Regime Switches, Agents' Beliefs, and Post-World War II U.S. Macroeconomic Dynamics*

Francesco Bianchi
Duke University

September 30, 2009

Abstract

The evolution of inflation and output over the last 50 years is examined through the lens of a micro-founded model that allows for changes in the behavior of the Federal Reserve and in the volatility of structural shocks. Agents are aware of the possibility of regime changes and their beliefs have an impact on the law of motion underlying the macroeconomy. The results support the view that there were regime switches in the conduct of monetary policy. However, behavior of the Federal Reserve is identified by repeated fluctuations between a *Hawk*- and a *Dove*- regime, instead of by the traditional pre- and post- Volcker structure. Counterfactual simulations show that if agents had anticipated the appointment of an extremely conservative Chairman, inflation would not have reached the peaks of the late '70s and the inflation-output trade-off would have been less severe. These "beliefs counterfactuals" are new in the literature. Finally, the paper provides a Bayesian algorithm to handle the technical difficulties that arise in rational expectations model with Markov-switching regimes.

*I am grateful to Chris Sims, Mark Watson, Efrem Castelnuovo, Nobuhiro Kiyotaki, David Lucca, Esteban Rossi-Hansberg, Barbara Rossi, Tom Sargent, and all seminar participants at Princeton University, the Board of Governors, the Federal Reserve Bank of Chicago, the New York Area Workshop on Monetary Policy, University of Pennsylvania, University of Virginia, Northwestern University, Duke University, UCLA, New York University, Philadelphia Fed, Penn State, CREI-Pompeu Fabra, University of Bern, NBER Summer Institute, Society for Economic Dynamics Annual Meeting for useful suggestions and comments. Part of this paper was completed while the author was visiting the Board of Governors, whose hospitality is gratefully acknowledged. Correspondence: Duke University, 213 Social Sciences Building, Box 90097, Durham, NC 27708-0097. E-mail: francesco.bianchi@duke.edu.

1 Introduction

The importance of agents' expectations in determining equilibrium outcomes became apparent to macroeconomists following the rational expectations hypothesis revolution that occurred in the '70s. Practitioners learned that empirical work revolving around policy changes cannot abstract from modeling agents' expectations about these very same changes. With this lesson in mind, this paper aims to explain the evolution of inflation and output dynamics over the last 50 years taking into account not only the possibility of regime switches in the behavior of the Federal Reserve, but also agents' beliefs around these changes. To this end, a general equilibrium model in which the behavior of the Federal Reserve is allowed to change over time is fit to the data. In such a model, regime changes are regarded as stochastic and reversible, agents are aware of this, and their beliefs matter for the law of motion governing the evolution of the economy.

Two main results emerge from this analysis. First, the estimates support the view that there were regime switches in the conduct of monetary policy. However, the idea that US monetary policy can be described in terms of pre- and post- Volcker proves to be misleading. The behavior of the Federal Reserve has instead repeatedly fluctuated between a *Hawk*- and a *Dove*- regime. Following an adverse technology shock, the Fed is willing to cause a deep recession to fight inflation only under the *Hawk* regime. Under the *Dove* regime, the Fed tries to minimize output fluctuations. Second, counterfactual simulations show that if in the '70s agents had anticipated the appointment of an extremely conservative Chairman, inflation would not have reached the peaks of those years and the inflation-output trade-off would have been less severe. These "beliefs counterfactuals" are new in the literature and take full advantage of the potential of Markov-switching rational expectations models.

In order to contextualize the results, I shall start with a brief description of the events that this paper intends to interrelate. Figure 1 shows the series for output gap, annualized quarterly inflation, and the Federal Funds rate (FFR) for the period 1954:IV-2008:I. The shaded areas represent the NBER recessions and the vertical lines mark the appointment dates of the Federal Reserve chairmen. Some stylized facts stand out. Over the early years of the sample inflation was relatively low and stable. Then, inflation started rising during the late '60s and spun out of control in the late '70s. At the same time the economy experienced a deep and long recession following the oil crisis of 1974. During the first half of the '80s the economy went through a painful disinflation. Inflation went back to the levels that were prevailing before the '70s at the cost of two severe recessions. From the mid-80s, until the recent financial crisis the economy has been characterized by remarkable economic stability. Economists like to refer to this last period with the term "Great Moderation", while the

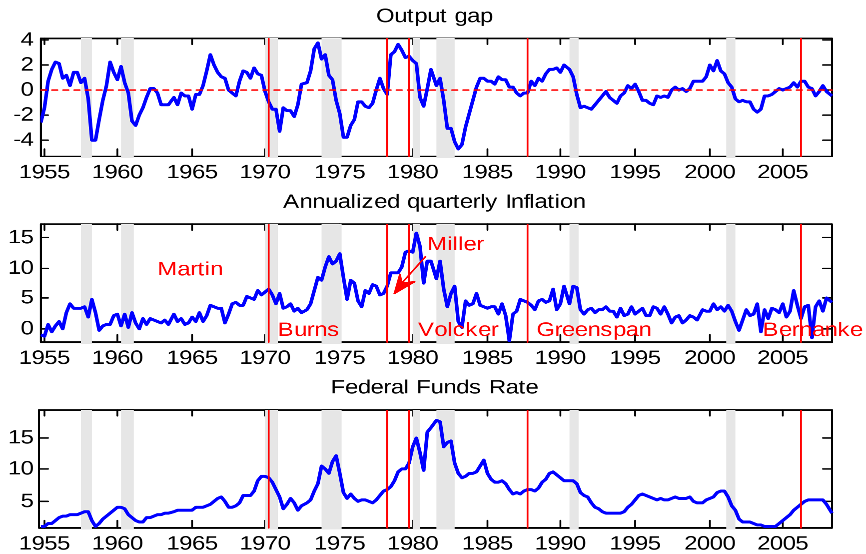

Figure 1: Output gap, inflation, and policy interest rate for the US. Output gap is obtained HP filtering the series of real per capita GDP. The shaded areas represent NBER recessions, while the vertical lines mark the appointment dates of the Chairmen.

name "Great Inflation" is often used to label the turmoil of the '70s. The sharp contrast between the two periods is evident. Understanding the causes of these remarkable changes in the reduced form properties of the macroeconomy is crucial, particularly now that policy makers are facing a severe crisis, along with concerns around the path of future inflation. If these changes are the result of exogenous shocks, events similar to those of the Great Inflation could occur again. If, on the other hand, policy makers currently posses a better understanding of the economy, then we could be somewhat optimistic about the long run consequences of the current economic crisis.

Economists who tend to establish a clear link between the behavior of the Fed and the performance of the economy would argue that the changes described above are the result of a substantial switch in the anti-inflationary stance of the Federal Reserve. Therefore, the Fed would be blamed for the high and volatile inflation of the '70s and praised for the stability that has characterized the **subsequent period** ("Good Policy"). The two most prominent examples of this school of thought are Clarida et al. (2000) and Lubik and Schorfheide (2004). These authors find that embedding the policy rule followed in the '70s in a general equilibrium model implies non-uniqueness of the equilibrium. On the other hand, several other authors find little evidence in favor of changes in the parameters describing policy rules and ample evidence supporting the idea that, at least to some extent, the Great Moderation would be due to "Good Luck", i.e. to a reduction in the magnitude of the shocks hitting the economy. As strong advocates of this alternative explanation, Sims and Zha (2006) use a Markov-switching VAR and identify changes in the volatilities of the structural disturbances as the key driver behind the stabilization of the U.S. economy.[1]

This debate often revolves around the sharp decline in inflation that started shortly after Paul Volcker was appointed chairman of the Federal Reserve in August 1979. It is tempting to draw a line between the two events and conclude that a substantial change in the conduct of monetary policy must have occurred in those years. Even if several economists would agree that this was in fact the case, there is much less consensus around the notion that this event represented an unprecedented and once-and-for-all regime change.

Thus, the first contribution of this paper is to shed new light on this controversy. I consider a Dynamic Stochastic General Equilibrium model in which the Taylor rule parameters characterizing the behavior of the Federal Reserve and the volatilities of the structural shocks are allowed to change over time. These changes are modeled as two independent Markov-switching processes in order to accommodate the two competing explanations. In the model agents are aware of the possibility of regime changes and they take this into account when forming expectations. Therefore, the law of motion of the variables of interest depends not only on the traditional microfounded parameters, but also on the beliefs around alternative regimes.

Two main results emerge from the estimates. First, the model supports the idea that US monetary policy was indeed subject to regime changes. The best performing model is one in which the Taylor rule is allowed to move between a *Hawk*- and a *Dove*- regime. The former implies a strong response to inflation and little concern for the output gap, whereas the latter comes with a weak response to inflation. In particular, while the *Hawk* regime, if taken in *isolation*, would satisfy the Taylor principle, the *Dove* regime would not.[2] Following an adverse technology shock, the Fed is willing to cause a deep recession to fight inflation only under the *Hawk* regime. Under the *Dove* regime, the Fed tries to minimize output fluctuations.

Second, the idea that US economic history can be divided into pre- and post- Volcker turns out to be misleading. Surely the results corroborate the widespread belief that the appointment of Volcker marked a change in the stance of the Fed toward inflation. In fact, around 1980, right after his appointment, the Fed moved from the *Dove* to the *Hawk* regime. However, the behavior of the Federal Reserve has repeatedly fluctuated between the

[1] Please refer to section 2 for a more in-depth overview of the Good Luck-Good Policy debate.

[2] The Taylor principle asserts that central banks can stabilize the macroeconomy by moving their interest rate instrument more than one-for-one in response to a change in inflation.

two alternative Taylor rules and regime changes have been relatively frequent. Specifically, the *Dove* regime was certainly in place during the second half of the '70s, but also during the first half of the '60s, again around the '91 recession, and with high probability toward the end of the sample.

The second contribution of the paper relates to the role of agents' beliefs in explaining the Great Inflation. I consider new counterfactual simulations that revolve around agents' beliefs about the evolution of monetary policy. How would things have been different if in the '70s agents had anticipated the appointment of an extremely conservative chairman such as Volcker? Were agents making decisions assuming that the Burns/Miller regime would have lasted for some time?

Counterfactual simulations suggest that this last hypothesis is more likely to explain what was occurring in the '70s. It seems that in those years the Fed was facing a severe credibility problem and beliefs about alternative monetary policy regimes were indeed playing a crucial role. To address this hypothesis, I introduce a third regime, the *Eagle* regime, that is even more hawkish than the *Hawk* regime. This regime is meant to describe the behavior of an extremely conservative chairman. It turns out that if agents had assigned a relatively large probability to this hypothetical regime, inflation would not have reached the peaks of the mid- and late- '70s, independent of whether or not the **Eagle regime occurred**. Furthermore, the cost in terms of lower output would not have been extremely large. Quite interestingly, simply imposing the *Hawk* regime throughout the entire sample would have implied modest gains in terms of inflation and a substantial output loss.

These last results point toward two important conclusions. First, beliefs about alternative regimes can go a long way in modifying equilibrium outcomes. Specifically, in the present model, the effective sacrifice ratio faced by the Federal Reserve depends on the alternative scenarios that agents have in mind. If agents had anticipated the appointment of a very conservative chairman, the cost of keeping inflation down would have been lower. Second, monetary policy does not need to be hawkish all the time in order to achieve the desired goal of low and stable inflation. What is truly necessary is a strong commitment to bring the economy back to equilibrium as soon as adverse shocks disappear. It seems that in the '70s the main problem was not simply that the Fed was accommodating a series of adverse technology shocks, but rather that there was a lack of commitment to restoring equilibrium once the economy had gone through the peak of the crisis. **In this context, the** Volcker disinflation might have been important exactly because it made such a commitment credible.

The last contribution of this paper is methodological. I propose a Bayesian algorithm to estimate a Markov Switching DSGE model (MS-DSGE) *via* Gibbs sampling. The algorithm allows for different assumptions regarding the transition matrix used by agents in the model.

Specifically, this matrix may or may not coincide with the one that is observed ex-post by the econometrician. The algorithm does not require approximating the likelihood and return estimates of the underlying DSGE states and structural shocks. **This is particularly convenient w**hen researchers are interested in investigating the behavior of those variables that are not directly observable or in conducting counterfactual simulations.

A MS-DSGE model represents a promising tool to better understand the Great Moderation, as well as the rise and fall of inflation, because it combines the advantages of the previous approaches while mitigating the drawbacks. Consider the Good Luck-Good Policy literature. It is quite striking that researchers tend to find *opposite* results moving from different starting points. The two most representative papers of the "Good Policy" view (Lubik and Schorfheide (2004) and Clarida et al. (2000)) are based on a subsample analysis: pre- and post-Volcker. Instead, authors supporting the "Good Luck" hypothesis draw their conclusions according to models in which parameter switches are modeled as stochastic and reversible. In other words, they do not impose a one-time-only regime change but they let the data decide *if* there was a break and *if* this break can be regarded as a permanent change. Furthermore, they also allow for the possibility that changes in the reduced form properties of the variables are the result of breaks in the volatility of the underlying shocks.

At the same time, both approaches have some limitations when taking into account the role of expectations. The Good Policy literature, based on subsample analysis, falls short in recognizing that if a regime change occurred once, it might occur again, and that agents should take this into account when forming expectations. At the same time, reduced form models do not allow for the presence of forward-looking variables that play a key role in dynamic stochastic general equilibrium models. This has important implications when interpreting those counterfactual exercises which show that little would have changed if more aggressive regimes had been in place during the '70s.

In a MS-DSGE, regime changes are not regarded as once-and-for-all and expectations are formed accordingly. Thus, the law of motion of the variables included in the model can change in response to changes in beliefs. These could deal with the nature of the alternative regimes or simply with the probabilities assigned to them. Consequently, counterfactual simulations are more meaningful and more robust to the Lucas critique (Lucas (1976)), because the model is re-solved not only incorporating eventual changes in the parameters of the model, but also taking into account the assumptions about what agents know or believe. This is particularly relevant, for example, when imposing that a single regime was in place throughout the sample: Under this assumption, the model is solved assuming that agents regard the regime in place as the only possible one and they form expectations accordingly.

Furthermore, given that the model is microfounded, all parameters have a clear economic interpretation. This implies that a given hypothesis around the source(s) of the Great Moderation can explicitly be tested against the others and also addresses the concerns recently raised by Benati and Surico (2008) about using structural VAR's to draw conclusions about the Good Luck-Good Policy debate.

Finally, the model considered in this paper accommodates both explanations of the Great Moderation given that it allows for a Markov-switching Taylor rule and heteroskedastic volatilities. As emphasized by Sims and Zha (2006) and Cogley and Sargent (2006), it is essential to account for the stochastic volatility of exogenous shocks when trying to identify shifts in monetary policy. In fact, it turns out that a change in the volatilities of the structural shocks contributes to the broad picture. A high volatility regime has been in place for a large part of the period that goes from the early '70s to the mid-80s. Interestingly, 1984 is often regarded as the year in which the Fed was finally able to gain control of inflation.

The content of this paper can be summarized as follows. Section 2 gives a brief summary of the related literature. Section 3 contains a description of the model. Section 4 describes the estimation algorithms. Section 5 presents parameter estimates, impulse responses, counterfactual exercises, and variance decomposition for the benchmark model in which the behavior of the Fed can switch between two Taylor rules. Section 6 considers alternative specifications that offer competing explanations for the source of the changes in macroeconomic dynamics. Section 7 confronts the different models with the data computing the marginal data densities. Section 8 concludes.

2 Related literature

This paper is related to the growing literature that allows for parameter instability in microfounded models. Justiniano and Primiceri (2008) consider a DSGE model allowing for time variation in the volatility of the structural innovations. Laforte (2005) models heteroskedasticity in a DSGE model according to a Markov-switching process. Liu et al. (2008) test empirical evidence of regime changes in the Federal Reserve's inflation target. They also allow for heteroskedastic shock disturbances. Along the same lines, Schorfheide (2005) estimates a dynamic stochastic general equilibrium model in which monetary policy follows a nominal interest rate rule that is subject to regime switches in the target inflation rate. Interestingly, he also considers the case in which agents use Bayesian updating to infer the policy regime. Ireland (2007) also estimates a New Keynesian model in which Federal Reserve's unobserved inflation target drifts over time. Davig and Doh (2008) consider a New-Keynesian model in which structural parameters can change across regimes to asses the sources that lead to a decline in inflation persistence. Inoue and Rossi (2009) suggest that the Great

Moderation could be erroneously attributed to "Good Luck" if alternative sources of instabilities cancel each other out. Davig and Leeper (2006b) estimate Markov-switching Taylor and Fiscal rules, plugging them into a calibrated DSGE model. Bikbov (2008) estimates a structural VAR with restrictions imposed according to an underlying New-Keynesian model with Markov-switching parameters. Regime changes are identified extracting information from the yield curve.

Fernández-Villaverde and Rubio-Ramírez (2007) consider a model with time-varying structural parameters. The model is solved using perturbation methods and estimated with particle filtering. They find substantial evidence of parameter instability. King (2007) proposes a method to estimate dynamic-equilibrium models subject to permanent shocks to the structural parameters. Time-varying structural parameters are treated as state variables that are both exogenous and unobservable, and the model is estimated with particle filtering. In a univariate framework, Castelnuovo et al. (2008) combine a regime-switching Taylor rule with a time-varying policy target, whereas Favero and Monacelli (2005) estimate fiscal policy feedback rules using Markov-switching regressions.

Finally, the paper is obviously related to the extensive literature that explores the evolution of output and inflation over the past fifty years. In their seminal contributions, Clarida et al. (2000) and Lubik and Schorfheide (2004) point out that the in the '70s the economy was subject to the possibility of self-fulfilling inflationary shocks because of the monetary policy rule that was followed at that time. Their estimated policy rule for the later period, on the other hand, implied no such indeterminacy. Boivin and Giannoni (2006) find that monetary policy has stabilized the economy more effectively in the post-1980 period. On the other hand, Bernanke and Mihov (1998), Leeper and Zha (2003), Stock and Watson (2003), Canova and Gambetti (2004), Kim and Nelson (2004), Cogley and Sargent (2006), and Primiceri (2005) provide little evidence in favor of the view that the monetary policy rule has changed drastically. Cogley and Sargent (2005), Primiceri (2006), and Sargent et al. (2006) propose interesting explanations for the evolution of Fed's beliefs about the structure of the economy.

3 The Model

The model used in this paper is based on Lubik and Schorfheide (2004). Nevertheless, the model is different to the extent that it allows for regime changes in the Taylor rule parameters and heteroskedasticity. This specification is chosen as the benchmark case because it nests the two alternative explanations of the Great Inflation and the Great Moderation. A change in the behavior of the Fed is often regarded as the keystone to explain the break in the volatility of macroeconomic variables, therefore the model allows for two distinct Taylor rules. At the same time, the Good Luck argument is captured by the Markov-switching volatilities. Appendix C presents the full model. **However, solving and estimating** the full non-linear model with regime switches is not a trivial task. In order **to make the estimation of the model a realistic goal,** the model is log-linearized around the steady state.

