LEARNING AND TIME-VARYING MACROECONOMIC VOLATILITY

FABIO MILANI

University of California, Irvine

Abstract. This paper presents a DSGE model in which agents' learning about the economy can endogenously generate time-varying macroeconomic volatility. Economic agents use simple models to form expectations and need to learn the relevant parameters. Their gain coefficient is endogenous and is adjusted according to past forecast errors.

The model is estimated using likelihood-based Bayesian methods. The endogenous gain is jointly estimated with the structural parameters of the system.

The estimation results show that private agents appear to have often switched to constant-gain learning, with a high constant gain, during most of the 1970s and until the early 1980s, while reverting to a decreasing gain later on. As a result, the model can generate a pattern of volatility, which is increasing in the 1970s and falling in the second half of the sample, with a decline that can roughly match the magnitude of the so-called "Great Moderation" in the 1984-2007 period. The paper also documents how a failure to incorporate learning into the estimation may lead econometricians to spuriously find time-varying volatility in the exogenous shocks, even when these have constant variance by construction.

Keywords: adaptive learning, constant gain, monetary policy, macroeconomic volatility, inflation dynamics.

JEL classification: C11, D84, E30, E50, E52, E58, E66.

I am grateful to two anonymous referees for their feedback, which definitely improved the quality of the paper. I would also like to thank Bill Branch, George Evans, Oscar Jorda, and especially discussants of the paper Jim Bullard and Andy Levin, participants at the 5th International Research Forum at the ECB, Frankfurt, the 'Learning Week' Workshop in St. Louis, MO, the Missouri Economics Conference at the University of Missouri, Columbia, the 8th Annual Bank of Finland/CEPR Conference on "Expectations and Business Cycle Dynamics", and seminar participants at the University of California, Davis, the Federal Reserve of St. Louis, and the University of Southern California, for comments and suggestions.

Address for correspondence: Department of Economics, 3151 Social Science Plaza, University of California, Irvine, CA 92697-5100. *Phone:* 949-824-4519. *Fax:* 949-824-2182. *E-mail:* fmilani@uci.edu. *Homepage:* http://www.socsci.uci.edu/~fmilani.

1. Introduction

1.1. **TV Macroeconomic Volatility.** Several studies have documented large changes in the volatility of macroeconomic fluctuations in the U.S. over the post-war period. Kim and Nelson (1999), McConnell and Pérez-Quiròs (2000), Blanchard and Simon (2001), and Stock and Watson (2002), among several others, have identified a large decline of output growth volatility in the years post-1984 and before the 2007 financial crisis, compared to the previous two decades (the large shift in volatility is commonly referred to as "The Great Moderation"). The reduction in volatility is apparent if one looks at simple measures as the variances of output growth and inflation in the 1950-1980 versus the 1980-2007 samples. Slightly more sophisticated approaches yield a similar message: Figure 1, for example, shows the conditional standard deviations from GARCH models for inflation and output gap over time. The conditional standard deviations for both series increase in the 1970s and substantially decline after the early 1980s.

Correctly modeling changes in volatility has been shown to be important for understanding macroeconomic fluctuations. Sims and Zha (2006) find that incorporating regime changes in the volatilities of disturbances in a Bayesian VAR overturns the evidence of large regime switches in US monetary policy. Primiceri (2005), instead, estimates a VAR in which he allows for a continuously changing variance-covariance matrix: he similarly concludes that the role played by the falling volatility of exogenous shocks seems more important than monetary policy changes in explaining the recent behavior of US inflation and unemployment.

With few exceptions, however, estimated DSGE models still habitually assume that the shocks have maintained constant variance throughout the whole sample (e.g., Smets and Wouters 2003, 2007, Lubik and Schorfheide 2004, and An and Schorfheide 2007). The papers by Justiniano and Primiceri (2008) and Fernandez-Villaverde and Rubio-Ramirez (2007) were the first to relax this assumption. Both papers incorporate stochastic volatility in optimizing DSGE models. They find that the volatilities of the shocks have significantly changed over time and that accounting for those variations is important to improve the models' fit to the data.

The existence of time-varying volatility in the economy, therefore, can be now considered an empirical regularity. But what drives the changes in the volatility of macroeconomic fluctuations?

In Justiniano and Primiceri (2008) and Fernandez-Villaverde and Rubio-Ramirez (2007), the changes in volatility are modeled as *exogenous*. But if these are an important feature of the economy as they appear to be, it becomes crucial to try to understand their potential causes.

1.2. **Paper's Contribution.** This paper takes a step in this direction by presenting a model in which stochastic volatility arises endogenously in the economy. I present a stylized New-Keynesian DSGE model in which agents' learning about the economy has implications for macroeconomic volatility. Economic agents use simple models to form expectations and need to learn the relevant model parameters over time.[1] Their learning speed is endogenous and depends on previous forecast errors. When the forecast errors are large, agents become concerned that the economy may be experiencing a structural break and, therefore, they start assigning a larger weight to new information. When the forecast errors are, instead, relatively modest, economic agents remain confident about their model and turn less responsive to new information. The endogenous time-varying learning speed has implications for the volatility of the macroeconomic variables that agents are trying to learn about. In this way, agents' learning with an endogenous gain can generate stochastic volatility in the economy.

[1]See Evans and Honkapohja (2001) for a treatment of several models with adaptive learning.

The learning rule with an endogenously switching gain is in the same spirit as the rule assumed by Marcet and Nicolini (2003), who use a similar mechanism to study hyperinflations. Here, however, the gain is not fixed at a particular value, but estimated from time series data.

1.3. **Related Literature.** This paper is related to the recent works by Branch and Evans (2007) and Lansing (2009), both of which present frameworks in which changes in volatility can arise endogenously. In Branch and Evans (2007), it is model uncertainty that plays the key role role in generating time-varying volatility in a Lucas-type monetary model.[2] Lansing (2009), instead, presents a New Keynesian Phillips curve with boundedly-rational expectations, which can give rise to time-varying persistence and volatility. In a single-equation setting for inflation, he can derive the optimal variable gain as the fixed point of a nonlinear map that relates the gain to the autocorrelation of inflation changes. Cho and Kasa (2012) allow agents to test the specification of their models using econometric tests and study the dynamics induced by the process of learning with model validation. They find, for example, that the combination of a variable gain (a constant gain that is allowed to increase during turbulent times as in response to the oil price shocks in the 1970s) with model switching can produce changes in the volatility of target inflation. Other papers have taken steps to endogenize the gain (Evans and Ramey, 2006, and, more recently, Kostyshyna, 2012, and Gaus, 2011).

Although this paper proposes an endogenous explanation for changes in macroeconomic volatility, its main focus is not on investigating the sources

[2]This paper and the Branch and Evans' approaches should be seen as complementary. A more realistic model, in fact, would possibly include agents that endogenously adjust their gain in response to the previous forecast errors, but that, at the same time, consider different models and switch among them as the performance of one of them becomes superior. This is, however, left for future research. Moreover, learning as in this paper might be seen as a crude way to model economic agents who are concerned about potential changes in the model of the economy, but without having to specify the different possible models or the number of regimes.

of the Great Moderation. Points of contact, however, clearly exist with the Great Moderation literature. Within that vast literature, this paper is more closely connected to the work by Bullard and Singh (2012), who find that agents' learning about different technology regimes (high and low, corresponding to 'expansion' and 'recession' states) may have played a role for the "Great Moderation": they conclude that 30% of the decline in variance may be due to learning, rather than pure good luck. The current paper takes a very different approach, but it shares the emphasis on the evolution of agents' learning as a potential source of moderation in economic volatility.

Finally, the paper is more broadly related to the extensive literature on adaptive learning in monetary policy models (e.g., Evans and Honkapohja 2001) and, in particular, to the papers that exploit learning to explain features of macroeconomic data, such as persistence and volatility (Orphanides and Williams 2003, 2005a,b, 2006, 2007, Adam 2005, and Milani 2007, 2011).

1.4. **Main Results.** The simulation results show that time variation in the gain can potentially generate substantial time-varying volatility in the inflation and output gap series. The model is then taken to the data to judge whether changes in the learning process may have been a contributor to the evolution of macroeconomic volatility in the US. The Bayesian approach used in the paper facilitates the joint estimation of the learning gain coefficients together with the structural parameters in the economy. The estimation reveals that the endogenous gain appears to have switched to large constant gain values for most of the 1970s and early 1980s as a consequence of larger forecast errors by private agents in those periods. In the latest two decades, instead, the agents have switched to a decreasing gain. The estimated gain values in the 1970s are large and can justify a sizeable increase in volatility in the period. Simulation of the model, in fact, with the parameters fixed at the posterior mean estimates, implies that under the estimated evolving gain, the economy would observe higher volatility in the 1960s-1970s than later on. The magnitude of the model-implied decline in volatility roughly matches the size of the Great Moderation.

Moreover, the paper shows that even if the economy was subject to structural shocks with constant-variance over the whole sample, a failure to incorporate agents' learning in the estimation would lead econometricians to spuriously find the existence of ARCH/GARCH effects in the model innovations. The paper finally discusses how the evidence of time-varying volatility in the innovations, as they are measured by the econometrician, may itself be the result of monetary policy and, mainly, of the interaction between policy and agents' learning, and not just a matter of luck.

2. The Model

The economy is described by the following New-Keynesian model[3]

$$\pi_t = \beta \widehat{E}_t \pi_{t+1} + \kappa x_t + u_t \tag{2.1}$$

$$x_t = \widehat{E}_t x_{t+1} - \sigma(i_t - \widehat{E}_t \pi_{t+1}) + g_t \tag{2.2}$$

$$i_t = \rho_t i_{t-1} + (1-\rho_t)(\chi_{\pi,t}\pi_{t-1} + \chi_{x,t} x_{t-1}) + \varepsilon_t \tag{2.3}$$

where π_t denotes inflation, x_t the output gap, and i_t the nominal interest rate; u_t, g_t, and ε_t denote supply, demand, and monetary policy shocks. Equation (2.1) represents the forward-looking New Keynesian Phillips curve that can be derived from the optimizing behavior of monopolistically competitive firms under Calvo price setting or quadratic adjustment costs in nominal prices. Inflation depends on expected inflation in $t+1$ and on current output gap. The parameter $0 < \beta < 1$ represents the households' discount factor, while κ denotes the slope of the Phillips curve and is an inverse function of the Calvo price stickiness parameter. Equation (2.2) represents the log-linearized intertemporal Euler equation that derives from the households' optimal choice of consumption. The output gap depends on the expected one-period ahead output gap and on the ex-ante real interest

[3]See Woodford (2003) for a standard derivation.

rate. The coefficient $\sigma > 0$ represents the intertemporal elasticity of substitution in consumption. Equation (2.3) describes monetary policy. The central bank follows a Taylor-type rule by adjusting its policy instrument, a short-term nominal interest rate, in response to deviations in inflation and the output gap. In light of McCallum's argument that only information up to $t-1$ might be available in real-time, I assume that the central bank cannot respond to contemporaneous variables, but it responds only to lagged variables (i.e., the Taylor rule is "operational" using McCallum's terminology). The policy coefficients are allowed to vary over time and, in particular, they differ between the pre- and post-1979 samples:[4]

$$\rho_t = \begin{cases} \rho_{pre-79} & t < 1979:Q3 \\ \rho_{post-79} & t \geq 1979:Q3 \end{cases}$$

$$\chi_{\pi,t} = \begin{cases} \chi_{\pi,pre-79} & t < 1979:Q3 \\ \chi_{\pi,post-79} & t \geq 1979:Q3 \end{cases}, \chi_{x,t} = \begin{cases} \chi_{x,pre-79} & t < 1979:Q3 \\ \chi_{x,post-79} & t \geq 1979:Q3 \end{cases}.$$

The supply shock u_t may arise endogenously in the model by assuming a time-varying elasticity of substitution among differentiated goods, whereas g_t derives from shocks to preferences, technology, or government spending, for example (both shocks are assumed to evolve as AR(1) processes).

The majority of the papers that focus on the estimation of DSGE models assumes that such shocks maintain constant variance over the whole sample (e.g. Smets and Wouters 2007, An and Schorfheide 2007). But recent papers have suggested that the changing volatilities of these shocks can be important to understand macroeconomic fluctuations.

Typical state-of-the-art DSGE models cannot endogenously generate time-varying stochastic volatility, but they need to assume stochastic volatility as exogenous (Justiniano and Primiceri, 2008 and Fernandez-Villaverde and Rubio-Ramirez, 2007). This paper contributes to the literature by showing

[4]I assume that a regime switch in policy occurs in 1979, when Paul Volcker begins his term as Chairman of the Federal Reserve (August 1979). Duffy and Engle-Warnick (2005), using nonparametric methods, similarly identify a switch in policy exactly in the third quarter of 1979. Allowing for unknown changes in policy is beyond the scope of this paper.

how stochastic volatility can arise *endogenously* from agents' learning about the economy.

