Working Paper Series

Paolo Gelain, Simone Manganelli

Monetary policy with judgment

Abstract

Two approaches are considered to incorporate judgment in DSGE models. First, Bayesian estimation indirectly imposes judgment via priors on model parameters, which are then mapped into a judgmental interest rate decision. Standard priors are shown to be associated with highly unrealistic judgmental decisions. Second, judgmental interest rate decisions are directly provided by the decision maker, and incorporated into a formal statistical decision rule using frequentist procedures. When the observed interest rates are interpreted as judgmental decisions, they are found to be consistent with DSGE models for long stretches of time, but excessively tight in the 1980s and late 1990s and excessively loose in the late 1970s and early 2000s.

NON-TECHNICAL SUMMARY

The use of judgment is pervasive in central banking. Yet, modern dynamic stochastic general equilibrium (DSGE) models have no consideration for it. This paper is concerned with the question of how to incorporate central bankers' judgmental decisions in state of the art macroeconometric models.

We first derive the optimal monetary policy decision consistent with a three-equation New Keynesian DSGE model, which constitutes the backbone of the current DSGE literature. We derive that the central bank has a quadratic loss function, with a dual mandate of stabilizing inflation around a target level and output around a reference equilibrium level. The optimal decision rule minimizes the loss function subject to the constraints represented by stochastic properties of the DSGE model economy.

We next show that Bayesian estimation of DSGE models is an indirect way to introduce judgment in the monetary policy decision problem. Bayesian estimation relies on a set of prior distributions for the model parameters, which are then combined with the likelihood function to arrive at a posterior distribution by applying Bayes rule. The expectation of the central banker's loss function is computed using the posterior distribution. We evaluate the priors by computing the expectation of the central banker's loss function using the prior (instead of the posterior) distribution and then minimizing it with respect to the interest rate. Our interpretation is that choosing priors is equivalent to endowing the central banker with specific judgmental decisions. Our results are quite striking: Imposing standard priors from the Bayesian DSGE literature is equivalent to assuming that the central banker has extreme judgmental interest rate decisions, ranging from -10% to 45%. We hope that our findings will raise awareness among economists and central bankers that the choice of the priors for Bayesian estimation is not an innocuous exercise, but it shapes the type of judgmental information which is brought into the analysis and may have strong and unwarranted implications for the decision process.

We also propose an alternative, frequentist strategy to incorporate judgment. Starting from an exogenous judgmental decision, the hypothesis that this decision is optimal can be tested by checking whether the corresponding first order condition evaluated at the estimated parameters are different from zero. If the test fails to reject the null, the econometric evidence provided by the DSGE model is compatible with the judgmental decision. Rejection of the null hypothesis, on the other hand, suggests to marginally move away from the judgmental decision in the direction of the maximum likelihood decision, since we are testing the null hypothesis that the gradient is equal to zero. By construction of the test statistic, marginal moves will be statistically significant until the closest boundary of the confidence interval associated with the gradient is reached. The interest rate decision incorporating judgment and compatible with the DSGE model is the interest rate which sets the gradient equal to the closest boundary of the confidence interval.

We estimate the model using U.S. data, taking as judgmental decision of the central bank the observed interest rates. We find that there are long stretches of time where the observed interest rate is consistent with the model. There are periods, however, where the DSGE model prescribes considerably different decisions, even though not as extreme as in the case of Bayesian decisions. Our DSGE model would have prescribed a significantly looser monetary policy in the 1980s and the late 1990s, and significantly tighter in the late 1970s and early 2000s.

1 Introduction

Central bankers face, in principle, a conceptually straightforward decision problem: for a given stochastic process governing the economy (a model), choose the interest rate which minimizes the expected loss. In practice, central bankers' decisions are a subjective combination of qualitative and quantitative information, rather than the result of a formal optimization problem. We refer to those interest rate decisions as judgmental decisions. This paper is concerned with the question of how to formally incorporate judgmental decisions in modern dynamic stochastic general equilibrium (DSGE) models.

We make three main contributions, two methodological and one empirical. First, we show that Bayesian estimation of DSGE models imposes implicit choices of central bankers' judgmental decisions, via the prior distribution of model parameters. We find that standard priors from the DSGE literature are associated with extreme and unrealistic judgmental decisions. Second, we propose an alternative frequentist approach to incorporate judgmental decisions in a DSGE framework, taking the judgmental decisions themselves, rather than the prior distribution, as a primitive of the decision problem. Third, estimating a DSGE model with U.S. data over the sample 1965-2007, we find that the Federal Reserve judgmental decisions have been compatible with the estimated DSGE model for slightly more than half of the sample. On the contrary they have been excessively tight in the 1980s and late 1990s and excessively loose in the late 1970s and early 2000s.

More concretely, suppose that the economic environment has the following characteristics:

1. the DSGE model is a correctly specified representation of the macroeconomic stochastic process;

2. the underlying model parameters are known;

3. the central banker's loss function is known.

Under these assumptions, the DSGE model can be solved and the loss function minimized with respect to the interest rate (the decision variable), thus providing the optimal solution to the central banker's decision problem. In the real world, none of those assumptions holds, a fact which goes a long way in explaining why judgment continues to play the dominant role in arriving at interest rate decisions. In this paper, we show how to formally include judgment in the central banker's decision problem, by maintaining two of the three assumptions above (1 and 3) and relaxing the hypothesis that the underlying parameters of the model are known, but have to be estimated from data.

Since our objective is not to improve on the DSGE modeling front, we work with the simplest possible, yet meaningful, three-equation New Keynesian DSGE model (Clarida et al. 1999, Galí 2015, Woodford 2003). Utility maximizing households and profit maximizing firms give rise to equilibrium conditions that describe the stochastic evolution of output (IS equation) and inflation (New Keynesian Phillips curve). The nominal interest rate is set by a central bank following a Taylor-type rule. This simple three-equation model has also the advantage that all the objects of interest can be obtained analytically.

We next derive the optimal monetary policy decision by following the standard approach in the literature (Clarida et al., 1999). The central bank loss function is derived by taking the second order approximation of the utility function, so representing a welfare function. It turns out to be quadratic in inflation and output gap, in line with the spirit of a dual mandate of stabilizing inflation around a target level and output around a reference equilibrium level. The optimal decision rule minimizes the loss function subject to the constraints represented by the behavior of private agents in the economy. We focus on the case of discretion, because of simplicity and because it better reflects the real world central banks' behavior.[1]

---

[1]For an empirical investigation supporting our choice see Chen et al. (2017).

We compare two alternative empirical strategies to incorporate judgment in the optimal monetary policy decision process.

First, we show how, in Bayesian estimation, judgment is indirectly introduced with a prior distribution for the model parameters. The logic is the following. In standard optimal monetary policy decision problems, the central banker minimizes the expected loss function with respect to the interest rate, where the expectation is computed using the posterior distribution. We perform this analysis too, but we also propose to compute the expected loss function using only the prior distribution. We refer to the interest rate minimizing this expectation as the judgmental interest rate decision implied by the priors.

This exercise is reminiscent of the prior predictive checks advocated for instance by Lancaster (2004), Geweke (2005), Geweke (2010), and Jarociński and Marcet (2019) in time series models, and by Del Negro and Schorfheide (2008), Lombardi and Nicoletti (2012), and Faust and Gupta (2012) in DSGE models.[2] It is important to stress that our interpretation is that choosing priors is equivalent to endowing the central banker with specific judgmental decisions. On the contrary the cited DSGE literature focus on the selection of priors based on their ability to match some features of the data, like steady state values and selected moments, or impulse response functions.

The second empirical strategy to arrive at the optimal decision is frequentist in nature and relies on the following reasoning. It starts from an exogenous judgment, formed by a judgmental decision and a confidence level. Judgment is exogenous in the same sense of priors in Bayesian estimation. Given the global convexity of the loss function, a necessary and sufficient condition for the optimality of the central banker's judgmental decision is that its first derivative with respect to interest rate (also labeled as gradient henceforth) is equal to zero in population. Population parameters are

---

[2]Although not exactly classifiable into the category of prior predict checks, Reis (2009) compares the response of endogenous variables to shocks computed on the basis of prior means and posterior means within a DSGE model with sticky information.

however unknown and the gradient can only be evaluated at their estimated value. We estimate the model's parameters by maximum likelihood. The gradient evaluated at the judgmental decision and at the estimated parameters will be different from zero with probability one. But the inference apparatus associated with the maximum likelihood procedure (White, 1996) allows us to test the null hypothesis that the judgmental decision is optimal by testing whether the empirical gradient is statistically different from zero, using the given confidence level. If the test fails to reject the null, we conclude that the econometric evidence provided by the DSGE model is not sufficient to suggest significant departures from the judgmental decision. Rejection of the null hypothesis, on the other hand, suggests to marginally move away from the judgmental decision in the direction of the maximum likelihood decision, since we are testing the null hypothesis that the gradient is equal to zero. By construction of the test statistic, marginal moves will be statistically significant until the closest boundary of the confidence interval associated with the gradient is reached. The implication of this reasoning is that the optimal decision compatible with the data and the DSGE model is the interest rate which sets the gradient not equal to zero (as in standard maximum likelihood applications), but to the closest boundary of the confidence interval.

We implement the two empirical strategies using U.S. data. Not surprisingly, we find that the Bayesian decisions differ from the observed ones and from those obtained by maximum likelihood with judgment, at times quite substantially. What is striking is that imposing standard priors from the Bayesian DSGE literature is equivalent to assuming that the central banker has extreme judgmental interest rate decisions, ranging from -10% to 45% in annual terms. An important insight of our findings is that the choice of the priors for Bayesian estimation is not an innocuous exercise, but it shapes the type of judgmental information which is brought into the analysis and may have strong and unwarranted implications for the decision process. Given the widespread use of DSGE Bayesian estimation in the central banking commu-

nity, we hope that our findings raise awareness about the hidden behavioral assumptions implicit in this practice.

As regards the frequentist approach of incorporating judgment, we take as judgmental decision the observed interest rates (but it could be implemented with any other arbitrary judgment). One way to read the empirical results of this paper is to check to what extent the history of Federal Reserve's decisions have been compatible with the three-equation DSGE model we are estimating. We find that there are long stretches of time where the observed interest rate is consistent with the model. There are periods, however, where the DSGE model prescribes considerably different decisions, even though not as extreme as in the case of Bayesian decisions. Our DSGE model would have prescribed a significantly looser monetary policy in the 1980s and the late 1990s, and significantly tighter in the late 1970s and early 2000s.

One important caveat of our findings is that the analysis rests on the assumption of correct model specification. In practice, the discrepancy between the judgmental and the model based decisions may be due randomness, bad judgment or model misspecification. In this paper, we cannot distinguish among these possibilities.

In terms of the relationship of our paper with the literature, there are only few contributions that consider judgment. Svensson (2005) explicitly models the role of judgment in a monetary policy decision problem. He represents judgment as the central bank's superior ability to estimate the conditional mean of stochastic deviations in the model equations. In particular, central bankers are assumed to know the reduced form parameters governing the dynamics of the economy, so that they don't face any estimation issue. There is also a group of papers which model judgment as the ability to produce more accurate forecasts on the basis of a larger information set coming from different sources (Tallman and Zaman in press, Boneva et al. 2019, Domit et al. 2019 , Del Negro and Schorfheide 2013, and Monti 2010). All existing approaches, unlike ours, build on a specific modeling of judgment. In this

paper, instead, judgment is taken as an exogenous variable to the decision problem, in the same spirit of the prior formulation in the Bayesian analysis. In that we follow Manganelli (2009).