3.1 A New-Keynesian model

The private sector can be described by a system of two equations:

$$\widetilde{\pi}_t = \beta E_t(\widetilde{\pi}_{t+1}) + \kappa(\widetilde{y}_t - z_t) \tag{1}$$

$$\widetilde{y}_t = E_t(\widetilde{y}_{t+1}) - \tau^{-1}(\widetilde{R}_t - E_t(\widetilde{\pi}_{t+1})) + g_t \tag{2}$$

The tilde denotes percentage deviations from the steady state or, in the case of output, from a trend path. The process z_t, captures exogenous shifts of the marginal costs of production and can be interpreted as a supply shock. The process g_t summarizes changes in preferences and the time-varying government spending. The two shocks evolve according to:

$$g_t = \rho_g g_{t-1} + \epsilon_{g,t} \tag{3}$$

$$z_t = \rho_z z_{t-1} + \epsilon_{z,t} \tag{4}$$

Inflation dynamics are described by the expectational Phillips curve (1) with slope κ. Intuitively, a boom defined as a positive value for $\widetilde{y}_t$ is only inflationary when it is not supported by a (temporary) technology improvement ($z_t > 0$). This relation can be derived assuming a quadratic adjustment cost or Calvo pricing.

Equation (2) is an intertemporal Euler equation describing the households' optimal choice of consumption and bond holdings. Since the underlying model has no investment, output is proportional to consumption up to the exogenous process g_t. The parameter $\tau^{-1} > 0$ can be interpreted as intertemporal substitution elasticity and $0 < \beta = 1/(1 + r^*) < 1$ is the households' discount factor, where r^* is the steady state real interest rate.

Conditional on a particular regime, the behavior of the monetary authority is described by:

$$\widetilde{R}_t = \rho_R(\xi_t^{sp})\widetilde{R}_{t-1} + (1 - \rho_R(\xi_t^{sp}))(\psi_1(\xi_t^{sp})\widetilde{\pi}_t + \psi_2(\xi_t^{sp})\widetilde{y}_t) + \epsilon_{R,t} \tag{5}$$

The central bank responds to deviations of inflation and output from their respective target levels adjusting the monetary policy interest rate.[3] Unanticipated deviation from the systematic component of the monetary policy rule are captured by $\epsilon_{R,t}$. ξ_t^{sp} is an unobserved state variable capturing the monetary policy regime that is in place at time t. The unobserved state takes on a finite number of values $j = 1, ..., m^{sp}$ and follows a Markov chain that evolves according to the transition matrix H^{sp}.

Agents in the model know the probability of moving across regimes and they use this information when forming expectations. The probability of moving across regimes does depend only on the regime that is in place at time t. It would seem natural to make the transition probabilities endogenous, for example, depending on the level of inflation. However, two orders of problems arise in this context. First, solving the model becomes much more complicated and computationally intensive (Davig and Leeper (2006a)). Second, this would require the estimation of threshold values. However, the number of regime switches is not likely to be high enough to pin down these thresholds.

Note that the Central Bank tries to stabilize $\widetilde{y}_t$, instead of $\widetilde{y}_t - z_t$. Therefore, following a technology shock, a trade-off arises: It is not possible to keep inflation stable and at the same time have output close to the target. Woodford (2003) (chapter 6) shows that it is fluctuations in $\widetilde{y}_t - z_t$ rather than $\widetilde{y}_t$ that are relevant for welfare. However, Woodford also (Woodford (2003), chapter 4) points out that there are reasons to doubt that the measure of output gap used in practice would coincide with $\widetilde{y}_t - z_t$. There are several measures of output gap and a Central Bank is likely to look at all of them when making decisions. More importantly, the assumption that the Fed responds to $\widetilde{y}_t - z_t$ is at odds with some recent contributions in the macro literature: Both Primiceri (2006) and Orphanides (2002) show that during the '70s there were important misjudgments around the path of potential output. Admittedly, the ideal solution would be to assume that the Fed faces a filtering problem, perhaps along the lines of Boivin and Giannoni (2008) and Svensson and Woodford (2003). However, this approach would add a substantial computational burden. Therefore, at this stage, the Taylor rule as specified in (5) is preferred.

Heteroskedasticity is modeled as an independent Markov-switching process:

$$\epsilon_t \sim N\left(0, Q\left(\xi_t^{vo}\right)\right), \; Q\left(\xi_t^{vo}\right) = diag\left(\sigma_R^2, \sigma_g^2, \sigma_z^2\right) \tag{6}$$

where $Q\left(\xi_t^{vo}\right)$ is the covariance matrix of the shocks and ξ_t^{vo} is an unobserved state that describes the evolution of the stochastic volatility regime. This Markov switching process

[3] Currently, the model does not allow for a time-varying inflation target. When allowing for both a MS Taylor rule and a drifting inflation target, the model cannot distinguish a change in the target from a regime change in the Taylor rule parameters. Furthermore, Liu et al. (2008) find that changes in the inflation target play little role in explaining macroeconomic volatility in a MS-DSGE model that allows for heteroskedastic shocks.

evolves according to an independent transition matrix H^{vo}.

If we define three vectors θ^{sp}, θ^{ss} and θ^{vo} containing the structural parameters, the steady state values, and the standard deviations of the shocks:[4]

$$\theta^{sp} = \left[\tau, \kappa, \psi_1, \psi_2, \rho_r, \rho_g, \rho_z\right]', \ \theta^{ss} = \left[r^*, \pi^*\right]', \ \theta^{vo} = \left[\sigma_R, \sigma_g, \sigma_z\right]'$$

and the DSGE state vector S_t as:

$$S_t = \left[\widetilde{y}_t, \widetilde{\pi}_t, \widetilde{R}_t, g_t, z_t, E_t(\widetilde{y}_{t+1}), E_t(\widetilde{\pi}_{t+1})\right]'$$

we can rewrite the system of equations (1)-(5) in a more compact form:

$$\Gamma_0\left(\xi_t^{sp}, \theta^{sp}\right) S_t = \Gamma_1\left(\xi_t^{sp}, \theta^{sp}\right) S_{t-1} + \Psi\left(\xi_t^{sp}, \theta^{sp}\right) \epsilon_t + \Pi \eta_t \tag{7}$$

with η_t a vector containing the expectations errors.

A model in this form in which there are no regime changes could be easily solved using the solution method for linear rational expectations models descibed in Sims (2002). In the current context computations become more complicated because the model is quasi-linear, i.e. it is linear only conditioning on ξ_t^{sp}.

3.2 Solving the MS-DSGE model

The solution method used in this paper is based on the work of Farmer et al. (2006). The idea is to expand the state space of a Markov-switching rational expectations model and to write an equivalent model with fixed parameters in this expanded space. The authors consider the class of Minimal State Variable solutions (McCallum (1983), MSV) to the expanded model and they prove that any MSV solution is also a solution to the original Markov-switching rational expectations model.[5] The class of MSV solutions is large, but it is not exhaustive. The authors argue that MSV solution is likely to be the most interesting class to study given that it is often stable under real time learning (Evans and Honkapohja (2001), McCallum (2003)). Furthermore, the problem of indeterminacy/determinacy in a MS-DSGE model is a very complicated one and, at least for the model considered in this paper, it has not yet been solved.[6]

[4] Here and later on *sp*, *vo*, and *ss* stand respectively for structural parameters, volatilities, and steady state.

[5] Appendix E provides a description of the solution algorithm.

[6] Davig and Leeper (2007) and Farmer et al. (2008) make important steps in this direction. Davig and Leeper (2007) rewrite expectations by distributing probability mass for the two possible regimes to obtain a *fully linear* representation of the original non-linear model. This allows them to derive a *Generalized Taylor Principle*. On the other hand, the model considered here retains the non-linearity arising from having regime

Farmer et al. (2006) show that when a MSV solution exists, it can be characterized as a vector-autoregression with regime switching, of the kind studied by Hamilton (1989) and Sims and Zha (2006):

$$S_t = T\left(\xi_t^{sp}, \theta^{sp}, H^{sp}\right) S_{t-1} + R\left(\xi_t^{sp}, \theta^{sp}, H^{sp}\right) \epsilon_t \tag{8}$$

It is worth emphasizing that the law of motion of the DSGE states depends on the structural parameters (θ^{sp}), the regime in place (ξ_t^{sp}), and the probability of moving across regimes (H^{sp}). This means that what happens under regime i does not only depend on the structural parameters describing that particular regime, but also on what agents expect it is going to happen under alternative regimes and on how likely it is that a regime change will occur in the future. In other words, agents' beliefs matter for the law of motion governing the economy. In principle, agents in the model might form expectations using a transition matrix (H^m) that differs from the *objective* transition matrix that is observed ex-post by the econometrician (H^{sp}). For example, it might take some time for agents to learn about this transition matrix. However, the benchmark model assumes $H^{sp} = H^m$, i.e. agents share the same information set of the econometrician.

From now on, a more compact notation will be used: $T\left(\xi_t^{sp}\right) = T\left(\xi_t^{sp}, \theta^{sp}, H^{sp}\right)$ and $R\left(\xi_t^{sp}\right) = R\left(\xi_t^{sp}, \theta^{sp}, H^{sp}\right)$.

4 Estimation strategy

The law of motion (8) can be combined with a system of observation equations.[7] The result is a model cast in state space form:

$$\begin{aligned}
Y_t &= D(\theta^{ss}) + Z S_t + v_t && (9)\\
S_t &= T\left(\xi_t^{sp}\right) S_{t-1} + R\left(\xi_t^{sp}\right) \epsilon_t && (10)\\
\epsilon_t &\sim N\left(0, Q\left(\xi_t^{vo}\right)\right),\ Q\left(\xi_t^{vo}\right) = diag\left(\theta^{vo}\left(\xi_t^{vo}\right)\right)^2 && (11)\\
v_t &\sim N\left(0, U\right),\ U = diag\left(\sigma_y^2, \sigma_\pi^2, \sigma_F^2\right) && (12)\\
p\left(\xi_t^{sp} = i | \xi_{t-1}^{sp} = j\right) &= H^{sp}(i,j),\ p\left(\xi_t^{vo} = i | \xi_{t-1}^{vo} = j\right) = H^{vo}(i,j) && (13)
\end{aligned}$$

changes. Farmer et al. (2008) take a different route relying on mean-square stability, a concept of equilibrium popular in the engineering literature. Both methods only apply to purely forward looking models, whereas the Taylor rule considered in this paper has an autoregressive component.

[7] The time series are extracted from the Global Insight database. Output gap is measured as the percentage deviations of real per capita GDP from a trend obtained with the HP filter. Inflation is annualized quarterly percentage change of CPI (Urban, all items). Nominal interest rate is the average Federal Funds Rate in percent.

where:

$$Y_t = \begin{bmatrix} GDP_t \\ INFL_t^A \\ FFR_t \end{bmatrix}, \; D(\theta^{ss}) = \begin{bmatrix} 0 \\ 4\pi^* \\ 4(\pi^* + r^*) \end{bmatrix}, \; Z = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and v_t is a vector containing observation errors.[8]

For a DSGE model with fixed parameters the likelihood can be easily evaluated using the Kalman filter and then combined with a prior distribution for the parameters. When dealing with a MS-DSGE model the Kalman filter cannot be applied in its standard form because, given an observation for Y_t, the estimate of the underlying DSGE state vector distribution is not unique. At the same time, the standard Bayesian updating procedure that is generally used to evaluate the likelihood of Markov-switching models, cannot be applied because it relies on the assumption that Markov states are history independent. This does not occur here: Given that we do not observe S_t, the probability assigned to a particular Markov state depends on the value of S_{t-1}, whose distribution depends on the realization of $\xi^{sp,t-1}$.[9]

Generally researchers rely on an approximation of the likelihood in order to estimate the model. In this paper a new method is proposed. Note that if we could observe $\xi^{sp,T}$ and $\xi^{vo,T}$, then it would be straightforward to apply the Kalman filter because given Y_t it would be possible to unequivocally update the distribution of S_t. In the same way, if S^T were observable, then standard methods could be applied to the Markov-switching VAR described by (10), (11) and (13). These considerations suggest that it is possible to sample from the posterior using a Gibbs sampling algorithm. A detailed description of the algorithm is given in appendix B. In what follows I briefly summarize the key step:[10]

At the beginning of iteration n we have: θ^{sp}_{n-1}, θ^{ss}_{n-1}, θ^{vo}_{n-1}, S^T_{n-1}, $\xi^{sp,T}_{n-1}$, $\xi^{vo,T}_{n-1}$, H^m_{n-1}, H^{sp}_{n-1}, and H^{vo}_{n-1}.

1. Given S^T_{n-1}, H^{sp}_{n-1} and H^{vo}_{n-1}, (10), (11) and (13) form a Markov-switching VAR. Use Bayesian updating to get a filtered estimate of the probability assigned to the Markov

[8] Observation errors turn out to be important only for inflation, consistently with the findings of Justiniano and Primiceri (2009), whereas they are virtually zero for the FFR and output.

[9] Here and later on $\xi^{sp,t-1}$ stands for $\{\xi^{sp}_s\}_{s=1}^{t-1}$.

[10] The chain is started making a draw from $N\left(\overline{\theta}, c^2\overline{\Sigma}\right)$, where $\overline{\theta}$ is the posterior mode for the model parameters and $\overline{\Sigma}$ is the inverse of the Hessian computed at the posterior mode. S_0^T is obtained filtering the data according to the most likely path for the MS states. Because the posterior density function is very non-Gaussian and complicated in shape, it is extremely important to find the posterior mode. The standard method to approximate the posterior is based on Kim's approximate evaluation of the likelihood (Kim and Nelson (1999)) and relies on an approximation of the DSGE state vector distribution. I also consider an alternative method: Instead of approximating the DSGE state vector distribution, I keep track of a limited number of alternative paths for the Markov-switching states. Both methods are described in appendix D.

switching states and the then use the backward drawing method to get $\xi_n^{sp,T}$ and $\xi_n^{vo,T}$.

2. Given $\xi_n^{sp,T}$ and $\xi_n^{vo,T}$, draw H_n^{sp} and H_n^{vo} according to a Dirichlet distribution.

3. Conditional on $\xi_n^{sp,T}$ and $\xi_n^{vo,T}$, the likelihood of the state space form model (9)-(12) can be evaluated using the Kalman filter. Draw $\widetilde{H}^m$, ϑ^{sp}, ϑ^{ss}, and ϑ^{vo} from the proposal distributions. The proposal parameters are accepted or rejected according to a Metropolis-Hastings algorithm. This step also returns filtered estimates of the DSGE states: $\widetilde{S}_n^T$.

4. Start drawing the last DSGE state $S_{T,n}$ from the terminal density $p\left(\widetilde{S}_{T,n}|Y^T,...\right)$ and then use a backward recursion to draw $p\left(S_{t,n}|S_{t+1,n},Y^T,...\right)$.

5. If $n < n_{sim}$, go back to 1, otherwise stop, where n_{sim} is the desired number of iterations.

In the algorithm described above no approximation of the likelihood is required, given that the DSGE parameters are drawn conditional on the Markov-switching states. If agents in the model know the transition matrix observed ex-post by the econometrician (i.e. $H^{sp} = H^m$), as is the case in the benchmark model, steps 1 and 2 need to be modified to take into account that a change in the transition matrix also implies a change in the law of motion of the DSGE states. In this case, I employ a Metropolis-Hastings step in which the DSGE states are regarded as observed variables. Please refer to appendix B for further details.

5 The Benchmark Model

The benchmark model allows for a total of four regimes, two for the structural parameters and two for the stochastic volatilities. The transition matrix that enters the model and is used by agents to form expectations, H^m, is assumed to coincide with the one observed by the econometrician, H^{sp}. The priors for the model parameters are reported in appendix A.[11] The results shown below are based on 1,200,000 Gibbs sampling replications. The first 400,000 draws are disregarded as burn-in and of the remaining 800,000 one every ten draws is retained. The posterior moments vary little over the retained draws providing evidence of convergence.[12]

[11] The priors for the Taylor rule parameters and the stochastic volatilities are consistent with the normalization used to identify the regimes: $\psi_1\left(\xi_t^{sp}=1\right) > \psi_1\left(\xi_t^{sp}=2\right)$ and $\sigma_z\left(\xi_t^{vo}=1\right) > \sigma_z\left(\xi_t^{vo}=2\right)$. Virtually identical results are obtained when assuming the "same" (truncated) priors for the two regimes.

[12] See appendix G.

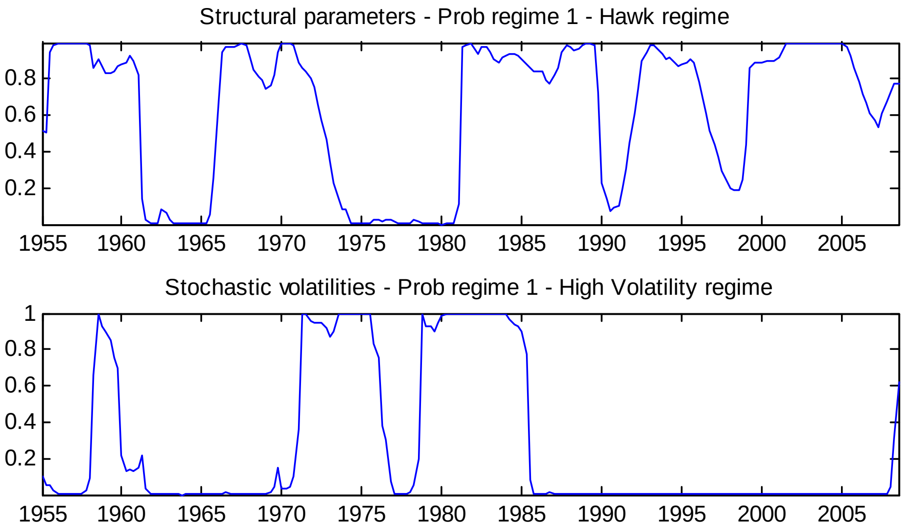

Figure 2: MS-DSGE model, posterior mode estimates. Top panel, probability of regime 1 for the structural parameters, the *Hawk* regime; lower panel, probability of regime 1 for the stochastic volatilities, *high volatility* regime.

Parameter	$\xi_t^{sp} = 1$	$\xi_t^{sp} = 2$
ψ_1	2.0528 (1.3721,2.5916)	0.5907 (0.3505,0.9892)
ψ_2	0.2744 (0.1088,0.4529)	0.3824 (0.2112,0.7882)
ρ_R	0.7530 (0.6299,0.8323)	0.7881 (0.6994,0.8798)
τ	2.8744 (2.0625,3.8259)	
κ	0.0257 (0.0152,0.0379)	
ρ_g	0.8404 (0.7957,0.8820)	
ρ_z	0.9071 (0.8677,0.9423)	
r^*	0.4822 (0.3873,0.5757)	
π^*	0.8444 (0.7207,0.9789)	

Parameter	$\xi^{vo} = 1$	$\xi^{vo} = 2$
σ_R	0.3211 (0.2555,0.4097)	0.0741 (0.0616,0.0883)
σ_g	0.3522 (0.2689,0.4552)	0.1483 (0.1184,0.1821)
σ_z	1.8538 (1.2719,2.6622)	0.5842 (0.3961,0.8352)
σ_y	0.0637 (0.0316,0.1172)	
σ_π	0.2929 (0.2565,0.3305)	
σ_F	0.0286 (0.0148,0.0505)	

$diag\,(H^{sp})$	$diag\,(H^{vo})$
0.9186 (0.8245,0.9785)	0.8948 (0.8152,0.9554)
0.9211 (0.8512,0.9781)	0.9556 (0.9224,0.9811)

Table 1: Means and 90% error bands of the DSGE parameters and of the transition matrix diagonal elements.

