# How Optimal is US Monetary Policy?

Xiaoshan Chen[a], Tatiana Kirsanova[b], Campbell Leith[b,*,†]

[a] University of Durham; [b] University of Glasgow

Received Date; Received in Revised Form Date; Accepted Date

---

## Abstract

Using a small-scale microfounded DSGE model with Markov switching in shock variances and policy parameters, we show that the data-preferred description of US monetary policy is a time-consistent targeting rule with a marked increase in conservatism after the 1970s. However, the Fed lost its conservatism temporarily in the aftermath of the 1987 stock market crash, and again following the 2000 dot-com crash and has not subsequently regained it. The high inflation of the 1970s would have been avoided had the Fed been able to commit, even without the appointment of Paul Volcker or the reduction in shock volatilities.

*Keywords:* Bayesian Estimation, Interest Rate Rules, Optimal Monetary Policy, Great Moderation

*JEL classification:* E58, E32, C11, C51, C52, C54

---

*Corresponding author. Address: ASBS, University of Glasgow, University Avenue, Glasgow, G12 0QL, UK; phone: +441413303702; email: campbell.leith@glasgow.ac.uk

†We would like to thank Davide Debortoli, Richard Dennis, Eric Leeper, Thomas Lubik, Charles Nolan, John Tsoukalas and seminar participants at the Riksbank, Universities of Birmingham and Surrey for helpful comments. All errors are ours.

## 1. Introduction

It is common practice to adopt a simple Taylor (1999)-type instrument rule to describe monetary policy when estimating DSGE models. This practice, however, is inconsistent with the claim of practitioners, that no central bank actually adopts such instrument rules, but rather prefer to set clear objectives and follow 'elaborate decision making-processes, in which huge amounts of data are elaborated and processed' (Svensson, 2003, pp. 428) in attempting to achieve those objectives. By specifying policy objectives the central bank adopts – using Svensson's terminology – a general targeting rule. This general targeting rule is then developed into a specific targeting rule by maximizing these objectives subject to the equations describing the decentralized equilibrium of the economy. The targeting rule that emerges is dependent on the degree of commitment the central bank possesses. What that degree of commitment is in practice, and whether or not we can develop data-coherent targeting rules, remains an open question, with the literature containing mixed results.

This paper considers various descriptions of policy – both instrument and targeting rules – and takes seriously the notion that policy making and the shocks hitting the US economy have been subject to shifts over the years. Doing so gives a far clearer indication as to which policy description best fits the data. This in turn has significant policy implications both in terms of designing monetary policy institutions and contributing to the debate on the source of the 'Great Moderation'.

The estimation demonstrates that the US monetary policy is best described by a time consistent targeting rule, labelled as discretion throughout the paper. This policy strongly dominates conventional simple instrument rules, as well as alternative forms of targeting rule with higher degrees of precommitment. This implies that during the post-WWII period the US Fed has not been making any credible policy commitments, either by following the Ramsey plan or following a simple instrument rule. The data also reveal that there have been changes in the Fed's degree of anti-inflation conservatism and in the volatilities of shocks hitting the economy. Ignoring these changes reduces the models' ability to fit the data and distorts the ranking of models.

The results imply that the inferences about shock processes, habit persistence and inflation indexation change significantly across different policy specifications. Under targeting rules, relative to instrument rules, we find that there is a shift in emphasis away from preference shocks towards cost-push shocks in driving the US business cycle. Under discretion this greater emphasis on cost-push shocks is not implausible, but is dramatic under commitment. Differences in the estimates of structural parameters under targeting rules further reflect the need to generate a meaningful policy trade-off, resulting in the degree of habits and inflation indexation being higher under commitment. In contrast, discretion tends to downplay the extent of habits to prevent implausibly aggressive policy responses to the associated externality.

The findings contribute to the literature in two respects. First, they add to the small but growing research on the empirical validity of targeting rules. While there are papers which estimate models under commitment (Adolfson et al., 2011 and Ilbas, 2010), discretion (Dennis, 2004) and an intermediate case of limited commitment, also known as quasi-commitment, as in Debortoli and Lakdawala (2016), very few compare the empirical relevance across these different targeting rules and with simple instrument rules.[1] In contrast to these papers, we consider a wide range of policy descriptions, and allow for potential regime switches in the monetary policy specification. Doing so explains how different policies interact with inferences about shock processes and structural parameters of the model.

Second, the analysis presented extends the 'good luck' and 'good policy' debate to the framework of targeting rules. There is a large literature on the 'Great Moderation' based on simple instrument rules, which finds that breaks in estimated policy rules (Lubik and

[1] Adolfson et al. (2011) find that commitment is preferred to a simple instrument rule using Swedish data. Givens (2012) and Le Roux and Kirsanova (2013) suggest that discretion is marginally preferred to commitment in the US and UK respectively.

Schorfheide, 2005, and Boivin and Giannoni, 2006), the implicit inflation target (Favero and Rovelli, 2003, Erceg and Levin, 2003 and Ireland, 2007) and/or the volatility of the underlying shock processes (Sims and Zha, 2006) help to explain the evolution of inflation dynamics across time. Given these findings, we allow for variation in the policy-maker's degree of anti-inflation conservatism, and for switches in the variance of the shock processes, when estimating different forms of targeting rule. The best-fitting model implies that US monetary policy is best described as being conducted under discretion, with an increase in central bank conservatism following the Volcker disinflation period, which is found to have occurred in 1982. More importantly, it identifies additional periods of policy change: the Fed relaxed policy temporarily in the aftermath of the 1987 stock market crash, and also lost conservatism following the 2000 dot-com crash, which it has never regained.

Finally, the counterfactual analysis using the best-fitting model suggests that the 'Great Moderation' in output and inflation volatility is due to both a reduction in shock variances and an increase in central bank anti-inflation conservatism. Decomposing the relative contribution of both effects implies that the far greater part of the 'Great Moderation' stems from the reduction in shock volatilities. More importantly, the counterfactuals show that inflation would never have breached 2% in the 1970s had the policy maker had access to a commitment technology. The potential gains from moving from discretion to commitment are substantial and dominate the gains from increasing central bank conservatism. Ensuring that the US Fed has access to commitment technologies and that they act to use such mechanisms is the 'good policy' that policymakers should focus on.

The plan of the paper is as follows. Section 2 outlines our model and the policy maker's preferences. The various descriptions of policy are discussed in Section 3. Section 4 considers data, priors and identification of the model, before presenting the estimation results in Section 5. Section 6 contrasts the results to those of Debortoli and Lakdawala (2016). Section 7 then undertakes various counterfactual simulation exercises which facilitate an exploration of both the sources and welfare consequences of the 'Great Moderation', and also an assessment of the potential benefits of further improvements in the conduct of monetary policy. Section 8 concludes.

## 2. The Model

The economy is comprised of households, a monopolistically competitive production sector, and the government. Full details of the underlying microfoundations of the model are given in the online Appendix A and only the linearized model is presented here.[2]

The household's optimization gives rise to the labor supply decision

$$\sigma \hat{X}_t + \varphi(\hat{y}_t - \hat{z}_t) = \hat{w}_t - \hat{\mu}_t, \tag{1}$$

and consumption Euler equation

$$\hat{X}_t = \mathbb{E}_t \hat{X}_{t+1} - \frac{1}{\sigma}\left(\hat{R}_t - \mathbb{E}_t \hat{\pi}_{t+1} - \mathbb{E}_t \hat{z}_{t+1}\right) - \hat{\xi}_t + \mathbb{E}_t \hat{\xi}_{t+1}, \tag{2}$$

where $\hat{X}_t$ is habits-adjusted consumption

$$\hat{X}_t = (1-\theta)^{-1}(\hat{y}_t - \theta \hat{y}_{t-1}), \tag{3}$$

and $\hat{y}_t$ denotes output, $\hat{w}_t$ is real wages, $\hat{\pi}_t$ is inflation and $\hat{R}_t$ is the nominal interest rate. Here $\sigma$ is the inverse of the intertemporal elasticity of substitution, $\varphi$ is the inverse of the Frisch elasticity and $\theta$ is the habit persistence parameter. The process $\hat{\mu}_t = \tau \hat{\tau}_t / (1-\tau)$ represents fluctuations in the labor income tax rate which serves as a cost-push shock, $\hat{z}_t$ is an innovation to non-stationary technology process which serves as a technology shock and $\hat{\xi}_t$ is a preference shock.

The firms' optimization decisions, in presence of both price and inflation inertia, give rise to a hybrid New Keynesian Phillips curve

$$\hat{\pi}_t = \chi_f \beta E_t \hat{\pi}_{t+1} + \chi_b \hat{\pi}_{t-1} + \kappa_c \hat{w}_t, \tag{4}$$

[2] An on-line Appendix contains information on the microfoundations of the model, solution algorithms, estimation and identification tests.

where the reduced form parameters are $\chi_f = \alpha/\Phi$, $\chi_b = \zeta/\Phi$, $\kappa_c = (1-\alpha)(1-\zeta)(1-\alpha\beta)/\Phi$, with $\Phi = \alpha(1+\beta\zeta)+(1-\alpha)\zeta$, where $1-\alpha$ is the Calvo (1983) probability of price change, $\beta$ is the households' discount factor and $\zeta$ is the proportion of firms setting prices who follow a backward-looking rule of thumb, rather than setting prices optimally.[3]

Hatted variables indicate that they have been linearized relative to their steady-states. The stationarity of the model's steady state is achieved by scaling by a non-stationary technology process discussed in Appendix A. The technology, cost-push and preference shocks follow AR(1) processes:

$$\hat{z}_t = \rho^z \hat{z}_{t-1} + \sigma_z \varepsilon_t^z, \qquad \varepsilon_t^z \sim N(0,1), \tag{5}$$

$$\hat{\mu}_t = \rho^\mu \hat{\mu}_{t-1} + \sigma_\mu \varepsilon_t^\mu, \qquad \varepsilon_t^\mu \sim N(0,1), \tag{6}$$

$$\hat{\xi}_t = \rho^\xi \hat{\xi}_{t-1} + \sigma_\xi \varepsilon_t^\xi, \qquad \varepsilon_t^\xi \sim N(0,1). \tag{7}$$

The model is then closed with one of the instrument or targeting rules considered in Section 3. The Fed's targeting rule can be inferred from their objectives.

## 3. Policy

The four basic forms of policy considered are a simple instrument rule and three types of targeting rule: discretion, timeless commitment and the intermediate case labelled 'quasi-commitment'. Across these alternative policies, the estimation permits changes in inflation conservatism by allowing Markov switching in instrument rule parameters, as well as in the relative weight given to inflation in the policy objective underpinning targeting rules, as detailed in this section.

[3] All parameters in this Phillips curve are assumed to be structural, see Gali and Gertler (1999).

### 3.1. *Instrument Rules*

The instrument rule is a generalized Taylor rule which, following An and Schorfheide (2007), is specified as

$$\widehat{R}_t = \rho^R \widehat{R}_{t-1} + (1-\rho^R)[\psi_1 \hat{\pi}_t + \psi_2(\Delta \hat{y}_t + \hat{z}_t)] + \varepsilon_t^R, \tag{8}$$

where the Fed adjusts interest rates in response to movements in inflation and deviations of output growth from trend.[4]

Within the framework of a generalized Taylor rule, potential changes in US monetary policy are accounted for by allowing for either changes in the Fed's inflation target or rule parameters. In the former case, following Schorfheide (2005), the measure of excess inflation in the Taylor rule, $\hat{\pi}_t$, removes the inflation target from the data, where that target follows a two-state Markov-switching process. In the latter case, when the policy changes are described as shifts in rule parameters $(\rho^R, \psi_1, \psi_2)$ between two regimes, the procedure developed by Farmer et al. (2011) is applied to solve the model.[5]

### 3.2. *Targeting Rules*

In the empirical analysis it is assumed that the Fed's objective function takes the micro-founded form, although the coefficients on the quadratic terms are freely estimated. Specifically, the empirical loss function can be written as

$$L = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left( \omega_1 \left( \hat{X}_t + \hat{\xi}_t \right)^2 + \omega_2 \left( \hat{y}_t - \frac{\sigma}{\varphi} \hat{\xi}_t \right)^2 + \omega_\pi \hat{\pi}_t^2 + \omega_3 \left( \hat{\pi}_t - \hat{\pi}_{t-1} \right)^2 \right), \tag{9}$$

see Appendix B for its microfoundations. This allows us to flexibly capture Svensson's (2003) notion of a general targeting rule by allowing the central bank to define the relative

[4] Rules of this form have not only been found to be empirically useful, but, when suitably parameterized, can often mimic optimal policy, see, for example, Schmitt-Grohe and Uribe (2007). Moreover, by allowing for an additional policy shock in the interest rate rule relative to the cases of optimal policy, we are further supporting the simple rule's ability to fit the data. As we shall see, despite this, discretionary policy is 'strongly' preferred by the data.

[5] The details of the solution algorithm are provided in Appendix C.

importance of welfare-relevant terms. Strictly speaking, it should not be interpreted as a welfare function unless the estimated coefficients coincide with the microfounded weights.

Given that much of the literature on estimated instrument rules finds that there have been significant changes in the conduct of policy over time, targeting rules derived under an assumption of unchanging policy maker preferences may be too stylized to capture such changes. Therefore, the relative weight on inflation, $\omega_\pi$, is allowed to be subject to regime switching between 1 and a value lower than 1 to capture policy regimes with lower conservatism. The estimation can therefore assess whether or not the Fed's attitudes to inflation targeting have varied over time. For example, has monetary policy been more conservative since the Volcker disinflation? Moreover, accounting for independent regime switching in the variances of shocks, $\sigma_z, \sigma_\mu$, and $\sigma_\zeta$ helps to assess whether the lower interest rates observed during 2001-2007 were due to economic conditions, or the result of the Fed putting less emphasis on inflation targeting relative to its other objectives.

When implementing targeting rules, the central bank selects interest rates to minimize loss function (9) subject to the structural equations describing private sector behavior, equations (1)-(4), and the evolution of shocks. The targeting rules considered include the standard cases of discretion and timeless commitment, which are the two polar cases of how well the central bank can manage the expectations of the private sector. Under timeless commitment the policy maker can make credible promises about the setting of the policy instrument in future periods, while under discretion they re-optimize and are expected to re-optimize in each period. This implies that under timeless commitment there is a history-dependence in policy making arising from these past commitments, which is absent under discretion. The empirical implementation of timeless commitment assumes that the targeting rule has been in place for a prolonged period.

The remaining form of targeting rule is quasi-commitment, as developed in Schaumburg and Tambalotti (2007) and Debortoli and Nunes (2010). The policymaker may deviate from commitment-based plans with a fixed exogenous probability, known by all agents. The current policy maker forms a commitment plan to be followed until randomly ejected, with a given probability, from office. At which point a new policy maker will be appointed, and a new plan formulated until that policy maker is, in turn, removed. Therefore, the central bank can neither completely control the expectations of the private sector, nor perfectly coordinate the actions of all future policy makers. This implies that, in contrast to the cases of discretion and timeless commitment, in each period there is a policy surprise resulting from the fact that expectations are formed as a probability-weighted average of policy with and without reneging, while actual policy will either renege or not. Such policy surprises imply that outcomes under quasi-commitment are not a probability-weighted average of those under discretion and timeless commitment.

The procedure described by Svensson and Williams (2007) is used to solve for the equilibrium dynamics under discretion and timeless commitment with Markov-switching in objectives.[6] In addition, this solution method is modified to incorporate the case of quasi-commitment, as Schaumburg and Tambalotti (2007) and Debortoli and Nunes (2010) do not allow for Markov switching in objectives. Appendix C presents the new algorithm.

## 4. Data, Priors and Identification

The empirical analysis uses US data on output growth, inflation, and nominal interest rates from 1961Q1 up to 2008Q3, just before nominal interest rates were reduced to their effective lower bound of 0.5% and the first round of quantitative easing was implemented. The data used in the estimation are plotted in Figure 3, alongside various counterfactual

[6] The Svensson and Williams (2007) algorithm implies that although policy makers can anticipate any changes in their objectives, they do not attempt to tie the hands of their future selves by altering today's policy plan as part of a strategic game, instead they set today's policy cooperatively with their future selves. We consider that this algorithm is in line with the conduct of US Fed policy as there may be some evolution in the consensus surrounding the objectives of monetary policy. However, in other policy making environments, where interest rate decisions are made by partisan politicians who may alternate in office, this would be less defensible and the approach of Debortoli and Nunes (2010) would be applicable.

simulation results which will be discussed below. The estimation strategy is standard and is described in Appendix D.

The priors are presented in Table 1. These are set to be broadly consistent with the literature on the estimation of New Keynesian models, in particular for the structural model parameters we follow Smets and Wouters (2003). For the Markov-switching instrument rule parameters, in line with Bianchi (2013), the priors for the response to output growth and the smoothing term are set to be symmetric across regimes, while asymmetric priors are chosen for the response to inflation.[7] For targeting rules, the relative weights (i.e. $\omega_1, \omega_2,$ and $\omega_3$) on the objective function are assumed to be distributed following beta distributions and $\omega_\pi$ is allowed to switch between 1 and a value lower than 1, where the beta distribution is used for the latter with a mean of 0.5. The prior for the probability of reneging on past promises under quasi-commitment policy, $\upsilon$, follows Debortoli and Lakdawala (2016) with a uniform prior on the interval [0,1]. The parameters, $\gamma^Q$, $\pi^A$ and $r^A$ represent the values of output growth, inflation and interest rates, respectively, when the economy is in its steady state. The prior means of $\gamma^Q$, $\pi^A$ and $r^A$ are set to be broadly consistent with their data averages during this pre-sample period from 1950Q1 to 1960Q4. Parameter $\pi^A$ is interpreted as an inflation target, and it is assumed to be constant for all models except the instrument rule model with Markov-switching inflation target, where the priors for $\pi^A$ are set in line with Schorfheide (2005). The average real interest rate, $r^A$, determines the discount factor, $\beta = \left(1 + r^A/400\right)^{-1}$.

[Table 1 around here]

Finally, it is important to note that all model parameters are identifiable. To demonstrate this, the identification tests of Komunjer and Ng (2011a) and Koop et al. (2013) were applied to the models which feature both policy and volatility switches. In all cases model parameters

[7] This way of setting priors for the switching parameters is also discussed by Davig and Doh (2014), as a means of introducing a natural ordering of regime-dependent parameters in order to avoid the potential risk of 'label switching', as noted in Hamilton, Waggoner, and Zha (2007).

are identified, see details in Appendix E. This is in contrast to the identification of parameters in larger models, (see the application of these tests to the Smets and Wouters model in Iskrev (2010), Caglar et al. (2011) and Komunjer and Ng (2011b), respectively) which is one reason why we prefer to work with a simpler model.

## 5. Results

This section presents the results of the estimation. It begins by identifying which description of policy best fits the data. It then discusses the implications of this for inferences about structural parameters of the economy, which shocks drive the business cycle in the US, and whether the Fed's preferences have changed over time.

### *5.1. Policy, Structural Parameters and Shocks*

The posterior means and the 90% confidence intervals are presented in Table 2 where each column corresponds to an alternative policy description, and these columns are ordered according to the log marginal likelihood values calculated using Geweke (1999) and Sims et al. (2008), respectively.

[Table 2 around here]

The first column of results in Table 2 is for the best-fitting model, which is discretionary policy. Following Kass and Raftery (1995) the evidence in favor of discretion relative to instrument rules with switches in rule parameters is identified as 'strong', and relative to timeless commitment as 'decisive'. The probability of reneging on policy promises under the quasi-commitment policy is $\upsilon = 0.29$, which implies that the commitment plan is expected to last for just 10 months. These estimates suggest that the discretionary form of targeting rule best fits the data, and there is no evidence of any commitment behavior on the part of the Fed.

The estimates obtained under the conventional instrument rule are broadly in line with other studies: an intertemporal elasticity of substitution, $\sigma = 2.9$, a measure of price stickiness, $\alpha = 0.77$, implying that price contracts typically last for one year; a relatively modest degree of price indexation, $\zeta = 0.09$, a sizeable estimate of the degree of habits, $\theta = 0.83$ and an inverse Frisch labor supply elasticity of $\varphi = 2.4$. Moving to the case of discretion, these deep parameter estimates remain largely the same, except that there is a significant decline in the degree of habits in the model, which falls to $\theta = 0.39$, and a modest increase in the degree of indexation in price setting to $\zeta = 0.16$. The quasi-commitment policy delivers similar values for these parameters. However, with a further increase in the degree of precommitment to the case of strict timeless commitment, the degree of indexation rises to $\zeta = 0.26$, while the extent of habit persistence increases to a level closer to that observed under instrument rules, $\theta = 0.69$.