I relax here the assumption of rational expectations. I assume that agents use a linear economic model to form their expectations. The agents do not know the model coefficients and need to learn them over time. Therefore, $\widehat{E}_t$ refers to subjective expectations and may differ from the mathematical expectations operator conditional on the true model of the economy E_t.[5]

2.1. **Learning.** Economic agents need to form expectations about the future aggregate inflation rate $\widehat{E}_t\pi_{t+1}$ and the future output gap $\widehat{E}_t x_{t+1}$ to solve their optimal consumption and price-setting decisions. I assume that agents use a perceived linear model of the economy and that they need to learn the relevant reduced-form coefficients. As in Evans and Honkapohja (2001), they behave as econometricians by estimating the model and updating their estimates as new data become available. In the benchmark case, agents use the following 'Perceived Law of Motion' (PLM):

$$Z_t = a_t + b_t Z_{t-1} + \eta_t \tag{2.4}$$

where $Z_t \equiv [\pi_t, x_t, i_t]'$, and where a_t and b_t are coefficient vectors and matrices of appropriate dimensions. The agents' PLM, therefore, is a simple VAR(1) in the endogenous variables π_t, x_t, and i_t. Notice that although the true constants in the model equal zero, agents are not endowed with this information. In this way, they also need to learn the steady-state of the variables. The PLM is, therefore, similar to the Minimum State Variable solution of the system under rational expectations. Agents, however, do not know the reduced-form model parameters and they cannot observe the exogenous shocks.[6] For each equation in the PLM, agents learn the model

[5] I have assumed a simple small-scale New-Keynesian model without adding "mechanical" sources of persistence as habit formation in consumption or inflation indexation; Milani (2006, 2007), in fact, shows that, under learning, those may become redundant as learning is successful in inducing persistence in the model.

[6] That is, I assume for now that agents estimate VARs in the endogenous variables, rather than VARMAs, as this is a more common practice in econometrics. It seems more

coefficients according to the following updating rules

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + g_{t,y} R_{t-1}^{-1} X_t (Z_t - \widehat{\phi}_{t-1}' X_t)' \tag{2.5}$$

$$R_t = R_{t-1} + g_{t,y} (X_t X_t' - R_{t-1}) \tag{2.6}$$

where $\widehat{\phi}_t = (a_t, b_t)'$ collects the coefficients, and $X_t \equiv \left\{1, Z_{t-1}'\right\}'$ is an appropriately stacked matrix of regressors. The first line describes the updating of the learning rule coefficients, whereas the second describes the updating of the matrix of second moments of the regressors, denoted by R_t. The coefficient $g_{t,y}$ denotes the gain, which in the paper will be endogenously determined and time-varying. I allow agents to learn about inflation, output, and interest rates at different rates, letting the gain $g_{t,y}$ differ for $y = \pi_t$, x_t, i_t (as Branch and Evans, 2006, discuss, in fact, if the degree of structural change can be expected to differ across series, the optimal gains should also differ).[7]

The gain endogenously adjusts according to past forecast errors as follows

$$g_{t,y} = \begin{cases} \left(g_{t-1,y}^{-1} + 1\right)^{-1} & \text{if } \frac{\sum_{j=0}^{J}(|y_{t-j} - E_{t-j-1} y_{t-j}|)}{J} < \upsilon_t^y \\ \overline{g}_y & \text{if } \frac{\sum_{j=0}^{J}(|y_{t-j} - E_{t-j-1} y_{t-j}|)}{J} \geq \upsilon_t^y, \end{cases} \tag{2.7}$$

where $y = \pi_t,\ x_t,\ i_t$. When the average of the past forecast errors (in absolute value) is below a certain threshold υ_t^y, the agents use a decreasing gain. The decreasing gain expression is equivalent to t^{-1}, with the only exception that, when agents revert from constant gain to decreasing gain, they do not restart with a fully new sample with $t = 1$, but they restart from the previous constant gain value $\overline{g}_y$ (the same assumption is made in Marcet

---

realistic to assume that agents do not observe the shocks; the results in the paper, however, do not hinge on this assumption. One can for example estimate the higher-order VAR(2) to approximate the VARMA(1,1). The paper's results remain in line with those obtained for the VAR(1) case. In section 4, however, I will also re-estimate the model under the assumption that agents fully observe the disturbances in their PLM.

[7] As more common in the literature, agents are learning as classical econometricians (constant-gain learning corresponds to a particular case of weighted-least squares). An alternative would be to assume that agents act as Bayesian econometricians, who update their priors in light of new sample observations. The latter case has been studied in Bullard and Suda (2011), who show how extra terms may appear in the economy's law of motion.

and Nicolini, 2003). When the average of previous forecast errors is above the threshold υ_t, instead, the agents become concerned that the economy may be experiencing a structural break. In the proximity of a structural break, a decreasing gain would be inefficient: the agents therefore switch to a constant gain, which allows them to better track the break by assigning a larger weight to new information. When the forecast errors fall again below the threshold, agents switch back to a decreasing gain, which is reset to the value it had in the previous period.

The endogenous switching gain is in the spirit of the gain assumed by Marcet and Nicolini (2003).[8] I assume that the threshold υ_t^y is given by the mean absolute deviation of historical forecast errors, which is recursively updated.[9] Notice that the degrees of freedom from this mechanism are the gain coefficients $\overline{g}_y$, the window length J for past forecast errors, as well as υ_t^y. The gain will be estimated from the data, whereas J will be initially fixed (later in the paper I will also treat J as a parameter and estimate its value, in addition to performing other robustness checks).

I assume that economic agents dispose of information only up to $t-1$ when forming expectations for next period (this is a common assumption in the adaptive learning literature). Therefore, economic agents use (2.4) and the updated parameter estimates in (2.5) and (2.6) to form their expectations for $t+1$ as

$$\widehat{E}_{t-1}Z_{t+1} = a_{t-1}(I + b_{t-1}) + b_{t-1}^2 Z_{t-1}, \tag{2.8}$$

[8] Although agents' learning with the described endogenous gain is by no means optimal, it can be expected to provide a fairly good approximation to the optimal forecasting behavior of agents who are concerned about possible unknown breaks in the economy, but who do not want to take a stand on the nature or timing of the breaks, or on the existence or number of regimes, and assuming that the agents, in their loss function, are much more concerned about very large forecast errors than relatively small ones.

[9] I have tried estimations in which the agents were assumed to compare past root mean squared forecast errors with the standard deviation of previous forecast errors, rather than using forecast errors in absolute value and the mean absolute deviation, and the results were similar. Other robustness checks regarding the threshold choices are discussed in Section 4.

where I denotes a 3×3 identity matrix. The resulting expectations can be substituted in the original system (2.1)-(2.3) to obtain the Actual Law of Motion of the Economy (ALM). The ALM can be written as:

$$\Upsilon_t = A_{t-1} + B_{t-1}\Upsilon_{t-1} + \epsilon_t \tag{2.9}$$

where $\Upsilon_t = [Z_t', w_t']'$, $w_t = [u_t, g_t]'$, and where ϵ_t contains the innovation components of the supply, demand, and monetary policy disturbances. The parameter vector A_{t-1} and matrix B_{t-1} are time-varying, and their elements are functions of both the structural coefficients in (2.1)-(2.3) and of the learning beliefs entering through (2.8). The ALM can be paired with the following set of observation equations

$$Observables_t = H\Upsilon_t, \tag{2.10}$$

to yield a state-space system that is ready for estimation. The vector of observables will include measures of inflation, the output gap, and interest rates; the matrix H selects the corresponding variables from the state vector Υ_t. The likelihood of the system (2.9)-(2.10) can be computed using the Kalman filter. At each iteration t, $t = 1, ..., T$, of the Kalman filter, the coefficient matrices are updated through the learning algorithms (2.5)-(2.6).

The literature on adaptive learning has often focused on studying the convergence of economic systems with less-than-fully-rational beliefs to the same equilibrium that would be reached under rational expectations. In our New Keynesian economy, if we restrict agents to use the correctly-specified MSV solution as their PLM, and learn using a decreasing gain, the conditions for determinacy and E-stability of the system would be standard (a version of the "Taylor principle"). Checking E-Stability in an economy with an endogenous switching gain, as the one assumed here, is a more complicated issue, which has already been thoroughly investigated in Gaus (2013). In the baseline case in this paper, however, agents are assumed to lack knowledge of the exogenous disturbances in their PLMs; therefore, the system with learning will not converge to the rational expectations equilibrium, or to a distribution around it in the case of a constant gain, but it may converge to a Restricted Perception Equilibrium (e.g., Branch, 2006).

The estimation of the ALM (2.9), however, can be performed regardless of whether the beliefs lead to convergence to a restricted-perception equilibrium or not. Later, the model will be estimated under the assumption that agents use a correctly-specified MSV solution as their PLM, i.e. one including the exogenous disturbances as regressors, and restricting the parameter space to guarantee equilibrium determinacy and E-stability. The implied evolution of the time-varying gain and the main conclusions of the paper will remain similar in those cases.

3. Endogenous Gain and Endogenous Time-Varying Volatility

The value of the gain coefficient affects the volatility in the economy. In a simple empirical model of inflation and unemployment dynamics, Orphanides and Williams (2005, 2007), for example, have shown that the volatility of those variables is a positive function of the gain (they consider only gain values between 0.01 and 0.04). I simulate the model (for now with constant, exogenously set, gain values) to show that this is also the case here.[10] Figure 2 makes clear that the standard deviations of inflation and the output gap would increase as a function of the constant gain value.

Changes in the gain over time, therefore, may potentially be an important determinant of the observed movements in macroeconomic volatility. To show the potential role of the gain, I turn now to the simulation of the model under an endogenous gain, which is allowed to switch as described in (2.7). Therefore, agents adopt a decreasing gain as long as their forecast errors are 'small'. They switch to a constant gain when those become larger

[10] I fix the following values for the parameters: $\beta = 0.99$, $\kappa = 0.05$, $\sigma = 0.1$, $\rho = 0.95$, $\chi_\pi = 1.5$, $\chi_x = 0.5$, $\rho_u = 0.9$, $\rho_g = 0.9$. Agents use the MSV solution of the system to form expectations. The economy is simulated for 1,000 periods using a grid of constant gain values from 0 to 0.15.

and above ν_t, the mean absolute deviation of past forecast errors. In the simulation, I assume $\overline{g}_\pi = \overline{g}_x = 0.15$. The choice of such high gain values is, for the moment, purely for illustrative purposes and it is meant to make the effects more striking in the graph (the value will be estimated later in the paper).[11] Figure 3 shows the time-varying endogenous gain together with the rolling standard deviations of inflation and output gap, obtained from a typical simulation. As the gain changes over time, the degree of volatility in the economy also experiences large shifts. The figure displays sizable time variation in volatility and various episodes characterized by volatility clustering, although the exogenous shocks had constant variance by construction. The persistence of the volatility series and the duration of the clusters obviously depend on the assumed window that agents use to compute past forecast errors; by varying the window size one could in principle mimic a wide range of changing volatility series.[12] For example, decreasing J to 500 would imply more frequent changes in volatility and shorter clusters, as shown in the upper panels of Figure 4. The lower panels, instead, plot the case when agents only adopt a constant gain (fixed at the lower value of 0.05).

The next section will take the model to the data. The estimation aims to infer the evolution of the endogenous gain from time series observations. The simulation can then be repeated in an artificial economy in which the learning process is calibrated to resemble the one estimated from U.S. data.

[11] I simulate the economy for 13,000 periods, allowing agents to use a window of 3,000 observations when computing the mean of past forecast errors, and discarding the first 3,000 periods. The large number of observations is again meant to make the time-varying volatility more apparent in the graph. The parameters are: $\beta = 0.99$, $\kappa = 0.05$, $\sigma = 0.1$, $\rho = 0.95$, $\chi_\pi = 1.5$, $\chi_x = 0.5$, $\rho_u = 0.5$, $\rho_g = 0.5$.

[12] Allowing the gain to change in a more 'continuous' fashion, rather than abruptly jumping from t^{-1} to $\overline{g}$ would imply more gradual movements in the volatility series. The case of a gradually changing gain, possibly along the lines proposed by Colucci and Valori (2004, 2005), is left for future research.

4. Bayesian Estimation

I estimate the model using likelihood-based Bayesian methods. The estimation follows Milani (2007), who extends the techniques reviewed in An and Schorfheide (2007) to allow for near-rational expectations and learning. The vector Θ collects the structural parameters of the model:

$$\Theta = \{\beta, \kappa, \sigma, \rho_t, \chi_{\pi,t}, \chi_{x,t}, \rho_u, \rho_g, \sigma_u, \sigma_g, \sigma_\varepsilon, \overline{g}_\pi, \overline{g}_x, \overline{g}_i\} \tag{4.1}$$

Differently from Milani (2007), the gain coefficient is now endogenous, being allowed to vary over time depending on the magnitude of the past forecast errors that agents make (as made clear by expression 2.7). The gain switches from decreasing (equal to t^{-1}) in 'stable' times to constant ($\overline{g}_y$), when past forecast errors become large and hence suggestive that a break might be occurring. The constant gain coefficient to which agents switch is not fixed to an ad-hoc value, rather its value is jointly estimated with the rest of the model parameters. I use the Metropolis-Hastings algorithm to generate 200,000 draws from the posterior distribution.[13] As discussed, the likelihood of the system (2.9)-(2.10) is evaluated after each MH draw using the Kalman Filter.[14] I use quarterly US data for the 1960:I-2006:I sample in the estimation to fit the series for inflation, output gap, and nominal interest rates.[15] Data from the pre-sample period 1954:III-1959:I were, instead, used to initialize the learning algorithm.

4.1. **Priors.** The priors for the model coefficients are reported in Table 1. Most prior choices follow Milani (2007). To minimize the influence of

[13] I discard a burn-in of 40,000 draws. See appendix in Milani (2007) for more details on the estimation.