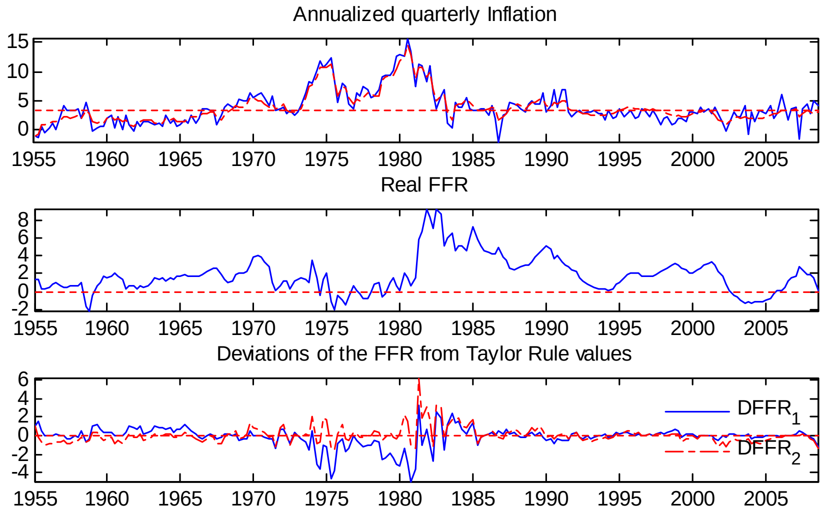

Figure 3: The top panel reports annualized quaterly inflation (observed and filtered) and the inflation target. The second panel contains the annualized real FFR as implied by the model, computed as $R_t - E_t(\pi_{t+1})$. The last panel displays the differences between the observed FFR and the ones implied by the two alternative Taylor rules (i.e., $DFFR_1 = FFR - FF\widehat{R(\xi_t^{sp} = 1)}$).

5.1 Parameters estimates and regime probabilities

Table 1 reports the means and 90% error bands for the DSGE parameters and transition matrices diagonal elements. Concerning the parameters of the Taylor rule, we find that under regime 1 ($\xi_t^{sp} = 1$) the Federal Funds Rate reacts strongly to deviations of inflation from its target, while output gap does not seem to be a major concern. The opposite occurs under regime 2. The degree of interest rate smoothing turns out to be similar across regimes. For obvious reasons, I shall refer to regime 1 as the *Hawk* regime, while regime 2 will be the *Dove* regime. Interestingly enough, if the two regimes were taken in isolation and embedded in a fixed coefficient DSGE model, only the former would imply determinacy.[13]

The point estimate of the inflation target is 0.8444, implying a target/steady state for annual inflation around 3.38%. The top panel of figure 3 displays the series of quarterly annualized inflation and the corresponding target/steady state value. There are some notable deviations, especially during the '60s and the '70s.

As for the other parameters, the low value of the slope of the Phillips curve ($\kappa = 0.0257$)

[13] As explained in section (3.2), the problem of determinacy/indeterminacy in the context of a MS-DSGE model is a very complicated one. At this stage, there is no solution method that can be used to establish determinacy for the model considered in this paper.

is particularly relevant, since such a small value implies a very high sacrifice ratio. In other words, in order to bring inflation down the Federal Reserve needs to generate a severe recession.

Figure 2 shows the (smoothed) probabilities assigned to $\xi_t^{sp} = 1$ (top panel) and $\xi_t^{vo} = 1$ (lower panel). Confronting these probabilities with narrative accounts of monetary policy history is a way to understand whether the results are reasonable. However, before proceeding, a *caveat* is in order. In interpreting the probabilities assigned to the two regimes the reader should take into account how these are related to the estimate of the inflation target. In other words, a high probability assigned to the *Dove* regime does not automatically imply a loose monetary policy, but only that the Fed is relatively unresponsive to deviations of inflation from the target. To facilitate the interpretation of the results, the third panel of figure 3 reports the difference between the observed Federal Funds rate and the interest rate that would be implied by the two Taylor rules. **During a period** characterized by high inflation, a large negative difference between the observed interest rate and its counterfactual value under regime 1 ($DFFR_1 = FFR - \widehat{FFR(\xi_t^{sp} = 1)}$), implies that the Fed is not responding strongly enough to inflation deviations. On the other hand, a large positive value of this same variable during a period of low inflation suggests that the FFR is relatively high.

Monetary policy turns out to be active during the early years of the sample, from 1955 to 1958, and with high probability during the following three years. Romer and Romer (2002) provide narrative evidence in favor of the idea that the stance of the Fed toward inflation during this period was substantially similar to that of the '90s. They also show that a Taylor rule estimated over the sample 1952:1-1958:4 would imply determinacy. Furthermore, after the presidential election of 1960, Richard Nixon blamed his defeat on excessively tight monetary policy implemented by the Fed. At that time, Fed chairman Martin had a clear goal in mind, that the Fed was "to take away the punch bowl just as the party gets going", i.e. to raise interest rates in response to an overheated economy.

Over the period 1961-1965 the *Dove* regime was the rule. This should not be interpreted as evidence of a lack of commitment to low inflation. In fact, the opposite is true. The *Dove* regime prevails because, **given a level of inflation very cl**ose to zero, the *Hawk* regime would have required the Fed to lower the FFR. The *Hawk* regime regains the lead during the last five years of Martin's chairmanship.

In February 1970, Arthur F. Burns was appointed chairman by Richard Nixon. Burns is often regarded as responsible for the high and variable inflation that prevailed during the '70s. It is commonly accepted that on several occasions he had to succumb to the requests of the White House. In fact, for almost the entire duration of his mandate, the Fed followed a passive Taylor rule. During these years, the *Hawk* regime would have required a much higher monetary policy interest rate.[14]

This long period of passive monetary policy ended in 1980, shortly after Paul Volcker took office in August 1979. Volcker was appointed with the precise goal of ending the high inflation. The high probability of the *Hawk* regime during these years confirms the widespread belief that he delivered on his commitment.

The middle panel of figure 3 contains the pattern of real interest rates as implied by the model (computed as $R_t - E_t(\pi_{t+1})$). During Burns' chairmanship real interest rates were negative or very close to zero, whereas, right after the appointment of Volcker, they suddenly increased to unprecedented heights. During the following years, inflation started moving down and the economy experienced a deep recession, while the Fed was still keeping the FFR high. Note that the probability of the *Dove* regime increases from zero to slightly positive values, implying that, given the target for inflation, a lower FFR would have been desirable. In other words, there is a non-zero probability, that Volcker set the FFR in a manner *less* responsive to changes in inflation: Regardless of inflation being on a downward sloping path and a severe recession, monetary policy was still remarkably tight.

For the remainder of the sample the *Hawk* regime has been the rule with a couple of important exceptions. The first one occurred during the 1991 recession: The Fed decided to mitigate the recession even if inflation was above the target. On the other hand, the relatively high values for the probability of the *Dove* regime during the second half of the '90s and toward the end of sample, indicate a FFR that was too high compared to what would have been implied by the *Hawk* regime.[15] It is worth emphasizing that the probability of the *Hawk* regime is basically 1 during the period 2001-2005. Several commentators have argued that during those years the FFR had been too low for too long. These results break a lance in favor of the Fed, showing that the remarkably low FFR can be justified in light of a discrete risk of deflation.

These results strongly support the idea that the appointment of Volcker marked a change in Fed's inflation stance and that the '70s were characterized by a passive monetary policy regime. At the same time, they question the wide-spread belief that US monetary policy history can be described in terms of a permanent and one-time-only regime change: pre- and post-Volcker. While a single regime prevails constantly during the chairmanships of Burns and Volcker, the same cannot be said for the remainder of the sample.

Up to this point nothing has been said about the Good Luck hypothesis. Looking at the second panel of figure 2, it emerges that regime 1, characterized by large volatilities for all

[14] Here the use of the words *active* and *passive* follows Leeper (1991). Monetary policy is active when the interest rate is highly responsive to inflation.

[15] The probability of the Hawk regime turns out to be very close to one when imposing tight priors on the target/steady state inflation, forcing it to be around 2.8%. See appendix G.

shocks, prevails for a long period that goes from the early '70s to 1985, with a break between the two oil crises. This result is quite informative because 1984 is regarded as a turning point in US economic history. There are two alternative ways to interpret this finding. On the one hand, even if a regime change occurred well before 1984, perhaps *the conquest of American inflation* was actually determined by a break in the uncertainty characterizing the macroeconomy. On the other hand, this same break might have occurred in response to the renewed commitment of the Federal Reserve to a low and stable inflation. Both interpretations suggest that the uncertainty characterizing the economy and the behavior of the Fed are likely to be interdependent. The Great Inflation was characterized by high volatilities and loose monetary policy, in a similar vein the Great Moderation emerged after a reduction in the volatilities of the structural shocks and a drastic change in the conduct of monetary policy.

Quite interestingly the probability of the high volatility regime rises again at the end of the sample, supporting the idea that the Great Moderation came to an end with the recent economic crisis.

5.2 Impulse response analysis

In order to assess the importance of changes in the behavior of the Federal Reserve for the dynamics of the economy, it is useful to look at differences in the propagation of the shocks across the two monetary policy regimes. Impulse response functions are computed conditional to one regime being in place. Nevertheless, the paths of the variables differ from the ones that would be obtained solving a fixed coefficient DSGE model because agents' expectations reflect the possibility of a regime change.

The first two rows of figure 4 show the impulse responses to a monetary policy shock under the *Hawk* and *Dove* regimes respectively. The initial shock is equal to the standard deviation of the monetary policy shock under the high volatility regime. Both inflation and output decrease following an increase in the FFR. The responses are remarkably similar across the two regimes.

The third and the fourth rows illustrate the impulse responses to a demand shock. Output and inflation increase under both regimes but their responses are slightly stronger under the *Dove* regime. This is consistent with the response of the Federal Funds rate that is larger under the *Hawk* regime, both on impact and over time. Note that the dynamics of the variables are otherwise similar across the two regimes. The Fed does not face any trade-off when deciding how to respond to a demand shock, therefore, the only difference lies in the magnitude of the response.

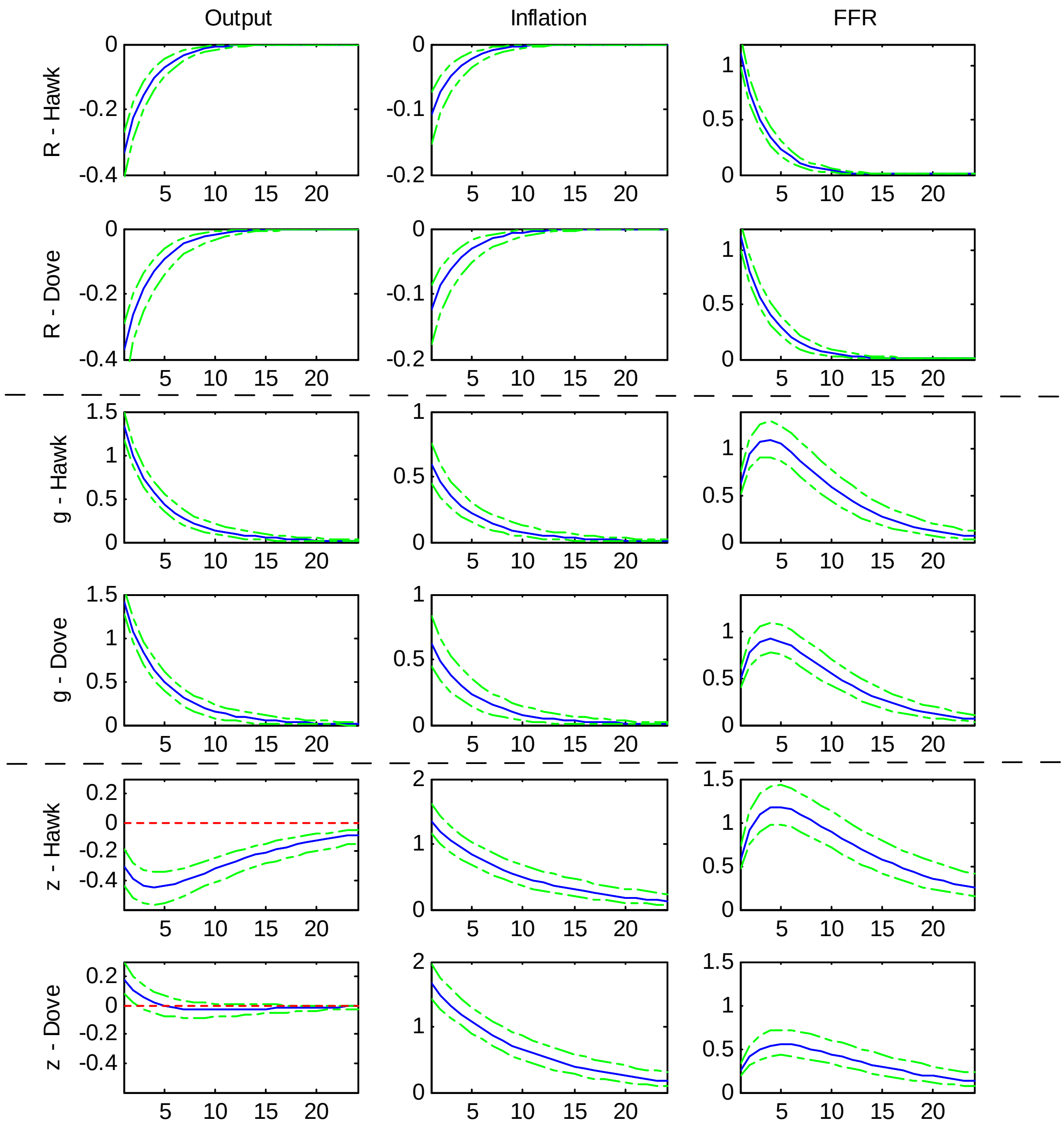

Figure 4: Impulse response functions. The graph is divided in three blocks of two rows each. The three blocks display respectively the impulse responses to a monetary policy shock (R), a demand shock (g), and an adverse technology shock (z). For each block, the first row shows the response under the *Hawk* regime, whereas the second one assumes that the *Dove* regime is in place.

Finally, the last two rows contain the impulse responses to an adverse supply shock, i.e. to an unexpected decrease in z_t. This last set of results is particularly interesting given that, as several economists would agree, one of the causes of the high inflation of the '70s was a series of unfavorable supply-side shocks. The behavior of the Federal Reserve differs substantially across the two regimes. Under the *Hawk* regime the Fed is willing to accept a recession in order to fight inflation. The FFR reacts strongly on impact and keeps rising for one year. On the contrary, under the *Dove* regime the response of the policy rate is much weaker because the Fed tries to keep the output gap around zero, at the cost of higher inflation. Note that on impact the economy experiences a boom: the increase in expected inflation determines a negative real interest rate that boosts the economy in the short run.

Three considerations are in order. First, it is quite evident that the gains in terms of lower inflation achieved under the *Hawk* regime are modest. This can be explained in light of the low value of κ, the slope of the Phillips curve. Second, under the *Dove* regime the Fed is not able to completely avoid a recession, but the contraction in output turns out to be significantly milder. Third, it is commonly accepted that the '70s were characterized by important supply shocks. At the same time, the results of the previous section show that the *Dove* regime has been in place for a large part of those years. Therefore, it might well be that in those years a dovish monetary policy was perceived as "optimal" in consideration of the particular kind of shocks hitting the economy. This seems plausible especially if the Fed was regarding the *sacrifice ratio* as particularly high, as suggested by Primiceri (2006). However, to explore this argument in more detail the probability of moving across regimes should be endogenized. As explained in section 3, this extension would further complicate the model. I regard it as a fascinating area for future research.

5.3 Counterfactual analysis

When working with models that allow for regime changes it is interesting to simulate what would have happened if regime changes had not occurred, or had occurred at different points in time, or had occurred when they otherwise did not. This kind of analysis is even more meaningful in the context of the MS-DSGE model employed in this paper. First of all, like a standard DSGE model, the MS-DSGE can be re-solved for alternative policy rules to address the effects of fundamental changes in the policy regime. The entire law of motion changes in a way that is consistent with the new assumptions around the behavior of the monetary policy authority. Furthermore, the solution also depends on the transition matrix used by agents when forming expectations and on the nature of the alternative regimes. Therefore, we can investigate what would have happened if agents' beliefs about the probability of moving across regimes had been different. This has important implications for counterfactual simulations in which a regime is assumed to have been in place throughout the sample because the expectation mechanism and the law of motion are consistent with the fact that no other regime would have been observed.

Last but not least, new counterfactual simulations can be explored: *Beliefs counterfactuals*. In these counterfactuals agents are endowed with specific beliefs about alternative regimes. These regimes might never occur, but they would still have important effects on the dynamics of the variables. An example that I will explore concerns the appointment of a very conservative Chairman whose behavior can be described by a remarkably hawkish Taylor rule. This particular kind of counterfactual analysis is not possible in the context of a Markov-switching or Time-varying VAR in which there is no explicit role for agents' expectations and forward looking variables.

Two main conclusions can be drawn according to the results of this section. First, little would have changed for the dynamics of inflation if the *Hawk* regime had been in place through the entire sample or if agents had put a large probability on going back to it. According to the results shown above, the only way to avoid high inflation would have been to cause a long and deep recession. The reason is quite simple: The model attributes the large increase in inflation to a technological slowdown that was not under the direct control of the Fed. Second, if agents had put a large enough probability on the occurrence of an even more hawkish regime, inflation would not have reached peaks as high as the ones observed in the late '70s. Furthermore, the cost of keeping inflation low would have been smaller with respect to the counterfactual hypothesis of the *Hawk* regime being in place over the entire sample, suggesting that expectations around alternative regimes can have important effects on the behavior of the economy. Considering that the Volcker era was characterized by a remarkably hawkish monetary policy, we might want to rephrase this result in a suggestive way: If agents had anticipated the appointment of Volcker, the Great Inflation would have been a much less spectacular phenomenon.

5.3.1 No Monetary Policy Shocks

The first set of counterfactual series is obtained by shutting down the monetary policy shocks. For each draw from the posterior the disturbance in the Taylor rule is set to zero independent of the regime in place. The parameters of the model, the sequence for the monetary policy regimes, and the remaining disturbances are left unchanged. Therefore, if the policy rule disturbances had not been set to zero, the simulations would have coincided with the actual series.

