Figuring Out the Fed - Online Appendix

Christian Matthes
Federal Reserve Bank of Richmond *

November 28, 2013

Abstract

This appendix contains additional results and robustness checks for "Figuring Out the Fed".

*email: christian.matthes@rich.frb.org. The views expressed in this paper are those of the author and should not be interpreted as those of the Federal Reserve Bank of Richmond or the Federal Reserve System.

1 Calculating Expectations (and Perceived Steady States)

One-period-ahead expectations of the output gap and inflation are calculated using the model probabilities derived in the main text and equilibrium laws of motion for the relevant variables that are part of the solution of the commitment and discretion optimal policy problems. In particular, I assume that

$$E_t(\pi_{t+1}) = p^c_{t-1} E^c_{t-1}(\pi_{t+1}) + p^d_{t-1} E^d_{t-1}(\pi_{t+1}) \tag{1}$$

$$E_t(y_{t+1}) = p^c_{t-1} E^c_{t-1}(y_{t+1}) + p^d_{t-1} E^d_{t-1}(y_{t+1}) \tag{2}$$

The probabilities are calculated using the Quasi-Bayesian approach described in the main text, while the expectations are calculated using the solution to the two optimal policy problems.

Expectations only depend on observables dated $t-1$ or earlier to avoid having to solve a non-linear fixed point problem when estimating the model[1]. The perceived steady state of inflation is calculated as follows:[2]

$$\overline{\pi}_t = p^c_{t-1}\pi^c + (1 - p^c_{t-1})\pi^d \tag{3}$$

[1]While this sort of assumption is common in the literature on learning in macroeconomics, it is used here for a slightly different reason: it is not the agents in this model who would have to solve a non-linear fixed point problem, as the expectations in the equilibrium conditions would only be functions of contemporaneous exogenous variables z_t and g_t and predetermined endogenous variables π_{t-1}, $\lambda_{NKPC,t-1}$ and $\lambda_{IS,t-1}$. Rather, the fixed point problem would be one of jointly solving for z_t, g_t and expectations consistent with those values when trying to calculate the likelihood of the model. If the expectations appearing in equations (1) and (2) depend on date t state variables of the two policy problems (including z_t and g_t) then by the equilibrium conditions outlined in the main text z_t and g_t would have to be jointly determined with model probabilities and conditional expectations.

[2]The equations in this section can be derived by log-linearizing the equilibrium conditions of a non-linear New Keynesian model around the perceived steady state, which is governed by the estimated model probabilities (that are taken as given in the log-linearization).

2 Difference in policy prescriptions

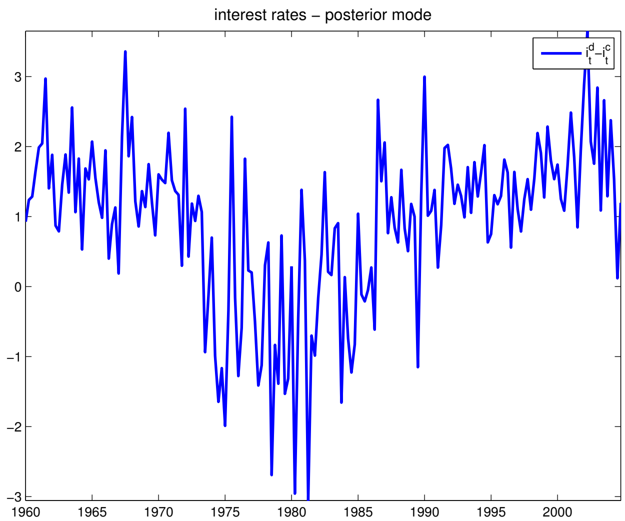

Figure 1: Difference in interest rate prescriptions

3 Posterior distribution of model probabilities

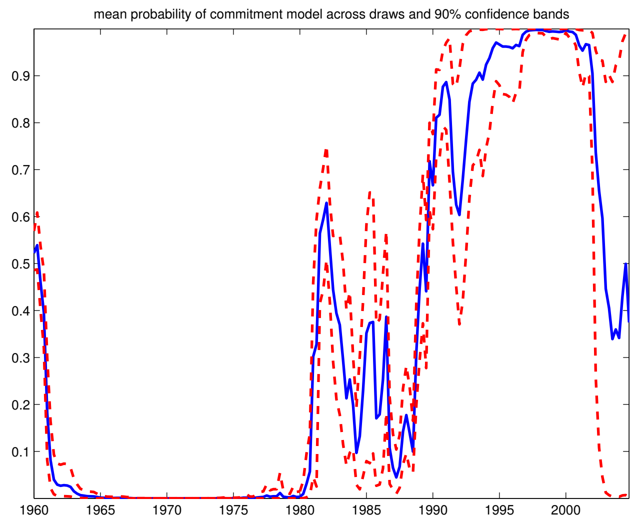

Figure 2: Posterior distribution of p_t^c

4 Contribution of different state variables to policy prescriptions

It is also useful to ask how different the policy prescriptions coming out of each model are. After all, this difference plays a major role in determining model probabilities. [3]

For the beginning and the end of the sample, the discretionary central bank would have set higher interest rates on average than its committed counterpart. At the end of the 1970s and the beginning of the 1980s, the commitment central bank would have set interest rates higher to combat inflation [4]. An interesting question is why the discretionary central bank would have set higher interest rates for most of the sample. To do so, it is instructive to look at the two policy rules $f^c(X_t^c)$ and $f^d(X_t^d)$: [5]

$$\begin{aligned} i_t^c &= 0.0103 - 0.56 z_t + 0.41 g_t + 0.38 \pi_{t-1} \\ &\quad + \text{terms depending on } \lambda_{NKPC,t-1} \text{ and } \lambda_{IS,t-1} \qquad (4) \\ i_t^d &= 0.0136 - 1.51 z_t + 0.92 g_t + 0.56 \pi_{t-1} \qquad (5) \end{aligned}$$