[14] Since stochastic volatility arises endogenously from the adjustment of expectations in the model and it is not assumed, instead, in the exogenous shocks, the estimation can be performed using the Kalman Filter rather than the more computationally-intensive particle filter employed in Fernandez-Villaverde and Rubio-Ramirez (2007).

[15] Inflation is defined as the annualized quarterly rate of change of the GDP Implicit Price Deflator, output gap as the log difference between GDP and Potential GDP (Congressional Budget Office estimate), and the federal funds rate is the measure for the nominal interest rate. The series are obtained from FRED, the Federal Reserve Bank of Saint Louis economic database.

the priors on the main parameters of interest, I assume a Uniform prior distribution in the $[0, 0.3]$ interval for the constant gain coefficients. I assume a dogmatic prior for β, which is fixed at 0.99, a Gamma prior for σ and κ, a Beta prior for the autoregressive coefficients, and Normal prior distributions for the feedback coefficients to inflation and output gap in the policy rule. I assume for now $J = 4$, i.e. agents care about forecast errors over the previous year (this restriction will be later relaxed) when deciding how much weight to assign to more recent information. I will point out in describing the results the situations in which the priors have important effects on the estimates.

4.2. **Empirical Results.** Figure 5 shows the evolution of the forecast errors (in absolute value) about inflation, output gap, and the federal funds rate over the sample under the estimated learning rules. Inflation and output were typically harder to predict during the 1970s and until the early 1980s. The forecast errors for both inflation and output gap were on average lower in the 1990s. Monetary policy, instead, was harder to forecast in the late 1960s, in most of the 1970s, and during Volcker's disinflation. Figure 6 shows the episodes in which the rolling means of the absolute forecast errors exceed the updated values of ν_t^y, which imply switches to learning with a constant gain.

Table 2 presents the parameter estimates, along with 95% credible sets. The corresponding prior and posterior distributions for each parameter are shown in Figure 7. The value of the constant gain to which private agents switch when their forecast errors jump above threshold is estimated equal to 0.082 for inflation and to 0.073 for output (a very low gain coefficient is, instead, found for the interest rate equation). Those values are substantially larger than the estimates in Milani (2007), but of course here they refer only to particular periods in the sample.[16] It appears, therefore, that agents, on average, adopt low gain coefficients, but they switch to considerably higher gains in periods of instability. Figure 8 plots the evolution of the time-varying gain coefficients estimated for inflation and the output gap.[17] The learning process for inflation often switches to a constant gain in the 1970s until the early 1980s and it reverts to a decreasing gain shortly after 1985 and for most of the latest part of the sample. Learning about the output gap is also characterized by frequent switches to a constant gain from the 1960s until 1985, and by a decreasing gain for most of the recent two decades (only two switches are identified from 1985 to 2006).

Turning to the other parameters, I estimate $\sigma^{-1} = 5.92$ and $\kappa = 0.022$. The posterior means for the monetary policy rule coefficients indicate a more aggressive response to inflation and a less active response towards the output gap in the second part of the sample than in the first (χ_π goes from 1.37 to 1.53, and χ_x declines from 0.58 to 0.48). The estimated monetary policy rules would, therefore, satisfy the Taylor principle in both sub-samples (a similar result in a model with learning is found in Milani 2006). However, as shown in Figure 7 and quite typical in estimated DSGE models (e.g., Smets and Wouters, 2007), the posterior distributions for the inflation feedback parameters in the policy rule are not far from the assumed prior distributions, suggesting that the data may not be informative enough along these dimensions to overturn relatively tight priors (the sensitivity to uninformative, uniform, priors will be analyzed in the next section). The data appear informative, instead, about the values of the gain coefficients. Although

[16]The larger gain for inflation than output is consistent with results in Branch and Evans (2007) and Milani (2006, 2007).

[17]I focus in this paper on inflation and output gap. I do not try, instead, to explain the time-varying volatility in the Taylor rule equation with learning. The estimated higher volatility of monetary policy shocks in some sub-periods can be more realistically attributed to misspecification of the Taylor rule in the 1979-1982 years and in few other episodes in the 1970s than to a time-varying gain story.

Uniform priors were assumed, the estimation has no trouble identifying the gains that imply the best fit of the data.

4.3. **Robustness.**

4.3.1. *Different Windows.* The results might depend on the particular choice of the number of observations J (in expression 2.7) that agents are assumed to use when computing the average of recent forecast errors. To check for robustness, I repeat the estimation using a longer window, i.e. $J = 20$ (now agents compute the mean absolute forecast errors over the past five years). The estimates are reported in Table 3a. Figure 9 shows the evolution of the endogenous gain coefficients in this case. The estimated gains equal 0.065 for inflation and 0.064 for output gap. The other estimates are not substantially different.

The window for the mean forecast errors can also be interpreted as a parameter that can be estimated from data. Table 3a reports the results when the estimation is repeated treating J as a free parameter. A gamma distribution with mean 12 and standard deviation 4.9 is assumed as prior for J. The estimated posterior mean is 4 (in the estimation, J is rounded to the closest integer, since agents need to use the previous J periods as described in 2.7), implying that agents care about forecast errors over the previous year (the time-varying gains are therefore similar to those shown in Figure 8). Overall, the results are not too sensitive to the choice of J. The finding of frequent switches to a constant gain coefficient in the 1970s is especially robust to the different J's. Switches to constant gain in the later part of the sample, instead, are more sensitive to its choice.

4.3.2. *Switching between high and low constant gains.* Since agents are unsure about the model of the economy and whether this is changing over time, one might argue that agents may be better off always using constant-gain learning, rather than reverting to a decreasing gain when their forecasting performance is satisfactory. I follow this argument here and assume that agents always adopt constant-gain learning, only switching from a 'low' to a 'high' gain when the conditions in (2.7) are met (the 'low' gain is fixed at 0.02, whereas the 'high' gain is estimated). The switches to the higher gain occur in similar periods to those found under the baseline case (Figure 10); the estimated gains equal 0.096 for inflation and 0.042 for the output gap.

4.3.3. *Alternative switching thresholds.* In the benchmark estimation, the gain switches when forecast errors fall above a historical volatility threshold, which is recursively updated. It can be argued that, in this way, the threshold may be more volatile in the first part of the sample, in which only few observations are used in its construction, than later on. We check the robustness of the results to that assumption here, by re-estimating the model now allowing for the threshold to be updated using a rolling window including ten years of data. The switches in the gains for output gap and inflation beliefs are again similar to those in the benchmark estimation, although switches are clearly more common (Figure 11, panel (i)).[18]

4.3.4. *Estimation Sample including the 'Great Recession'.* We now extend the sample's ending date from 2006:Q1 to 2013:Q4 to examine whether the learning mechanism also worked in the instability of the "Great Recession". The model is re-estimated on the expanded sample.[19] The evolution of the endogenous switching gain is shown in Figure 11, panel (ii). The gain for the output gap switches again to the high constant value in 2008 and it declines starting from late 2010. The findings, therefore, suggest that the uncertainty and larger forecast errors at the onset of the crisis affected

[18]I have also tried an estimation with fixed threshold ν over the full sample. The gain paths are similar.

[19]We abstract here from the zero-lower-bound constraint in the estimation in the later part of the sample. Deviations between the desired interest rate implied by the Taylor rule and zero are thus simply attributed to the monetary policy shock ε_t.

learning dynamics, which, in turn, likely induced higher volatility in the aggregate economy.

4.3.5. *Policy Rule and Policy Coefficient Priors.* The estimation is now repeated assuming a Taylor rule that responds to contemporaneous, rather than lagged, variables. The implied posterior estimates are not far from those shown in Table 2.

In the benchmark estimation, the prior for the inflation feedback coefficient in the Taylor rule was assumed to be Normally distributed with mean 1.5 and standard deviation 0.25 and the prior for the feedback to the output gap to be Normally distributed with mean 0.25 and standard deviation 0.125. Such priors assign most of the probability mass to values that would lead the system to satisfy the Taylor principle and that would, therefore, be consistent with determinacy (in a model with rational expectations) and E-stability (in the case of decreasing-gain learning and with a PLM that corresponds to the MSV solution under rational expectations). We now repeat the estimation with uninformative priors for the policy coefficients: we assume a Uniform distribution between 0 and 3 for the inflation feedback coefficient and a Uniform between 0 and 2 for the output gap coefficient. Rather than inducing determinacy and E-stability through the priors for these parameters, we let the data speak freely about their values, but directly impose the dogmatic prior that the system is consistent with a determinate and E-stable equilibrium. This is done in the estimation by rejecting every MCMC draw characterized by parameter combinations that fail to satisfy determinacy and E-stability conditions.

The posterior estimates, shown in Table 3b, indicate a larger switch in the inflation feedback coefficient between the pre-1979 and post-1979 samples (now from a posterior mean of 1.19 to a posterior mean equal to 2.03). The resulting time-variation in the endogenous gain coefficients for both cases of a contemporaneous Taylor rule and uninformative policy priors closely track each other and remain largely unchanged with respect to the baseline case (Figure 11, panel (iii) and (iv)).

4.3.6. *Observed Disturbances in PLM.* The benchmark estimation has assumed what is arguably the most empirically realistic forecasting model: one in which agents observe endogenous variables such as output, inflation, and interest rates, but cannot observe exogenous disturbances. For robustness, however, the model is re-estimated using the complete MSV solution, endowing agents with knowledge of the disturbances at each point in the sample as well. Their PLM now takes the form

$$Z_t = a_t + b_t Z_{t-1} + c_t w_t + \eta_t. \tag{4.2}$$

The implied time-varying gains maintain constant values in the 1960s and 1970s and switch to decreasing gains for most of the post-1980 period (Figure 11, panel (v)). The estimated values for the constant gain coefficients equal 0.117 for inflation beliefs, 0.058 for output beliefs, and 0.024 for interest rates. The properties of the disturbances change somewhat, with autoregressive coefficients increasing to 0.91, for the supply, and to 0.94, for the demand shock. The results under this scenario, therefore, also do not overturn the main conclusion that learning gain coefficients were higher before 1980 than after.

4.3.7. *TV Volatilities of structural shocks.* The benchmark estimation assumes that structural shocks have constant variance over the whole sample. It is interesting to consider whether the results are robust to allowing the volatility of the shocks to vary over time, as already assumed for the monetary policy coefficients. Motivated by the Great Moderation literature, the standard deviations of the supply and demand shocks are allowed to switch between the pre-1984 and post-1984 periods (with a break in 1984:Q2). For the standard deviation of the monetary policy shock, we allow it, instead, to be drawn from a distribution that has a different variance in the non-borrowed-reserve targeting period (1979:Q3-1982:Q4) than in the remaining samples. Therefore, the standard deviations evolve as follows

$$\sigma_{\varepsilon,t} = \begin{cases} \sigma_{\varepsilon,1979-82} & 1979 : Q3 \leq t \leq 1982 : Q4 \\ \sigma_{\varepsilon} & o/w \end{cases}$$

$$\sigma_{u,t} = \begin{cases} \sigma_{u,pre-1984} & t < 1984 : Q2 \\ \sigma_{u,post-1984} & t \geq 1984 : Q2 \end{cases}, \sigma_{g,t} = \begin{cases} \sigma_{g,pre-1984} & t < 1984 : Q2 \\ \sigma_{g,post-1984} & t \geq 1984 : Q2 \end{cases}$$

Table 4 reports the estimation results. There is evidence that the volatility of the shocks has decreased over time. The posterior mean estimate for σ_u and σ_g decline from 1.02 to 0.73 and from 0.76 to 0.50 after 1984. Figure 11 displays the evolution of the switching gains related to inflation and output in this context (panel (vi)). The gains switch to high constant gains in the first half of the sample and do so less often in the second half, as found in the benchmark estimation.

5. Simulation

I repeat the simulation of the model, but now using the estimated parameter values (shown in Table 2) from the previous section and fixing the agents' learning to resemble the one estimated from US data (i.e, assuming that the endogenous gain switches as in Figure 8). I simulate an economy with 185 periods (the same length as the estimated New Keynesian model) for 10,000 times. The shocks that hit the economy are drawn from distributions with constant variance over the whole sample.[20] We have previously seen that learning can imply time-varying volatility in the variables about which agents are forming expectations. But suppose that learning is neglected in an empirical exercise. Let's consider the following experiment. Suppose that an econometrician would estimate inflation and output

[20] I use a 'projection facility' in the simulation to ensure that the economy does not become unstable. As in Orphanides and Williams (2005, 2007), in fact, I assume that agents recognize that the economy is stable and every time the matrix of autoregressive coefficients in their VAR has an eigenvalue larger than 1 in absolute value, they do not update their estimates, keeping $\widehat{\phi}_t = \widehat{\phi}_{t-1}$ and $R_t = R_{t-1}$. If this is not enough to guarantee non-explosive dynamics, I reject the specific draw.

equations on the simulated data, but without taking learning into account. Would the econometrician find evidence of time-varying volatility, even if the true data-generating process had shocks with *constant* variance through the whole sample?

