Policy Rules, Regime Switches, and Trend Inflation: An Empirical Investigation for the U.S.*

Efrem Castelnuovo
University of Padua and Bank of Finland

Luciano Greco
University of Padua

Davide Raggi
University of Bologna

January 2010

Abstract

This paper estimates Taylor rules featuring instabilities in policy parameters, switches in policy shocks' volatility, and time-varying trend inflation using post-WWII U.S. data. The model embedding the stochastic target performs better in terms of data-fit and identification of the changes in the FOMC's chairmanships. Policy breaks are found not to be synchronized with variations in policy shocks' volatilities. Finally, we detect a negative correlation between systematic monetary policy aggressiveness and inflation gap persistence.

*First draft: June 2008. We thank Gianni Amisano, Francesco Bianchi, Giacomo Carboni, Matteo Ciccarelli, Patrick Crowley, Troy Davig, Esa Jokivuolle, David Mayes, Jouko Vilmunen, Paolo Zagaglia as well as participants at presentations held at University of Padua, CEF 2008 (Paris), Bank of Finland, ICEEE 2009 (Ancona), SNDE 2009 (Atlanta), and RES 2009 (Guildford) for useful comments and suggestions. The opinions expressed in this paper do not necessarily reflect those of the Bank of Finland. All remaining errors are ours. Authors' e-mail accounts: efrem.castelnuovo@unipd.it, luciano.greco@unipd.it, davide.raggi@unibo.it .

Keywords: Policy switches, heteroskedasticity, trend inflation, inflation gap persistence, Markov-Switching models.

JEL Classification: E52, E61, E62.

1 Introduction

Taylor rules are often employed to track the U.S. monetary policy behavior (Taylor (1993)). This paper describes the post-WWII Federal Reserve's conduct with a flexible simple rule in which *policy parameters* are free to *switch* between *states*, *policy shocks* are *heteroskedastic*, and *trend inflation* is *stochastic*.[1] We pay particular attention to this last object. According to Cogley, Primiceri, and Sargent (2009), trend inflation is the single most important factor behind the U.S. inflation dynamics. We then aim at assessing the role played by the low frequency component of inflation in the estimation of the *dating* of the policy shifts, as well as in the investigation of the relationship between policy regimes and inflation persistence. To our knowledge, this is the first contribution dealing with a drifting trend inflation process *jointly* with systematic policy shifts and policy shocks' heteroskedasticity.

We find evidence supporting policy switches and time variation in the estimated inflation target. Crucially, the *timing* of the *switches* in the systematic policy conduct is *sensitive* to the *definition* of the *inflation gap* (the difference between raw and trend inflation) policymakers are assumed to react to. In particular, a constant inflation target leads to *underestimate* the length of the regimes during which the Fed has *actively* tackled inflation (gap) oscillations.[2] Moreover, in presence of a passive regime, the model with a

[1] In this paper we will use the terms "time-varying trend inflation" and "trend-inflation" as synonymous. Ascari (2004) coined the term "trend inflation" to indicate a strictly positive level of steady state inflation around which to approximate firms' non-linear price-setting first order condition. For a related theoretical investigation on the role played by trend inflation in a new-Keynesian framework, see Yun (2005).

[2] Following Leeper (1991), "active" ("passive") monetary policy refers to a policy that adjusts the nominal interest rate more (less) than one-to-one with movements in inflation.

constant inflation target indicates a systematic reaction in line with previous findings (see Clarida, Gali, and Gertler (2000) and Lubik and Schorfheide (2004), among others). However, when embedding trend inflation in the model, such a reaction gets downsized. This indicates that models with a constant inflation target may *underestimate* the magnitude of the shift towards a more hawkish policy, so possibly downplaying the relevance of "good policy" as a driver of the U.S. great moderation.

Interestingly, the model with trend inflation appears to capture the historical changes in the Fed's chairmanship more precisely than its restricted fixed-target counterpart. Regarding the timing issue, shifts in the systematic monetary policy conduct are found *not* to be synchronized with the shifts in the volatility of the monetary policy shocks. This result suggests that the popular "pre- vs. post-Volcker" working hypothesis, often employed to interpret the post-WWII U.S. inflation swings, may offer a misleading (or, at least, incomplete) picture of the different phases characterizing the Fed's monetary policy conduct. Finally, we find a *regime-dependent* inverse correlation between the *systematic* monetary policy aggressiveness and the *persistence* of the inflation gap.

A string of papers has already dealt with instabilities in the U.S. monetary policy rule along several dimensions. Clarida, Gali, and Gertler (2000), Lubik and Schorfheide (2004), Boivin and Giannoni (2006), and Mavroeidis (2009) document breaks in the Federal Reserve's systematic reaction to inflation. Elaborating on this hypothesis, Benati and Surico (2009) and Lubik and Surico (2008) show that a small AD/AS model is able to replicate the great moderation given a policy break in the late 1970s. With a similar strategy, Benati and Surico (2008) mimic the reduction in inflation predictability occurred since 1985 as documented by, among others, Stock and Watson (2007) and D'Agostino, Giannone, and Surico (2006). Castelnuovo and Surico (2009) employ such a policy switch to interpret the instability in the estimated inflation response to a monetary policy shock in standard

VARs. Ireland (2007), Stock and Watson (2007), and Cogley and Sbordone (2008) lend support to time-varying trend inflation. Using different techniques, Schorfheide (2005), Sims and Zha (2006), Liu, Waggoner, and Zha (2007), and Justiniano and Primiceri (2008) support the relevance of shocks' heteroskedasticity to explain the U.S. macroeconomic dynamics. Most of these papers deal with *one* out of the three possibly relevant ingredients (systematic regime shifts, policy shocks' heteroskedasticity, trend inflation) we consider in our analysis.

Another group of contributions is very closely related to our investigation. Bianchi (2009) and Davig and Doh (2009) find evidence in favor of recurrent regime-switches in the systematic policy conduct and the volatility of policy innovations with Markov-Switching DSGE models. However, they assume the inflation target to be fixed in the entire post-WWII period, so falling short of modeling the low frequency evolution of inflation. Cogley, Primiceri, and Sargent (2009) fit multivariate VARs with drifting coefficients and stochastic volatility to a handful of macroeconomic post-WWII U.S. series, and establish that the predictability of the inflation gap has evolved over time and has fallen in the great moderation period. Then, they estimate a new-Keynesian framework over different, predetermined subsamples, and find the enhanced stability of the Fed's inflation target as the single most important factor behind the reductions in the inflation gap volatility and persistence. Differently, we are agnostic as regards the timing of the policy breaks, and let the data free to speak along this dimension. With respect to these contributions, we deal with a partial equilibrium approach, which admittedly misses to exploit possibly informative cross-equation restrictions, but it is obviously more robust to model misspecification. Sims and Zha (2006) and Davig and Leeper (2006a) also perform single-equation estimations, and find evidence in favor of regime-switches and heteroskedastic monetary policy shocks.[3] Again, while they assume a stable inflation target,

[3]Notably, Sims and Zha (2006) find that their best fitting model displays no changes

we uncover the role played by trend inflation in these empirical exercises. Schorfheide (2005) finds evidence in favor of a *switching* inflation target in the context of a small scale macroeconomic model. Liu, Waggoner, and Zha (2007) embed Schorfheide's (2005) inflation target process in a medium scale model featuring heteroskedastic macroeconomic shocks, and find support for a constant trend inflation target over the post-WWII period. Differently, we focus on a *drifting* inflation target process as the one analyzed by Ireland (2007), Stock and Watson (2007), Cogley and Sbordone (2008), and Cogley, Primiceri, and Sargent (2009).

Before moving to our analysis, one caveat is in order. Ascari and Ropele (2009) show that under trend inflation a "modified Taylor principle" may apply, since a more aggressive systematic response to inflation fluctuations is needed to induce equilibrium uniqueness.[4] This implies that, in conducting our empirical exercise, we run the risk to label as - say - "active" a policy that, in Ascari and Ropele's (2009) framework, would possibly lead to multiple equilibria. However, this does not harm our evidence in favor of an unstable systematic monetary policy. Moreover, while being theoretically sound, the quantification of the impact of trend inflation on the indeterminacy region in Ascari and Ropele's (2009) framework has still to be established.[5]

The structure of the paper reads as follows. The next Section presents the regime-switching models we estimate. Section 3 proposes and discusses our empirical findings. Section 4 concludes.

---

in the policy rule coefficients and heteroskedastic policy shocks. Nevertheless, they cannot reject models with an unstable policy rule on the basis of marginal likelihood comparisons.

[4] In presence of trend inflation, the aggregate price index displays an upward trend. Then, firms able to reoptimize i) set higher prices and ii) assign a larger weight on expected - relative to current - realizations of the business cycle to avoid the erosion of relative prices and real profits. This flattens the Phillips curve and induces an increase in the sacrifice ratio. Consequently, a widening of the indeterminacy territory occurs.

[5] The investigation on this issue is left to future research.