Part of the answer is the higher inflation target of the discretionary central bank. It is worth noting, however, that this higher inflation target does *not* translate one-to-one to a higher constant in the policy rule for that central bank. Instead, while the policy rule for that central bank does imply a steady state value for the nominal interest rate equal to $\pi^d + r^*$, part of that steady state interest rate comes from the coefficient on lagged inflation. In fact, the difference in the constants of the two policy rules in annualized percentage terms is only 1.32%. The discre-

[3] It is the relative squared distance to the actual observed interest rate, though, that ultimately determines model probabilities. This can be seen in a graph in the appendix.

[4] A graph of the difference in policy prescriptions is available in the online appendix.

[5] The negative sign on z_t in both policy rules is a result of that shock entering with a minus sign in the New Keynesian Phillips Curve, which is a standard way of writing this equation (see, for example, Lubik & Schorfheide (2004)).

tionary central bank reacts more strongly to both shocks and the lagged inflation rate than the commitment central bank, but it obviously lacks the terms depending on the Lagrange multipliers $\lambda_{NKPC,t-1}$ and $\lambda_{IS,t-1}$, coming from the nature of the commitment solution. This stronger contemporaneous response is a standard feature of discretionary policymaking, as described in, for example, Woodford (2003).

To come back to the question posed above, the combination of the stronger response to lagged inflation and the higher inflation target lead to the higher interest rate prescriptions of the opportunistic central bank for most of the sample[6]. The following graphs plot the contributions of the relevant state variables to the policy prescriptions of the committed and discretionary policymakers.

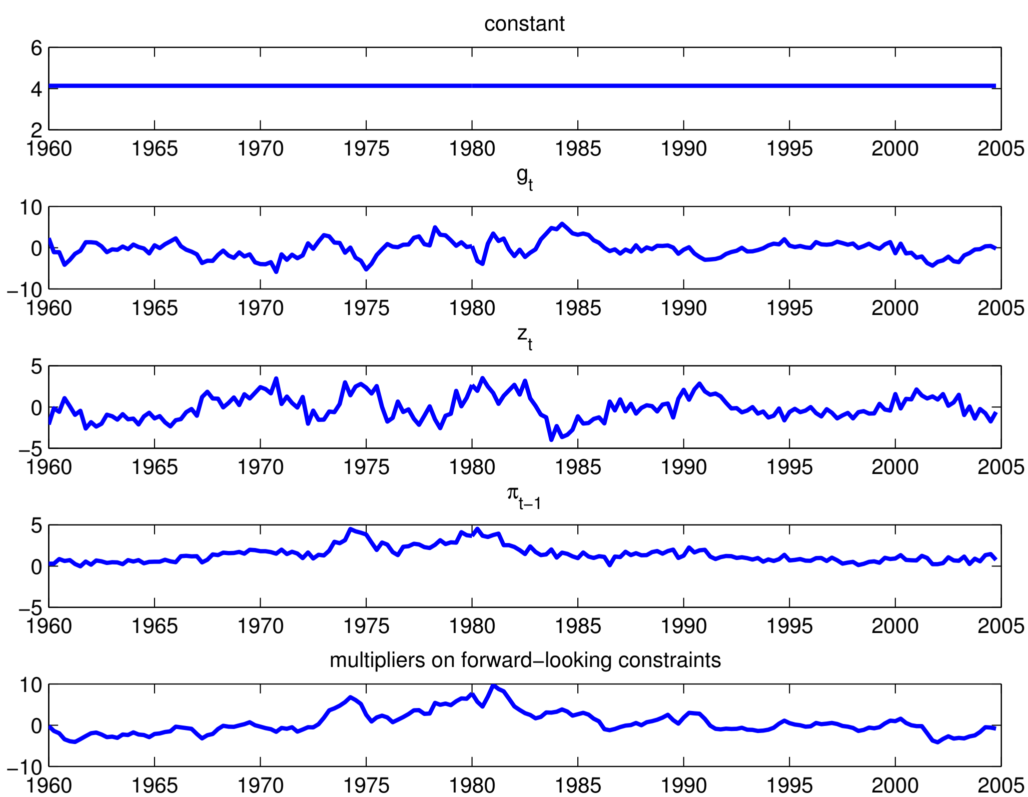

Figure 3: Contributions of different state variables to i_t^c (in annualized percent)

[6]The online appendix contains graphs that plot the contribution of different state variables to the policy prescriptions.