To answer this question, I regress artificially-generated inflation and output gap series on a constant and their first lag (a similar regression on actual data gives the plot for conditional standard deviations in Figure 1); then I take the implied residuals and perform a test on the existence of ARCH/GARCH effects (at the 5% significance level). Table 5 reports the percentages of rejections of the null hypothesis of no ARCH/GARCH effects from simulated data. In the case that the data derive from an economy with no learning (i.e., imposing $g_{t,y} = 0$ at all times), the test rejects the null of no ARCH effects only about 5% of the times. In the case with learning, even though the variances of the shocks are constant by construction, the test concludes that ARCH/GARCH effects are a feature of the data in 52% of the cases for inflation and 78% for output gap (see Table 5 for more results under different cases).[21]

These results are suggestive that estimations that abstract from agents' learning can significantly overestimate the time variation in the volatility of exogenous shocks.[22]

5.1. **Time-Varying Volatility.** The model can generate time-varying volatility in macroeconomic variables even if the exogenous disturbances have maintained constant variance over the whole sample. I try to verify whether

[21]A more sophisticated version of the same experiment would imply estimating the full DSGE model under Rational Expectations, hence disregarding learning dynamics, to test for the existence of spurious ARCH effects or stochastic volatility in the exogenous shocks. This would be, however, extremely computationally cumbersome, as it would require running a full set of MH draws for each of the 10,000 sets of artificially simulated series. A similar, more-reduced-form, approach as the one used in this paper was also followed in Lubik and Surico (2010) in a different application.

[22]As mentioned in the introduction, a similar point, albeit made in a largely different context, is reached by Bullard and Singh (2012).

the model can generate a pattern of volatility that roughly resembles what is observed in actual data. That is, I aim to verify whether the outcomes from the simulations imply volatility series that increase in the first part of the sample (when the endogenous gain switches to constant) and decline in the second part (when the endogenous gain reverts to decreasing). In each simulation, I take the residuals from the inflation and output equations and I look at the point in the sample in correspondence of which the maximum rolling standard deviation is obtained (using a rolling window of 20 periods). I then estimate the Kernel density of the maxima across all simulations. Figure 12 displays the results. The standard deviations often increase near those observations that would correspond to the late 1960s and 1970s, and become typically lower in the second part of the sample. If the economy was simulated without learning, instead, one would find that the maxima of the volatility series are uniformly spread across all observations in the sample.

Therefore, as in the previous section, neglecting learning in empirical work, if learning is a feature of the data-generating process of the economy, may lead researchers to spuriously find time variation in the volatility of residuals from estimated univariate equations (or possibly of exogenous shocks in DSGE models under RE). The pattern of volatility is in the ballpark of that estimated on actual data and reported in Figure 1. One would, in fact, conclude that the volatility of shocks has increased in periods in which agents' forecast errors are generally large and in which learning introduces more noise in the economy.

5.2. **The Great Moderation.** The model estimation has shown that the gain coefficients were typically larger in the 1960s and 1970s than in the following decades, and this has affected the degree of volatility in the economy. The volatility of output gap and inflation has fallen in the post-1984 sample. Is the model with learning able to generate the Great Moderation? For each simulation, I compute the ratio of the standard deviations of inflation and output gap in the second part of the sample (corresponding to 1985-2006 observations) versus the first part (which correspond to 1960-1984). Taking the median from the simulations, for the baseline model, the ratios between post-1984 and pre-1984 standard deviations equal 0.39 for inflation (versus 0.35 on actual data) and 0.42 for output gap (versus 0.50 on actual data). The model with learning, with parameters fixed at their posterior mean estimates, is, therefore, in principle capable of endogenously generating a reduction in the volatility of the main macroeconomic variables of a magnitude comparable to the Great Moderation (Table 6 shows that the results remain similar under the other cases).

We can provide some complementary evidence on the ability of the model to explain portions of macroeconomic volatility through the estimation of the model as well. In section 4, the estimated model allowed for a switch in the volatilities of the shocks in 1984 and included the endogenous switching gain. To gain intuition on the relative contributions of learning and exogenous luck in driving changes in volatility, the model can be re-estimated now shutting down the switching gain. The model now assumes a recursive-least-square learning with a decreasing gain over the full sample. The results in Table 4 indicate that the volatility of the demand shock has fallen from 1.16 to 0.52, while the volatility for the supply shock has fallen from 1.31 to 1.02. The decline in the standard deviation of the cost-push shock in the second sample persists in both estimations, with and without the endogenous switching gain; learning, however, can account for some of the exogenous volatility both in the pre- and post-1984 samples (the required variances for the exogenous shocks are reduced from 1.31 to 1.02 and from 1.02 to 0.73 when the endogenous gain channel is re-introduced).

Learning with a switching gain can also explain a portion of the moderation in output volatility after 1984. Figure 13 shows the posterior distributions of the standard deviation coefficients related to the demand shock. The figure makes clear that the change in (exogenous) luck needed to fit the data is much larger, more than double, when the endogenous switching gain learning is replaced by learning with a decreasing gain. The posterior means for the demand shock standard deviations go from shifting from 1.16 to around 0.5 to shifting from 0.75 to around 0.5, under switching-gain learning.[23]

5.3. **The Effects of Policy.** The literature that studies the main sources of the Great Moderation has often focused on testing explanations based on changes in monetary policy versus explanations based on reductions in the volatility of the exogenous shocks that hit the economy. A decline in the estimated volatility of shocks is usually taken as evidence in favor of the bad luck-good luck hypothesis versus the alternative hypothesis of transition from bad to good policy.

Changes in volatility, however, may not be unrelated to the monetary policy regime. Chairman Bernanke, in a speech about the Great Moderation in February 2004, for example, argued

> [*I am not convinced that the decline in macroeconomic volatility of the past two decades was primarily the result of good luck.*]

> [*...changes in monetary policy could conceivably affect the size and frequency of shocks hitting the economy, at least as an econometrician would measure those shocks*]

He continues:

[23]Here, I have assumed that the standard deviations of the shocks, as well as the monetary policy coefficients, change at discrete points in the sample. We have used the size of the estimated switch to obtain intuition about what portion of volatility can be explained by the endogenous-gain learning process. A more sophisticated approach would consist of re-estimating the model, but allowing for both learning and stochastic volatility. That would, however, imply moving to a nonlinear non-Gaussian framework, including both continuously-changing parameters and volatilities, the estimation of which could be performed using the particle filter. The estimation of such a model is beyond the scope of this paper, though.

[*changes in inflation expectations, which are ultimately the product of the monetary policy regime, can also be confused with truly exogenous shocks in conventional econometric analyses.*]

Therefore,

[*some of the effects of improved monetary policies may have been misidentified as exogenous changes in economic structure or in the distribution of economic shocks.*]

This paper makes time-varying volatility endogenous. In this way, it is easier to study how the volatility of the shocks (as they are measured by the econometrician) may itself depend on policy. This section can, therefore, provide an initial evaluation of the extent to which changes in the volatility of inferred economic shocks may, as in Bernanke's claims, simply reflect better monetary policies.

To examine the interaction between volatility and monetary policy, I simulate the economy with an endogenous gain as in section 3, but now under a wide range of policy feedback coefficients to inflation (using a grid from 0 to 5 in 0.5 increments). Figure 14 shows the result under the model parametrization obtained in Table 2. A more aggressive monetary policy reduces agents' forecast errors and, therefore, it affects the frequency of switches in their gain coefficients (from almost 70% of the times when $\chi_\pi = 0$ to less than 30% when $\chi_\pi = 5$, with an average gain in the sample that goes from above 0.07 to below 0.05), and, through this channel, it affects also changes in volatility in the economy.[24] The more aggressive monetary policy, the less often the econometrician would spuriously find time-varying volatility in the reduced-form residuals (from more than 80 to 55% of the times). Changes in the volatility of estimated shocks, therefore, may not in principle be a matter of luck after all, but an implication of better policy (notice, however, that under the estimated coefficients in this paper, large changes in volatility due to policy changes are unlikely).

[24] The importance of monetary policy in reducing agents' forecast errors is also discussed in Orphanides and Williams (2003).

This result echoes, although in a different setting and with a different focus, the argument in a recent paper by Benati and Surico (2009). They artificially generate data from a New Keynesian model assuming a policy change from a 'passive' to an 'active' policy rule and they ask whether a common VAR estimation would be able to recover the change in policy. They show that the estimated VAR would lead researchers to inaccurately conclude that the variances of the shocks have changed, but not the policy coefficients.

This paper's results similarly suggest caution: simply finding that the variances have changed from reduced-form regressions may not necessarily imply changes in luck, but it might be an effect of policy changes or, as in this paper, of the interaction between changing policies and private agents' learning.

6. Conclusions and Future Directions

The paper has presented a New Keynesian model in which agents' learning with a switching gain coefficient endogenously generates time-varying volatility in the economy. The estimation of the model has shown that there is evidence of large changes in the gain over the post-war US sample. The changes in the gain can imply important changes in macroeconomic volatility, which can roughly match the magnitude of the Great Moderation. An econometrician that would abstract from such learning dynamics, however, would be lead to overestimate the importance of changes in the volatility of exogenous shocks. Moreover, time variation in volatility may not be simply a matter of luck, but it may itself be affected by changes in monetary policy and, in particular, it can stem from the interaction between policy and learning by private agents.

An important direction for future research will be to test whether extensions of the model would be able to generate endogenous stochastic volatility series able to match those estimated in DGSE models by Justiniano and Primiceri (2008) and Fernandez-Villaverde and Rubio-Ramirez (2007). More generally, this would allow researchers to test what fraction of the changes in volatility may be due to learning, rather than to the decline in the volatility of exogenous disturbances, at a finer level of detail than what reported in this paper.

References

[1] Adam, K., (2005). "Learning to Forecast and Cyclical Behavior of Output and Inflation", *Macroeconomic Dynamics*, Vol. 9(1), 1-27.

[2] Benati, L., and P. Surico, (2009). "VAR Analysis and the Great Moderation", *American Economic Review*, 99(4), 1636-52.

[3] Blanchard, O.J., and J. Simon, (2001). "The Long and Large Decline in U.S. Output Volatility", *Brookings Papers on Economic Activity*, 1, pp. 135-174.

[4] Branch, W., (2006). "Restricted Perceptions Equilibria and Learning in Macroeconomics ", in David Colander ed., *Post Walrasian Macroeconomics. Beyond the Dynamic Stochastic General Equilibrium Model*, Cambridge: Cambridge University Press, 135-160.

[5] Branch, W., and G.W. Evans, (2006). "A Simple Recursive Forecasting Model", *Economics Letters*, vol. 91(2), 158-166.

[6] Branch, W., and G.W. Evans, (2007). "Model Uncertainty and Endogenous Volatility", *Review of Economic Dynamics*, vol. 10(2), 207-237.

[7] Bullard, J., and K. Mitra, (2007). "Determinacy, Learnability, and Monetary Policy Inertia", *Journal of Money, Credit, and Banking*, 39(5), 1177-1212.

[8] Bullard, J., and A. Singh, (2012). "Learning and the Great Moderation", *International Economic Review*, 53(2), 375-397.

[9] Bullard, J., and J. Suda, (2011). "The Stability of Macroeconomic Systems with Bayesian Learners," mimeo, Federal Reserve Bank of St. Louis and Banque de France.

[10] Cho, I.-K. and K. Kasa, (2012). "Model Validation and Learning," mimeo, University of Illinois and Simon Fraser University.

[11] Colucci, D., and V. Valori, (2004). "Adaptive Learning in the Cobweb Model with an Endogenous Gain Sequence", mimeo, Universita' di Firenze.

[12] Colucci, D., and V. Valori, (2005). "Error Learning Behaviour and Stability Revisited", *Journal of Economic Dynamics and Control*, Vol. 29/03, 371-388.

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

[14] Evans, G. W., and G. Ramey (2006). "Adaptive Expectations, Underparameterization and the Lucas Critique", *Journal of Monetary Economics*, Vol. 53, 249-264.

[15] Fernandez-Villaverde, J., and J. Rubio-Ramirez, (2007). "Estimating Macroeconomic Models: A Likelihood Approach", *Review of Economic Studies*, vol. 74(4), pp. 1059-1087.

[16] Gaus, E., (2011). "DSGE Estimation of Models with Learning", mimeo, Ursinus College.

[17] Gaus, E., (2013). "Robust Stability of Monetary Policy Rules under Adaptive Learning", *Southern Economic Journal*, forthcoming.

[18] Honkapohja, S., and K. Mitra, (2003). "Learning with Bounded Memory in Stochastic Models", *Journal of Economic Dynamics and Control*, vol. 27(8), 1437-1457.

[19] Justiniano A. and G.E. Primiceri (2008). "The Time Varying Volatility of Macroeconomic Fluctuations", *American Economic Review*, 98(3), pp. 604-41.

[20] Kim, C.J., and C.R. Nelson, (1999). "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *The Review of Economics and Statistics*, vol. 81(4), pages 608-616.

[21] Kostyshyna, O., (2012). "Application of an Adaptive Step-Size Algorithm in Models of Hyperinflation", Macroeconomic Dynamics, 16(S3), 355-375.

[22] Lansing, K., (2009). "Time Varying U.S. Inflation Dynamics and the New Keynesian Phillips Curve," *Review of Economic Dynamics*, 12(2), 304-326.

[23] Lubik, T., and F. Schorfheide, (2004). "Testing for Indeterminacy: an Application to U.S. Monetary Policy", *American Economic Review*, 94(1), pp. 190-217.

[24] Lubik, T., and P. Surico, (2010). "The Lucas critique and the stability of empirical models", *Journal of Applied Econometrics*, 25(1), 177194.

