Online Supplementary Materials

Modeling Changes in U.S. Monetary Policy with a Time-Varying Nonlinear Taylor Rule

Anh D. M. Nguyen, Efthymios G. Pavlidis, and David A. Peel*

1 Asymmetric Monetary Policy Rule

The Lagrangian of the policy problem is written as follows

$$\underset{\pi_t, y_t, i_t}{Min} E_t\Big\{\frac{e^{\alpha(\pi_t-\pi^*)}-\alpha(\pi_t-\pi^*)-1}{\alpha^2}+\frac{\mu}{2}(y_t)^2+\frac{\gamma}{2}(i_t-i^*)^2 \\ -\phi_t^{\pi}(\pi_t-\kappa y_t-\varepsilon_t^s)-\phi_t^{y}(y_t+\varphi i_t-\varepsilon_t^d)\Big\},$$

where ϕ_t^{π} and ϕ_t^{y} are the Lagrange multipliers.

It is straightforward to obtain the following first-order conditions

$$E_t(\mu y_t+\phi_t^{\pi}\kappa-\phi_t^{y})=0,$$
$$E_t[\gamma(i_t-i^*)-\phi_t^{y}\varphi]=0,$$
$$E_t\Big\{\frac{e^{\alpha(\pi_t-\pi^*)}-1}{\alpha}-\phi_t^{\pi}\Big\}=0.$$

Using these conditions, we obtain

$$E_t\Big\{\frac{e^{\alpha(\pi_t-\pi^*)}-1}{\alpha}+\frac{\mu}{\kappa}y_t-\frac{\gamma}{\varphi\kappa}(i_t-i^*)\Big\}=0. \tag{1.1}$$

Based on (1.1), the central bank sets the interest rate according to

$$i_t=i^*+E_t\Big\{\frac{e^{\alpha(\pi_t-\pi^*)}-1}{\alpha}\frac{\varphi\kappa}{\gamma}+\frac{\mu\varphi}{\gamma}y_t\Big\},$$

*Nguyen (corresponding author): Economics Department, Bank of Lithuania and Faculty of Economics, Vilnius University, Lithuania, Email: anguyen@lb.lt; Pavlidis: Department of Economics, Lancaster University, Lancaster, UK; and Peel: Department of Economics, Lancaster University, Lancaster, UK.

and the above expression can be approximated as

$$
\begin{aligned}
i_t &= i^* + E_t\Big\{(\pi_t - \pi^*)\frac{\varphi\kappa}{\gamma} + (\pi_t - \pi^*)^2\frac{\alpha\varphi\kappa}{2\gamma} + \frac{\mu\varphi}{\gamma}y_t\Big\} \\
&= i^* + \frac{\varphi\kappa}{\gamma}E_t(\pi_t - \pi^*) + \frac{\alpha\varphi\kappa}{2\gamma}E_t(\pi_t - \pi^*)^2 + \frac{\mu\varphi}{\gamma}E_t y_t.
\end{aligned} \tag{1.2}
$$

2 Particle-filtering algorithm

The state-space system is written in its probabilistic form as follows

$$\mathbf{x}_t = h(\mathbf{x}_{t-1}, \mathbf{w}_t; \varpi), \tag{2.1}$$

$$i_t = g(\mathbf{x}_t, \varepsilon_t; \varpi, \mathbf{A}_t), \tag{2.2}$$

where $\mathbf{w}_t = [\varepsilon_{0,t}, \varepsilon_{1,t}, \varepsilon_{2,t}, \varepsilon_{3,t}, \varepsilon_{4,t}, \varepsilon_{5,t}]'$ is the vector of state noises, which has a multivariate normal distribution with zero mean and identity covariance matrix, $\varpi = [\sigma_{a_0}, \sigma_{a_1}, \sigma_{a_2}, \sigma_{a_3}, \sigma_{a_4}, \sigma_{a_5}]'$ presents the set of time-invariant parameters, and $\mathbf{A}_t = [i_{t-1}, \pi_{t|t}, y_{t|t}, \sigma^2_{\pi_t|t}]$ includes observed inputs. To ease notation, in what follows we drop $\mathbf{A}_t$ without any loss of generality. The functions $h(\cdot)$ and $g(\cdot)$ come from the equations that characterize the behavior of the model respectively.