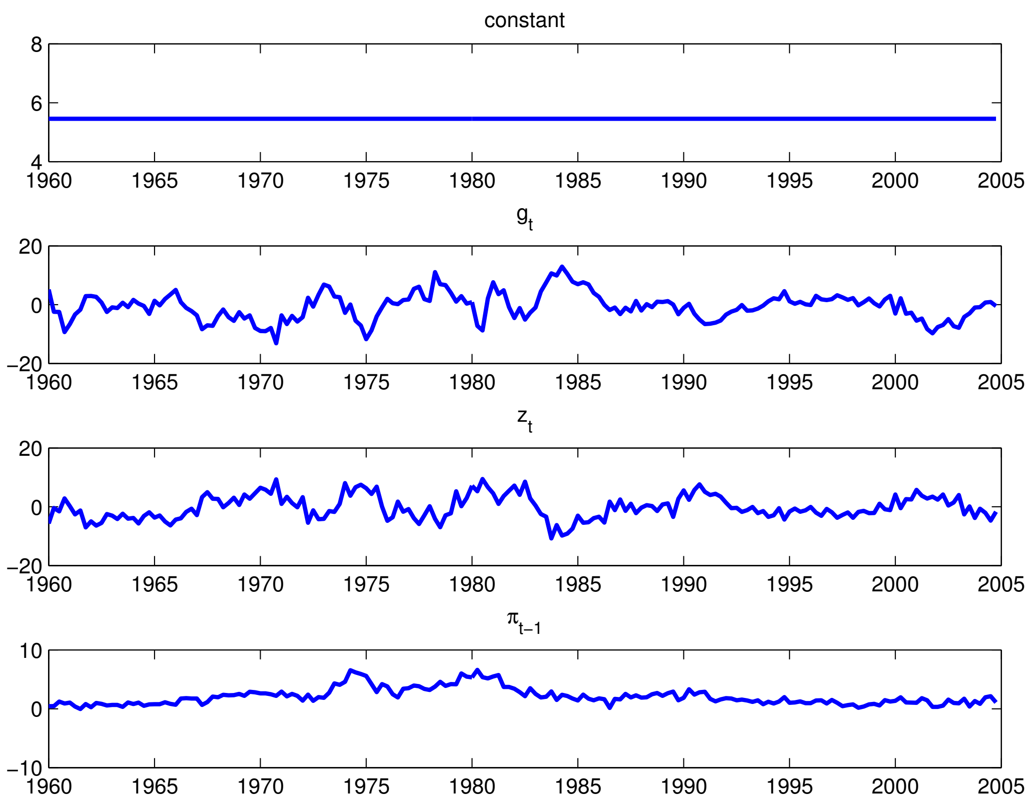

Figure 4: Contributions of different state variables to i_t^d (in annualized percent)

5 10 year average of squared prediction errors

To gain further insight into how the private sector arrives at the posterior model probabilities, it is useful to analyze squared prediction errors coming from the two submodels, as they govern the Gaussian conditional likelihoods l_t^c and l_t^d:

$$(400i_t - 400i_t^d)^2 - (400i_t - 400i_t^c)^2 \tag{6}$$

A value of this statistic less than 0 implies a better fit of the discretion submodel. This statistic is very volatile and periods where one submodel is preferred by the date are often followed by a period where the other submodel is preferred. What determines the model probabilities is a 10 year average of this difference in squared prediction errors. This moving average is plotted in the online appendix.

A positive value of the moving average implies evidence in favor of the commitment submodel. Because I set the prior model probabilities to 0.5, a positive value of this moving average implies that p_t^c is larger than 0.5. The evidence for or against one of the two submodels is strongest in the 1970s, where the magnitude of the moving average is much larger than what it is at any point after 1980.

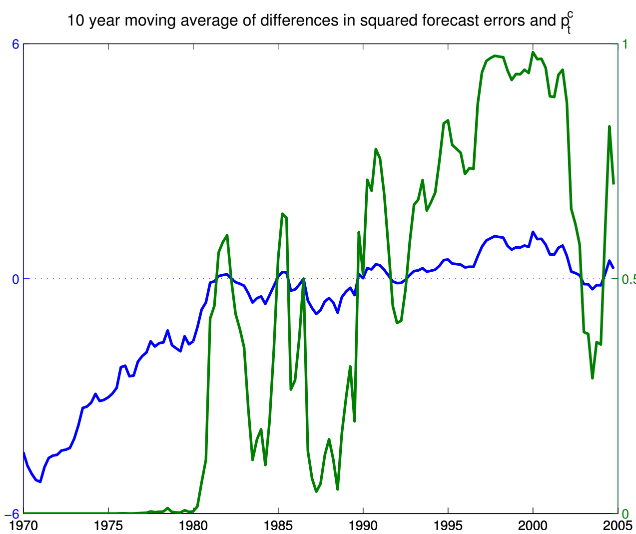

Figure 5: Moving average of difference in squared forecast errors vs p_t^c

6 Data

In this paper I use a quarterly sample starting in the first quarter of 1960 and ending in the third quarter of 2004. The following data series are used:

- quarterly PCE inflation
- the quarterly average Federal Funds rate.
- deviations of log per capita real output from a trend calculated using the Hodrick-Prescott Filter (Hodrick & Prescott (1997)) in real time. I calculate the trend using the Hodrick-Prescott filter for every period in the sample (using only data up until that time period), then calculate the deviations of the most recent observation from the most recent value for the trend and build up a sample for the output gap that way. I do so since I use conditional one-step-ahead forecast densities to calculate the likelihood, which could be problematic if the trend is calculated only once using the entire sample.

The raw data for the Federal Funds rate and per capita output are the same that are used in Smets & Wouters (2007) and more information about the data can be found in that paper. The source for the PCE price index used for calculating PCE inflation is the Bureau of Economic Analysis. Per capita real output and the PCE price index are seasonally adjusted.