The paper is structured as follows. Section 2 presents the DSGE model. Section 3 solves the decision problem of the central banker and shows how to incorporate judgment. Section 4 contains the empirical evidence. Section 5 concludes.

2 Maximum likelihood estimation of a stylized DSGE model

We first present a stylized New Keynesian DSGE model in its log-linearized form. Next we show how to construct the likelihood function and derive the asymptotic approximation to the variance covariance matrix of the model parameters.

2.1 The model

The model we consider is mainly taken from Clarida et al. (1999) with some elements from Smets and Wouters (2007), with households, intermediate and final good producers and a monetary and fiscal authority. The macroeconomic variables of interest are output y_t, the inflation rate π_t, and the nominal interest rate r_t. The economy is driven by three exogenous shocks: a price mark-up shock u_t, a government spending shock g_t and a monetary policy shock ε_t^r. The optimizing behavior of the agents in the economy implies the following equilibrium conditions:

$$y_t^* = -\frac{1}{\sigma}\left[r_t - E_t\pi_{t+1}\right] + E_t y_{t+1}^* - \frac{\nu}{(\nu + \sigma)}\Delta E_t g_{t+1} \tag{1}$$

$$\pi_t = \beta E_t \pi_{t+1} + \lambda\left(\sigma + \nu\right)y_t^* + \lambda u_t \tag{2}$$

where σ is the inverse elasticity of intertemporal substitution, ν the inverse of the elasticity of work effort with respect to the real wage (or inverse Frisch elasticity), β is the households' discount factor, and $\lambda = \frac{(1-\xi_p)(1-\xi_p\beta)}{\xi_p}$, with ξ_p the Calvo parameter that regulates the degree of price rigidities.

Equation 1 is the IS curve. It is expressed in terms of output gap y_t^*, which is the difference between actual and potential output y_t^n. The latter is defined as the counterfactual level of output that emerges in the absence of price rigidities and in the absence of inefficient shocks, in our case the price mark-up shock. It is equal to $\frac{\sigma}{(\nu+\sigma)}g_t$. Equation 2 is the new Keynesian Phillips curve. These two equilibrium conditions can be used to solve the model, together with the log-linear approximation of the Taylor rule:

$$r_t = \rho r_{t-1} + (1-\rho)\left(\phi_\pi \pi_t + \phi_{y^*} y_t^*\right) + \phi_{\Delta y^*}\left(y_t^* - y_{t-1}^*\right) + \sigma_r \varepsilon_t^r \qquad (3)$$

where ρ measures the interest rate smoothing, ϕ_π the systematic response to inflation, ϕ_{y^*} to the output gap, and $\phi_{\Delta y^*}$ to the output gap growth. The last term is a monetary policy shock and it is assumed to be $i.i.d.N(0,1)$.

The other two structural shocks are assumed to follow an AR(1) process, evolving according to the following parameterization:

$$\log g_t = \rho_g \log g_{t-1} + \sigma_g \varepsilon_t^g \qquad (4)$$

$$\log u_t = \rho_u \log u_{t-1} + \sigma_u \varepsilon_t^u \qquad (5)$$

with $\varepsilon_t \equiv [\varepsilon_t^g, \varepsilon_t^u]' \sim i.i.d.N(0,I)$.

Finally, we assume to observe the real GDP quarter-on-quarter growth rate $\Delta \log y_t^{obs}$, the quarter-on-quarter GDP deflator inflation π_t^{obs}, and the effective federal funds rate r_t^{obs} for the period $1965q1 - 2007q3$.[3] The link between model and observed variables is created through the following mea-

---

[3]See Appendix A for details.

surement equations

$$\Delta \log y_t^{obs} = y_t - y_{t-1}$$
$$\pi_t^{obs} = \pi_t$$
$$r_t^{obs} = r_t$$

Define the vector of observables and the vector of econometric parameter to be estimated:[4]

$$x_t \equiv [\Delta \log y_t^{obs}, \pi_t^{obs}, r_t^{obs}]' \tag{6}$$

$$\theta \equiv [\sigma_g, \sigma_r, \sigma_u, \rho_g, \rho_u, \sigma, \nu]' \tag{7}$$

Define also the vector of state variables $s_t = [y_{t-1}, g_t, u_t, r_t, y_t^*]'$ and the vector of structural shocks innovations $\epsilon_t \equiv [\varepsilon_t^u, \varepsilon_t^g, \varepsilon_t^r]'$. Then the state space representation of this economy is:

$$x_t = \phi(\theta)s_t \tag{8}$$

$$s_t = \psi_0(\theta)s_{t-1} + \psi_1(\theta)\epsilon_t \tag{9}$$

for suitable choices of the matrices $\phi(\theta)$, $\psi_0(\theta)$ and $\psi_1(\theta)$.

2.2 Estimation and asymptotics

Assuming that the structural shocks are jointly normally distributed, it is possible to construct the log-likelihood function for this economy.

Let us first set up some notation, following the general framework of White (1996).[5] The history of observations available at time n is $x^n =$

---

[4]In principle, abstracting from identification issues one can estimate all parameters. Our choice of estimating 7 parameters only is explained in section 2.3.

[5]In the DSGE literature, the convention is to use upper case letters to denote levels of variables and lower case letters for the logs. In econometrics, upper case letters denote the random variable, while lower case letters denote the realization of the random variable.

$(x'_1, ..., x'_n)'$. The observed data x_t are assumed to be a realization of a stochastic process with cdf F_t, so that $F_t(x_t^1, ..., x_t^v) = P(X_t^1 < x_t^1, ..., X_t^v < x_t^v | x^{t-1})$, $t = 1, 2, ...$ In our DSGE model x_t is defined in (6), so that $v = 3$.

The log-likelihood function is $\ell_n(X^n, \theta) \equiv n^{-1} \sum_{t=1}^{n} \log f_t(X^t, \theta)$, where $f_t(\cdot, \theta) : \mathbb{R}^{vt} \to \mathbb{R}^+$ is the pdf of the multivariate normal distribution and $\theta \in \mathbb{R}^p$, $p \in \mathbb{N}$, is a finite dimensional parameters vector. In our case, θ is defined in (7) and therefore $p = 7$. The pdf $f_t(X^t, \theta)$ can be constructed recursively from the system (8)-(9), for a given initial value s_0 (see for instance chapter 10 of Fernandez-Villaverde et al., 2016). At each point in time t, it is the pdf of a normal distribution with mean $\phi(\theta)\psi_0(\theta)s_{t-1}$ and variance $[\phi(\theta)\psi_1(\theta)][\phi(\theta)\psi_1(\theta)]'$. The maximum likelihood estimator is $\theta(X^n) = \arg\max_\theta \ell_n(X^n, \theta)$.

Assuming also that the conditions for consistency and asymptotic normality are satisfied (White, 1996) gives:

$$B^{-1/2}A\sqrt{n}(\theta(X^n) - \theta) \overset{A}{\sim} N(0, I_p) \tag{10}$$

where $A \equiv E(\nabla^2 \ell_n(X^n, \theta))$, $B^* \equiv \text{var}(\sqrt{n}\nabla\ell_n(X^n, \theta))$ and I_p is the identity matrix of dimension p. The asymptotic covariance matrix $A^{-1}BA^{-1}$ is consistently estimated by $\hat{A}_n^{-1}\hat{B}_n\hat{A}_n^{-1}$, where:

$$\hat{A}_n \equiv n^{-1} \sum_{t=1}^{n} \nabla^2 \log f_t(X^t, \theta(X^n))$$

$$\hat{B}_n \equiv n^{-1} \sum_{t=1}^{n} \nabla \log f_t(X^t, \theta(X^n)) \nabla' \log f_t(X^t, \theta(X^n))$$

2.3 Implementation

Estimating the model via maximum likelihood can be challenging, mainly because of parameters identifiability issues, also in the context of our simple model (see for instance Canova and Sala, 2009 and Iskrev, 2010 for details). We refer to Mickelsson (2015), Andreasan (2010), and Ireland (2004) for

details on maximum likelihood estimation of DSGE models.

To ensure that the estimated parameters are invariant to the optimization algorithm used and to the parameters initial values, we restrict the number of parameters to be estimated. To choose the subset, we first estimate all the 13 parameters in the model with Bayesian methods.[6] We then use the modes of the posterior distributions to initialize the maximum likelihood estimation. We start by estimating the parameters related to the shocks and fixing the others to the posterior modes. We progressively add the other parameters until when different optimizers deliver the same estimates. At this point, we test for different initial values. Our benchmark specification is reported in Table 1.[7] All the comparisons below between frequentist and Bayesian approach are based on the same set of estimated parameters.

3 Optimal monetary policy with judgment

The problem of optimal monetary policy is to choose the interest rate r_t which minimizes the central banker's expected loss. Following Galí (2015), the loss function is derived from the second order approximation of the households utility function. It can be shown that it can be expressed in terms of inflation

---

[6]Identification issues are independent of the estimation methodology. Bayesian estimation produces estimates for all parameters, even those not identified. In maximum likelihood estimation, unidentified parameters cannot be estimated and must be given specific values. In our simple model, the parameter β is clearly not identified given that the posterior distribution is identical to the prior.

[7]There is another subset of parameters that fulfills our criteria. Our results hold also with that alternative.

and output gap.[8] Hence, the central banker's problem at time t is:

$$\min_{r_t} \mathcal{L}(\theta, r_t) \equiv \frac{1}{2} E_t \left\{ \sum_{i=0}^{\infty} \beta^i \left[\omega_y y_{t+i}^{*2} + \omega_\pi \pi_{t+i}^2 \right] \right\} \tag{11}$$

$$s.t. \quad \pi_{t+i} = \beta E_t \pi_{t+1+i} + \lambda \left(\sigma + \nu \right) y_{t+i}^* + \lambda u_{t+i}$$

where $\omega_y = \sigma + \nu$ and $\omega_\pi = \frac{\varepsilon}{\lambda}$.[9] This problem can be solved analytically.[10] Under discretion it gives the following first order condition for the interest rate:

$$\nabla_r \mathcal{L}(\theta, r_t) \equiv r_t - h_t(\theta) \tag{12}$$
$$= 0$$

where

$$h_t(\theta) \equiv \left[1 + \frac{(1 - \rho_u) \sigma \varepsilon}{\rho_u} \right] \rho_u \omega_y q \lambda u_t + \frac{\sigma \nu}{(\nu + \sigma)} \left(\rho_g - 1 \right) g_t \tag{13}$$

is a function of the model parameters.

If θ were known, the solution to the central banker's problem is $\hat{r}_t = h_t(\theta)$. In other words, equation (12) establishes for every quarter a unique optimal value for the interest rate. Any deviation from that would be sub-optimal, so the central banker should always decide to set $\hat{r}_t$ if she wants to minimize her loss.

---

[8]See Galí (2015), chapter 4, for details. This assumption slightly deviates from Clarida et al. (1999). They assume an ad-hoc loss function with an identical functional form, but where the weights on inflation and output gap are not function of the model's deep parameters. We have performed our analysis under that alternative environment and results are qualitatively similar. We believe that the current specification is more consistent with the rest of the analysis.