2 Model and estimation strategy

2.1 Model

We consider a Markov-Switching model (Hamilton (1989)) characterized by two latent processes, $\boldsymbol{S}_1 = \{S_{1,t}, t = 1, ..., T\}$ and $\boldsymbol{S}_2 = \{S_{2,t}, t = 1, ..., T\}$, which identify the relevant states regarding, respectively, policy conduct and conditional variances of the model's shocks. Our empirical model reads as follows:

$$
\begin{aligned}
i_t &= (1-\rho_{S_1,t})(\overline{r} + \pi_t + \alpha_{S_1,t} z_t + \beta_{S_1,t} y_t) + \rho_{S_1,t} i_{t-1} + \epsilon_{S_2,t} && (1)\\
\pi_t^* &= (1-\rho_\pi)\pi^{LR} + \rho_\pi \pi_{t-1}^* + \xi_{S_2,t} && (2)\\
\pi_t &= \pi_t^* + z_t && (3)\\
z_t &= \phi_{S_1,t} z_{t-1} + \eta_{S_2,t} && (4)\\
\epsilon_{S_2,t} &\sim \mathcal{N}(0, \sigma^2_{\epsilon S_2,t}), \xi_{S_2,t} \sim \mathcal{N}(0, \sigma^2_{\xi S_2,t}), \eta_{S_2,t} \sim \mathcal{N}(0, \sigma^2_{\eta S_2,t}) && (5)
\end{aligned}
$$

where i is the short-term nominal interest rate, π is the inflation rate, y is the output gap, $\overline{r}$ is the unknown long-run real interest rate, π^* is the unobservable (possibly) time-varying inflation target, $z = \pi - \pi^*$ is the inflation gap. In particular, $\boldsymbol{S}_1$ and $\boldsymbol{S}_2$ are defined as two *independent* binary Markov chains with transition probabilities from state i to state j $Pr[S_{k,t} = i \mid S_{k,t-1} = j] = p_{k,ij}$, $k \in \{1, 2\}$, and $i, j \in \{0, 1\}$.

We relax the original Taylor rule along different dimensions. First, we allow (but do not necessarily require) policy parameters to be state-dependent. In particular, in the light of the instability of the U.S. monetary policy conduct documented by several empirical investigations, we enable the "Taylor parameter" α as well as the remaining policy parameters β and ρ to switch between the two states. Second, we assume the time-varying inflation target to follow the autoregressive process (2), whose persistence is captured by the parameter ρ_π and whose unconditional mean is π^{LR}.[6] We interpret π_t^* as

[6] We assume the variances of the "standard" monetary policy shock and the time-

trend inflation underlying the Fed's policy stance conditional on its understanding of the monetary policy transmission mechanism. By contrast, the long-run inflation π^{LR} is meant to be consistent with long-run goals such as sustainable growth and employment. We assume the inflation gap process z_t to follow the autoregressive process (4). This enables us to estimate the persistence of the inflation gap, which is the relevant policy objective (as opposed to raw inflation) in presence of a drifting inflation target. Consistently, we allow the persistence parameter ϕ_{S_1} to be regime-dependent. Finally, we assume shocks to monetary policy, trend inflation, and the inflation gap to be heteroskedastic, mutually uncorrelated independently distributed martingale differences.

2.2 Estimation strategy

We estimate our model with Bayesian techniques by implementing an efficient Markov Chain Monte Carlo (MCMC) strategy through the Gibbs sampler.[7] To deal with identification issues related to "label-switching" (Hamilton, Waggoner, and Zha (2007)), we impose some constraints on the parameters' space. In particular, the "Taylor principle" asserts that monetary authorities can stabilize the economy by moving the policy rate more than one-to-one in response to inflation swings. As in the original rule (Taylor (1993)), eq. (1) has a built-in one-to-one reaction due to the "Fisher relationship" component.[8] We then impose $\alpha_{S_1=0} > 0 > \alpha_{S_1=1}$, i.e. we represent "active"

varying inflation target to be related to the Markov chain $S_{2,t}$. In fact, one may instead interpret them as "policy-related" innovations. Further exercises, not shown for the sake of brevity but available upon request, confirm the robustness of our results to linking such shocks to the chain $S_{1,t}$.

[7] We confine the details of our estimation algorithm to a Technical Appendix, which is available upon request. A comprehensive review on MCMC methods for regime-switching models may be found in Frühwirth-Schnatter (2006).

[8] It is straightforward to rewrite the Taylor rule (1) in the following (somewhat more conventional) version: $i_t = (1-\rho_{S_{1t}})(\gamma_{S_{1t}} + \widetilde{\alpha}_{S_{1t}}\pi_t + \beta_{S_{1t}} y_t) + \rho_{S_t} i_{t-1} + \epsilon_{S_2,t}$ where $\gamma_{S_{1t}} \equiv \overline{r} - \alpha_{S_{1t}}\pi_t^*$, $\widetilde{\alpha}_{S_{1t}} \equiv 1 + \alpha_{S_{1t}}$. When conditioning to a single state, Woodford (2003) shows that the systematic monetary policy reaction to inflation required to pin

systematic monetary policy with $\alpha_{S_1=0}$. Following Cogley, Primiceri, and Sargent (2009), we also assume a lower mean of the inflation gap persistence in correspondence to a tighter systematic policy. Consistently, we employ the priors $\phi_{S_1=0} \sim \mathcal{N}(0.4, 0.1)$ and $\phi_{S_1=1} \sim \mathcal{N}(0.6, 0.1)$. Notice that, these priors notwithstanding, the data are left free to speak as regards the relationship between policy conduct and the inflation gap persistence. As for the second chain, we impose $\sigma^2_{\epsilon S_2=0} > \sigma^2_{\epsilon S_2=1}$ (again, to avoid label switching).

Importantly, we do *not* take any *a-priori* stand on the timing of any of the switches we model. When investigating parameter instabilities with fixed coefficient-models, researchers typically split the U.S. samples at dates corresponding to the beginning of Paul Volcker's FOMC chairmanship or the great moderation. This choice is somewhat "natural" and clearly based on historical events. However, it has some drawbacks. First, it rules out the possibility of regime switches possibly occurred in the 1960s and during the great moderation. Second, it forces policy parameters and volatility shocks to switch *contemporaneously*. A sample-splitting strategy implemented with fixed coefficients is, by construction, ill-suited to study this situation. By contrast, with our approach, which allows policy parameters and shocks to *independently* shift over time, we are well-equipped to detect heteroskedasticity in the policy shocks *within an active systematic policy regime*.

We focus on two different policy rules. Our more flexible model allows the inflation target to follow the process (2). To appreciate the contribution of such stochastic process in our regime-switching context, we also estimate

down a unique equilibrium in a new-Keynesian framework, i.e. the "Taylor principle", is $\widetilde{\alpha}_{S_{1t}} > 1-(1-\delta)\beta_{S_t}/\kappa$, where δ is firms' discount factor and κ is the slope of the Phillips curve. Notably, if policymakers' reaction to business cycle fluctuations $\beta_{S_t} > 0$, a value of $\widetilde{\alpha}_{S_{1t}}$ lower than one (i.e. a value of $\alpha_{S_{1t}}$ lower than zero) can still be consistent with a unique equilibrium. Therefore, the set of constraints we impose to identify the states might induce an overestimation of the "passive" monetary policy phases. We performed an *ex-post* check based on our estimated $\alpha_{S_{1t}}$ and β_{S_t} and conditional to $\delta = 0.99$ and $\kappa = 0.1$ (a calibration widely adopted in the literature), and verified that such overestimation does not occur, i.e. the "passive" states we obtain remain unchanged when considering the exact Woodford's (2003) uniqueness condition.

a version of the model conditional on a fixed inflation target, i.e. $\pi_t^* = \pi^{LR}$ at all times. We fit these two models to U.S. quarterly data spanning the sample 1955:I-2007:II.[9]

3 Empirical evidence

3.1 Trend inflation

Figure 1 displays the U.S. inflation rate along with our two different estimated inflation targets in the post-WWII period. The estimate (posterior mean) of the fixed inflation target is 2.09%. A very different picture emerges when allowing the target to vary over time. The minimum value is estimated to be 1.12 in 1962:II, a phase during which inflation was well under control. Not surprisingly, the highest estimate for trend inflation falls in 1980:IV, at the end of the inflation drift occurred in the 1970s and just before the dramatic disinflation of the early 1980s. In first approximation, trend inflation oscillates around a relatively moderate value during the mid-1950s up to mid-1960s, then it drifts up in correlation with raw inflation, somewhat stabilizes around a new mean value in the second part of the 1970s, dramatically falls at the beginning of the 1980s, and finally stabilizes during the great moderation. Our estimated inflation target is very close to the one proposed by Ireland (2007) and Stock and Watson (2007), belongs to the 90% credible set put forth by Cogley and Sbordone (2008), and it is in line - at least, in first approximation - with the latent factor identified by Cogley, Primiceri, and Sargent (2009).