7 Recursive Computation of the Likelihood

The components of the likelihood function can be recursively computed as follows:

1. Using prior model probabilities p^c and p^d, initial state variables for both submodels, X_0^c and X_0^d and initial observations π_0, y_0 and i_0 as

well as initial values for z_0 and g_0 (set to 0 in the actual estimation) and a parameter vector Θ, do the following:

- calculate the optimal policies under discretion and commitment
- solve for the policy prescriptions of both models at time 0 using the policy rules and initial state vectors for both models X_0^c and X_0^d
- use those policy prescriptions and i_0 to calculate the likelihoods of the submodels at time 0
- use the likelihoods, p^c and p^d, and equation (10) to calculate p_0^c and p_0^d
- use X_0^c and X_0^d to calculate $E_0^c(\pi_2)$, $E_0^c(y_2)$, $E_0^d(\pi_2)$ and $E_0^d(y_2)$ which can be computed using the solution to the two optimal policy problems
- use the calculated expectations and time 1 data to back out z_1 and g_1 using (1) and (2)
- given observations for z and g dated 0 and 1 back out ε_1 and use that to update the multipliers using (14)
- calculate the time 1 contribution to the likelihood of the entire model using ε_1 and its distributional assumption

2. for all $t = 2, ..., T$ do the following:

   - solve for the policy prescriptions of both models at time $t-1$ using the policy rules and state vectors for both models X_{t-1}^c and X_{t-1}^d
   - use those policy prescriptions and i_{t-1} to calculate the likelihoods of the submodels at time $t-1$

- use the likelihoods, p^c and p^d, and equation (10) to calculate p^c_{t-1} and p^d_{t-1}
- use X^c_{t-1} and X^d_{t-1} to calculate $E^c_{t-1}(\pi_{t+1})$, $E^c_{t-1}(y_{t+1})$, $E^d_{t-1}(\pi_{t+1})$ and $E^d_{t-1}(y_{t+1})$ which can be computed using the solution to the two optimal policy problems
- use the calculated expectations and time t data to back out z_t and g_t using (1) and (2)
- given observations for z and g dated $t-1$ and t back out ε_t and use that to update the multipliers using (14)
- calculate the time t contribution to the likelihood of the entire model using ε_t and its distributional assumption

8 Numerical Implementation of the Metropolis-Hastings Algorithm

I use a standard Random Walk Metropolis-Hastings algorithm to generate draws from the posterior distribution of the parameters. I generate $N = 1100000$ draws, where I discard the first 100000 to allow for effects of the initial condition to wear off.

To generate a starting value Θ_0 I start off different optimizers [7] at 1000 randomly chosen points in the parameter space to get a preliminary estimate of the posterior mode [8]. I also calculated the negative of the inverse Hessian at that point. Usually a scaled version of that matrix is used as the innovation matrix in the random walk proposal, but I found for my purposes that a weighted average (with weight .5) of that matrix and the prior covariance matrix worked better numerically. Note that nothing in the theory of the Metropolis-Hastings algorithm requires the innovation matrix to be a scaled version of the negative inverse Hessian.

A detailed description of the Metropolis-Hastings algorithm can be found in, for example, An & Schorfheide (2007). The posterior distributions of all counterfactuals and statistics such as p_t^c are calculated using 50000 draws, which are drawn uniformly from the entire set of 1000000 draws.

[7] I used the Knitro suite of optimizers.

[8] I maximized the posterior kernel since the normalizing constant is notoriously hard to estimate.

9 Posterior Distribution

Table 1: Statistics of Posterior Distribution

Variable	Posterior Mode	Posterior Mean	5th Percentile	95th Percentile
κ	0.6994	0.7304	0.6818	0.7581
σ	1.6055	1.6026	1.4855	1.6906
ρ_g	0.4021	0.4105	0.3851	0.4359
ρ_z	0.5727	0.5926	0.5651	0.6211
σ_ν	0.0061	0.0063	0.0057	0.0071
σ_z	0.0078	0.0081	0.0075	0.0086
σ_g	0.0157	0.0155	0.0148	0.0161
ρ_{gz}	0.6074	0.6326	0.5656	0.6815
$400 * \pi^c$	1.76	1.92	1.76	2.08
$400 * \pi^d$	5.48	5.12	4.76	5.48
$16 * \lambda^c$	0.0656	0.3024	0.0224	0.6768
$16 * \lambda^d$	0.4944	0.5280	0.2880	0.7216
λ_i^c	0.1128	0.1099	0.0997	0.1199
λ_i^d	0.3149	0.3123	0.2608	0.3689
r	0.0076	0.0059	0.0043	0.0081

10 Alternative Specifications of the Hypothetical Policymakers

So far, the private sector in this model has considered two hypothetical policymakers that differ across two dimensions: the timing protocol of policymaking (discretion or commitment) and their preferences over paths of the inflation rate, the nominal interest rate and the output gap. This section describes various robustness checks with respect to the set of hypothetical policymaker considered by the private sector. The learning setup is the same as in the benchmark model: private agents consider two models of monetary policymaking and use observed data to figure out which of the two policy prescriptions is more likely to have generated the data. Three alternatives are considered:

1. Two policymakers with different preferences who solve for the optimal policy under commitment (not taking into account that the private sector is learning, just as in the benchmark case).

2. The corresponding case where two discretionary policymakers with different preferences solve for the optimal policy (not taking into account that the private sector is learning, just as in the benchmark case).

3. A committed and a discretionary policymaker who share the same preferences solve for the optimal monetary policy under their respective timing protocols (not taking into account that the private sector is learning, just as in the benchmark case).

As outlined in the main text, a discretionary and a committed policymaker generally differ because acting under discretion (i.e. in Markov-perfect equilibrium) there is both an average inflation bias and a stabilization bias - the discretionary central bank has less power to smooth the effects of shocks over time. The third scenario checks whether these two biases alone can induce enough differences between a committed and a discretionary policymaker to make the data fit as well as the benchmark case in which I impose differences in average inflation outcomes (whereas in the third scenario those differences are endogenously coming from the average inflation bias induced by different output targets) and differences in preferences that help fit the data. In the first and second case the priors in the Bayesian estimation of the model are the same as for the original model, i.e. the priors on the preference parameters for one of the policymakers is equal to the priors for the committed policymaker in the baseline model, while the priors for the preference parameter of the other policymaker correspond to the priors attached to the preference parameters of the discretionary policymaker in the baseline model.