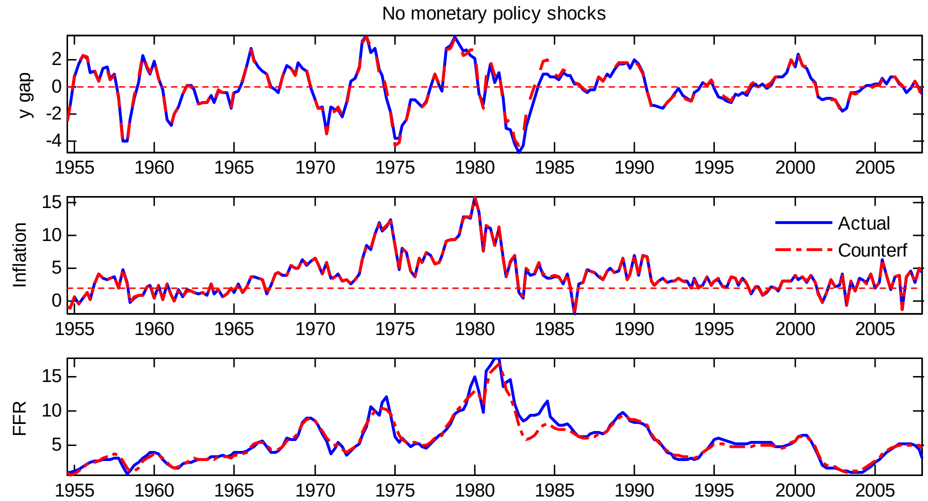

Figure 5: Counterfactual simulation obtained setting the Taylor rule distrurbances to zero.

Figure 5 shows the actual and counterfactual series.[16] The path for inflation is virtually identical to the one that is observed. Deviations can be detected in the series for the output gap, but they are negligible. Interestingly, the FFR would have been lower around the years 1983-1984, suggesting that during those years monetary policy was extremely tight, even under the assumption that the *Hawk* regime was in place. This result corroborates the findings of section 5.1: To some extent Volcker made monetary policy less responsive to inflation. Note that this is in line with the intent of building credibility for a renewed commitment to low and stable inflation.

5.3.2 A Fixed *Hawk* regime

Figure 6 shows the results for the counterfactual simulations obtained by imposing the *Hawk* regime over the entire sample. To make the results consistent with this assumption, the model is solved assuming that agents regard the *Hawk* regime as the only one that is possible. In other words, I solve a fixed coefficient version of the model in which the behavior of the Fed is described by the *Hawk* regime parameters. It is apparent that the Fed would not have been able to completely avoid the rise in inflation, but would only have managed to partially contain it, at the cost of a substantial and prolonged loss in terms of output. In particular, annualized quarterly inflation would not have reached a peak as high as 15%, as it did in the first half of 1980.

[16] For clarity, the figures report only the median of the counterfactual series. Analogous graphs endowed with error bands can be found in appendix G.

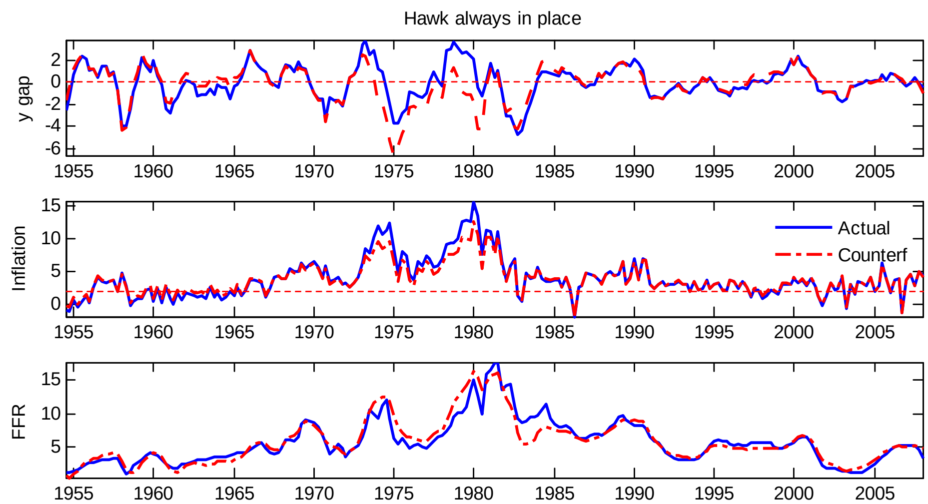

Figure 6: Counterfactual simulation based on the *Hawk* regime being in place over the entire sample.

During the mid-60s, output would have been slightly larger. This is in line with the finding that during those years monetary policy was too tight given a target for inflation around 3%. On the other hand, output would have been lower during the '91 recession. However, these differences are not significant, given that the 90% error bands for the counterfactual series contain the actual ones.

Summarizing, the model does not attribute the rise in inflation to changes in the conduct of monetary policy. It seems that the Fed could only have contained the rise in inflation causing a deep recession. Moreover, while the loss in terms of output would have been certain and large, the gain in terms of inflation would have been quite modest. This is consistent with the idea that the high inflation was driven by a series of shocks on which the Fed had little, if any, control.

5.3.3 An *Eagle* behind the scenes

From what has been shown so far it seems that no reduction in inflation could have been achieved without a substantial output loss. However, the role of agents' beliefs about alternative monetary policy regimes has not yet been explored. The simple and intriguing exercise conducted in this section asks what would have happened if during the high inflation of the '70s agents had put a relatively large probability on the appointment of a very conservative Chairman, willing to fight inflation without any real concern for the state of the real economy. I shall label this hypothetical third scenario *Eagle* regime. The *Eagle* regime

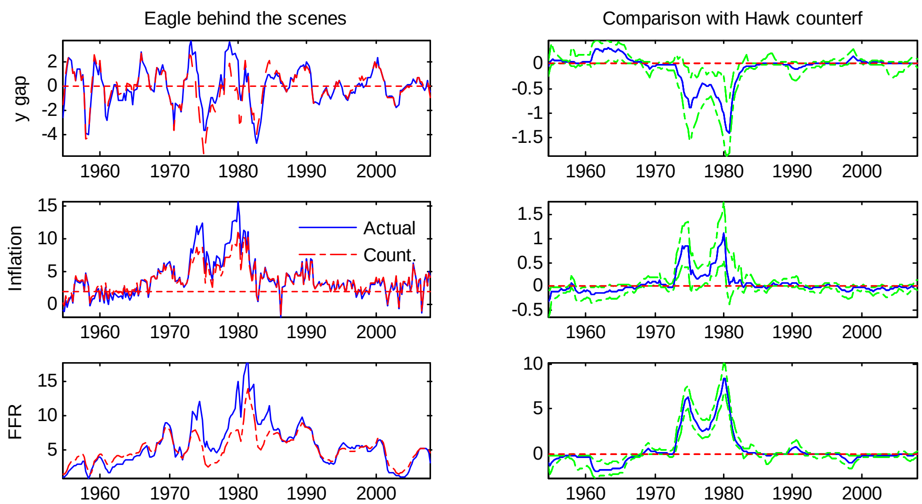

Figure 7: Counterfactual simulation based on having a very hawkish regime, the *Eagle* regime, that receives a very high probability whenever the *Dove* regime is in place, even if it never occurred on sample. The left column reports actual and counterfactual series. For each variable, the right column reports the median and the 90% error bands for the difference between the *Hawk* counterfactual and the *Eagle behind the scenes* counterfactual.

differs from the *Hawk* regime in terms of the response to inflation, that is assumed to be twice as large, and to output, that is half as large. Note that this implies a strong response to deviations of inflation from the target and makes the role of output gap secondary. The *Eagle* regime never occurs over the sample, but I assume that when agents observe the *Dove* regime, they regard the *Eagle* regime as the alternative scenario and they put a relatively large probability on its occurrence. To that end, the probability of staying in the *Dove* regime is reduced by 30 percent. The probability of staying in the *Eagle* regime is equal to the persistence of the *Hawk* regime. From the *Eagle* regime the economy can move only to the *Hawk* regime. These assumptions imply an interesting interpretation of the *Eagle* regime: It is a regime that occurs with high probability after a period of passive monetary policy in order to restore credibility, leading the way to the ordinary active regime.[17]

The left column of figure 7 contains the actual and counterfactual series. The results for inflation look somehow similar to the ones obtained in the previous section. However, there are some notable differences in the output gap and the Federal Funds rate. The former turns out to be larger, while the latter is remarkably lower over the second half of the '70s, the years during which the *Dove*/*Eagle* regime prevails. To make this point stronger, the

[17] Ideally, it would be nice to make the probability of moving to the *Eagle* regime endogenous, but the algorithm used to solve the model is based on the assumption that the transition matrix is exogenous.

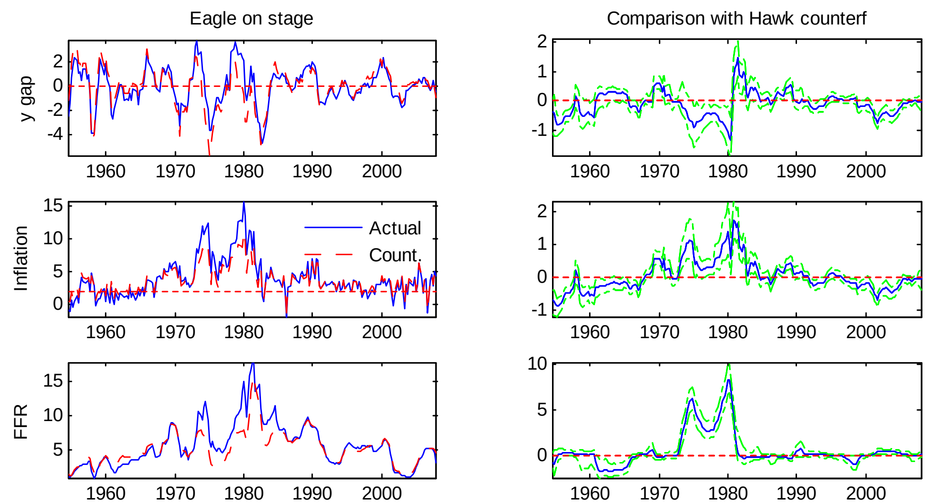

Figure 8: Counterfactual simulation in which the *Hawk* regime is replaced with the *Eagle* regime and the persistence of the *Dove* regime is decreased by 30%. The left column reports actual and counterfactual series. For each variable, the right column reports the median and the 90% error bands for the difference between the *Hawk* counterfactual and the *Eagle on stage* counterfactual.

right column of figure 5 displays, for each series, the difference between the *Hawk*- and the *Eagle*- counterfactual. It turns out that the threat of the *Eagle* regime is enough to deliver the same, if not better, results in terms of low inflation, with a substantial reduction in the output loss. Note that all the results are driven by the high probability that agents assign to the *Eagle* regime. The FFR is low because inflation is relatively low and the *Dove* regime is in place. What keeps inflation down is the fact that agents anticipate the possibility of a drastic change in the conduct of monetary policy.

The goal of this exercise is not to propose a new way to conduct monetary policy: Maintain loose policy today while trying to persuade the public that the Fed will be extremely active in the future. This kind of strategy clearly presents a problem of credibility. However, two lessons can be learned from this experiment. First, it is quite possible that the problem in the '70s was not that the Fed was not reacting strongly enough to inflation, but that there was a lack of confidence around the possibility of a substantial change in the conduct of monetary policy. Second, this exercise suggests that the alternative scenarios that agents have in mind are at least as important as the regime that is in place.

5.3.4 An *Eagle* on stage

The final counterfactual simulation replaces the *Hawk* regime with the *Eagle* regime described in the previous section. Even in this case, the transition matrix is twisted: The probability of staying in the active regime is kept unchanged, while the persistence of the passive regime is lowered by 30 percent. Therefore, agents have in mind a transition matrix that implies only short lasting deviations from this very hawkish regime. This is not consistent with the long lasting dovish period characterizing the '70s, but it might appear in line with the behavior of the Fed after the Volcker disinflation (see figure 2).

The left column of figure 8 contains the counterfactual and actual series. Note that inflation and output would have been lower during the '70s, without substantial increases in the FFR. Even in this case, the result is driven largely by the expectation mechanism. Then, in the early '80s the *Eagle* regime becomes effective and we observe a jump in the FFR and a further reduction in inflation. Quite interestingly, during the early '80s, the path for the FFR is hardly distinguishable from the actual one, suggesting that the *Eagle* regime does a good in job in replicating the behavior of the Federal Reserve during the early years of Volcker's chairmanship.

These outcomes differ from the case in which the *Hawk* regime is assumed to be in place throughout the sample. The right column of figure 8 compares the two counterfactual simulations. If the *Hawk* regime had been replaced by the *Eagle* regime, inflation would have been lower and the slowdown of the early '80s would have been more abrupt. However, it is not clear if the final cost in terms of output would have been different: Output is lower in the early '80s, but it is higher in the second half of the '70s, when the *Dove* regime is in place. In fact, it seems that the gains and costs are likely to cancel out. Therefore, the *Eagle-Dove* combination could be preferred, given that it delivers lower inflation with a similar cost in terms of output.

This last counterfactual simulation points toward an important conclusion: If a Central Bank were able to commit to a flexible inflation targeting, in which severe shocks are temporarily accommodated and followed by a strong commitment to bring the economy back to the steady state, then it would be possible to achieve low inflation with a substantially smaller cost in terms of output. In other words, the effective sacrifice ratio would be much smaller. Admittedly, this kind of policy is not readily practicable. Among other things, the duration of the passive regime matters a lot. When supply-side shocks are large and persistent, as they were in the '70s, if the Central Bank decides to implement a dovish monetary policy, agents are likely to be discouraged about the possibility of moving back to an active regime. In this context, there is no immediate way to persuade agents that a regime change is around the corner.

5.3.5 Quantifying gains and losses

Table 2 quantifies gains and losses arising from the different counterfactuals. The first column contains the *counterfactual sacrifice ratio*, a temptative measure of the cost of bringing inflation down during the Great Inflation of the '70s. This is defined as:

$$CFSR_{T_0,T_1} = \left[\sum_{t=T_0}^{T_1} (y_t - \widehat{y}_t) \right] / \left[\sum_{t=T_0}^{T_1} (\pi_t - \widehat{\pi}_t) \right] \tag{14}$$

where the numerator and the denominator are respectively the cumulative difference between realized and counterfactual output and realized and counterfactual inflation. Intuitively, the larger this number, the larger the cost of lowering inflation. Notice how simply imposing the *Hawk* regime would have implied the largest sacrifice ratio, whereas the *Eagle on stage* counterfactual delivers a much more favorable value. Not surprisingly, the *Eagle behind the scenes* counterfactual returns the lowest value: Here inflation is kept low by agents' expectations.

However, it might be more enlightening to look at the relative performance of the different counterfactuals over the entire sample. The second and third columns contain the counterfactual sum of squared deviations of output and inflation from their respective targets, while the last two columns display the percentage change of these measures with respect to the data. The *Eagle on stage* counterfactual imply smaller losses both in terms of output and inflation compared to the *Hawk* counterfactual and it does better than the *Eagle behind the scenes* in reducing inflation losses. When compared with the actual losses, all counterfactuals imply substantial gains in terms of inflation (up to 66% with the *Eagle on stage* counterfactual) and relatively small losses in terms of output. Desirability of these different outcomes depends on the relative weights given to output and inflation stability, whereas attainability is linked to the ability of the Fed to commit. With regard to this, Bernanke (3 Feb 2003) points out that US monetary policy is conducted according to *constrained discretion:* The Fed can accommodate temporary shocks because, starting with the Volcker disinflation, it has built credibility for its commitment to keeping inflation low and stable. In the context of this paper, this is like saying that the *Eagle on stage* counterfactual is a more credible way to conduct monetary policy after the events of the early '80s.

5.4 Variance decomposition

In this section, I compute the contributions of the structural shocks to the volatility of the macroeconomic variables for all possible combinations of the monetary policy and volatility regimes. It is well known that high inflation is often associated with high volatility. This

Counterfactual	$CFSR_{1970-1984}$	$\widehat{SSD_y}$	$\widehat{SSD_{\pi^A}}$	$\widehat{D\%SSD_y}$	$\widehat{D\%SSD_{\pi^A}}$
Hawk	1.1133 (0.8843,1.4349)	592.9 (513.1,700.1)	928.6 (738.2,1224.1)	+17.77% (+1.99%,+38.97%)	−45.76% (−56.57%,−29.79%)
Eagle behind	0.5292 (0.4184,0.6736)	508.8 (475.0,555.5)	755.3 (596.3,946.7)	+1.12% (−5.59,+10.36)	−55.89% (−64.88,−45.78%)
Eagle on stage	0.6256 (0.4976,0.7835)	589.0 (542.2,648.8)	576.1 (438.55,739.24)	+16.99% (+7.68%,+28.80)	−66.36% (−74.18%,−57.61%)

Table 2: For each counterfactual simulation, the first column reports the counterfactual sacrifice ratio as defined in formula (14), the second and third columns contain the sum of squared deviations of output and inflation from their respective targets, whereas the last two columns contain the percentage changes of these statistics with respect to what observed in the data.

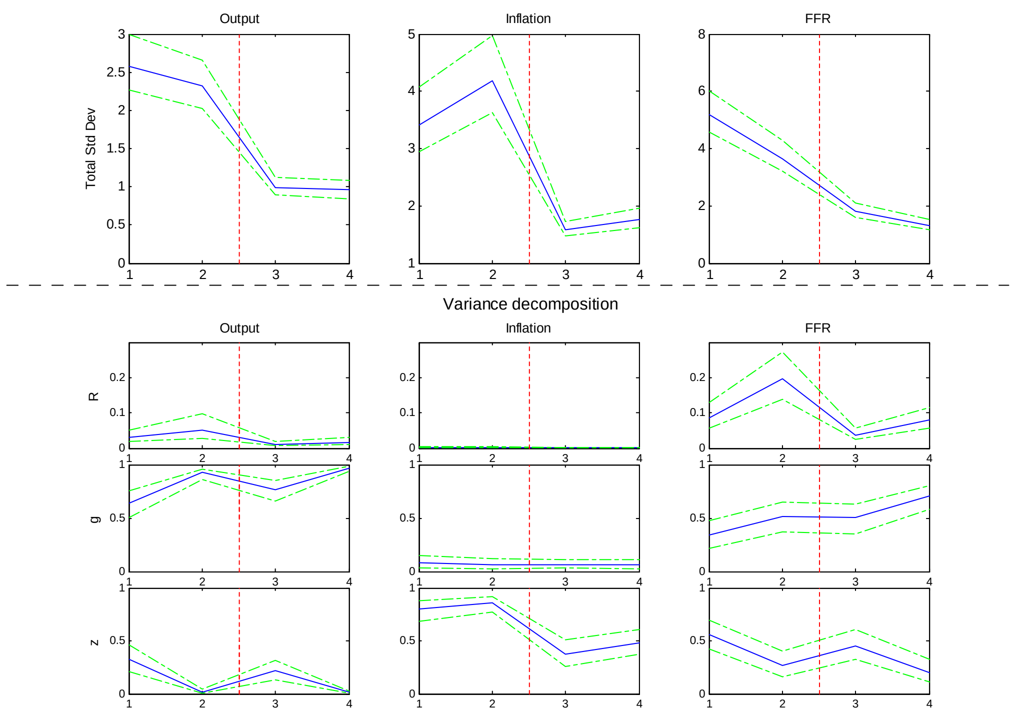

Figure 9: The first row presents analytical standard deviations of the macroeconomic variables for different regime combinations (1 →[High volatility, *Hawk*], 2 →[High, *Dove*], 3 →[Low, *Hawk*], and 4 →[Low, *Dove*]). For each of the regime combinations, the contributions of the monetary policy shock (second row), demand shock (third row), and supply shock (fourth row) to the total volatility are shown in the lower part of the graph. Note that the sum does not need to be 1 because the model allows for an observation error. Each graph reports the median and the 68% error bands.

was surely the case in the '70s. This exercise provides a way to understand what would have changed if the *Hawk* regime had been in place during those years.