[25] Marcet, A., and J.P. Nicolini, (2003). "Recurrent Hyperinflations and Learning," *American Economic Review*, vol. 93(5), pages 1476-1498.

[26] McConnell, M.M., and G. Perez-Quiros, (2000). Output Fluctuations in the United States: What Has Changed Since the Early 1980s? *American Economic Review*, 90, 1464-1476.

[27] Milani, F., (2008). "Learning, Monetary Policy Rules, and Macroeconomic Stability", *Journal of Economic Dynamics and Control*, 32(10), 3148-3165.

[28] Milani, F., (2007). "Expectations, Learning and Macroeconomic Persistence", *Journal of Monetary Economics*, 54(7), 2065-2082.

[29] Milani, F., (2011). "Expectation Shocks and Learning as Drivers of the Business Cycle", *Economic Journal*, 121(552), 379-401.

[30] Orphanides, A., and J. Williams, (2003). "Imperfect Knowledge, Inflation Expectations and Monetary Policy", in Ben Bernanke and Michael Woodford, eds., *Inflation Targeting*. Chicago: University of Chicago Press.

[31] Orphanides, A., and J. Williams,(2005a). "Inflation Scares and Forecast-Based Monetary Policy", *Review of Economic Dynamics*, Vol. 8 (2), pages 498-527.

[32] Orphanides, A., and J. Williams, (2005b). "The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations", *Journal of Economic Dynamics and Control*, vol. 29(11), pages 1927-1950.

[33] Orphanides, A., and J. Williams, (2006). "Monetary Policy with Imperfect Knowledge", *Journal of the European Economic Association*, 4(2-3), 366-375.

[34] Orphanides, A., and J. Williams, (2007). "Inflation Targeting and Imperfect Knowledge", *Federal Reserve of San Francisco Economic Review*.

[35] Owyang, M.T., (2001). "Persistence, Excess Volatility, and Volatility Clusters in Inflation", *Federal Reserve Bank of St. Louis Review*, 83(6), pp. 41-52.

[36] Primiceri, G.E., (2005). Time Varying Structural Vector Autoregressions and Monetary Policy", *Review of Economic Studies*, 72, 821-852.

[37] Smets, F., and R. Wouters, (2003). "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, vol. 1(5), pages 1123-1175.

[38] Smets, F., and R. Wouters, (2007). "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach", *American Economic Review*, 97(3), 586-606.

[39] Stock, J.H., and M.W. Watson, (2002). "Has the Business Cycle Changed, and Why?", *NBER Macroeconomics Annual*, 17, 159-218.

[40] Woodford, M., (2003). *Interest and Prices*. Princeton University Press.

		Prior Distribution			
Description	Param.	Range	Distr.	Mean	95% Int.
Inverse IES	σ^{-1}	$\mathbb{R}^+$	G	1	$[.12, 2.78]$
Slope PC	κ	$\mathbb{R}^+$	G	.25	$[.03, .7]$
Discount Rate	β	.99	–	.99	–
Interest-Rate Smooth	ρ_{pre79}	$[0, 1]$	B	.8	$[.46, .99]$
Feedback to Infl.	$\chi_{\pi,pre79}$	$\mathbb{R}$	N	1.5	$[1.01, 1.99]$
Feedback to Output	$\chi_{x,pre79}$	$\mathbb{R}$	N	.25	$[.01, .49]$
Interest-Rate Smooth	ρ_{post79}	$[0, 1]$	B	.8	$[.46, .99]$
Feedback to Infl.	$\chi_{\pi,post79}$	$\mathbb{R}$	N	1.5	$[1.01, 1.99]$
Feedback to Output	$\chi_{x,post79}$	$\mathbb{R}$	N	.25	$[.01, .49]$
Autoregr. Cost-push shock	ρ_u	$[0, 1]$	B	0.8	$[.57, .95]$
Autoregr. Demand shock	ρ_g	$[0, 1]$	B	0.8	$[.57, .95]$
Std. MP shock	σ_ε	$\mathbb{R}^+$	IG	0.5	$[.17, 1.34]$
Std. g_t	σ_g	$\mathbb{R}^+$	IG	0.5	$[.17, 1.34]$
Std. u_t	σ_u	$\mathbb{R}^+$	IG	0.5	$[.17, 1.34]$
Constant Gain infl.	$\overline{\mathbf{g}}_\pi$	$[0, 0.3]$	U	.15	$[.007, .294]$
Constant Gain gap	$\overline{\mathbf{g}}_x$	$[0, 0.3]$	U	.15	$[.007, .294]$
Constant Gain FFR	$\overline{\mathbf{g}}_i$	$[0, 0.3]$	U	.15	$[.007, .294]$

Table 1 - Prior Distributions. *Notes*: the table displays prior means and 95% posterior probability intervals for the model parameters. N denotes Normal distribution, G denotes Gamma distribution, B denotes Beta distribution, IG denotes Inverse-Gamma distribution, and U denotes Uniform distribution.

		Posterior Distribution	
Description	Parameter	Mean	95% Post. Prob. Int.
Inverse IES	σ^{-1}	5.92	[4.23-8.34]
Slope PC	κ	0.022	[0.003-0.06]
Discount Factor	β	0.99	-
IRS pre-79	ρ_{pre79}	0.938	[0.85-0.99]
Feedback Infl. pre79	$\chi_{\pi,pre-79}$	1.37	[0.87-1.89]
Feedback Gap pre79	$\chi_{x,pre-79}$	0.58	[0.18-1.02]
IRS post-79	ρ_{post79}	0.93	[0.88-0.97]
Feedback Infl. post79	$\chi_{\pi,post-79}$	1.53	[1.05-2.05]
Feedback Gap post79	$\chi_{x,post-79}$	0.48	[0.04-0.92]
Autoregr. Cost-push shock	ρ_u	0.40	[0.28-0.52]
Autoregr. Demand shock	ρ_g	0.84	[0.75-0.92]
Std. Cost-push shock	σ_u	0.89	[0.8-1.00]
Std. Demand shock	σ_g	0.65	[0.58-0.72]
Std. MP shock	σ_ε	0.97	[0.87-1.07]
Constant gain (Infl.)	$\overline{\mathbf{g}}_\pi$	0.082	[0.07-0.09]
Decreasing gain (Infl.)	t^{-1}	-	-
Constant gain (Gap)	$\overline{\mathbf{g}}_x$	0.073	[0.06-0.083]
Decreasing gain (Gap)	t^{-1}	-	-
Constant gain (FFR)	$\overline{\mathbf{g}}_i$	0.001	[0,0.01]
Decreasing gain (FFR)	t^{-1}	-	-

Table 2 - Posterior Estimates. *Note*: the table shows posterior mean estimates and 95% posterior probability intervals obtained for the baseline case, with forecast window $J = 4$.

		Posterior Distributions							
		$J = 20$		Estim. $J = \hat{J}$		High/low CG		Rolling ν_t	
Description	Parameter	Mean	95% PPI	Mean	95% PPI	Mean	95% PPI	Mean	95% PPI
Inverse IES	σ^{-1}	6.89	[4.62,10]	6.37	[4.43,9.13]	7.44	[5.22-10.34]	6.50	[4.64-8.82]
Slope PC	κ	0.02	[0.003,0.06]	0.02	[0.003,0.054]	0.025	[0.003-0.07]	0.022	[0.003-0.05]
Discount Factor	β	0.99	-	0.99	-	0.99	-	0.99	-
IRS pre-79	ρ_{pre79}	0.936	[0.86-0.99]	0.94	[0.87-0.99]	0.932	[0.856-0.99]	0.94	[0.80-0.99]
Feedback Infl. pre79	$\chi_{\pi,pre-79}$	1.31	[0.91-1.76]	1.37	[0.91-1.82]	1.33	[0.86-1.82]	1.33	[0.85-1.90]
Feedback Gap pre79	$\chi_{x,pre-79}$	0.61	[0.13-0.98]	0.63	[0.26-1.11]	0.62	[0.22-1]	0.30	[0.07-0.54]
IRS post-79	ρ_{post79}	0.914	[0.86-0.96]	0.93	[0.88-0.97]	0.91	[0.86-0.96]	0.92	[0.87-0.97]
Feedback Infl. post79	$\chi_{\pi,post-79}$	1.54	[1.09-1.93]	1.57	[1.04-2.04]	1.56	[1.1-2.02]	1.60	[1.18-2.01]
Feedback Gap post79	$\chi_{x,post-79}$	0.41	[0.006-0.87]	0.49	[0.04-0.88]	0.44	[-0.04-0.85]	0.25	[0.006-0.48]
Autoregr. Cost-push shock	ρ_u	0.39	[0.26-0.5]	0.402	[0.28-0.52]	0.31	[0.18-0.46]	0.40	[0.28-0.55]
Autoregr. Demand shock	ρ_g	0.83	[0.74-0.91]	0.83	[0.73-0.92]	0.79	[0.7-0.87]	0.84	[0.75-0.91]
Std. Cost-push shock	σ_u	0.97	[0.87-1.07]	0.90	[0.81-0.99]	0.91	[0.82-1.01]	0.91	[0.83-1.01]
Std. Demand shock	σ_g	0.67	[0.61-0.75]	0.64	[0.58-0.71]	0.74	[0.66-0.82]	0.62	[0.55-0.69]
Std. MP shock	σ_ε	0.96	[0.87-1.07]	0.97	[0.87-1.07]	0.96	[0.87-1.07]	0.97	[0.87-1.08]
F.E. Window	J	20	-	4	[1-5]	4	-	4	-
Constant gain (Infl.)	$\overline{\mathbf{g}}_\pi$	0.065	[0.055-0.075]	0.078	[0.062-0.097]	0.096	[0.09-0.104]	0.066	[0.054-0.075]
Decreasing gain (Infl.)	t^{-1}	-	-	-	-	-	-	-	-
Constant gain (Gap)	$\overline{\mathbf{g}}_x$	0.064	[0.049-0.072]	0.073	[0.06-0.084]	0.042	[0.036-0.052]	0.072	[0.063-0.080]
Decreasing gain (Gap)	t^{-1}	-	-	-	-	-	-	-	-
Constant gain (FFR)	$\overline{\mathbf{g}}_i$	0.005	[0,0.032]	0.001	[0,0.01]	0.01	-	0.001	[0-0.009]
Decreasing gain (FFR)	t^{-1}	-	-	-	-	-	-	-	-
Low Constant gain (Infl.)	$\overline{\mathbf{g}}_\pi^L$					0.02	-		
Low Constant gain (Gap)	$\overline{\mathbf{g}}_x^L$					0.02	-		

Table 3a - Posterior Distributions: Robustness checks. *Notes*: the posterior estimates refer to the following cases: i) estimation with forecast window fixed at $J = 20$; ii) estimated forecast window $\widehat{J}$; iii) estimation with endogenous gain switching between high and low constant gain values; iv) estimation using rolling thresholds as benchmark to which historical forecast errors are compared.

		Posterior Distributions							
		Sample End. 2014		Contemp. Policy Rule		Diffuse Policy Priors		Dist. in PLM	
Description	Parameter	Mean	95% PPI	Mean	95% PPI	Mean	95% PPI	Mean	95% PPI
Inverse IES	σ^{-1}	7.32	[5.06-10.09]	5.70	[4.03-7.82]	6.07	[4.31-8.39]	12.51	[10.07-15.47]
Slope PC	κ	0.018	[0.002-0.05]	0.022	[0.002-0.061]	0.025	[0.003-0.07]	0.021	[0.002-0.05]
Discount Factor	β	0.99	-	0.99	-	0.99	-	0.99	-
IRS pre-79	ρ_{pre79}	0.94	[0.85-0.99]	0.90	[0.79-0.98]	0.94	[0.88-0.99]	0.98	[0.94-0.99]
Feedback Infl. pre79	$\chi_{\pi,pre-79}$	1.32	[0.85-1.75]	1.32	[0.82-1.87]	1.19	[0.18-2.51]	1.46	[1.00-1.94]
Feedback Gap pre79	$\chi_{x,pre-79}$	0.30	[0.05-0.52]	0.36	[0.13-0.57]	0.74	[0.03-1.95]	0.26	[0.06-0.50]
IRS post-79	ρ_{post79}	0.93	[0.88-0.98]	0.90	[0.85-0.94]	0.93	[0.88-0.97]	0.96	[0.92-0.99]
Feedback Infl. post79	$\chi_{\pi,post-79}$	1.60	[1.16-2.03]	1.63	[1.20-2.17]	2.03	[0.19-2.96]	1.48	[0.95-1.96]
Feedback Gap post79	$\chi_{x,post-79}$	0.27	[0.03-0.52]	0.33	[0.10-0.55]	0.60	[0.03-0.64]	0.24	[0.01-0.48]
Autoregr. Cost-push shock	ρ_u	0.34	[0.22-0.46]	0.43	[0.31-0.54]	0.41	[0.29-0.55]	0.91	[0.87-0.95]
Autoregr. Demand shock	ρ_g	0.80	[0.71-0.88]	0.83	[0.75-0.91]	0.84	[0.76-0.92]	0.94	[0.87-0.99]
Std. Cost-push shock	σ_u	0.89	[0.81-1.00]	0.90	[0.81-1.00]	0.89	[0.81-0.99]	1.12	[0.98-1.29]
Std. Demand shock	σ_g	0.68	[0.62-0.74]	0.64	[0.58-0.71]	0.65	[0.58-0.72]	0.41	[0.35-0.47]
Std. MP shock	σ_ε	0.92	[0.83-1.00]	0.95	[0.85-1.05]	0.96	[0.86-1.07]	0.97	[0.88-1.08]
F.E. Window	J	4	-	4	-	4	-	4	-
Constant gain (Infl.)	$\overline{\mathbf{g}}_\pi$	0.075	[0.057-0.094]	0.082	[0.078-0.09]	0.083	[0.078-0.091]	0.117	[0.101-0.135]
Decreasing gain (Infl.)	t^{-1}	-	-	-	-	-	-	-	-
Constant gain (Gap)	$\overline{\mathbf{g}}_x$	0.059	[0.048-0.068]	0.077	[0.067-0.086]	0.072	[0.06-0.081]	0.058	[0.052-0.065]
Decreasing gain (Gap)	t^{-1}	-	-	-	-	-	-	-	-
Constant gain (FFR)	$\overline{\mathbf{g}}_i$	0.006	[0-0.035]	0.001	[0,0.004]	0.001	[0-0.003]	0.024	[0.005-0.04]
Decreasing gain (FFR)	t^{-1}	-	-	-	-	-	-	-	-

Table 3b - Posterior Distributions: Robustness checks. *Notes*: the posterior estimates refer to the following cases: v) estimation with sample ending in 2014; vi) model with Taylor rule responding to contemporaneous variables; vii) estimation with Uniform priors for Taylor rule coefficients and determinacy/E-stability prior; viii) model with observed disturbances in the agents' PLM.

