $s_t$ . The estimated value of  $\rho$  decreased in the period under analysis, especially after the second oil price shock of 1979 – 1980 and was always lower than 0.97, the value chosen by Blanchard and Gali (2007). As for  $\sigma_\varepsilon^2$  and  $\sigma_o^2$ , the estimated patterns of  $\delta_t$  and  $\rho$  can be interpreted in terms of policymakers’ beliefs over the behavior of the process for the real price of oil.

In this view, we can infer that since the second large oil price increase of 1979 – 80, the monetary authority has been attributing a larger fraction of the variation in  $s_t$  to permanent changes in the structure of this process, while the perceived persistence of the transitory shock  $o_t$  has decreased.

Figure 3 - Learning on the history of the time-varying intercept  $\delta_t$



Note: Each line is the estimated history  $\delta_t^t$ , using information up to time  $t$ . The dotted line is the estimated history computed at  $T = 2009 : III$  using all the available information.

The learning process assumed in this framework implies that, at each time  $t$ , policymakers re-estimate the variances  $\sigma_\varepsilon^2$  and  $\sigma_o^2$ , and the entire history of  $\theta_{s,t}$  up to the current period. Given that the prior on  $\theta_{s,0}$  is not updated over time, the adjustments of the history  $\theta_s^t$  are due to the arrival of new data and to policymakers' changing beliefs over the covariance matrix  $V$ . Figure 2 provides some information on this learning process, by showing how the estimated values of  $\delta_t$ ,  $\rho$ ,  $\sigma_o^2$  and  $\sigma_\varepsilon^2$  were modified during the sample period under analysis. Figure 3 offers some additional insights by reporting the estimated history of the time-varying intercept  $\delta_t$  for each  $t = 1, \dots, T$ . The dotted line represent the estimate obtained at  $T = 2009 : III$ , using all the available information. From this figure, it is evident that in more than one occasion the believed history  $\delta^t$  was subject to significant revisions in later periods. Thus, figure 2 and 3 provide evidence on the fact that the process of learning on the real price of oil was substantial.

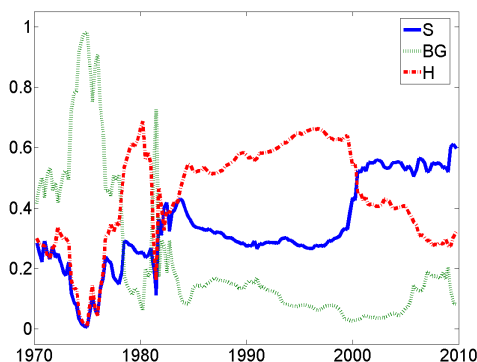
### 3.2 The robust policy response to oil prices

The robust policy recommendation was obtained using the approach described in Section 2. As I already mentioned, this approach requires the policymaker to define a composite model of the economy, in which each of the approximating specifications is weighted based on its relative posterior probability. In interpreting the results, these weights can be thought of as policymakers' beliefs over the extent to which the considered specifications approximate the true data generating process.

Figure 4 reports the pattern of the weights attached to each of the approximating models in the policy decision process. The first observation that we can make is that over time none of the three models became clearly predominant in terms of posterior probability. The period until the mid 1980s is characterized by large and rapid changes, especially for the  $BG$  and  $H$  models. In particular, while during the first large increase in the real price of oil of 1973 – 74 the data seems to favor the  $BG$  model, the second oil price shock of 1979 – 80 induces a change in the relative posteriors in favor of the  $H$  model. The decrease in the real price of oil in the first half of the 1980s is characterized by the  $H$  and  $BG$  models rapidly alternating as the model with the highest posterior, followed by a few periods in which the relative posterior probabilities of the three models are

very similar in magnitude. Since the mid 1980s the weights attached to each model have remained relatively more stable. The  $H$  model has maintained the highest relative posterior from the mid 1980s until around 2000, while during the long and steady increase in the real price of oil that started around 2000 the data seems to have favored the  $S$  model. Finally, during the rapid drop in the real price of oil that characterized the last few periods of the sample, the relative posteriors of the  $H$  and  $S$  models have increased at the expenses of the  $BG$  model.

Figure 4 - Model weights



Note: Relative posterior probability of each approximating model, computed using the procedure explained in Appendix 2.

We can interpret the pattern of the model weights in terms of policymakers' beliefs over the specification that best approximates the true data generating process. This interpretation allows to make some conjectures about how these weights could have affected the optimal policy recommended by the regulator problem. In the years around the first oil price shock, the policymaker believed that the  $BG$  model closely approximated the true model of the economy. As a consequence, he believed that monetary policy could play an important role in contrasting the effects of a change in oil prices on the economy, and in particular on the output gap. On the other hand, in the years around the second oil price shock, and from the second half of the 1980s to around 2000, policymakers' beliefs over the true data generating process changed in favor of the  $H$  model, in which the policy instruments cannot alter the impact of oil prices on output. Then, we might expect that in these years the response of monetary policy to changes in the real price of oil was mostly directed to reduce its consequences on the inflation rate. Finally, since 2000 the data seems to favor the  $S$  model, in which monetary policy has an indirect effect on the output gap through unanticipated inflation. Depending on the estimated magnitude of this effect, the monetary authority might believe that policy has again a role in limiting the effects of a change in oil prices on output. In the remainder of this section, I will show that these conjectures on the relationship between policymakers' beliefs and decisions are actually not always correct. This happens because one of the approximating models under consideration, namely the  $BG$  model, tends to dominate policy decisions, even when its weight is relatively small compared to the other two specifications.

Figure 5 reports the optimal (and robust to model uncertainty) policy recommendations obtained from the regulator problem described in section 2. The left-hand panel shows the value of the policy instrument  $x_{t|t-1}$ ,<sup>20</sup> while the right-hand panel focuses on the fraction of the optimal policy that is due to the direct response to the real price of oil.<sup>21</sup> There are a few observations that can be made from this figure. First, the optimal interest

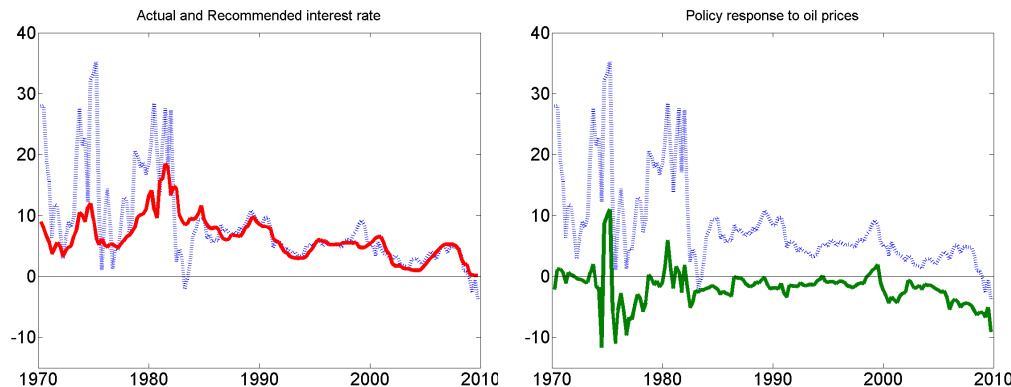
<sup>20</sup>I imposed no restrictions on the value of the control variable in the regulator problem; this is the reason why the optimal interest rate is negative in a few periods. This problem could be avoided by incorporating some interest rate target level, which I didn't do for simplicity of the analysis.

<sup>21</sup>In this setup, the time-varying intercept  $\delta_t$  enters in the state for each model (see Appendix 3 for a description of the state



rate recommended by the composite regulator problem reproduces much of the increases and decreases in the Federal funds rate during the entire sample under analysis. In addition, starting from 1985, it is also very close in magnitude. This result persists even if I eliminate the interest rate smoothing term in the policymaker's loss function. Therefore, the model uncertainty and learning framework employed in this paper seems to be able to capture some relevant features of the actual decision process during this time period. Second, both the optimal policy recommendation and the direct response to oil prices are a lot more volatile until the first half of the 1980s. This is due to several reasons. First, all the variables included in the approximating models are more volatile in this first part of the sample (the period after 1985 is usually referred to as "the great moderation"). The procedure described in section 2 delivers an optimal policy that is a linear function of the models' variables and that, as a consequence, follows their behavior. In addition, in this framework a higher volatility of the models' variables implies that the estimates of the models' parameters are subject to rapid and large adjustments, also because a smaller number of observations are available in this first portion of the sample. These changes are reflected in the matrices of the state space representation, so they affect the optimal policy recommendation. Finally, the pattern of the model weights is also changing quite rapidly in the period until the mid 1980s, and this has an effect on the linear regulator problem as well, since these weights appear in the policymaker's loss function.

Figure 5 - Policy recommendation - composite model



Note: The left-hand panel reports the policy recommended by the optimization problem described in section 2 (dashed line) together with the annual Federal Funds rate (continuous line). The right-hand panel reports the fraction of the optimal policy that represents the direct response to the real price of oil.

An important factor that emerges from the right-hand panel of figure 5 is that during almost the entire sample period, the direct policy response to the real price of oil is negative. As I mentioned, the optimal policy recommended by the regulator problem described in section 2 is a linear function of the models' variables. Since the real price of oil, with the exception of a handful of data points in the early 1970s and late 1990s, is almost always positive, this implies that in the optimal policy rule the coefficients attached to the oil variables are negative in most periods. Therefore, in most periods the robust regulator problem urged policymakers to react to an increase in the real price of oil by decreasing the nominal interest rate, that is, by implementing an expansionary policy. Two relevant exceptions to this conclusion are the large oil price shocks of 1973 – 74 and 1979 – 80, in which the direct response to oil prices becomes positive. In the next section I will propose

---

space representation for each model). Therefore, the direct policy response to the real price of oil has been computed as the response to  $\delta_t$  in addition to the response to  $s_t$ . However, since the real price of oil has an impact on output and on the inflation rate as well, the overall effect of oil prices on the policy recommendation might actually be larger than the simple response to  $\delta_t$  and  $s_t$ . It is in principle possible to study the overall impact of the real price of oil on the optimal policy, but I preferred to focus on the direct response for clarity of exposition.

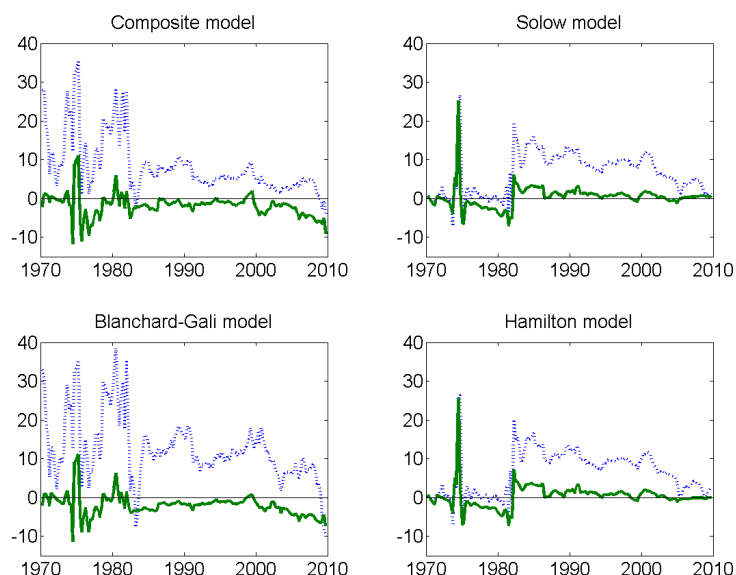
an explanation to these exceptions based on the optimal policies emerging from the individual models during these two historical events.