[9] We retrieved the short-term policy rate (effective federal funds rate), the seasonally adjusted real GDP level Y_t, the estimate of the potential output made by the Congressional Budget Office Y_t^*, and the seasonally adjusted GDP deflator P_t from the Federal Reserve Bank of St. Louis' website. Quarterly observations of the federal funds rate were obtained by averaging monthly observations. The output gap was computed as the percentualized log-deviation of the real GDP from its potential level, i.e. $y_t \equiv 100 \log(Y_t/Y_t^*)$. The inflation rate was calculated as the quarter-by-quarter annualized growth rate of the GDP deflator, i.e. $\pi_t \equiv 400 \log(P_t/P_{t-1})$.

3.2 Estimated coefficients and policy shifts: Model comparison

Table 1 reports the posterior mean estimates - along with the $[5th, 95th]$ percentiles - of our policy rules. Several considerations are in order.

First, in presence of a fixed inflation target, the estimated Taylor rule parameters under the active and passive regime line up with those proposed in the literature.[10] By contrast, we observe a dramatic fall in the Taylor parameter under passive monetary policy when trend inflation is stochastic.

The inflation reaction is basically halved with respect to that estimated under the fixed inflation target assumption, i.e. it moves from 0.79 to 0.35, a value much lower than those proposed in the literature. Interestingly, under the active monetary policy regime the posterior mean of the Taylor parameter is quite comparable between models in terms of magnitude. Consequently, the presence of a time-varying inflation target appears to *augment* the difference between the reactions to inflation in the two regimes. Notably, Cogley, Primiceri, and Sargent (2009) find a much larger value of α when focusing on the 1970s. The difference between their results and ours is likely to be due to their choice of performing full-system estimations conditional to equilibrium uniqueness, a choice that forces the parameter α to assume values satisfying the Taylor principle.

Moving to the remaining parameters, the policy reaction to business cycle oscillations is clearly larger than zero, and the policy moves are implemented with substantial gradualism. The persistence of the inflation target is large, with an estimated posterior mean of ρ_π equal to 0.97, i.e. a process close to the random walk employed by Ireland (2007), Cogley and Sbordone (2008), Stock and Watson (2007), and Cogley, Primiceri, and Sargent (2009).[11] Monetary policy shocks and shocks to the inflation gap are clearly heteroskedastic.

[10] See Clarida, Gali, and Gertler (2000), Table II page 157; Lubik and Schorfheide (2004), Table 3 (first and third columns), page 206.

[11] Consequently, the impact of the long run target π^{LR} on the estimated trend inflation is *de facto* negligible.

Also trend inflation shocks turn out to be heteroskedastic, thus corroborating the evidence recently put forth by Cogley, Primiceri, and Sargent (2009).

Second, our dating of the active/passive regimes displays several departures with respect to the "pre- vs. post-Volcker appointment" reading of the post-WWII U.S. inflation dynamics. Figure 2 - top panel - plots the evolution of the probability of being in an active monetary policy regime. While the 1970s are clearly associated to a passive monetary policy regime, the 1960s are actually related to an active systematic policy conduct. The 1980s and 1990s are also characterized by switches in the policy stance. In particular, the 1991 recession is clearly pin-pointed as a phase in which the Fed acted passively.

Third, Figure 2 - top panel - reveals relevant differences in the dating suggested by the fixed vs. trend inflation target models. In particular, the former suggests a switch towards an aggressive monetary policy in 1983:I, temporally well over the beginning of Paul Volcker's chairmanship in August 1979. By contrast, the trend inflation model suggests a shift to active policy in 1980:IV, shortly after Volcker's appointment. Going back in time, on February 1970, Burns was appointed by Richard Nixon as chairman of the Fed. As stressed by Bianchi (2009), Burns is typically regarded as responsible for the great inflation drift of the 1970s, possibly due to political pressures coming from the White House. Our trend inflation model dates the switch from active to passive monetary policy in 1969:II, very close to Burns' appointment. By contrast, the fixed target model clearly anticipates such event, with a dating suggesting 1968:I. As regards the 1991 recession, the trend inflation model identifies more precisely such an event, suggesting a passive policy in correspondence to the subsample 1989:III-1993:III, much shorter than the 1987:IV-1994:I implied by the fixed target model. The passive conduct might be interpreted in the light of possible fears for future deflation and economic recession leading the Fed to loosen systematic monetary policy so to influence agents' expectations. In general, the model with a constant inflation target

appears to *underestimate* the length of the subsamples in which the Fed has *actively* fought inflation swings.

Importantly, the superiority of the trend inflation model is given formal statistical support. We compute the marginal likelihood *via* the method developed by Chib and Jeliazkov (2001).[12] This method combines a sequence of reduced runs of the MCMC algorithm, performed to evaluate the posterior distributions of the parameters, with a particle filter step, used to compute the likelihood function (see Doucet, de Freitas, and Gordon (2001) for a review of the literature). Our marginal likelihood-based comparison, engaged on the basis of the figures reported in Table 1 (last row), suggests a difference of about 22 log-points. In the language of Kass and Raftery (1995), this is "very strong" evidence in favor of time-varying trend inflation.

Fourth, our posterior probabilities provide evidence against contemporaneous instabilities in the systematic monetary policy vs. policy shocks' volatility. In particular, Figure 2 shows that the recession occurred at the beginning of the 1960s correlates with a passive monetary policy conduct featured by low volatile policy shocks. By contrast, the two recessions in the 1970s are associated to passive monetary policy / high policy shocks, an evidence consistent with the high inflation and output volatilities in the 1970s according to standard AD/AS models. Another interesting asynchrony emerges when looking at the 2001 recession, which occurred in the aftermath of the 'dot-com' bubble burst. In fact, we find no systematic policy switch, a result we interpret as the Fed's attempt to stabilize inflation expectations. However, there is evidence of a switch towards high policy shocks' volatility. Notably, this is a very different situation with respect to the *systematically* passive monetary policy that occurred in the 1970s.

The difference in these two latter policy conducts (i.e. active monetary policy with high policy shocks vs. passive monetary policy) may be bet-

[12] A detailed explanation of our algorithm is presented in a Technical Appendix available upon request.

ter appreciated when thinking of a regime-contingent Taylor rule coupled with a standard *linearized* new-Keynesian AD/AS framework. Working on the standard hypothesis of rational expectations formed on the basis of the policy regime in place, a *systematically active* monetary policy is able to anchor inflation expectations no matter what the volatility of the (white noise) policy shock is. Consequently, if a linearized model returns a reliable approximation of the U.S. economic dynamics, one may interpret the 2001 episode as a reaction to the stock market crash leading to a *temporary* deviation with respect to the Taylor rate *conditional* on a long-run reaction Taylor parameter satisfying the Taylor principle. Of course, a higher-order of approximation of the new-Keynesian framework would acknowledge the role of second moments (namely, switches in the policy shock volatilities) in influencing agents' expectations. To our knowledge, the effects of policy switches and heteroskedasticity in a non-linear monetary policy framework are still to be explored.

Finally, we find statistical evidence in favor of the two-chain set up as opposed to the single chain framework. When conditioning to a single chain $\boldsymbol{S}_1 = \boldsymbol{S}_2 = \boldsymbol{S}$, we obtain marginal likelihoods reading -466.81 for the trend inflation model, and -467.79 for the fixed target framework, i.e much worse than -432.62, which is the marginal likelihood of our best model. Again, our model embedding the stochastic inflation target turns out to be statistically superior. More importantly, the comparison between the two models with trend inflation substantially favors the more flexible two-chain framework, with a difference in terms of log-points of the marginal likelihood of about 34. This result corroborates Bianchi's (2009) and Davig and Doh's (2009) findings on the relevance of modeling different chains for policy parameters and policy shocks, and casts doubts on the standard "pre- vs. post-Volcker" interpretation of the post-WWII U.S. monetary policy conduct.

3.3 Inflation gap persistence

In the light of our empirical evidence, the evolution of the inflation *gap* is of great interest from a policy standpoint. Figure 3 displays the scatter plot of the realizations of ϕ_{S_1} occurring under the passive monetary policy regime against those obtained under the more aggressive one.[13] Pairs located near the 45 degree line indicate constant persistence between the two regimes. By contrast, pairs clustered below the 45-degree line represent a decrease in persistence in correspondence to a systematic policy switch towards active monetary policy.

The inflation gap persistence turns out to be substantially lower in correspondence to an active monetary policy. As much as 96.59% of the pairs are located below the 45-degree line. The percentages collected by Table 2, and computed on the basis of different thresholds (that basically shift the line downward), suggest a remarkable drop in the inflation gap persistence in correspondence to a switch towards a more aggressive systematic monetary policy. In terms of posterior means, the inflation gap persistence falls from $\phi_{S_1=1} = 0.55$ to $\phi_{S_1=0} = 0.25$. These figures corroborate some recent evidence put forward by Cogley and Sbordone (2008) and Cogley, Primiceri, and Sargent (2009), are similar to the inflation gap's normalized spectrum at zero frequency (median values) proposed by Benati and Surico (2008), and support their findings on the negative correlation between monetary policy aggressiveness and the persistence of the inflation gap. Differently from these contributions, our result is obtained by considering subsamples identified with a *data-driven* procedure, which limits the weight of subjective *a priori* assumptions on the timing of the regime shifts.