These differences in the estimated structural parameters across targeting rules reflect the need to ensure the policy maker faces a meaningful trade-off. In the benchmark New Keynesian model it is only cost-push shocks which present a trade-off between output and inflation stabilization for the policy maker. All other shocks would result in policy responses which perfectly stabilize inflation. Introducing a habits externality breaks this ‘divine coincidence’ and implies other shocks will matter to the policy maker. Therefore, in order to explain the observed volatility in inflation, the estimation under timeless commitment retains the degree of habits relative to instrument rules. This increases the ability of shocks, other than the cost-push shock, to generate inflation volatility. In such an environment the degree of inflation indexation is also likely to affect these policy trade-offs.

The case of discretion is more subtle. The inability to commit to a small but sustained response to shocks implies that in the presence of the habits externality the policy maker will react aggressively to such shocks, see Leith, Moldovan, and Rossi (2012). This would imply higher interest rate volatility than is observed in the data. Therefore, the estimation downplays the extent of habits under discretion, relative to timeless commitment.

In addition to variations in the degree of habits and inflation indexation across the estimates obtained under targeting rules, the balance between different shocks also changes. Again, this reflects the need to generate meaningful policy trade-offs in order to explain the inflation volatility observed in the data. Therefore, we see a reduction in both the persistence and standard deviation of preference shocks under targeting rules relative to instrument rules. At the same time, the persistence and standard deviation of cost push shocks increase, dramatically so in the case of timeless commitment. However, it is important to note that under discretion the unconditional variance of this shock is not dissimilar to those found in other studies employing instrument rules as their description of policy.[8]

To summarize, relative to conventional instrument rules, our preferred targeting rule adjusts structural and shock parameter estimates to create a meaningful trade-off for policy when explaining macroeconomic volatility. This includes a shift from preference to cost push shocks in explaining the US business cycle.

### *5.2. Inflation Conservatism*

The results suggest that the Fed's stance on inflation targeting has varied over the sample period. Taking into account potential switches in shock volatilities, for each policy specification, the estimation identifies two distinct inflation targeting regimes with a different degree of conservatism. We label them 'more' and 'less' conservative regimes, depending on the size of the weight on inflation, $\omega_\pi$, under targeting rules. Under all targeting rules, $\omega_\pi$ is more than halved in the less conservative regime from the default level of one in the more conservative regime.

[8]It should be noted that the cost-push shock enters the Phillips curve with the reduced form coefficient $\kappa_c$, which lies in the range 0.036-0.065 across our estimates. Calculating the unconditional variance of the normalized cost-push process $\kappa_c \hat{\mu}_t$ for discretion implies that the variance of 0.002 and 0.017 in low and high volatility regimes, respectively, is lower than that estimated by Smets and Wouters, 2003 for a single volatility regime (0.0217). For the case of quasi-commitment the corresponding numbers are 0.0012 and 0.014. However, commitment requires substantial increases in the unconditional variance of the cost push shock to 0.16 and 0.65 for the low and high volatility regime, respectively.

As for instrument rules with either Markov-switching rule parameters or inflation targets, a 'less' conservative inflation regime can be also identified by observing a reduction in the size of the coefficient on excess inflation, $\psi_1$, or an increase in inflation target, $\pi^A$, respectively. In the former case, although policy satisfies Taylor principle across both regimes, $\psi_1$ falls from 2.124 to 1.219, while for the latter case, $\pi^A$ rises from 3.34% to 4.33%.

We now explore when these less conservative inflation regimes were estimated to have occurred. Figure 1 plots the smoothed probabilities of being in the less conservative targeting regime, as well as being in the high volatility regime. In the case of quasi-commitment, the plot also shows the probability that the policy maker has reneged on previous commitments.

[Figure 1 around here]

The best-fitting model, discretion, provides more information than the instrument rule-based models on the conduct of monetary policy over recent years, as the smoothed probabilities show. The estimation finds the relaxation of monetary policy in the 1970s that is well documented in the existing literature following Clarida et al. (1998). However, unlike the vast majority of the literature our estimates date the Volcker disinflation as occurring in 1982 rather than 1979.[9] Additionally, the smoothed probabilities from this model also suggest that policy was relaxed briefly following the stock market crash of October 1987. More interestingly, a prolonged reduction in the Fed's weight on the inflation target is identified as occurring at the time of dot-com crash and persisting all the way through to the financial crisis. Such a pattern is not so apparent in the instrument rule-based models. Similarly, the less conservative policy episodes are largely confined to the mid to late 1970s under timeless commitment. Quasi-commitment utilizes two mechanisms to capture a relaxation in the Fed's anti-inflation stance. Specifically, we may observe a reduction in the weight attached to inflation stabilization in the objective function (lost conservatism) or periods

[9] More recent papers also find that the date of the Volcker disinflation is later than previously thought. See, for example, Bianchi (2013), Schorfheide (2005).

of reneging on past policy commitments. Relative to discretion, quasi-commitment relies on extensive periods of lost conservatism to such an extent that it is easier to define when conservatism was not lost under this policy description – briefly in the early 1980s and a few years prior to the bursting of the dot-com bubble – and even then, not fully. In addition, the quasi-commitment estimates imply that the Fed reneged on policy commitments relatively frequently in the 1970s, and was showing signs of having possibly done so in the lead up to the financial crisis too.

### *5.3. The Importance of Switches in Policy and Volatilities*

Turning to explore how important accounting for both the switches in policy and shock volatilities are for our estimated results, Table 3 re-estimates our models without allowing for either form of switching.[10] In this case, the simple instrument rule is preferred by the data, but only marginally. This is because targeting rules are heavily penalized by being prevented from accounting for the less conservative policy in the 1970s. The ranking amongst targeting rules also changes: quasi-commitment is preferred to discretion with timeless commitment struggling to fit the data. The apparent superiority of quasi-commitment relative to other forms of targeting rule is due to the presence of policy surprises. Without allowing for switches in shock volatilities these policy surprises, largely identified during the 1970s, serve as an additional shock to increase the ability of the model to fit the data. Once switches in shock volatilities are introduced in Table 2, quasi-commitment loses this advantage over instrument rules and discretion.

[Table 3 around here]

Introducing the possibility of policy switches, but not switches in the volatility of shocks, highlights several interesting features of the benchmark estimates that would be otherwise

[10] Here we only present selected parameters. The complete set of parameter estimates is given in Appendix G.

missed, see Table 4.

[Table 4 around here]

First, the ranking of policies changes again: the policy switches can account for the less conservative regime in the 1970s enabling discretion and quasi-commitment to dominate instrument rules.

Second, the differences between the less and more conservative regimes are greater than in the case in Table 2, where the switches in shock volatilities are present. Without volatility switching, as shown in Table 4, the instrument rule does not satisfy the Taylor principle in the less conservative regime. This mirrors the findings of Sims and Zha (2006) who warn of the biases that may be introduced by failing to account for heteroscedasticity in the error terms. For the instrument rule with switches in the inflation target, the differences in the targets are also widened. Similarly, for the targeting rules, the relative weight on inflation falls by more across all policy descriptions in the less conservative regime. These results support a generalization of the arguments in Sims and Zha (2006) that failure to account for shifts in shock volatility may overstate the apparent weakness in policy during certain periods.

Third, not including switching in shock volatilities also leads to a loss of nuance in the identification of periods with less conservative regime under discretion. Without volatility switches all policy descriptions pick up the high inflation in the 1970s as being the result of a less conservative targeting regime, and that this episode ends with the Volcker disinflation somewhere between 1979 and 1982, see Figure G1 in Appendix G. However, when we combine volatility shifts with policy shifts there are additional periods where the Fed appears to have lost conservatism.[11] These are often associated with well known periods of stock market volatility, specifically in 1987 and following the bursting of the dot-com bubble.

[11] There are less extensive periods of reduced conservatism under quasi-commitment when we do not allow for switches in shock volatilities. In essence, the less conservative regime under quasi-commitment allows the estimation to accommodate higher shock volatilities without inducing an overly aggressive and therefore data-incoherent policy response during reneging periods.

To summarize, with no switching in objectives the targeting rules find it more difficult to account for the inflation of 1970s than instrument rules. Adding switches in policy objectives results in discretion dominating all other forms of policy, see Table 4. Allowing for switches in shock volatilities, policy surprises generated by quasi-commitment policy become relatively less effective in explaining the data. As a result, quasi-commitment moves further down in the ranking of the data-preferred policies as shown in Table 2.

## 6. Comparison with Debortoli and Lakdawala (2016)

Our estimates imply that discretion dominates all other descriptions of policy. This is in contrast to the conclusions of Debortoli and Lakdawala (2016) who argue that the data reject both discretion and timeless commitment, preferring quasi-commitment. They reach this conclusion based on the fact that the estimated probability of reneging on past promises does not tend to either zero or one in estimation. This section seeks to explore the reasons underpinning the apparent disparity in conclusions.

The first thing to note is that our estimates of the probability of reneging on past policy commitments are not dissimilar to theirs. However, the fact that the estimates do not tend to the limiting case of discretion does not imply that quasi-commitment dominates discretion in terms of its ability to explain the data. Instead, the Bayes factor implies that discretion is decisively preferred to quasi-commitment. The reason for this is that the quasi-commitment model is not actually an intermediate case lying between the cases of timeless commitment and discretion, as discussed before in Section 3.2. Instead, it introduces policy surprises – serving as a new kind of policy shock – which arise from the fact that economic agents form expectations based on the probability of experiencing a reneging regime in the next period. The realization or otherwise of the reneging regime is then always a shock relative to these expectations. When the probability of reneging is low, economic agents expect the policy maker to keep their promises so that reneging offers the policy maker the opportunity to exploit those expectations generating a sizeable policy shock. Conversely, when there is a high probability of reneging, the policy maker makes more extreme policy promises to retain a desirable influence over expectations which, in turn, imply a large policy shock whenever the policy maker keeps that promise (see Schaumburg and Tambalotti, 2007). The estimated probability of reneging needs to balance these two scenarios to produce policy shocks that match the volatility in the data. As discussed in Section 5.3, once switches in shock volatilities are allowed, there is less need to rely on such policy shocks to fit the data.[12]

Finally, we can check that our results are not driven by adopting an objective function which takes the form of the microfounded objective function (9) rather than the simpler specification used in Debortoli and Lakdawala (2016). We consider two forms of *ad hoc* objective function based on

$$L = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left( \hat{\pi}_t^2 + \omega_y \hat{y}_t^2 + \omega_R \left( \Delta \hat{R}_t \right)^2 \right). \tag{10}$$

Loss function type I excludes the term in the interest rate smoothing, $\omega_R = 0$, and only retains terms in inflation and the output gap. Loss function type II allows the interest rate smoothing term $\omega_R$ to be estimated.

[Table 5 is around here]

In Table 5 we compare four policies, all excluding Markov switching in policy objectives but including switching in shock volatilities, as in Debortoli and Lakdawala (2016). Discretion with objectives in a microfounded form dominates the three cases of quasi-commitment, with type I and II ad hoc objectives and the case with objectives in a microfounded form (9) used throughout the paper.

Three clear messages emerge from this comparison. First, with objective function (9) discretion dominates quasi-commitment. Again, this confirms that when switches in shock

[12] The quasi-commitment estimation also adds complexity to the model in the form of an additional estimated parameter, the need to estimate the probability that we have observed a reneging regime in each period and the scale of the state-space representation of the model relative to the discretionary case. This complexity is penalized in the construction of the Bayes factors.

volatilities are accounted for the policy-surprise shocks generated by quasi-commitment are less effective in fitting the data.

Second, adding the interest rate smoothing term significantly raises the ability of quasi-commitment to fit the data. Quasi-commitment using the type II ad hoc objective achieves a better fit compared to the other two cases of quasi-commitment. However, the estimated weight on the smoothing term is implausibly large, $\omega_R = 1.5$ as the estimation seeks to limit the sharp movements in the policy instrument implied by the policy shocks described above. If we remove the interest rate smoothing term (type I ad hoc), then such policy results in the worst fit out of the four cases considered in Table 5.

Third, quasi-commitment policies with ad hoc welfare objectives identify similar probabilities of reneging and periods of high volatility as the quasi-commitment policy presented in Table 2, see Figure G2 in Appendix G.

## 7. Counterfactuals

The best-fitting model is obtained under discretionary policy with Markov switching in the weight on inflation target in the policy maker's objectives, as well as switches in the volatility of shocks hitting the economy. This allows us to undertake various counterfactual exercises. For example, exploring what the outcomes would have been if shock volatilities had not declined in the 1980s, or what would have happened had the Fed adopted a tougher anti-inflation stance in the 1970s. Moreover, this section explores how much further economic outcomes would have improved had the policy maker not only adopted tougher anti-inflation policies in the 1980s, but also been able to act under timeless commitment.

### *7.1. Good Luck*

The series of counterfactuals begins by analyzing the role of good luck in stabilizing US output and inflation. To do so the pattern of switches in policy regimes is fixed as estimated, but the counterfactual sets the volatility of shocks at their high or low values. The estimated shocks are therefore re-scaled by the relative standard deviations from the high and low volatility regimes. Panel A of Figure 2 plots the actual and counterfactual series for inflation, interest rates and output growth. We can see that the high volatility of shocks plays a significant role in raising inflation during the 1970s. In the absence of these high volatility shocks, inflation would never have risen above 5%. In addition, it is apparent that output growth fluctuations could have been dampened if policy makers had had the 'good luck' of experiencing the low shock volatility regime during the 1970s and early 1980s. Moreover, it is also notable that under the policy regimes estimated in the post-Volcker period, inflation and output fluctuations would not have changed too dramatically regardless of the magnitude of shocks. This may be an indication that tougher anti-inflation policies in the 1980s helped stabilize the US economy.

[Figure 2 around here]

### *7.2. Conservative Monetary Policy*

The second set of counterfactual analyses assesses the impact that increased conservatism would have had on US inflation and output, especially during the 1970s. To simulate the set of counterfactual variables we subject the economy to the sequence of estimated shocks, but set the weight on inflation in the policy maker's objective function, $\omega_\pi$, to either its default value of one in the more conservative regime, or to 0.436 in the less conservative regime, throughout the sample period. The first two pictures in Panel B of Figure 2 plot the actual and counterfactual series for inflation and interest rates. The third picture plots the output loss, which is the difference between model implied output with estimated objective function weights and the counterfactual output when the policy maker is more conservative.

Panel B of Figure 2 shows that even if the Fed had adopted a tougher anti-inflation stance in the 1970s, it would not have been able to completely avoid higher inflation, but observed inflation would have been significantly lowered at a cost of higher output losses. Similarly, the two periods of rising inflation that occurred following the stock market crash of 1987 and the bursting of the dot-com bubble could also have been mitigated if the Fed had maintained its stance on inflation targeting. The counterfactual paths for interest rates largely reflect the tightness or slackness of policy implied by the alternative scenarios. However, since the effective stance of monetary policy is reflected in the real interest rate, the path for nominal interest rates under the less conservative policy are above those implied by the more conservative policy, reflecting the latter's success in controlling inflation.

### *7.3. The Value of Commitment*

Finally, Panel C of Figure 2 assesses the implications of moving from discretion to timeless commitment. Both the shock volatility and policy switches follow their estimated realizations, but we change whether or not the policy maker has access to a timeless commitment technology. The results are striking. If the Fed had been able to make credible policy commitments in the 1970s, even although it was subject to high volatility shocks and had a reduced weight on the inflation target in that period, inflation would have remained below 2% throughout the sample period. Although it appears that there would have been non-trivial losses in output with a peak loss of around 1% by the mid 1970s, the welfare analysis in the next section suggests that these losses are more than compensated for by the reduction in inflation volatility.

### *7.4. Welfare Analysis*

In addition to providing the counterfactual figures above, it is insightful to compute the unconditional variances of key variables and the value of unconditional welfare (using both the estimated policy objective (9) and the fully microfounded objectives where the weights are microfounded functions of the estimated structural parameters of the model) under alternative counterfactuals.

As a benchmark case we consider the worst case scenario where the economy is permanently in the high shock volatility regime and adopt a less conservative policy with $\omega_\pi = 0.436$ under discretion. We can then consider the extent to which 'good policy' or 'good luck' alone would be able to stabilize inflation, output and interest rates and improve welfare.

Table 6 presents variances of output, inflation and interest rate under different conservatism – volatility scenarios. The degree of conservatism ranges from that estimated under the less conservative through the more conservative regimes, both of which are using estimated policy objective weights, to the extreme level implied by the fully microfounded welfare function. Two welfare metrics are used to measure losses, one with estimated weights and the other with microfounded weights.

[Table 6 around here]

Panel A in Table 6 shows that under discretion either implementing the 'more' conservative regime, or enjoying a reduction in shock volatility alone, would reduce by more than half the volatility in inflation and interest rates implied by the worst case scenario. However, it is the 'good luck' that would lead to significant output stabilization and, therefore, achieve bigger gains as measured by either the central bank's estimated or the microfounded welfare metrics.

If the policy maker further increases the level of conservatism to the levels implied by the microfounded objectives, there is a striking reduction in inflation volatility to negligible levels. However, it significantly worsens output volatility in the high volatility regime. Clearly, the Fed has not implemented monetary policy with a degree of inflation conservatism anywhere near that implied by microfounded objectives.

Turning to Panel B of Table 6 we consider the same experiment, but now assume that policy is conducted under timeless commitment. In the absence of 'good luck', being able to act with timeless commitment allows the central bank to almost completely stabilize inflation volatility, but at the cost of moderate increases in output fluctuations. It is also important to note that welfare is clearly improved regardless of the degree of central bank conservatism. This result suggests that the reduction in inflation volatility achieved by being able to act under timeless commitment is such that the issue of conservatism becomes of second-order importance. Therefore, the dimension of 'good policy' policymakers should be concerned with is not the weight given to inflation stabilization in the policy maker's objective function, i.e. the conservatism of the central bank, but rather that they have the tools and credibility to effectively pursue a *timeless* commitment policy and to make time-inconsistent promises which they will keep. Finally, under timeless commitment we again see substantial decreases in output volatility when there is good luck.

## 8. Conclusions

A time consistent targeting rule – discretionary policy – provides the best fit to the data, outperforming conventional instrument rules and the other forms of optimal policy with different degrees of precommitment. Bayes factors reveal that there is 'strong' evidence in favor of this description of policy relative to simple instrument rules, and 'decisive' evidence relative to targeting rules formed under either timeless commitment or quasi-commitment. However, the ranking of policies in terms of fitting the data crucially depends on whether or not we account for potential changes in the Fed's degree of inflation conservatism and in shock volatilities. A failure to take into account policy switches hinders the ability of targeting rules to account for the monetary policy response to the high inflation of the 1970s relative to instrument rules. The absence of variation in shock volatilities exaggerates the fit of quasi-commitment because it can rely on policy surprises as a source of volatility. We demonstrate how inferences about shock processes, habit persistence and inflation indexation change across different policy specifications.

The preferred model implies that there was an increase in central bank conservatism following the Volcker disinflation period, which is estimated to occur in 1982. This description of policy also finds that the Fed relaxed policy temporarily in the aftermath of the 1987 stock market crash, and also lost conservatism following the 2000 dot-com crash, which it has never regained.

Based on estimates from the best-fit model, a range of counterfactual simulations are undertaken which throw light on various aspects of policy. First, there have been significant welfare gains to the conservatism in policy making that was adopted following the Volcker disinflation. However, these gains are small compared to those attained from the estimated reduction in shock volatilities. Relative to the average rate of inflation of 6.51% in the 1970s, a policy maker acting under discretion, but with the higher degree of conservatism observed later on in the sample, would have reduced average inflation to 4.71%. In contrast, inflation would have been expected to be 3.39% in the same period had the economy been lucky enough to have been in the low volatility regime. Second, had the US Fed been able to commit, rather than acting under discretion, then in the 1970s the average rate of inflation would have been below 2%, regardless of the level of conservatism. Taken together, this suggests that attempts to improve monetary policy outcomes should concentrate on ensuring that the Fed is able to make and communicate credible promises concerning future policy, and that this is of more importance than altering the preferences of the central banker.

The model employed in the paper was deliberately small scale, capturing the essential features of the larger scale models often employed in empirical analyses while facilitating the development of intuition. Future research could usefully analyse different countries and extend the analysis to larger scale models, using more refined models of the Phillip curve.

## References

Adolfson, M., Laseen, S., Linde, J., Svensson, L. E., 2011. Optimal Monetary Policy in an Operational Medium-Sized DSGE Model. Journal of Money, Credit and Banking 43 (7), 1287–1331.

An, S., Schorfheide, F., 2007. Bayesian Analysis of DSGE Models. Econometric Reviews 26 (2-4), 113–172.

Bianchi, F., 2013. Regime Switches, Agents' Beliefs, and Post-World War II U.S. Macroeconomic Dynamics. Review of Economic Studies 80 (2), 491–511.

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

Caglar, E., Chadha, J. S., Shibayama, K., 2011. Bayesian Estimation of DSGE models: Is the Workhorse Model Identified?, working Paper No 1125, University of Kent.