		Posterior Distributions			
		TV variances, switching gain		TV variances, switching gain shut off	
Description	Parameter	Mean	95% PPI	Mean	95% PPI
Inverse IES	σ^{-1}	6.51	[4.50-9.21]	6.98	[4.65-10.1]
Slope PC	κ	0.021	[0.003-0.06]	0.035	[0.005-0.09]
Discount Factor	β	0.99	-	0.99	-
IRS pre-79	ρ_{pre79}	0.95	[0.87-0.99]	0.94	[0.85-0.99]
Feedback Infl. pre79	$\chi_{\pi,pre-79}$	1.31	[0.83-1.82]	1.22	[0.70-1.79]
Feedback Gap pre79	$\chi_{x,pre-79}$	0.31	[0.05-0.55]	0.33	[0.07-0.57]
IRS post-79	ρ_{post79}	0.93	[0.89-0.97]	0.91	[0.86-0.96]
Feedback Infl. post79	$\chi_{\pi,post-79}$	1.48	[0.99-1.95]	1.45	[1.05-1.88]
Feedback Gap post79	$\chi_{x,post-79}$	0.30	[0.06-0.53]	0.30	[0.06-0.53]
Autoregr. Cost-push shock	ρ_u	0.38	[0.26-0.50]	0.45	[0.32-0.56]
Autoregr. Demand shock	ρ_g	0.81	[0.71-0.91]	0.76	[0.66-0.86]
Std. Cost-push shock pre84	$\sigma_{u,pre-84}$	1.02	[0.88-1.17]	1.31	[1.14-1.51]
Std. Demand shock pre84	$\sigma_{g,pre-84}$	0.76	[0.66-0.88]	1.16	[1.00-1.33]
Std. MP shock 1979-82	$\sigma_{\varepsilon,79-82}$	2.30	[1.64-3.18]	2.26	[1.57-3.20]
Std. Cost-push shock post84	$\sigma_{u,post-84}$	0.73	[0.63-0.85]	1.02	[0.88-1.17]
Std. Demand shock post84	$\sigma_{g,post-84}$	0.50	[0.43-0.58]	0.52	[0.45-0.59]
Std. MP shock 1960-79&1982-06	σ_{ε}	0.76	[0.68-0.84]	0.75	[0.67-0.83]
Constant gain (Infl.)	$\overline{\mathbf{g}}_{\pi}$	0.084	[0.078-0.095]	-	-
Decreasing gain (Infl.)	t^{-1}	-	-	-	-
Constant gain (Gap)	$\overline{\mathbf{g}}_{x}$	0.073	[0.06-0.085]	-	-
Decreasing gain (Gap)	t^{-1}	-	-	-	-
Constant gain (FFR)	$\overline{\mathbf{g}}_{i}$	0.001	[0-0.008]	-	-
Decreasing gain (FFR)	t^{-1}	-	-	-	-

Table 4 - Posterior Distributions: Estimation with shock variances switching in 1984:Q2. *Notes*: The first set of estimates in the left columns refers to the model with switching variances and with the TV endogenous switching gain; the second set of estimates refers to the same model with switching variances, but with the possibility of endogenous switches in the gain shut down.

	Endogenous TV Gain				No Learning	
	$J = 4$		$J = 20$			
	ARCH(1)	GARCH(1,1)	ARCH(1)	GARCH(1,1)	ARCH(1)	GARCH(1,1)
Inflation	0.517	0.61	0.48	0.56	0.05	0.06
Output Gap	0.785	0.89	0.85	0.90	0.045	0.05

Table 5 - Test for the existence of ARCH/GARCH effects (5% significance): proportion of rejections of the null hypothesis of no ARCH/GARCH effects. *Notes:* To test for evidence of ARCH(q) effects against the hypothesis of no ARCH effects, the squared residuals from inflation and output gap regressions are regressed on a constant and q lagged values and the statistic $\chi^2 = TR^2$ is computed. Under the null hypothesis of no ARCH effects this statistic has a limiting chi-square distribution with q degrees of freedom. The test for GARCH(p, q) is equivalent to a test for ARCH($p + q$).

	Endogenous TV Gain Baseline	$J = 20$	CG	No Learning	Data
Ratio $\frac{Std.\ Infl.\ 1985-2006}{Std.\ Infl.\ 1960-1984}$	0.39	0.42	0.43	1.00	**0.35**
Ratio $\frac{(Std.\ OutputGap\ 1985-2006)}{(Std.\ Output\ Gap\ 1960-1984)}$	0.42	0.52	0.54	1.00	**0.50**

Table 6 - The Great Moderation: ratio of standard deviations for inflation and output gap in the second versus the first part of the simulated samples (median across simulations).

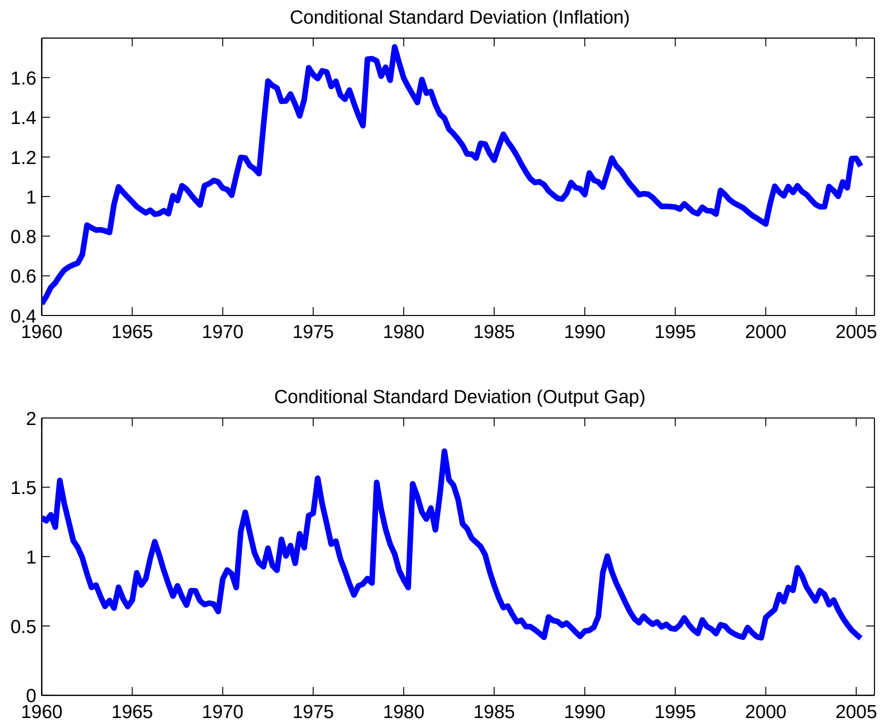

Figure 1. Conditional Standard Deviation series for Inflation and Output Gap.

Notes: to compute the conditional standard deviation series, I have estimated AR(1) models for inflation and output gap series (the latter calculated using the deviation of real GDP from the CBO's potential GDP series), allowing for a GARCH(1,1) specification for the residuals.

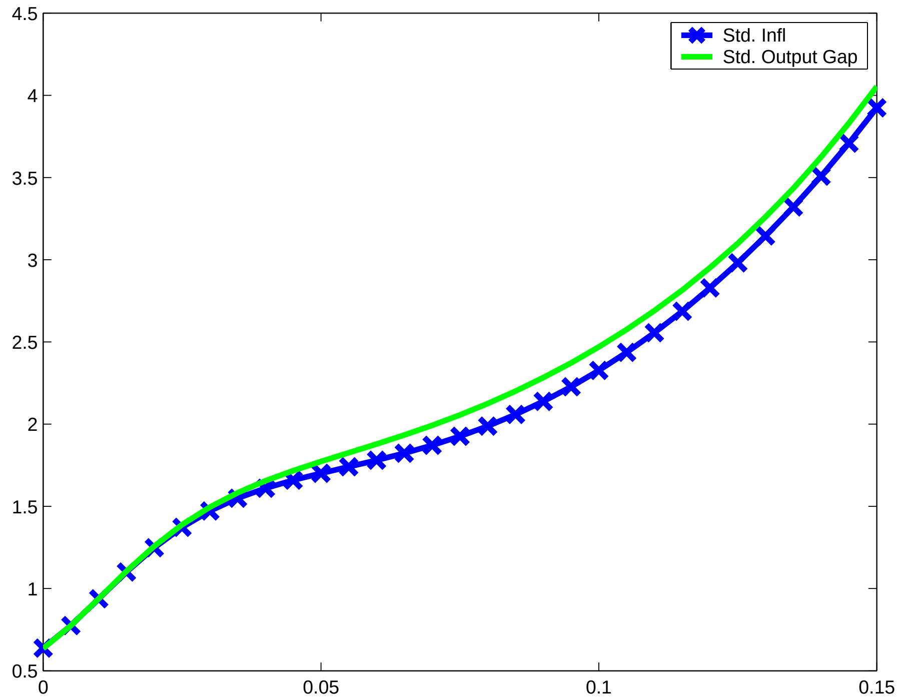

FIGURE 2. Volatility of simulated Inflation and Output Gap as a function of the constant gain coefficient.

Notes: The economy is simulated for 1,000 periods using a grid of constant gain values from 0 to 0.15. The calibration is: $\beta = 0.99$, $\kappa = 0.05$, $\sigma = 0.1$, $\rho = 0.95$, $\chi_\pi = 1.5$, $\chi_x = 0.5$, $\rho_u = 0.9$, $\rho_g = 0.9$.

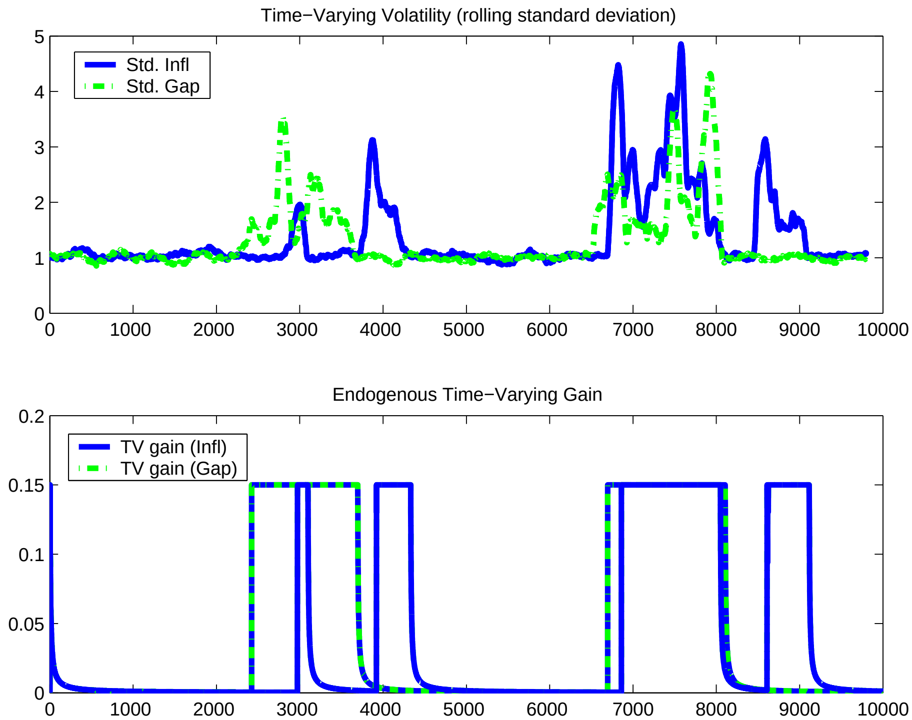

Figure 3. Time-Varying Volatility with Time-Varying Endogenous Gain Coefficient.