The third case features a smaller parameter space since the preference parameters are the same across the two hypothetical policy makers. For that case, I have adjusted the prior mean of the preference parameters of the policymakers to be the mean of the corresponding preference parameters in the other cases [9]. In the third case I allow for a non-zero output gap target. Since preferences are the same across policymakers the identification argument for not including such an output gap target is no longer valid. The introduction of this additional parameter introduces the possibility of an average inflation bias arising for the discretionary policymaker (if the output gap target is not equal to zero): the private sector thinks that the average inflation under the discretionary policymaker will be higher than under the committed policymaker even though both share the same preferences. I use a uniform prior on the output gap target, with a lower bound of 0 and the upper bound of $\max(y_t)$. The only prior restrictions I thus enforce on this parameter estimate are non-

[9] The prior standard deviations are the same for corresponding policy parameters in the other cases. Therefore I did not modify any prior standard deviations.

negativity and an upper bound of the maximum observed output gap. As we will see below it is important to impose an upper bound on this parameter estimate.

For the sake of brevity, I focus on the posterior mode estimates. All estimation results reported in this section also use the prior on reasonable policy modes as described above. Estimates labeled 1 in the table below are estimates of the preference parameters based on the prior used for the commitment central bank in the original model, while estimates labeled 2 are based on the prior for the discretionary policymaker.

TABLE 2 HERE

The preference parameter estimates differ substantially across the specifications, with the hypothetical policymakers in the case of two committed policymakers not caring at all about the output gap and very little about interest rate deviations from target. In the case of two discretionary hypothetical policymakers, these two policymakers turn out to have the same target for inflation[10]. The non-policy parameters are very similar across the first two specifications (and also similar to the estimates of the original model). The third specification leads to substantially different parameter estimates for that set of parameters, though. In particular, the z process is almost white noise and the estimate of κ is substantially lower, implying more price rigidity is needed to fit the data. In addition, the estimate of the output gap target is equal to the upper bound of the uniform prior distribution discussed above. To fit the data, the third specification calls for an unreasonably high output gap target

[10]This is an artefact of a restriction imposed during estimation: the inflation target of the policymaker with the lower prior mean on the inflation target has to be less or equal to the other policymakers inflation target. If that restriction is not imposed the two inflation targets are still very similar, but the policymaker that starts out with a lower prior mean on the inflation target has a posterior mode estimate for the inflation target that is 0.2 annualized percentage points higher. Other estimates are basically unchanged in that case. This switching in the ordering of inflation targets can happen because the prior distributions for the other preference parameters of the two policymakers are also different.

and more price rigidity than what is otherwise needed.

In order to give a first estimate of the relative fit of the specifications considered here, table 2 reports the value of the posterior kernel at the posterior mode. The corresponding number for the main model is 1908, a number 10 log points higher than the highest number reported in table 2. While this exercise can not replace the calculation of Bayesian model probabilities, it is hopefully at least indicative of the relative empirical performance of the various model specifications.

Table 2: Posterior Modes, Alternative Specifications of Hypothetical Policymakers

Variable	First Case	Second Case	Third Case
κ	0.7459	0.7512	0.0944
σ	1.7020	1.3912	2.0577
ρ_g	0.3133	0.4072	0.4673
ρ_z	0.6478	0.6176	0.0406
σ_ν	0.0070	0.0059	0.0060
σ_z	0.0078	0.0082	0.0058
σ_g	0.0150	0.0154	0.0151
ρ_{gz}	0.6500	0.6134	0.2956
$400 * \pi^1$	2.3326	4.0967	3.2587
$400 * \pi^2$	7.5872	4.0967	-
$16 * \lambda^1$	0	0	0.1280
$16 * \lambda^2$	0	1.2112	-
λ_i^1	0.0530	0.4214	0.0917
λ_i^2	0.0398	0.3056	-
r	0.0073	0.0069	0.0057
log likelihood + log (ddp)	1860.2	1898.3	1890.8

11 Inference Using a Conditional Likelihood Function

As described in the main text, I use a likelihood function conditional on the path of nominal interest rates to shield myself against possible misspecification of the actual monetary policy rule. To see the advantages and disadvantages of such an approach, let's consider for a moment building a joint model for $(y^T, \pi^T, i^T) = X^T$. As in the main text, the parameters governing private agents' behavior are called Θ. The parameters of the actual policy rule are called Γ. Note that Γ could be very high dimensional if time variation in the monetary policy rule coefficients is allowed for. For convenience, I will drop the conditioning argument I (the initial values used in estimation) throughout this exposition. We can think of all densities possibly being conditioned on I. It is crucial to realize that the posterior density for Θ derived in this section implicitly conditions on the model for the actual (and not only the perceived) evolution of i^T, even after we integrate over the parameter values for that actual policy rule. In contrast, the posterior used in the paper does not depend on a specific model for the actual evolution of nominal interest rates. The posterior of the joint model is

$$p(\Theta, \Gamma | X^T) \propto p(\Theta, \Gamma) p(X^T | \Theta, \Gamma) \tag{7}$$

Next, let's make the common assumption of independent priors:

$$p(\Theta, \Gamma) p(X^T | \Theta, \Gamma) = p(\Theta) p(\Gamma) p(X^T | \Theta, \Gamma) \tag{8}$$