Finally, the right hand panel of figure 5 also shows that the robust response to oil prices is reduced in magnitude, and in average closer to zero, in the period that goes from the mid 1980s until around 2005. This result is actually consistent with the empirical literature studying the reaction of the Federal Reserve to oil price shocks. For instance, Bernanke, Gertler and Watson (1997) show that the response of the Federal funds rate to a 1% oil price shock is significantly smaller in the period 1986 – 1995 compared to the previous portion of their sample, and Kilian and Lewis (2009) argue that there is no evidence of systematic monetary policy responses to oil price shocks after 1987.

### 3.3 The role of uncertainty

The previous subsection studied the optimal policy recommended by a composite problem that explicitly accounts for model uncertainty and learning. How is this policy affected by policymakers’ changing beliefs on the model that best approximates the true data generating process? In other words, how does the pattern of the model weights reproduced in figure 4 relate to the optimal policy and direct response to the real price of oil reported in figure 5? To provide more insight into this question, I compare the policy obtained from the composite model with the optimal interest rate recommended by each individual approximating model. The results of this exercise are reported in figure 6 and figure 7.

Figure 6 - Policy recommendation - composite and individual models



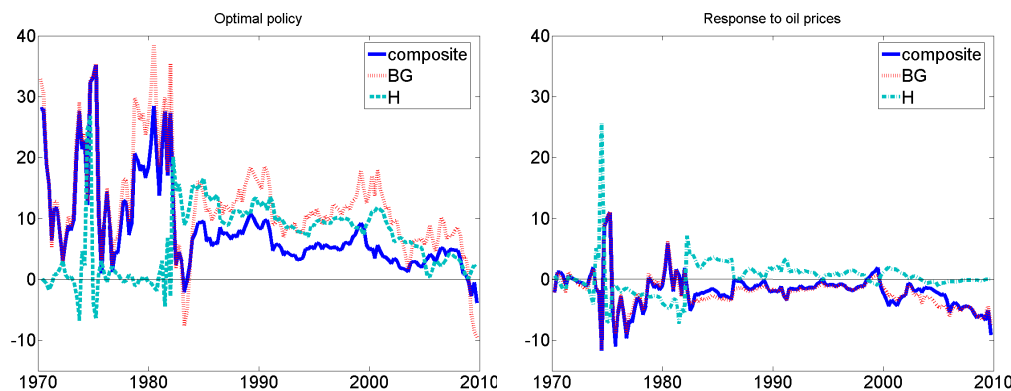
Note: Optimal policy recommendation (dashed line) and direct response to the real price of oil (continuous line) in the composite and individual models.

A first observation that we can draw from figure 6 is that both the optimal policy and the direct response to the real price of oil are very similar in the  $S$  and  $H$  models.<sup>22</sup> This happens because the estimated coefficients in the lag polynomial  $\alpha_{\pi}^S(L)$  turn out to be very small, even if significantly different from zero. Since the term  $\alpha_{\pi}^S(L) [\pi_{t-1} - E_{t-2}(\pi_{t-1})]$  is the only difference between the  $S$  and the  $H$  approximating models, this

<sup>22</sup>There are some differences only in the early 1980s and after 2005.

implies that when estimated these two models are indeed very similar, and therefore imply similar policies recommendations. Thus, for clarity of exposition, the  $S$  model has been omitted from figure 7.

Figure 7 - Policy recommendation - composite,  $BG$  and  $H$  models



Note: The left-hand panel reports the policies recommended by the composite problem described in section 2, the  $BG$  model and the  $H$  model. The right-hand panel reports the direct response to the real price of oil for the same models.

The main conclusion that we can infer from figures 6 and 7 is that the  $BG$  model seems to dominate the optimal policy obtained from the composite model, regardless of the pattern of the model weights. For the policy instrument, this is particularly true in the first part of the sample, up to around 1985. On the other hand, in the direct policy response to the real price of oil, the composite model closely follows the  $BG$  model during the entire sample, except for a few time periods in the 1980s and in between 2000 and 2005. In the 1980s, this difference might be due to the rapid increase in the weights attached to the  $S$  and  $H$  models, together with the fact that these two models were recommending a quite different response to oil prices in this period (a strong contractionary instead of the expansionary response suggested by the  $BG$  model). In between 2000 and 2005, the weight attached to the  $BG$  model reached near zero values, and this likely pushed the direct response to oil prices in the composite model closer to the recommendations of the  $H$  and  $S$  models.

What is the reason why the  $BG$  model drives policymakers' response to the real price of oil? The conclusion that one or a few of the specifications in the model space play a central role in the regulator problem is not uncommon in the literature using the Bayesian approach to account for model uncertainty.<sup>23</sup> Typically, this happens because some model specifications are more difficult to stabilize, and generate high losses under a wide range of policies. For this reason, in a model uncertainty framework the concern for robustness motivates policymakers to grant these specifications a particular treatment. In other words, in a regulator problem like the one described in section 2, models that imply very high (sometimes close to infinity) losses in a large set of policy choices will tend to dominate the optimal policy recommendation, even when their weight in the composite loss function is low. In this framework, it is difficult to provide an exhaustive analysis of the stability features of the  $S$ ,  $BG$  and  $H$  specifications, not only because they depend on the set of policies under consideration but also because they change over time as the policymaker updates his estimates of the models' parameters. However, there are some pieces of evidence from which it is possible to infer that the greater difficulty in stabilizing the  $BG$  model is an important part of the story in this environment as well. For instance, it is easy to show that the adoption of a fixed interest rate, that does not respond to changes in the state of the economy, generates dominant eigenvalues that exceed the stability boundary of  $\beta^{-1/2}$  in the  $BG$  model, but not in the  $S$  and  $H$  models. More generally, in the sample period under analysis the  $BG$  model is the only one that is never stable

<sup>23</sup>For instance, Cogley and Sargent (2005b) found that their "Samuelson-Solow" model was strongly affecting the optimal policy in the composite problem even when its weight was relatively low compared to the other models.

without policymakers' intervention, thus requiring an active policy in order to be stabilized. It follows that the *BG* specification can be viewed as a worst-case model, to which the Bayesian policymaker needs to devote particular attention, even when its weight in the decision process is very low. This conclusion also reveals a connection with the minimax approach proposed by Hansen and Sargent (2001a; 2001b; 2008)<sup>24</sup>, in which the regulator always implements the policy that allows him to protect against the worst-case scenario, regardless of the probability of this scenario to actually happen.

The characteristics of the *BG* model can be used to provide an economic interpretation of its central role in the policy decision process. Among the three approximating models, the *BG* is the one in which oil prices have potentially the greatest effect on the variables of interest. This is because in addition to the direct effect on output and inflation (which given the estimated coefficients is comparable in magnitude to the *S* and *H* models), this specification also implies a feedback of output on inflation and of (expected) inflation on output. This last effect is not present in the *H* model, and it is negligible in the *S* model. At the same time, the *BG* model is the one in which monetary policy can potentially control the impact of oil prices on the economy to the greatest extent. In particular, it is the only model in which monetary policy can directly reduce the impact on output, and the subsequent transmission to inflation through output and back to output through expected inflation.<sup>25</sup> Thus, the economic relationships implied by the *BG* model provide an additional interpretation of the reasons why policymakers might want to give this specification primary attention in the policy decision process.

The dominance of the *BG* model in the direct response to oil prices facilitates the understanding of the recommended reaction to specific events in the postwar period. In general, the part of the sample going from the first oil price shock to the early 1980s is characterized by rapid changes in the response to oil prices, for the reasons that I have already described above. In most periods this response is negative, thus suggesting that an increase in the real price of oil should be followed by a decrease in the nominal interest rate. This result can be viewed as a consequence of the structure of the *BG* model, in which regulators can contrast the negative effects of oil prices on output by implementing an expansionary policy. However, for a few periods, in particular right after the large oil price increases of 1973 – 74 and 1979 – 80, the response to oil prices is positive. In these periods the estimated parameters in the *BG* model imply an overall positive impact of the real price of oil on the variables of interest to policymakers. In these circumstances, an increase in oil prices will require a contractionary policy response in order to contrast the origin of potentially explosive patterns.<sup>26</sup> In the remainder of the sample starting from the early 1980s, the policy response to oil prices in the composite model is almost always negative, but more moderate and stable than in the previous years, because of the greater stability of both economic variables and estimated parameters in this period.

One last important result emerges from the left panel of figure 7. For a long period of time, from around 1985 until roughly 2005, the optimal interest rate obtained from the composite problem is lower than the interest rate recommended by each of the approximating models. At first impact, this result might seem puzzling; why should model uncertainty make policymakers less precautionary in their policy choices? The answer to this question can be obtained by analyzing the optimal regulator problem described in section 2, together with the individual characteristics of the approximating model considered in this work. In the *BG* model, the policymaker can affect output directly and inflation indirectly through lagged output. In the *H* model, the policymaker can affect inflation but not output, nor directly or indirectly. In the *S* model, the

<sup>24</sup>Cogley and Sargent (2005b) already noticed this connection in the framework under analysis in their paper.

<sup>25</sup>Some evidence on the larger role of policy in the *BG* model is provided by Rondina (2010) which, using a framework similar to the one adopted in this paper, presents some impulse-response exercises comparing the reaction of the models' variables to an oil price shock in the different approximating models under different policy rules.

<sup>26</sup>More specifically, in these periods the increase in inflation caused by oil prices has a larger impact on output (through the real interest rate) than oil prices themselves. This might originate trajectories in which the values of both  $y_t$  and  $\pi_t$  diverge to infinity.

policymaker can affect output indirectly through inflation, but as I already mentioned this effect is quite small. The composite model is a weighted average of these three approximating models, in which the optimal policy instrument simultaneously responds to all variables in the state vector, that is to all the variables included in all models. As a consequence, in the robust problem the policy instrument can generate (and use) relationships between variables that are not possible in the underlying models. This rationalizes departures from the policies recommended by the individual models, and explains the optimal policy in the years 1985 – 2005.

The ability of the policy instrument to create correlations between variables in the different models is a consequence of the approach used to incorporate model uncertainty in this environment and it is not related to the introduction of oil prices in the economy. From a formal point of view, this outcome is caused by the fact that the transition law in (28) looses the original block diagonal form implied by (21) once the optimal policy rule for  $x_{t|t-1}$  is obtained from the linear regulator problem. The additional relationships between variables that characterize (28) relative to the individual models are of two forms. First, the recommendation for the policy instrument  $x_{t|t-1}$  is a linear function of the composite state  $S_{E,t-1}$ , which includes all variables present in all models. Therefore, by implementing the optimal policy the regulator will make each individual model depend also on those variables that were originally included only in other models. In this sense, the Bayesian optimal control problem allows policymakers to reduce model uncertainty by equalizing the set of variables that affect each individual model. Second, the state vector  $S_{E,t-1}$  also contains variables that are included in more than one model. Again, in the optimal policy rule the instrument  $x_{t|t-1}$  is a function of  $S_{E,t-1}$ , and this potentially generates relationships between all variables in all models, even between the same variable in different models. In other words, the robust regulator problem assumes that policymakers act as perfect forecasters of possible future patterns for the variables of interest, and these patterns account for the uncertainty on the model from which future variables will be generated. In expectation, these future trajectories can exhibit interactions between variables in the different models, which are generated and affected by the optimal policy rule. Thus, the Bayesian optimal control problem allows for potential gains from model uncertainty, in the sense that (in expectation) policymakers will be able to exploit correlations that do not exist in the individual models. In practice, whether the composite regulator problem described in section 2 will recommend a less or more precautionary optimal policy relative to the individual models depends on the specific models under analysis, and on the channels of transmission of monetary policy that they imply. In this framework, for a number of periods the correlations between variables allow regulators to implement a robust interest rate that is lower than the optimal interest rate obtained from the  $S$ ,  $BG$  and  $H$  models. My conjecture is that this result is likely to emerge in other environments in which the specifications in the model space are characterized by a very different impact of the policy instrument on the variables of interest to the policymaker.