Notes: The economy is simulated for 13,000 periods, with agents using a window of 3,000 observations to compute the mean of past forecast errors. The calibration is as follows: $\beta = 0.99$, $\kappa = 0.05$, $\sigma = 0.1$, $\rho = 0.95$, $\chi_\pi = 1.5$, $\chi_x = 0.5$, $\rho_u = 0.5$, $\rho_g = 0.5$, $\bar{g}_\pi = \bar{g}_x = 0.15$.

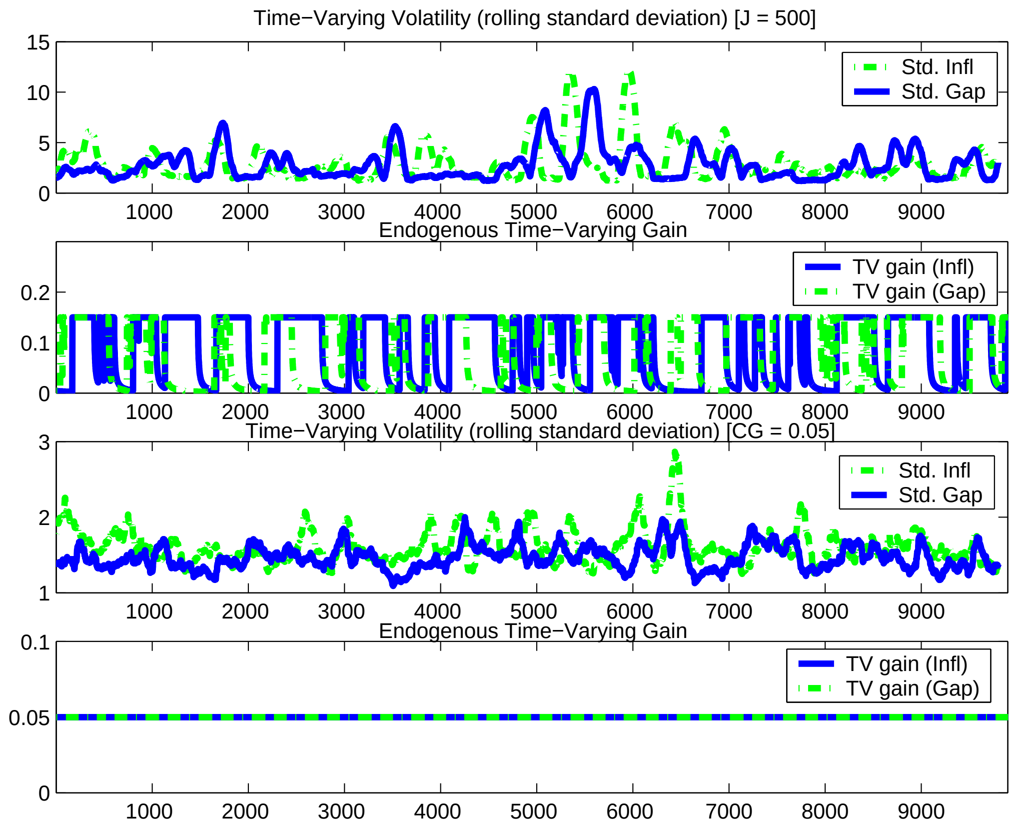

FIGURE 4. Time-Varying Volatility: Additional Cases.

Upper Plots: Time-Varying Endogenous Gain with $J = 500$ and $\overline{g} = 0.15$.
Lower Plots: Constant Gain with $\overline{g} = 0.05$.

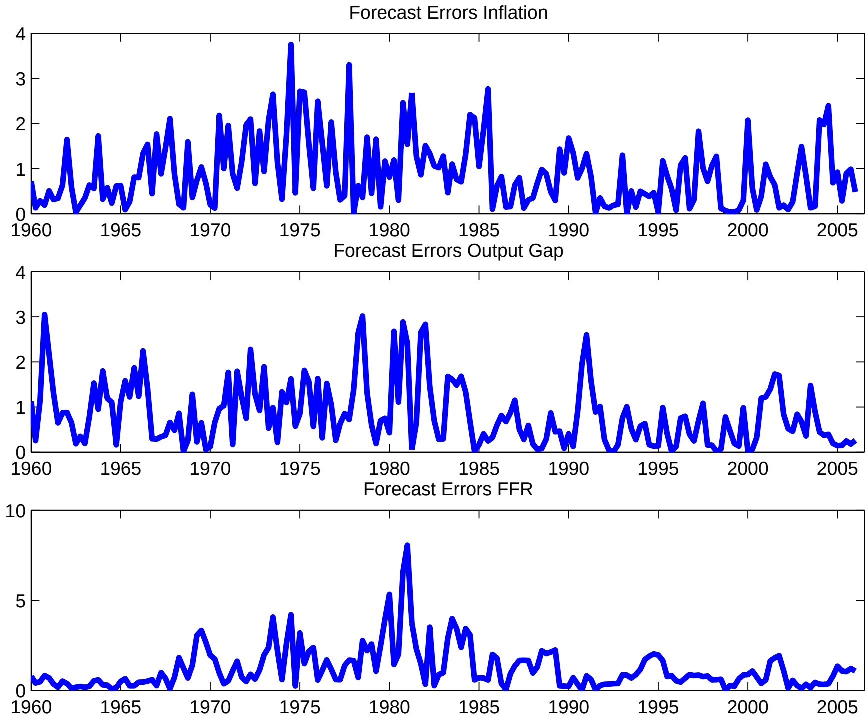

FIGURE 5. Forecast errors for inflation, output gap, and federal funds rate (absolute values, baseline estimation).

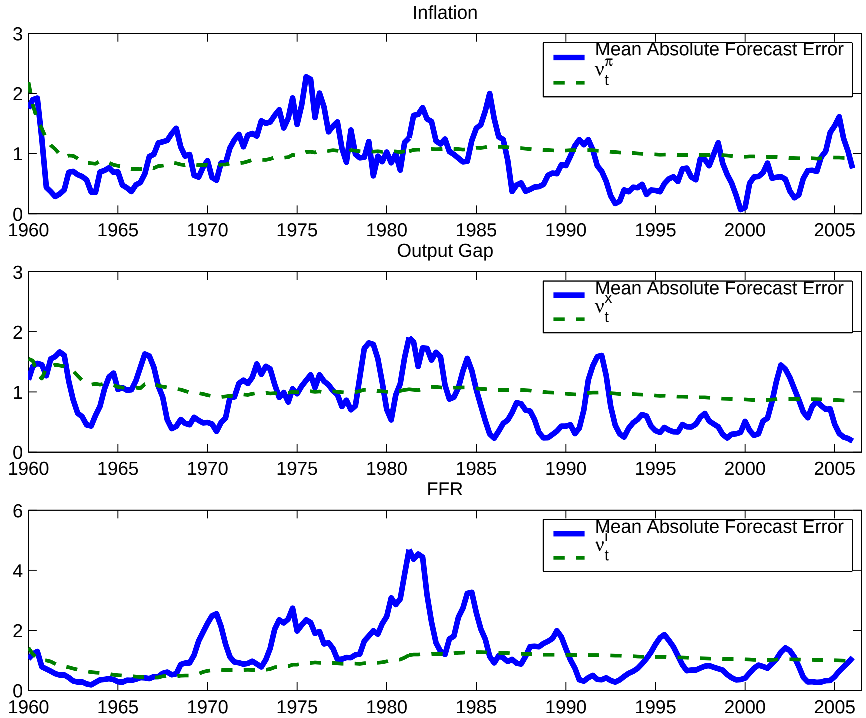

FIGURE 6. Rolling Mean Absolute Forecast errors vs. Updated ν_t for inflation, output gap, and federal funds rate series (baseline estimation).

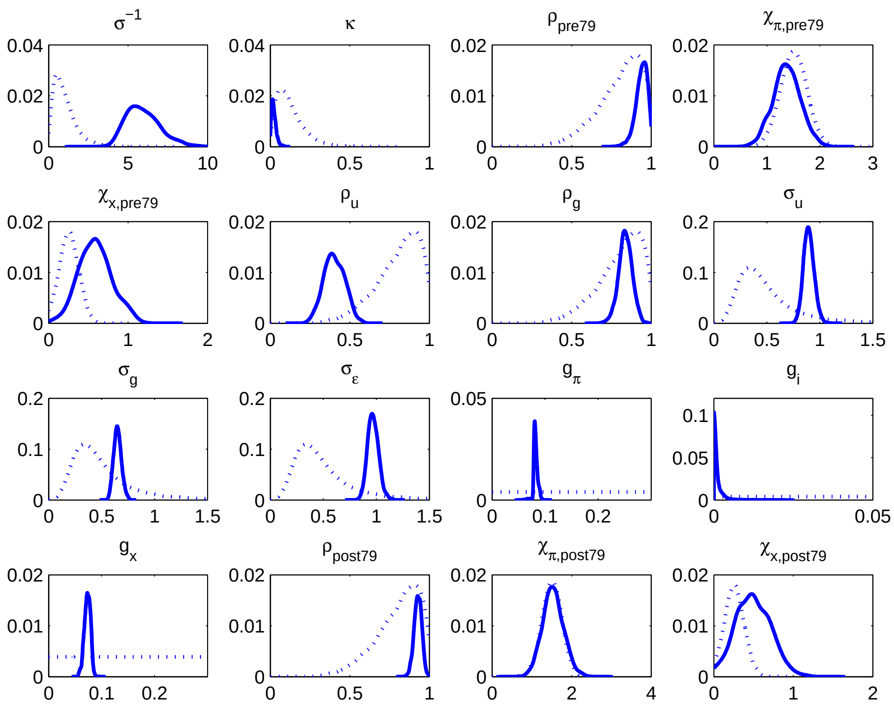

FIGURE 7. Prior and Posterior Distributions: all coefficients in benchmark estimation (the corresponding posterior means and 95% probability intervals are shown in Table 2).

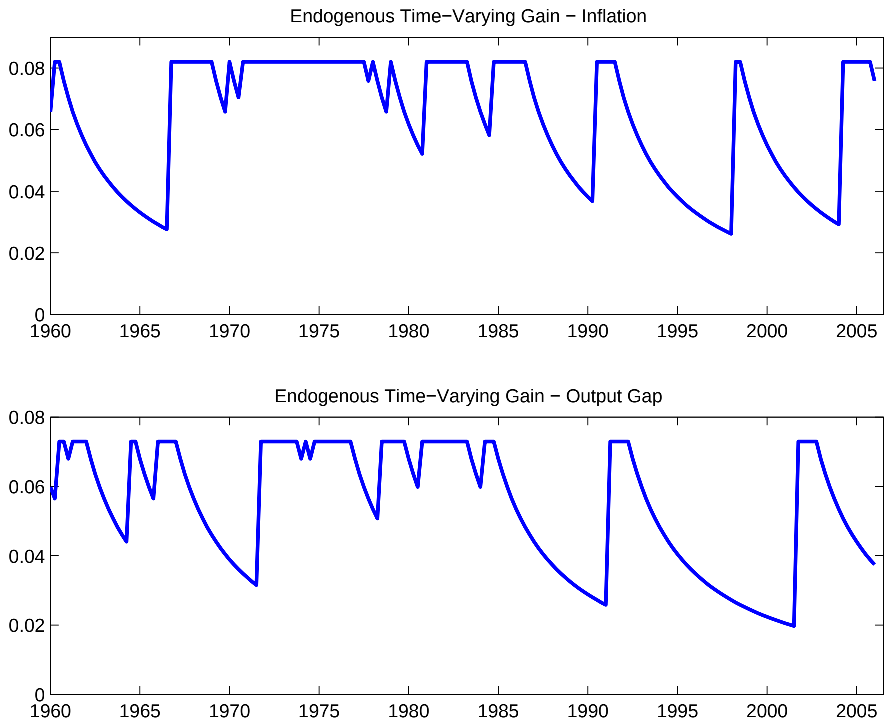

FIGURE 8. Endogenous Time-Varying Gain Coefficients (estimated constant gain): Baseline Estimation.

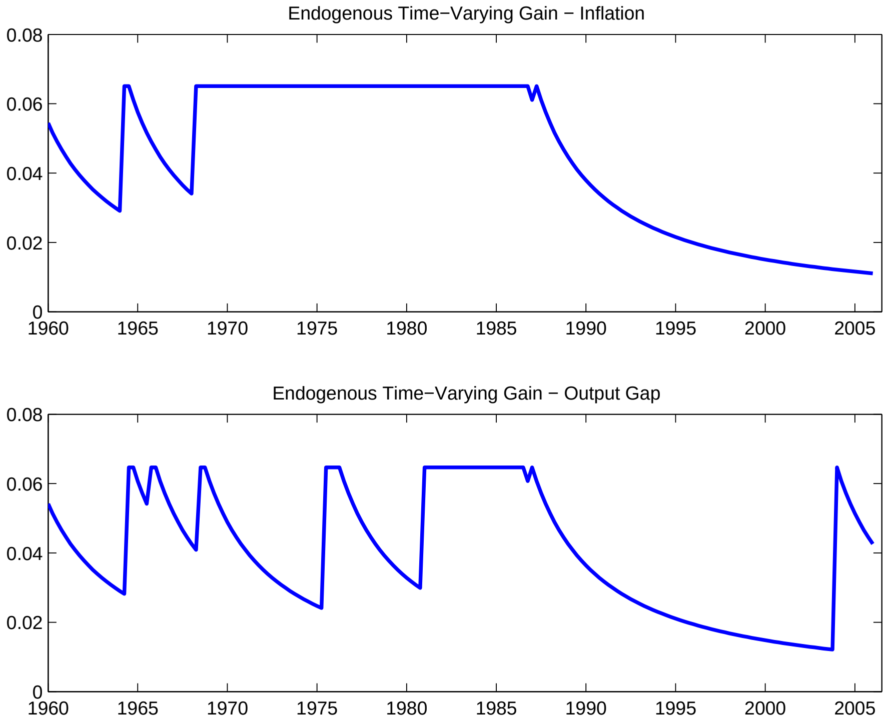

FIGURE 9. Endogenous Time-Varying Gain Coefficients (estimated constant gain): Estimation with $J = 20$.