Consider the model in state space form (9)-(13). For each draw of the Gibbs sampling algorithm we can compute the conditional covariance matrix as implied by the different regime combinations (ξ^{sp}, ξ^{vo}):[18]

$$
\begin{aligned}
V(S_t|\cdot) &= T(\xi_t^{sp})V(S_t|\cdot)\,T(\xi_t^{sp})' + R(\xi_t^{sp})Q(\xi^{vo})\,R(\xi_t^{sp})' \\
V(Y_t|\cdot) &= ZV\left(S_t|\theta^{sp}, \theta^{vo}, \xi_t^{sp}, \xi_t^{vo}, H^{sp}\right) Z' + U
\end{aligned}
$$

where for each variable x_t, $V(x_t|\cdot)$ stands for $V\left(x_t|\theta^{sp}, \theta^{vo}, \xi_t^{sp}, \xi_t^{vo}, H^{sp}\right)$ and $V(S_t|\cdot)$ is obtained solving the discrete Lyapunov equation. The contribution of the shock i is obtained by replacing $Q(\xi^{vo})$ with $Q_i(\xi^{vo})$, a diagonal matrix in which the only element different from zero is the one corresponding to the variance of the shock i (under regime ξ^{vo}).

The first row of figure 9 plots the analytical standard deviations for the three macroeconomic variables. The first two values, on the left of the dashed line, refer to the high volatility regime, while the third and the fourth values assume that the low volatility regime is in place. In each sub-group, the first point marks the standard deviation under the *Hawk* regime. It is evident that the overall volatility is largely determined by the variance of the underlying structural shocks: Moving from the left to the right side of the dashed line implies a remarkable reduction in the volatility of all macroeconomic variables. Nevertheless, being in the *Hawk* regime implies a reduction in inflation volatility.

The remaining rows of figure 9 present the variance decomposition for the four possible regime combinations. It is quite evident that for inflation the monetary policy regime does not matter much: A large fraction of volatility comes from the supply shocks independent of the behavior of the Federal Reserve. Furthermore, monetary policy shocks play a marginal role. On the other hand, the monetary policy regime is definitely important in explaining the volatility of output. Demand shocks account for almost the entire output volatility when the *Dove* regime is in place. More importantly, supply shocks are relevant only under the *Hawk* regime. Under the *Hawk-high volatility* combination, supply shocks explain around 30% of output volatility, while when the *Dove* regime is in place, their contribution is basically null, independent of the volatility of the supply shock. This result is quite interesting and in line with the impulse response analysis of section 5.2. Under the *Dove* regime, the Fed accommodates supply shocks in order to minimize output fluctuations. This seems to accurately describe what was occurring in the '70s. As for the FFR, the volatility is largely

[18] Here the term "conditional" refers to the regime combination. Note that in fact I am computing an unconditional variance using the law of motion implied by a particular regime combination.

determined by the systematic component of the Taylor rule. Obviously, under the *Hawk* regime monetary policy shocks explain a smaller fraction of the FFR volatility, given that the Fed has a stronger incentive to bring the economy back to the steady state.

6 Alternative specifications

In this section I consider two alternative specifications to capture competing explanations of the macroeconomic dynamics observed over the last fifty years.

6.1 Just Good Luck

A natural alternative to the benchmark specification is represented by a model that allows for heteroskedasticity but assumes no change at all in the behavior of the Federal Reserve. Such a model would explain the Great Moderation invoking Good Luck, i.e. a substantial reduction in the volatility of macroeconomic shocks. Table 3 reports the means and 90% error bands for the DSGE parameters and the transition matrices, while figure 10 plots the probability of regime 1 ($\xi^{vo} = 1$). Once again, regime 1 is the high volatility regime. It prevails around 1958 and between 1970 and 1985, with a break between the two oil crises. Even the estimates of the volatilities are remarkably similar to the ones obtained under the benchmark case.

As for the structural parameters, the response to inflation turns out to be modest but larger than 1, while the output gap coefficient and the level of interest rate smoothing are relatively large. Moreover, the steady state real interest rate and the target for inflation are substantially unaffected. The point estimates for the autocorrelation parameters of the shocks are also very close to the ones obtained in the benchmark model, while the degree of interest smoothing is somewhat larger. The remaining structural parameters are substantially unchanged when compared to the estimates obtained under the benchmark specification.

6.2 One-time-only switch

In their seminal contribution Lubik and Schorfheide (2004) consider a fixed parameters DSGE model analogous to the one employed in this paper extending the solution for the case of indeterminacy. They construct posterior weights for the determinacy and indeterminacy region of the parameter space and estimates for the propagation of fundamental and sunspot shocks. According to their results, U.S. monetary policy post-1982 is consistent with determinacy, whereas the pre-Volcker policy is not.

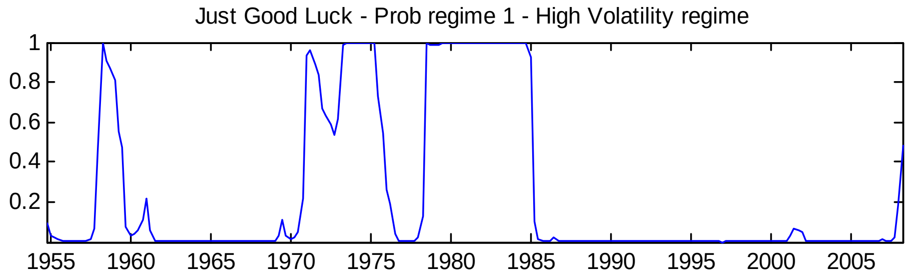

Figure 10: Just Good Luck specification: Probability of regime 1 (High Volatility) at the posterior mode.

Parameter	$\overline{\xi}_t^{sp} = 1$
ψ_1	1.1289 (0.9528,1.3568)
ψ_2	0.4107 (0.3020,0.5390)
ρ_R	0.8380 (0.8032,0.8698)
τ	3.0375 (2.1982,4.0108)
κ	0.0247 (0.0143,0.0378)
ρ_g	0.8378 (0.7938,0.8787)
ρ_z	0.9098 (0.8699,0.9437)
r^*	0.4206 (0.3314,0.5111)
π^*	0.8022 (0.6803,0.9293)

Parameter	$\xi^{vo} = 1$	$\xi^{vo} = 2$
σ_R	0.3707 (0.3106,0.4446)	0.0986 (0.0857,0.1126)
σ_g	0.3642 (0.2792,0.4670)	0.1615 (0.1285,0.1983)
σ_z	1.9862 (1.3424,2.8970)	0.6222 (0.4125,0.9146)
σ_y	0.0617 (0.0314,0.1131)	
σ_π	0.2748 (0.2419,0.3103)	
σ_F	0.0297 (0.0152,0.0539)	

$diag\left(H^{vo}\right)$
0.9048 (0.8258,0.9351)
0.9656 (0.9650,0.9877)

Table 3: Just Good Luck specification: Means and 90% error bands of DSGE parameters and transition matrix diagonal elements.

Here I consider a specification that is in the same spirit but with some important modifications. First, I do not impose a turning date. I let the data decide when the regime change occurred using a Markov-switching model with an absorbing state. Second, I consider a larger sample, spanning the entire WWII postwar era (1954:IV-2008:I). On the other hand, in line with the authors, I assume that: 1) There is only one regime change 2) The regime change is once-for-all and fully credible 3) All parameters of the model are allowed to change. This last assumption allows the steady state levels to change across regimes. I impose that regime 1 implies indeterminacy and I use the results of Lubik and Schorfheide (2004) to compute the likelihood under this hypothesis. The solution under indeterminacy is characterized by some additional parameters.

Table 4 contains the parameter estimates. The change across regimes is somehow more extreme than the one found by Lubik and Schorfheide (2004). The response to inflation jumps from 0.6065 to 3.0897 while the target for (annualized) inflation decreases from 3.7776 to 3.052. Along the same lines, the response to output gap is substantially reduced: from 0.49 to 0.17. Furthermore, the slope of the Phillips curve is remarkably larger under the current regime (0.17 and 0.36). The values of the other structural parameters of the model do not present dramatic changes across regimes and are also quite similar to the ones obtained under the previous specifications. On the other hand, the change in the volatilities of the shocks is remarkable.

The time of the change is quite interesting. Figure 11 plots the probability of regime 2, the current regime. This probability does not start moving before 1982 and hits 1 in 1985. In section 4.1 the MS-DSGE picked up the appointment of Volcker with remarkable precision. Here, the regime change seems to occur several years later. This shows a potential advantage of the benchmark model that allows volatilities and monetary policy rules to evolve according to two independent chains. The MS-DSGE model seems to be able to recognize when the change in the *intents* of the Fed occurred, even if the control over inflation and the break in the volatility of the shocks took place only some years later.

7 Model comparison

Different specifications provide competing explanations regarding the causes of the changes in the reduced form properties of the macroeconomy. The most sensible way to determine which of them returns the most accurate description of the data is to conduct model comparison.

Bayesian model comparison automatically penalizes models with a larger number of parameters and it is based on the posterior odds ratio:

$$\frac{P(M_i|Y_T)}{P(M_j|Y_T)} = \frac{P(Y_T|M_i)}{P(Y_T|M_j)} \frac{P(M_i)}{P(M_j)}$$

where M_i and M_j are two competing models.

The second term on the RHS is the prior odds ratio, i.e. the relative probability assigned to the two models before observing the data, while the first term is the Bayes factor, the ratio of marginal likelihoods. Assuming that all models are regarded as equally likely *a priori*, the Bayes factor is all we need to conduct model comparison.

The marginal data density is computed according to the method described in Sims et al. (2008) who modify the approach proposed by Geweke (1999) to deal with the fact that when allowing for time-varying parameters the posterior tends to be non-Gaussian.

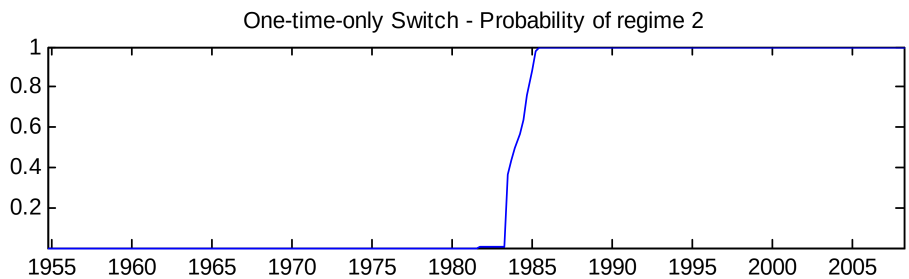

Figure 11: One-time-only switch specification: Probability of regime 2 at the posterior mode.

Parameter	$\xi_t^{sp}=1$	$\xi_t^{sp}=2$
ψ_1	0.6065 (0.3479,0.8441)	3.0897 (2.1839,4.1750)
ψ_2	0.4906 (0.2829,0.7381)	0.1678 (0.0437,0.3368)
ρ_R	0.8590 (0.8048,0.9021)	0.7646 (0.6732,0.8374)
τ	2.3948 (1.6702,3.2460)	1.8398 (1.1929,2.7198)
κ	0.1741 (0.0894,0.2981)	0.3631 (0.1250,0.6989)
ρ_g	0.7927 (0.7073,0.8636)	0.8902 (0.8405,0.9347)
ρ_z	0.7935 (0.7167,0.8652)	0.8110 (0.7379,0.8718)
r^*	0.4564 (0.2724,0.6562)	0.4607 (0.3073,0.6303)
π^*	0.9444 (0.6637,1.2454)	0.7630 (0.6722,0.8576)

Parameter	$\xi^{vo}=1$	$\xi^{vo}=2$
σ_R	0.2490 (0.2221,0.2792)	0.0887 (0.0670,0.1162)
σ_g	0.2684 (0.1953,0.3525)	0.1356 (0.1074,0.1683)
σ_z	1.5328 (1.8171,2.0209)	0.4471 (0.3757,0.5340)
ρ_{gz}	0.3467 (0.0173,0.6389)	0.6911 (0.4079,0.8618)
σ_η	0.0391 (0.0164,0.0869)	–
$M_{\eta r}$	1.8463 (1.1221,2.5719)	–
$M_{\eta g}$	−1.6753 (−2.6471,−0.8612)	–
$M_{\eta z}$	0.8446 (0.4998,1.2621)	–
σ_y	0.0594 (0.0307,0.1074)	
σ_π	0.3099 (0.2789,0.3424)	
σ_F	0.0307 (0.0152,0.0559)	

Table 4: One-time-only switch specification: Means and 90% error bands of DSGE parameters.

Model	$p=0.1$	$p=0.3$	$p=0.5$	$p=0.7$
MS T.R.+heter.+ind H^m	2,391.6	2,390.5	2,390.4	2,390.3
MS T.R.+heter.	2,390.1	2,390.1	2,390.0	2,390.0
Just Good Luck	2,379.0	2,379.0	2,379.0	2,379.0
One-time-only switch	2,349.4	2,349.1	2,349.1	2,349.0

Table 5: Marginal data density (log) for different values of p, the fraction of draws used in the numerical approximation.

Table 5 reports the log marginal data density for different values of p. A smaller value of p implies that less draws are used to compute the marginal data density. This means a better behavior of the weigthing function over the parameter space, but also a greater simulation error. Taking the exponential of the difference between two rows delivers an estimate of the Bayes factor for the corresponding models.

The best performing model coincides with the benchmark specification in which the Taylor rule parameters are allowed to switch across regimes. I consider two versions of this model. In one case agents are assumed to know the transition matrix observed ex-post by the econometrician ($H^m = H^{sp}$), while in the other the two matrices are allowed to differ. The second specification returns slightly better results, but not enough to conclude that it has to be preferred to the benchmark model presented in the paper. The third and fourth models correspond respectively to the "Just Good Luck" and "one-time-only switch" specifications. Quite interestingly, the former dominates the latter. This last result is in line with the findings of Justiniano and Primiceri (2008) and it suggests that there are important gains from allowing for heteroskedastic disturbances. On the other hand, it should not be taken as conclusive evidence against the hypothesis that in the '70s the economy was subject to the possibility of sunspot shocks, as suggested by Lubik and Schorfheide (2004) and Clarida et al. (2000). To make such an argument, we would need to formally *test for indeterminacy* in the context of a Markov-switching model. However, what it does suggest is that a break in the volatility of structural shocks is at least an alternative worth noting (or concause) in explaining the high volatility of the '70s even in the context of a DSGE model.

8 Conclusions

Many economists like to think about US monetary policy history in terms of pre- and post-Volcker. The underlying idea is that since the Volcker disinflation the Fed has acquired a better understanding of how to manage the economy and provide a stable and reliable anchor for agents expectations.

This paper has shown that in fact the appointment of Volcker came with a substantial change in the conduct of monetary policy, with the Fed moving from a passive to an active regime. However, the assumption that this represented an unprecedented and once-and-for-all regime change turns out to be misleading. According to a Markov-switching model in which agents form expectations taking into account the possibility of regime changes, the Fed has moved back and forth between a *Hawk* and a *Dove* regime. Under the *Hawk* regime the Fed reacts strongly to deviations of inflation from the target, while under the *Dove* regime output stability turns out to be at least equally important. The two regimes have very different implications for the dynamics of the economy. In particular, given an adverse technology shock, the Fed is willing to cause a large recession to lower inflation only under the *Hawk* regime.

The '70s were surely dominated by the *Dove* regime, with the Fed trying to minimize output losses. However, this is not enough to explain the rise in inflation that occurred in those years. In fact, little would have changed if the *Hawk* regime had been in place over the entire sample: Inflation would have been slightly lower, but with important losses in terms of output. The Great Inflation might well have been the result of a high volatility regime being in place starting from the early '70s to 1984.

However, the appointment of Volcker might have been important for its impact on agents' expectations. Through counterfactual simulations, I have shown that if agents had put a large probability on the appointment of an extremely conservative Chairman, inflation would not have reached the peaks of the late '70s-early '80s. Moreover, the cost in terms of lower output would have been relatively low compared to the case in which the *Hawk* regime is assumed to be in place over the entire sample. Therefore, it seems that the main problem in the '70s might have been a lack of confidence in the possibility of quickly moving back to an active regime. The results point towards the fascinating conclusion that if agents had anticipated the appointment of Volcker, the Great Inflation would have been a less extreme event.

These results imply that there could be important gains in terms of low inflation and stable output from committing to a flexible inflation targeting regime. In such a regime the Fed would accommodate those shocks that would otherwise have pervasive effects on the economy. At the same time, once the shocks are gone, there should be a clear commitment to generate a recession large enough to bring the economy back to equilibrium. Compared to the case in which the Fed simply follows a hawkish regime, the final disinflation could be more painful, but the cumulative cost is likely to be smaller.

Even if with the Volcker disinflation the US did not enter an absorbing state, there is some hope that events like the Great Inflation will not occur again. Not because the Fed is likely to behave differently on impact, but because agents have now seen what follows a period of loose monetary policy. Obviously, this is an optimistic view. First of all, it is not clear to what extent agents learn from the past. More importantly, the probabilities attached to the different regimes are likely to depend on the persistence of the shocks. Policy makers should avoid trying to accommodate those shocks that are likely to persist for a long time because this would determine a change in the probabilities that agents attach to the different regimes. These considerations seem particularly relevant in light of the recent economic turmoil. In the past year, the Federal Reserve has dealt with a pervasive and severe financial crisis. This led to substantial deviations from common monetary policy practice, a massive quantity of liquidity has been injected into the market, and monetary policy has been remarkably loose. In light of the results of this paper, these events do not represent a problem as long as agents do not revise their beliefs about the commitment to low and stable inflation.

References

Benati, L. and Surico, P.: 2008, VAR analysis and the Great moderation, *American Economic Review* . forthcoming.

Bernanke, B.: 3 Feb 2003, Constrained Discretion and monetary policy, Remarks before Money Marketeers of New York University.

Bernanke, B. S. and Mihov, I.: 1998, Measuring monetary policy, *Quarterly Journal of Economics* **113**(3), 862–902.

Bikbov, R.: 2008, Monetary policy regimes and the term structure of interest rates, Columbia Business School.

Boivin, J. and Giannoni, M.: 2006, Has monetary policy become more effective?, *Review of Economics and Statistics* **88**(3).

Boivin, J. and Giannoni, M.: 2008, Optimal monetary policy in a data-rich environment, HEC and Columbia Business School.

Canova, F. and Gambetti, L.: 2004, On the time variations of US monetary policy: Who is right?, *Money Macro and Finance (MMF) Research Group Conference 2004 96*, Money Macro and Finance Research Group. available at http://ideas.repec.org/p/mmf/mmfc04/96.html.

Carter, C. and Kohn, P.: 1994, On Gibbs sampling for state space models, *Biometrika* **81**, 541–553.

Castelnuovo, E., Greco, L. and Raggi, D.: 2008, Time-varying targets and policy regimes: An empirical investigation for the U.S., University of Padua, mimeo.

Cho, S. and Moreno, A.: 2005, A small-sample study of the new-keynesian macro model, *Journal of Money, Credit, and Banking* **38**(6), 1–45.

Clarida, R., Gali, J. and Gertler, M.: 2000, Monetary policy rules and macroeconomic stability: Evidence and some theory, *Quarterly Journal of Economics* **115**, 147–180.