[9]ε is the elasticity of substitution among differentiated intermediate goods. It is related to the steady state value of the mark-up shock U as follows: $U = \frac{1}{\varepsilon - 1}$. We set it at 6, a common value in the literature, which implies a steady state value of the net mark-up of 20%.

[10]See Appendix D.

In this paper, though, we are interested in the situation when θ is unknown and needs to be estimated from the data. We introduce additional structure to the problem in the form of judgment. We assume that the central banker has judgment summarized by the pair $\{\tilde{r}_t, \alpha\}$. The judgmental decision $\tilde{r}_t \in \mathbb{R}$ can be thought of as the decision the central banker would take at time t on the basis of unmodeled private information. The other component of judgment, $\alpha \in [0,1]$, is the confidence level that the central bank has on her judgmental decision. It also represents the probability used to conduct the hypothesis testing at the base of our analysis that we discuss in the following paragraphs.

Since the decision maker wishes to minimize the loss function in population, she can test that the population gradient evaluated at the judgmental decision is zero:

$$H_0 : \nabla_r \mathcal{L}(\theta, \tilde{r}_t) = 0 \tag{14}$$

Given that the estimated parameters are now random variables, the empirical gradient is too. Hence, hypothesis (14) can be tested by using the test statistic $\nabla_r \mathcal{L}_t(\theta(X^n), \tilde{r}_t)$ and noting that by the asymptotic properties of $\theta(X^n)$ described in section 2.2:

$$\sqrt{n} \Sigma_t^{-1} \nabla_r \mathcal{L}(\theta(X^n), \tilde{r}_t) \sim N(0,1) \tag{15}$$

where $\Sigma_t^2 \equiv \nabla_\theta h_t(\theta) A_n^{-1} B_n A_n^{-1} \nabla'_\theta h_t(\theta)$.[11] We denote with $\nabla_\theta h_t(\theta)$ the vector of first derivatives of the first order condition (12) with respect to the estimated parameters. Individual entries are reported in appendix F.

Testing the null (14) is equivalent to testing whether the judgmental decision $\tilde{r}_t$ is optimal. If the null is not rejected, statistical evidence is not strong enough to suggest any deviation from $\tilde{r}_t$. Rejection at the confidence level α, however, implies that *marginal* moves away from $\tilde{r}_t$ in the direction of the maximum likelihood decision do not increase the loss function with

---

[11]See appendix E for a formal derivation of this result.

probability $1 - \alpha$. Since any statistical procedure involves uncertainty, there is still the possibility that moves toward the maximum likelihood increase the loss function. However, this statistical risk is bounded above by the chosen confidence level and happens only with probability less than or equal to α, the probability of a Type I error. Iterating forward this reasoning leads to the conclusion that the decision with judgment lies at the boundary of the $(1 - \alpha)$ confidence interval associated with (15). We refer to Manganelli (2009) for details.

Given the observed sample realization x^n and using (12), the decision rule incorporating the central banker's judgment is the one which, in case of rejection, sets the gradient equal to the boundary of the confidence interval:

$$\hat{r}_t(x^n|\tilde{r}_t, \alpha) = \begin{cases} h_t(\theta(x^n)) + n^{-1/2}\hat{\Sigma}_t c_{\alpha/2} & \text{if } z_t < c_{\alpha/2} \\ \tilde{r}_t & \text{if } c_{\alpha/2} \leq z_t \leq c_{1-\alpha/2} \\ h_t(\theta(x^n)) + n^{-1/2}\hat{\Sigma}_t c_{1-\alpha/2} & \text{if } z_t > c_{1-\alpha/2} \end{cases} \qquad (16)$$

where $z_t \equiv \sqrt{n}\hat{\Sigma}_t^{-1}(\tilde{r}_t - h_t(\theta(x^n)))$ is the sample realization of the asymptotically normally distributed gradient in (15), $\hat{\Sigma}_t$ is a consistent estimator of the asymptotic variance and $c_\alpha = \Phi^{-1}(\alpha)$, $\Phi(\cdot)$ being the cdf of the standard normal distribution.

A plain English interpretation of this rule is the following. Check whether the gradient (12) evaluated at the estimated parameters $\theta(x^n)$ and at the judgmental interest rate $\tilde{r}_t$ falls within the confidence interval. If it does, retain the judgmental decision $\tilde{r}_t$. If it doesn't, choose the interest rate that moves the empirical gradient to the closest boundary of the confidence interval. A graphical representation of this reasoning is reported in Figure 1.

This decision coincides with the judgmental decision if there is not enough statistical evidence against it, and shrinks towards the maximum likelihood decision otherwise.

4 Have Fed interest rate decisions been compatible with the DSGE model?

In this section we use the estimated models to compare the decisions that a frequentist and a Bayesian decision maker would reach by following their respective statistical approach. We take an historical perspective. We run a counterfactual experiment aimed at evaluating how different central banker types would have set the policy rate compared to the observed rate over our sample.

We first need to describe how the Bayesian decision is constructed. The econometrician has first to choose a prior distribution for the parameters θ of the DSGE model. Let us denote it with $p(\theta)$. This prior distribution is then combined with the likelihood function to arrive at a posterior distribution $p(\theta|x^n)$. The posterior distribution is finally used to compute the expectation of (11), which is then minimized with respect to the interest rate. The Bayesian decision is therefore:

$$\hat{r}_t(x^n|p(\theta)) = \arg\min_{r_t} \int \mathcal{L}(\theta, r_t)dp(\theta|x^n) \tag{17}$$

This way of considering parameter uncertainty in the Bayesian framework is common in the optimal monetary policy DSGE literature, like for instance in Levin et al. (2006), Reis (2009), and Edge et al. (2010).

Turning to the frequentist decision with judgment discussed in section 3, the key input is the judgment, which is defined as a judgmental decision and a confidence level associated with it. For the confidence level, we take a standard 5% probability, which results in 95% confidence intervals. For the judgmental decision, we take the observed interest rate.

It is important to stress that our framework could be implemented with any arbitrary judgmental decision. Nevertheless, using the observed rate gives us the advantage to make our counterfactual more comparable with

the Bayesian case. But, most importantly, it clarifies the relationship of our framework with the literature on optimal monetary policy. In fact, even if parameter uncertainty has been considered in some Bayesian studies, it is more common to analyze the counterfactual under optimal policy on the basis of point estimates, either maximum likelihood values or a selected moment of the posterior distributions. This is true in the three papers mentioned above, where the case of no-parameter uncertainty is the baseline and the alternative is a robustness check. But also in those papers where parameter uncertainty is ignored tout court, like for instance Smets and Wouters (2002), Adjemian et al. (2008), Justiniano et al. (2013), Chen et al. (2017), Bodenstein and Zhao (2019), and Furlanetto et al. (2020).

Our counterfactual is similar in spirit, but it takes parameter uncertainty explicitly into account in the frequentist case. It is a shrinkage towards the maximum likelihood decision, i.e towards the decision that the literature would emphasize first. It is worth noting that we can recover that decision also within our framework. Obviously in the trivial case of completely abstracting from considering parameter uncertainty. But also in the case in which the policy maker sets α equal to 1. The confidence interval then collapses to a single value, i.e. exactly the maximum likelihood decision. The latter is reported as a blue-dotted line in the bottom panel of Figure 2.

Hence, given the observed FOMC interest rate decision, we ask the following questions: Is this decision compatible with the DSGE model discussed in this paper? If not, how can it be improved? This can be checked by testing the null hypothesis (14) at each point in time, which in turn can be implemented by setting up the confidence intervals for the empirical gradient $\nabla_r \mathcal{L}(\theta(X^n), r_t)$, where we have imposed $r_t = \tilde{r}_t$. The decision rule (16) prescribes that if the empirical gradient falls within the confidence interval, the judgmental decision is retained. If, instead, the empirical gradient falls outside the confidence interval (that is, if the null hypothesis is rejected), the optimal decision is the interest rate which moves the empirical gradient to

the closest boundary of the confidence interval. This procedure is depicted in the upper panel of Figure 2. A clear interpretation of that figure follows by bearing in mind the procedure as described in Figure 1. The latter is repeated for every quarter. One has to imagine a normal distribution with zero mean and a different variance for every quarter t.

In order to fully appreciate the difference between the judgment and our suggested optimal decision, we report in the bottom panel of the same figure both those series in percentage annual terms. Our optimal decision deviates from judgment 45% of times. Discrepancies are quite large at a time, reaching values of 3.34% (0.83% in quarterly terms) in a single quarter (1972q1). The mean deviation is 0.95%. Looking at the same panel in Figure 2, it is possible to effectively visualize the shrinkage of our decision towards the maximum likelihood.

The comparison between observed, frequentist and the Bayesian decisions is reported in Figure 3, where the observed interest rates are plotted together with the optimal decisions according to the DSGE model for the two decisions.[12] As already emphasized, when the maximum likelihood decision incorporating judgment advocated in this paper coincides with the judgmental decision itself, the optimal decision is identical to the observed one. The Bayesian decision, on the other hand, while correlating substantially with the observed one (the correlation is 0.87), prescribes at times and on average significant departures. For instance, in the mid-1970s the Bayesian interest rate is up to 5 percentage points higher than the observed one, while in the late 1990s it is 2 percentage point lower. The literature also shows sizable effects of parameter uncertainty on the optimal policy outcomes compared to the no-uncertainty case.

As for the comparison with the frequentist decision, differences are also very large in those periods. But they are quite large also on average. The

---

[12]According to our estimated parameters, the implied weights in the loss function are ω_π=407.76 and ω_y=8.04. Hence, we are *de facto* considering the case of a strict inflation targeting central bank, although this is not exactly equivalent to the case ω_π=1 and ω_y=0.

mean difference along the sample is 1.05%. We can overall conclude that the discrepancies between the two estimates are qualitatively small (the correlation is 0.94), but quantitatively sizable.

Finally, it is possible to establish a link between the judgmental decision $\tilde{r}_t$ and the prior $p(\theta)$. The judgmental decision implied by the prior is the optimal interest rate when the optimization problem in (17) is computed with respect to the prior:

$$\tilde{r}_t = \arg\min_{r_t} \int \mathcal{L}(\theta, r_t) dp(\theta) \tag{18}$$

This allows us to highlight some other important aspect of our analysis. Our choice of priors follows the standards in the literature and is reported in Table 1. We find it difficult to assess whether these are sensible priors on the basis of our knowledge of the parameters. And we find it even more difficult to convey to the decision maker (the central banker) the type of non sample information that is incorporated in the Bayesian decision. Equation (18), however, provides a simple way to evaluate the prior, by asking the following question: What interest rate decision is associated with the prior distribution over the DSGE parameters described in Table 1? This exercise is similar in spirit to the *prior predictive checks*, as discussed in Lancaster (2004), Geweke (2005), Geweke (2010), and Jarociński and Marcet (2019) in time series models, and in Del Negro and Schorfheide (2008), Lombardi and Nicoletti (2012), and Faust and Gupta (2012) in DSGE models. The difference is that instead of checking whether the model and the priors are compatible with specific moments of the observables, we compute the optimal interest rate decision which minimizes the central banker's expected loss using the prior distribution.