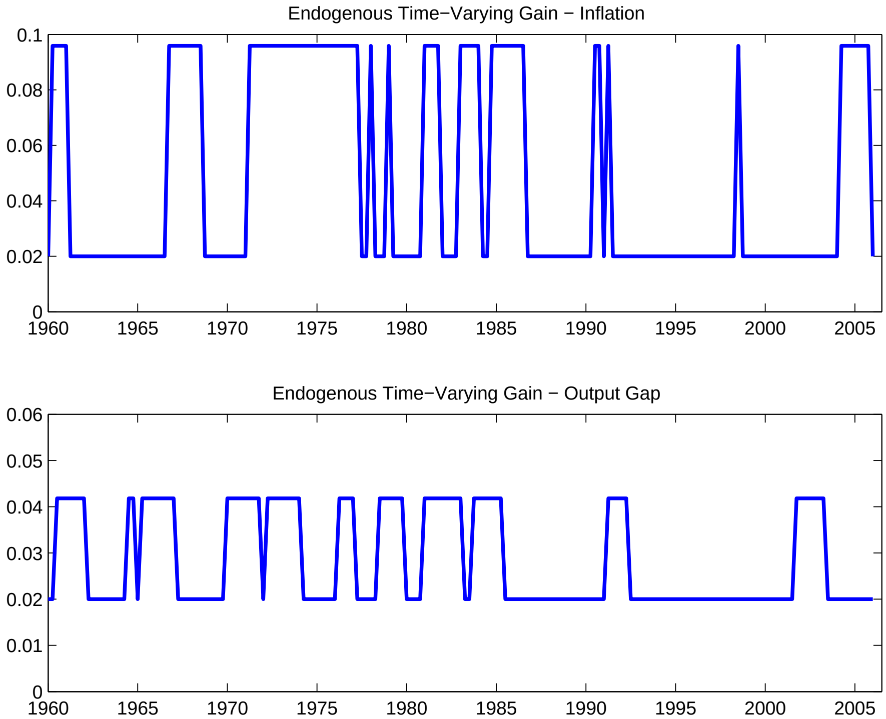

FIGURE 10. Endogenous Time-Varying Gain Coefficients: Estimation with gain coefficients switching between low and high constant gain values.

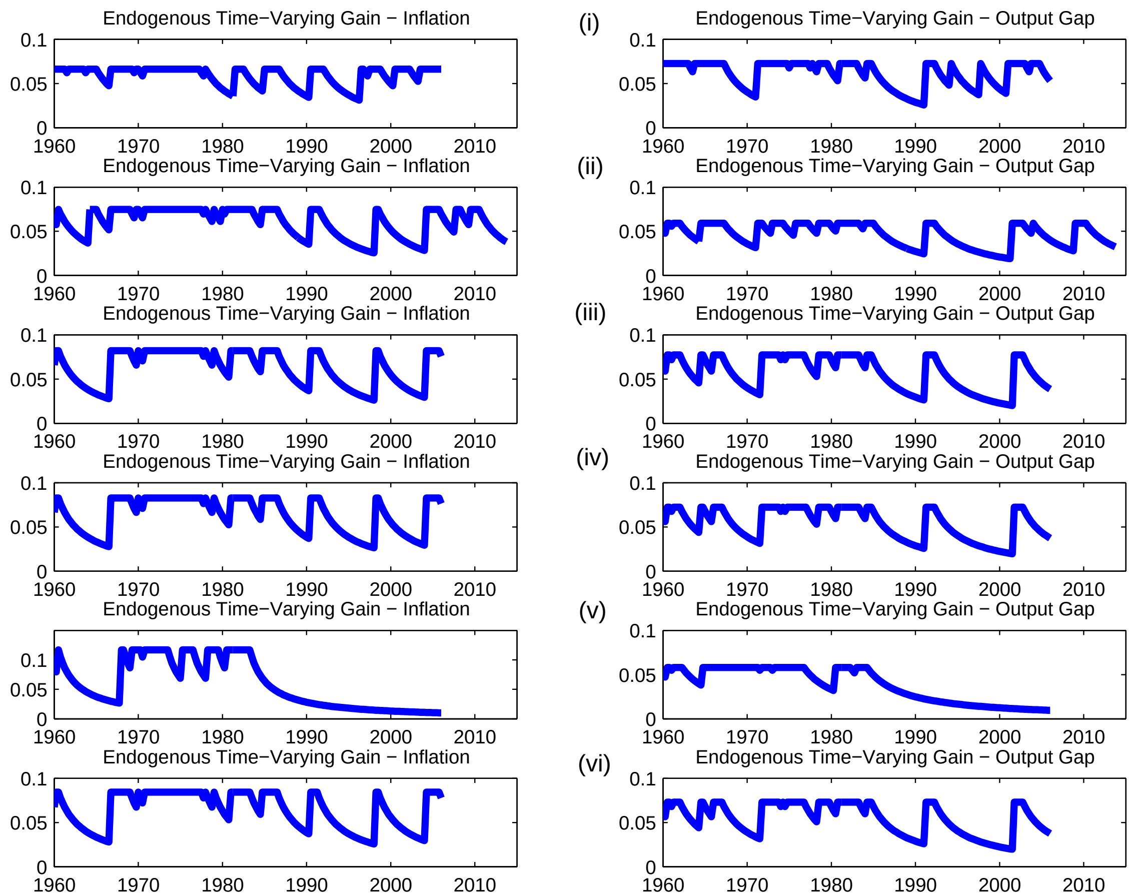

FIGURE 11. Endogenous Time-Varying Gain Coefficients: Robustness Checks.

Notes: The figure shows the evolution of the time-varying gain coefficients for inflation and the output gap over the sample, across different robustness checks. Panel (i) refers to the case with a rolling, rather than recursive, threshold for the historical forecast errors; panel (ii) refers to the estimation on the updated sample ending in 2014; panel (iii) refers to the estimation with the contemporaneous Taylor rule; panel (iv) refers to the estimation with uninformative priors for the policy rule feedback coefficients and a determinacy/E-stability prior; panel (v) refers to the case with a PLM that allows agents to observe structural disturbances; panel (vi) refers to the case of switching variances for the structural shocks.

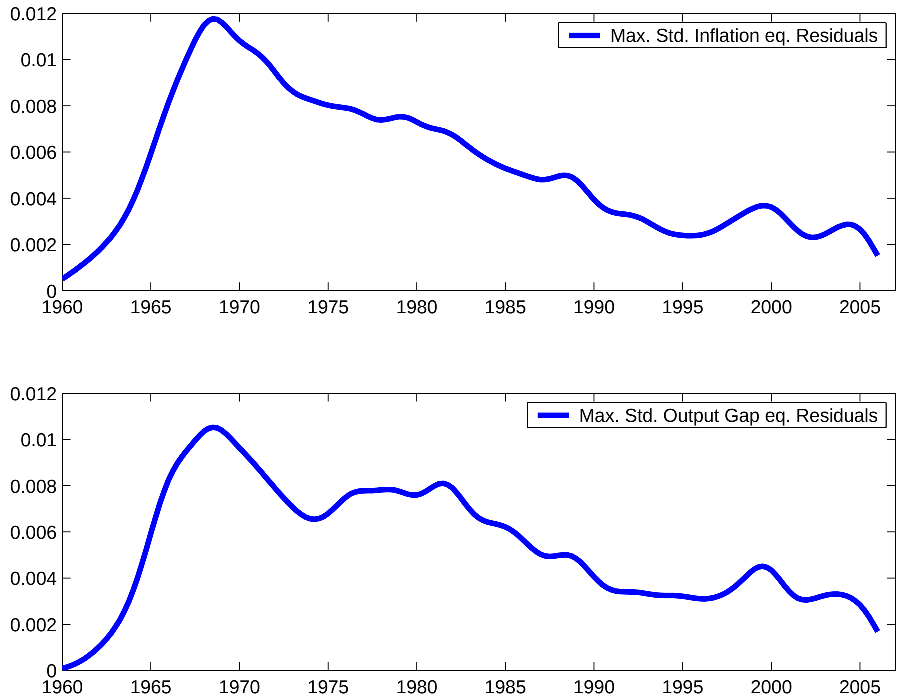

Figure 12. Kernel Density Estimation: sample observation in correspondence of which the Maximum Rolling Standard Deviation of residuals in the sample is obtained, across simulations.

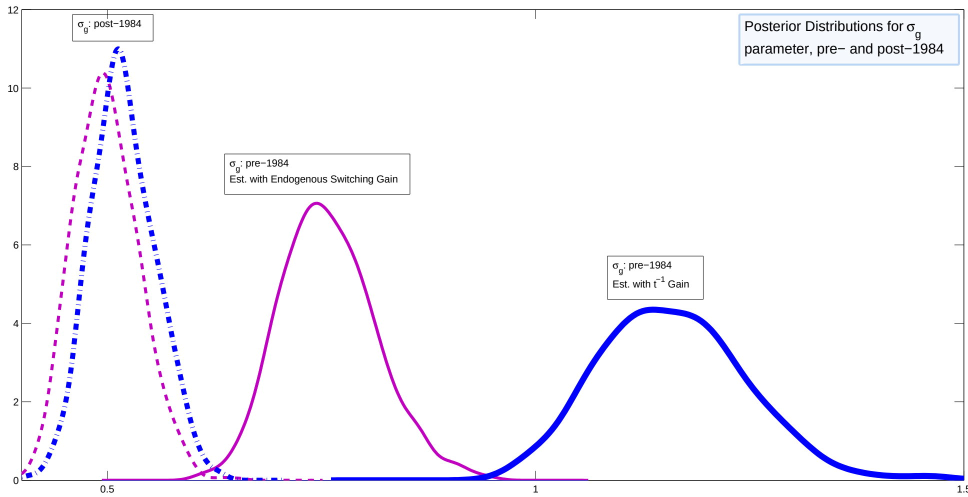

FIGURE 13. Posterior distributions: standard deviation coefficients for demand shock g_t.

Notes: the magenta lines refer to the estimation with the endogenous switching gain; the blue lines refer to the estimation with the endogenous switching gain shut down and replaced by learning with a decreasing gain. In each estimation, a solid line denotes the posterior distribution for the standard deviation before 1984, a dashed line denotes the posterior distribution for the standard deviation after 1984.

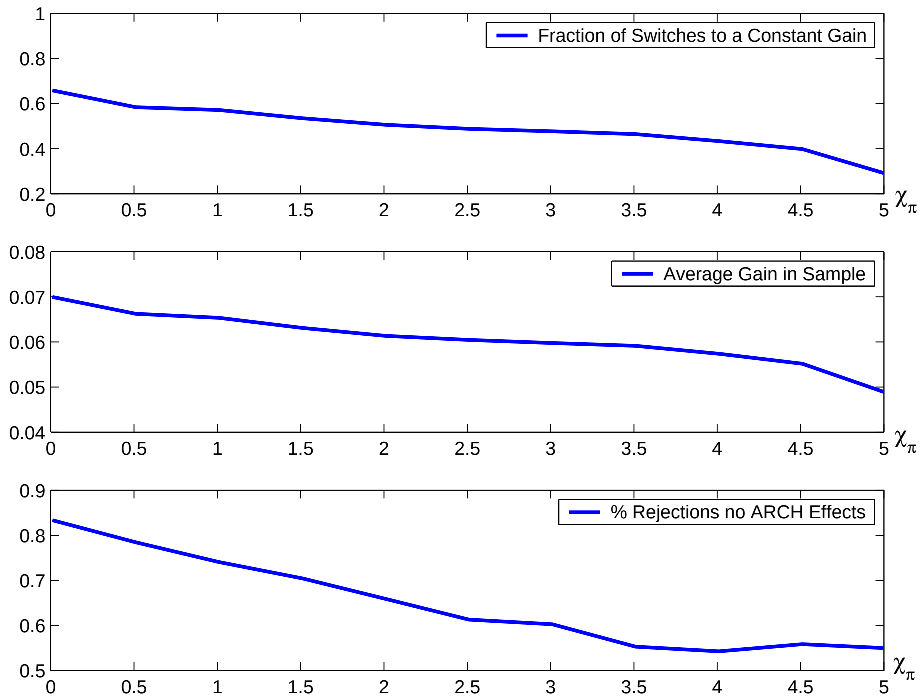

Figure 14. Effects of Monetary Policy on Volatility (Simulation under estimated parameters).

Graph 1: Fraction of sample periods in which agents' learning switches to constant gain as function of policy feedback to inflation χ_π; Graph 2: Average gain in the sample as function of policy feedback to inflation χ_π; Graph 3: Percentage of rejections of the null of no ARCH effects on the residuals as function of policy feedback to inflation χ_π.