[13] These scatter-plots are constructed by employing 25,000 draws from our empirical posterior densities.

4 Conclusions

This paper estimates regime-switching Taylor rules with trend inflation for the post-WWII U.S. economy. We find evidence in favor of regime shifts, monetary policy shocks heteroskedasticity, and time-variation of the inflation target. Policy breaks are found not to be synchronized with variations in policy shocks' volatilities. Importantly, the estimated Taylor rule parameters and regime switches are sensitive to the choice of the inflation definition entering the Taylor rule. In particular, a framework embedding an inflation *gap* modeled with a stochastic trend inflation process returns a representation of the Fed's policy conduct closer to historical events (as opposed to that suggested by a fixed inflation target rule). Interestingly from a policy standpoint, we find that a more aggressive systematic policy is associated to a lower inflation gap persistence. We also find evidence of "temporary" deviations from the Taylor rate conditional on active systematic policy, i.e. situations in which highly volatile policy shocks are consistent with the Fed meeting the standard Taylor principle. Possibly, this last finding calls for the employment of non-linear models to assess the impact of such highly volatile policy shocks on private sector's inflation expectations.

Our findings offer support to four quite exciting lines of research. The first regards the estimation of DSGE models with regime switching techniques. Given the instability we found in the U.S. policy rule, two issues are in order. First, how to solve the model in which agents are endowed with a density over future possible policy outcomes? Second, how do models embedding this ingredient square with the data? Theoretical contributions by Davig and Leeper (2007), Davig and Leeper (2009), Farmer, Waggoner, and Zha (2008), and Farmer, Waggoner, and Zha (2009),[14] and empirical endeavors by Bianchi (2009) and Davig and Doh (2009) offer first answers to these questions.

[14] Further considerations on this issue are proposed by Zha (2010).

The second research line refers to the empirical role played by the time-varying inflation target. Our exercise shows that the use of a stochastic process for trend inflation like the one in Ireland (2007), Cogley and Sbordone (2008), Stock and Watson (2007), and Cogley, Primiceri, and Sargent (2009), leads to a Taylor rule enjoying a superior fit of the data. This moves the attention from raw inflation to the inflation gap as the relevant policy object. First investigations on its properties - mainly, its persistence - have been performed by Cogley and Sbordone (2008), Benati and Surico (2008), and Cogley, Primiceri, and Sargent (2009).

Third, it is key to understand *why* the inflation target evolved over time. Imperfect knowledge of the economic structure leading to a changing perception of the inflation-output trade-off by the Fed is one of the candidate explanations. Interesting efforts in this direction have already been undertaken by Cogley and Sargent (2005b), Primiceri (2006), Sargent, Williams, and Zha (2006), and Carboni and Ellison (2009).

Finally, once established that the time-varying inflation target is an important ingredient to describe the U.S. inflation rise and fall in the post-WWII U.S. inflation, the switch from the positive to the normative standpoint appears to be warranted. How should monetary policy be conducted in presence of trend inflation? Interesting first attempts to tackle this issue are Ascari and Ropele (2007) and Ascari and Ropele (2009), but more research in this largely unexplored territory is clearly needed.

References

Ascari, G. (2004): "Staggered Prices and Trend Inflation: Some Nuisances," *Review of Economic Dynamics*, 7, 642–667.

Ascari, G., and T. Ropele (2007): "Optimal Monetary Policy under Low Trend Inflation," *Journal of Monetary Economics*, 54, 2568–2583.

——— (2009): "Trend Inflation, Taylor Principle and Indeterminacy," *Journal of Money, Credit and Banking*, forthcoming.

Benati, L., and P. Surico (2008): "Evolving U.S. Monetary Policy and the Decline of Inflation Predictability," *Journal of the European Economic Association*, 6(2-3), 634–646.

——— (2009): "VAR Analysis and the Great Moderation," *The American Economic Review*, 99(4), 1636–1652.

Bianchi, F. (2009): "Regime Switches, Agents' Beliefs, and Post-World War II U.S. Macroeconomic Dynamics," Duke University, mimeo.

Boivin, J., and M. Giannoni (2006): "Has Monetary Policy Become More Effective?," *The Review of Economics and Statistics*, 88(3), 445–462.

Carboni, G., and M. Ellison (2009): "The Great Inflation and the Greenbook," *Journal of Monetary Economics*, forthcoming.

Castelnuovo, E., and P. Surico (2009): "Monetary Policy Shifts, Inflation Expectations and the Price Puzzle," *Economic Journal*, forthcoming.

Chib, S., and I. Jeliazkov (2001): "Marginal Likelihood from the Metropolis-Hastings Output," *Journal of the American Statistical Association*, 96, 270–281.

Clarida, R., J. Gali, and M. Gertler (2000): "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115, 147–180.

Cogley, T., G. E. Primiceri, and T. Sargent (2009): "Inflation-Gap Persistence in the U.S.," *American Economic Journal: Macroeconomics*, forthcoming.

Cogley, T., and T. Sargent (2005b): "The Conquest of U.S. Inflation: Learning and Robustness to Model Uncertainty," *Review of Economic Dynamics*, 8, 528–563.

Cogley, T., and A. Sbordone (2008): "Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve," *The American Economic Review*, 98(5), 2101–2126.

D'Agostino, A., D. Giannone, and P. Surico (2006): "(Un)Predictability and Macroeconomic Stability," European Central Bank Working Paper Series 605.

Davig, T., and T. Doh (2009): "Monetary Policy Regime Shifts and Inflation Persistence," *Federal Reserve Bank of Kansas City Working Paper*, 08-16, January.

Davig, T., and E. M. Leeper (2009): "Reply to "Generalizing the Taylor Principle: A Comment"," *The American Economic Review*, forthcoming.

Davig, Y., and E. Leeper (2006a): "Fluctuating Macro Policies and the Fiscal Theory," *NBER Macroeconomics Annual*, pp. 247–298.

——— (2007): "Generalizing the Taylor Principle," *The American Economic Review*, 97(3), 607–635.

Doucet, A., N. de Freitas, and N. Gordon (2001): *Sequential Monte Carlo Methods in Practice*, vol. Springer, New York. Springer, New York.

Farmer, R., D. Waggoner, and T. Zha (2008): "Generalizing the Taylor Principle: Comment," *The American Economic Review*, forthcoming.

——— (2009): "Understanding Regime-Switching Rational Expectations Models," *Journal of Economic Theory*, 144, 1849–1867.

Frühwirth-Schnatter, S. (2006): *Finite Mixtures and Markov Switching Models*. Springer, New York.

Hamilton, J. (1989): "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57, 357–384.

Hamilton, J., D. Waggoner, and T. Zha (2007): "Normalization in Econometrics," *Econometric Reviews*, 26, 221–252.

Ireland, P. (2007): "Changes in Federal Reserve's Inflation Target: Causes and Consequences," *Journal of Money, Credit and Banking*, 39(8), 1851–1882.

Justiniano, A., and G. Primiceri (2008): "The Time-Varying Volatility of Macroeconomic Fluctuations," *The American Economic Review*, 98(3), 604–641.

Kass, R., and A. Raftery (1995): "Bayes Factors," *Journal of the American Statistical Association*, 90, 773–795.

Leeper, E. (1991): "Equilibria Under 'Active' and 'Passive' Monetary and Fiscal Policies," *Journal of Monetary Economics*, 27, 129–147.

Liu, Z., D. Waggoner, and T. Zha (2007): "Has the Federal Reserve's Inflation Target Changed?," mimeo.

Lubik, T., and F. Schorfheide (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *The American Economic Review*, 94(1), 190–217.

Lubik, T., and P. Surico (2008): "The Lucas Critique and the Stability of Empirical Models," *Journal of Applied Econometrics*, forthcoming.

Mavroeidis, S. (2009): "Monetary Policy Rules and Macroeconomic Stability: Some New Evidence," *The American Economic Review*, forthcoming.

PRIMICERI, G. (2006): "Why Inflation Rose and Fell: Policymakers' Beliefs and U.S. Postwar Stabilization Policy," *Quarterly Journal of Economics*, 121, 867–901.

SARGENT, T., N. WILLIAMS, AND T. ZHA (2006): "Shocks and Government Beliefs: The Rise and Fall of American Inflation," *The 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., AND T. ZHA (2006): "Were There Regime Switches in U.S. Monetary Policy?," *The American Economic Review*, 96(1), 54–81.

STOCK, J., AND M. WATSON (2007): "Why Has Inflation Become Harder to Forecast?," *Journal of Money, Credit and Banking*, 39(1), 3–33.

TAYLOR, J. (1993): "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.

WOODFORD, M. (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

YUN, T. (2005): "Optimal Monetary Policy with Relative Price Distortions," *The American Economic Review*, 95, 89–108.

ZHA, T. (2010): "Why Do We Need to Go Beyond Gaussianity in Structural Modeling?," 2nd International Conference in Memory of Carlo Giannini, Invited Lecture, Bank of Italy.