As I already mentioned, the possibility to use interactions between the variables in the composite state vector is due to the specific approach proposed by Cogley and Sargent (2005b) and employed in this paper, which incorporates Bayesian model averaging to an optimal control type of problem. This result is not present in previous works in which Bayesian model averaging was adopted to select a robust policy among a predefined set of simple policy rules (see, for instance, Brock, Durlauf and West, 2007; Brock et al., 2007; Cogley et al., 2009; Rondina, 2010). It follows that a composite regulator problem as the one described in this work has relevant implications on the way policymakers view model uncertainty, and important consequences on the interpretation of the optimal policy recommendation. However, the literature is lacking a comprehensive treatment of the properties of this approach, particularly with respect to the differences with the other techniques that have been proposed to account for model uncertainty in economic policy. For this reason, I believe that further investigation in this direction could offer interesting insights, and provide a fuller understanding of the implications that different methods have for policymakers' behavior in face of model uncertainty.

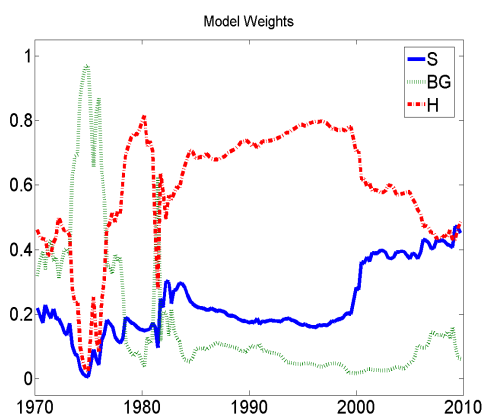
## 4 Sensitivity analysis

In this section, I explore the sensitivity of the results to changes in some of the baseline assumptions. First, I study the effects of different model priors on the pattern of the posterior probabilities, and the resulting impact on the optimal policy rule. Second, I investigate the robustness of the results to changes in the process for  $s_t$  and in the procedure used to estimate its parameters. Last, I analyze the role of policymakers' preferences, and in particular I examine how the recommended policy would have been different had the regulator only focused on a target for core CPI inflation.

### 4.1 Model priors

In the baseline case, an equal initial prior of  $1/3$  was attached to each of the approximating models. As I already mentioned, the literature in economics has proposed many different mechanisms through which oil prices can affect economic performance. In particular, some of these studies have focused on the demand-side effects of changes in the price of oil, while some other works have analyzed the impact of oil prices on the supply side of the economy. The  $S$  and  $BG$  models fall into the first group, while the  $H$  model can be considered as part of the latter. Therefore, it seems natural to consider a combination of model priors that takes into account this classification. In this spirit, half of the initial probability weight can be split equally between the demand-side specifications, and half can be attached to the supply-side specification. This originates the following initial priors:  $\alpha_{S,0} = \alpha_{BG,0} = 1/4$ ;  $\alpha_{H,0} = 1/2$ .

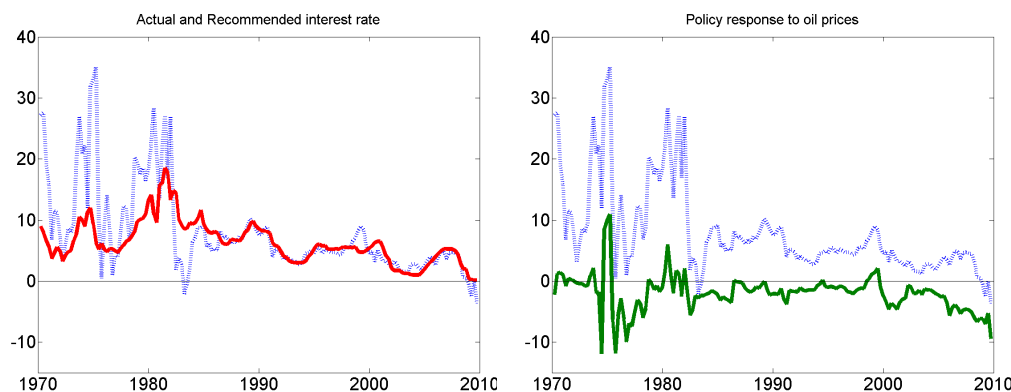
Figure 8 - Model weights -  $\alpha_{S,0} = \alpha_{BG,0} = 1/4$ ;  $\alpha_{H,0} = 1/2$



Note: Relative posterior probability of each approximating model, computed using the procedure explained in Appendix 2, with the alternative prior specification  $\alpha_{S,0} = \alpha_{BG,0} = 1/4$ ;  $\alpha_{H,0} = 1/2$ .

Figure 8 reports the pattern of the model weights when this different set of prior probabilities is adopted. Figure 9 shows the corresponding optimal policy recommendation and direct response to oil prices. The pattern of the relative posterior probabilities is very similar to the benchmark case, but the higher prior of the  $H$  model is reflected on a higher weight attached to this model in the entire sample period. The policy recommendation and the direct response to the real price of oil are almost the same as in the baseline case. This result was somehow expected, since the pattern of the model weights is not greatly affected by the change in prior probabilities, so the  $BG$  model still dominates the optimal regulator problem.

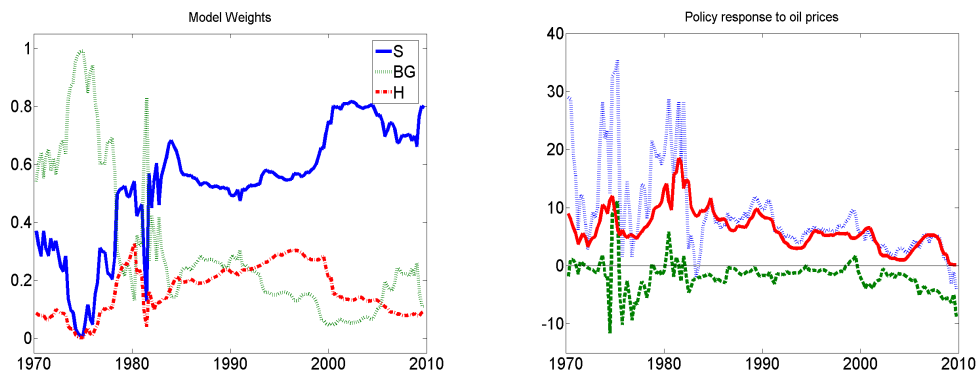
Figure 9 - Policy recommendation - composite model,  $\alpha_{S,0} = \alpha_{BG,0} = 1/4$ ;  $\alpha_{H,0} = 1/2$ .



Note: The left-hand panel reports the policy recommended by the optimization problem described in section 2 (dashed line) together with the annual Federal Funds rate (continuous line). The right-hand panel reports the fraction of the optimal policy that represents the direct response to the real price of oil. The initial priors are  $\alpha_{S,0} = \alpha_{BG,0} = 1/4$ ;  $\alpha_{H,0} = 1/2$ .

Finally, I considered two more sets of prior probabilities. In the first case, I attached a low initial prior (1/10) to the supply-side specification. In the second case, the low prior was attached to the demand-side specifications. The pattern of the model weights and the optimal policy recommend by the regulator problem in these two cases are reported in figure 10 and figure 11. These figures show that more extreme departures from the baseline set of priors considerably affect the behavior of the model weights, even if some patterns, like the increase in the relative posterior of the *BG* model around the first oil price shock, remain. In general, a higher prior increases the relative posterior of the approximating model for the entire length of the sample. This result is due to the fact that the changes in posterior probabilities generated by the new data are quite similar across the approximating models.<sup>27</sup> Thus, the relative posterior probabilities are only slightly affected by the new information that becomes available over time, and the impact of the initial priors on the model weights is carried throughout the sample.

Figure 10 - Model weights and policy recommendation -  $\alpha_{S,0} = \alpha_{BG,0} = 9/20$ ;  $\alpha_{H,0} = 1/10$



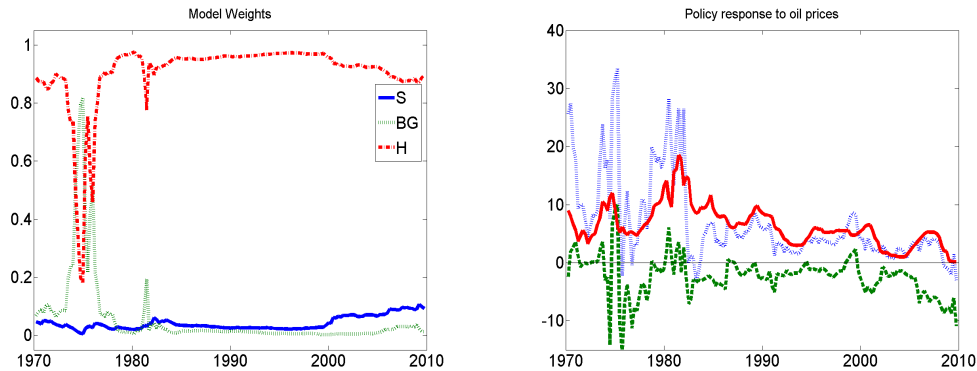
Note: The left-hand panel reports the relative posterior probability of each model specification, computed using the procedure explained in Appendix 2. The right-hand panel reports the optimal policy (dotted line), the direct response to the real price of oil (dashed line), and the actual Federal Funds rate (continuous line). The initial priors are:  $\alpha_{S,0} = \alpha_{BG,0} = 9/20$ ;  $\alpha_{H,0} = 1/10$ .

The different patterns of the model weights implied by the different assumptions on the initial priors affect both the optimal interest rate and the direct policy response to oil prices. This statement can be verified by comparing the right panels in figures 10 and 11 with figure 5, that reported the results for the baseline scenario.

<sup>27</sup>This finding actually confirms the evidence that this environment is characterized by a problem of model uncertainty.

However, while the model weights change significantly based on the chosen set of priors, the optimal policy is affected to a much smaller extent. In particular, the policy recommendations are subject to some variations in magnitude but their general pattern in the sample period under analysis remains almost unchanged.<sup>28</sup> Therefore, we can conclude that the *BG* model keeps playing a dominant role in the regulator problem, even when its weight is close to zero. In this sense, in this environment the presence of model uncertainty itself is more important than the actual pattern of the weights attached to each approximating model.

Figure 11 - Model weights and policy recommendation -  $\alpha_{S,0} = \alpha_{BG,0} = 1/20$ ;  $\alpha_{H,0} = 9/10$



Note: The left-hand panel reports the relative posterior probability of each model specification, computed using the procedure explained in Appendix 2. The right-hand panel reports the optimal policy (dashed line), the direct response to the real price of oil (dashed line), and the actual Federal Funds rate (continuous line). The initial priors are:  $\alpha_{S,0} = \alpha_{BG,0} = 1/20$ ;  $\alpha_{H,0} = 9/10$ .

## 4.2 The process for the real price of oil

In this subsection I investigate the extent to which the optimal policy is affected by the definition of the process for the real price of oil and by the assumptions used to estimate its parameters. Notice that alternative formulations of  $s_t$  will not modify the models' relative posterior probabilities, since this process does not differ across specifications. Thus, none of the variations that I consider in this section will change the pattern of the weights reported in figure 4. However, the real price of oil affects the behavior of the variables of interest to policymakers, so that changes in (7) and in the pattern of its estimated parameters have consequences on the regulator problem and can potentially alter the optimal policy recommendations.

### 4.2.1 Estimation priors

Previous works using MCMC methods to estimate processes in the form of (7) have shown that the assumptions on the initial prior distributions are important for the speed of convergence of the algorithm, and for the value of the estimated parameters (see, for instance, Cogley and Sargent, 2001 and 2005a; Primiceri, 2005). Thus, in this subsection I investigate how the baseline results are affected by changes in the inverse Wishart priors  $IW(B_{o,0|0}, v_{o,0})$  and  $IW(B_{\varepsilon,0|0}, v_{\varepsilon,0})$ , and in the Gaussian prior  $N(\theta_{s,0}, P_{s,0})$ .