The results of this exercise are reported in Figure 4 and are quite striking. Imposing the standard priors of the literature is equivalent to endowing the central banker (our decision maker) with nonsensical judgmental decisions. The judgmental interest rate implied by our priors ranges from -10% in the

beginning of 2000 to more than 45% in the beginning of the 1980.

These results should warn the reader about the hidden implications of blindly applying Bayesian techniques. They impose implicit restrictions on the behavior of the decision maker.

5 Robustness

The results obtained so far depend, among other things, on the empirical Taylor rule, as specified in equation (3). Even though none of the estimated parameters in that equation directly affects the solution of the model under optimal policy, the specification has an impact on the other estimated parameters.

Typically the literature does not address that issue. All papers cited in the previous section are no exception. Nevertheless, we address it by assessing the robustness of our results to different specifications of the Taylor rule. This is to be sure that our counterfactual experiments satisfy as much as possible the principle of *Counterfactual Equivalence* as in Beraja (2019). We select two alternatives:

$$r_t = \rho r_{t-1} + (1 - \rho) \phi_\pi \pi_t + \sigma_r \varepsilon_t^r \tag{19}$$

$$r_t = \rho r_{t-1} + \frac{(1 - \rho)}{4} \left[\phi_\pi \left(\pi_t + \pi_{t-1} + \pi_{t-2} + \pi_{t-3} \right) + \phi_{\Delta y} \left(y_t - y_{t-4} \right) \right] + \sigma_r \varepsilon_t^r \tag{20}$$

Specification (20) is often used in empirical papers, as for instance in Justiniano et al. (2013) and Christiano et al. (2014).

We estimate the models with those rules following the same procedure used to estimate our baseline model. We report our results in Figure 5. The upper panels show the gradients, the 95% confidence intervals, and the resulting optimal decision for the model with equation (19) (left panel) and (20) (right panel). The maximum likelihood estimate of the parameter σ

continues to be very high for both Taylor rule specifications. This implies that the solution of the model under optimal monetary policy continues to be similar to the case of a strict inflation targeting central bank. That explains why the gradients are similar to the one in our baseline case in Figure 2. As for the confidence intervals, they depend on other factors too, so they do not necessarily have to be similar.

Given the small differences in the gradients and despite the somewhat large differences in the confidence intervals across different specifications, the implied optimal decisions reached by the central bank under the three Taylor rules do not show substantial discrepancies (Figure 5, bottom panel). We conclude that our results do not depend on the Taylor rule selected to estimate the model.

According to the log-marginal data density of the three models, our baseline specification is by far the one that fits the data better. Those values are -318.91 for our baseline model, -331.70 for the model with equation (19), and -344.16 for the model with equation (20).

6 Conclusion

Judgment plays an important role in the decision making process of a central bank. There are many quantitative and qualitative considerations which eventually lead to a decision on interest rates. For the purpose of this paper, we refer to the observed decisions as to the judgmental decisions. We then ask whether these judgmental decisions are compatible with an off-the-shelf DSGE model estimated with maximum likelihood. For a given loss function, which is derived to be quadratic in inflation and output gap, the optimal interest rate sets the gradient of the loss function equal to zero. Since the gradient depends on unknown estimated parameters, it is possible to test whether the gradient evaluated at the judgmental decision is equal to zero. If it is, we conclude that the statistical evidence provided by the estimated

DSGE model is not sufficiently strong to reject the judgmental decision. If on the other hand the gradient is statistically different from zero, our statistical decision rule prescribes to change the interest rate up to the point where the gradient is no longer statistically different from zero, a point which coincides with the closest boundary of the confidence interval.

We estimate this decision rule on US data from 1965 to 2007. We find that the Federal Reserve judgmental decisions have been compatible with the estimated DSGE model for slightly more than half of the sample. On the contrary, the judgmental interest rate decisions were tighter than what the model would have prescribed in the 1980s and late 1990s, and looser than what the model would have prescribed in the late 1970s and early 2000s.

We compare our frequentist decision rule with decisions obtained with standard Bayesian estimation methods. Bayesian optimal monetary policy decisions are derived by minimizing the expected central banker's loss function with respect to the interest rate, using the posterior distribution to compute the expectation. We highlight how Bayesian estimation of DSGE models incorporates judgment in an indirect way, via priors on statistical parameters. Priors implicitly assume a judgmental decision, which is obtained by minimizing the expected loss function with respect to the interest rate, using the prior distribution to compute the expectation. We show how standard priors from the mainstream DSGE literature are associated with unrealistic judgmental decisions on interest rates.

References

Adjemian, S., M. Darracq Pariés, and F. Smets (2008), *A Quantitative Perspective on Optimal Monetary Policy Cooperation Between the US and the Euro Area*, Working Paper Series 884, European Central Bank.

Andreasan, A. (2010), How to Maximize the Likelihood Function for a DSGE Model, *Computational Economics*, 35: 127-154.

Beraja, M. (2019), *Counterfactual Equivalence in Macroeconomics*, unpublished manuscript.

Bodenstein, M. and J. Zhao (2019), Employment, Wages and Optimal Monetary Policy, *Journal of Monetary Economics*, in press.

Boneva, L., N. Fawcett, R. M. Masolo, and M. Waldron (2019), Forecasting the UK Economy: Alternative Forecasting Methodologies and the Role of Off-model Information, *International Journal of Forecasting*, 35, 1, pages 100-120.

Canova, F. and L. Sala (2009), Back to Square One: Identification Issues in DSGE Models, *Journal of Monetary Economics*, 56: 431-449.

Chen, F., T. Kirsanova, and C. Leith (2017), How Optimal is US Monetary Policy?, *Journal of Monetary Economics*, 92: 96-111.

Christiano, L., R. Motto, and M. Rostagno (2014), Risk Shocks, *American Economic Review*, 104(1): 27-65.

Clarida, R., J. Gali, and M. Gertler (1999), The Science of Monetary Policy: A New Keynesian Perspective, *Journal of Economic Literature*, 37(4): 1661-1707.

Del Negro, M., and F. Schorfheide (2008), Forming Priors for DSGE Models (and How it Affects the Assessment of Nominal Rigidities), *Journal of Monetary Economics*, 55(7): 1191-1208.

Del Negro, M., and F. Schorfheide (2013), DSGE Model-Based Forecasting, In: G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, Volume 2, Part A, p.57 – 140, (2013), Elsevier, New York.

Domit, S., F. Monti, and A. Sokol (2019), Forecasting the UK Economy with a Medium-Scale Bayesian VAR, *International Journal of Forecasting*, 35, 4, pages 1669-1678.

Edge, R., T. Laubach, and J. C. Williams (2010), Welfare-Maximizing Monetary Policy Under Parameter Uncertainty, *Journal of Applied Econometrics*, 25(1): 129-143.

Faust, J., and A. Gupta (2012), *Posterior Predictive Analysis for Evaluating DSGE Models*, NBER Working Papers 17906, National Bureau of Economic Research, Inc.

Fernandez-Villaverde, J., J. Rubio-Ramirez, and F. Schorfheide (2016), Solution and Estimation Methods for DSGE Models, in H. Uhlig and J. Taylor (eds.), *Handbook of Macroeconomics*, Vol. 2, 527-724, Elsevier, New York.

Furlanetto, F., P. Gelain, and M. Taheri Sanjani (2020), *Output Gap, Monetary Policy Trade-offs and Financial Frictions*, Federal Reserve Bank of Cleveland Working Paper.

Galí, J. (2015), *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework*, Princeton University Press.

Geweke, J. (2005), *Contemporary Bayesian Econometrics and Statistics*, John Wiley & Sons, Inc.

Geweke, J. (2010), *Complete and Incomplete Econometric Model*, Princeton University Press.

Ireland, P. (2004), A Method for Taking Models to the Data, *Journal of Economic Dynamics and Control*, 28: 1205-1226.

Iskrev, N. (2010), Local Identification in DSGE Models, *Journal of Monetary Economics*, 57(2): 189-202.

Jarociński, M. and A. Marcet (2019), Priors about Observables in Vector Autoregressions, *Journal of Econometrics*, Vol. 209(2): 238-255.

Justiniano, A., G. E. Primiceri, and A. Tambalotti (2013), Is There a Trade-Off Between Inflation and Output Stabilization?, *American Economic Journal: Macroeconomics*, 5(2): 1-31.

Lancaster, T. (2004), *An Introduction to Modern Bayesian Econometrics*, Blackwell Publishing.

Levin, A. T., Onatski, A., Williams, J. C., Williams, N. (2006), Monetary Policy Under Uncertainty in Micro-Founded Macroeconometric Models, in *NBER Macroeconomics Annual 2005*, ed. Mark Gertler and Kenneth Rogoff, 229-97. Cambridge: MIT Press.

Lombardi, M. J., and G. Nicoletti (2012), Bayesian Prior Elicitation in DSGE Models: Macro- vs Micropriors, *Journal of Economic Dynamics and Control*, 36(2): 294-313.

Manganelli, S. (2009), Forecasting With Judgment, *Journal of Business & Economic Statistics*, 27(4): 553-563.

Mickelsson, G. (2015), *Estimation of DSGE Models: Maximum Likelihood vs.Bayesian Methods*, ZWB WP No. 2015:16.

Monti, F. (2010), Combining Judgment and Models, *Journal of Money, Credit and Banking*, 42(8): 1642-1662.

Reis, R. (2009), Optimal Monetary Policy Rules in an Estimated Sticky-Information Model, *American Economic Journal: Macroeconomics*, 1(2): 1-28.

Sims, C. A. (2001), Pitfalls of a Minimax Approach to Model Uncertainty, *American Economic Review*, 91(2): 51–4.

Smets, F. and R. Wouters (2002),*Monetary Policy in an Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area*, Proceedings, Federal Reserve Bank of San Francisco, March.

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

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

Svensson, L.E.O. (2005), Monetary Policy with Judgment: Forecast Targeting, *International Journal of Central Banking*, 1(1): 1-54.

Tallman, E. and S. Zaman (in press), Combining Survey Long-run Forecasts and Nowcasts with BVAR Forecasts Using Relative Entropy, *International Journal of Forecasting*.

White, H. (1996), *Estimation, Inference and Specification Analysis*, Cambridge Books, Cambridge University Press.