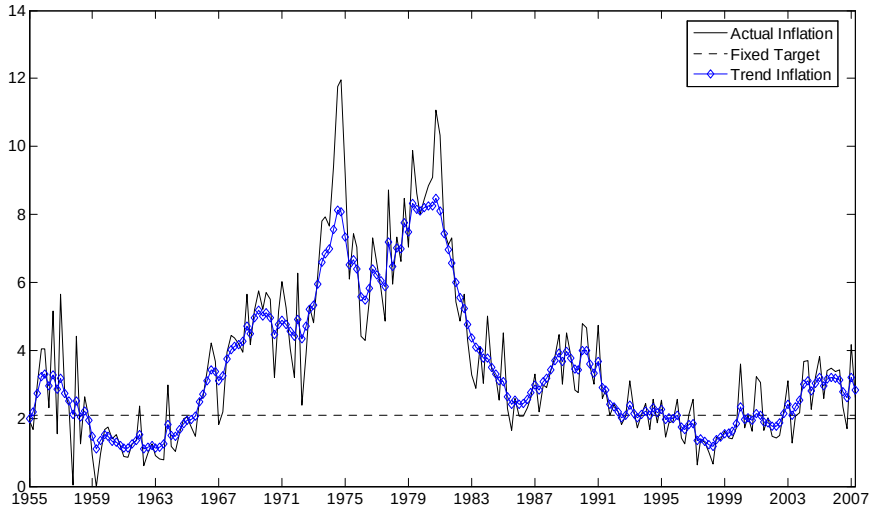

Figure 1: **Raw Inflation Rate and Estimated Trend Inflation.**

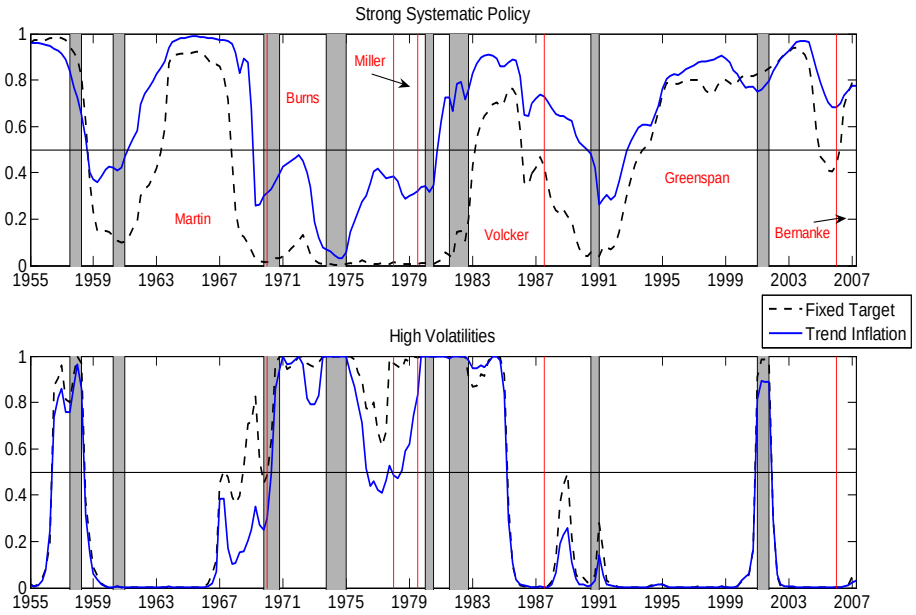

Figure 2: **Fixed Inflation Target vs. Trend Inflation: Estimated Probabilities.** Description of the models provided in the text. Smoothed estimates of the probability of being in the 'active' monetary policy regime (upper panel) and in the 'high' volatility regime (lower panel). Grey vertical bars identify NBER recessions, red vertical bars identify Fed's chairmanships.

Param.	*Prior*	*Fixed Target*	*Time-Var. Target*
$\overline{r}$	$\mathcal{N}(2.0, 1.0)$	3.07 [2.19,3.88]	2.95 [1.99,3.76]
$1+\alpha_{S_1=0}$	$\mathcal{N}(1.7, 0.1)\mathbf{I}_{[1,5]}$	1.69 [1.12,2.30]	1.67 [1.14,2.28]
$1+\alpha_{S_1=1}$	$\mathcal{N}(0.7, 0.1)\mathbf{I}_{[0,1]}$	0.79 [0.47,0.99]	0.35 [0.01,0.88]
$\beta_{S_1=0}$	$\mathcal{N}(0.25, 0.15)$	0.50 [0.05,0.95]	0.57 [0.09,0.96]
$\beta_{S_1=1}$	$\mathcal{N}(0.25, 0.15)$	0.84 [0.46,0.99]	0.86 [0.56,0.99]
$\rho_{S_1=0}$	$\mathcal{N}(0.8, 1.0)$	0.96 [0.92,0.99]	0.94 [0.90,0.97]
$\rho_{S_1=1}$	$\mathcal{N}(0.8, 1.0)$	0.85 [0.77,0.91]	0.74 [0.52,0.88]
$\sigma^2_{\epsilon S_2=0}$	$IG(2.5, 0.75)$	1.76 [1.23,2.554]	1.80 [1.22,2.68]
$\sigma^2_{\epsilon S_2=1}$	$IG(2.5, 0.75)$	0.11 [0.08,0.16]	0.11 [0.07,0.16]
π^{LR}	$\mathcal{N}(2.0, 0.075)$	2.09 [1.78,2.36]	2.03 [1.74,2.33]
ρ_π	$\mathcal{N}(0.9, 0.1)$	–	0.97 [0.92,0.99]
$\sigma^2_{\xi S_2=0}$	$IG(2.5, 0.75)$	–	0.63 [0.20,1.60]
$\sigma^2_{\xi S_2=1}$	$IG(2.5, 0.75)$	–	0.18 [0.09,0.32]
$\phi_{S_1=0}$	$\mathcal{N}(0.4, 0.01)$	0.42 [0.25,0.61]	0.25 [0.07,0.45]
$\phi_{S_1=1}$	$\mathcal{N}(0.6, 0.01)$	0.87 [0.79,0.95]	0.55 [0.33,0.78]
$\sigma^2_{\eta S_2=0}$	$IG(2.5, 0.75)$	2.76 [1.96,3.91]	1.66 [0.71,2.71]
$\sigma^2_{\eta S_2=1}$	$IG(2.5, 0.75)$	0.68 [0.49,0.95]	0.45 [0.27,0.74]
p_{01,S_1}	$Beta(5, 95)$	0.05 [0.02,0.09]	0.05 [0.02,0.09]
p_{01,S_1}	$Beta(5, 95)$	0.05 [0.02,0.08]	0.06 [0.02,0.10]
p_{01,S_2}	$Beta(5, 95)$	0.06 [0.03,0.10]	0.06 [0.03,0.10]
p_{01,S_2}	$Beta(5, 95)$	0.05 [0.02,0.08]	0.04 [0.02,0.07]
$log(ML)$	–	-454.87	-432.62

Table 1: **Estimated Monetary Policy Rules: Fixed vs. Time-Varying Inflation Target.** Figures reported in the Table are posterior means; [5th,95th] percentiles of the simulated posteriors in squared brackets. Moments of the priors generally report the mean and standard deviation of the density in brackets, with the exception of the Beta priors for the switching probabilities, which are defined by their shape parameters. Log Marginal data density computed using the output of the MCMC simulators. The model favored by the data attains the highest marginal data density. Description of the different models: See Figure 1. Details on estimation procedure reported in the text.

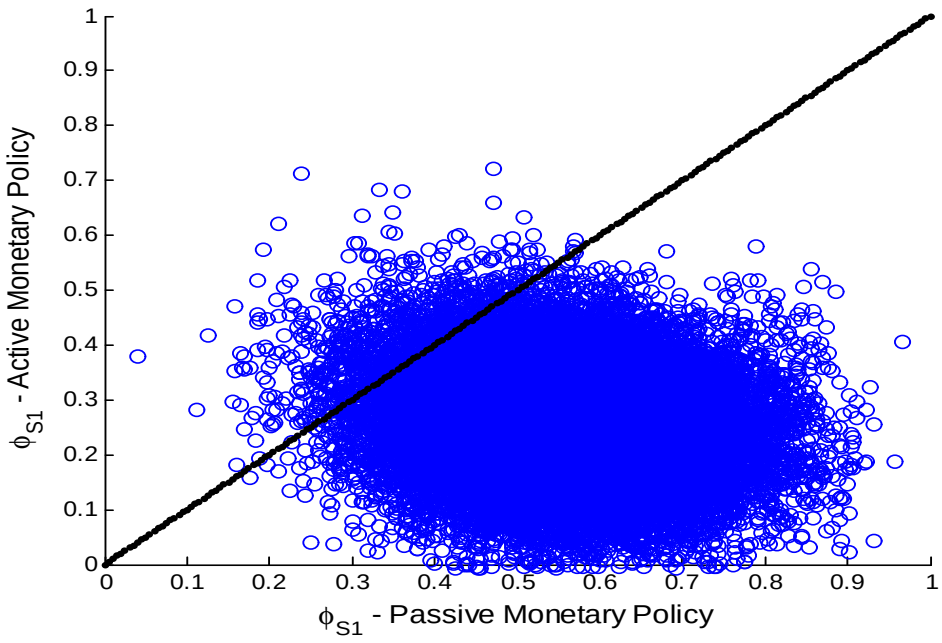

Figure 3: **Inflation Gap Persistence: Scatter Plots.** Figure scatter-plotting realizations of the persistence parameter, x-axis: Passive monetary policy vs. y-axis: Active monetary policy. Pairs constructed by drawing 25,000 realizations from the estimated posterior densities.