As explained above, in the baseline scenario the prior  $p(V)$  is tighter than those usually adopted in the literature. A decrease in the degrees of freedom  $v_{o,0}$  and  $v_{\varepsilon,0}$  would make this prior more disperse and less informative, hence allowing the data to play a larger role in the estimation of  $\sigma_o^2$  and  $\sigma_\varepsilon^2$ . However, given

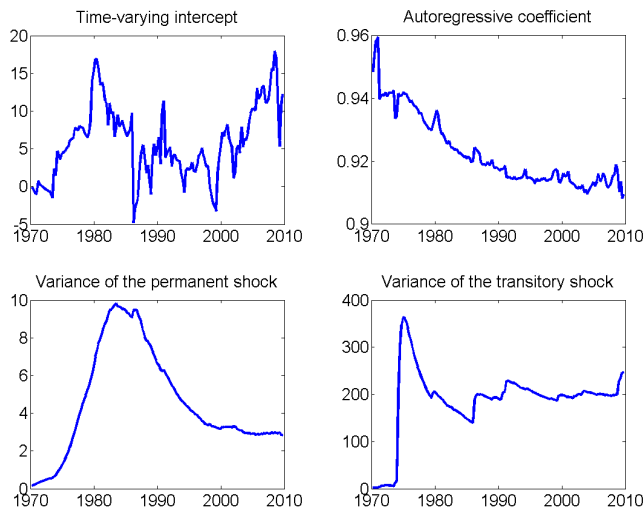
<sup>28</sup>Notice that the robust interest rate is lower in periods in which the relative posterior probability of the *H* model moves closer to one. In these circumstances, the composite optimization problem attaches a near one weight to a specification in which monetary policy has a very limited impact on the variables of interest to policymakers. As a consequence, this is a situation in which the relationships between models created by the robust policy rule have the potential to deliver the greatest benefits.



the high volatility in the historical series for  $s_t$ , reducing the initial degrees of freedom could slow the speed of convergence of the algorithm and require additional adjustments that do not seem consistent with the assumption of learning on the side of the policymaker.<sup>29</sup> In the baseline scenario,  $v_{o,0} = v_{\varepsilon,0} = 43$  (the degrees of freedom in the training sample). I experimented with a few alternative numbers for the initial degrees of freedom, decreasing them as long as the convergence of the algorithm was providing acceptable results. Setting the degrees of freedom to 15 would already raise the convergence issues mentioned above. Therefore, in figures 12 I report the results for  $v_{o,0} = v_{\varepsilon,0} = 20$ .

The main difference with the baseline scenario is that in this case the policymaker substantially overestimates the variance  $\sigma_\varepsilon^2$  in the first part of the sample. Then, he reviews his beliefs over time, and eventually converges to values that are similar to those obtained in the baseline estimation. The consequences on the estimated autoregressive coefficient and time-varying intercept are that  $\rho$  decreases more quickly, while the pattern of  $\delta_t$  reaches higher values in the initial part of the sample, when  $\sigma_\varepsilon^2$  is higher. However, the differences with the values of  $\rho$  and  $\delta_t$  displayed in figure 2 are not very large, so the resulting optimal policy and direct response to the real price of oil are similar to the baseline scenario, and are not reported.

Figure 12 - Parameters in the process for the real price of oil



Note: Pattern of the parameters in the process for the real price of oil, in the case  $v_{o,0} = v_{\varepsilon,0} = 20$ . The top panel reports the time-varying intercept  $\delta_t$  (left) and autoregressive coefficient  $\rho$  (right) in (7). The bottom panel reports the variances  $\sigma_\varepsilon^2$  (right) and  $\sigma_o^2$  (left). These parameters were estimated using the Gibbs sampler procedure described in Appendix 2.

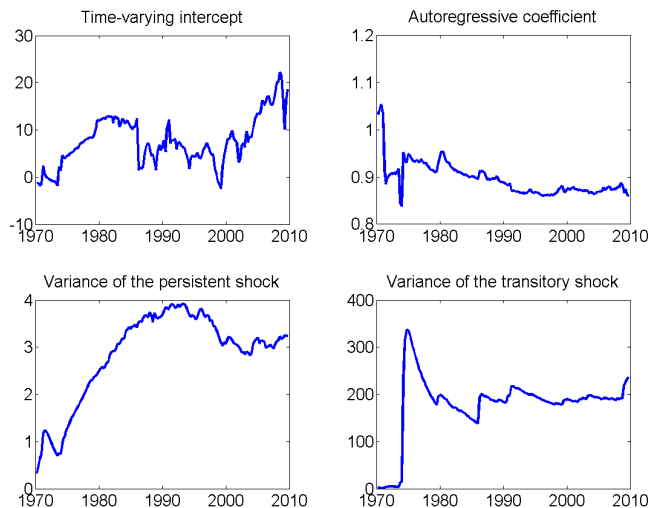
In a second exercise, I set  $v_{o,0} = v_{\varepsilon,0} = 15$ ,  $IW(0.9^2 B_{o,0|t}, v_{o,0})$ ,  $IW(0.9^2 B_{\varepsilon,0|t}, v_{\varepsilon,0})$  and  $N(\theta_{s,0}, 5^2 P_{s,0})$ , for any  $t = 0, \dots, T$  with  $T = 2009$ : *III*. In this way, the prior on  $V$  attaches a smaller weight to the estimates of  $\sigma_o^2$  and  $\sigma_\varepsilon^2$  coming from the previous period, and the Gaussian prior on the value of  $\theta_s$  at time 0 is a lot more disperse than in the baseline scenario. Given the lower degrees of freedom, the decrease in the scale parameters in the inverse Wishart priors was introduced to facilitate the convergence of the sampler. On the other hand, the increase in the variance of the Gaussian prior ensures that the total time variation is consistent with that in the baseline scenario. This choice of prior distributions allows the data to play a greater role in

<sup>29</sup>More specifically, reducing the degrees of freedom might require the policymaker to significantly reduce  $\widehat{B}_{o,0|0}$  and  $\widehat{B}_{\varepsilon,0|0}$  in order to guarantee a quick convergence of the Gibbs sampler. At each time  $t$ , this would make the estimation of the joint posterior  $p(\theta_s^t, V | s^t)$  heavily dependent on the data alone, with almost no role for the value of the covariance matrix  $V$  estimated at time  $t - 1$ . This result does not appear to be consistent with the idea of a policymaker learning over time and it seems more suggestive of an environment in which the monetary authority changes in every period and decides not to take into account the estimates of their predecessors.

the estimation of the joint posterior  $p(\theta_s^t, V | s^t)$ . Figure 13 shows that in this case  $\sigma_\varepsilon^2$  moves more rapidly to values in the range of its final estimate and, as a consequence,  $\rho$  decreases more quickly than in the baseline scenario. However, in the last part of the sample, the autocorrelations between the draws obtained from the Gibbs sampler are not as small as in the previous choices of priors, especially for  $\sigma_\varepsilon^2$ . This suggests that a higher number of draws might have been necessary to obtain a better convergence of the algorithm. In addition, this prior parameterization generates estimates of  $\sigma_\varepsilon^2$  that are more volatile. In the first few periods, this produces a value of  $\rho$  that is higher than in the baseline scenario, and that makes the composite problem not stabilizable. Apart from this initial part of the sample, the optimal policy and direct response to oil prices obtained using the estimates depicted in figure 13 are again very close to those in the baseline scenario.

A lower number of degrees of freedom would require even smaller values of the scale parameters in the inverse Wishart priors which, as I previously mentioned, does not seem consistent with the idea of learning. On the other hand, leaving  $IW(B_{o,0|t}, v_{o,0})$  and  $IW(B_{\varepsilon,0|t}, v_{\varepsilon,0})$  as in the baseline scenario while increasing the variance in the Gaussian prior would just increase the amount of time variation without changing the pattern of  $\sigma_o^2$  and  $\sigma_\varepsilon^2$ .

Figure 13 - Parameters in the process for the real price of oil



Note: Pattern of the parameters in the process for the real price of oil, in the case  $v_{o,0} = v_{\varepsilon,0} = 15$ ,  $IW(0.9^2 B_{o,0|t}, v_{o,0})$ ,  $IW(0.9^2 B_{\varepsilon,0|t}, v_{\varepsilon,0})$  and  $N(\theta_{s,0}, 5^2 P_{s,0})$ . The top panel reports the time-varying intercept  $\delta_t$  (left) and autoregressive coefficient  $\rho$  (right) in (7). The bottom panel reports the variances  $\sigma_\varepsilon^2$  (right) and  $\sigma_o^2$  (left). These parameters were estimated using the Gibbs sampler procedure described in Appendix 2.

Finally, I also considered asymmetric changes in  $IW(B_{o,0|0}, v_{o,0})$  and  $IW(B_{\varepsilon,0|0}, v_{\varepsilon,0})$  by allowing one of the two prior distributions be more disperse than the other. The choice:  $IW(B_{o,0|t}, v_{o,0})$ ,  $IW(0.9^2 B_{\varepsilon,0|t}, v_{\varepsilon,0})$  and  $N(\theta_{s,0}, 5^2 P_{s,0})$ , with  $v_{o,0} = 43$  and  $v_{\varepsilon,0} = 20$  generated results that were similar to those reported in figure 13, even if the estimated value of  $\sigma_\varepsilon^2$  was smaller in magnitude. Allowing for less degrees of freedom did not guarantee convergence of the algorithm, especially around the large oil price shocks and sometimes at the end of the sample period. Other asymmetric parameterizations of the joint prior  $p(\theta_{s,0}, V)$  delivered results that were either quite similar to the baseline scenario or uninteresting, because one of the two variances was capturing the entire variation in  $s_t$ .

### 4.2.2 The mean shifting component

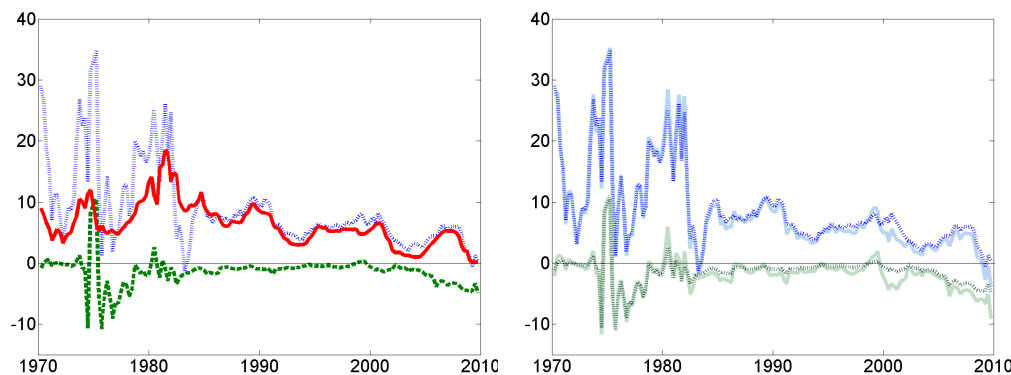
In the baseline framework, the process for the real price of oil was described by (7) – (10). In this section, I investigate whether the results reported in section 3 change if I assume that  $s_t$  follows an exogenous zero-mean AR(1) process in the form:

$$s_t = \rho s_{t-1} + o_t \quad (30)$$

with  $\rho = 0.97$ . This is the specification adopted by Blanchard and Gali (2007).