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

Appendices

A Data

In the estimation we use quarterly data on real GDP (GDPC1), GDP deflator (GDPDEF), and the federal funds rate (FEDFUNDS). All series are available in FRED, the online database maintained by the Federal Reserve Bank of St. Louis at https://fred.stlouisfed.org/. FRED's acronyms are in parenthesis. GDP growth rates is the quarter-on-quarter log-difference of real GDP. Inflation is the quarter-on-quarter log-difference of the GDP deflator. The fed funds rate is divided by 4. All series are demeaned. In the figure below we show them as they enter the estimation.
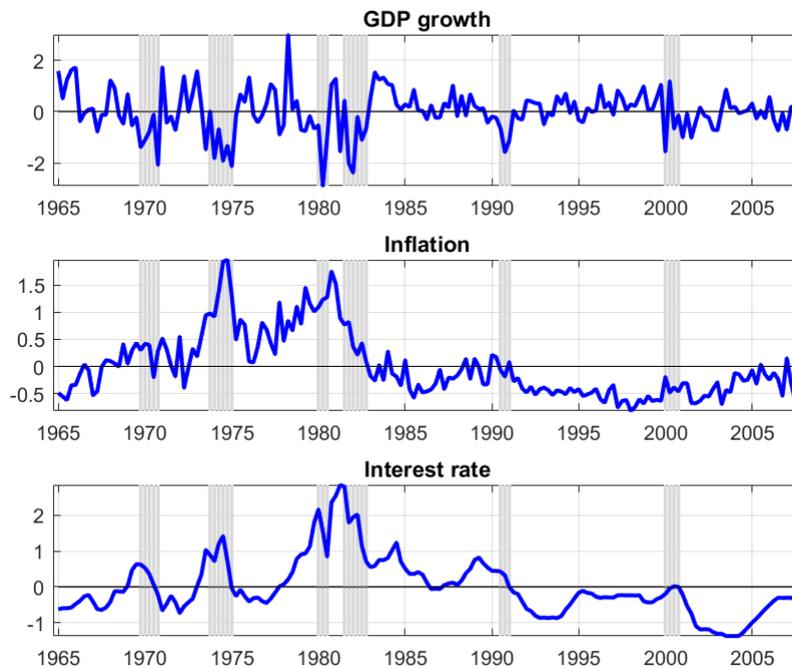
B Table and Figures

| Parameters | | Prior | | | Posterior | | | Maximum likelihood | |
	Shape	Mean	Std.	Mode	5%	95%	Estimate	Std. error
σ_g	Std. gov. spending	IG	0.10	2.00	1.1728	1.0365	1.3595	1.0626	0.1198
σ_r	Std. mon. policy	IG	0.10	2.00	0.3173	0.2947	0.3544	0.3101	0.0185
σ_u	Std. price mark-up	IG	0.10	2.00	5.8592	5.2897	6.8297	6.4164	0.7829
ρ_g	Auto. gov. spending	B	0.50	0.20	0.8891	0.8503	0.9182	0.9176	0.0236
ρ_u	Auto. price mark-up	B	0.50	0.20	0.9244	0.8885	0.9557	0.9528	0.0197
σ	Inverse elast. sub.	N	1.50	0.25	2.2436	1.7831	2.6399	3.6197	1.1641
ν	Frisch	G	2.00	0.75	3.4518	2.6668	4.5220	4.4197	1.0942

Table 1: IG = inverse gamma, B = beta, N = normal, G = gamma. The remaining parameters are set to the modes of the posterior distribution of the Bayesian estimation performed on all parameters of the model. They are: $\xi_p = 0.8868$, $\phi_\pi = 1.7469$, $\rho = 0.7496$, $\phi_y^* = 0.0958$, $\phi_{\Delta y}^* = 0.2554$, $\beta = 0.2342$.

Figure 1: Example of decision rule with judgment

$\nabla_r \mathcal{L}(\theta(x^n), \hat{r}_t)$ $\nabla_r \mathcal{L}(\theta(x^n), \tilde{r}_t)$

$1 - \alpha$ confidence interval

$\nabla_r \mathcal{L}(\theta(x^n), \tilde{r}_t)$ $\nabla_r \mathcal{L}(\theta(x^n), \hat{r}_t)$

$\nabla_r \mathcal{L}(\theta(x^n), \tilde{r}_t)$
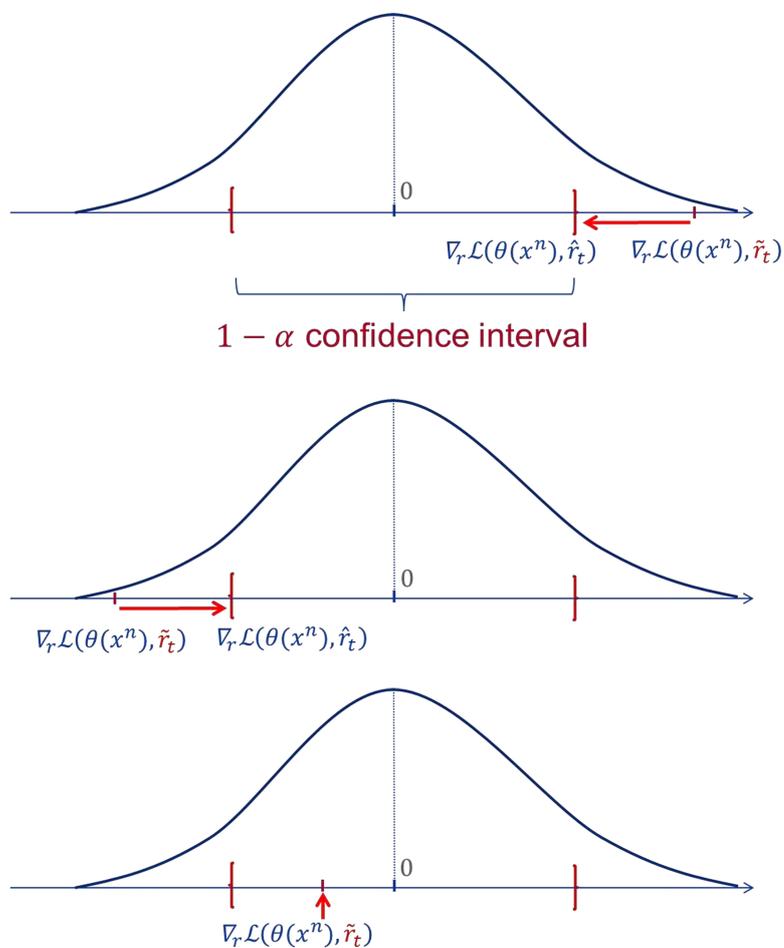
Note: The decision rule with judgment tests whether the gradient of the loss function at the judgmental decision $\tilde{r}_t$ is statistically different from zero. If it is, the optimal decision is the one associated with the closest boundary of confidence interval (depicted in the top two charts). If it is not, the judgmental decision is retained (bottom chart). The maximum likelihood decision chooses the action setting the empirical gradient to zero. It corresponds to a decision with judgment by setting $\alpha = 1$, as in this case the confidence interval degenerates to a point, the judgmental decision is always rejected and the closest boundary always coincides with the empirical gradient being set to zero.

Figure 2: Empirical gradient associated with the judgmental decision
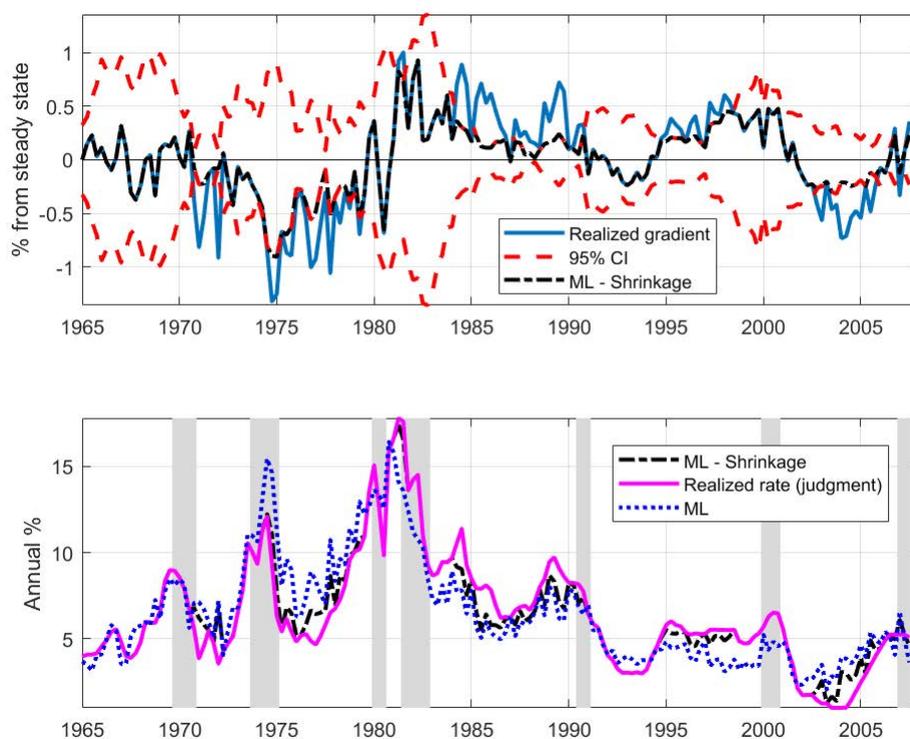
Note: The figure reports in the upper panel the 95% confidence intervals associated with the empirical gradient (dashed red line), the realized empirical gradient (blue line) and the optimal decision (dashed-dotted black line). In the lower panel in % annual terms, the realized interest rate (judgment, magenta-solid line), the optimal interest rate obtained from the shrinkage (dashed-dotted black line), and the maximum likelihood decision (dotted blue line).

Figure 3: Rate levels implied by different estimation methodologies
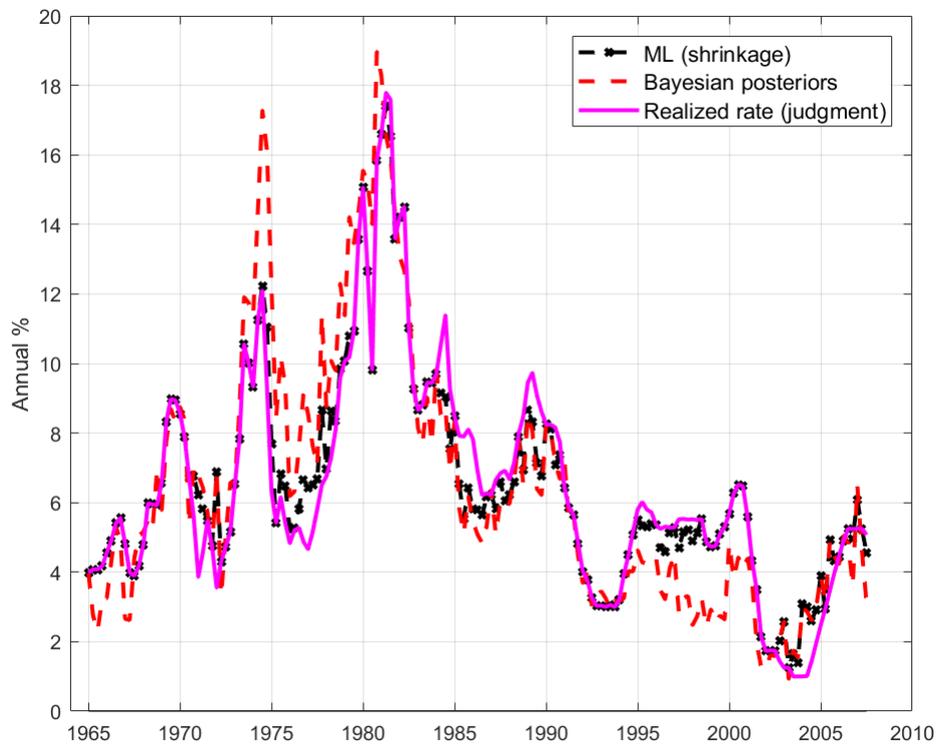
Note: The figure reports the realized interest rates, together with the optimal interest rates obtained from a maximum likelihood and a Bayesian estimation of the DSGE model.