We can rewrite the likelihood to arrive at:

$$p(\Theta, \Gamma) p(X^T | \Theta, \Gamma) = p(\Theta) p(\Gamma) p(y^T, \pi^T | \Theta, \Gamma, i^T) p(i^T | \Theta, \Gamma) \tag{9}$$

Given the assumption made on the learning model, $p(y^T, \pi^T|\Theta, \Gamma, i^T)$ does not depend on Γ once we condition on the path of nominal interest rates: private sector behavior only depends on monetary policy through the actual path of the monetary policy instruments since that is what determines private agents' beliefs and influences their actions. Thus the posterior can be rewritten as (up to a constant factor)[11]:

$$p(\Theta)p(\Gamma)p(y^T, \pi^T|\Theta, i^T)p(i^T|\Theta, \Gamma) \tag{10}$$

Integrating this expression over Γ yields:

$$p(\Theta|y^T, \pi^T, i^T) \propto p(\Theta)p(y^T, \pi^T|\Theta, i^T) \int p(\Gamma)p(i^T|\Theta, \Gamma)d\Gamma \tag{11}$$

In this paper I am only interested in inference on Θ. Building a joint model of output, inflation and nominal interest rates only helps me in this endeavor through $p(i^T|\Theta, \Gamma)$[12]. The behavior of nominal interest rates will tell me something about Θ because in general any reasonable model of nominal interest rates will depend on output and inflation, which are determined by private agents (and thus depend on Θ). This information, which is most likely not as useful as that contained directly in output and inflation data, comes at the cost of possible misspecification. If $p(i^T|\Theta, \Gamma)$ is misspecified, then any gains from modeling the monetary policy rule will most likely vanish quickly. In this paper, I follow a substantial part of the learning literature (e.g. Sargent, Williams & Zha (2006)) and do not assume that private agents necessarily have the correct model of the

[11]Integrating this posterior with respect to Γ will give the posterior of Θ conditional on a specific model for the evolution of i^T. Below I describe why this posterior will in general differ from the posterior when I do not make an assumption on what model determines i^T, as is done in the paper. Note that even when one integrates out the specific parameters for the monetary policy rule Γ, the form of that policy rule will still matter for the marginal posterior of Θ if one takes a stand on a model for the policy instrument.

[12]If $p(i^T|\Theta, \Gamma)$ does not depend on Θ, then the posterior above and the posterior used in the paper will exactly coincide since $\int p(\Gamma)p(i^T|\Theta, \Gamma)d\Gamma$ can then be subsumed into the factor of proportionality in the equation determining the marginal posterior of Θ.

economy, which is why I introduced a separate vector of parameters Γ for the actual policy rule in this section. If the parameters of the perceived policy rule would be a part of Γ, then the above derivation would obviously not hold. If we were to impose that the true policy rule would switch between the two perceived types according to an exogenous (and persistent) two-state Markov Chain[13] (after all we should allow for some switching between the two types), but if the real policy rule is not of this type, then we introduce severe misspecification not present in the estimation algorithm in the paper. This is exactly the misspecification I want to avoid. On the other hand, if the proposed policy rule is the actual data generating process then the estimated model probabilities of the agents will give a good approximation to filtered estimates of the model probabilities and the gains from efficiency coming from explicitly modeling the policy rule are likely to be small. This happens because in the case in which the discrete state Markov Chain is persistent, looking at the last 10 years of data and calculating which model is more likely to have generated the data (as private agents do in the model), will be a good approximation to filtered model probabilities calculated using the Markov chain approach.

12 Using a Simulation-Based Prior

This section uses an alternative set of prior distributions to assess the robustness of the estimates presented above. I keep the "standard" priors described earlier, but replace the data-driven prior with a prior on the

[13]Having this feature in addition to the learning would make inference substantially harder from a computational perspective, but let's use this setup for our thought experiment.

mean and variance of the average policy prescription[14]:

$$i_t^{av} \equiv \frac{i_t^c + i_t^d}{2} \tag{12}$$

If we denote the vector of the sample mean and sample variance of i_t^{av} by χ, then the prior distribution I will be using has the form

$$\begin{aligned} \underbrace{p(\Theta, \chi | y^T, \pi^T, I, i^T)}_{posterior} &\propto \underbrace{p(\Theta)}_{prior} \underbrace{p(y^T, \pi^T | \Theta, I, i^T)}_{likelihood} \\ = \underbrace{p_1(\Theta)}_{\text{"}standard\text{" } prior} & \underbrace{p_2^{sim}(\chi(\Theta)|\Theta, I, i^T)}_{simulation\ based\ prior} \underbrace{p(y^T, \pi^T | \Theta, I, i^T)}_{likelihood} \end{aligned} \tag{13}$$

The above expression emphasizes the dependence of χ on the parameter vector Θ. Furthermore, p_2^{sim} conditions on the path of observed interest rates and the initial values used for the computation of the likelihood. Since the likelihood also conditions on these variables, an application of Bayes' Law shows that this conditioning is valid. As mentioned before, there are in general no closed form expressions for χ as a function of θ and the condition arguments of p_2^{sim} available. Therefore, following Gallant & McCulloch (2009), I use a simulation-based prior. When simulating, I take as given the same initial values I used to compute the likelihood and the path of actual nominal interest rates. For every evaluation of the prior, I use $N = 500$ simulated paths of the exogenous disturbances and then calculate the mean and variance of i_t^{av} across those simulations and across time[15]. I then choose a shape for p_2^{sim} (in this case two independent Gaussian distributions, one of which is truncated since the variance is non-negative) and evaluate p_2^{sim} at the given value of Θ and the calculated value for χ. This approach to modeling prior information

[14] i_t^{av} is defined as the simple average of the policy prescriptions instead of the probability weighted average to decrease identification problems in areas of the parameter space where one model of policymaking has very low probability. Remember that a similar argument was made when defining the data driven prior earlier.