Calvo, G., 1983. Staggered Prices in a Utility-Maximising Framework. Journal of Monetary Economics 12, 383–398.

Clarida, R. H., Galí, J., Gertler, M., 1998. Monetary Policy Rules in Practice. Some International Evidence. European Economic Review 42, 1033–1067.

Davig, T., Doh, T., 2014. Monetary policy regime shifts and inflation persistence. The Review of Economics and Statistics 96 (5), 862–875.

Debortoli, D., Lakdawala, A., 2016. How Credible is the Federal Reserve? A Structural Estimation of Policy Re-optimizations. AEJ: Macroeconomics 8 (3), 42–76.

Debortoli, D., Nunes, R., 2010. Fiscal Policy under Loose Commitment. Journal of Economic Theory 145 (3), 1005–1032.

Dennis, R., 2004. Inferring Policy Objectives from Economic Outcomes. Oxford Bulletin of Economics and Statistics 66 (S1), 735–64.

Erceg, C. J., Levin, A. T., 2003. Imperfect Credibility and Inflation Persistence. Journal of Monetary Economics 50, 915–944.

Farmer, R. E. A., Waggoner, D. F., Zha, T., 2011. Minimal State Variable Solutions to Markov-switching Rational Expectations Models. Journal of Economic Dynamics and Control 35 (12), 2150–2166.

Favero, C. A., Rovelli, R., 2003. Macroeconomic stability and the preferences of the Fed. A formal analysis, 1961-98. Journal of Money, Credit and Banking 35 (4), 545–556.

Gali, J., Gertler, M., 1999. Inflation dynamics: a structural econometric analysis. Journal of Monetary Economics 44, 195–222.

Geweke, J., 1999. Using simulation methods for Bayesian econometric models: inference, development, and communication. Econometric Reviews 18 (1), 1–73.

Givens, G. E., 2012. Estimating Central Bank Preferences under Commitment and Discretion. Journal of Money, Credit and Banking 44 (6), 1033–1061.

Hamilton, J. D., Waggoner, D. F., Zha, T., 2007. Normalisation in Econometrics. Econometric Reviews 26, 221–252.

Ilbas, P., 2010. Estimation of monetary policy preferences in a forward-looking model: a Bayesian approach. International Journal of Central Banking 6 (3), 169–209.

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

Iskrev, N., 2010. Local identification in DSGE models. Journal of Monetary Economics 52 (7), 189–202.

Kass, R. E., Raftery, A. E., 1995. Bayes Factors. Journal of the American Statistical Association 90 (430), 773–795.

Komunjer, I., Ng, S., 2011a. Dynamic Identification of Dynamic Stochastic General Equilibrium Models. Econometrica 79 (6), 1995–2032.

Komunjer, I., Ng, S., 2011b. Supplementary Material to 'Dynamic Identification of Dynamic Stochastic General Equilibrium Models'. Econometrica 79 (6), 1995–2032.

Koop, G., Pesaran, M. H., Smith, R. P., 2013. On Identification of Bayesian DSGE Models. Journal of Business and Economic Statistics 31 (3), 300–314.

Le Roux, S., Kirsanova, T., 2013. Commitment vs. Discretion in the UK: An Empirical Investigation of the Monetary and Fiscal Policy Regime. International Journal of Central Banking 9 (4), 99–151.

Leith, C., Moldovan, I., Rossi, R., 2012. Optimal Monetary Policy in a New Keynesian Model with Habits in Consumption. Review of Economic Dynamics 15 (3), 416–435.

Lubik, T., Schorfheide, F., 2005. A Bayesian Look at the New Open Economy Macroeconomics. In: NBER Macroeconomics Annual 2005. MIT Press, pp. 313–379.

Schaumburg, E., Tambalotti, A., 2007. An Investigation of the Gains from Commitment in Monetary Policy. Journal of Monetary Economics 54 (2), 302–324.

Schmitt-Grohe, S., Uribe, M., 2007. Optimal Simple and Implementable Monetary and Fiscal Rules. Journal of Monetary Economics 54 (6), 1702–1725.

Schorfheide, F., 2005. Learning and Monetary Policy Shifts. Review of Economic Dynamics 8, 392–419.

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

Sims, C. A., Zha, T., 2006. Were There Regime Switches in US Monetary Policy? American Economic Review 96 (1), 54–81.

Smets, F., Wouters, R., 2003. An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area. Journal of the European Economic Association 1 (1), 1123–75.

Svensson, L., 2003. What Is Wrong with Taylor Rules? Using Judgment in Monetary Policy through Targeting Rules. Journal of Economic Literature 41, 426–477.

Svensson, L., Williams, N., 2007. Monetary Policy with Model Uncertainty: Distribution Forecast Targeting, CEPR Discussion Papers, 6331.

Taylor, J. B., 1999. Staggered Price and Wage Setting in Macroeconomics. In: Taylor, J. B., Woodford, M. (Eds.), Handbook of Macroeconomics. North Holland, pp. 1009–1050.

Table 1: Distribution of Priors

| Parameters | | Range | Density | Mean | Std Dev |
|---|---|---|---|---|---|
| inv. of intertemp. elas. of subst. | $\sigma$ | $\mathbb{R}$ | Normal | 2.50 | 0.25 |
| Calvo parameter | $\alpha$ | $[0,1)$ | Beta | 0.75 | 0.02 |
| inflation inertia | $\zeta$ | $[0,1)$ | Beta | 0.50 | 0.15 |
| habit persistence | $\theta$ | $[0,1)$ | Beta | 0.50 | 0.15 |
| inverse of Frisch elasticity | $\varphi$ | $\mathbb{R}$ | Normal | 2.50 | 0.25 |
| AR coeff., taste shock | $\rho^{\xi}$ | $[0,1)$ | Beta | 0.50 | 0.15 |
| AR coeff., cost-push shock | $\rho^{\mu}$ | $[0,1)$ | Beta | 0.50 | 0.15 |
| AR coeff., productivity shock | $\rho^{z}$ | $[0,1)$ | Beta | 0.50 | 0.15 |
| steady state interest rate | $r^A$ | $\mathbb{R}^+$ | Gamma | 3.5 | 2 |
| inflation target | $\pi^A$ | $\mathbb{R}^+$ | Gamma | 3.5 | 2 |
| steady state growth rate | $\gamma^Q$ | $\mathbb{R}$ | Nomal | 0.52 | 1 |
| probability of reneging | $\upsilon$ | $[0,1]$ | Uniform | 0.5 | 0.25 |
| Markov Switching s.d. of shocks | | | | | |
| preference shocks | $\sigma_{\xi(S=1=2)}$ | $\mathbb{R}^+$ | Inv. Gamma | 0.50 | 5 |
| cost-push shocks | $\sigma_{\mu(S=1=2)}$ | $\mathbb{R}^+$ | Inv. Gamma | 0.50 | 5 |
| technology shocks | $\sigma_{z(S=1=2)}$ | $\mathbb{R}^+$ | Inv. Gamma | 0.50 | 5 |
| policy shocks | $\sigma_{R(S=1=2)}$ | $\mathbb{R}^+$ | Inv. Gamma | 0.50 | 5 |
| Markov switching rule parameters | | | | | |
| interest rate smoothing | $\rho^R_{(s=1=2)}$ | $[0,1)$ | Beta | 0.50 | 0.25 |
| inflation (more conservative) | $\psi_{1(s=1)}$ | $\mathbb{R}^+$ | Gamma | 1.50 | 0.50 |
| inflation (less conservative) | $\psi_{1(s=2)}$ | $\mathbb{R}^+$ | Gamma | 1.0 | 0.50 |
| output | $\psi_{2(s=1=2)}$ | $\mathbb{R}^+$ | Gamma | 0.50 | 0.25 |
| Weights on Objectives | | | | | |
| gap term, $\hat{X}_t - \hat{\xi}_t$ | $\omega_1$ | $[0,1)$ | Beta | 0.50 | 0.15 |
| gap term, $\hat{y}_t - \frac{\sigma}{\varphi}\hat{\xi}_t$ | $\omega_2$ | $[0,1)$ | Beta | 0.50 | 0.15 |
| change in inflation, $\hat{\pi}_t - \hat{\pi}_{t-1}$ | $\omega_3$ | $[0,1)$ | Beta | 0.50 | 0.15 |
| inflation, $\hat{\pi}_t$ | $\omega_{\pi(s=2)}$ | $[0,1)$ | Beta | 0.50 | 0.15 |
| Markov switching in Inflation Target | | | | | |
| inflation target $(s=1)$ | $\pi^A_{(s=1)}$ | $\mathbb{R}^+$ | Gamma | 6 | 2 |
| inflation target $(s=2)$ | $\pi^A_{(s=2)}$ | $\mathbb{R}^+$ | Gamma | 3 | 2 |
| Transition Probabilities | | | | | |
| policy: remains more conservative | $p_{11}$ | $[0,1)$ | Beta | 0.90 | 0.05 |
| policy: remains less conservative | $p_{22}$ | $[0,1)$ | Beta | 0.90 | 0.05 |
| volatility: remains with low volatility | $q_{11}$ | $[0,1)$ | Beta | 0.90 | 0.05 |
| volatility: remains with high volatility | $q_{22}$ | $[0,1)$ | Beta | 0.90 | 0.05 |

*Notes:* For policy switches $s = 1$ is more conservative regime and $s = 2$ is less conservative regime. For volatility switches $S = 1$ is less volatile regime and $S = 2$ is more volatile regime.

Table 2: Estimation Results

| Parameters | Discretion | Rule: Parameters | Rule: Target | Quasi-Commitment | Timeless Commitment |
|---|---|---|---|---|---|
| | | | Model Parameters | | |
| $\sigma$ | 2.901 [2.526,3.244] | 2.937 [2.564,3.309] | 2.934 [2.556,3.301] | 2.692 [2.356,3.038] | 2.912 [2.480,3.338] |
| $\alpha$ | 0.735 [0.708,0.763] | 0.770 [0.742,0.799] | 0.775 [0.746,0.804] | 0.754 [0.732,0.776] | 0.775 [0.748,0.803] |
| $\zeta$ | 0.165 [0.069,0.254] | 0.088 [0.031,0.142] | 0.084 [0.030,0.138] | 0.182 [0.096,0.270] | 0.262 [0.114,0.419] |
| $\theta$ | 0.387 [0.206,0.560] | 0.827 [0.702,0.956] | 0.790 [0.631,0.950] | 0.372 [0.201,0.544] | 0.694 [0.304,0.953] |
| $\varphi$ | 2.459 [2.060,2.844] | 2.442 [2.030,2.855] | 2.424 [2.004,2.838] | 2.286 [1.889,2.672] | 2.199 [1.782,2.638] |
| | | | Shock Processes | | |
| $\rho^{\xi}$ | 0.830 [0.791,0.870] | 0.890 [0.853,0.927] | 0.901 [0.866,0.938] | 0.893 [0.869,0.919] | 0.919 [0.898,0.941] |
| $\rho^{\mu}$ | 0.939 [0.914,0.963] | 0.504 [0.262,0.759] | 0.502 [0.252,0.751] | 0.923 [0.900,0.948] | 0.992 [0.986,0.998] |
| $\rho^{z}$ | 0.195 [0.141,0.248] | 0.329 [0.228,0.427] | 0.359 [0.257,0.462] | 0.186 [0.134,0.238] | 0.162 [0.106,0.218] |
| $\sigma_{\xi(S=1)}$ | 0.425 [0.297,0.546] | 0.682 [0.527,0.837] | 0.545 [0.390,0.690] | 0.495 [0.334,0.649] | 0.404 [0.249,0.555] |
| $\sigma_{\xi(S=2)}$ | 0.873 [0.599,1.139] | 1.467 [1.040,1.888] | 1.346 [0.958,1.721] | 0.909 [0.652,1.167] | 1.224 [0.720,1.757] |
| $\sigma_{\mu(S=1)}$ | 0.236 [0.182,0.292] | 0.277 [0.169,0.381] | 0.276 [0.169,0.383] | 0.251 [0.188,0.315] | 1.329 [0.737,1.905] |
| $\sigma_{\mu(S=2)}$ | 0.684 [0.527,0.840] | 0.546 [0.343,0.751] | 0.545 [0.390,0.690] | 0.864 [0.658,1.065] | 2.806 [1.697,3.913] |
| $\sigma_{z(S=1)}$ | 0.512 [0.391,0.622] | 0.601 [0.540,0.660] | 0.603 [0.542,0.664] | 0.433 [0.352,0.515] | 0.452 [0.372,0.526] |
| $\sigma_{z(S=2)}$ | 1.064 [0.932,1.193] | 1.184 [0.981,1.380] | 1.156 [0.977,1.329] | 1.034 [0.918,1.148] | 0.989 [0.870,1.103] |
| $\sigma_{R(S=1)}$ | – | 0.140 [0.124,0.156] | 0.146 [0.129,0.162] | – | – |
| $\sigma_{R(S=2)}$ | – | 0.412 [0.332,0.489] | 0.455 [0.379,0.529] | – | – |
| | | | Data Means | | |
| $r^{A}$ | 0.802 [0.294,1.282] | 0.541 [0.189,0.873] | 0.509 [0.165,0.828] | 0.803 [0.330,1.266] | 0.722 [0.257,1.184] |
| $\pi^{A}_{(s=1)}$ | 1.305 [0.629,1.943] | 3.558 [2.986,4.122] | 3.336 [2.745,3.948] | 1.962 [1.588,2.326] | 2.755 [2.303,3.189] |
| $\pi^{A}_{(s=2)}$ | – | – | 4.329 [3.662,5.001] | – | – |
| $\gamma^{Q}$ | 0.773 [0.669,0.897] | 0.713 [0.592,0.832] | 0.700 [0.566,0.829] | 0.790 [0.697,0.885] | 0.828 [0.721,0.931] |

continued on the next page

Table 2: Estimation Results – continued

| Parameters | Discretion | Rule: Parameters | Rule: Target | Quasi-Commitment | Timeless Commitment |
|---|---|---|---|---|---|
| | | | Policy Parameters | | |
| $\upsilon$ | – | – | – | 0.290 [0.227,0.355] | – |
| $\rho^R_{(s=1)}$ | – | 0.825 [0.793,0.858] | 0.821 [0.793,0.851] | – | – |
| $\rho^R_{(s=2)}$ | – | 0.868 [0.779,0.946] | – | – | – |
| $\psi_{1(s=1)}$ | – | 2.124 [1.798,2.447] | 2.014 [1.655,2.370] | – | – |
| $\psi_{1(s=2)}$ | – | 1.219 [0.809,1.635] | – | – | – |
| $\psi_{2(s=1)}$ | – | 0.511 [0.327,0.692] | 0.587 [0.381,0.784] | – | – |
| $\psi_{2(s=2)}$ | – | 0.274 [0.102,0.438] | – | – | – |
| $\omega_1$ | 0.380 [0.232,0.534] | – | – | 0.624 [0.476,0.777] | 0.503 [0.320,0.690] |
| $\omega_2$ | 0.635 [0.468,0.800] | – | – | 0.749 [0.618,0.884] | 0.559 [0.280,0.843] |
| $\omega_3$ | 0.436 [0.200,0.667] | – | – | 0.369 [0.141,0.586] | 0.454 [0.195,0.695] |
| $\omega_{\pi(s=1)}$ | 1 | – | – | 1 | 1 |
| $\omega_{\pi(s=2)}$ | 0.436 [0.279,0.589] | – | – | 0.301 [0.204,0.395] | 0.373 [0.216,0.527] |
| | | | Markov Transition Probabilities | | |
| $p_{11}$ | 0.947 [0.903,0.989] | 0.964 [0.942,0.988] | 0.902 [0.840,0.964] | 0.798 [0.715,0.882] | 0.978 [0.959,0.997] |
| $p_{22}$ | 0.918 [0.876,0.962] | 0.846 [0.812,0.880] | 0.812 [0.740,0.889] | 0.914 [0.865,0.966] | 0.798 [0.722,0.877] |
| $q_{11}$ | 0.952 [0.919,0.986] | 0.956 [0.928,0.985] | 0.979 [0.960,0.998] | 0.907 [0.852,0.962] | 0.958 [0.931,0.986] |
| $q_{22}$ | 0.955 [0.910,0.997] | 0.843 [0.779,0.910] | 0.946 [0.902,0.992] | 0.941 [0.905,0.977] | 0.933 [0.887,0.976] |
| | | | Log Marginal Data Densities and Bayes Factors | | |
| Geweke | −759.78 (1.00) | −764.16 (80.29) | −765.83 (425.76) | −770.29 (3.67$e$+4) | −793.62 (4.98$e$+14) |
| Sims et.al. | −759.91 (1.00) | −764.21 (74.08) | −765.95 (422.76) | −770.34 (3.40$e$+4) | −793.95 (6.12$e$+14) |

Notes: Here and in Tables 3-5 for each parameter the posterior distribution is described by its mean and 90% confidence interval in square brackets. Bayes Factors for marginal data densities are in parentheses.

Table 3: Selected Parameter Estimates - No Switching

| Parameters | Simple Rule | Quasi-Commitment | Discretion | Timeless Commitment |
|---|---|---|---|---|
| | | Selected Model Parameters | | |
| $\zeta$ | 0.103 [0.039,0.166] | 0.170 [0.087,0.252] | 0.156 [0.066,0.241] | 0.594 [0.489,0.737] |
| $\theta$ | 0.823 [0.685,0.964] | 0.421 [0.210,0.627] | 0.476 [0.267,0.680] | 0.643 [0.444,0.782] |
| | | Policy Parameters | | |
| $\upsilon$ | – | 0.556 [0.329,0.816] | – | – |
| $\rho^R$ | 0.791 [0.756,0.826] | – | – | – |
| $\psi_1$ | 1.716 [1.455,1.972] | – | – | – |
| $\psi_2$ | 0.492 [0.290,0.697] | – | – | – |
| $\omega_1$ | – | 0.703 [0.552,0.861] | 0.458 [0.287,0.627] | 0.627 [0.490,0.808] |
| $\omega_2$ | – | 0.828 [0.727,0.935] | 0.758 [0.628,0.901] | 0.446 [0.316,0.620] |
| $\omega_3$ | – | 0.390 [0.163,0.619] | 0.451 [0.213,0.692] | 0.489 [0.268,0.712] |
| | | Data Means | | |
| $r^A$ | 0.706 [0.246,1.139] | 0.759 [0.143,1.330] | 0.966 [0.352,1.569] | 1.088 [0.459,1.540] |
| $\pi^A$ | 4.746 [3.800,5.677] | 2.586 [1.899,3.095] | 2.656 [1.008,4.221] | 4.050 [3.642,4.674] |
| $\gamma^Q$ | 0.688 [0.547,0.826] | 0.737 [0.613,0.861] | 0.716 [0.593,0.835] | 0.726 [0.594,0.797] |
| | | Log Marginal Data Densities and Bayes Factors | | |
| Geweke | $-841.01$ (1.00) | $-841.67$ (1.94) | $-842.49$ (4.41) | $-855.43$ ($1.84e+6$) |
| Sims et.al | $-841.09$ (1.00) | $-841.54$ (1.57) | $-842.69$ (4.96) | $-858.26$ ($2.85e+7$) |