Figure 14 show that the optimal policy, and especially the direct response to the real price of oil, are less volatile when (7) is substituted by (30). This result was somehow expected, since the process for the real price of oil described by (30) is not subject to changes in mean, and because its autoregressive coefficient is not adjusted over time. The differences with the baseline scenario are emphasized in the right-hand panel of figure 14. These differences are larger (a few percentage points) in some periods; however, they are in general not too big in magnitude. This happens for two reasons. First, policymakers are still assumed to estimate and update the parameters of (1) – (6) using the approach described in section 2. Therefore, even if the process for the real price of oil is different, the way it affects the other variables of interest is not. Second, even if there is no permanent innovation  $\varepsilon_t$  in (30),  $o_t$  is still extremely persistent, and this feature seems to have a large impact on the magnitude of the optimal policy recommendation.<sup>30</sup>

Figure 14 - Policy recommendation - process for  $s_t$  as in Blanchard and Gali (2007)



Note: The left-hand panel reports the optimal policy (dotted line), the direct response to the real price of oil (dashed line), and the actual Federal Funds rate (continuous line), in the case the process for the real price of oil is as in Blanchard and Gali (2007). The right-hand panel compares the results for this case (dotted lines) with those obtained in the baseline scenario (continuous lines).

### 4.3 Policymakers' preferences

Last, I examined the effects of a change in policymakers' preferences. In particular, I investigated the extent to which the optimal policy recommendations would have been different had the monetary authority decided to adopt an inflation targeting type of objective. Given that in this framework the different approximating models vary in terms of the variables that can be directly affected by monetary policy, changes in the relative weight of the output gap and inflation rate in policymakers' loss function have the potential to deliver results that are different from the baseline case. Notice, however, that policy decisions do not affect the parameter distributions; for this reason, variations in regulators' preferences will not alter the pattern of the model weights reported in figure 4.

<sup>30</sup>Indeed, I experimented with values of  $\rho$  in the range 0.6 – 0.8 and I obtained that both the optimal policy and the direct response to oil prices were several percentage points apart from the results of the baseline scenario.

Inflation targeting was introduced in the optimization problem by changing the matrix  $Q$  in (27). In particular, I defined an alternative matrix  $Q'$  as:

$$Q' = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \lambda_\pi & 0 \\ 0 & 0 & \lambda_i \end{bmatrix}$$

with  $\lambda_\pi = 1$  and  $\lambda_i = 0.1$  as in the baseline scenario.<sup>31</sup> If  $\pi^* = 2\%$ , substituting  $Q'$  for  $Q$  in the regulator problem implies that policymakers are aiming for a 2% target on core CPI inflation, while stabilizing the output gap is no longer in their objective function.

I found that the optimal policy and the direct response to the real price of oil in the case of inflation targeting are almost the same as in the baseline scenario. This suggests that the policy recommendations obtained in section 3 are mostly driven by policymakers' concern for stabilizing the core inflation rate, so that assuming an inflation targeting type of policy does not change their priorities relative to the baseline problem. Moreover, given the asymmetric role of the three approximating models in the composite problem, this result implies that, ultimately, the central factor influencing the robust policy rule is the need to control core CPI inflation in the Blanchard-Gali model. The conclusion that the optimal policy recommendations are primarily determined by the behavior of the inflation variable is supported by the fact that, on the other hand, the value of the policy instrument changes significantly if the monetary authority is assumed to care more about the output gap than core CPI inflation. This statement can be easily verified by considering an optimization problem in which the original  $Q$  matrix with a smaller value of  $\lambda_\pi$  is used instead.<sup>32</sup>

I also considered a few additional variations in policymakers' preferences. I found that the results in the baseline scenario are robust to changes in the value of the target inflation  $\pi^*$  in the range 0.5% – 4%. Finally, setting  $\lambda_i = 0$  only makes the policy recommendations more volatile in the fraction of the sample before 1985, while it leaves both the optimal policy and the direct response to oil prices almost unchanged after this year.

## 5 Expectations formation

The formalization of expectations on future variables in an environment characterized by model uncertainty is not straightforward. The main problem refers to the definition of private agents' information set, with regard in particular to the inclusion in this set of the uncertainty on the true data generating process. In other words, should model uncertainty be incorporated in the public's expectations of future inflation? And should the public's uncertainty be of the same form as the one that characterizes policymakers' decision process? And more, if the public's and the policymaker's information sets, upon which expectations and policy decisions are based, are different, how should we think about possible iterations between them?

The literature on policy design and evaluation in environments characterized by model uncertainty hasn't provided an organic answer to these questions, yet. A recent contribution by Cogley et al. (2009) has opted for expectations to be formed rationally, but internally in each model. In other words, each of the approximating models included in the policy decision process is estimated under the assumption that the public has rational expectations and believes that this specific model generates the data. This approach eliminates all issues (for instance, signal extraction and/or higher order expectations) that might arise when the public and policymakers both account for model uncertainty, but using different information on the true data generating

<sup>31</sup> Thus, I assume that the policymaker still has a preference for interest rate smoothing.

<sup>32</sup> The differences with the baseline scenario become particularly evident if  $\lambda_\pi$  is decreased to less than 0.5.

process.<sup>33</sup> However, it generates expectations on the same future variable that are different in the different model specifications, and this result might be undesirable in some circumstances. An alternative approach is the use of survey data, as in Brock, Durlauf and West (2007). However, this technique is sometimes not feasible, especially when the focus of the analysis is the interpretation of historical behavior, because survey data is often only available for relatively recent years.<sup>34</sup>

A different procedure, proposed again by Brock, Durlauf and West (2007), is to allow the regulator problem to account for alternative ways in which expectations might be formed. More specifically, policymakers can treat the imperfect knowledge on the way private agents form their expectations as an additional form of model uncertainty, and incorporate it into policy evaluation. This is a sound way to proceed, but it might significantly increase the dimension of the model space, especially in environments in which the monetary authority is already concerned about several other forms of uncertainty.

The issue of expectations formation in an environment characterized by model uncertainty and learning is not minor. Indeed, different assumptions on the way the public forms its expectations can affect not only the models' estimated parameters, but also their posterior probabilities and policymakers' learning on the true data generating process. Intuitively, there might be circumstances in which expectations will help the regulator discriminate among different model specifications, and others in which the opposite will happen. The purpose of this work was to focus on the consequences for monetary policy of the uncertainty about the channel of transmission of oil prices into the economy. For this reason, I decided to sidestep the issues related to the expectation formation process, and I assumed that the public adopts a simple backward rule for predicting the value of future variables. However, I am aware of the fact that this assumption is quite restrictive, and I do believe that further research in this direction could provide interesting insights.

## 6 Conclusions

Using US postwar data, this paper examined the optimal policy recommended by a regulator problem that explicitly accounts for the uncertainty on the channel through which oil prices affect the economy, and in which policymakers learn from the new data that becomes available over time. In this environment, I find that model uncertainty plays a relevant role on two dimensions. First, in terms of the direct response to the real price of oil, this paper shows that one of the approximating models under consideration, namely the *BG* model, drives policymakers' decisions. This result is robust to changes in the pattern of the weights attached to the individual models in the optimization process, and can therefore be attributed to the presence of model uncertainty itself. Second, I show that for a large portion of the sample period under analysis the optimal interest rate for the composite model encompassing model uncertainty is lower than the optimal interest rate for any of the underlying models of the economy. This result is due to the specific approach adopted in this work to compute the robust policy rule, i.e. a linear quadratic dynamic programming problem incorporating Bayesian model averaging, which allows regulators to exploit correlations between variables that do not exist in the individual models.

The role of learning over the parameters and posterior probabilities of the approximating models is also investigated. I show that policymakers' beliefs over the model that best approximates the data, which correspond

---

<sup>33</sup>This issue seems to be particularly relevant in environments characterized by learning, as in this paper.

<sup>34</sup>For instance, Brock, Durlauf and West (2007) use price expectations from the Survey of Professional Forecasters to compute expected inflation. This data only goes back to 1968, so in the framework adopted in this paper it would not provide enough information to be able to estimate the initial beliefs on the parameters of each model. One possible solution would be to set the initial beliefs to some arbitrary value, and use the data from the Survey of Professional Forecasters in the learning process starting from 1970 : *I*. This extension of the paper is left for future research.

to the models' relative posterior probabilities, are affected by the chosen set of priors. This happens because the changes in posteriors are very similar across models, which suggests that the arrival of new data does not offer much support in discriminating among the different specifications. This conclusion confirms previous findings about the presence of model uncertainty in this environment. In the process for the real price of oil, I similarly find that policymakers' learning pattern is influenced by the chosen prior distributions. However, the perceived persistence of this variable, which seems to be a leading factor for policy decisions, is only marginally affected by these changes.

This work could be extended in a few directions. First, I examined the measure in which different assumptions on the process for the real price of oil can change the optimal policy recommendation. However, all the specifications that I considered in this study modeled oil prices as an exogenous variable. Given that, as I mentioned in the main text, there is some debate in the literature about the direction of the contemporaneous relationship between oil prices and domestic variables, I think it would be interesting to consider alternative definitions of the process for the real price of oil that can account for this issue.

Second, this paper shows that the implementation of Bayesian model averaging in a optimal linear regulator type of problem as proposed by Cogley and Sargent (2005b) has important consequences on policymakers' approach to model uncertainty. In particular, the way in which the composite regulator problem creates relationships between variables in the different models is new in the literature using Bayesian methods to incorporate model uncertainty into policy evaluation. I believe that further investigation in this direction could provide a better understanding of the implications of the different techniques, and offer important insights for the interpretation of the robust policies.

Finally, as discussed in the last section of the paper, I think that it would be very interesting to examine possible ways of embodying expectations formation in this framework. This would require taking a position on whether, and to what extent, private agents should share policymakers' concern for model uncertainty.

# Appendix 1

## Data description

The variables used in the empirical analysis are the following:

- $y_t$  is the output gap, computed as the difference between real GDP and the CBO estimate of potential GDP, both expressed in logs.
- $\pi_t$  is the annualized difference in log core CPI, where core CPI is the "CPI for all urban consumers: all items less energy products".
- $s_t$  is the real price of oil, defined as the difference between the nominal price of oil and core CPI, both expressed in logs. The nominal price of oil is the West Texas Intermediate spot oil price, while core CPI is the same used to compute  $\pi_t$ . The variable was rescaled so that  $s_{1970-I} = 0$ .
- $i_t$  is the annualized average Federal Funds rate.

All the data was obtained from the Federal Reserve Bank of St. Louis web site. The data is quarterly and includes observations from 1957 – II to 2009 – III. The training sample used to estimate the parameters in the initial priors (11) and (17) goes from 1957 – II to 1969 – IV, with data from 1957 – II to 1958 – III used to provide lags. The optimization and learning recursion starts in 1970 – I.

# Appendix 2

## Updating the parameters in the process for $s_t$

I use a MCMC algorithm in the form of a Gibbs sampler to simulate draws from the joint posterior  $p(\theta_s^t, V | s^t)$ , which are then used to obtain estimates of  $\theta_s^t$ ,  $\sigma_o^2$  and  $\sigma_\varepsilon^2$ . This algorithm is repeated in each period to account for the new information that becomes available. More specifically, the Gibbs sampler and learning procedure that I use in this paper can be summarized in the following three steps. The first two steps are the same as in Cogley and Sargent (2001); the third step is added to incorporate policymakers' learning from the new data.

*Step 1: drawing  $\theta_s^t$  given the history of data  $s^t$  and the covariance matrix  $V$ .*

Given information up to time  $t - 1$  and before information for time  $t$  is revealed, the policymaker starts with the prior  $p(\theta_{s,0})$  specified by (13), and with the prior  $p(V) = IW(B_{o,0|t-1}, v_{o,0}) IW(B_{\varepsilon,0|t-1}, v_{\varepsilon,0})$ . At time  $t = 1$ ,  $p(V)$  assumes the parameterization defined by (12).