	$\phi_{S_1=0} < \phi_{S_1=1} - k, \%$
$k = 0.05$	93.78
$k = 0.10$	89.43
$k = 0.20$	74.06

Table 2: **Inflation Gap Persistence, Joint Density: Probabilities.** Furher explanations on the statistics reported in the table are detailed in the text.

Technical Appendix

1 MCMC methodology

In the following, we describe our inferential solution for the model

$$
\begin{aligned}
i_t &= (1-\rho_{S_1,t})[\bar{r}+\pi_t+\alpha_{S_1,t}z_t+\beta_{S_1,t}y_t]+\rho_{S_1,t}i_{t-1}+\epsilon_{S_2,t} && (1)\\
\pi_t^* &= (1-\rho_\pi)\pi_{LR}+\rho_\pi\pi_{t-1}^*+\xi_{S_2,t} && (2)\\
\pi_t &= \pi_t^*+z_t && (3)\\
z_t &= \phi_{S_1,t}z_{t-1}+\eta_{S_2,t} && (4)\\
\epsilon_{S_2,t} &\sim \mathcal{N}(0,\sigma^2_{\epsilon S_2,t}), \xi_{S_2,t}\sim\mathcal{N}(0,\sigma^2_{\xi S_2,t}), \eta_{S_2,t}\sim\mathcal{N}(0,\sigma^2_{\eta S_2,t}) && (5)
\end{aligned}
$$

In particular $S_{1,t},\ S_{2,t}$ are the unobservable two-state first order Markov chains with transition probability matrix $\Pi = \{p_{k,ij}\}\,,\ k\in\{1,2\}\,,\ i,j\in\{0,1\}$.

The goal of the inferential procedure is to estimate the latent switching regimes processes $\boldsymbol{S}_k = \{S_{k,t},\ t=1,\ldots,T\}$, the unobservable inflation target $\boldsymbol{\pi}^* = \{\pi_t^*,\ t=1,\ldots,T\}$ and the parameters $\boldsymbol{\theta}$. The observed interest rate is denoted by $\boldsymbol{I} = \{i_t,\ t=1,\ldots,T\}$, while $\boldsymbol{X} = (\boldsymbol{\pi},\boldsymbol{y})$, where $\boldsymbol{\pi} = \{\pi_t,\ t=1,\ldots,T\}$ is the inflation rate and $\boldsymbol{y} = \{y_t,\ t=1,\ldots,T\}$ is the output gap.

Since for non-linear latent factor models the likelihood function is not available in closed form, inference has to be based on approximations or

numerical evaluations. Markov Chain Monte Carlo (MCMC) basically avoids the need to directly computing the likelihood function that is expressed in terms of a highly multivariate integral. The basic idea behind MCMC is to build a Markov chain transition kernel starting from a given initial point and with limiting invariant distribution equal to the posterior distribution of the quantities of interest. Under suitable conditions (see Robert and Casella, 1999, ch. 6-7), such a transition kernel converges in distribution to the target posterior density. In practice, the chain is updated to a new position by first simulating from a given proposal distribution, then by eventually accepting the move.

In our framework, given the starting vector $(\boldsymbol{\theta}^{(0)}, \boldsymbol{S}_1^{(0)}, \boldsymbol{S}_2^{(0)}, \boldsymbol{\pi}^{*\ (0)})$ we simulate through MCMC the trajectory of the Markov chain $\{\boldsymbol{\theta}^{(j)}, \boldsymbol{S}^{(j)}, \boldsymbol{S}_2^{(j)}, \boldsymbol{\pi}^{*\ (j)},\ j = 1, \ldots, n\}$ whose draws converge to the posterior distribution $p(\boldsymbol{\theta}, \boldsymbol{S}_1, \boldsymbol{S}_2, \boldsymbol{\pi}^* | \boldsymbol{I}, \boldsymbol{X})$. Once convergence is achieved, inference can be based on the serially dependent sample generated. More precisely, estimates of the latent factors and of the posterior mean of $\boldsymbol{\theta}$ are given respectively by averaging over the realization of the chain, i.e., $\hat{Pr}[S_{k,t} = 1 | \boldsymbol{I}, \boldsymbol{X}] = n^{-1} \sum_{j=1}^{n} S_{k,t}^{(j)}$, $\hat{\pi}_t^* = n^{-1} \sum_{j=1}^{n} \pi_t^{*(j)}$ and $\hat{\boldsymbol{\theta}} = n^{-1} \sum_{j=1}^{n} \boldsymbol{\theta}^{(j)}$. To account for serial correlation in the draws, we estimated the numerical standard error of the sample posterior mean using the approach implemented in Kim, Shephard, and Chib (1998). MCMC for switching regime ARMA models have been introduced in Albert and Chib (1993) and in McCulloch and Tsay (1993) whereas their inferential approach have been successively generalized in Billio, Monfort, and Robert (1999) and in Kim and Nelson (1999). In particular Frühwirth-Schnatter (2001) proposed an MCMC based solution to address for the parameters identifi-

cation problems related to the switching regime models (see also Frühwirth-Schnatter, 2006 for a comprehensive treatment of this topic).

Moving the whole vector $(\boldsymbol{\theta}, \boldsymbol{S}_1, \boldsymbol{S}_2, \boldsymbol{\pi}^*)$ in block can be inefficient, since it is highly multivariate. We implement our MCMC strategy through the Gibbs sampler, an algorithm that efficiently simulates each element or block of $(\boldsymbol{\theta}, \boldsymbol{S}_1, \boldsymbol{S}_2, \boldsymbol{\pi}^*)$ in turn from its full conditional distribution, i.e., the distribution conditioned on the data and the remaining components of the vector. In this case the acceptance probability of moving each sub-block of the vector is 1. As suggested in Shephard (1994) and Carter and Kohn (1994) amongst others, updating the whole latent process $\boldsymbol{S}_k$ in block from its joint distribution given the data and the other parameters should reduce the autocorrelation between states and then speed up the convergence of the chain to its invariant distribution.

To take care of the regime switches, we provide an efficient algorithm based on the multi-move Gibbs sampler proposed in Chib (1996) to update the states $S_{1,t}$ and $S_{2,t},\ t = 1, \ldots, T$, whereas we update the parameters and the latent process π_t^* one component at a time.

To simplify the notation, in the following we call θ_j the generic j-th block[1] of the vector $\boldsymbol{\theta} = \{\theta_j, j = 1, \ldots, J\}$, $\boldsymbol{\theta}_{-j^-} = (\theta_1, \ldots, \theta_{j-1})$ and $\boldsymbol{\theta}_{-j^+} = (\theta_{j+1}, \ldots, \theta_J)$. We summarize the algorithm as follows:

[1]It is worth noting that a block of parameters can be a sub-vector of parameters or a singleton.

MCMC algorithm

- Initialize the chain at $(\boldsymbol{\theta}^{(0)}, \boldsymbol{S}_1^{(0)}, \boldsymbol{S}_2^{(0)}, \boldsymbol{\pi}^{*\,(0)})$
- At step $j = 1, \ldots, n$
  - Update $\boldsymbol{\theta}$ one-at-a-time from the full conditional $p(\boldsymbol{\theta}_i | \boldsymbol{S}_1^{(j-1)}, \boldsymbol{S}_2^{(j-1)}, \boldsymbol{\pi}^{(j-1)}, \boldsymbol{\theta}_{-i^-}^{(j)}, \boldsymbol{\theta}_{-i^+}^{(j-1)}, \boldsymbol{I}, \boldsymbol{X})$ through the Gibbs sampler algorithm;
  - Draw $(S_{1,1}^{(j)}, \ldots, S_{1,T}^{(j)})$ in block from $p(\boldsymbol{S}_1 | \boldsymbol{S}_2^{(j-1)}, \boldsymbol{\theta}^{(j)}, \boldsymbol{\pi}^{(j-1)}, \boldsymbol{I}, \boldsymbol{X})$;
  - Draw $(S_{2,1}^{(j)}, \ldots, S_{2,T}^{(j)})$ in block from $p(\boldsymbol{S}_2 | \boldsymbol{\pi}^{(j-1)}, \boldsymbol{\theta}^{(j)}, \boldsymbol{S}_1^{(j)}, \boldsymbol{I}, \boldsymbol{X})$;
  - Draw $(\pi_1^{*\,(j)}, \ldots, \pi_T^{*\,(j)})$ one-at-a-time from the conditional $p(\pi_t^* | \boldsymbol{S}_1^{(j)}, \boldsymbol{S}_2^{(j)}, \pi_{t-1}^{*\,(j)}, \pi_{t+1}^{*\,(j-1)}, \boldsymbol{\theta}^{(j)}, \boldsymbol{I}, \boldsymbol{X})$, through Gibbs sampler.
- $j = j + 1$

In the next subsections we describe the details of the algorithm.