Table 4: Selected Parameter Estimates - Switches in Policy Only

| Parameters | Discretion | Quasi-Commitment | Rule: Parameters | Rule: Target | Timeless Commitment |
|---|---|---|---|---|---|
| | | Selected Model Parameters | | | |
| $\zeta$ | 0.155 [0.069,0.239] | 0.182 [0.091,0.274] | 0.102 [0.038,0.163] | 0.123 [0.054,0.195] | 0.229 [0.078,0.366] |
| $\theta$ | 0.479 [0.286,0.835] | 0.371 [0.192,0.543] | 0.825 [0.698,0.954] | 0.810 [0.658,0.961] | 0.606 [0.388,0.843] |
| | | Policy Parameters | | | |
| $\upsilon$ | – | 0.325 [0.239,0.411] | – | – | – |
| $\rho^R_{(s=1)}$ | – | – | 0.746 [0.708,0.786] | 0.797 [0.762,0.831] | – |
| $\rho^R_{(s=2)}$ | – | – | 0.845 [0.794,0.900] | – | – |
| $\psi_{1(s=1)}$ | – | – | 2.075 [1.824,2.315] | 1.805 [1.507,2.097] | – |
| $\psi_{1(s=2)}$ | – | – | 0.909 [0.621,1.189] | – | – |
| $\psi_{2(s=1)}$ | – | – | 0.483 [0.309,0.645] | 0.498 [0.285,0.714] | – |
| $\psi_{2(s=2)}$ | – | – | 0.245 [0.098,0.393] | – | – |
| $\omega_1$ | 0.259 [0.035,0.414] | 0.633 [0.480,0.785] | – | – | 0.502 [0.331,0.666] |
| $\omega_2$ | 0.650 [0.460,0.847] | 0.759 [0.631,0.893] | – | – | 0.523 [0.295,0.732] |
| $\omega_3$ | 0.442 [0.164,0.698] | 0.349 [0.126,0.559] | – | – | 0.460 [0.205,0.710] |
| $\omega_{\pi(s=1)}$ | 1 | 1 | – | – | 1 |
| $\omega_{\pi(s=2)}$ | 0.347 [0.219,0.477] | 0.348 [0.254,0.440] | – | – | 0.302 [0.194,0.414] |
| | | Data Means | | | |
| $r^A$ | 0.766 [0.303,1.213] | 0.997 [0.377,1.591] | 0.695 [0.276,1.105] | 0.662 [0.239,1.054] | 0.975 [0.358,1.561] |
| $\pi^A_{(s=1)}$ | 2.683 [1.275,4.022] | 2.097 [1.770,2.431] | 3.736 [3.183,4.299] | 4.234 [3.470,4.995] | 3.064 [2.733,3.411] |
| $\pi^A_{(s=2)}$ | – | – | – | 6.058 [5.217,6.862] | – |
| $\gamma^Q$ | 0.683 [0.567,0.800] | 0.722 [0.598,0.842] | 0.677 [0.540,0.808] | 0.681 [0.544,0.822] | 0.741 [0.619,0.862] |
| | | Log Marginal Data Densities and Bayes Factors | | | |
| Geweke | $-810.98$ (1.00) | $-814.83$ (47.0) | $-825.33$ (1.72$e$+6) | $-831.74$ (1.04$e$+9) | $-832.85$ (3.14$e$+9) |
| Sims et.al. | $-811.24$ (1.00) | $-814.30$ (21.21) | $-825.44$ (1.46$e$+6) | $-831.81$ (8.52$e$+8) | $-832.98$ (2.75$e$+9) |

Table 5: Estimation Results - MS Shocks only

| Parameters | Discretion micro-founded objective | Quasi-commitment Ad Hoc objective type II | Quasi-commitment micro-founded objective | Quasi-commitment Ad Hoc objective type I |
|---|---|---|---|---|
| | | Model Parameters | | |
| $\sigma$ | 2.866 [2.503,3.227] | 2.588 [2.220,2.944] | 2.375 [2.054,2.688] | 2.186 [1.851,2.509] |
| $\alpha$ | 0.751 [0.724,0.779] | 0.808 [0.788,0.827] | 0.787 [0.765,0.808] | 0.811 [0.793,0.828] |
| $\zeta$ | 0.173 [0.075,0.261] | 0.173 [0.089,0.256] | 0.194 [0.142,0.243] | 0.153 [0.080,0.224] |
| $\theta$ | 0.459 [0.220,0.715] | 0.495 [0.409,0.580] | 0.478 [0.383,0.570] | 0.293 [0.233,0.349] |
| $\varphi$ | 2.274 [1.872,2.675] | 2.020 [1.577,2.439] | 1.793 [1.453,2.137] | 2.031 [1.614,2.436] |
| | | Shock Processes | | |
| $\rho^{\xi}$ | 0.843 [0.810,0.877] | 0.822 [0.758,0.891] | 0.898 [0.875,0.922] | 0.875 [0.842,0.910] |
| $\rho^{\mu}$ | 0.936 [0.911,0.961] | 0.926 [0.891,0.963] | 0.930 [0.903,0.957] | 0.936 [0.907,0.968] |
| $\rho^{z}$ | 0.183 [0.132,0.239] | 0.300 [0.215,0.386] | 0.194 [0.142,0.243] | 0.201 [0.142,0.258] |
| $\sigma_{\xi(S=1)}$ | 0.443 [0.311,0.575] | 0.510 [0.315,0.709] | 0.510 [0.332,0.681] | 0.480 [0.340,0.616] |
| $\sigma_{\xi(S=2)}$ | 0.898 [0.622,1.171] | 1.905 [1.187,2.657] | 1.082 [0.747,1.404] | 1.186 [0.846,1.509] |
| $\sigma_{\mu(S=1)}$ | 0.234 [0.178,0.286] | 0.829 [0.433,1.260] | 0.317 [0.219,0.411] | 0.583 [0.372,0.781] |
| $\sigma_{\mu(S=2)}$ | 0.769 [0.579,0.951] | 2.247 [1.557,2.910] | 1.094 [0.773,1.431] | 1.801 [1.345,2.239] |
| $\sigma_{z(S=1)}$ | 0.476 [0.380,0.569] | 0.526 [0.441,0.610] | 0.450 [0.358,0.542] | 0.438 [0.347,0.526] |
| $\sigma_{z(S=2)}$ | 1.064 [0.361,1.189] | 0.962 [0.794,1.111] | 1.061 [0.937,1.184] | 1.024 [0.893,1.148] |
| | | Data Means | | |
| $r^{A}$ | 0.763 [0.277,1.213] | 0.732 [0.249,1.218] | 0.666 [0.245,1.082] | 0.662 [0.241,1.071] |
| $\pi^{A}$ | 1.706 [0.693,2.643] | 2.276 [1.879,2.678] | 2.150 [1.674,2.636] | 2.481 [2.178,2.793] |
| $\gamma^{Q}$ | 0.789 [0.692,0.885] | 0.761 [0.645,0.882] | 0.783 [0.682,0.883] | 0.786 [0.684,0.887] |

continued on the next page

Table 5: Estimation Results - MS Shocks only – continued

| Parameters | Discretion micro-founded objective | Quasi-commitment Ad Hoc objective type II | Quasi-commitment micro-founded objective | Quasi-commitment Ad Hoc objective type I |
|---|---|---|---|---|
| | | Policy Parameters | | |
| $\upsilon$ | – | 0.194 [0.127,0.261] | 0.260 [0.187,0.333] | 0.144 [0.110,0.177] |
| $\rho^R$ | – | – | – | – |
| $\psi_1$ | – | – | – | – |
| $\psi_2$ | – | – | – | – |
| $\omega_1$ | 0.454 [0.275,0.642] | – | 0.746 [0.609,0.882] | – |
| $\omega_2$ | 0.715 [0.569,0.867] | – | 0.819 [0.714,0.927] | – |
| $\omega_3$ | 0.444 [0.198,0.676] | – | 0.402 [0.167,0.633] | – |
| $\omega_\pi$ | 1 | 1 | 1 | 1 |
| $\omega_y$ | – | 0.819 [0.711,0.933] | – | 0.866 [0.781,0.952] |
| $\omega_R$ | – | 1.533 [0.734,2.349] | – | – |
| | | Markov Transition Probabilities | | |
| $p_{11}$ | 0.916 [0.866,0.968] | 0.948 [0.902,0.997] | 0.900 [0.840,0.964] | 0.879 [0.813,0.944] |
| $p_{22}$ | 0.892 [0.849,0.934] | 0.959 [0.931,0.986] | 0.939 [0.904,0.973] | 0.940 [0.903,0.978] |
| | | Log Marginal Data Densities and Bayes Factors | | |
| Geweke | $-776.22$ (1.0) | $-782.97$ (854.06) | $-792.73$ (1.48$e$+7) | $-837.80$ (5.54$e$+26) |
| Sims et.al | $-776.23$ (1.0) | $-782.81$ (718.38) | $-792.74$ (1.49$e$+07) | $-837.64$ (4.67$e$+26) |

Note: The prior for $\omega_R$ is Gamma (1,1) and for $\omega_y$ it is Beta (0.5,0.15).

Table 6: Unconditional Variances and Welfare under Alternative Policies and Volatilities

| Regime: (conservatism, volatility) | Output | Inflation | Interest Rate | Welfare Cost (est. weights) | Welfare Cost (micro. weights) |
|---|---|---|---|---|---|
| A: Discretion | | | | | |
| (less, high)* | 0.147 [0.092,0.228] | 2.044 [1.413,3.157] | 1.452 [0.936,2.459] | 3.726 [2.250,6.554] | 1.05% [0.69%,1.54%] |
| (more, high) | 0.151 [0.100,0.234] | 0.698 [0.467,1.00] | 0.593 [0.449,0.844] | 3.584 [2.126,6.397] | 0.41% [0.30%,0.60%] |
| (micro, high) | 0.177 [0.127,0.259] | 0.002 [0.001,0.003] | 0.480 [0.403,0.566] | – | 0.08% [0.05%,0.15%] |
| (less, low) | 0.060 [0.036,0.093] | 0.798 [0.541,1.231] | 0.509 [0.311,0.893] | 0.811 [0.485,1.451] | 0.17% [0.11%,0.26%] |
| (high, low) | 0.057 [0.035,0.089] | 0.281 [0.179,0.407] | 0.223 [0.166,0.322] | 0.793 [0.470,1.435] | 0.07% [0.05%,0.115%] |
| (micro, low) | 0.061 [0.042,0.094] | 0.001 [0.000,0.001] | 0.232 [0.193,0.276] | – | 0.02% [0.01%,0.03%] |
| B: Timeless Commitment | | | | | |
| (less, high) | 0.166 [0.112,0.250] | 0.053 [0.037,0.081] | 0.746 [0.624,0.893] | 2.982 [1.588,5.720] | 0.13% [0.09%,0.20%] |
| (more, high) | 0.168 [0.117,0.251] | 0.018 [0.012,0.026] | 0.697 [0.0.586,0.829] | 3.009 [1.616,5.753] | 0.10% [0.07%,0.17%] |
| (micro, high) | 0.179 [0.129,0.261] | 0.000 [0.000,0.000] | 0.463 [0.387,0.547] | – | 0.08% [0.05%,0.15%] |
| (less, low) | 0.062 [0.040,0.095] | 0.023 [0.015,0.033] | 0.364 [0.296,0.446] | 0.688 [0.377,1.319] | 0.03% [0.02%,0.04%] |
| (more, low) | 0.061 [0.040,0.094] | 0.008 [0.005,0.012] | 0.341 [0.279,0.414] | 0.694 [0.383,1.326] | 0.02% [0.02%,0.04%] |
| (micro, low) | 0.062 [0.042,0.095] | 0.000 [0.000,0.000] | 0.225 [0.187,0.268] | – | 0.02% [0.01%,0.03%] |

Notes: The welfare costs are computed using equation (9) where weights are either estimated or microfounded functions of estimated structural parameters. The microfounded welfare costs are expressed as a percentage of steady-state consumption. For both timeless commitment and discretionary policy we compute social welfare using regimes and regime parameters identified for discretionary policy.

Figure 1: Markov Switching Probabilities: Policy and Volatility Switches

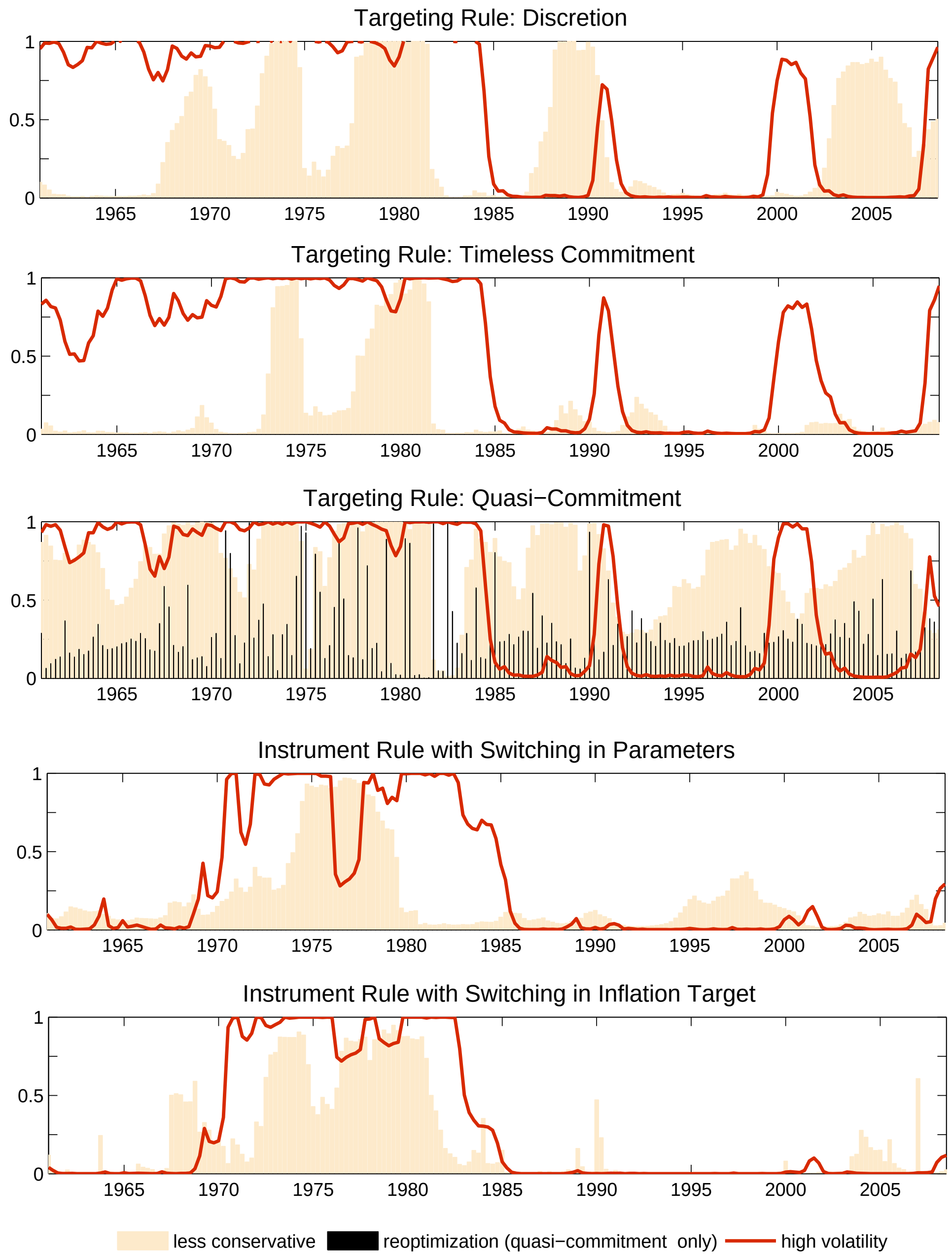
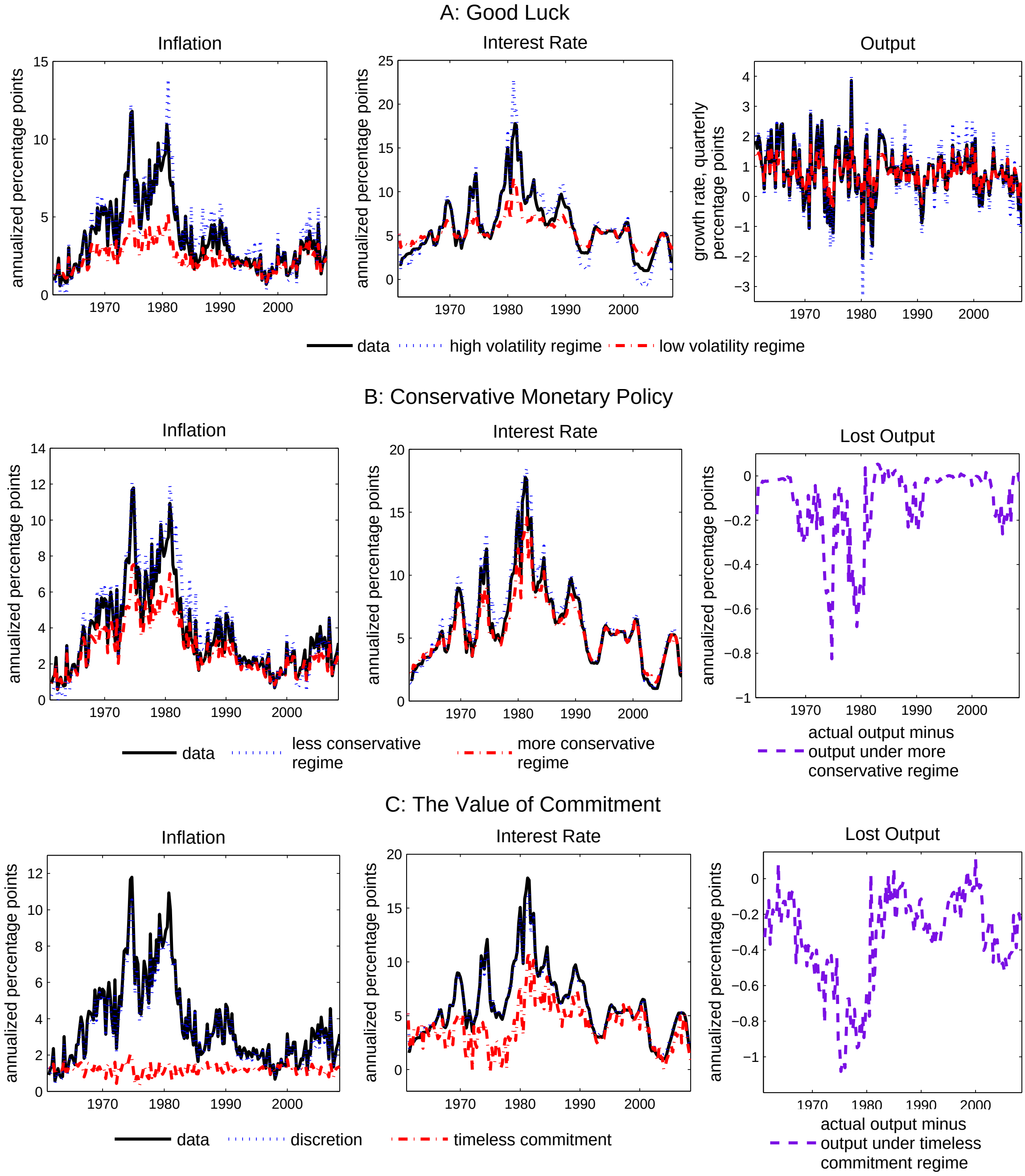


Figure 2: Counterfactuals



# How Optimal is US Monetary Policy?
# Online Appendix

Xiaoshan Chen *
University of Durham

Tatiana Kirsanova†
University of Glasgow

Campbell Leith‡
University of Glasgow

August 7, 2017

*Address: Department of Economics, Durham University Business School, University of Durham, Durham, DH1 3LB; e-mail xiaoshan.chen@durham.ac.uk

†Address: Economics, Adam Smith Business School, Gilbert Scott Building, University of Glasgow, Glasgow G12 8QQ; e-mail tatiana.kirsanova@glasgow.ac.uk

‡Address: Economics, Adam Smith Business School, Gilbert Scott Building, University of Glasgow, Glasgow G12 8QQ; e-mail campbell.leith@glasgow.ac.uk

This Appendix contains the detailed explanation of the model (Appendix A), derivation of the microfounded objective function (Appendix B), description of the algorithms used to solve the policy problem (Appendix C), details of the estimation strategy (Appendix D), discussion and application of identification tests (Appendix E), details of the approach to model selection (Appendix F) and some additional results (Appendix G).

# A The Model

The model is a small scale New Keynesian model featuring households who supply labour to imperfectly competitive firms which are subject to nominal inertia in the form of Calvo (1983) contracts. In order to introduce the possibility of some intrinsic persistence in the model, households are assumed to be subject to a habits externality, while some firms may employ simple rules of thumb when setting prices in a manner which introduces inflation inertia to the NKPC. Below we describe the micro-foundations of the model in more detail.

## A.1 Households

The economy is populated by a continuum of households, indexed by $k$ and of measure 1. Households derive utility from consumption of a composite good, $C_t^k = \left(\int_0^1 \left(C_{it}^k\right)^{\frac{\eta-1}{\eta}} di\right)^{\frac{\eta}{\eta-1}}$ where $\eta$ is the elasticity of substitution between the goods in this basket and suffer disutility from hours spent working, $N_t^k$. Habits are both superficial and external implying that they are formed at the level of the aggregate consumption good, and that households fail to take account of the impact of their consumption decisions on the utility of others. To facilitate data-consistent detrending around a balanced growth path without restricting preferences to be logarithmic in form, we also follow Lubik and Schorfheide (2005) and An and Schorfheide (2007) in assuming that the consumption that enters the utility function is scaled by the economy wide technology trend, implying that household's consumption norms rise with technology as well as being affected by more familiar habits externalities. Accordingly, households derive utility from the habit-adjusted composite good,

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[ \frac{\left(C_t^k/A_t - \theta C_{t-1}/A_{t-1}\right)^{1-\sigma} \xi_t^{-\sigma}}{1-\sigma} - \frac{\left(N_t^k\right)^{1+\varphi} \xi_t^{-\sigma}}{1+\varphi} \right],$$

where $C_{t-1} \equiv \int_0^1 C_{t-1}^k dk$ is the cross-sectional average of consumption.[1] In other words households gain utility from consuming more than other households, and are disappointed if their consumption doesn't grow in line with technical progress, $A_t$, and are subject to a time-preference or taste-shock, $\xi_t$. $\mathbb{E}_t$ is the mathematical expectation conditional on information available at time $t$, $\beta$ is the discount factor $(0 < \beta < 1)$, and $\sigma$ and $\varphi$ are the inverses of the intertemporal elasticities of habit-adjusted consumption and work $(\sigma, \varphi > 0;\ \sigma \neq 1)$.