Given the history of data and  $V$ ,  $p(\theta_s^t | s^t, V)$  is Gaussian, and can be factored as:

$$p(\theta_s^t | s^t, V) = p(\theta_{s,t} | s^t, V) \prod_{j=1}^{t-1} p(\theta_{s,j} | \theta_{s,j+1}, s^j, V) \quad (31)$$

For  $j = 1, \dots, t$ , let  $\theta_{s,j|j} = E(\theta_{s,j} | s^j, V)$ ;  $P_{s,j|j-1} = Var(\theta_{s,j} | s^{j-1}, V)$ ;  $P_{s,j|j} = Var(\theta_{s,j} | s^j, V)$ . These conditional means and variances can be updated recursively using the Kalman filter, given the initial values  $\theta_{s,0}$  and  $P_{s,0}$ , and given the values of  $\sigma_o^2$  and  $\sigma_\varepsilon^2$  estimated in the previous period. In the framework adopted in this paper, the updating formulas are:

$$\begin{aligned}
K_{s,j} &= P_{s,j|j-1} X_{s,j} (X'_{s,j} P_{s,j|j-1} X_{s,j} + \sigma_o^2)^{-1} \\
\theta_{s,j|j} &= \theta_{s,j-1|j-1} + K_{s,j} (s_j - X'_{s,j} \theta_{s,j-1|j-1}) \\
P_{s,j|j-1} &= P_{s,j-1|j-1} + Q \\
P_{s,j|j} &= P_{s,j|j-1} - K_{s,j} X'_{s,j} P_{s,j|j-1}
\end{aligned}$$

This forward recursion delivers the mean and variance of the vector of coefficients  $\theta_{s,t}$  at the terminal period  $t$ :

$$p(\theta_{s,t} | s^t, V) = N(\theta_{s,t|t}, P_{s,t|t}) \quad (32)$$

A value for  $\theta_{s,t}$  can be drawn from this distribution. This value, together with the output of the forward Kalman recursion, can be used to compute the mean and variances of  $p(\theta_{s,j} | \theta_{s,j+1}, s^j, V)$ , for any  $j = 1, \dots, t-1$ . More specifically, define:  $\theta_{s,j|j+1} = E(\theta_{s,j} | \theta_{s,j+1}, s^j, V)$ ;  $P_{s,j|j+1} = \text{Var}(\theta_{s,j} | \theta_{s,j+1}, s^j, V)$ . Then, we have:

$$p(\theta_{s,j} | \theta_{s,j+1}, s^j, V) = N(\theta_{s,j|j+1}, P_{s,j|j+1}) \quad j = 1, \dots, t-1 \quad (33)$$

The parameters in these distributions can be computed using the following backward recursion:

$$\begin{aligned}
\theta_{s,j|j+1} &= \theta_{s,j|j} + P_{s,j|j} P_{s,j+1|j}^{-1} (\theta_{s,j+1} - \theta_{s,j|j}) \\
P_{s,j|j+1} &= P_{s,j|j} - P_{s,j|j} P_{s,j+1|j}^{-1} P_{s,j|j}
\end{aligned}$$

Then, starting from a draw of  $\theta_{s,t}$  from (32), a complete history of the vector of parameters  $\theta_s$  can be drawn using (33).

Notice that here, because of the particular structure of the framework under analysis, this backward recursion is much simplified. It is easy to show that the updating formulas in the Kalman recursion imply that  $P_{s,j|j+1}$  is a matrix of zeros, except for the term in the first row, first column. In other words,  $P_{s,j|j+1}$  has the same structure as  $Q$ . This implies that we will obtain an history only for the element in the first row of  $\theta_s$ , i.e.  $\delta$ , while for  $\rho$  we have:  $\rho_j = \rho_{j|j+1}$ . In addition, from the backward recursion formulas, it can be shown that for  $\rho$  we have:  $\rho_{j|j+1} = \rho_{j+1}$ . Then,  $\rho_j = \rho_{j+1}$  for any  $j = 1, \dots, t-1$ , which is consistent with the assumptions on the parameters in the process for the real price of oil. In other words, at each time  $t$ , the sampler delivers a complete history  $\delta^t$  and a single value for  $\rho$ . This also implies that (33) can be rewritten as:

$$p(\theta_{s,j} | \theta_{s,j+1}, s^j, V) = N(\delta_{j|j+1}, P_{s,j|j+1}^{11}) \quad j = 1, \dots, t-1 \quad (34)$$

where  $\delta_{j|j+1}$  is the first element in the vector  $\theta_{s,t|t+1}$ , and  $P_{s,t|t+1}^{11}$  is the element in the first row, first column of the matrix  $P_{s,t|t+1}$ . The fact that, in each iteration of the Gibbs sampler, we can draw a backward history only for one element of the vector of coefficients reduces the computational burden of the estimation.<sup>35</sup>

*Step 2: drawing  $V$  given the history of data  $s^t$  and  $\theta_s^t$ .*

Given the history of data and parameters, and given the transition law in (9), the innovations are i.i.d. conditionally normal random variables. In addition, given  $s^t$  and  $\theta_s^t$ , they are observable. Therefore, a draw for  $V$  can be obtained from:

$$p(V | \theta_s^t, s^t) = p(\sigma_o^2 | \theta_s^t, s^t) \cdot p(\sigma_\varepsilon^2 | \theta_s^t, s^t)$$

<sup>35</sup>Notice, however, that even if we need to draw a complete backward history only for the first element of  $\theta_s$ , because of the covariances between coefficients, we still need to carry the complete vectors  $\theta_{s,j+1}$  and  $\theta_{s,j|j}$  for computing the first element of  $\theta_{s,j|j+1}$ .



where:

$$\begin{aligned}
p(\sigma_o^2 | \theta_s^t, s^t) &= IW(B_{o,t|t-1}, v_{o,t}) \\
B_{o,t|t-1} &= B_{o,0|t-1} + \sum_{j=1}^t \widehat{\sigma}_j \widehat{\sigma}_j' \\
v_{o,t} &= v_{o,0} + t
\end{aligned}$$

and

$$\begin{aligned}
p(\sigma_\varepsilon^2 | \theta_s^t, s^t) &= IW(B_{\varepsilon,t|t-1}, v_{\varepsilon,t}) \\
B_{\varepsilon,t|t-1} &= B_{\varepsilon,0|t-1} + \sum_{j=1}^t \widehat{\varepsilon}_j \widehat{\varepsilon}_j' \\
v_{\varepsilon,t} &= v_{\varepsilon,0} + t
\end{aligned}$$

where  $\widehat{\sigma}_j$  and  $\widehat{\varepsilon}_j$  are the observed innovations for period  $j$ .

Under regularity conditions<sup>36</sup>, the sequence of draws from the Gibbs sampler converges to a draw from the joint distribution  $p(\theta_s^t, V | s^t)$ .

*Step 3: updating the prior  $p(V)$ .*

The first two steps of the algorithm provide a number of draws from the posterior  $p(\theta_s^t, V | s^t)$ . We can compute the estimated history  $\theta_s^t$  and variances  $\sigma_o^2$  and  $\sigma_\varepsilon^2$  as the averages of the retained draws. Then, we can use this information to update the prior  $p(V)$ . More specifically, let  $\sigma_{o,t}^2$  and  $\sigma_{\varepsilon,t}^2$  be the estimated values obtained in period  $t$  using information up to that time. Then, the prior for period  $t+1$  will be set as:

$$p(V) = IW(B_{o,0|t}, v_{o,0}) \cdot IW(B_{\varepsilon,0|t}, v_{\varepsilon,0})$$

where:  $B_{o,0|t} = v_{o,0} \cdot \sigma_{o,t}^2$  and  $B_{\varepsilon,0|t} = v_{\varepsilon,0} \cdot \sigma_{\varepsilon,t}^2$ .

## Updating the parameters of the approximating models

For each model, the prior distribution on the model parameters before the information about time  $t$  is revealed, is given by:

$$\begin{aligned}
p(\theta, \Sigma | Z^{t-1}) &= p(\theta | \Sigma, Z^{t-1}) \cdot p(\Sigma | Z^{t-1}) \\
&= N(\theta_{t-1}, P_{t-1}) \cdot IW(B_{t-1}, v_{t-1})
\end{aligned}$$

As discussed in the main text, in each model the conditional likelihood function is a multivariate normal distribution. Thus, the posterior distribution on the model parameters after data for time  $t$  is observed will also be a normal-inverse Wishart distribution:

$$\begin{aligned}
p(\theta, \Sigma | Z^t) &= p(\theta | \Sigma, Z^t) \cdot p(\Sigma | Z^t) \\
&= N(\theta_t, P_t) \cdot IW(B_t, v_t)
\end{aligned}$$

Starting from the initial values  $\theta_0$ ,  $P_0$ ,  $B_0$  and  $v_0$ , the updating formulas for the parameters of these distributions are:

---

<sup>36</sup>Further details can be found in Roberts and Smith (1992).

$$B_t = \sum_{j=1}^t [Y_j - X_j' \theta_t] [Y_j - X_j' \theta_t]' = \begin{bmatrix} \sum_{j=1}^t (y_j - X_{y,j}' \theta_{y,t})^2 & \sum_{j=1}^t (y_j - X_{y,j}' \theta_{y,t}) (\pi_j - X_{\pi,j}' \theta_{\pi,t}) \\ \sum_{j=1}^t (y_j - X_{y,j}' \theta_{y,t}) (\pi_j - X_{\pi,j}' \theta_{\pi,t}) & \sum_{j=1}^t (\pi_j - X_{\pi,j}' \theta_{\pi,t})^2 \end{bmatrix}$$

$$v_t = v_{t-1} + 1$$

The covariance matrix  $\Sigma_t$  can be estimated as:  $\Sigma_t = \begin{bmatrix} \sigma_{y,t}^2 & \sigma_{y\pi,t} \\ \sigma_{y\pi,t} & \sigma_{\pi,t}^2 \end{bmatrix} = B_t/v_t$

For the vector of coefficients we have:

$$\theta_t = R_t^{-1} (R_{t-1} \theta_{t-1} + X_t Y_t)$$

or

$$\theta_t = \begin{bmatrix} \theta'_{y,t} & \theta'_{\pi,t} \end{bmatrix}'$$

$$\theta_{y,t} = R_{y,t}^{-1} (R_{y,t-1} \theta_{y,t-1} + X_{y,t} y_t)$$

$$\theta_{\pi,t} = R_{\pi,t}^{-1} (R_{\pi,t-1} \theta_{\pi,t-1} + X_{\pi,t} \pi_t)$$

where:

$$R_{y,t} = R_{y,t-1} + X_{y,t} X_{y,t}'$$

$$R_{\pi,t} = R_{\pi,t-1} + X_{\pi,t} X_{\pi,t}'$$

$$R_t = R_{y,t} \oplus R_{\pi,t} = R_{t-1} + X_t X_t'$$

Here,  $R_t$  is a  $k \times k$  matrix,  $R_{y,t}$  a  $k_y \times k_y$  matrix, and  $R_{\pi,t}$  a  $k_\pi \times k_\pi$  matrix. The initial values of these matrices were computed using data from the training sample:  $R_{y,0} = \sum_{j=1}^{T_0} X_{y,j} X_{y,j}'$ ,  $R_{\pi,0} = \sum_{j=1}^{T_0} X_{\pi,j} X_{\pi,j}'$  and  $R_0 = R_{y,0} \oplus R_{\pi,0}$  where  $T_0$  is the length of the training sample.