Cogley, T. and Sargent, T. J.: 2005, The conquest of American inflation: Learning and robustness to model uncertainty, *Review of Economic Dynamics* **8**, 528–563.

Cogley, T. and Sargent, T. J.: 2006, Drifts and volatilities: Monetary policies and outcomes in the post WWII U.S., *Review of Economic Dynamics* **8**, 262–302.

Coleman, I. W.: 1991, Equilibrium in a production economy with an income tax, *Econometrica* **59**(4), 1091–1101.

Davig, T., Chung, H. and Leeper, E. M.: 2007, Monetary and fiscal policy switching, *Journal of Money, Credit, and Banking* **39**(4), 607–635.

Davig, T. and Doh, T.: 2008, Changes in inflation persistence: Lessons from estimated markov-switching new keynesian models, Federal Reserve Bank of Kansas City, mimeo.

Davig, T. and Leeper, E. M.: 2006a, Endogenous monetary policy regime change, Indiana University working paper.

Davig, T. and Leeper, E. M.: 2006b, Fluctuating macro policies and the fiscal theory, *NBER Macroeconomics Annual 2006* pp. 247–298.

Davig, T. and Leeper, E. M.: 2007, Generalizing the Taylor principle, *American Economic Review* **97**(3), 607–635.

Evans, G. and Honkapohja, S.: 2001, *Learning and Expectations In Macroeconims*, Princeton University Press, Princeton.

Farmer, R. E. A., Waggoner, D. and Zha, T.: 2006, Minimal state variable solutions to markov-switching rational expectations models, Federal Reserve Bank of Atlanta, mimeo.

Farmer, R. E. A., Waggoner, D. and Zha, T.: 2008, Understanding Markov-switching rational expectations models, Federal Reserve Bank of Atlanta, mimeo.

Favero, C. and Monacelli, T.: 2005, Fiscal policy rules and regime (in)stability: Evidence from the U.S., IGIER-Bocconi University working paper.

Fernández-Villaverde, J. and Rubio-Ramírez, J. F.: 2007, How structural are structural parameters, *2007 NBER Macroeconomics Annual* .

Gelfand, A. and Dey, D.: 1994, Bayesian model choice: Asymptotics and exact calculations, *Journal of the Royal Statistical Society Series B* **56**, 501–514.

Geweke, J.: 1999, Using simulation methods for bayesian econometric models: Inference, development, and communication, *Econometric Reviews* **18**(1), 1–126.

Hamilton, J. D.: 1989, A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica* **57**, 357–384.

Inoue, A. and Rossi, B.: 2009, Identifying the sources of instabilities in macroeconomic fluctuations, Duke University, mimeo.

Ireland, P.: 2007, Changes in the federal reserve's inflation target: Causes and consequences, working paper.

Justiniano, A. and Primiceri, G.: 2008, The time varying volatility of macroeconomic fluctuations, *American Economic Review* . forthcoming.

Justiniano, A. and Primiceri, G.: 2009, Potential and natural output, Northwestern University, mimeo.

Kim, C.-J. and Nelson, C. R.: 1999, *State-Space Models with Regime Switching*, MIT Press, Cambridge, Massachusetts.

Kim, C.-J. and Nelson, C. R.: 2004, Estimation of a forward-looking monetary policy rule: A time-varying parameter model using ex-post data, *Journal of Monetary Economics* **53**(8), 1949–1966.

King, T.: 2007, Stochastic trends and time-varying structural parameters in macroeconomic models, Board of Governors of the FRS, mimeo.

Laforte, J.-P.: 2005, DSGE models and heteroskedasticity: A Markov-switching approach, Board of Governors of the Federal Reserve System, mimeo.

Leeper, E.: 1991, Equilibria under active and passive monetary and fiscal policies, *Journal of Monetary Economics* **27**, 129–147.

Leeper, E. and Zha, T.: 2003, Modest policy intervantion, *Journal of Monetary Economics* **50**, 1673–1700.

Liu, Z., Waggoner, D. and Zha, T.: 2008, Sources of the Great moderation: Shocks, frictions, or monetary policy?, Federal Reserve Bank of Atlanta, mimeo.

Lubik, T. and Schorfheide, F.: 2004, Testing for indeterminacy: An application to U.S. monetary policy, *American Economic Review* **94**(1), 190–217.

Lucas, R. E.: 1976, Econometric policy evaluation: A critique, *in* K. Brunner and A. Meltzer (eds), *The Phillips Curve and Labor Markets*, Vol. 1 of *Carnegie-Rochester Series on Public Policy, Vol. 1*, North Holland, Amsterdam.

McCallum, B.: 1983, On non-uniqueness in rational expectations models: An attempt at perspective, *Journal of Monetary Economics* **11**, 139–168.

McCallum, B.: 2003, The unique minimum state variable RE solution is E-stable in all well formulated linear models, NBER Working Paper 9960.

Orphanides, A.: 2002, Monetary policy rules and the Great Inflation, *American Economic Review* **92**(2), 115–120. (Proceedings issue).

Primiceri, G.: 2005, Time varying structural vector autoregressions and monetary policy, *The Review of Economic Studies* **72**(3), 821–852.

Primiceri, G.: 2006, Why inflation rose and fell: Policymakers' beliefs and US postwar stabilization policy, *The Quarterly Journal of Economics* **121**(August), 867–901.

Romer, C. D. and Romer, D.: 2002, A rehabilitation of monetary policy in the 1950's, *American Economic Review* **92**(2), 121–127.

Sargent, T., Williams, N. and Zha, T.: 2006, Shocks and government beliefs: The rise and fall of American inflation, *American Economic Review* **96**(4), 1193–1224.

Schorfheide, F.: 2005, Learning and monetary policy shifts, *Review of Economic Dynamics* **8**(2), 392–419.

Sims, C. A.: 2002, Solving linear rational expectations models, *Computational Economics* **20**(1), 1–20.

Sims, C. A., Waggoner, D. F. and Zha, T.: 2008, Methods for inference in large multiple-equation Markov-switching models, *Journal of Econometrics* **146**, 113–144.

Sims, C. A. and Zha, T.: 2006, Were there regime switches in US monetary policy?, *American Economic Review* **91**(1), 54–81.

Stock, J. H. and Watson, M. W.: 2003, Has the business cycle changed? Evidence and explanations, Presented at Monetary Policy and Uncertainty: Adapting to a Changing Economy, Federal Reserve Bank of Kansas City Symposium, Jackson Hole, Wyoming.

Svensson, L. E. O. and Williams, N.: 2007, Monetary policy with model uncertainty: Distribution forecast targeting, Princeton University, mimeo.

Svensson, L. E. O. and Woodford, M.: 2003, Indicator variables for optimal policy, *Journal of Monetary Economics* **50**, 691–720.

Woodford, M.: 2003, *Interest and prices: Foundations of a thory of monetary policy*, Princeton University Press, Princeton, New Jersey.

A Priors

The specification of the prior distribution for the DSGE parameters is summarized in Table 6, which reports prior densities, means, and standard deviations. Each column of H^{sp}, H^{m}, and H^{vo} is modeled according to a Dirichlet distribution: $H^{x}(\cdot,i) \sim D(a_{ii}^{x}, a_{ij}^{x})$ with $a_{ii}^{x} = 10$, and $a_{ij}^{x} = 1$.

B Gibbs sampling algorithm

At the beginning of iteration n we have: θ_{n-1}^{sp}, θ_{n-1}^{ss}, θ_{n-1}^{vo}, S_{n-1}^{T}, $\xi_{n-1}^{sp,T}$, $\xi_{n-1}^{vo,T}$, H_{n-1}^{sp}, H_{n-1}^{m}, and H_{n-1}^{vo}.

Step 1: Sampling the Markov-switching states $\left(\xi_{n}^{sp,T} \text{ and } \xi_{n}^{vo,T}\right)$

Conditional on the DSGE parameters and on S_{n-1}^{T}, we have a Markov-switching VAR with known hyperparameters:

$$S_t = T\left(\xi_t^{sp}\right) S_{t-1} + R\left(\xi_t^{sp}\right)\epsilon_t \tag{15}$$

$$\epsilon_t \sim N\left(0, Q\left(\xi_t^{vo}\right)\right),\ Q\left(\xi_t^{vo}\right) = diag\left(\theta^{vo}\left(\xi_t^{vo}\right)\right) \tag{16}$$

$$H^{sp}(\cdot,i) \sim D(a_{ii}^{sp}, a_{ij}^{sp}),\ H^{vo}(\cdot,i) \sim D(a_{ii}^{vo}, a_{ij}^{vo}) \tag{17}$$

Parameter	Density	Range	Mean	Std. deviation	90% Bands
$\psi_1\,(\xi^{sp}=1)$	Gamma	$\mathbb{R}^+$	1.5	0.5	$[0.7838, 2.4055]$
$\psi_1\,(\xi^{sp}=2)$	Gamma	$\mathbb{R}^+$	0.7	0.4	$[0, 1949, 1.4955]$
ψ_2	Gamma	$\mathbb{R}^+$	0.25	0.15	$[0.0635, 0.5365]$
ρ_R	Beta	$[0,1)$	0.5	0.2	$[0.1718, 0.8220]$
τ	Gamma	$\mathbb{R}^+$	2	0.5	$[1.2542, 2.8885]$
κ	Gamma	$\mathbb{R}^+$	0.3	0.15	$[0.2224, 0.8676]$
ρ_g	Beta	$[0,1)$	0.8	0.1	$[0.6146, 0.9388]$
ρ_z	Beta	$[0,1)$	0.7	0.15	$[0.4274, 0.9162]$
r^*	Gamma	$\mathbb{R}^+$	0.6	0.3	$[0.2047, 1.1631]$
π^*	Gamma	$\mathbb{R}^+$	0.8	0.1	$[0.6428, 0.9714]$
σ_R	Inv. Gamma	$\mathbb{R}^+$	0.31	0.4	$[0.1180, 0.7237]$
σ_g	Inv. Gamma	$\mathbb{R}^+$	0.38	0.4	$[0.1525, 0.8619]$
$\sigma_z\,(\xi^{vo}=1)$	Inv. Gamma	$\mathbb{R}^+$	2	0.8	$[1.1550, 3.4475]$
$\sigma_z\,(\xi^{vo}=2)$	Inv. Gamma	$\mathbb{R}^+$	1	0.8	$[0.4397, 2.1411]$
σ_y	Inv. Gamma	$\mathbb{R}^+$	0.1	0.3	$[0.0338, 0.2462]$
σ_p	Inv. Gamma	$\mathbb{R}^+$	0.15	0.1	$[0.0527, 0.3649]$
σ_r	Inv. Gamma	$\mathbb{R}^+$	0.05	0.3	$[0.0164, 0.1241]$

Table 6: Prior distributions for DSGE model parameters

Therefore, for given H_{n-1}^{sp} and H_{n-1}^{vo}, Bayesian updating can be used to derive the filtered probabilities of the different regimes. Then, the multimove Gibbs-sampling of Carter and Kohn (1994) can be used to draw $\xi_n^{sp,T}$ and $\xi_n^{vo,T}$ (see step 4 for a description of method).

Step 2: Sampling the transition matrices (H_n^{sp} and H_n^{vo})

Given the draws for the MS state variables $\xi_n^{sp,T}$ and $\xi_n^{vo,T}$, the transition probabilities are independent of S_{n-1}^T and the other parameters of the model and have a Dirichlet distribution. For each column of H_n^{sp} and H_n^{vo} the posterior distribution is given by

$$\begin{aligned} H_n^{sp}(\cdot,i) &\sim D(a_{ii}^{sp}+\eta_{ii}^{sp}, a_{ij}^{sp}+\eta_{ij}^{sp}) \\ H_n^{vo}(\cdot,i) &\sim D(a_{ii}^{vo}+\eta_{ii}^{vo}, a_{ij}^{vo}+\eta_{ij}^{er}) \end{aligned}$$

where η_{ij}^{sp} and η_{ij}^{vo} denote respectively the numbers of transitions from state i^{sp} to state j^{sp} and from state i^{vo} to state j^{vo} and $\left(a_{ii}^{sp}, a_{ij}^{sp}, a_{ii}^{vo}, a_{ij}^{vo}\right)$ are the parameters describing the prior.

Step 3.a: Sampling the DSGE parameters ($\theta_n = \{\theta_n^{sp}, \theta_n^{vo}, \theta_n^{ss}\}$)

Start drawing a new set of parameters from the proposal distribution: $\vartheta_n^{sp} \sim N\left(\theta_{n-1}^{sp}, c^{sp}\overline{\Sigma}^{sp}\right)$, $\vartheta_n^{vo} \sim N\left(\theta_{n-1}^{vo}, c^{vo}\overline{\Sigma}^{vo}\right)$, $\vartheta_n^{oe} \sim N\left(\theta_{n-1}^{oe}, c^{oe}\overline{\Sigma}^{oe}\right)$ (if a block optimization algorithm has been used to find the posterior mode) or $vec\,(\vartheta) \sim N\left(\theta_{n-1}, c\overline{\Sigma}\right)$. Here $\overline{\Sigma}$ is the inverse of the Hessian computed at the posterior mode and c is a scale factor. If $n=1$, set $\theta_{n-1} = \overline{\theta} + c\overline{\Sigma}$, where $\overline{\theta}$ is the posterior mode estimate of the DSGE parameters. A Metropolis-Hastings algorithm is used to accept/reject ϑ. Conditional on $\xi_n^{sp,T}$ and $\xi_n^{vo,T}$ there is no uncertainty around the hyperparameters characterizing the state space form model:

$$y_t = D(\theta^{ss}) + ZS_t + v_t \tag{18}$$

$$S_t = T\left(\xi_t^{sp}\right) S_{t-1} + R\left(\xi_t^{sp}\right)\epsilon_t \tag{19}$$

$$\epsilon_t \sim N\left(0, Q\left(\xi_t^{vo}\right)\right),\ Q\left(\xi_t^{vo}\right) = diag\left(\theta^{vo}\left(\xi_t^{vo}\right)\right) \tag{20}$$

$$v_t \sim N\left(0, U\right),\ U = diag\left(\sigma_y^2, \sigma_\pi^2, \sigma_F^2\right) \tag{21}$$

Therefore, the Kalman filter can be used to evaluate the conditional likelihood according to θ_{n-1}, the old set of parameters, and ϑ, the proposed set of parameters. Then the conditional likelihood is combined with the prior distributions of the DSGE parameters. Compute $cut = min\{1, r\}$ where

$$r = \frac{\ell\left(\vartheta^{sp}, \vartheta^{vo}, \vartheta^{ss}|Y^T, \xi_n^{sp,T}, \xi_n^{vo,T}, ...\right) p\left(\vartheta^{sp}, \vartheta^{vo}, \vartheta^{ss}\right)}{\ell\left(\theta_{n-1}^{sp}, \theta_{n-1}^{vo}, \theta_{n-1}^{ss}|Y^T, \xi_n^{sp,T}, \xi_n^{vo,T}, ...\right) p\left(\theta_{n-1}^{sp}, \theta_{n-1}^{vo}, \theta_{n-1}^{ss}\right)}$$

Draw a random number d from an uniform distribution defined over the interval $[0, 1]$. If $d < r$, $(\theta_n^{sp}, \theta_n^{ss}, \theta_n^{vo}) = (\vartheta^{sp}, \vartheta^{vo}, \vartheta^{ss})$, otherwise set $(\theta_n^{sp}, \theta_n^{ss}, \theta_n^{vo}) = \left(\theta_{n-1}^{sp}, \theta_{n-1}^{ss}, \theta_{n-1}^{vo}\right)$.

Step 3.b: Sampling the transition matrix used by agents H_n^m

Start drawing a new set of values for the columns of H^m using a Dirichlet distribution: $\widetilde{H}^m(\cdot, i) \sim D(b_{ii,n-1}^m, b_{ij,n-1}^m)$, where $b_{ii,n-1}^m$ and $b_{ii,n-1}^m$ depend on the columns of H_{n-1}^m. This step defines the transition probability $q\left(\widetilde{H}^m|H_{n-1}^m\right)$. Then, use a Metropolis-Hastings algorithm to accept/reject $\widetilde{H}^m$. Compute $cut = min\{1, r\}$ where

$$r = \frac{\ell\left(\widetilde{H}^m|Y^T, \theta_n, \xi_n^{sp,T}, ...\right) p\left(\widetilde{H}^m\right) q\left(H_{n-1}^m|\widetilde{H}^m\right)}{\ell\left(H_{n-1}^m|Y^T, \theta_n, \xi_n^{sp,T}, ...\right) p\left(H_{n-1}^m\right) q\left(\widetilde{H}^m|H_{n-1}^m\right)}$$

Draw a random number d from an uniform distribution defined over the interval $[0, 1]$. If $d < r$, $H_n^m = \widetilde{H}^m$, otherwise set $H_n^m = H_{n-1}^m$.

Step 4: Sampling the DSGE state vector $\left(S_n^T\right)$

For a given set of DSGE parameters and MS states, (18)-(21) form a state-space model with known hyperparameters. Step 3 returns a filtered estimate of the state variable: $S_n^T|Y^T$. The multimove Gibbs-sampling of Carter and Kohn (1994) can be used to draw the whole vector of S_n^T. Note that:

$$p\left(S_n^T|Y^T\right) = p\left(S_{T,n}|Y^T\right)\prod_{t=1}^{T-1} p\left(S_t|S_{t+1}, Y^T\right)$$

Therefore, the whole vector $S_n^T|Y^T$ can be obtained drawing $S_{T,n}$ from $p\left(S_{T,n}|Y^T\right)$ and then using a backward algorithm to draw $S_{t,n}$, $t = 1...T-1$. Note that the state space model (18)-(21) is linear and Gaussian. It follows that:

$$\begin{aligned} S_{T,n}|Y^T &\sim N\left(S_{T,n|T}, P_{T,n|T}\right) \\ S_t|Y^T, S_{t+1} &\sim N\left(S_{t,n|t,S_{t+1},}, P_{t,n|t,S_{t+1}}\right) \end{aligned}$$

where

$$S_{T,n|T} = E\left(S_{T,n}|Y^T\right) \tag{22}$$

$$P_{T,n|T} = Cov\left(S_{T,n}|Y^T\right) \tag{23}$$

$$S_{t,n|t,S_{t+1}} = E\left(S_t|Y^T, S_{t+1}\right) \tag{24}$$

$$P_{t,n|t,S_{t+1}} = Cov\left(S_t|Y^T, S_{t+1}\right) \tag{25}$$

Step 3 returns $S_{T,n|T}$ and $P_{T,n|T}$, while $S_{t,n|t,S_{t+1}}$ and $P_{t,n|t,S_{t+1}}$ can be obtained updating the estimate of $S_{t,n}$ combining $S_{t,n|t}$, the filtered estimate from step 3, with the new information contained in $S_{t+1,n}$. See Kim and Nelson (1999) for further details.

Step 5

If $n < n_{sim}$, go back to 1, otherwise stop, where n_{sim} is the desired number of iterations.