Figure 4: Judgmental decision associated with the prior distributions
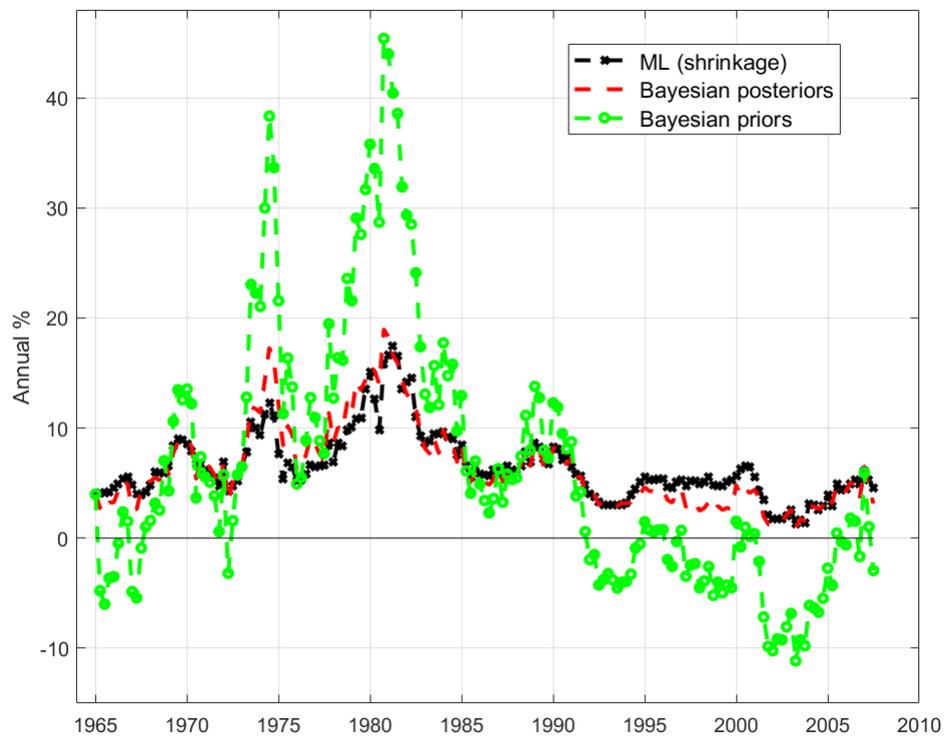
Note: The figure reports the interest rate decision solving the optimization problem (18) using only the prior distribution, together with the optimal interest rates obtained from a maximum likelihood and a Bayesian estimation of the DSGE model.

Figure 5: Robustness
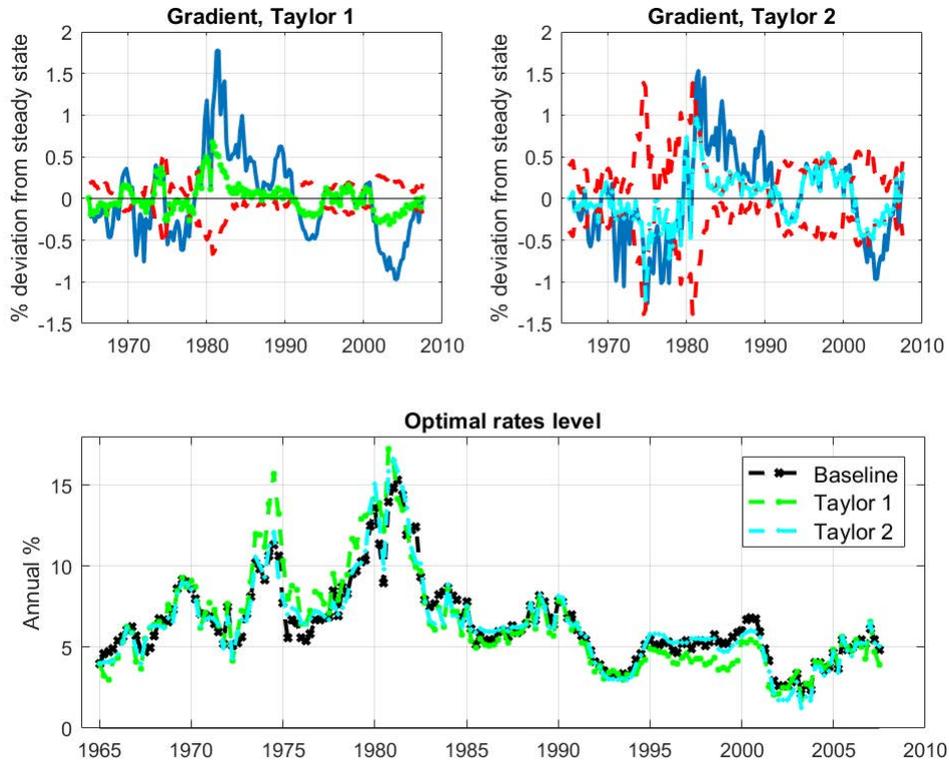
Note: The figure reports the gradients and the implied interest rate levels obtained by imposing imposing alternative interest rate rules.

C The model

C.1 Production sector

Final good producers. This type of producers operate in a perfectly competitive market. They produce the final good Y_t, to be sold to households, by combining a continuum of intermediate goods $Y_t(i)$, $i \in [0,1]$, according to the technology

$$Y_t = \left[\int_0^1 Y_t(i)^{\frac{1}{1+U_t}} \, di \right]^{1+U_t} \tag{21}$$

where U_t determines the degree of substitutability across intermediate goods in the production of the final good and hence the elasticity of demand for each of these intermediates. It is modelled as an exogenous stochastic process

$$\log \frac{1+U_t}{1+U} = \log u_t = \rho_u \log u_{t-1} + \sigma_u \varepsilon_t^u$$

where U is the steady state level of U_t and ε_t^u is an $i.i.d.N(0,1)$ innovation.[13] The latter is defined as price mark-up shock.

Profit maximization and the zero profit condition imply that the price of the final good, P_t, is a CES aggregate of the prices of the intermediate goods, $P_t(i)$

$$P_t = \left[\int_0^1 P_t(i)^{-\frac{1}{U_t}} \, di \right]^{-U_t}$$

and that the demand function for intermediate good i is

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\frac{1+U_t}{U_t}} Y_t$$

Intermediate goods producers. A monopolistic competitive firm produces the intermediate good i using labor input L_t according to the following

---

[13]Throughout this Appendix, small case letters always indicate log-deviation from steady state, while letters without subscript t indicate the steady state value.

production function

$$Y_t(i) = L_t(i) \tag{22}$$

The cost minimization problem is

$$\min_{L_t(i)} \pounds_t(i) = W_t^n L_t(i) + MC_t \left[Y_t(i) - L_t(i) \right]$$

where W_t^n represents nominal wages, and MC_t nominal marginal cost.

The first order condition is

$$\frac{\partial \pounds_t(i)}{\partial L_t(i)} : \qquad W_t^n - MC_t = 0 \tag{23}$$

which also implies that $W^n/P_t = MC_t/P_t = Mc_t$, where Mc_t is real marginal cost.

As for the pricing decision, following the formalism proposed in Calvo (1983), each firm may reset its price only with probability $1 - \xi_p$ in any given period. Thus, each period a measure $1 - \xi_p$ of producers reset their prices, while a fraction ξ_p keep their prices unchanged.

The above environment implies that the aggregate price dynamics is described by the equation

$$P_t = \left[\xi_p P_{t-1}^{-\frac{1}{U_t}} + (1 - \xi_p) \left(P_t' \right)^{-\frac{1}{U_t}} \right]^{-U_t}$$

where P_t' is the price set in period t by firms reoptimizing their price in that period.

A firm reoptimizing in period t will choose the price P_t' that maximizes the current market value of the profits generated while that price remains effective. Formally, it solves the following problem

$$\max_{P_t(i)} E_t \sum_{k=0}^{\infty} \xi_p^k \beta^k \frac{\Lambda_{t+k}}{\Lambda_t} \left[\left(P_t(i) \prod_{s=1}^{k} \Pi_{t+s-1} - MC_{t+k} \right) Y_{t+k}(i) \right]$$

subject to the sequence of demand constraints

$$Y_{t+k}(i) = \left(\frac{P_t(i) \prod_{s=1}^{k} \Pi_{t+s-1}}{P_{t+k}} \right)^{-\frac{1+U_{t+k}}{U_{t+k}}} Y_{t+k}$$

for $k = 0, 1, 2, \ldots$ where β and Λ_t are respectively the discount factor and the marginal utility of the representative household that owns the firm, and $\Pi_{t+1} \equiv \frac{P_t}{P_{t+1}}$ is the gross inflation rate. The first order condition associated with the problem above takes the form

$$E_t \sum_{k=0}^{\infty} \xi_p^k \beta^k \Lambda_{t+k} Y_{t+k}' \left[P_t' \prod_{s=1}^{k} \Pi_{t+s-1} - (1 + U_{t+k}) MC_{t+k} \right] = 0 \qquad (24)$$

where P_t' is the optimally chosen price, which is the same for all producers, and Y_{t+k}' is the demand they face in $t + k$.

C.2   Households

The economy is populated by a continuum of households. The representative household j derives utility from consuming the final good C_t and disutility from labor L_t. It maximizes its discounted expected utility

$$\max_{C(j)_t, L(j)_t} E_t \sum_{t=0}^{\infty} \beta^t \left[\frac{C_t(j)^{1-\sigma}}{1-\sigma} - \frac{L_t(j)^{1+\nu}}{1+\nu} \right]$$

where $\sigma > 0$ is the inverse of the intertemporal elasticity of substitution, $\nu > 0$ is the inverse of the labor supply elasticity (or inverse Frisch elasticity). The maximization is subject to the following budget constraint

$$P_t C_t(j) + B_t(j) + T_t = W_t^n(j) L_t(j) + R_{t-1} B_{t-1}(j) + Div_t$$

where T_t is the lump-sum tax levied on households, Div_t are dividends derived from firms. Households have access to the bonds market where they can buy

risk-free bonds B_t with a risk-free gross nominal return $R_t \equiv (1 + r_t)$, where r_t is the net nominal return.