For the covariance matrix  $P_t$  we have:

$$P_t = R_t^{-1} \begin{bmatrix} \sum_{j=1}^t X_j \Sigma_t X_j' \end{bmatrix} R_t^{-1} = \begin{bmatrix} \sigma_{y,t}^2 R_{y,t}^{-1} & \sigma_{y\pi,t} \left[ R_{y,t}^{-1} \left( \sum_{j=1}^t X_{y,j} X_{\pi,j}' \right) R_{\pi,t}^{-1} \right] \\ \sigma_{y\pi,t} \left[ R_{y,t}^{-1} \left( \sum_{j=1}^t X_{y,j} X_{\pi,j}' \right) R_{\pi,t}^{-1} \right]' & \sigma_{\pi,t}^2 R_{\pi,t}^{-1} \end{bmatrix}$$

## Updating the model weights

The procedure used to update the model weights follows Cogley and Sargent (2005b). The marginalized likelihood function  $m_{it}$  in (18) is:

$$m_{it} = \int \int l(Y_i^t, X_i^t, \theta_i, \Sigma_i) p(\theta_i, \Sigma_i) d\theta_i d\Sigma_i$$

where  $l(Y_i^t, X_i^t, \theta_i, \Sigma_i)$  is the conditional likelihood for model  $i$  through date  $t$ , defined as:

$$l(Y^t, X^t, \theta, \Sigma) = \prod_{j=1}^t p(Y_j | X_j, \theta, \Sigma) \quad (35)$$

Using Bayes's Theorem,  $m_{it}$  can also be obtained as (see Cogley and Sargent, 2005b, for the derivation of this expression):

$$m_{it} = \frac{l(Y_i^t, X_i^t, \theta_i, \Sigma_i) p(\theta_i, \Sigma_i)}{p(\theta_i, \Sigma_i | Z_i^t)} \quad (36)$$

In the regulator problem, each model is weighted based on its posterior probability, which is computed in every period given the available data. From (18), the unnormalized weight attached to model  $i$  is defined as:

$$w_{it} = m_{it} \cdot p(M_i)$$

Then, following Cogley and Sargent (2005b), we can develop a recursion for  $w_{it}$  :

$$\begin{aligned} \frac{w_{it+1}}{w_{it}} &= \frac{m_{it+1} \cdot p(M_i)}{m_{it} \cdot p(M_i)} \\ &= p(Y_{it+1} | X_{it+1}, \theta_i, \Sigma_i) \frac{p(\theta_i, \Sigma_i | Z_i^{t+1})}{p(\theta_i, \Sigma_i | Z_i^t)} \end{aligned}$$

and taking logs:

$$\log w_{it+1} = \log w_{it} + \log p(Y_{it+1} | X_{it+1}, \theta_i, \Sigma_i) - \log \frac{p(\theta_i, \Sigma_i | Y^{t+1}, X^{t+1})}{p(\theta_i, \Sigma_i | Y^t, X^t)} \quad (37)$$

where  $p(Y_{it+1} | X_{it+1}, \theta_i, \Sigma_i)$  is the conditional likelihood for observation  $t + 1$  and  $p(\theta_i, \Sigma_i | Y^t, X^t)$  is the posterior distribution of the model's parameters given information up to date  $t$ . We know that the conditional likelihood  $p(Y_{it+1} | X_{it+1}, \theta_i, \Sigma_i)$  is a bivariate normal distribution, while the posterior  $p(\theta_i, \Sigma_i | Y^t, X^t)$  is a normal-inverse Wishart distribution. In each period, I computed  $\log w_{it+1}$  by evaluating these distribution at the mean values of  $\theta_i$  and  $\Sigma_i$ .

Finally, the value of the normalized model weights  $\alpha_{it}$  can be obtained from (19) or using the log posteriors odd ratios with the formula reported in Cogley and Sargent (2005b).

# Appendix 3

## State Space representation

### Solow model

The full equations for the Solow model are:

$$\begin{aligned}
 y_t &= \alpha_0^S + \alpha_{y_1}^S y_{t-1} + \alpha_{y_2}^S y_{t-2} + \alpha_{y_3}^S y_{t-3} + \alpha_{\pi_1}^S \left[ \pi_{t-1} - \frac{1}{4} \sum_{j=1}^4 \pi_{t-1-j} \right] + \\
 &\quad \alpha_{\pi_2}^S \left[ \pi_{t-2} - \frac{1}{4} \sum_{j=1}^4 \pi_{t-2-j} \right] + \alpha_{s_1}^S s_{t-1} + \omega_{y,t}^S
 \end{aligned} \tag{38}$$

$$\begin{aligned}
 \pi_t &= \beta_0^S + \beta_{y_1}^S y_{t-1} + \beta_{\pi_1}^S \pi_{t-1} + \beta_{\pi_2}^S \pi_{t-2} + \beta_{\pi_3}^S \pi_{t-3} + \beta_{i_1}^S i_{t-1} + \beta_{i_2}^S i_{t-2} \\
 &\quad + \beta_{s_1}^S s_{t-1} + \omega_{\pi,t}^S
 \end{aligned} \tag{39}$$

The process for the real price of oil is described by (7). The state space representation takes the form:

$$A_{0,S} S_{S,t} = A_{1,S} S_{S,t-1} + B_{1,S} i_t + \omega_{S,t} \tag{40}$$

where the state vector is composed of 14 variables and it is defined as:

$$S_{S,t} = [y_t \ y_{t-1} \ y_{t-2} \ \pi_t \ \pi_{t-1} \ \pi_{t-2} \ \pi_{t-3} \ \pi_{t-4} \ \pi_{t-5} \ i_t \ i_{t-1} \ s_t \ \delta_t \ 1]' \tag{41}$$

the vector attached to the policy instrument is  $B_{1,S} = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0]'$  and the vector of residuals is:

$$\omega_{S,t} = [\omega_{y,t}^S \ 0 \ 0 \ \omega_{\pi,t}^S \ 0 \ 0 \ 0 \ 0 \ \xi_t \ 0 \ o_t \ \varepsilon_t \ 0]'$$

The matrix  $A_{0,S}$  is a  $14 \times 14$  matrix, with  $A_{0,S}(12, 13) = -1$ , ones in the main diagonal, and zeros elsewhere.

The matrix  $A_{1,S}$  is defined as:

$$\begin{bmatrix}
 \alpha_{y_1}^S & \alpha_{y_2}^S & \alpha_{y_3}^S & \alpha_{\pi_1}^S & \alpha_{\pi_2}^S & -\frac{\alpha_{\pi_1}^S}{4} & -\frac{\alpha_{\pi_2}^S + \alpha_{\pi_1}^S}{4} & -\frac{\alpha_{\pi_2}^S + \alpha_{\pi_1}^S}{4} & -\frac{\alpha_{\pi_2}^S + \alpha_{\pi_1}^S}{4} & -\frac{\alpha_{\pi_2}^S}{4} & 0 & 0 & \alpha_{s_1}^S & 0 & \alpha_0^S \\
 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 \beta_{y_1}^S & 0 & 0 & \beta_{\pi_1}^S & \beta_{\pi_2}^S & \beta_{\pi_3}^S & 0 & 0 & 0 & 0 & \beta_{i_1}^S & \beta_{i_2}^S & \beta_{s_1}^S & 0 & \beta_0^S \\
 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho & 0 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
 \end{bmatrix}$$

Using this information we can rewrite the state space representation in (40) as:

$$S_{S,t} = A_S S_{S,t-1} + B_S i_t + C_S \omega_{S,t} \quad (42)$$

where  $A_S = (A_{0,S})^{-1} A_{1,S}$ ,  $B_S = (A_{0,S})^{-1} B_{1,S}$  and  $C_S = (A_{0,S})^{-1}$ .

Finally, the selection matrix  $M_S$  in the policymaker's loss function is defined as:

$$M_S = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\pi^* \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix}$$

### Blanchard-Gali model

The full equations for the Blanchard-Gali model are:

$$y_t = \alpha_0^{BG} + \alpha_{y_1}^{BG} y_{t-1} + \alpha_{y_2}^{BG} y_{t-2} + \alpha_{y_3}^{BG} y_{t-3} + \alpha_i^{BG} \left( i_{t-1} - \frac{1}{4} \sum_{j=1}^4 \pi_{t-j} \right) + \alpha_{s_1}^{BG} s_{t-1} + \omega_{y,t}^{BG} \quad (43)$$

$$\Delta \pi_t = \beta_0^{BG} + \beta_{y_1}^{BG} y_{t-1} + \beta_{y_2}^{BG} y_{t-2} + \beta_{\pi_1}^{BG} \Delta \pi_{t-1} + \beta_{\pi_2}^{BG} \Delta \pi_{t-2} + \beta_{s_1}^{BG} s_{t-1} + \omega_{\pi,t}^{BG} \quad (44)$$

The process for the real price of oil is described by (7). The state space representation takes the form:

$$A_{0,BG} S_{BG,t} = A_{1,BG} S_{BG,t-1} + B_{1,BG} i_t + \omega_{BG,t} \quad (45)$$

In this case, the state vector is composed of 12 variables:  $S_{BG,t} = [y_t \ y_{t-1} \ y_{t-2} \ \pi_t \ \pi_{t-1} \ \pi_{t-2} \ \pi_{t-3} \ i_t \ i_{t-1} \ s_t \ \delta_t \ 1]'$ . The vector of coefficients attached to the policy instrument  $i_t$  is  $B_{1,BG} = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0]'$ . The vector of residuals is:  $\omega_{BG,t} = [\omega_{y,t}^{BG} \ 0 \ 0 \ \omega_{\pi,t}^{BG} \ 0 \ 0 \ 0 \ \xi_t \ 0 \ \omega_t \ \varepsilon_t \ 0]'$ .

The matrix  $A_{0,BG}$  is a  $12 \times 12$  matrix, with  $A_{0,BG}(10,11) = -1$ , ones in the main diagonal, and zeros elsewhere. The matrix  $A_{1,BG}$  is defined as:

$$\begin{bmatrix} \alpha_{y_1}^{BG} & \alpha_{y_2}^{BG} & \alpha_{y_3}^{BG} & -\frac{\alpha_i^{BG}}{4} & -\frac{\alpha_i^{BG}}{4} & -\frac{\alpha_i^{BG}}{4} & -\frac{\alpha_i^{BG}}{4} & \alpha_i^{BG} & 0 & \alpha_{s_1}^{BG} & 0 & \alpha_0^{BG} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \beta_{y_1}^{BG} & \beta_{y_2}^{BG} & 0 & \left(1 + \beta_{\pi_1}^{BG}\right) & \left(\beta_{\pi_2}^{BG} - \beta_{\pi_1}^{BG}\right) & -\beta_{\pi_2}^{BG} & 0 & 0 & 0 & \beta_{s_1}^{BG} & 0 & \beta_0^{BG} \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Again, using this information we can rewrite the state space representation in (45) as:

$$S_{BG,t} = A_{BG}S_{BG,t-1} + B_{BG}i_t + C_{BG}\omega_{BG,t} \quad (46)$$

where  $A_{BG} = (A_{0,BG})^{-1} A_{1,BG}$ ,  $B_{BG} = (A_{0,BG})^{-1} B_{1,BG}$  and  $C_{BG} = (A_{0,BG})^{-1}$ .