1.1 Updating the parameters

The use of conjugate priors makes it straightforward to update $\boldsymbol{\theta}$. It is easy to show that the conditional posteriors of $\sigma_{\epsilon,i}^2, \sigma_{\xi,i}^2$ and $\sigma_{\eta,i}^2$ are Inverse Gamma, $p_{k,ij}$ are Beta, whereas all the remaining parameters are Gaussian. Since the full conditional distribution can be simulated directly, each sub-movement of the chain is accepted with probability 1.

The main issue when estimating regime switching models is identification, since the labeling of the states can be permuted without modifying

the marginal likelihood. It is thus common in the literature to impose some constraints on the parameters' space. In particular, we find it useful to identify $S_{2,t}$ by imposing $\sigma^2_{\epsilon S_2=0} > \sigma^2_{\epsilon S_2=1}$ and we implement it in our algorithm through the permutation sampler of Frühwirth-Schnatter (2001). Furthermore, to identify the first regime $S_{1,t}$ we hypothesize:

(i) $0 < \alpha_{S_1=0} < 5$ and $-1 < \alpha_{S_1=1} < 0$;

(ii) $\phi_{S_1=0} \sim \mathcal{N}(0.4, 0.1)$ and $\phi_{S_1=1} \sim \mathcal{N}(0.6, 0.1)$

We notice that (i)-(ii) are informative to identify the two states. If for instance we reverse the constraint, i.e. $0 < \alpha_{S_1=1} < 5$ and $-1 < \alpha_{S_1=0} < 0$ we obtain the exact reverse labeling of the probability of being in S_1.

1.2 Updating $\boldsymbol{S}_k$

We update $\boldsymbol{S}_k = (S_{k,1}, \ldots, S_{k,T})$ in block, moving the two vectors according to the algorithm proposed in Chib (1996). In the following we focus on $\boldsymbol{S}_1$ since the procedure for $\boldsymbol{S}_2$ is in principle the same. To keep the notation concise, we define a generic vector $\boldsymbol{S}_{k,t_1:t_2} = (S_{t_1}, \ldots, S_{t_2})$. The method exploits the following decomposition

$$p(\boldsymbol{S}_1|\boldsymbol{\pi}^*, \boldsymbol{I}, \boldsymbol{X}, \boldsymbol{S}_2, \boldsymbol{\theta}) = p(S_{1,T}|\boldsymbol{\pi}^*, \boldsymbol{I}, \boldsymbol{X}, \boldsymbol{S}_2, \boldsymbol{\theta}) \prod_{t=1}^{T-1} p(S_{1,t}|\boldsymbol{\pi}^*, \boldsymbol{I}, \boldsymbol{X}, \boldsymbol{S}_2, \boldsymbol{S}_{1,t+1:T}, \boldsymbol{\theta}), \tag{6}$$

in which the generic term of the product is

$$p(S_{1,t}|\boldsymbol{\pi}^*, \boldsymbol{I}, \boldsymbol{X}, \boldsymbol{S}_{1,t+1:T}, \boldsymbol{S}_2, \boldsymbol{\theta}) \propto p(S_{1,t}|\boldsymbol{I}_{1:t}, \boldsymbol{\pi}^*_{1:t}, \boldsymbol{X}_{1:t}, \boldsymbol{S}_{2,1:t}, \boldsymbol{\theta}) p(S_{1,t+1}|S_{1,t}, \boldsymbol{\theta}). \tag{7}$$

In particular $p(S_{1,t}|\boldsymbol{I}_{1:t}, \boldsymbol{\pi}^*_{1:t}, \boldsymbol{X}_{1:t}, \boldsymbol{S}_{2,1:t}, \boldsymbol{\theta})$ can be written as

$$\begin{aligned} &p(S_{1,t}|\boldsymbol{I}_{1:t}, \boldsymbol{\pi}^*_{1:t}, \boldsymbol{X}_{1:t}, \boldsymbol{S}_{2,1:t}, \boldsymbol{\theta}) \propto \\ &\quad p(S_{1,t}|\boldsymbol{I}_{1:t-1}, \boldsymbol{\pi}^*_{1:t-1}, \boldsymbol{X}_{1:t-1}, \boldsymbol{S}_{2,1:t-1}, \boldsymbol{\theta}) p(i_t, \pi_t, \pi^*_t, S_{2,t}|\boldsymbol{\pi}^*_{1:t-1}, \boldsymbol{X}_{1:t-1}, \boldsymbol{S}_{2,1:t-1}, \boldsymbol{\theta}) \end{aligned} \tag{8}$$

where

$$\begin{aligned} &p(S_{1,t}|\boldsymbol{I}_{1:t-1}, \boldsymbol{\pi}^*_{1:t-1}, \boldsymbol{X}_{1:t-1}, \boldsymbol{S}_{2,1:t-1}, \boldsymbol{\theta}) = \\ &\quad = \sum_{j=0}^{1} p(S_{1,t}|S_{1,t-1} = j, \boldsymbol{\theta}) p(S_{1,t-1} = j|\boldsymbol{I}_{1:t-1}, \boldsymbol{\pi}^*_{1:t-1}, \boldsymbol{X}_{1:t-1}, \boldsymbol{S}_{2,1:t-1}, \boldsymbol{\theta}) \end{aligned} \tag{9}$$

These two latter distributions can be numerically evaluated in a recursive fashion by setting the distribution of the initial state $S_{1,1}$ as the stationary distribution of the Markov chain, $p(S_{1,1}|\boldsymbol{\theta})$. Once computed all these quantities, $S_{1,T}$ is sampled from $p(S_{1,T}|\boldsymbol{\pi}^*, \boldsymbol{I}, \boldsymbol{X}, \boldsymbol{S}_2, \boldsymbol{\theta})$, that is a Binomial random variable, while the remaining states can be directly simulated from $p(S_{1,t}|\boldsymbol{\pi}^*, \boldsymbol{I}, \boldsymbol{X}, \boldsymbol{S}_2, S_{1,t+1:T}, \boldsymbol{\theta})$, starting from $S_{1,T-1}$ until $S_{1,1}$.

2 Marginal Likelihood computation

In Bayesian statistics it is common practice to use the marginal likelihood to measure the goodness-of-fit. This quantity is defined as

$$\begin{aligned} &m(\boldsymbol{I}, \boldsymbol{X}|\mathcal{M}) = \\ &\quad = \int p(\boldsymbol{I}, \boldsymbol{X}|\boldsymbol{\pi}^*, \boldsymbol{S}_1, \boldsymbol{S}_2, \boldsymbol{\theta}, \mathcal{M}) p(\boldsymbol{\pi}^*, \boldsymbol{S}_1, \boldsymbol{S}_2|\boldsymbol{\theta}, \mathcal{M}) p(\boldsymbol{\theta}|\mathcal{M}) d\boldsymbol{\theta} d\boldsymbol{\pi}^* d\boldsymbol{S}_1 d\boldsymbol{S}_2, \end{aligned} \tag{10}$$

where $(\mathcal{M}, \boldsymbol{\theta})$ indicate a given model $\mathcal{M}$ and its parameters and $p(\boldsymbol{\theta}|\mathcal{M})$ are the prior distributions. In the following we suppress the model index $\mathcal{M}$ for conciseness. Many techniques have been proposed in the literature to evaluate the marginal likelihood. A review on some Monte Carlo alternatives is given in Han and Carlin (2001). In many cases it is useful to recur to the following decomposition

$$m(\boldsymbol{I}, \boldsymbol{X}) = \frac{p(\boldsymbol{I}, \boldsymbol{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{\theta}|\boldsymbol{I}, \boldsymbol{X})}, \tag{11}$$

that is valid for each point $\boldsymbol{\theta}$ on the parameter' space. Since this ratio is not dependent on the specific choice of $\boldsymbol{\theta}$, to compute the marginal likelihood is sufficient to estimate the posterior $p(\boldsymbol{\theta}|\boldsymbol{I}, \boldsymbol{X})$ and the likelihood $p(\boldsymbol{I}, \boldsymbol{X}|\boldsymbol{\theta})$ at a fixed point $\boldsymbol{\theta}^*$.