The constrained problem can be written as follows

$$
\max_{C_t(j), L_t(j), B_t(j)} \pounds_t(j) = E_t \sum_{t=0}^{\infty} \beta^t \left\{ \frac{C_t(j)^{1-\sigma}}{1-\sigma} - \frac{L_t(j)^{1+\nu}}{1+\nu} \right\}
$$
$$
+ \beta^t \Lambda_t \left[W_t^n(j) L_t(j) + R_{t-1} B_{t-1}(j) + Div_t - P_t C_t(j) - B_t(j) - T_t \right]
$$

The first order conditions are the following

$$
\frac{\partial \pounds_t(j)}{\partial C_t(j)} : \qquad \beta^t C_t(j)^{-\sigma} - \beta^t \Lambda_t P_t = 0 \tag{25}
$$

$$
\frac{\partial \pounds_t(j)}{\partial L_t(j)} : \qquad -\beta^t L_t(j)^{\nu} + \beta^t \Lambda_t W_t^n(j) = 0 \tag{26}
$$

$$
\frac{\partial \pounds_t(j)}{\partial B_t(j)} : \qquad \beta^{t+1} \Lambda_{t+1} R_t - \beta^t \Lambda_t = 0 \tag{27}
$$

Combining equations 25 and 27 we can derive the consumption Euler equation (where we drop index j)

$$
E_t \left[\beta \left(\frac{C_{t+1}}{C_t} \right)^{-\sigma} \frac{R_t}{\Pi_{t+1}} \right] = 1 \tag{28}
$$

The labor supply condition, from 25 and 26, writes as follows

$$
\frac{L_t^{\nu}}{C_t^{-\sigma}} = \frac{W_t^n}{P_t} \tag{29}
$$

C.3 Monetary and government policy

We assume that the monetary authority controls the nominal interest rate and sets it according to a feedback rule, of the type that has been found to provide a good description of actual monetary policy in the United States at least since Taylor (1993). We chose the same specification of Smets and

Wouters (2007), hence the policy rule features interest rate smoothing, governed by parameter ρ, a systematic response to inflation, measured by parameter ϕ_π, a systematic response to the output gap, measured by parameter ϕ_{y^*}, and a response to the output gap growth with intensity $\phi_{\Delta y^*}$

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^\rho \left[\Pi_t^{\phi_\pi} (Y_t^*)^{\phi_{y^*}}\right]^{1-\rho} \left(\frac{Y_t^*}{Y_{t-1}^*}\right)^{\phi_{\Delta y^*}} e^{\sigma_r \varepsilon_t^r}$$

where R is the steady state value of the nominal interest rate, Y_t^* is the output gap (to be defined in the next section), and $\varepsilon_t^r \sim i.i.d.\ N(0,1)$ is a monetary policy shock.

Public spending E_t is determined exogenously as a time-varying fraction G_t of output

$$E_t = \left(1 - \frac{1}{G_t}\right) Y_t \tag{30}$$

where

$$\log \frac{G_t}{G} = \log g_t = \rho_g \log g_{t-1} + \sigma_g \varepsilon_t^g$$

with ε_t^g an $i.i.d.N(0,1)$ innovation. The latter is defined as government spending shock.

C.4   Equilibrium and Log-linearized model

Market clearing in the goods market requires

$$Y_t = C_t + E_t$$

Using expression 30

$$Y_t = C_t + \left(1 - \frac{1}{G_t}\right) Y_t$$

$$\frac{Y_t}{G_t} = C_t$$

and taking logs

$$c_t = y_t - g_t \tag{31}$$

Log-linearizing the Euler equation 28 yields

$$c_t = -\frac{1}{\sigma}\left[r_t - E_t\pi_{t+1}\right] + E_tc_{t+1}$$

and using 31

$$y_t - g_t = -\frac{1}{\sigma}\left[r_t - E_t\pi_{t+1}\right] + E_ty_{t+1} - E_tg_{t+1}$$

$$y_t = -\frac{1}{\sigma}\left[r_t - E_t\pi_{t+1}\right] + E_ty_{t+1} - \Delta E_tg_{t+1} \tag{32}$$

Market clearing in the labor market requires

$$L_t = \int_0^1 L(i)\,di$$

Using 22 we have

$$L_t = \int_0^1 Y(i)\,di$$

$$L_t = Y_t \int_0^1 \left(\frac{P_t(i)}{P_t}\right)^{-\frac{1+U_t}{U_t}} di$$

where the second equality follows from 21 and goods market clearing. Taking logs

$$l_t = y_t + d_t$$

where $d_t \equiv \log \int_0^1 \left(P_t(i)/P_t\right)^{-\frac{1+U_t}{U_t}}$ is a measure of price (and, hence, output) dispersion across firms. In a neighborhood of the zero inflation steady state, d_t is equal to zero up to a first order approximation. Hence one can write the following approximate relation between aggregate output and employment

$$l_t = y_t$$

Turning to the price equations, the log-linear approximation of the price setting equation for firms changing prices (equation 24) becomes

$$E_t \sum_{k=0}^{\infty} \xi_p^k \beta^k \left[p_t^{'} - \sum_{s=1}^{k} \pi_{t+s} - u_{t+k} - mc_{t+k} \right] = 0$$

Solving for the summation

$$\frac{1}{1-\xi_p\beta} p_t^{'} \qquad = E_t \sum_{k=0}^{\infty} \xi_p^k \beta^k \left[\sum_{s=1}^{k} \pi_{t+s} + u_{t+k} + mc_{t+k} \right]$$

$$= \sum_{s=1}^{0} \pi_{t+s} + u_t + mc_t + \frac{\xi_p\beta}{1-\xi_p\beta} E_t \sum_{s=1}^{1} \pi_{t+s} +$$

$$+ \xi_p^k \beta^k E_t \sum_{k=1}^{\infty} \xi_p^{k-1} \beta^{k-1} \left[\sum_{s=1}^{k} \pi_{t+s+1} + u_{t+k} + mc_{t+k} \right]$$

$$= u_t + mc_t + \frac{\xi_p\beta}{1-\xi_p\beta} E_t \left[p_{t+1}^{'} + \pi_{t+1} \right]$$

Prices evolves as

$$0 = (1-\xi_p) \, p_t^{'} - \xi_p \pi_t$$

from which we obtain the New Keynesian Phillips curve

$$\pi_t = \beta E_t \pi_{t+1} + \lambda mc_t + \lambda u_t \tag{33}$$

where $\lambda = \frac{(1-\xi_p)(1-\xi_p\beta)}{\xi_p}$.

We now want to derive an expression for the output gap Y_t^*, namely the difference between actual output Y_t and the natural level of output Y_t^n. The latter is defined as the equilibrium level of output under flexible prices and no price mark-up shocks. The model will be expressed then in terms of output gap rather than actual output. We start by taking logs of equation 23. That yields

$$\log Mc_t = \log W_t^n - \log P_t$$

From 29

$$\nu \log L_t + \sigma \log C_t = \log W_t^n - \log P_t$$

So

$$\log Mc_t = \nu \log L_t + \sigma \log C_t$$

Using 22 and 31

$$\log Mc_t = \nu \log Y_t + \sigma \left(\log Y_t - \log G_t \right)$$

$$\log Mc_t = \left(\nu + \sigma \right) \log Y_t - \sigma \log G_t \tag{34}$$

It follows that in the natural economy

$$\log Mc = \left(\nu + \sigma \right) \log Y_t^n - \sigma \log G_t \tag{35}$$

Given that in the natural economy $logMc = -U$, 35 implies the following expression

$$\log Y_t^n = \frac{\sigma}{\left(\nu + \sigma \right)} \log G_t - \frac{U}{\left(\nu + \sigma \right)} \tag{36}$$

And in log-deviation from steady state

$$y_t^n = \frac{\sigma}{\left(\nu + \sigma \right)} g_t$$

Equation 34 minus 35 provides the relationship between marginal costs and output gap log-deviation from steady state terms

$$mc_t = \left(\nu + \sigma \right) y_t^*$$

where $y_t^* = \left(y_t - y_t^n \right)$ is the output gap.

We can now re-write the IS curve (equation 32) in terms of output gap

$$y_t^* + y_t^n = -\frac{1}{\sigma} \left[r_t - E_t \pi_{t+1} \right] + E_t y_{t+1}^* + E_t y_{t+1}^n - \Delta E_t g_{t+1}$$

$$y_t^* = -\frac{1}{\sigma} \left[r_t - E_t \pi_{t+1} \right] + E_t y_{t+1}^* + E_t y_{t+1}^n - y_t^n - \Delta E_t g_{t+1}$$

$$y_t^* = -\frac{1}{\sigma}\left[r_t - E_t\pi_{t+1}\right] + E_t y_{t+1}^* + \frac{\sigma}{(\nu+\sigma)}\Delta E_t g_{t+1} - \Delta E_t g_{t+1}$$

$$y_t^* = -\frac{1}{\sigma}\left[r_t - E_t\pi_{t+1}\right] + E_t y_{t+1}^* - \frac{\nu}{(\nu+\sigma)}\Delta E_t g_{t+1} \tag{37}$$

And also the New Keynesian Phillips curve

$$\pi_t = \beta E_t\pi_{t+1} + \lambda\left(\sigma+\nu\right)y_t^* + \lambda u_t \tag{38}$$

Finally the log-linear approximation of the Taylor rule

$$r_t = \rho r_{t-1} + (1-\rho)\left(\phi_\pi \pi_t + \phi_{y^*} y_t^*\right) + \phi_{\Delta y^*}\left(y_t^* - y_{t-1}^*\right) + \sigma_r \varepsilon_t^r$$

D   Derivation of optimal simple rule

The central bank problem is

$$\min_{r_t} \mathcal{L}(\theta, r_t) \equiv \min_{y_t^*, \pi_t} \frac{1}{2} E_t \left\{ \sum_{i=0}^{\infty} \beta^i \left[\omega_y y_{t+i}^{*2} + \omega_\pi \pi_{t+i}^2 \right] \right\}$$

$$s.t. \qquad \pi_{t+i} = \beta E_t \pi_{t+1+i} + \lambda (\sigma + \nu) y_{t+i}^* + \lambda u_{t+i}$$

where $\omega_y = \sigma + \nu$, $\omega_\pi = \frac{\varepsilon}{\lambda}$, and ε is the elasticity of substitution among differentiated intermediate goods. It is related to the steady state value of the mark-up shock U as follows: $U = \frac{1}{\varepsilon - 1}$. Under discretion expectations are taken as given so the Lagrangian is

$$\min_{y_t^*, \pi_t} \mathcal{L}(\theta, r_t) = \frac{1}{2} \left[\omega_y y_t^{*2} + \omega_\pi \pi_t^2 \right] + F_t + \eta_t \left[\pi_t - \lambda (\sigma + \nu) y_t^* - f_t \right]$$

where $F_t \equiv \frac{1}{2} E_t \left\{ \sum_{i=1}^{\infty} \beta^i \left[\omega_y y_{t+i}^{*2} + \omega_\pi \pi_{t+i}^2 \right] \right\}$ and $f_t \equiv \beta E_t \pi_{t+1+i} + \lambda u_{t+i}$. The f.o.c.s are

$$\frac{\partial \mathcal{L}_t}{\partial y_t^*} : \qquad -\omega_y y_t^* - \eta_t \lambda (\sigma + \nu) = 0 \qquad (39)$$

$$\frac{\partial \mathcal{L}_t}{\partial \pi_t} : \qquad -\omega_\pi \pi_t + \eta_t = 0 \qquad (40)$$

Combining the two f.o.c.s we obtain

$$y_t^* = -\frac{\lambda (\sigma + \nu) \omega_\pi}{\omega_y} \pi_t \qquad (41)$$

Using this to substitute y_t^* in the Phillips curve 2 and solving forward we obtain[14]

$$\pi_t = \omega_y q \lambda u_t \qquad (42)$$

where $q = \frac{1}{\lambda^2 (\sigma+\nu)^2 \omega_\pi + \omega_y (1 - \beta \rho_u)} = \frac{1}{\lambda^2 (\sigma+\nu)^2 \frac{\varepsilon}{\lambda} + \omega_y (1 - \beta \rho_u)} = \frac{1}{\lambda (\sigma+\nu)^2 \varepsilon + \omega_y (1 - \beta \rho_u)}$.

---

[14]See appendix for derivations.