Finally, the selection matrix  $M_{BG}$  in the policymaker's loss function is similar to  $M_S$ , with changes due to the different state vector in the two model:

$$M_{BG} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\pi^* \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix}$$

### Hamilton model

The full equations for the Hamilton model are:

$$y_t = \alpha_0^H + \alpha_{y_1}^H y_{t-1} + \alpha_{y_2}^H y_{t-2} + \alpha_{y_3}^H y_{t-3} + \alpha_{s_1}^H s_{t-1} + \omega_{y,t}^H \quad (47)$$

$$\begin{aligned} \pi_t &= \beta_0^H + \beta_{y_1}^H y_{t-1} + \beta_{\pi_1}^H \pi_{t-1} + \beta_{\pi_2}^H \pi_{t-2} + \beta_{\pi_3}^H \pi_{t-3} \\ &\quad + \beta_{i_1}^H i_{t-1} + \beta_{i_2}^H i_{t-2} + \beta_{s_1}^H s_{t-1} + \omega_{\pi,t}^H \end{aligned} \quad (48)$$

with the process for the real price of oil described by (7). The state space representation takes the form:

$$A_{0,H}S_{H,t} = A_{1,H}S_{H,t-1} + B_{1,H}i_t + \omega_{H,t} \quad (49)$$

In this case, the state vector is composed of 11 variables:  $S_{H,t} = [y_t \ y_{t-1} \ y_{t-2} \ \pi_t \ \pi_{t-1} \ \pi_{t-2} \ i_t \ i_{t-1} \ s_t \ \delta_t \ 1]'$ . The vector of coefficients attached to the policy instrument  $i_t$  is  $B_H = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1 \ 0 \ 0 \ 0 \ 0]'$ . The vector of residuals is:  $\omega_{H,t} = [\omega_{y,t}^H \ 0 \ 0 \ \omega_{\pi,t}^H \ 0 \ 0 \ \xi_t \ 0 \ \omega_t \ \varepsilon_t \ 0]'$ .

Similarly to the other models, the matrix  $A_{0,H}$  is a  $11 \times 11$  matrix, with  $A_{0,H}(9,10) = -1$ , ones in the main diagonal, and zeros elsewhere. The matrix of coefficients  $A_{1,H}$  is:

$$\begin{bmatrix} \alpha_{y_1}^H & \alpha_{y_2}^H & \alpha_{y_3}^H & 0 & 0 & 0 & 0 & 0 & \alpha_{s_1}^H & 0 & \alpha_0^H \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \beta_{y_1}^H & 0 & 0 & \beta_{\pi_1}^H & \beta_{\pi_2}^H & \beta_{\pi_3}^H & \beta_{i_1}^H & \beta_{i_2}^H & \beta_{s_1}^H & 0 & \beta_0^H \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \rho & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Again, we can rewrite the state space representation in (49) as:

$$S_{H,t} = A_H S_{H,t-1} + B_H i_t + C_H \omega_{H,t} \quad (50)$$

where  $A_H = (A_{0,H})^{-1} A_{1,H}$ ,  $B_H = (A_{0,H})^{-1} B_{1,H}$  and  $C_H = (A_{0,H})^{-1}$ .

Finally, the selection matrix  $M_H$  in the policymaker's loss function is similar to the other two models, with differences due to the differences in the state vectors:

$$M_H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -\pi^* \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 & 0 \end{bmatrix}$$

## References

- [1] Baumeister, C., Peersman, G., 2008. Time-Varying Effects of Oil Supply Shocks on the US Economy. Mimeo, Faculty of Economics and Business Administration, Ghent University.
- [2] Beck, G., Wieland, V., 2002. Learning and control in a changing environment. *Journal of Economic Dynamics and Control* 26, 1359–1377.
- [3] Bernanke, B.S., 1983. Irreversibility, uncertainty, and cyclical investment. *Quarterly Journal of Economics* 98, 85–106.
- [4] Bernanke, B. S., Gertler, M., Watson, M., 1997. Systematic monetary policy and the effects of oil price shocks. *Brookings Papers on Economic Activity*, no. 1, pp. 91-142.
- [5] Bernanke, B. S., Gertler, M., Watson, M., 2004. Oil shocks and aggregate macroeconomic behavior: the role of monetary policy: reply. *Journal of Money, Credit, and Banking*, vol. 36, no. 2, April, pp. 287-91.
- [6] Blanchard, O. J., Gali, J., 2007. The Macroeconomic Effects of Oil Shocks: Why are the 2000s So Different from the 1970s? National Bureau of Economic Research, Working Papers: 13368.
- [7] Brainard, W., 1967. Uncertainty and the effectiveness of policy. *American Economic Review*, vol. 57, n. 2, 411-425.
- [8] Brock, W. A., Durlauf, S. N, West, K. D., 2003. Policy Evaluation in Uncertain Economic Environments. *Brookings Papers on Economic Activity*, pp.235-301.
- [9] Brock, W. A., Durlauf, S. N, West, K. D., 2007. Model Uncertainty and Policy Evaluation: Some Theory and Empirics. *Journal of Econometrics*, vol. 136, no. 2, February, pp. 629-64.
- [10] Brock, W. A., Durlauf, S. N., Nason, J. M., Rondina, G., 2007. Simple versus Optimal Rules as Guides to Policy. *Journal of Monetary Economics*, vol. 54, no. 5, July, pp. 1372-96.
- [11] Cogley, T., Colacito, R. , Sargent, T.J. , 2007. Benefits from U.S. monetary policy experimentation in the days of Samuelson and Solow and Lucas. *Journal of Money, Credit and Banking*, supplement to Vol. 39, pp. 67-99.
- [12] Cogley, T., De Paoli, B., Matthes, C., Nikolov, K., Yates, T., 2009. A Bayesian Approach to Optimal Monetary Policy with Parameter and Model Uncertainty. Mimeo, department of economics, New York University.
- [13] Cogley, T., Sargent, T.J., 2001. Evolving post world war II US inflation dynamics. In: *NBER Macroeconomics Annual*, vol. 16, pp. 331–373.
- [14] Cogley, T., Sargent, T.J. 2005a. Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. *Review of Economic Dynamics* 8, 262-302.
- [15] Cogley, T., Sargent, T.J., 2005b. The Conquest of U.S. inflation: Learning and Robustness to Model Uncertainty. *Review of Economic Dynamics*, 8, 528-563.
- [16] Finn, M.G., 2000. Perfect competition and the effects of energy price increases on economic activity. *Journal of Money, Credit, and Banking* 32, 400–416.



- [17] Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2003. Bayesian Data Analysis, 2nd edition. CRC Press.
- [18] Giannoni, M. P., 2007. Robust Optimal Monetary Policy in a Forward-Looking Model with Parameter and Shock Uncertainty. *Journal of Applied Econometrics*, vol. 22, no. 1, January-February, pp. 179-213.
- [19] Hamilton, J.D., 1983. Oil and the Macroeconomy since World War II. *Journal of Political Economy*, vol. 91, pp. 228-248.
- [20] Hamilton, J.D., 1988. A neoclassical model of unemployment and the business cycle. *Journal of Political Economy* 96, 593–617.
- [21] Hamilton, J.D., 1996. This is What Happened to the Oil Price-Macroeconomy Relationship. *Journal of Monetary Economics*, vol. 38, pp. 215-220.
- [22] Hamilton, J.D., 2005. Oil and the Macroeconomy. Mimeo, Department of Economics, University of California, San Diego, prepared for Palgrave Dictionary of Economics.
- [23] Hamilton, J. D., Herrera A. M., 2004. Oil shocks and aggregate macroeconomic behavior: the role of monetary policy: comment. *Journal of Money, Credit, and Banking*, vol. 36, no. 2, April, pp. 265-86.
- [24] Hansen, L., Sargent, T., 2001a. Robust control and model uncertainty. *American Economic Review*, vol. 91, no. 2, May, pp. 60-66.
- [25] Hansen, L., Sargent, T., 2001b. Acknowledging Misspecification in Macroeconomic Theory. *Review of Economic Dynamics*, vol. 4, no. 3, July, pp. 519-35.
- [26] Hansen, L., Sargent, T., 2008. Robustness. Princeton University Press, Princeton, New Jersey.
- [27] Herrera, A., Pesavento, E., 2009. Oil Price Shocks, Systematic Monetary Policy, And The "Great Moderation". *Macroeconomic Dynamics*, vol. 13, n. 1, February, pp. 107-137.
- [28] Kilian, L., 2008. The Economic Effects of Energy Price Shocks. *Journal of Economic Literature*, vol. 46, n.4, December, pp. 871-909.
- [29] Kilian, L., Lewis, L., 2009. Does the Fed Respond to Oil Price Shocks?. C.E.P.R. Discussion Papers: 7594.
- [30] Kim, I., Loungani, P., 1992. The role of energy in real business cycle models. *Journal of Monetary Economics* 29, 173–189.
- [31] Levin, A., Williams, J., 2003. Robust monetary policy with competing reference models. *Journal of Monetary Economics* 50, 945–975.
- [32] Lippi, F., Nobili, A., 2008. Oil and the Macroeconomy: A Structural VAR Analysis with Sign Restrictions. CEPR Discussion Papers n.6830.
- [33] Mork, K. A., 1994. Business cycles and the oil market (special issue). *The Energy Journal*, vol. 15, pp. 15–37.
- [34] Pindyck, R.S., 1980. Energy price increases and macroeconomic policy. *The Energy Journal* 1, 1–20.
- [35] Pindyck, R.S., 1999. The Long-Run Evolutions of Energy Prices. *The Energy Journal*, vol. 20, n.2, pp. 1-28.

- [36] Primiceri, G., 2005. Time Varying Structural Vector Autoregression and Monetary Policy. *Review of Economic Studies* 72, 821-852.
- [37] Primiceri, G. E., 2006. Why inflation rose and fell: Policymakers' beliefs and US postwar stabilization policy. *The Quarterly Journal of Economics*, vol. 121, August, pp. 867-901.
- [38] Rasche, R. H., Tatom, J. A., 1977. Energy resources and potential GNP. *Federal Reserve Bank of St. Louis Review* 59 (June), 10-24.
- [39] Roberts, G.O., Smith, A.F.M., 1992. Simple conditions for the convergence of the Gibbs sampler and Metropolis-Hastings algorithms. *Stochastic Processes and Their Applications* 49, 207-216.
- [40] Rondina, F., 2010. Policy evaluation and uncertainty about the effects of oil prices on economic activity. Mimeo, Department of Economics, University of Wisconsin, Madison.
- [41] Rotemberg, J.J., Woodford, M., 1996. Imperfect competition and the effects of energy price increases. *Journal of Money, Credit, and Banking* 28, 549-577.
- [42] Rudebusch, G. D., 2001. Is the Fed too timid? Monetary Policy in an uncertain world. *Review of Economics and Statistics*, vol. 83, n. 2, 203-217.
- [43] Rudebusch, G., Svensson, L., 1999. Policy rules for inflation targeting. In: Taylor, J. (Ed.), *Monetary Policy Rules*. University of Chicago Press, Chicago, pp. 203-246.
- [44] Sargent, T. J., Williams, N., 2005. Impacts of Priors on Convergence and Escapes from Nash Inflation. *Review of Economic Dynamics*, vol. 8, n. 2, pp. 360-91.
- [45] Segal, P., 2007. Why do oil price shocks no longer shock? Oxford Institute for Energy Studies Working Paper.
- [46] Söderström, U., 2000. Should central banks be more aggressive? Mimeo, Sveriges Riksbank and Stockholm School of Economics.
- [47] Söderström, U., 2002. Monetary policy with uncertain parameters. *Scandinavian Journal of Economics*, vol. 104, n. 1, 125-145.
- [48] Solow, R. M., 1980. What to do (macroeconomically) when OPEC comes. In: Fischer, S. (Ed.), *Rational Expectations and Economic Policy*. University of Chicago Press, Chicago.
- [49] Svensson, L. E. O., 1999. Inflation targeting: some extensions. *Scandinavian Journal of Economics*, vol. 101, n. 3, 337-361.
- [50] Tetlow, R.J., von zur Muehlen, P., 2001. Robust monetary policy with misspecified models: does model uncertainty always call for attenuated policy? *Journal of Economic Dynamics and Control*, 25, 911-949.
- [51] Wieland, V., 2000a. Monetary policy, parameter uncertainty, and optimal learning. *Journal of Monetary Economics* 46, 199-228.
- [52] Wieland, V., 2000b. Learning by doing and the value of optimal experimentation. *Journal of Economic Dynamics and Control* 24, 501-534.