2.1 Computing the posterior distribution of $\boldsymbol{\theta}$ at $\boldsymbol{\theta}^*$

The first problem is to estimate the posterior $p(\boldsymbol{\theta}^*|\boldsymbol{I}, \boldsymbol{X})$, that is computed through the method proposed in Chib and Jeliazkov (2001) and is based on a sequence of reduced run of the same MCMC algorithm used for the inference. The method consists in dividing the parameters vector $\boldsymbol{\theta}$ in blocks, and at each step of the algorithm associate to the i-th block, a given value θ_i^*. In this way we split the vector into two parts, $\boldsymbol{\theta}^*_{-i^-}$ and $\boldsymbol{\theta}_{-i^+}$. The estimate of the posterior at $\boldsymbol{\theta}^*$ is then given by

$$\hat{p}(\boldsymbol{\theta}^*|\boldsymbol{I}, \boldsymbol{X}) = \prod_{i=1}^{J} \hat{p}(\theta_i^*|\boldsymbol{I}, \boldsymbol{X}, \boldsymbol{\theta}^*_{-i^-}), \tag{12}$$

where each factor $\hat{p}(\theta_i^*|\boldsymbol{I}, \boldsymbol{X}, \boldsymbol{\theta}^*_{-i^-})$ can be computed as

$$\hat{p}(\theta_i^*|\boldsymbol{I}, \boldsymbol{X}, \boldsymbol{\theta}^*_{-i^-}) = \frac{1}{M}\sum_{j=1}^{M} p(\theta_i^*|\boldsymbol{I}, \boldsymbol{X}, \boldsymbol{\pi}^{*(j)}, \boldsymbol{S}_1^{(j)}, \boldsymbol{S}_2^{(j)}, \boldsymbol{\theta}^*_{-i^-}, \boldsymbol{\theta}^{(j)}_{-i^+}) \tag{13}$$

in which $\boldsymbol{\theta}_{-i^+}^{(j)}$, $\boldsymbol{S}_k^{(j)}$ and $\boldsymbol{\pi}^{*(j)}$ are M draws from the same MCMC algorithm used for the inference but with $(\theta_1, \ldots, \theta_i)$ fixed to $(\theta_1^*, \ldots, \theta_i^*)$.

2.2 Computing the likelihood

The second issue is to compute $p(\boldsymbol{I}, \boldsymbol{X}|\boldsymbol{\theta}^*)$. In fact, the likelihood function is not known in closed form due to the presence of latent factors, and has to be evaluated by integrating their effects out. Integration has been performed by a particle filtering procedure (see Doucet, de Freitas, and Gordon, 2001 for a detailed description on this topic). Particle filter algorithms provide a sub-optimal but feasible solution to the Bayesian filtering problem. Consider the general state-space model defined by the density $p(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \boldsymbol{\theta})$ that describes the evolution of the latent states $\boldsymbol{x}_t$ and by $p(w_t|\boldsymbol{x}_t, \boldsymbol{\theta})$ that specifies the observable w_t. Our goal is to estimate the distribution $p(\boldsymbol{x}_{t+1}|\boldsymbol{w}_{1:t+1}, \boldsymbol{\theta})$ given $p(\boldsymbol{x}_t|\boldsymbol{w}_{1:t}, \boldsymbol{\theta})$ in which, as before, $\boldsymbol{w}_{1:t} = (w_1, \ldots, w_t)$ is the past history of the observable process up to time t. We also require the knowledge of the initial distribution $p(\boldsymbol{x}_0|\boldsymbol{\theta})$, of the transition distribution $p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t, \boldsymbol{\theta})$ $t \geq 0$, and of the measurement distribution $p(w_{t+1}|\boldsymbol{x}_{t+1}, \boldsymbol{\theta})$, $t \geq 1$. The key idea is to approximate the filtering density $p(\boldsymbol{x}_{t+1}|\boldsymbol{w}_{1:t+1}, \boldsymbol{\theta})$ by a discrete cloud of points called particles, $\boldsymbol{x}_{t+1}^{(j)}$, $j = 1, \ldots, N$, and a set of weights $\omega_{t+1}^{(j)}$ as follows

$$\hat{p}(\boldsymbol{x}_{t+1}|\boldsymbol{w}_{1:t+1}, \boldsymbol{\theta}) = \sum_{j=1}^{N} \omega_{t+1}^{(j)} \delta(\boldsymbol{x}_{t+1} - \boldsymbol{x}_{t+1}^{(j)}), \tag{14}$$

where $\delta(\cdot)$ is the Dirac delta measure. The cloud of points at time $t+1$ are chosen by using the importance sampling principle, in which the importance

density is $q(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t^{(j)}, w_t)$ and the weights are

$$\omega_{t+1}^{(j)} \propto \omega_t^{(j)} \frac{p(w_{t+1}|\boldsymbol{x}_{t+1}, \boldsymbol{\theta}) p(\boldsymbol{x}_{t+1}^{(j)}|\boldsymbol{x}_t^{(j)}, \boldsymbol{\theta})}{q(\boldsymbol{x}_{t+1}^{(j)}|\boldsymbol{x}_t^{(j)}, y_t)} \qquad j = 1, \ldots, N. \tag{15}$$

In our context $\boldsymbol{x}_t = (\pi_t^*, S_{1,t}, S_{2,t})$, whereas the observable vector at t is $w_t = (i_t, \pi_t, y_t)$. The algorithm can be summarized as follows:

`Particle filter algorithm`

- At time t, given a weighted sample $(\pi_t^*, S_{1,t}, S_{2,t}, \omega_t)^{(j)}$, $j = 1, \ldots, N$ from $p(\pi_t^*, S_{1,t}, S_{2,t}|\boldsymbol{I}_{1:t}, \boldsymbol{X}_{1:t}, \boldsymbol{\theta}^*)$
  - Draw $(\pi_{t+1}^*, S_{1,t+1}, S_{2,t+1})^{(j)}$, $j = 1, \ldots, N$ from the importance density, that in our case is $p(\pi_{t+1}^*, S_{1,t+1}, S_{2,t+1}|\pi_t^{*\,(j)}, S_{1,t}^{(j)}, S_{2,t}^{(j)}, \boldsymbol{\theta}^*)$.
  - Compute $\omega_{t+1}^{(j)}$ using eq. (15).
  - Store $\left(\pi_{t+1}^*, S_{1,t+1}, S_{2,t+1}, \omega_{t+1}\right)^{(j)}$, $j = 1 \ldots, N$ that provide an approximation of $p(\pi_{t+1}^*, S_{1,t+1}, S_{2,t+1}|\boldsymbol{I}_{1:t+1}, \boldsymbol{X}_{1:t+1}, \boldsymbol{\theta}^*)$;
- *t=t+1*

Once the states are filtered, it is easy to evaluate the likelihood by

$$p(\boldsymbol{I}, \boldsymbol{X}|\boldsymbol{\theta}^*) = \prod_{t=1}^{T} \hat{p}(i_t, \pi_t, y_t|\boldsymbol{I}_{1:t-1}, \boldsymbol{X}_{1:t-1}, \boldsymbol{\theta}^*) \tag{16}$$

where $p(i_t, \pi_t, y_t|\boldsymbol{I}_{1:t-1}, \boldsymbol{\pi}_{1:t-1}, \boldsymbol{\theta}^*)$ can be estimated by integrating out the latent processes through a Monte Carlo procedure, for instance by simulating the state $(S_{1,t}, S_{2,t}, \pi_t^*)$ from $p(S_{1,t}, S_{2,t}, \pi_t^*|(S_{1,t-1}, S_{2,t-1}, \pi_{t-1}^*)^{(j)}, \boldsymbol{\theta}^*)$, $j = 1, \ldots, N$ and in which $(S_{1,t-1}, S_{2,t-1}, \pi_{t-1}^*)^{(j)}$ is the outcome of the filtering procedure at time $t-1$.

References

Albert, J. H., and S. Chib (1993): "Bayes Inference via Gibbs Sampling of Autoregressive Time Series Subject to Markov Mean and Variance Shifts," *Journal of Business & Economic Statistics*, 11, 1–15.

Billio, M., A. Monfort, and C. Robert (1999): "Bayesian estimation of Switching ARMA models," *Journal of Econometrics*, 93, 229–255.

Carter, C. K., and R. Kohn (1994): "On Gibbs Sampling for State Space Models," *Biometrika*, 81, 541–553.

Chib, S. (1996): "Calculating posterior distributions and modal estimates in Markov mixture models," *Journal of Econometrics*, 75, 79–97.

Chib, S., and I. Jeliazkov (2001): "Marginal Likelihood from the Metropolis-Hastings Output," *Journal of the American Statistical Association*, 96, 270–281.

Doucet, A., N. de Freitas, and N. Gordon (2001): *Sequential Monte Carlo Methods in Practice*. Springer, New York.

Frühwirth-Schnatter, S. (2001): "Markov Chain Monte Carlo Estimation of Classical and Dynamic Switching and Mixture Models," *Journal of the American Statistical Association*, 96, 194–209.

——— (2006): *Finite Mixture and Markov Switching Models*. Springer, Berlin.

Han, C., and B. Carlin (2001): "MCMC Methods for Computing Bayes Factors: A Comparative Review," *Journal of the American Statistical Association*, 96, 1122–1132.

Kim, C. J., and C. Nelson (1999): *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. MIT Press.

Kim, S., N. Shephard, and S. Chib (1998): "Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models," *Review of Economic Studies*, 65, 361–393.

McCulloch, R., and R. S. Tsay (1993): "Statistical Inference of Macroeconoomic time series via Markov Switching models," *Journal of Time Series Analysis*, 15, 523–539.

Robert, C., and G. Casella (1999): *Monte Carlo Statistical Methods*. Springer, Berlin.

Shephard, N. (1994): "Partial Non-Gaussian State Space," *Biometrika*, 81, 115–131.