The process for technology is non-stationary:

$$\begin{aligned} \ln A_t &= \ln \gamma + \ln A_{t-1} + \ln z_t, \\ \ln z_t &= \rho_z \ln z_{t-1} + \varepsilon_{z,t}. \end{aligned}$$

Households decide the composition of the consumption basket to minimize expenditures, and the demand for individual good $i$ is

$$C_{it}^k = \left(\frac{P_{it}}{P_t}\right)^{-\eta} C_t^k = \left(\frac{P_{it}}{P_t}\right)^{-\eta} \left(X_t^k + \theta C_{t-1}\right).$$

By aggregating across all households, we obtain the overall demand for good $i$ as

$$C_{it} = \int_0^1 C_{it}^k dk = \left(\frac{P_{it}}{P_t}\right)^{-\eta} C_t. \tag{1}$$

The remainder of the household's problem is standard. Specifically, households choose the habit-adjusted consumption aggregate, $X_t^k = C_t^k/A_t - \theta C_{t-1}/A_{t-1}$, hours worked, $N_t^k$, and the portfolio allocation, $D_{t+1}^k$, to maximize expected lifetime utility

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{\left(X_t^k\right)^{1-\sigma} (\xi_t)^{-\sigma}}{1-\sigma} - \frac{\left(N_t^k\right)^{1+\varphi} (\xi_t)^{-\sigma}}{1+\varphi}\right],$$

subject to the budget constraint

$$\int_0^1 P_{it} C_{it}^k di + E_t Q_{t,t+1} D_{t+1}^k = W_t N_t^k (1 - \tau_t) + D_t^k + \Phi_t + T_t,$$

and the usual transversality condition. The household's period-$t$ income includes: wage income from providing labor services to goods producing firms, $W_t N_t^k$, which is subject to a time-varying tax rate, $\tau_t$ , dividends from the monopolistically competitive firms, $\Phi_t$, and payments on the

[1] Note that this utility specification is slightly different from that in Lubik and Schorfheide (2005) who adopt the following specification, $(C_t - \theta\gamma C_{t-1})/A_t)^{1-\sigma} (\xi_t)^{-\sigma}/(1-\sigma)$. Their specification introduces a technology shock into the definition of habits adjusted consumption which then complicates the derivation of welfare. Therefore we adopt a specification which implies habits in detrended variables, which means that the only place the technology shock appears is in the consumption Euler equation.

portfolio of assets, $D_t^k$. Financial markets are complete. Lump-sum transfers, $T_t$, are paid by the government. The tax rate, $\tau_t$, will be used to finance lump-sum transfers, and can be designed to ensure that the long-run equilibrium is efficient in the presence of the habits and monopolistic competition externalities. However, we shall assume that the tax rate fluctuates around this efficient level such that it is responsible for generating an autocorrelated cost-push shock.

In the maximization problem, households take as given the processes for $C_{t-1}$, $W_t$, $\Phi_t$, and $T_t$, as well as the initial asset position $D_{-1}^k$. The first order conditions for labor and habit-adjusted consumption are

$$\frac{\left(N_t^k\right)^{\varphi}}{\left(X_t^k\right)^{-\sigma}} = \frac{W_t}{P_t A_t}(1-\tau_t),$$

and

$$1 = \beta \mathbb{E}_t \left[ \left( \frac{X_{t+1}^k \xi_{t+1}}{X_t^k \xi_t} \right)^{-\sigma} \frac{A_t}{A_{t+1}} \frac{P_t}{P_{t+1}} \right] R_t,$$

where

$$Q_{t,t+1} = \beta \left( \frac{X_{t+s}^k \xi_{t+1}}{X_t^k \xi_t} \right)^{-\sigma} \frac{A_t}{A_{t+1}} \frac{P_t}{P_{t+1}}.$$

is the one-period stochastic discount factor for nominal payoffs and $R_t^{-1} = \mathbb{E}_t \left[Q_{t,t+1}\right]$ denotes the inverse of the risk-free gross nominal interest rate between periods $t$ and $t+1$.

## A.2 Firms

We further assume that intermediate goods producers are subject to the constraints of Calvo (1983)-contracts such that, with fixed probability $(1-\alpha)$ in each period, a firm can reset its price and with probability $\alpha$ the firm retains the price of the previous period, but where, following Yun (1996) that price is indexed to the steady-state rate of inflation. When a firm can set the price, it can either do so in order to maximize the present discounted value of profits, $\mathbb{E}_t \sum_{s=0}^{\infty} \alpha^s Q_{t,t+s} \Phi_{it+s}$, or it can follow a simple rule of thumb as in (Galí and Gertler, 1999). The constraints facing the forward looking profit maximizers are the demand for their own good (1) and the constraint that all demand be satisfied at the chosen price. Profits are discounted by the $s$-step ahead stochastic discount factor $Q_{t,t+s}$ and by the probability of not being able to set prices in future periods. The firm's optimization problem is

$$\max_{\{P_{it},\, Y_{it}\}} \mathbb{E}_t \sum_{s=0}^{\infty} \alpha^s Q_{t,t+s} \left(P_{it} \pi^s - MC_{t+s}\right) Y_{it+s},$$

subject to the demand system

$$Y_{it+s} = \left(\frac{P_{it}\pi^s}{P_{t+s}}\right)^{-\eta} Y_{t+s},$$

and nominal rigidity.

The relative price set by firms able to reset prices optimally in a forward-looking manner, satisfies the following relationship

$$\frac{P_t^f}{P_t} = \frac{\eta}{\eta-1} \frac{\mathbb{E}_t \sum_{s=0}^{\infty} (\alpha\beta)^s \left(X_{t+s}\xi_{t+s}\right)^{-\sigma} mc_{t+s} \left(\frac{P_{t+s}\pi^{-s}}{P_t}\right)^{\eta} \frac{Y_{t+s}}{A_{t+s}}}{\mathbb{E}_t \sum_{s=0}^{\infty} (\alpha\beta)^s \left(X_{t+s}\xi_{t+s}\right)^{-\sigma} \left(\frac{P_{t+s}\pi^{-s}}{P_t}\right)^{\eta-1} \frac{Y_{t+s}}{A_{t+s}}}, \tag{2}$$

where $mc_t = MC_t/P_t$ is the real marginal cost and $P_t^f$ denotes the price set by all firms who are able to reset prices in period $t$ and choose to do so in a profit maximizing way.

To introduce inflation inertia we allow some firms to follow simple rules of thumb when setting prices. Specifically, when a firm is given the opportunity of posting a new price, we assume that rather than posting the profit-maximizing price (2), a proportion of those firms, $\zeta$, follow a simple rule of thumb in resetting that price

$$P_t^b = P_{t-1}^* \pi_{t-1}, \tag{3}$$

such that they update their price in line with last period's rate of inflation rather than steady-state inflation, where $P_{t-1}^*$ denotes an index of the reset prices given by

$$\ln P_{t-1}^* = (1-\zeta) \ln P_{t-1}^f + \zeta P_{t-1}^b.$$

With $\alpha$ of firms keeping last period's price (but indexed to steady-state inflation) and $(1-\alpha)$ of firms setting a new price, the law of motion of this price index is,

$$(P_t)^{1-\eta} = \alpha \left(P_{t-1}\pi\right)^{1-\eta} + (1-\alpha) \left(P_t^*\right)^{1-\eta}.$$

We denote the fixed share of price-setters following the rule of thumb (3) by $\zeta$, we derive a price inflation Phillips curve, as detailed in Leith and Malley (2005). Combining the rule of thumb of price setters with the optimal price setting described above leads to the price Phillips curve

$$\hat{\pi}_t = \chi_f \beta \mathbb{E}_t \hat{\pi}_{t+1} + \chi_b \hat{\pi}_{t-1} + \kappa_c \hat{w}_t,$$

where $\hat{\pi}_t = \ln(P_t) - \ln(P_{t-1}) - \ln(\pi)$ is the deviation of inflation from its steady state value, $\hat{w}_t = \ln(W_t/P_t) - \ln A_t - \ln((\eta-1)/\eta) = \hat{mc}_t$, are log-linearized real marginal costs (wages), and the reduced-form parameters are defined as $\chi_f \equiv \alpha/\Phi$, $\chi_b \equiv \zeta/\Phi$, $\kappa_c \equiv (1-\alpha)(1-\zeta)(1-\alpha\beta)/\Phi$, with $\Phi \equiv \alpha(1+\beta\zeta) + (1-\alpha)\zeta$.

### A.3 The Government

The government collects a distortionary tax on labor income which it rebates to households as a lump-sum transfer. The steady-state value of this distortionary tax will be set at a level which offsets the combined effect of the monopolistic competition distortion and the effects of the habits externality, as in Levine, McAdam, and Pearlman (2008), see Appendix B.1. However, shocks to the tax rate described by

$$\ln(1-\tau_t) = \rho^{\mu}\ln(1-\tau_{t-1}) + (1-\rho^{\mu})\ln(1-\tau) - \varepsilon_t^{\mu}$$

serve as autocorrelated cost-push shocks to the NKPC. There is no government spending per se. The government budget constraint is given by

$$\tau_t W_t N_t = -T_t.$$

### A.4 The Complete Model

The complete system of non-linear equations describing the equilibrium are given by

$$N_t^{\varphi}\left(\frac{X_t}{A_t}\right)^{\sigma} = \frac{W_t}{A_t P_t}(1-\tau_t) \equiv w_t(1-\tau_t), \tag{4}$$

$$\left(\frac{X_t}{A_t}\right)^{-\sigma}\xi_t^{-\sigma} = \beta\mathbb{E}_t\left[\left(\frac{X_{t+1}}{A_{t+1}}\right)^{-\sigma}\frac{A_t}{A_{t+1}}\;\xi_{t+1}^{-\sigma}R_t\pi_{t+1}^{-1}\right], \tag{5}$$

$$N_t = \frac{Y_t}{A_t}\int_0^1\left(\frac{P_t(i)}{P_t}\right)^{-\eta}di, \tag{6}$$

$$X_t = C_t - \theta C_{t-1}, \tag{7}$$

$$Y_t = C_t, \tag{8}$$

$$\tau_t W_t N_t = -T_t, \tag{9}$$

$$\frac{P_t^f}{P_t} = \frac{\eta}{\eta-1}\frac{\mathbb{E}_t\sum_{s=0}^{\infty}(\alpha\beta)^s\left(\frac{X_{t+s}\xi_{t+s}}{A_{t+s}}\right)^{-\sigma}mc_{t+s}\left(\frac{P_{t+s}\pi^{-s}}{P_t}\right)^{\eta}\frac{Y_{t+s}}{A_{t+s}}}{\mathbb{E}_t\sum_{s=0}^{\infty}(\alpha\beta)^s\left(\frac{X_{t+s}\xi_{t+s}}{A_{t+s}}\right)^{-\sigma}\left(\frac{P_{t+s}\pi^{-s}}{P_t}\right)^{\eta-1}\frac{Y_{t+s}}{A_{t+s}}}, \tag{10}$$

$$mc_t = \frac{W_t}{A_t P_t}, \tag{11}$$

$$P_t^b = P_{t-1}^*\pi_{t-1}, \tag{12}$$

$$\begin{aligned}
\ln P_{t-1}^{*} &= (1-\zeta)\ln P_{t-1}^{f} + \zeta P_{t-1}^{b}, & (13)\\
P_t^{1-\eta} &= \alpha\left(\pi P_{t-1}\right)^{1-\eta} + (1-\alpha)\left(P_t^{*}\right)^{1-\eta}, & (14)\\
\ln A_t &= \ln\gamma + \ln A_{t-1} + \ln z_t, & (15)\\
\ln z_t &= \rho_z \ln z_{t-1} + \varepsilon_t^{z}, & (16)\\
\ln(1-\tau_t) &= \rho^{\mu}\ln(1-\tau_{t-1}) + (1-\rho^{\mu})\ln(1-\tau) - \varepsilon_t^{\mu}, & (17)
\end{aligned}$$

with an associated equation describing the evolution of price dispersion, $\Delta_t = \int_0^1 \left(\frac{P_t(i)}{P_t}\right)^{-\eta} di$, which is not needed to tie down the equilibrium upon log-linearization. The model is then closed with instrument or targeting policy rule as discussed in the main text.

In order to render this model stationary we scale certain variables by the non-stationary level of technology, $A_t$ such that $k_t = K_t/A_t$ where $K_t = \{Y_t, C_t, W_t/P_t\}$. All other real variables are naturally stationary. Applying this scaling, the steady-state equilibrium conditions reduce to:

$$\begin{aligned}
N^{\varphi}X^{\sigma} &= w(1-\tau),\\
1 &= \beta R\pi^{-1}/\gamma = \beta r/\gamma,\\
y &= N = c,\\
X &= c(1-\theta),\\
\frac{\eta}{\eta-1} &= \frac{1}{w}.
\end{aligned}$$

This system yields

$$N^{\sigma+\varphi}(1-\theta)^{\sigma} = w(1-\tau). \tag{18}$$

which can be solved for $N$. Note that this expression depends on the real wage $w$, which can be obtained from the steady-state pricing decision of our monopolistically competitive firms. In Appendix B.1 we contrast this with the labor allocation that would be chosen by a social planner in order to fix the steady-state tax rate required to offset the net distortion implied by monopolistic competition and the consumption habits externality.

# B Derivation of Objective Functional Form

## B.1 The Social Planner's Problem

The subsidy level that ensures an efficient long-run equilibrium is obtained by comparing the steady-state solution of the social planner's problem with the steady state obtained in the decentralized equilibrium. The social planner ignores the nominal inertia and all other inefficiencies and chooses real allocations that maximize the representative consumer's utility subject to the aggregate resource constraint, the aggregate production function, and the law of motion for habit-adjusted consumption:

$$\max_{\{X_t^*, C_t^*, N_t^*\}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t u\left(X_t^*, N_t^*, \xi_t, A_t\right)$$

subject to

$$\begin{aligned} Y_t^* &= C_t^*, \qquad Y_t^* = A_t N_t^*, \\ X_t^* &= C_t^*/A_t - \theta C_{t-1}^*/A_{t-1}. \end{aligned}$$

The optimal choice implies the following relationship between the marginal rate of substitution between labor and habit-adjusted consumption and the intertemporal marginal rate of substitution in habit-adjusted consumption

$$\chi\left(N_t^*\right)^{\varphi}\left(X_t^*\right)^{\sigma} = \left(1-\theta\beta\right)\mathbb{E}_t\left(\frac{X_{t+1}^*\xi_{t+1}}{X_t^*\xi_t}\right)^{-\sigma}.$$

The steady state equivalent of this expression can be written as

$$\chi\left(N^*\right)^{\varphi+\sigma}\left(1-\theta\right)^{\sigma} = \left(1-\theta\beta\right).$$

If we contrast this with the allocation achieved in the steady-state of our decentralized equilibrium (18) we can see that the two will be identical whenever the tax rate is set optimally to be

$$\tau^* \equiv 1 - \frac{\eta}{\eta-1}(1-\theta\beta).$$

Notice that in the absence of habits the optimal tax rate would be negative, such that it is effectively a subsidy which offsets the monopolistic competition distortion. However, for the estimated values of the habits parameter the optimal tax rate is positive as the policy maker wishes to prevent households from overconsuming.

### B.2 Quadratic Representation of Social Welfare

Individual utility in period $t$ is

$$\Gamma_0 = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{X_t^{1-\sigma}\xi_t^{-\sigma}}{1-\sigma} - \frac{N_t^{1+\varphi}\xi_t^{-\sigma}}{1+\varphi}\right),$$

where $X_t = c_t - \theta c_{t-1}$ is habit-adjusted aggregate consumption after adjusting consumption for the level of productivity, $c_t = C_t/A_t$.

Linearization up to second order yields

$$\Gamma_0 = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left( \overline{X}^{1-\sigma} \left\{ \frac{1-\theta\beta}{1-\theta} \left( \hat{c}_t + \frac{1}{2}\hat{c}_t^2 \right) - \frac{1}{2}\sigma \hat{X}_t^2 - \sigma \hat{X}_t \hat{\xi}_t \right\} - \overline{N}^{1+\varphi} \left\{ \hat{N}_t + \frac{1}{2}(1+\varphi)\hat{N}_t^2 - \sigma \hat{N}_t \hat{\xi}_t \right\} \right) + tip(3),$$

where $tip(3)$ includes terms independent of policy and terms of third order and higher, and for every variable $Z_t$ with steady state value $Z$ we denote $\hat{Z}_t = \log(Z_t/Z)$.

The second order approximation to the production function yields the exact relationship $\hat{N}_t = \hat{\Delta}_t + \hat{y}_t$ , where $y_t = Y_t/A_t$ and $\Delta_t = \int_0^1 \left(\frac{P_t(i)}{P_t}\right)^{-\eta} di$. We substitute $\hat{N}_t$ out and follow Eser et al. (2009) in using

$$\sum_{t=0}^{\infty} \beta^t \hat{\Delta}_t = \frac{\alpha}{1-\alpha\beta}\hat{\Delta}_{-1} + \frac{1}{2}\sum_{t=0}^{\infty} \beta^t \frac{\alpha\eta}{(1-\beta\alpha)(1-\alpha)} \left( \hat{\pi}_t^2 + \frac{\zeta\alpha^{-1}}{(1-\zeta)} \left[\hat{\pi}_t - \hat{\pi}_{t-1}\right]^2 \right)$$

to yield

$$\Gamma_0 = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left( \overline{X}^{1-\sigma} \left\{ \frac{1-\theta\beta}{1-\theta} \left( \hat{c}_t + \frac{1}{2}\hat{c}_t^2 \right) - \frac{1}{2}\sigma \hat{X}_t^2 - \sigma \hat{X}_t \hat{\xi}_t \right\} - \overline{N}^{1+\varphi} \left( \begin{array}{c} \hat{y}_t + \frac{1}{2}\frac{\alpha\eta}{(1-\beta\alpha)(1-\alpha)} \left( \hat{\pi}_t^2 + \frac{\zeta\alpha^{-1}}{(1-\zeta)} \left[\hat{\pi}_t - \hat{\pi}_{t-1}\right]^2 \right) \\ + \frac{1}{2}(1+\varphi)\hat{y}_t^2 - \sigma \hat{y}_t \hat{\xi}_t \end{array} \right) \right) + tip(3).$$

The second order approximation to the national income identity yields

$$\hat{c}_t + \frac{1}{2}\hat{c}_t^2 = \hat{y}_t + \frac{1}{2}\hat{y}_t^2 + tip\,(3)\,.$$

Finally, we use the fact that in the efficient steady-state $\overline{X}^{1-\sigma}(1-\theta\beta) = (1-\theta)\overline{N}^{1+\varphi}$ and collect terms to arrive at

$$\Gamma_0 = -\frac{1}{2}\overline{N}^{1+\varphi}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \frac{\sigma(1-\theta)}{1-\theta\beta} \left( \hat{X}_t + \hat{\xi}_t \right)^2 + \varphi \left( \hat{y}_t - \frac{\sigma}{\varphi}\hat{\xi}_t \right)^2 + \frac{\alpha\eta}{(1-\beta\alpha)(1-\alpha)} \left( \hat{\pi}_t^2 + \frac{\zeta\alpha^{-1}}{(1-\zeta)} \left[\hat{\pi}_t - \hat{\pi}_{t-1}\right]^2 \right) \right\} + tip\,(3)\,.$$

Normalizing the coefficient on inflation to one and changing sign, we arrive at the following quadratic approximation to the loss function:

$$L_0 = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \frac{(1-\alpha\beta)(1-\alpha)\sigma(1-\theta)}{\alpha\eta(1-\theta\beta)} \left( \hat{X}_t + \hat{\xi}_t \right)^2 + \frac{(1-\beta\alpha)(1-\alpha)\varphi}{\alpha\eta} \left( \hat{y}_t - \frac{\sigma}{\varphi}\hat{\xi}_t \right)^2 + \hat{\pi}_t^2 + \frac{\zeta\alpha^{-1}}{(1-\zeta)} \left[\hat{\pi}_t - \hat{\pi}_{t-1}\right]^2 \right\} + tip\,(3)\,.$$

# C Policy

## C.1 Instrument Rule Specification and Solution

The US monetary policy is described by the following instrument rule with Markov-switching rule parameters $(\rho^R, \psi_1, \psi_2)$ between two regimes:

$$R_t = \rho^R(s_t)R_{t-1} + (1 - \rho^R(s_t))[\psi_1(s_t)\hat{\pi}_t + \psi_2(s_t)(\Delta\hat{y}_t + \hat{z}_t)] + \varepsilon_t^R. \tag{19}$$