Substituting this last expression into 41 it follows

$$y_t^* = -\frac{\lambda(\sigma+\nu)\omega_\pi}{\omega_y}\omega_y q\lambda u_t$$

$$y_t^* = -\lambda^2(\sigma+\nu)\omega_\pi q u_t \qquad (43)$$

Finally using expression 42 and 43 in the Euler equation we derive the optimality condition for the nominal interest rate as follows

$$r_t - \left[1 + \frac{(1-\rho_u)\sigma\varepsilon}{\rho_u}\right]\rho_u\omega_y q\lambda u_t + \frac{\sigma\nu}{(\nu+\sigma)}(\rho_g-1)g_t = 0$$

To derive equation 42 we start by substituting expression 41 in the Phillips curve 2

$$\pi_t = \beta E_t\pi_{t+1} - \frac{\lambda^2(\sigma+\nu)^2\omega_\pi}{\omega_y}\pi_t + \lambda u_t$$

$$\left[1 + \frac{\lambda^2(\sigma+\nu)^2\omega_\pi}{\omega_y}\right]\pi_t = \beta E_t\pi_{t+1} + \lambda u_t$$

$$A\pi_t = \beta E_t\pi_{t+1} + \lambda u_t$$

where $A \equiv \frac{\omega_y+\lambda^2(\sigma+\nu)^2\omega_\pi}{\omega_y}$. Solving 1 step forward

$$\pi_t = \frac{\beta}{A}E_t\pi_{t+1} + \frac{\lambda}{A}u_t$$

$$\pi_t = \frac{\beta}{A}\left[\frac{\beta}{A}E_t\pi_{t+2} + \frac{\lambda}{A}E_t u_{t+1}\right] + \frac{\lambda}{A}u_t$$

from 5, $E_t u_{t+1} = \rho_u u_t$, hence

$$\pi_t = \frac{\beta^2}{A^2}E_t\pi_{t+2} + \frac{\beta}{A^2}\lambda\rho_u u_t + \frac{\lambda}{A}u_t$$

Solving 2 steps forward

$$\pi_t = \frac{\beta^2}{A^2}\left[\frac{\beta}{A}E_t\pi_{t+3} + \frac{\lambda}{A}E_t u_{t+2}\right] + \frac{\beta}{A^2}\lambda\rho_u u_t + \frac{\lambda}{A}u_t$$

$$\pi_t = \frac{\beta^3}{A^3}E_t\pi_{t+3} + \frac{\beta^2}{A^3}\lambda\rho_u^2 u_t + \frac{\beta}{A^2}\lambda\rho_u u_t + \frac{\lambda}{A}u_t$$

$$\pi_t = \frac{\beta^3}{A^3}E_t\pi_{t+3} + \left(\frac{\beta^2}{A^2}\rho_u^2 + \frac{\beta}{A}\rho_u + 1\right)\frac{\lambda}{A}u_t$$

Solving $j \to \infty$ steps forward

$$\pi_t = \frac{\beta^j}{A^j}E_t\pi_{t+j} + \frac{\lambda}{A}u_t\sum_{j=0}^{\infty}\frac{\beta^j}{A^j}\rho_u^j$$

The ratio $\frac{\beta^j}{A^j}$ goes to 0 for $j \to \infty$, while the series $\sum_{j=0}^{\infty}\frac{\beta^j}{A^j}\rho_u^j$ converges to $\frac{1}{1-\frac{\beta}{A}\rho_u}$. Recalling the value of $A \equiv \frac{\omega_y+\lambda^2(\sigma+\nu)^2\omega_\pi}{\omega_y}$

$$\pi_t = \frac{\lambda}{A}u_t\frac{1}{1-\frac{\beta}{A}\rho_u}$$

$$\pi_t = \lambda u_t\frac{1}{A-\beta\rho_u}$$

$$\pi_t = \lambda u_t\frac{1}{\frac{\omega_y+\lambda^2(\sigma+\nu)^2\omega_\pi}{\omega_y}-\beta\rho_u}$$

$$\pi_t = \omega_y q\lambda u_t \tag{44}$$

where $q = \frac{1}{\lambda^2(\sigma+\nu)^2\omega_\pi+\omega_y(1-\beta\rho_u)}$. The expression for 43 is obtained simply substituting 44 into 41.

Finally, we can get the optimality condition for the nominal interest rate by using in 42 and 43 in the IS curve 1

$$y_t^* = -\frac{1}{\sigma}\left[r_t - E_t\pi_{t+1}\right] + E_t y_{t+1}^* - \frac{\nu}{(\nu+\sigma)}\Delta E_t g_{t+1}$$

$$-\lambda^2 \left(\sigma + \nu\right) \omega_\pi q u_t = -\frac{1}{\sigma} r_t + \frac{1}{\sigma} \omega_y q \lambda E_t u_{t+1} - \lambda^2 \left(\sigma + \nu\right) \omega_\pi q E_t u_{t+1} - \frac{\nu}{\left(\nu + \sigma\right)} \Delta E_t g_{t+1}$$

$$r_t = \sigma \lambda^2 \left(\sigma + \nu\right) \omega_\pi q u_t + \omega_y q \lambda \rho_u u_t - \sigma \lambda^2 \left(\sigma + \nu\right) \omega_\pi q \rho_u u_t - \frac{\sigma \nu}{\left(\nu + \sigma\right)} \Delta E_t g_{t+1}$$

$$r_t - \left[1 + \frac{\left(1 - \rho_u\right) \sigma \lambda \left(\sigma + \nu\right) \omega_\pi}{\rho_u \omega_y}\right] \rho_u \omega_y q \lambda u_t + \frac{\sigma \nu}{\left(\nu + \sigma\right)} \left(\rho_g - 1\right) g_t = 0$$

$$r_t - \left[1 + \frac{\left(1 - \rho_u\right) \sigma \lambda \left(\sigma + \nu\right) \frac{\varepsilon}{\lambda}}{\rho_u \left(\sigma + \nu\right)}\right] \rho_u \omega_y q \lambda u_t + \frac{\sigma \nu}{\left(\nu + \sigma\right)} \left(\rho_g - 1\right) g_t = 0$$

$$r_t - \left[1 + \frac{\left(1 - \rho_u\right) \sigma \varepsilon}{\rho_u}\right] \rho_u \omega_y q \lambda u_t + \frac{\sigma \nu}{\left(\nu + \sigma\right)} \left(\rho_g - 1\right) g_t = 0$$

E   Derivation of the asymptotic distribution of the gradient

The gradient evaluated at the maximum likelihood estimator and at the judgmental decision $\tilde{r}_t$ is $\nabla_r \mathcal{L}(\theta(X^n), \tilde{r}_t) = \tilde{r}_t - h_t(\theta(X^n))$. A mean value expansion around the population value θ gives:

$$\tilde{r}_t - h_t(\theta(X^n)) = \tilde{r}_t - h_t(\theta) - \nabla_\theta h_t(\bar{\theta})(\theta(X^n) - \theta)$$

where $\bar{\theta}$ lies between θ and $\theta(X^n)$. Since under the null hypothesis $\tilde{r}_t - h_t(\theta) = 0$, multiplying by $\sqrt{n}$ and using (10) gives the result.

F   Gradient's derivatives

In this appendix we show the individual entries of the vector $\nabla_\theta h_t(\theta)$, the vector of the first derivatives with respect to the estimated parameters of the central banker's first order condition $\nabla_r \mathcal{L}(\theta, r_t)$. The size of that vector is $p = 7$. We report here only 4 derivatives, with respect to ρ_g, ρ_u, σ, and ν, because those with respect to the shock variances σ_g, σ_r, and σ_u, are all zero in every quarter t.

$$\nabla_{\rho_g} h_t(\theta) \equiv \frac{\partial h_t(\theta)}{\partial \rho_g} = -\frac{\nu \sigma}{\sigma + \nu} g_t$$

$$
\begin{aligned}
\frac{\partial h_t(\theta)}{\partial \rho_u} = {}& -\frac{\varepsilon \sigma}{\rho_u \left[\frac{\theta}{1-\theta} \frac{(1-\beta\rho_u)}{(1-\beta\theta)} + \varepsilon (\sigma + \nu) \right]} u_t \\
& + \frac{\varepsilon \sigma (1 - \rho_u) + \rho_u}{\rho_u \left[\frac{\theta}{1-\theta} \frac{(1-\beta\rho_u)}{(1-\beta\theta)} + \varepsilon (\sigma + \nu) \right]} u_t \\
& + \frac{\beta (\sigma + \nu)^2 (1-\theta)(1-\beta\theta) \left[\varepsilon \sigma (1 - \rho_u) + \rho_u \right]}{\theta \left[(\sigma + \nu)(1 - \beta\rho_u) + \frac{\varepsilon(\sigma+\nu)^2(1-\theta)(1-\beta\theta)}{\theta} \right]^2} u_t
\end{aligned}
$$

$$\frac{\partial h_t(\theta)}{\partial \sigma} = \frac{\varepsilon\sigma\left(1-\rho_u\right)+\rho_u}{\left(\sigma+\nu\right)\left[\frac{\theta}{(1-\theta)}\frac{(1-\beta\rho_u)}{(1-\beta\theta)}+\varepsilon\left(\sigma+\nu\right)\right]}u_t$$

$$+\frac{\varepsilon\left(1-\rho_u\right)}{\frac{\theta}{(1-\theta)}\frac{(1-\beta\rho_u)}{(1-\beta\theta)}+\varepsilon\left(\sigma+\nu\right)}u_t$$

$$-\frac{\frac{(1-\theta)^2(1-\beta\theta)^2(\nu+\sigma)^2[\varepsilon\sigma(1-\rho_u)+\rho_u]\left[2\varepsilon-\frac{\beta\rho_u\theta-\theta}{(1-\theta)(1-\beta\theta)(\nu+\sigma)}\right]}{\theta}}{\theta\left[\frac{\varepsilon(1-\theta)(1-\beta\theta)(\sigma+\nu)^2}{\theta}+\left(1-\beta\rho_u\right)\left(\sigma+\nu\right)\right]^2}u_t$$

$$-\frac{\left(\rho_g-1\right)\nu^2}{\left(\sigma+\nu\right)^2}g_t$$

$$\frac{\partial h_t(\theta)}{\partial \nu} = \frac{\varepsilon\sigma\left(1-\rho_u\right)+\rho_u}{\left(\nu+\sigma\right)\left[\frac{\theta}{(1-\theta)}\frac{(1-\beta\rho_u)}{(1-\beta\theta)}+\varepsilon\left(\nu+\sigma\right)\right]}u_t$$

$$-\frac{\frac{(1-\theta)^2(1-\beta\theta)^2(\nu+\sigma)^2[\varepsilon\sigma(1-\rho_u)+\rho_u]\left[2\varepsilon-\frac{\beta\rho_u\theta-\theta}{(1-\theta)(1-\beta\theta)(\nu+\sigma)}\right]}{\theta}}{\theta\left[\frac{\varepsilon(1-\theta)(1-\beta\theta)(\nu+\sigma)^2}{\theta}+\left(1-\beta\rho_u\right)\left(\nu+\sigma\right)\right]^2}u_t$$

$$-\frac{\left(\rho_g-1\right)\sigma^2}{\left(\nu+\sigma\right)^2}g_t$$

Paolo Gelain
Federal Reserve Bank of Cleveland, Cleveland, United States; email: paolo.gelain@clev.frb.org

Simone Manganelli
European Central Bank, Frankfurt am Main, Germany; email: simone.manganelli@ecb.europa.eu