Step 1, step 2 and step 3.b when $H^m = H^{sp} = H^{sp,m}$

In this case we cannot draw H_n^{sp} simply counting the number of transitions across the MS states, because a change in the transition matrix implies also a change in the law of motion of the DSGE states. Instead, we can apply a Metropolis-Hastings algorithm treating S_{n-1}^T as observed data and using the Hamilton filter to evaluate the likelihood. In this case, define $cut = min\{1, r\}$ where

$$r = \frac{\ell\left(\widetilde{H}^{sp,m}|S_{n-1}^T, \theta_{n-1}, ...\right) p\left(\widetilde{H}^{sp,m}\right) q\left(H_{n-1}^{sp,m}|\widetilde{H}^{sp,m}\right)}{\ell\left(H_{n-1}^{sp,m}|S_{n-1}^T, \theta_{n-1}, ...\right) p\left(H_{n-1}^{sp,m}\right) q\left(\widetilde{H}^{sp,m}|H_{n-1}^{sp,m}\right)}$$

As a side product, we obtain filtered estimates for the MS states and we can use them to draw $\xi_n^{sp,T}$ and $\xi_n^{vo,T}$ with the usual backward drawing algorithm. Finally, H^{vo} can be drawn jointly with $H^{sp,m}$ or according to the standard procedure described above, given that the law of motion does not depend on H^{vo}.

C The model

Here I present a model that when solved a linearized returns the model presented in the paper.

C.1 Private sector

The economy consists of a continuum of monopolistic firms, a representative household, and a monetary policy authority. The household maximizes the following utility function:

$$E_t \left[\sum_{s=t}^{\infty} \beta^{s-t} \left(\frac{C_s^{1-\tau} - 1}{1-\tau} + \chi \log \frac{M_s}{P_s} - h_s \right) \right] \tag{26}$$

where $C_s^{1-\tau}$ denotes consumption of a composite good, h_s are hours worked, M_s/P_s is the real balance of money, β is the discount factor and $\tau > 0$ is the coefficient of relative risk aversion. The household budget constraint is:

$$C_t + \frac{B_t}{P_t} + \frac{M_t}{P_t} + \frac{T_t}{P_t} = W_t h_t + \frac{M_{t-1}}{P_t} + R_{t-1} \frac{B_{t-1}}{P_t} + D_t \tag{27}$$

The production sector is characterized by a continuum of monopolistically competitive firms facing a downward-sloping demand curve:

$$Y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-1/\upsilon} Y_t \tag{28}$$

The parameter $1/\upsilon$ is the elasticity of substitution between two differentiated goods. The firms take as given the general price level, P_t, and level of activity, Y_t. Whenever a firm wants to change its price, it faces quadratic adjustment costs represented by an output loss:

$$AC_t(j) = \frac{\varphi}{2} \left(\frac{P_t(j)}{P_{t-1}(j)} - \pi \right)^2 Y_t(j) \tag{29}$$

Labor is the only input in a linear production function:

$$Y_t(j) = A_t h_t\,(j) \tag{30}$$

where total factor productivity A_t evolves according to:[19]

$$\ln A_t = \ln \overline{A} + \widetilde{a}_t \tag{31}$$
$$\widetilde{a}_t = \widetilde{a}_{t-1} + \epsilon_{a,t} \tag{32}$$

Here $\widetilde{a}_t$ can be interpreted as an aggregate technology shock.

Therefore, the firm's problem consists in choosing the price $P_t(j)$ to maximize the present value of future profits subject :

$$E_t \left[\sum_{s=t}^{\infty} Q_s \left(\frac{P_s(j)}{P_s} Y_s(j) - W_s h_s(j) - \frac{\varphi}{2} \left(\frac{P_s(j)}{P_{s-1}(j)} - \pi \right)^2 Y_s(j) \right) \right]$$

Where Q_s is the marginal value of a unit of the consumption good: $Q_s/Q_t = \beta \left[u_c(s)/u_c(t) \right] = \beta^{s-t}(C_t/C_s)^{\tau}$.

C.2 Government

The behavior of the central bank is captured by a Taylor rule whose parameters are allowed to change over time according to a Markov-switching process. The central bank sets the nominal interest rate in response to deviations of inflation and output from their target levels:

$$\frac{R_t}{R^*} = \left(\frac{R_{t-1}}{R^*} \right)^{\rho_R(\xi_t^{sp})} \left[\left(\frac{\pi_t}{\pi^*} \right)^{\psi_1(\xi_t^{sp})} \left(\frac{Y_t}{Y_t^*} \right)^{\psi_2(\xi_t^{sp})} \right]^{\left(1-\rho_R(\xi_t^{sp})\right)} e^{\epsilon_{R,t}}$$

where R^* is the steady-state nominal rate, Y_t^* is the target for output, π^* is the target level for inflation. ξ_t^{sp} is an unobserved state variable capturing the monetary policy regime that is in place at time t and evolves according to the transition matrix H^{sp}:

$$H^{sp} = \begin{bmatrix} h_{11}^{sp} & 1 - h_{22}^{sp} \\ 1 - h_{11}^{sp} & h_{22}^{sp} \end{bmatrix}$$

Agents in the model know the probability of moving across regimes and they use this information when forming expectations.

Government expenditure is a fraction ζ_t of total output and it is equally divided among the J different goods. We define $d_t = 1/(1 - \zeta_t)$ and we assume that $\widetilde{d}_t = \ln(d_t/d^*)$ follows

[19] An alternative assumption would be $\ln A_t = \gamma + \ln A_{t-1} + \widetilde{a}_t$. However, this would introduce the additional parameter γ that is hard to identify when using detrended output, given that the parameter would enter only through the discount factor β. The results for such a model are virtually identical to the ones presented in the paper and are available upon request.

a stationary AR(1) process:

$$\widetilde{d}_t = \rho_d \widetilde{d}_{t-1} + \epsilon_{g,t} \tag{33}$$

Therefore $\epsilon_{g,t}$ can be interpreted as a shock to Government expenditure. The government collects a lump-sum tax (or provides a subsidy) to balance the fiscal deficit:

$$\zeta_t Y_t + R_{t-1}\frac{B_{t-1}}{P_t} + \frac{M_{t-1}}{P_t} = \frac{B_t}{P_t} + \frac{M_t}{P_t} + \frac{T_t}{P_t}$$

C.3 Solution and linearization

Solving the household's and firm's optimization problems and then expressing the FOC's, the economy-wide resource constraint, and the Taylor rule in deviations from the steady state, we get:

$$\begin{aligned}
1 &= E_t\left[e^{\widetilde{R}_t-\widetilde{\pi}_{t+1}-\tau(\widetilde{c}_{t+1}-\widetilde{c}_t)}\right] && (34)\\
\frac{(1-\upsilon)(e^{\tau\widetilde{c}_t-\widetilde{a}_t}-1)}{\upsilon\varphi\pi^2} &= \left(e^{\widetilde{\pi}_t}-1\right)\left[e^{\widetilde{\pi}_t}\left(1-\frac{1}{2\upsilon}\right)+\frac{1}{2\upsilon}\right] && (35)\\
&\quad -\beta E_t\left[\left(e^{\widetilde{\pi}_{t+1}}-1\right)e^{\widetilde{\pi}_{t+1}+\widetilde{y}_{t+1}-\widetilde{y}_t-\tau(\widetilde{c}_{t+1}-\widetilde{c}_t)}\right]\\
e^{\widetilde{c}_t-\widetilde{y}_t} &= e^{-\widetilde{d}_t}-\frac{g\varphi\pi^2}{2}\left(e^{\widetilde{\pi}_t}-1\right)^2 && (36)\\
e^{\widetilde{R}_t} &= e^{\rho_R\widetilde{R}_{t-1}+(1-\rho_R)[\psi_1\widetilde{\pi}_t+\psi_2\widetilde{y}_t]+\epsilon_{R,t}} && (37)
\end{aligned}$$

To obtain (5), just take logs on both sides of (37):

$$\widetilde{R}_t = \rho_R\widetilde{R}_{t-1} + (1-\rho_R)\left[\psi_1\widetilde{\pi}_t + \psi_2\widetilde{y}_t\right] + \epsilon_{R,t}$$

Now take a first order Taylor expansion on both sides of (36):

$$\widetilde{c}_t = \widetilde{y}_t - \widetilde{d}_t \tag{38}$$

Then, take a first order Taylor expansion on both sides of (35) and use (38):

$$\kappa\left(\widetilde{y}_t - \widetilde{y}_t^*\right) = \widetilde{\pi}_t - \beta E_t\left[\widetilde{\pi}_{t+1}\right]$$

where we have defined $\kappa = \tau\frac{(1-\upsilon)}{\upsilon\varphi\pi^2}$ and used the fact that in absence of nominal rigidities output would be given by:

$$\widetilde{y}_t^* = \widetilde{d}_t + \frac{1}{\tau}\widetilde{a}_t$$

Rearranging and re-labeling $\widetilde{y}_t^*$ with z_t we obtain (2).

Finally to obtain (2), we take logs of (34) and we use again (38):

$$\widetilde{y}_t = E_t\left[\widetilde{y}_{t+1}\right] - \frac{1}{\tau}\left(\widetilde{R}_t - E_t\left[\widetilde{\pi}_{t+1}\right]\right) + (1-\rho_{\widetilde{d}})\widetilde{d}_t$$

Following Lubik and Schorfheide (2004), I assume that the net effect of shocks to government expenditure/preferences can be summarized by the process (3) and I obtain (1):

$$\widetilde{y}_t = E_t\left[\widetilde{y}_{t+1}\right] - \frac{1}{\tau}\left(\widetilde{R}_t - E_t\left[\widetilde{\pi}_{t+1}\right]\right) + g_t$$

D Approximations of the likelihood

This appendix contains a description of the two algorithms used to approximate the likelihood when maximizing the posterior mode and computing the marginal data density. When combined with Kim's approximation, the *trimming approximation* is, by definition, more accurate. This approximation requires a larger computational burden, but might be more appropriate when dealing with switches in the structural parameters of a DSGE model since the laws of motion can vary quite a lot across regimes.

D.1 Kim's approximation of the Likelihood

In this section I describe Kim's approximation of the likelihood (Kim and Nelson (1999)). Consider the model described by (9)-(13). Combine the MS states of the structural parameters and of the heteroskedastic shocks in a unique chain, ξ_t. ξ_t can assume m different values, with $m = m^{sp} * m^{vo}$, and evolves according to the transition matrix $H = H^{sp} \otimes H^{vo}$. For a given set of parameters, and some assumptions about the initial DSGE state variables and MS latent variables, we can recursively run the following filter:

$$\begin{aligned} S_{t|t-1}^{(i,j)} &= T_j S_{t-1|t-1}^i \\ T_j &= T\left(\xi_t = j\right) \end{aligned}$$

$$\begin{aligned} P_{t|t-1}^{(i,j)} &= T_j P_{t-1|t-1}^i T_j' + R_j Q_j R_j' \\ Q_j &= Q\left(\xi_t = j\right),\ R_j = R\left(\xi_t = j\right) \end{aligned}$$

$$e_{t|t-1}^{(i,j)} = y_t - D - Z S_{t|t-1}^{(i,j)}$$

$$f_{t|t-1}^{(i,j)} = ZP_{t|t-1}^{(i,j)}Z' + U$$

$$\begin{aligned} S_{t|t}^{(i,j)} &= S_{t|t-1}^{(i,j)} + P_{t|t-1}^{(i,j)} Z' \left(f_{t|t-1}^{(i,j)} \right)^{-1} e_{t|t-1}^{(i,j)} \\ P_{t|t}^{(i,j)} &= P_{t|t-1}^{(i,j)} - P_{t|t-1}^{(i,j)} Z' \left(f_{t|t-1}^{(i,j)} \right)^{-1} Z e_{t|t-1}^{(i,j)} \end{aligned}$$

At end of each iteration the $M \times M$ elements of $S_{t|t}^{(i,j)}$ and $P_{t|t}^{(i,j)}$ are collapsed into M elements which are represented by $S_{t|t}^{j}$ and $P_{t|t}^{j}$:

$$S_{t|t}^{j} = \frac{\sum_{i=1}^{M} \Pr\left[\xi_{t-1} = i, \xi_t = j | Y_t\right] S_{t|t}^{(i,j)}}{\Pr\left[\xi_t = j | Y_t\right]}$$

$$P_{t|t}^{j} = \frac{\sum_{i=1}^{M} \Pr\left[\xi_{t-1} = i, \xi_t = j | Y_t\right] \left(P_{t|t}^{(i,j)} + \left(S_{t|t}^{j} - S_{t|t}^{(i,j)} \right) \left(S_{t|t}^{j} - S_{t|t}^{(i,j)} \right)' \right)}{\Pr\left[\xi_t = j | Y_t\right]}$$

Finally, the likelihood density of observation y_t is given by:

$$\ell\left(y_t | Y_{t-1}\right) = \sum_{j=1}^{m} \sum_{i=1}^{m} f\left(y_t | \xi_{t-1} = i, \xi_t = j, Y_{t-1}\right) \Pr\left[\xi_{t-1} = i, \xi_t = j | Y_t\right]$$

$$f\left(y_t | \xi_{t-1} = i, \xi_t = j, Y_{t-1}\right) = (2\pi)^{-N/2} |f_{t|t-1}^{(i,j)}|^{-1/2} \exp\left\{ -\frac{1}{2} e_{t|t-1}^{(i,j)\prime} f_{t|t-1}^{(i,j)} e_{t|t-1}^{(i,j)} \right\}$$

D.2 Trimming approximation

This section proposes an alternative algorithm to approximate the likelihood of a MS-DSGE model. This approach is computationally more intensive, but returns a better approximation of the likelihood, especially when dealing with structural breaks. The idea is to keep track of a limited number of alternative paths for the Markov-switching states. Paths that have been assigned a low probability are trimmed or approximated using Kim's algorithm.

Combine ξ_t^{sp} and ξ_t^{vo} to obtain ξ_t. ξ_t can assume all values from 1 to m, where $m = m^{sp} * m^{vo}$, and it evolves according to the transition matrix $H = H^{sp} \otimes H^{vo}$. Suppose the algorithm has reached time t. From previous steps, we have a $((t-1) \times l_{t-1})$ matrix L containing the l_{t-1} retained paths, a vector L_p collecting the probabilities assigned to the different paths, and a $(n \times l_{t-1})$ matrix L_S and a $(n \times n \times l_{t-1})$ matrix L_P containing respectively means and covariance matrices of the DSGE state vector corresponding to each of the l_{t-1} paths.

The goal is to approximate the likelihood for time t, $\ell\left(y_t | Y^{t-1}\right)$ for a given a set of parameters:

1. $\forall\ i = 1...l_{t-1}$, $\forall\ j = 1...m$, compute a one-step-ahaed Kalman filter with $S^i_{t-1|t-1} = L_s(:,i)$ and $P^i_{t-1|t-1} = L_P(:,:,i)$. This will return $f\left(y_t|\xi^{t-1} = i, \xi_t = j, Y_{t-1}\right)$, i.e. the probability of observing y_t given *history* i and $\xi_t = j$. At the end of this step we will have a total of $l_{t-1} * m$ possible histories that are stored in L'. $\forall\ i$ and $\forall\ j$ save $\widetilde{S}^{(i,j)}_{t|t}$ and $\widetilde{P}^{(i,j)}_{t|t}$ and store them in L'_S and L'_P.

2. Compute the ex-ante probabilities for each of the $l_{t-1} * m$ possible paths using the transition matrix H:

$$\begin{aligned} p_{t|t-1}(j,i) &= p_{t-1|t-1}(i) * H(j,i) \\ p_{t-1|t-1}(i) &= L_p(i) \end{aligned}$$

3. Compute the likelihood density of observation y_t as a weighted average of the conditional likelihoods:

$$f(y_t|Y_{t-1}) = \sum_{j=1}^{m}\sum_{i=1}^{l_t} p_{t|t-1}(j,i)\, f\left(y_t|\xi^{t-1} = i, \xi_t = j, Y_{t-1}\right)$$

4. Update the probabilities for the different paths:

$$\begin{aligned} \widetilde{p}_{t|t}(i') &= \frac{p_{t|t-1}(j,i)\, f\left(y_t|\xi^{t-1} = i, \xi_t = j, Y_{t-1}\right)}{f(y_t|Y_{t-1})} \\ i' &= 1...l_{t-1} * m \end{aligned}$$

   and store them in L'_p.

5. Reorder L'_p in decreasing order and rearrange L'_S, L'_P and L' accordingly. Retain l_t of the possible paths where $l_t = \min\{B, l\}$, where B is an arbitrary integer and $l > 0$ is such that

$$\sum_{i'=1}^{l} \widetilde{p}_{t|t}(i') \geq tr$$

   where $tr > 0$ is an arbitrary threshold (for example: $B = 100$, $tr = 0.99$). Update the matrices L_P, L_S, and L:

$$\begin{aligned} L_P &= L'_P(:,:,1:l_t) \\ L_S &= L'_S(:,1:l_t) \\ L &= L'(:,1:l_t) \end{aligned}$$

6. Rescale the probabilities of the retained paths and update L_p:

$$L_p(i) = p_{t|t}(i) = \frac{\widetilde{p}_{t|t}(i)}{\sum_{j=1}^{l_t} \widetilde{p}_{t|t}(j)}, \quad i = 1...l_t$$

Note that Kim's approximation can be applied to the trimmed paths. In this case, the algorithm explicitly keeps track of those paths that turn out to have the largest probability, whereas all the others are approximated.

E Solving the MS-DSGE model

E.1 FWZ solution method

In what follows I provide an outline of the solution method used in the paper that should suffice for those readers interested in using the algorithm for applied work. Please refer to Farmer et al. (2006) for further details.

As a first step, partition the vector of parameters θ in three subvectors: θ^{sp}, θ^{ss} and θ^{vo} contain respectively the structural parameters, the steady state values and the standard deviations of the shocks:

$$\begin{aligned} \theta^{sp} &= \left[\tau, \kappa, \psi_1, \psi_2, \rho_r, \rho_g, \rho_z\right]' \\ \theta^{ss} &= \left[\ln r^*, \ln \pi^*\right]', \ \theta^{vo} = \left[\sigma_R, \sigma_g, \sigma_z\right]' \end{aligned}$$

and define the DSGE state vector S_t as:

$$S_t = \left[\widetilde{y}_t, \widetilde{\pi}_t, \widetilde{R}_t, g_t, z_t, E_t(\widetilde{y}_{t+1}), E_t(\widetilde{\pi}_{t+1})\right]'$$

The model described by equations (1)-(5) can be rewritten as:

$$\overbrace{\begin{bmatrix} \underset{(n-l)\times n}{\Gamma_{0,1}\left(\xi_t^{sp}\right)} \\ \underset{l\times n}{\Gamma_{0,2}} \end{bmatrix}}^{\Gamma_0\left(\xi_t^{sp}\right)} \underset{n\times 1}{S_t} = \overbrace{\begin{bmatrix} \underset{(n-l)\times n}{\Gamma_{1,1}\left(\xi_t^{sp}\right)} \\ \underset{l\times n}{\Gamma_{1,2}} \end{bmatrix}}^{\Gamma_1\left(\xi_t^{sp}\right)} \underset{n\times 1}{S_{t-1}} + \overbrace{\begin{bmatrix} \underset{(n-l)\times k}{\psi\left(\xi_t^{sp}\right)} \\ \underset{l\times k}{0} \end{bmatrix}}^{\Psi\left(\xi_t^{sp}\right)} \underset{k\times 1}{\epsilon_t} + \overbrace{\begin{bmatrix} \underset{(n-l)\times n}{0} \\ \underset{l\times n}{\pi} \end{bmatrix}}^{\Pi} \underset{l\times 1}{\eta_t} \tag{39}$$

where ξ_t^{sp} follows an m^{sp}-state Markov chain, $\xi_t^{sp} \in M^{sp} \equiv \{1, ..., m^{sp}\}$, with stationary transition matrix H^m, n is the number of endogenous variables ($n = 7$ in this case), k is the number of exogenous shocks ($k = 3$), and l is the number of endogenous shocks ($l = 2$). The fundamental equations of (39) are allowed to change across regimes but the parameters defining the non-fundemental shocks do not depend on ξ_t^{sp}.