[15] To reduce simulation noise, I keep the standardized values of the disturbances fixed across evaluations of the prior density.

in Bayesian inference is similar in spirit to the approach proposed in Del Negro & Schorfheide (2008)[16].

Table 3 shows the estimation results using this approach, focusing on the posterior mode for sake of comparability. For most parameters, the estimates are reasonably close to the estimates calculated using the data driven prior. The main difference are the preferences of the hypothetical committed central bank, which now cares substantially more about fluctuations in the output gap and the nominal interest rate. The main conclusions found using the original approach do carry over to this approach, as can be seen by looking at the estimated series for p_t^c (figure 6) implied by the posterior mode calculated using the simulation based prior. While this probability goes to zero slower at the beginning of the sample and does not rise as fast in 1980, the overall pattern does remain the same. To confirm this, figure 7 plots the private sector expectations using the posterior mode estimates calculates using the simulation-based prior, which are very similar to the expectations obtained using the original approach[17].

[16]The prior mean for the sample average is 6 % annualized and the standard deviation is .5% annualized. While this is a rather tight prior, I want to impose the strong prior belief that the average policy prescription is close to the levels of the interest rate that we have seen since 1960. The prior mean for the sample variance is 7e-5, which is roughly the variance of the nominal interest rate used in the estimation. The standard deviation of the normal distribution for the variance (before truncation) is 1e-5.

[17]Other results such as counterfactual outcomes are also broadly similar.

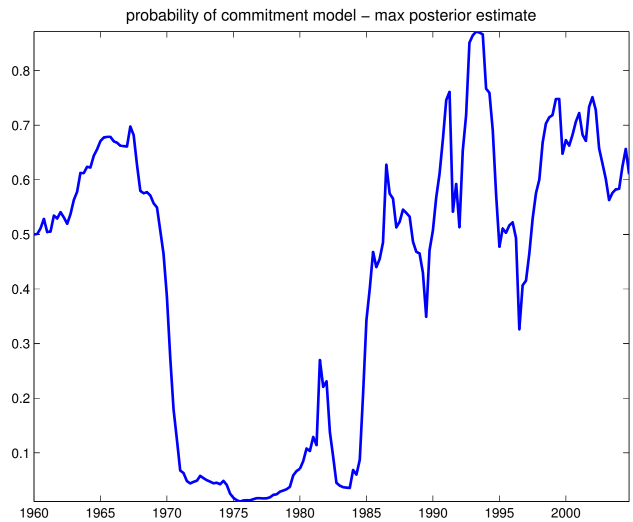

Figure 6: Posterior mode estimate of p_t^c, simulation-based prior

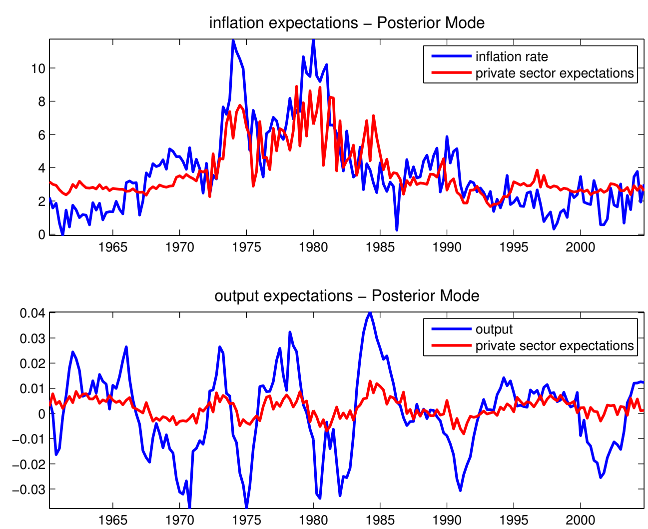

Figure 7: Private sector expectations, simulation-based prior

Table 3: Posterior Mode Estimates, Simulation-Based Prior

Variable	Posterior Mode
κ	0.4486
σ	1.4074
ρ_g	0.4334
ρ_z	0.3217
σ_ν	0.0126
σ_z	0.0074
σ_g	0.0160
ρ_{gz}	0.6570
$400\pi^c$	1.5008
$400\pi^d$	6.0911
$16 * \lambda^c$	2.67
$16 * \lambda^d$	0.0064
λ_i^c	0.2697
λ_i^d	0.3826
r	0.0067

13 Additional Robustness Exercises

This section briefly shows results for two additional robustness exercises. For the sake of brevity, I only show a subset of the results for these cases. First, what would happen if agents use the entire sample to form model expectations? Results are very similar to the benchmark case. Figures 8 and 9 show results for this case. The model probabilities evolve in a smoother fashion (as expected), but are qualitatively and quantitatively very similar. The same is true for the expectations of private agents. Other results (such as the counterfactuals) are not effected by this assumption either.

The paper shows that Bayes rule calls for a 'prior' conditional on i^T. To see the effect of this 'prior', I re-estimated the model without it. Figure 10 shows the expected interest rate at the posterior mode for that case. Agents hold highly unreasonable views in this case.