The algorithm proposed by Farmer et al. (2011) is applied to solve the model. The model can be recast in the following system

$$\overset{A(s_t)}{\begin{bmatrix} \underset{(n-l)\times n}{a_1(s_t)} \\ \underset{l\times n}{a_2(s_t)} \end{bmatrix}} \underset{n\times 1}{x_t} = \overset{B(s_t)}{\begin{bmatrix} \underset{(n-l)\times n}{b_1(s_t)} \\ \underset{l\times n}{b_2(s_t)} \end{bmatrix}} \underset{n\times 1}{x_{t-1}} + \overset{\Psi(s_t)}{\begin{bmatrix} \underset{(n-l)\times k}{\psi_1(s_t)} \\ \underset{l\times k}{\psi_2(s_t)} \end{bmatrix}} \underset{k\times 1}{z_t} + \overset{\Pi}{\begin{bmatrix} \underset{(n-l)\times l}{0} \\ \underset{l\times l}{I} \end{bmatrix}} \underset{l\times 1}{\eta_t}, \tag{20}$$

where $x_t = [\hat{z}_t, \hat{\mu}_t, \hat{\xi}_t, \hat{y}_t, \hat{\pi}_t, \hat{R}_t, \mathbb{E}_t\hat{y}_{t+1}, \mathbb{E}_t\hat{\pi}_{t+1}]'$ is a vector of state variables. Vector $z_t = [\varepsilon_t^z, \varepsilon_t^\mu, \varepsilon_t^\xi, \varepsilon_t^R]$ stacks the exogenous shocks and $\eta_t$ is composed of rational expectation forecast errors. The latent value $s_t$ follows an two-state Markov chain, $s_t \in \{1, 2\}$, with transition matrix $P = [p_{ij}]$ defined as

$$p_{ij} = \Pr(s_t = i | s_{t-1} = j).$$

Theorem 1 in Farmer et al. (2011) shows that if $\{X_t, \eta_t\}_{t=1}^\infty$ is an MSV solution of the system (20), then

$$\begin{aligned} x_t &= V(s_t)F_1(s_t)x_{t-1} + V(s_t)G_1(s_t)z_t, \\ \eta_t &= -\left(F_2(s_t)x_{t-1} + G_2(s_t)z_t\right), \end{aligned}$$

where the matrix $\begin{bmatrix} A(s_t)V(s_t) & \Pi \end{bmatrix}$ is invertible and

$$B(s_t) = \begin{bmatrix} A(s_t)V(s_t) & \Pi \end{bmatrix} \begin{bmatrix} F_1(s_t) & F_2(s_t) \end{bmatrix}', \tag{21}$$

$$\Psi(s_t) = \begin{bmatrix} A(s_t)V(s_t) & \Pi \end{bmatrix} \begin{bmatrix} G_1(s_t) & G_2(s_t) \end{bmatrix}', \tag{22}$$

$$\sum_{i=1}^{2} p_{ij}F_2(s_t = i)V(s_{t-1} = j) = \underset{l\times(n-l)}{0}. \tag{23}$$

Without loss of generality, Farmer et al. (2011) assume that

$$A(s_t)V(s_t) = \begin{bmatrix} I_{n-l} & -X(s_t) \end{bmatrix}', \tag{24}$$

for some $l \times (n-l)$ matrix $X(s_t)$. Since

$$F_2(s_t) = \begin{bmatrix} 0_{l\times(n-l)} & I_l \end{bmatrix} \begin{bmatrix} A(s_t)\,V(s_t) & \Pi \end{bmatrix}^{-1} B(s_t) = \begin{bmatrix} X(s_t) & I_l \end{bmatrix} B(s_t),$$

equation (23) becomes

$$\begin{aligned} &\sum_{i=1}^{2} p_{ij} \begin{bmatrix} X(s_t=i) & I_l \end{bmatrix} B(s_t=i)\,A(s_{t-1}=j)^{-1} \begin{bmatrix} I_{n-l} & -X(s_{t-1}=j) \end{bmatrix}' \\ = & \underset{l\times(n-l)}{0},\ j \in \{1,2\} \end{aligned} \tag{25}$$

Therefore, the problem of finding an MSV solution can be reduced to that of finding the roots of the quadratic polynomial equation (25). Farmer et al. (2011) use Newton's method to compute roots as shown in Algorithm 1 in their paper. Once $X(s_t)$ is found, $V(s_t)$ can be subsequently solved using equation (24). With $V(s_t)$ obtained, equations (21) and (22) can be used to find $F_1(s_t), F_2(s_t), G_1(s_t)$ and $G_2(s_t)$ to obtain an MSV solution of the system (20).

To apply this procedure we start with a large number of initial guesses of $X(s_t)$ to explore all possible MSV solutions. In addition, we check whether the solutions are mean-square stable and select the stationary one for our estimation.

## C.2 Targeting Rules and Solution

We only present the solution under Quasi-commitment, as discretion and commitment follow Svensson and Williams (2007). Moreover, both of them are limiting cases of the quasi-commitment model.

The model (1)-(4) in the main text belongs to the class of linear models

$$\begin{bmatrix} I & 0 \\ 0 & H_{t+1} \end{bmatrix} \begin{bmatrix} X_{t+1} \\ \mathbb{E}_t x_{t+1} \end{bmatrix} = \begin{bmatrix} A_{11,t+1} & A_{12,t+1} \\ A_{21,t} & A_{22,t} \end{bmatrix} \begin{bmatrix} X_t \\ x_t \end{bmatrix} + \begin{bmatrix} B_{1,t+1} \\ B_{2,t} \end{bmatrix} [u_t] + \begin{bmatrix} C_{t+1} \\ 0 \end{bmatrix} [\epsilon_{t+1}], \tag{26}$$

where $X_t = [\hat{z}_t, \hat{\mu}_t, \hat{\xi}_t, \hat{y}_{t-1}, \hat{\pi}_{t-1}, \hat{R}_{t-1}]'$ is an $n_1 \times 1$ vector of predetermined variables (the state) in period $t$, $x_t = [\hat{y}_t, \hat{\pi}_t]'$ is an $n_2 \times 1$ vector of forward-looking variables in period $t$, $u_t = [\hat{R}_t]$ is an $n_p \times 1$ vector of central-bank instruments (control variables) in period $t$, and $\epsilon_t = [\varepsilon_t^z, \varepsilon_t^\mu, \varepsilon_t^\xi]$ is an $k \times 1$ vector of zero-mean i.i.d. shocks realized in period $t$ with covariance matrix $\Sigma$. Matrix $A_{22,t}$ is nonsingular. We denote $z_t = [X_t', x_t']'$.

The policy objective is a quadratic form:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left( z_t' \mathcal{W}_t z_t + 2 z_t' \mathcal{U}_t u_t + u_t' \mathcal{R}_t u_t \right).$$

Matrices $A_{11,t}$, $A_{12,t}$, $B_{1,t}$, $C_t$, $H_t$, $A_{21,t}$, $A_{22,t}$, $B_{2,t}$, $\mathcal{W}_t$, $\mathcal{U}_t$, and $\mathcal{R}_t$ (assumed to be of appropriate dimension) are random and can each take $n$ different values in period $t$, corresponding to the $n$ modes $j_t = 1, 2, ..., n$ in period $t$. We denote these values $A_{11,t} = A_{11,j_t}$ , $A_{12,t} = A_{12,j_t}$, and so forth. The modes $j_t$ follow a Markov process with constant transition probabilities:

$$P_{j_k} \equiv Pr\{j_{t+1} = k | j_t = j\} \qquad (j, k = 1, ..., n).$$

Schaumburg and Tambalotti (2007) and Debortoli and Nunes (2010) demonstrate that the optimization problem can be written as

$$\min_{\{u_t\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} (\beta (1-\upsilon))^t \left( z_t' \mathcal{W}_t z_t + 2 z_t' \mathcal{U}_t u_t + u_t' \mathcal{R}_t u_t + \beta \upsilon X_{t+1}' V_{t+1}^d X_{t+1} \right),$$

subject to

$$\begin{aligned} X_{t+1} &= A_{11 j_{t+1}} X_t + A_{12 j_{t+1}} x_t + B_{1 j_{t+1}} u_t, & (27) \\ (1-\upsilon) H_{j_{t+1}} \mathbb{E}_t x_{t+1} + \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} X_{t+1} &= A_{21 j_t} X_t + A_{22 j_t} x_t + B_{2 j_t} u_t, & (28) \end{aligned}$$

where $\Phi_{j_t}$ and $S$ are components of the solution to the corresponding discretionary problem, $x_t = \Phi_{j_t} X_t$ and the loss is $L_t(y_t) = \frac{1}{2} X_t' V_{t+1}^d X_t$. Here $\upsilon$ is probability of reneging on the previously chosen plan.

Substitute (27) into (28) to yield

$$\begin{aligned} (1-\upsilon) H_{j_{t+1}} \mathbb{E}_t x_{t+1} &= \left( A_{21 j_t} - \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} A_{11 j_{t+1}} \right) X_t \\ &\quad + \left( A_{22 j_t} - \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} A_{12 j_{t+1}} \right) x_t \\ &\quad + \left( B_{2 j_t} - \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} B_{1 j_{t+1}} \right) u_t. \end{aligned}$$

It is straightforward to bring the optimization problem into the minmax form (see Marcet and Marimon (2011) for the general method and Svensson (2010) for linear-quadratic problems)

$$\max_{\{\gamma_t\}_{t=0}^{\infty}} \min_{\{x_t, u_t\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} (\beta (1-\upsilon))^t \tilde{L}_t$$

with

$$\begin{aligned} \tilde{L}_t &= z_t' \mathcal{W}_{j_t} z_t + 2 z_t' \mathcal{U}_{j_t} u_t + u_t' \mathcal{R}_{j_t} u_t + \beta \upsilon X_{t+1}' V_{j_{t+1}}^d X_{t+1} \\ &\quad - \gamma_t' \left( \left( A_{21 j_t} - \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} A_{11 j_{t+1}} \right) X_t \right. \\ &\quad \left. + \left( A_{22 j_t} - \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} A_{12 j_{t+1}} \right) x_t + \left( B_{2 j_t} - \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} B_{1 j_{t+1}} \right) u_t \right) \\ &\quad + \frac{1}{\beta} \Xi_{t-1}' H_{j_t} x_t, \end{aligned}$$

and

$$\begin{aligned}
\Xi_t &= \gamma_t, \\
X_{t+1} &= A_{11j_{t+1}} X_t + A_{12j_{t+1}} x_t + B_{1j_{t+1}} u_t.
\end{aligned}$$

The Bellman equation is:

$$V(X_t, \Xi_{t-1}) = \max_{\gamma_t} \min_{(x_t, u_t)} \left( \tilde{L}_t + \beta(1-\upsilon) V(X_{t+1}, \Xi_t) \right).$$

Denote

$$\begin{aligned}
\hat{L}_t &= z_t' \mathcal{W}_{j_t} z_t + 2 z_t' \mathcal{U}_{j_t} u_t + u_t' \mathcal{R}_{j_t} u_t + \frac{1}{\beta} \Xi_{t-1}' H_{j_t} x_t - \gamma_t' (A_{21j_t} X_t + A_{22j_t} x_t + B_{2j_t} u_t) \\
&\quad + \gamma_t' \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} \left( A_{11j_{t+1}} X_t + A_{12j_{t+1}} x_t + B_{1j_{t+1}} u_t \right),
\end{aligned}$$

and

$$\tilde{u}_t = \begin{bmatrix} x_t \\ \gamma_t \\ u_t \end{bmatrix}, \tilde{X}_t = \begin{bmatrix} X_t \\ \Xi_{t-1} \end{bmatrix},$$

so that

$$\tilde{L}_t = \hat{L}_t + \beta \upsilon X_{t+1}' V_{j_{t+1}}^d X_{t+1}$$

and the Bellman equation is

$$V(X_t, \Xi_{t-1}) = \max_{\gamma_t} \min_{(x_t, u_t)} \left( \hat{L}_t + \beta \upsilon X_{t+1}' V_{j_{t+1}}^d X_{t+1} + \beta(1-\upsilon) V(X_{t+1}, \Xi_t) \right). \tag{29}$$

Then

$$\hat{L}_t = \begin{bmatrix} \tilde{X}_t \\ \tilde{u}_t \end{bmatrix}' \begin{bmatrix} Q_{j_t} & N_{j_t} \\ N_{j_t}' & R_{j_t} \end{bmatrix} \begin{bmatrix} \tilde{X}_t \\ \tilde{u}_t \end{bmatrix},$$

where

$$\begin{aligned}
Q_{j_t} &= \begin{bmatrix} \mathcal{W}_{11,j_t} & 0 \\ 0 & 0 \end{bmatrix}, \\
N_{j_t} &= \begin{bmatrix} \mathcal{W}_{12,j_t} & -\left( A_{21j_t} - \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} A_{11j_{t+1}} \right)' & \mathcal{U}_{1,j_t} \\ \frac{1}{\beta} H_{j_t} & 0 & 0 \end{bmatrix}, \\
R_{j_t} &= \begin{bmatrix} \mathcal{W}_{22,j_t} & \left( \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} A_{12j_{t+1}} - A_{22j_t} \right)' & \mathcal{U}_{2,j_t} \\ \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} A_{12j_{t+1}} - A_{22j_t} & 0 & \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} B_{1j_{t+1}} - B_{2j_t} \\ \mathcal{U}_{2,j_t}' & \left( \upsilon H_{j_{t+1}} \Phi_{j_{t+1}} B_{1j_{t+1}} - B_{2j_t} \right)' & \mathcal{R}_{j_t} \end{bmatrix}.
\end{aligned}$$

The terms in the RHS of Bellman equation (29) can be written as

$$\beta \upsilon X_{t+1}' V_{j_{t+1}}^d X_{t+1} + \beta(1-\upsilon) \tilde{X}_{t+1}' V_{j_{t+1}} \tilde{X}_{t+1} = \beta \tilde{X}_t' \tilde{V}_{j_{t+1}} \tilde{X}_{t+1},$$

where

$$\tilde{V}_{j_{t+1}} = \left(\upsilon S' V^d_{j_{t+1}} S + (1-\upsilon) V_{j_{t+1}}\right),$$

and $S = \begin{bmatrix} I & 0 \end{bmatrix}$.

The Bellman equation (29), therefore, can be rewritten

$$\tilde{X}_t' V_{j_t} \tilde{X}_t = \max_{\tilde{u}_t} \min \left( \begin{bmatrix} \tilde{X}_t \\ \tilde{u}_t \end{bmatrix}' \begin{bmatrix} Q_{j_t} & N_{j_t} \\ N_{j_t}' & R_{j_t} \end{bmatrix} \begin{bmatrix} \tilde{X}_t \\ \tilde{u}_t \end{bmatrix} + \beta \tilde{X}_t' \tilde{V}_{j_{t+1}} \tilde{X}_{t+1} \right).$$

Note that $V_{j_t}$in the LHS is not the same as $\tilde{V}_{j_{t+1}}$ in the RHS, and the constraints can be rewritten as

$$\tilde{X}_{t+1} = \tilde{A}_{j_{t+1}} \tilde{X}_t + \tilde{B}_{j_{t+1}} \tilde{u}_t + \tilde{C}_{j_{t+1}} \epsilon_{t+1},$$

where

$$\begin{aligned} \tilde{A}_{j_t} &= \begin{bmatrix} A_{11j_t} & 0 \\ 0 & 0 \end{bmatrix}, \tilde{B}_{j_t} = \begin{bmatrix} A_{12j_t} & 0 & B_{1j_t} \\ 0 & I & 0 \end{bmatrix}, \\ \tilde{C}_{j_t} &= \begin{bmatrix} C \\ 0 \end{bmatrix} \end{aligned}$$

FOC wrt $\tilde{u}_t$ yields

$$0 = \left(N_{j_t}' + \beta \tilde{B}_{j_{t+1}}' \tilde{V}_{j_{t+1}} \tilde{A}_{j_{t+1}}\right) \tilde{X}_t + \left(R_{j_t} + \beta \tilde{B}_{j_{t+1}}' \tilde{V}_{j_{t+1}} \tilde{B}_{j_{t+1}}\right) \tilde{u}_t = K_{j_t} \tilde{X}_t + J_{j_t} \tilde{u}_t,$$

where

$$\begin{aligned} J_{j_t} &= R_{j_t} + \beta \tilde{B}_{j_{t+1}}' \tilde{V}_{j_{t+1}} \tilde{B}_{j_{t+1}}, \\ K_{j_t} &= N_{j_t}' + \beta \tilde{B}_{j_{t+1}}' \tilde{V}_{j_{t+1}} \tilde{A}_{j_{t+1}}, \end{aligned}$$

so that optimal policy is

$$\tilde{u}_t = -J_{j_t}^{-1} K_{j_t} \tilde{X}_t = F_{j_t} \tilde{X}_t.$$

The iterative algorithm can be written as follows. For each Markov state $j$ and at each iteration

1. Suppose $V = \begin{bmatrix} V_{11} & V_{12} \\ V_{12}' & V_{22} \end{bmatrix}$, $V^d = V_{11}$ and $\Phi$ are known. (This is an initial guess which will be updated in the following steps.)

2. Construct $\begin{bmatrix} Q_{j_t} & N_{j_t} \\ N_{j_t}' & R_{j_t} \end{bmatrix}$, taking into account future states $j_{t+1}$.

3. Construct $\tilde{A}_{j_{t+1}}, \tilde{B}_{j_{t+1}}$ (use probabilities).

4. Construct

$$\tilde{V}_{j_{t+1}} = \upsilon S' V^d_{j_{t+1}} S + (1-\upsilon) V_{j_{t+1}}.$$

5. Find

$$\begin{aligned} J_j &= R_j + \beta P_{jj} \tilde{B}_j' \tilde{V}_j \tilde{B}_j + \beta \sum_{k \neq j} P_{jk} \tilde{B}_k' \tilde{V}_k \tilde{B}_k, \\ K_j &= N_j' + \beta P_{jj} \tilde{B}_j' \tilde{V}_j \tilde{A}_j + \beta \sum_{k \neq j} P_{jk} \tilde{B}_k' \tilde{V}_k \tilde{A}_k, \\ V_{j_t} &= Q_{j_t} - K_j' J_j^{-1} K_j + \beta P_{jj} \tilde{A}_j' \tilde{V}_j \tilde{A}_j + \beta \sum_{k \neq j} P_{jk} \tilde{A}_k' \tilde{V}_k \tilde{A}_k, \end{aligned}$$

$$F_j = -J_j^{-1} K_j.$$

6. Update $\Phi = F_{xX}$ and $V = V_{j_t}, V^d = V_{11}$ and repeat steps 2-6 until the fixed point is obtained.

# D Estimation Strategy

Our empirical analysis uses the US data on output growth ($\Delta GDP_t$), inflation ($INF_t$), and nominal interest rates ($INT_t$) from 1961Q1 up to 2008Q3, just before nominal interest rates were reduced to their effective lower bound of 0.5%.[2] The measurement equations are specified as:

$$\begin{aligned} \Delta GDP_t &= \gamma^Q + \Delta \hat{y}_t + \hat{z}_t, \\ INF_t &= \pi^A + 4\hat{\pi}_t, \\ INT_t &= r^A + \pi^A + 4\gamma^Q + 4\hat{R}_t. \end{aligned}$$

In estimation, the likelihood function is approximated using Kim (1994)'s filter due to the presence of Markov-switching parameters, and is then combined with the prior distribution to obtain the posterior distribution. Sims (2002) optimization routine CSMINWEL is used to find the posterior modes. The inverse Hessian is then calculated at these posterior modes and used as the covariance matrix of the proposal distribution. It is scaled to yield a target acceptance rate

[2]The specific data series used are the Effective Federal Funds Rate - FEDFUNDS, Gross Domestic Product in United States-USARGDPQDSNAQ and the Gross Domestic Product: Implicit Price Deflator-GDPDEF. All data are seasonally adjusted and at quarterly frequencies. Output growth is the log difference of real GDP, multiplied by 100. Inflation is the log difference of the associated implicit price deflator, scaled by 400. All data are taken from the FRED database.

of 25%-40%. We adopt Schorfheide (2005)'s strategy that employs a random walk Metropolis-Hastings algorithm to generate 500,000 draws with the first 200,000 draws being discarded before saving every 20th draw from the remaining draws.[3]

Finally, the log marginal likelihood values for each model are computed using both the modified harmonic mean estimator of Geweke (1999) and Sims et al. (2008). The latter is designed for models with time-varying parameters, where the posterior density may be non-Gaussian.

When estimating commitment we follow Givens (2012) to update the Lagrange multiplies. Alternatively, Ilbas (2010) and Adolfson et al. (2011) use presample to initialize the Lagrange multipliers. Both approaches give very similar results.