The first step consists in rewriting (39) as a *fixed parameters* system of equations in the expanded state vector $\overline{S}_t$:

$$\overline{\Gamma}_0 \overline{S}_t = \overline{\Gamma}_1 \overline{S}_{t-1} + \overline{\Psi} u_t + \overline{\Pi} \eta_t \tag{40}$$

where:

$$\underset{np \times np}{\overline{\Gamma}_0} = \begin{bmatrix} diag\left(a_1\left(1\right), ..., a_1\left(m^{sp}\right)\right) \\ a_2, ..., a_2 \\ \Phi \end{bmatrix} \tag{41}$$

$$\underset{np \times np}{\overline{\Gamma}_1} = \begin{bmatrix} \left[diag\left(b_1\left(1\right), ..., b_1\left(m^{sp}\right)\right)\right]\left(H^m \otimes I_n\right) \\ b_2, ..., b_2 \\ 0 \end{bmatrix} \tag{42}$$

$$\underset{np \times l}{\Pi} = \begin{bmatrix} 0 \\ \pi \\ 0 \end{bmatrix}, \quad \underset{(m^{sp}-1)l \times np}{\Phi} = \begin{bmatrix} e_2' \otimes \phi_2 \\ \vdots \\ e_{m^{sp}}' \otimes \phi_{m^{sp}} \end{bmatrix} \tag{43}$$

$$\overline{\Psi} = \begin{bmatrix} I_{(n-l)m^{sp}} & diag\left(\psi\left(1\right), ..., \psi\left(m^{sp}\right)\right) \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\overline{S}_t = \begin{bmatrix} \iota_{(\xi_t^{sp}=1)} S_t \\ \vdots \\ \iota_{(\xi_t^{sp}=m^{sp})} S_t \end{bmatrix}$$

where Φ will be described later. The vector of shocks u_t is defined as:

$$u_t = \begin{bmatrix} \Xi_{\xi_t^{sp}} \left(e_{\xi_{t-1}^{sp}} \otimes \left(1'_{m^{sp}} \otimes I_n\right) \overline{S}_{t-1} \right) \\ e_{\xi_t^{sp}} \otimes \epsilon_t \end{bmatrix}$$

with

$$\underset{(n-l)h \times nh}{\Xi_i} = \left(diag\left[b_1\left(1\right), ..., b_1\left(m^{sp}\right)\right]\right) \times \left[\left(e_i \otimes 1'_{m^{sp}} - H^m\right) \otimes I_n\right]$$

The error term u_t contains two types of shocks: the *switching* shocks and the *normal* shocks. The normal shocks ($e_{\xi_t^{sp}} \otimes \epsilon_t$) carry the exogenous shocks that hit the structural equations, while the *switching* shocks turn on or off the appropriate blocks of the model to represent the Markov-switching dynamics. Note that both shocks are zero in expectation.

Definition 1 *A stochastic process* $\left\{\overline{S}_t, \eta_t\right\}_{t=1}^{\infty}$ *is a solution to the model if:*

1. $\left\{\overline{S}_t, \eta_t\right\}_{t=1}^{\infty}$ *jointly satisfy equation (39)*

2. *The endogenous stochastic process* $\{\eta_t\}$ *satisfies the property* $E_{t-1}\{\eta_t\} = 0$

3. S_t *is bounded in expectation in the sense that* $\left\|E_t\left\{\overline{S}_{t+s}\right\}\right\| < M_t$ *for all* $s > 0$

As mentioned above, FWZ focus on MSV solutions. They prove the equivalence between the MSV solution to the original model and the MSV solution to the expanded fixed coefficient model (40).

The matrix Φ plays a key role. Definition 1 requires boundness of the stochastic process in solving the model. To accomplish this the solution of the expanded system is required to lie in the stable linear subspace. This is accomplished by defining a matrix Z such that

$$Z'\overline{S}_t = 0 \tag{44}$$

To understand how the matrix Z and Φ are related, consider the impact of different regimes. Supposing regime 1 occurs, the third block of (40) imposes a series of zero restrictions on the variables referring to regimes $i = 2...m^{sp}$. These restrictions, combined with the ones arising from the first block of equations, set the correspondent element of $\overline{S}_t$ to zero. If regime $i = 2...m^{sp}$ occurs, we would like a similar block of zero restrictions imposed on regime 1. Here I describe the definition of Φ such that, using (44), it is possible to accomplish the desired result :

Algorithm 2 *Start with a set of matrices* $\left\{\phi_i^0\right\}_{i=2}^{m^{sp}}$ *and construct* $\overline{\Gamma}_0$*. Next compute the QZ decomposition of* $\{A^0, B\}$*:* $Q^0T^0Z^0 = B$ *and* $Q^0S^0Z^0 = A^0$*. Reorder the triangular matrices* $S = (s_{i,j})$ *and* $T = (t_{i,j})$ *in such a way that* $t_{i,i}/s_{i,i}$ *is in are in increasing order. Let* $q \in \{1, 2..., m^{sp}\}$ *be the integer such that* $t_{i,i}/s_{i,i} < 1$ *if* $i \leq q$ *and* $t_{i,i}/s_{i,i} > 1$ *if* $i > q$*. Let* Z_u *be the last* $np - q$ *rows of* Z*. Partition* Z_u *as* $Z_u = [z_1, ..., z_{m^{sp}}]$ *and set* $\phi_i^1 = z_i^1$*. Repeat the procedure until convergence.*

If convergence occurs the solution to (40) is also a solution to (39) and it can be written as a VAR with time dependent coefficients:

$$S_t = T\left(\xi_t^{sp}, \theta^{sp}, H^{sp}\right)S_{t-1} + R\left(\xi_t^{sp}, \theta^{sp}, H^{sp}\right)\epsilon_t$$

E.2 Alternative solution methods and determinacy

The solution method described in the previous section is not the only one available. Davig and Leeper (2006b) and Davig et al. (2007) consider models that are more general than the linear-in-variables model that are considered here and, in certain special cases, they can be solved explicitly. Their solution method makes use of the monotone map method, based on Coleman (1991). The algorithm requires a discretized state space and a set of initial decision rules that reduce the model to a set of non-linear expectational first-order difference equations. A solution consists of a set of functions that map the minimum set of state variables into values for the endogenous variables. Local uniqueness of a solution must be proved by perturbing the equilibrium decision rules. This solution method is appealing to the extent that it is well suited for a larger class of models, but it suffers from a computational burden. This makes the algorithm impractical when the estimation strategy requires solving the model several times, as is the case in this paper.

Another solution algorithm for a large class of linear-in-variables regime-switching models is provided by Svensson and Williams (2007). This method returns the same solution obtained with the FWZ algorithm when the equilibrium is unique.

Bikbov (2008) generalizes a method proposed by Cho and Moreno (2005) for fixed coefficient New-Keynesian models, to the case of regime switching dynamics. The method returns a solution in the form of a MS-VAR, as in FWZ. However, this is the only similarity between the two approaches. In Bikbov (2008) the solution is achieved by working directly on the original model through an iteration procedure. For the fixed coefficient case, Cho and Moreno (2005) report that, in the case of a unique stationary solution, their method delivers the same solution as obtained with the QZ decomposition method. If the rational expectations solution is not unique the method yields the minimum state variable solution. Unfortunately, it is not clear whether a similar argument applies to the case with Markov-switching dynamics and how to check if a unique stationary equilibrium exists. Furthermore, the algorithm imposes a "no-bubble condition" that, to the best of my knowledge, must be verified by simulation.

F Model comparison

Bayesian model comparison is based on the posterior odds ratio:

$$\frac{P(M_i|Y_T)}{P(M_j|Y_T)} = \frac{P(Y_T|M_i)}{P(Y_T|M_j)} \frac{P(M_i)}{P(M_j)}$$

where M_i and M_j are two competing models.

The second term on the RHS is the prior odds ratio, i.e. the relative probability assigned to the two models before observing the data, while the first term is the Bayes factor, the ratio of marginal likelihoods. Assuming that all models are regarded as equally likely *a priori*, the

Bayes factor is all we need to conduct model comparison.

Let θ be a $(k \times 1)$ vector containing all the parameters of model M_i. Moreover denote the likelihood function and the prior density by $p(Y_T|\theta)$ and $p(\theta)$ respectively. The marginal data density is given by:

$$p(Y_T) = \int p(Y_T|\theta)p(\theta)d\theta \tag{45}$$

The modified harmonic mean (MHM) method of Gelfand and Dey (1994) can be used to approximate (45) numerically. This method is based on the following result:

$$p(Y_T)^{-1} = \int_{\Theta} \frac{h(\theta)}{p(Y_T|\theta)p(\theta)} p(\theta|Y_T)d\theta \tag{46}$$

where Θ is the support of the posterior probability density. The weighting function $h(\theta)$ is a probability density whose support is contained in Θ. A numerical approximation of the integral on the right hand side of (46) can be obtained by montecarlo integration:

$$\begin{aligned} \widehat{p}(Y_T)^{-1} &= \frac{1}{N} \sum_{i=1}^{N} m\left(\theta^i\right) \\ m\left(\theta^i\right) &= \frac{h(\theta^i)}{p(Y_T|\theta^i, M_i)p(\theta^i)} \end{aligned}$$

where θ^i is the *ith* draw from the posterior distribution of $p(\theta|Y_T)$. As long as $m(\theta)$ is bounded above the montecarlo approximation converges at a reasonable rate.

Geweke (1999) suggests an implementation based on the posterior simulator. The weighting function $h(\theta)$ is a truncated multivariate Gaussian density. The mean $\overline{\theta}$ and the covariance $\overline{\Omega}$ are obtained from the posterior simulator. To ensure the boundness condition, choose $p \in (0,1)$ and take

$$h(\theta) = p^{-1} N\left(\theta; \overline{\theta}, \overline{\Omega}\right) I_{\widehat{\Theta}_M}$$

$$\widehat{\Theta}_M = \left\{\theta : \left(\theta - \overline{\theta}\right)' \overline{\Omega}^{-1} \left(\theta - \overline{\theta}\right) \leq \chi^2_{1-p}(k)\right\}$$

where $I_{\widehat{\Theta}_M}$ is an indicator function that is equal to one when $\theta \in \widehat{\Theta}_M$. If $\widehat{\Theta}_M \subsetneq \Theta$, the domain of integration needs to be redefined as $\widehat{\Theta}_M \cap \Theta$.

Sims et al. (2008) point out that while the approach proposed by Geweke works generally well when dealing with fixed coefficients models, problems can arise when it is applied to Markov-switching models. When allowing for time variation of the parameters the posterior tends to be Non-Gaussian. Therefore, they suggest replacing the Gaussian distribution with elliptical distributions centered at the posterior mode, $\overline{\overline{\theta}}$. Then, the sample covariance matrix $\overline{\Omega}$ is replaced with:

$$\overline{\overline{\Omega}} = \frac{1}{N}\sum_{i=1}^{N}\left(\theta^i - \overline{\overline{\theta}}\right)\left(\theta^i - \overline{\overline{\theta}}\right)'$$

The density form of an elliptical distribution centered at $\overline{\overline{\theta}}$ and scaled by $\overline{\overline{S}} = \sqrt{\overline{\overline{\Omega}}}$ is

$$g(\theta) = \frac{\Gamma(k/2)}{2\pi^{k/2}\left|\det\left(\overline{\overline{S}}\right)\right|}\frac{f(r)}{r^{k-1}}$$

where k is the dimension of θ, $r = \sqrt{\left(\theta^i - \overline{\overline{\theta}}\right)'\overline{\overline{\Omega}}^{-1}\left(\theta^i - \overline{\overline{\theta}}\right)}$, and $f()$ is any one-dimensional density defined on the positive reals. Sims et al. (2008) explain how to draw from the elliptical distribution.

G Additional graphs

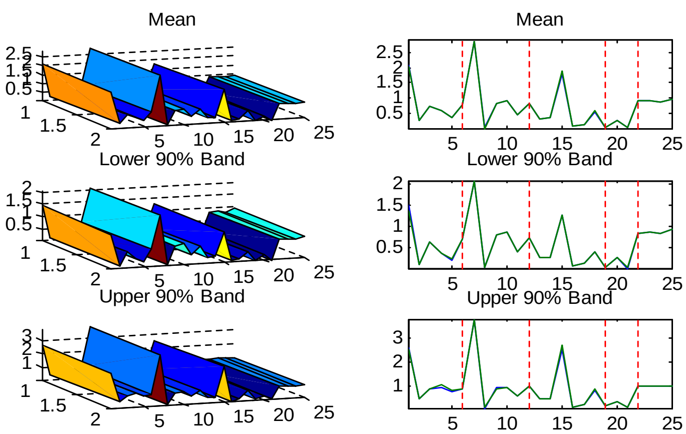

Figure 12: The figure shows means and 90% error bands computed on two not overlapping windows of 40000 Gibbs sampling draws obtained dividing in two parts the 80000 draws used in the paper.

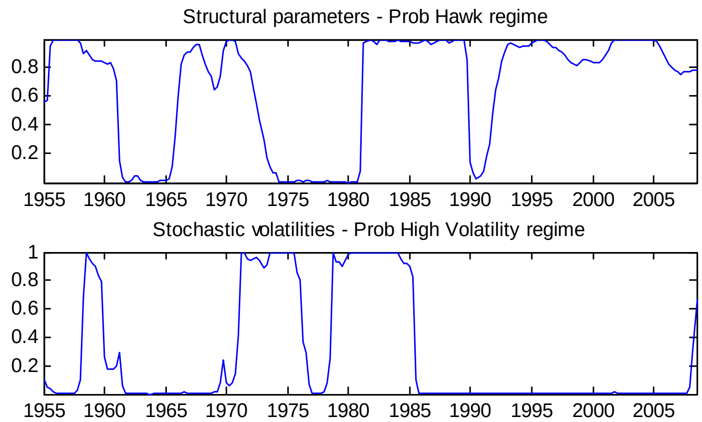

Figure 13: MS-DSGE model, posterior mode estimates when restricting the (annualized) inflation target/steady state to be 3%. Top panel, probability of regime 1 for the structural parameters, the *Hawk* regime; lower panel, probability of regime 1 for the stochastic volatilities, *high volatility* regime.

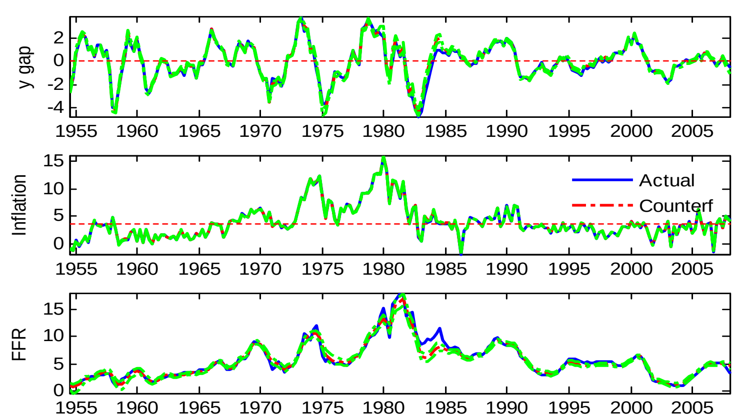

Figure 14: *No monetary policy shock*: Actual, counterfactual, and 68% error bands.

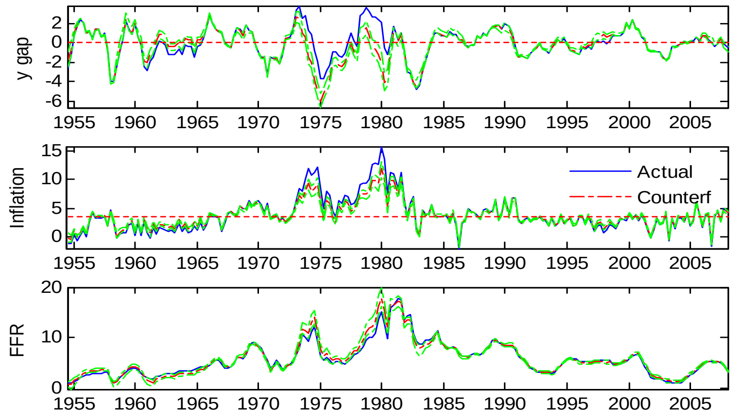

Figure 15: *Hawk always in place*: Actual, counterfactual, and 68% error bands.

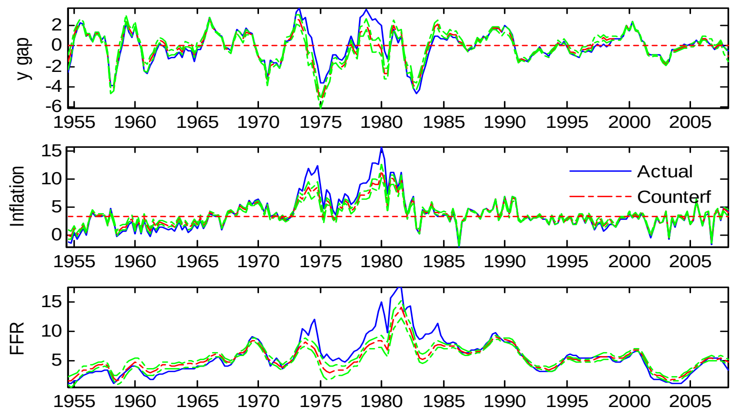

Figure 16: *Eagle behind the scenes*: Actual, counterfactual, and 68% error bands.

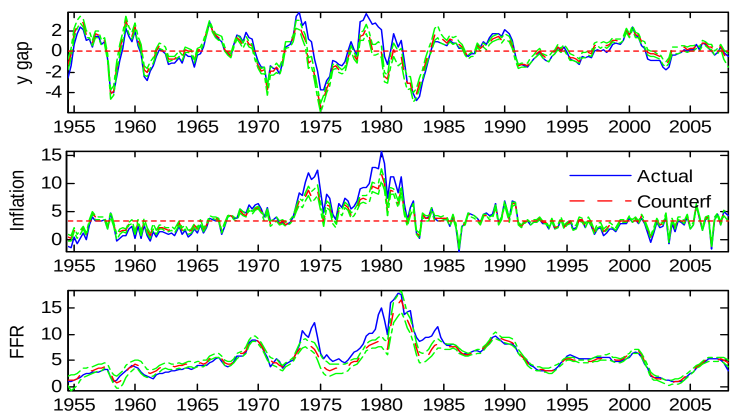

Figure 17: *Eagle on stage*: Actual, counterfactual, and 68% error bands.