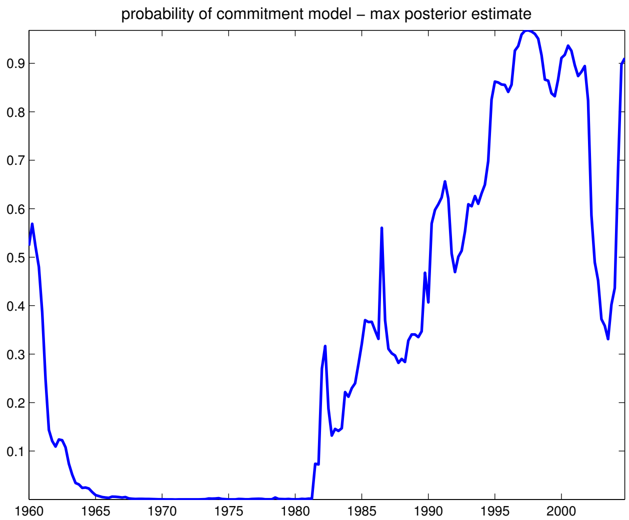

Figure 8: Probability associated with committed policymaker, agents use full sample

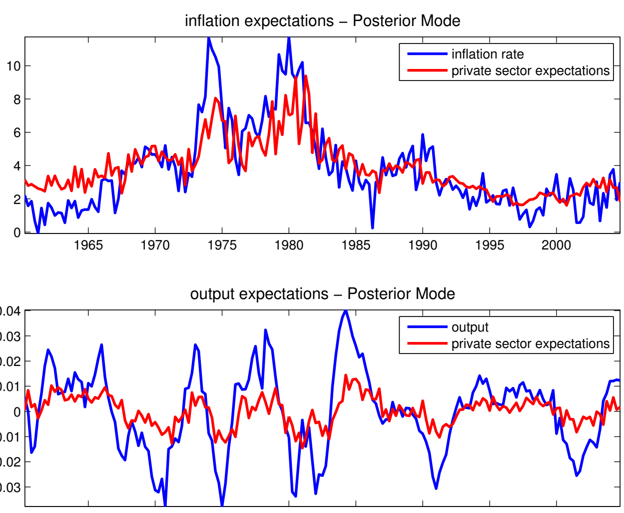

Figure 9: Private sector expectations, agents use full sample

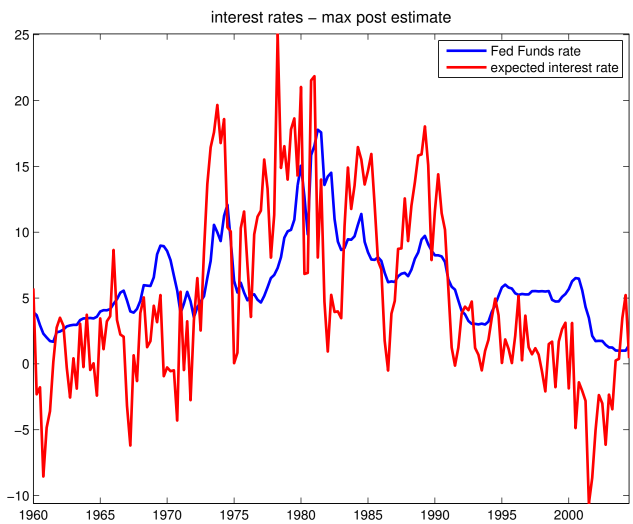

Figure 10: Expected interest rates, estimation without $p(\Theta|i^T)$

14 Proof that Bayesian Model Probabilities are Martingales

Let x be a variable of interest (in our case x corresponds to the nominal interest rate). Suppose a Bayesian entertains two models for x, which I will denote model A (with prior model probability $P(A)$ and likelihood $p(x|A)$) and model B (with prior model probability $P(B)$ and likelihood $p(x|B)$). Bayes' theorem tells us that

$$P(A|x) = \frac{p(x|A)P(A)}{p(x)}$$

where $p(x) = p(x|A)P(A) + p(x|B)P(B)$. Let's next calculate the expectation of $P(A|x)$:

$$E(P(A|x)) = E(\frac{p(x|A)P(A)}{p(x)}) = \int \frac{p(x|A)P(A)}{p(x)} p(x)dx$$

$$= \int p(x|A)P(A)dx = P(A) \int p(x|A)dx = P(A)$$

The same calculations go through when a Bayesian sequentially updates his prior (in which case we have to explicitly use conditional expectations in the equation above).

References

An, S. & Schorfheide, F. (2007), 'Bayesian analysis of DSGE models', *Econometric Reviews* **26**, 113–172.

Del Negro, M. & Schorfheide, F. (2008), 'Forming priors for DSGE models (and how it affects the assessment of nominal rigidities)', *Journal of Monetary Economics* **55**(7), 1191–1208.

Gallant, A. R. & McCulloch, R. E. (2009), 'On the determination of general scientific models with application to asset pricing', *Journal of the American Statistical Association* **104**(485), 117–131.

Hodrick, R. J. & Prescott, E. C. (1997), 'Postwar U.S. business cycles: An empirical investigation', *Journal of Money, Credit and Banking* **29**(1), 1–16.

Lubik, T. A. & Schorfheide, F. (2004), 'Testing for indeterminacy: An application to U.S. monetary policy', *American Economic Review* **94**(1), 190–217.

Sargent, T., Williams, N. & Zha, T. (2006), 'Shocks and government beliefs: The rise and fall of American inflation', *American Economic Review* **96**(4), 1193–1224.

Smets, F. & Wouters, R. (2007), 'Shocks and frictions in US business cycles: A Bayesian DSGE approach', *American Economic Review* **97**(3), 586–606.

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