# E Identification Tests

## E.1 Komunjer and Ng (2011)

We use the Komunjer and Ng (2011) identification test to analyze the main models presented in Table 2. Komunjer and Ng (2011) study the local identification of a DSGE model from its linearized solution. Their test uses the restrictions implied by equivalent spectral densities to obtain rank and order conditions for identification. Minimality and left-invertibility are necessary and sufficient conditions for identification. In addition, Komunjer and Ng (2011) discuss the identification conditions for both singular and nonsingular systems. The singular case applies to our targeting rule models where the number of shocks are equal to the number of observations, while the nonsingular case applies to the instrument rule based models where we have an additional unsystematic interest rate shock.

It is important to note that the Komunjer and Ng (2011) identification test only applies to covariance stationary processes. Therefore, the parameters associated with Markov-switching shock variances cannot be incorporated into the test.[4] As for monetary policy changes, we can solve our model assuming that monetary policy stays in one regime, even though the private agents in the economy are aware that there is a probability of monetary policy switching to a different regime. In addition, the test can only applied to the parameters that enter into the model solution. Therefore, we cannot test the identification of parameters that only contribute to data means such as $\pi^A$ and $\gamma^Q$, while the identification of the real interest rate, $r^A$, can be tested as it links to the discount factor, $\beta$. This also implies that we will not be able to apply Komunjer and Ng (2011) to asses the identification of the instrument rule which allows for switches in the

[3] Geweke (1992) convergence diagnostics indicate that convergence is achieved. These are available upon request.

[4] To consider the case of switches in shock volatility we employ Koop et al (2013) as discussed in Section E.2.

inflation target (i.e. Rule-Target in Table 2).

For the results under discretion and commitment presented in Table 2, we have the parameter vector $\boldsymbol{\theta} \equiv (\sigma,\, \alpha,\, \varphi,\, \zeta,\, \theta,\, \omega_1,\, \omega_2,\, \omega_3,\, \omega_{\pi(s=2)},\, \rho^z,\, \rho^\mu,\, \rho^\xi,\, \sigma_{z(S=1)},\, \sigma_{\mu(S=1)},\, \sigma_{\xi(S=1)},\, r^A,\, p_{11},\, p_{22})$ of dimension $n_{\boldsymbol{\theta}} = 18$. To apply the test, we solve for the model and find a minimal representation of the solution as follows

$$\begin{aligned} X_{1t+1} &= A(\boldsymbol{\theta})X_{1t} + B(\boldsymbol{\theta})\varepsilon_{t+1}, \\ Y_{t+1} &= C(\boldsymbol{\theta})X_{1t} + D(\boldsymbol{\theta})\varepsilon_{t+1}. \end{aligned}$$

In this new system, $X_{1t+1} \equiv \left(\hat{z}_t, \hat{\mu}_t, \hat{\xi}_t, \hat{y}_t, \hat{\pi}_t\right)'$ and $n_X = 5$. $Y_{t+1} \equiv (\Delta GDP_t, INF_t, INT_t)$ and $\varepsilon_t \equiv (\varepsilon_t^z, \varepsilon_t^\mu, \varepsilon_t^\xi)$ are the observables and shocks. As $n_Y = n_\varepsilon = 3$, the model is square. Proposition 2-S in Komunjer and Ng (2011) can be employed to assess identification. As for quasi-commitment, the parameter vector $\theta \equiv (\upsilon,\, \sigma,\, \alpha,\, \varphi,\, \zeta,\, \theta,\, \omega_1,\, \omega_2,\, \omega_3,\, \omega_{\pi(s=2)},\, \rho^z,\, \rho^\mu,\, \rho^\xi,\, \sigma_{z(S=1)},\, \sigma_{\mu(S=1)},\, \sigma_{\xi(S=1)},\, r^A,\, p_{11},\, p_{22})$ has dimension of $n_{\boldsymbol{\theta}} = 19$. However, its minimal representation is consistent with Discretion and Commitment.

Using the same notation as in Komunjer and Ng (2011), we check the rank of $\Delta^s(\theta_0)$ which is the matrix of partial derivatives of $\delta^s(\boldsymbol{\theta}, T, U)$ evaluated at $(\boldsymbol{\theta}_0, \mathbf{I}_{n_X}, \mathbf{I}_{n\varepsilon})$.

$$\begin{aligned} \Delta^s(\theta_0) &\equiv \left( \frac{\partial \delta^s(\boldsymbol{\theta}, \mathbf{I}_{n_X}, \mathbf{I}_{n_\varepsilon})}{\partial \boldsymbol{\theta}}\ \frac{\partial \delta^s(\boldsymbol{\theta}, \mathbf{I}_{n_X}, \mathbf{I}_{n_\varepsilon})}{\partial vecT}\ \frac{\partial \delta^s(\boldsymbol{\theta}, \mathbf{I}_{n_X}, \mathbf{I}_{n_\varepsilon})}{\partial vecU} \right)_{\boldsymbol{\theta}=\theta_0} \\ &\equiv \left(\Delta_\Lambda^s(\theta_0)\ \Delta_T^s(\theta_0)\ \Delta_U^s(\theta_0)\right). \end{aligned}$$

The rank of $\Delta^s(\theta_0)$ required for identification is

$$\begin{aligned} rank(\Delta^s(\theta_0)) &= rank\left(\Delta_\Lambda^s(\theta_0)\ \Delta_T^s(\theta_0)\ \Delta_U^s(\theta_0)\right) \\ &= n_{\boldsymbol{\theta}} + n_X^2 + n_\varepsilon^2. \end{aligned}$$

As for the rule-based model, we add an additional unsystematic interest rate shock, and thus $\varepsilon_t \equiv (\varepsilon_t^z, \varepsilon_t^\mu, \varepsilon_t^\xi, \varepsilon_t^R)$. Therefore, Proposition 2-NS in Komunjer and Ng (2011) is used to study identification when $n_Y < n_\varepsilon$. The parameter vector $\boldsymbol{\theta} \equiv (\sigma,\, \alpha,\, \varphi,\, \zeta,\, \theta,\, \rho_{s=1}^R,\, \psi_{1,s=1},\, \psi_{2,s=1},\, \rho_{s=2}^R,\, \psi_{1,s=2},\, \psi_{2,s=2},\, \rho^z,\, \rho^\mu,\, \rho^\xi,\, \sigma_{z(S=1)},\, \sigma_{\mu(S=1)},\, \sigma_{\xi(S=1)},\, \sigma_{R(S=1)},\, r^A,\, p_{11},\, p_{22})$ has a dimension of $n_{\boldsymbol{\theta}} = 21$. In addition, as we allow for a lagged interest rate in the generalized Taylor rule, $X_{1t+1} \equiv \left(\hat{z}_t, \hat{\mu}_t, \hat{\xi}_t, \hat{y}_t, \hat{\pi}_t, \hat{R}_t\right)'$ and $n_X = 6$.

Again, adopting the notation of Komunjer and Ng (2011), for the nonsingular system, we check the rank of $\Delta^s(\theta_0)$ which is the matrix of partial derivatives of $\delta^s(\boldsymbol{\theta}, T)$ evaluated at $(\boldsymbol{\theta}_0, \mathbf{I}_{n_X})$.

$$\begin{aligned}\Delta^{NS}(\theta_0) &\equiv \left(\frac{\partial\delta^{NS}(\boldsymbol{\theta}, \mathbf{I}_{n_X})}{\partial\boldsymbol{\theta}}\ \frac{\partial\delta^{NS}(\boldsymbol{\theta}, \mathbf{I}_{n_X})}{\partial vecT}\right)_{\boldsymbol{\theta}=\theta_0} \\ &\equiv \left(\Delta_\Lambda^{NS}(\theta_0)\ \Delta_T^{NS}(\theta_0)\right).\end{aligned}$$

The rank of $\Delta^s(\theta_0)$ required for identification is

$$\begin{aligned}rank(\Delta^{NS}(\theta_0)) &= rank\left(\Delta_\Lambda^{NS}(\theta_0)\ \Delta_T^{NS}(\theta_0)\right) \\ &= n_{\boldsymbol{\theta}} + n_X^2.\end{aligned}$$

As in Komunjer and Ng (2011). we also consider 11 levels of tolerance along with the Matlab default to analyze the ranks of $\Delta^s(\theta_0)$ and $\Delta^{NS}(\theta_0)$ for both the singular and nonsingular systems. We use the change in rank as tolerance tightens to identify problematic parameters. The Komunjer and Ng (2011) test does not indicate that any parameters are unidentified. In Table E1 the required rank for identification of each model is presented, along with the Tol at which the model passes the rank requirement.

Table E1: Komunjer and Ng's (2011) identification test

| | Tol | $\Delta_\Lambda^s$ | $\Delta_T^s$ | $\Delta_U^s$ | $\Delta^s$ | Pass |
|---|---|---|---|---|---|---|
| Discretion | | | | | | |
| | Required | 18 | 25 | 9 | 52 | |
| Regime 1 More Conservative | $e-10$ | 18 | 25 | 9 | 52 | YES |
| Regime 2 Less Conservative | $e-9$ | 18 | 25 | 9 | 52 | YES |
| Commitment | | | | | | |
| | Required | 18 | 25 | 9 | 52 | |
| Regime 1 More Conservative | $e-5$ | 18 | 25 | 9 | 52 | YES |
| Regime 2 Less Conservative | $e-5$ | 18 | 25 | 9 | 52 | YES |
| Quasi-Commitment | | | | | | |
| | Required | 19 | 25 | 9 | 53 | |
| Regime 1 More Conservative | $e-6$ | 19 | 25 | 9 | 53 | YES |
| Regime 2 More Conservative, reneging | $e-10$ | 19 | 25 | 9 | 53 | YES |
| Regime 3 Less Conservative | $e-6$ | 19 | 25 | 9 | 53 | YES |
| Regime 4 Less Conservative, reneging | $e-10$ | 19 | 25 | 9 | 53 | YES |
| Rule - Parameters | | | | | | |
| | Required | 21 | 36 | – | 57 | |
| Regime 1 More Conservative | $e-7$ | 21 | 36 | – | 57 | YES |
| Regime 2 Less Conservative | $e-7$ | 21 | 36 | – | 57 | YES |

## E.2 Koop, Pesaran, Smith (2013)

Given that Komunjer and Ng (2011) only applies to covariance stationary processes, we also analyze the identification of the remaining parameters in our model using the Koop et al. (2013) Bayesian learning-rate indicator. However, it is worth noting that this indicator does not propose a 'Yes/No' answer to the question of whether a given parameter is identified as in Komunjer and Ng (2011). It indicates the degree of identification. The advantage of this approach is that it can be easily applied to models with Markov-switching parameters, since it requires only a few additional steps beyond a conventional Bayesian estimation. This test is based on Bayesian asymptotic theory. The theory states that the role of the prior of a parameter vanishes with increasing sample size. As a result, the posterior of a parameter asymptotically converges to its true value. Implementing this indicator involves simulating artificial datasets of increasing size from a DSGE model, and then estimating this model using these datasets. A parameter is said to be identified if its posterior precision increases with the sample size. However, if a parameter is weakly or not identifiable, its posterior precision will be updated at a rate lower than the sample size.

It is important to note that a prerequisite in using Koop et al. (2013) Bayesian learning rate indicator is that a subset of parameters are known to be identifiable. A vector of model parameters $\boldsymbol{\theta} = [\boldsymbol{\theta}_u, \boldsymbol{\theta}_i]'$ is split into $\boldsymbol{\theta}_i$, known to be identified and $\boldsymbol{\theta}_u$ under question. Therefore, the results obtained from Komunjer and Ng (2011) are complementary in applying this indicator, such that we use Koop et al. (2013) to assess those parameters that are unable to be included in our application Komunjer and Ng (2011). For example, under discretion, $\boldsymbol{\theta}_i \equiv (\sigma,\ \alpha,\ \varphi,\ \zeta,\ \theta,\ \omega_1,\ \omega_2,\ \omega_3,\ \omega_{\pi(s=2)},\ \rho^z,\ \rho^\mu,\ \rho^\xi,\ \sigma_{z(S=1)},\ \sigma_{\mu(S=1)},\ \sigma_{\xi(S=1)},\ r^A,\ p_{11},\ p_{22})$ and $\boldsymbol{\theta}_u \equiv (q_{11},\ q_{22}, \sigma_{z(S=2)},\ \sigma_{\mu(S=2)},\ \sigma_{\xi(S=2)},\ \pi^A,\ \gamma^Q)$, where $\boldsymbol{\theta}_u$ includes parameters associated with Markov-switching shock variances.

In addition, Koop et al. (2013) assume Gaussian priors for model parameters in order to obtain an analytical solution for the posterior precision when the sample period reaches infinity. However, for most DSGE models in the literature, the priors are non-gaussian. Although the assumption of Gaussian priors can be relaxed, the exact expression of the posterior precision will differ from those illustrated in Koop et al. (2013). Caglar et al. (2011) demonstrate how to apply this indicator to a medium-sized DSGE model with more complicated priors. They suggest treating the Hessian at the posterior mode as the measure of posterior precision. We do not present the technical details of this test as these can be found in Koop et al. (2013).

To implement the Bayesian learning rate indicator, we first need to produce samples of artificial data from our models. Models that incorporate Markov-switching parameters complicate the data generating processes (DGPs). To simulate data from a Markov-switching model, we need to set not only the model parameters but also the probabilities of each regime. We set the parameters equal to the posteriors presented in Table 2. In addition, we generate the probabilities of different sample sizes of $T = 100, 1000, 10000, 20000$. Unlike the fixed parameter model used to generate artificial data discussed in Koop et al. (2013) and Caglar et al. (2011), we cannot generate a single artificial dataset of $T = 20000$ and then take subsets of it to produce smaller subsamples. This is because for our Markov-switching models the probabilities in each sample size should correspond to the estimated transition probabilities (i.e.$p_{11}, p_{22}, q_{11}, q_{22}$ ). Therefore, the generated artificial datasets for smaller $T$ are not subsamples of larger $T$. This may contribute to the variation in the results across samples. However, we try to make the use of this indicator as comparable as possible across models. To do so, we use the same seed for the random number generator for all DGPs from our models presented in Table 2.

Table E2: Posterior precision divided by sample size

| Parameters | $n = 100$ | $n = 1000$ | $n = 10000$ | $n = 20000$ |
|---|---|---|---|---|
| Discretion - Parameters associated with the MS shock volatility | | | | |
| $\sigma_{\xi(S=2)}$ | 2.646 | 0.924 | 2.773 | 1.323 |
| $\sigma_{\mu(S=2)}$ | 3.399 | 1.051 | 2.950 | 1.726 |
| $\sigma_{z(S=2)}$ | 4.331 | 1.209 | 2.907 | 1.677 |
| $q_{11}$ | 16.171 | 3.220 | 7.444 | 5.666 |
| $q_{22}$ | 12.066 | 16.622 | 11.746 | 11.687 |
| $\pi^A$ | 0.037 | 0.010 | 0.012 | 0.028 |
| $\gamma^Q$ | 4.455 | 4.103 | 5.011 | 3.568 |
| Rule-Target - Parameters associated with the MS inflation targets & shock volatility | | | | |
| $\sigma_{\xi(S=2)}$ | 0.876 | 0.659 | 0.783 | 0.879 |
| $\sigma_{\mu(S=2)}$ | 0.277 | 0.038 | 0.010 | 0.008 |
| $\sigma_{z(S=2)}$ | 0.325 | 0.965 | 0.772 | 0.859 |
| $\sigma_{R(S=2)}$ | 0.164 | 0.035 | 0.004 | 0.003 |
| $\pi^A_{s=1}$ | 0.154 | 0.180 | 0.248 | 0.188 |
| $\pi^A_{s=2}$ | 0.206 | 0.259 | 0.215 | 0.143 |
| $\gamma^Q$ | 2.638 | 3.641 | 2.816 | 3.104 |
| $p_{11}$ | 5.691 | 1.552 | 0.012 | 1.999 |
| $p_{22}$ | 5.513 | 2.486 | 0.283 | 0.127 |
| $q_{11}$ | 9.559 | 9.515 | 13.602 | 10.390 |
| $q_{22}$ | 2.590 | 2.090 | 1.216 | 1.755 |

In Table E2 we only present the parameters whose identification cannot be verified by using Komunjer and Ng (2011) under discretion and the instrument rule with Markov-switching inflation targets.[5] As discussed in Koop et al. (2013), the posterior precision of identified parameters need not to rise monotonically with $T$. The posterior precision may, in fact, fall before rising depending on the prior type. In addition, Koop et al. (2013) show that the posterior precision of an unidentified parameter will shrink to zero very quickly as $T$ increases. To make our results robust, we increase $T = 10000$,the largest sample size used in Koop et al. (2013) to $T = 20000$. It can be seen that no element of the normalized posterior precision of $\boldsymbol{\theta}_u$ collapses to zero when $T = 20000$. However, for the instrument rule with Markov-switching inflation targets, the posterior precision of standard deviations of the cost-push shock and interest rate rule shock (i.e.$\sigma_{\mu(S=2)}$ and $\sigma_{R(S=2)}$) decline quickly with the enlargement of $T$.

# F Model Selection

Bayes factors are used to rank descriptions of policy in the main text. Alternatively, model posterior probabilities can be calculated to evaluate relative model fit. Table F1 presents posterior probabilities for models in Table 2. An equal model prior probability is placed on the five competing models. The model posterior probability is then given by

$$p(M_k|Y) = \frac{p(Y_{1:T}, |M_k)}{\Sigma_{k=1}^{n} p(Y_{1:T}, |M_k)}, \ n = 5, \tag{30}$$

where $p\left(Y_{1:T}, |M_k\right)$ is the marginal data density for each model. It can be seen that even though equal model priors are placed on the five competing models, the data chooses to tilt the posterior probability towards discretion.

Table F1: model posterior probabilities

| Discretion | Rule - Parameters | Rule - Target | Quasi-Commitment | Commitment |
|---|---|---|---|---|
| 0.99 | 0.01 | 0.00 | 0.00 | 0.00 |

The posterior probabilities are often used as weights for model averaging to forecast future observations. Forecasting applications of Bayesian model averaging in the economics literature include Min and Zellner (1993), Wright (2008) and Del Negro et al. (2016).

While it is generally preferable to average across all models with nonzero posterior probabilities, we use the posterior probabilities for model selection to choose one model that has the highest posterior probability to conduct policy analysis. As discussed in Del Negro (2011), a

[5] The rule-based model with fixed parameters are identifiable as indicated by Komunjer and Ng (2011). We can provide the results of the Koop et.al. (2013) test for the other models in Table 2 upon request. In all cases all parameters are identified.

Table G1: Full Table 3 No Switching

| Parameters | Simple Rule | Quasi-Commitment | Discretion | Commitment |
|---|---|---|---|---|
| Model Parameters | | | | |
| $\sigma$ | 2.802 [2.407,3.188] | 2.254 [1.904,2.604] | 2.722 [2.338,3.101] | 2.832 [2.477,3.062] |
| $\alpha$ | 0.779 [0.751,0.807] | 0.798 [0.776,0.822] | 0.760 [0.734,0.786] | 0.768 [0.734,0.789] |
| $\zeta$ | 0.103 [0.039,0.166] | 0.170 [0.087,0.252] | 0.156 [0.066,0.241] | 0.594 [0.489,0.737] |
| $\theta$ | 0.823 [0.685,0.964] | 0.421 [0.210,0.627] | 0.476 [0.267,0.680] | 0.643 [0.444,0.782] |
| $\varphi$ | 2.417 [2.005,2.833] | 1.835 [1.443,2.219] | 2.387 [2.146,2.627] | 2.312 [2.046,2.749] |
| Shock Processes | | | | |
| $\rho^{\xi}$ | 0.899 [0.859 0.941] | 0.894 [0.862 0.927] | 0.845 [0.806 0.885] | 0.862 [0.793 0.903] |
| $\rho^{\mu}$ | 0.500 [0.246 0.747] | 0.951 [0.926 0.976] | 0.947 [0.922 0.974] | 0.968 [0.948 0.989] |
| $\rho^{z}$ | 0.320 [0.219 0.418] | 0.217 [0.168 0.265] | 0.223 [0.172 0.275] | 0.199 [0.132 0.241] |
| $\sigma_{\xi}$ | 0.987 [0.715 1.247] | 1.104 [0.647 1.550] | 0.763 [0.486 1.039] | 0.983 [0.706 1.165] |
| $\sigma_{\mu}$ | 0.567 [0.341 0.788] | 0.448 [0.255 0.629] | 0.489 [0.356 0.613] | 3.628 [2.816 4.916] |
| $\sigma_{z}$ | 0.797 [0.731 0.863] | 0.845 [0.771 0.915] | 0.827 [0.756 0.896] | 0.770 [0.701 0.815] |
| $\sigma_{R}$ | 0.251 [0.227 0.273] | – | – | – |

continued on the next page

model selection approach is likely to provide a good approximation if the posterior probability of one model is very close to one, the probabilities associated with all other specifications are very small.

# G Additional Results

This section contains additional results not reported in the main paper.

Table G1: Full Table 3 Estimation Results - No Switching – continued

| Parameters | Simple Rule | Quasi-Commitment | Discretion | Commitment |
|---|---|---|---|---|
| | | Data Means | | |
| $r^A$ | 0.706 [0.246,1.139] | 0.759 [0.143,1.330] | 0.966 [0.352,1.569] | 1.088 [0.459,1.540] |
| $\pi^A$ | 4.746 [3.800,5.677] | 2.586 [1.899,3.095] | 2.656 [1.008,4.221] | 4.050 [3.642,4.674] |
| $\gamma^Q$ | 0.688 [0.547,0.826] | 0.737 [0.613,0.861] | 0.716 [0.593,0.835] | 0.726 [0.594,0.797] |
| | | Policy Parameters | | |
| $\rho^R$ | 0.791 [0.756,0.826] | – | – | – |
| $\psi_1$ | 1.716 [1.455,1.972] | – | – | – |
| $\psi_2$ | 0.492 [0.290,0.697] | – | – | – |
| $\omega_1$ | – | 0.703 [0.552,0.861] | 0.458 [0.287,0.627] | 0.627 [0.490,0.808] |
| $\omega_2$ | – | 0.828 [0.727,0.935] | 0.758 [0.628,0.901] | 0.446 [0.316,0.620] |
| $\omega_3$ | – | 0.390 [0.163,0.619] | 0.451 [0.213,0.692] | 0.489 [0.268,0.712] |
| $\upsilon$ | – | 0.556 [0.329,0.816] | – | – |
| | | Log Marginal Data Densities and Bayes Factors | | |
| Geweke (1999) | −841.01 (1.00) | −841.67 (1.94) | −842.49 (4.41) | −855.43 (1.84$e$+6) |
| Sims et al. (2008) | −841.09 (1.00) | −841.54 (1.57) | −842.69 (4.96) | −858.26 (2.85$e$+7) |

Table G2: Full Table 4 Switches in Policy Only

| Parameters | Discretion | Quasi-Commitment | Rule - Parameters | Rule - Target | Commitment |
|---|---|---|---|---|---|
| **Model Parameters** | | | | | |
| $\sigma$ | 2.896 [2.500,3.288] | 2.550 [2.190,2.896] | 2.621 [2.382,2.861] | 2.791 [2.403,3.187] | 2.921 [2.560,3.277] |
| $\alpha$ | 0.731 [0.706,0.758] | 0.755 [0.733,0.777] | 0.775 [0.747,0.803] | 0.779 [0.750,0.807] | 0.770 [0.744,0.796] |
| $\zeta$ | 0.155 [0.069,0.239] | 0.182 [0.091,0.274] | 0.102 [0.038,0.163] | 0.123 [0.054,0.195] | 0.229 [0.078,0.366] |
| $\theta$ | 0.479 [0.286,0.835] | 0.371 [0.192,0.543] | 0.825 [0.698,0.954] | 0.810 [0.658,0.961] | 0.606 [0.388,0.843] |
| $\varphi$ | 2.331 [1.916,2.757] | 2.249 [1.861,2.632] | 2.425 [2.025,2.848] | 2.410 [2.003,2.846] | 2.271 [1.872,2.679] |
| **Shock Processes** | | | | | |
| $\rho^{\xi}$ | 0.805 [0.766,0.844] | 0.899 [0.868,0.929] | 0.887 [0.850,0.927] | 0.898 [0.858,0.941] | 0.904 [0.877,0.933] |
| $\rho^{\mu}$ | 0.957 [0.937,0.978] | 0.942 [0.916,0.967] | 0.501 [0.250,0.748] | 0.499 [0.250,0.751] | 0.986 [0.978,0.995] |
| $\rho^{z}$ | 0.213 [0.164,0.261] | 0.218 [0.167,0.269] | 0.307 [0.208,0.403] | 0.317 [0.218,0.417] | 0.210 [0.154,0.268] |
| $\sigma_{\xi}$ | 0.515 [0.289,0.719] | 0.851 [0.551,1.145] | 0.981 [0.755,1.199] | 0.848 [0.609,1.090] | 0.797 [0.511,1.069] |
| $\sigma_{\mu}$ | 0.444 [0.327,0.554] | 0.628 [0.440,0.812] | 0.275 [0.169,0.382] | 0.569 [0.340,0.795] | 2.325 [1.697,2.947] |
| $\sigma_{z}$ | 0.829 [0.755,0.896] | 0.831 [0.760,0.902] | 0.797 [0.169,0.382] | 0.795 [0.727,0.861] | 0.779 [0.711,0.846] |
| $\sigma_{R}$ | – | – | 0.235 [0.213,0.256] | 0.252 [0.229,0.275] | – |
| **Data Means** | | | | | |
| $r^{A}$ | 0.766 [0.303,1.213] | 0.997 [0.377,1.591] | 0.695 [0.276,1.105] | 0.662 [0.239,1.054] | 0.975 [0.358,1.561] |
| $\pi^{A}_{(s=1)}$ | 2.683 [1.275,4.022] | 2.097 [1.770,2.431] | 3.736 [3.183,4.299] | 4.234 [3.470,4.995] | 3.064 [2.733,3.411] |
| $\pi^{A}_{(s=2)}$ | – | – | – | 6.058 [5.217,6.862] | – |
| $\gamma^{Q}$ | 0.683 [0.567,0.800] | 0.722 [0.598,0.842] | 0.677 [0.540,0.808] | 0.681 [0.544,0.822] | 0.741 [0.619,0.862] |

continued on the next page

Table G2: Full Table 4 Switches in Policy Only – continued

| Parameters | Discretion | Quasi-Commitment | Rule - Parameters | Rule - Target | Commitment |
|---|---|---|---|---|---|
| | | | Policy Parameters | | |
| $\rho^R_{(s=1)}$ | – | – | 0.746 [0.708,0.786] | 0.797 [0.762,0.831] | – |
| $\rho^R_{(s=2)}$ | – | – | 0.845 [0.794,0.900] | – | – |
| $\psi_{1(s=1)}$ | – | – | 2.075 [1.824,2.315] | 1.805 [1.507,2.097] | – |
| $\psi_{1(s=2)}$ | – | – | 0.909 [0.621,1.189] | – | – |
| $\psi_{2(s=1)}$ | – | – | 0.483 [0.309,0.645] | 0.498 [0.285,0.714] | – |
| $\psi_{2(s=2)}$ | – | – | 0.245 [0.098,0.393] | – | – |
| $\omega_1$ | 0.259 [0.035,0.414] | 0.633 [0.480,0.785] | – | – | 0.502 [0.331,0.666] |
| $\omega_2$ | 0.650 [0.460,0.847] | 0.759 [0.631,0.893] | – | – | 0.523 [0.295,0.732] |
| $\omega_3$ | 0.442 [0.164,0.698] | 0.349 [0.126,0.559] | – | – | 0.460 [0.205,0.710] |
| $\omega_{\pi(s=1)}$ | 1 | 1 | – | – | 1 |
| $\omega_{\pi(s=2)}$ | 0.347 [0.219,0.477] | 0.348 [0.254,0.440] | – | – | 0.302 [0.194,0.414] |
| $p_{11}$ | 0.978 [0.962,0.994] | 0.860 [0.777,0.946] | 0.962 [0.939,0.989] | 0.956 [0.930,0.984] | 0.979 [0.962,0.996] |
| $p_{22}$ | 0.940 [0.900,0.981] | 0.879 [0.819,0.941] | 0.802 [0.734,0.870] | 0.796 [0.722,0.876] | 0.816 [0.735,0.901] |
| $\upsilon$ | – | 0.325 [0.239,0.411] | – | – | – |
| | | | Log Marginal Data Densities and Bayes Factors | | |
| Geweke (1999) | $-810.98$ (1.00) | $-814.83$ (47.0) | $-825.33$ (1.72$e$+6) | $-831.74$ (1.04$e$+9) | $-832.85$ (3.14$e$+9) |
| Sims et al. (2008) | $-811.24$ (1.00) | $-814.30$ (21.21) | $-825.44$ (1.46$e$+6) | $-831.81$ (8.52$e$+8) | $-832.98$ (2.75$e$+9) |

Figure G1: Markov Switching Probabilities: Policy Switches Only

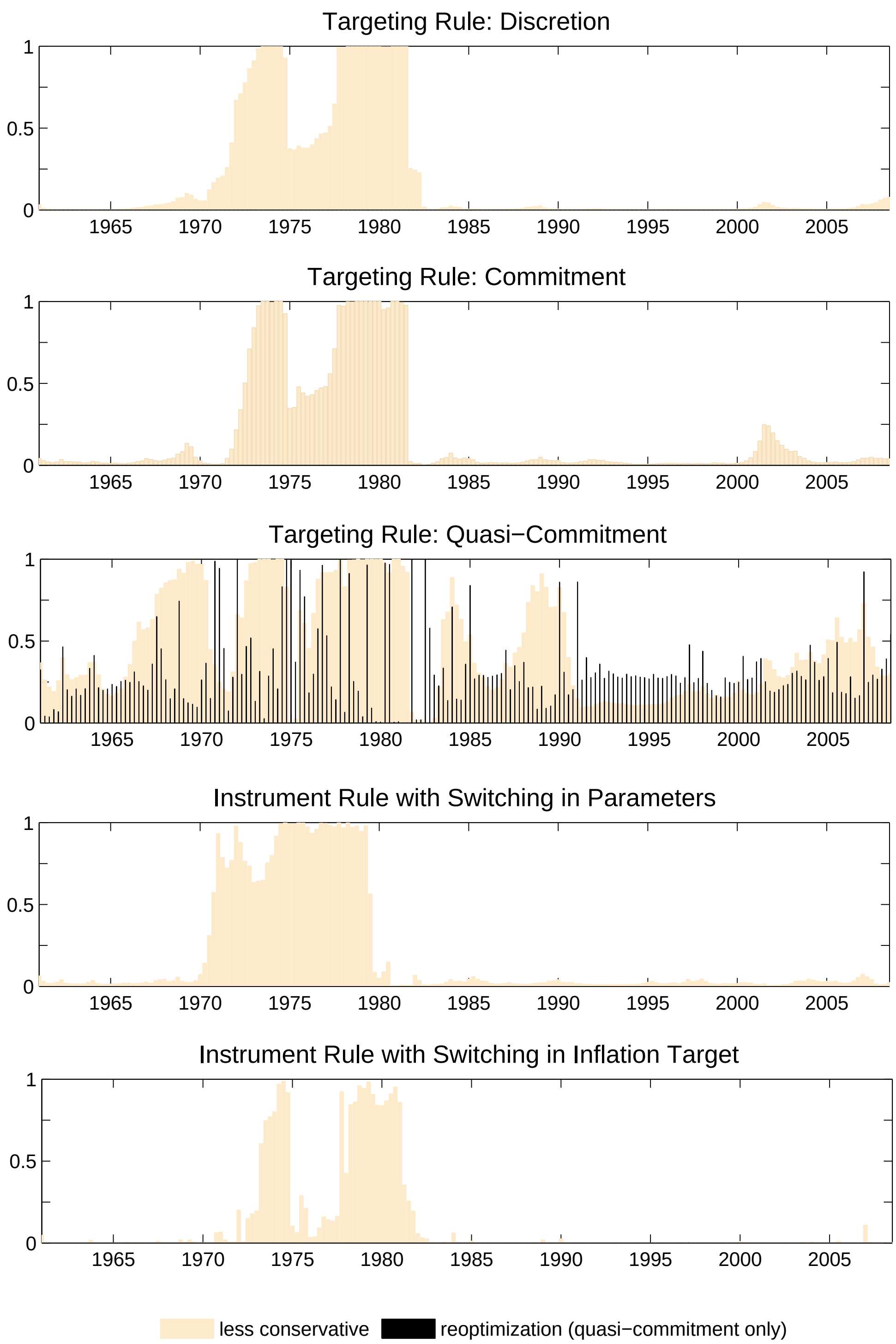
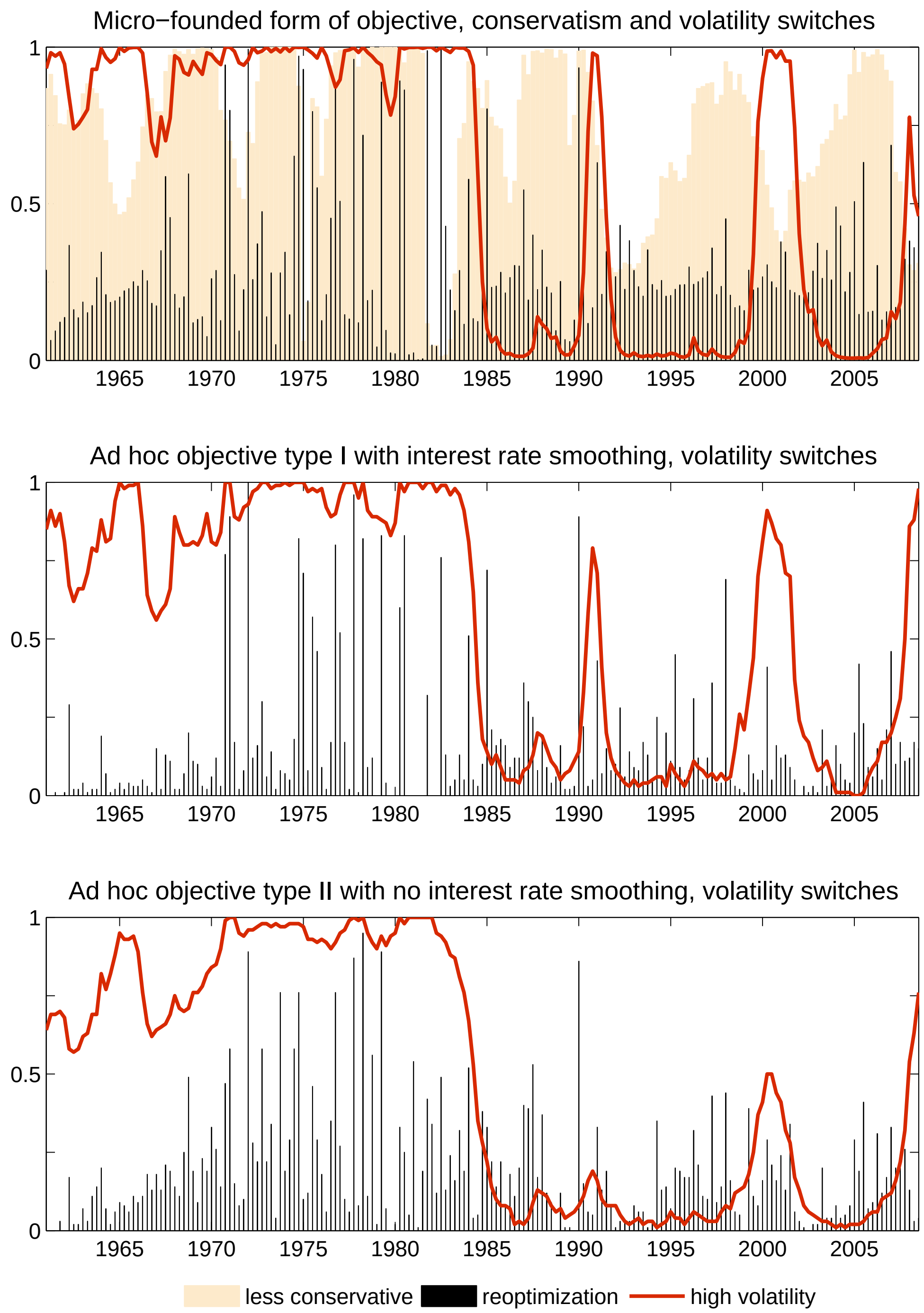


Figure G2: Quasi-commitment with different objectives



# References


Adolfson, M., S. Laseen, J. Linde, and L. E. Svensson (2011). Optimal Monetary Policy in an Operational Medium-Sized DSGE Model. *Journal of Money, Credit and Banking 43*(7), 1287–1331.

An, S. and F. Schorfheide (2007). Bayesian Analysis of DSGE Models. *Econometric Reviews 26*(2-4), 113–172.

Caglar, E., J. S. Chadha, and K. Shibayama (2011). Bayesian Estimation of DSGE models: Is the Workhorse Model Identified? Working Paper No 1125, University of Kent.

Calvo, G. (1983). Staggered Prices in a Utility-Maximising Framework. *Journal of Monetary Economics 12*, 383–398.

Debortoli, D. and R. Nunes (2010). Fiscal Policy under Loose Commitment. *Journal of Economic Theory 145*(3), 1005–1032.

Del Negro, M. (2011). Bayesian Macroeconometrics. In J. Geweke, G. Koop, and H. Van Dijk (Eds.), *The Oxford Handbook of Bayesian Econometrics*, pp. 293–389.

Del Negro, M., R. B. Hasegawa, and F. Schorfheide (2016). Dynamic prediction pools: An investigation of financial frictions and forecasting performance. *Journal of Econometrics 192*(2), 391–405.

Eser, F., C. Leith, and S. Wren-Lewis (2009). When is monetary policy all we need? University of Oxford Working Paper No.430.

Farmer, R. E. A., D. F. Waggoner, and T. Zha (2011). Minimal State Variable Solutions to Markov-switching Rational Expectations Models. *Journal of Economic Dynamics and Control 35*(12), 2150–2166.

Galí, J. and M. Gertler (1999). Inflation Dynamics: A Structural Econometric Analysis. *Journal of Monetary Economics 44*, 195–222.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In J. O. Berger, J. M. Bernardo, A. P. Dawid, and A. F. M. Smith (Eds.), *Proceedings of the Fourth Valencia International Meetings on Bayesian Statistics*, pp. 169–194. Oxford: Oxford University Press.

Geweke, J. (1999). Using simulation methods for Bayesian econometric models: inference, development, and communication. *Econometric Reviews 18*(1), 1–73.

Givens, G. E. (2012). Estimating Central Bank Preferences under Commitment and Discretion. *Journal of Money, Credit and Banking 44*(6), 1033–1061.

Ilbas, P. (2010). Estimation of monetary policy preferences in a forward-looking model: a Bayesian approach. *International Journal of Central Banking 6*(3), 169–209.

Kim, C.-J. (1994). Dynamic linear models with Markov-switching. *Journal of Econometrics 60*, 1–22.

Komunjer, I. and S. Ng (2011). Dynamic Identification of Dynamic Stochastic General Equilibrium Models. *Econometrica 79*(6), 1995–2032.

Koop, G., M. H. Pesaran, and R. P. Smith (2013). On Identification of Bayesian DSGE Models. *Journal of Business and Economic Statistics 31*(3), 300–314.

Leith, C. and J. Malley (2005). Estimated General Equilibrium Models for the Analysis of Monetary Policy in the US and Europe. *European Economic Review 49*(8), 2137–2159.

Levine, P., P. McAdam, and J. Pearlman (2008). Quantifying and sustaining welfare gains from monetary commitment. *Journal of Monetary Economics 55*, 1253–1276.

Lubik, T. and F. Schorfheide (2005). A Bayesian Look at the New Open Economy Macroeconomics. In *NBER Macroeconomics Annual 2005*, pp. 313–379.

Marcet, A. and R. Marimon (2011). Recursive Contracts. CEP Discussion Papers DP1055, Centre for Economic Performance, LSE.

Min, C. and A. Zellner (1993). Bayesian and Non-Bayesian Methods for Combining Models and Forecasts with Applications to Forecasting International Growth Rates. *Journal of Econometrics 56*(1-2), 89–118.

Schaumburg, E. and A. Tambalotti (2007). An Investigation of the Gains from Commitment in Monetary Policy. *Journal of Monetary Economics 54*(2), 302–324.

Schorfheide, F. (2005). Learning and Monetary Policy Shifts. *Review of Economic Dynamics 8*, 392–419.

Sims, C. A. (2002). Solving Linear Rational Expectations Models. *Computational Economics 20*(1-2), 1–20.

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

Svensson, L. and N. Williams (2007). Monetary Policy with Model Uncertainty: Distribution Forecast Targeting. CEPR Discussion Papers 6331.

Svensson, L. E. O. (2010). Optimization under Commitment and Discretion, the Recursive Saddlepoint Method, and Targeting Rules and Instrument Rules: Lecture Notes. available from http://larseosvensson.se/.

Wright, J. H. (2008). Bayesian Model Averaging and Exchange Rate Forecasting. *Journal of Econometrics 146*, 329–341.

Yun, T. (1996). Nominal price rigidity, money supply endogeneity, and business cycles. *Journal of Monetary Economics 37*(2-3), 345–370.