Let $X_t = \{\mathbf{x}_j, j = 0, ..., t\}$ and $I_t = \{i_j, j = 0, ..., t\}$ represent the sequences of all states and available measurements, respectively, up to time t. The joint posterior density at time t is denoted by $p(X_t|I_t)$ and its marginal is $p(\mathbf{x}_t|I_t)$. Let $\{X_t^k, \omega_t^k\}_{k=1}^N$ denote a random measure that describes the joint posterior $p(X_t|I_t)$ where $\{X_t^k, k = 1, ..., N\}$ is a set of support points with associated weights $\{\omega_t^k, k = 1, ..., N\}$. The weights are normalized by dividing each by their sum. Thus, the joint posterior distribution at t can be approximated by

$$p(X_t|I_t) \approx \sum_{k=1}^{N} \omega_t^k \delta(X_t - X_t^k), \tag{2.3}$$

where $\delta(.)$ is the Dirac delta measure. The normalized weights ω_t^k are chosen by applying the principle of importance sampling in which X_t^k is drawn from an importance density $q(X_t|I_t)$

$$\omega_t^k \propto \frac{p(X_t^k|I_t)}{q(X_t^k|I_t)}. \tag{2.4}$$

If the importance density is chosen so that it can be factorized by

$$q(X_t|I_t) \triangleq q(\mathbf{x}_t|X_{t-1}, I_t)q(X_{t-1}|I_{t-1}), \tag{2.5}$$

then the samples $X_t^k \sim q(X_t|I_t)$ can be achieved by augmenting each of the existing samples $X_{t-1}^k \sim q(X_{t-1}|I_{t-1})$ with the new state $\mathbf{x}_t^k \sim q(\mathbf{x}_t|X_{k-1}, I_k)$. At time step t when a measurement i_t becomes available, the posterior density $p(X_t|I_t)$ can be updated from

$p(X_{t-1}|I_{t-1})$ by

$$\begin{aligned} p(X_t|I_t) &= \frac{p(i_t|X_t, I_{t-1})p(X_t|I_{t-1})}{p(i_t|I_{t-1})} \\ &= \frac{p(i_t|X_t, I_{t-1})p(\mathbf{x}_t|X_{t-1}, I_{t-1})p(X_{t-1}|I_{t-1})}{p(i_t|I_{t-1})} \\ &= \frac{p(i_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1})p(X_{t-1}|I_{t-1})}{p(i_t|I_{t-1})} \\ &\propto p(i_t|\mathbf{x}_t)p(\mathbf{x}_t|\mathbf{x}_{t-1})p(X_{t-1}|I_{t-1}). \end{aligned} \tag{2.6}$$

Substituting (2.5) and (2.6) into (2.4) yields the weight update equation

$$\begin{aligned} \omega_t^k &\propto \frac{p(i_t|\mathbf{x}_t^k)p(\mathbf{x}_t^k|\mathbf{x}_{t-1}^k)p(X_{t-1}^k|I_{t-1})}{q(\mathbf{x}_t^k|X_{t-1}^k, I_t)q(X_{t-1}^k|I_{t-1})} \\ &= \omega_{t-1}^k \frac{p(i_t|\mathbf{x}_t^k)p(\mathbf{x}_t^k|\mathbf{x}_{t-1}^k)}{q(\mathbf{x}_t^k|X_{t-1}^k, I_t)}. \end{aligned}$$

Moreover, by choosing the importance density to depend only on $\mathbf{x}_{t-1}$ and i_t, the weights are given by

$$\omega_t^k \propto \omega_{t-1}^k \frac{p(i_t|\mathbf{x}_t^k)p(\mathbf{x}_t^k|\mathbf{x}_{t-1}^k)}{q(\mathbf{x}_t^k|\mathbf{x}_{t-1}^k, i_t)}.$$

In this study, we use the bootstrap filtering proposed by Gordon et al. (1993) in which $q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{z}_t) = p(\mathbf{x}_t|\mathbf{x}_{t-1})$. Therefore,

$$\omega_t^k \propto \omega_{t-1}^k p(i_t|\mathbf{x}_t^k).$$

Given these weights, the marginal posterior density $p(\mathbf{x}_t|I_t)$ can be approximated as

$$p(\mathbf{x}_t|I_t) \approx \sum_{k=1}^{N} \omega_t^k \delta(\mathbf{x}_t - \mathbf{x}_{t-1}^k). \tag{2.7}$$

As $N \to \infty$ the approximation (2.7) approaches the true marginal posterior density $p(\mathbf{x}_t|I_t)$ (Ristic et al., 2004).

It is however worth emphasizing that, given the importance function of the form (2.5), the variance of importance weights can only increase over time, thus leading to the degeneracy problem (Ristic et al., 2004). Therefore, resampling that replaces the samples with low importance weights by those with high importance weights is required. There are many different resampling schemes, which can be referred to Doucet and Johansen (2009), but we use the systematic resampling method because it is easy to apply and outperforms other resampling schemes in most cases (Doucet and Johansen, 2009).

3 An approximation for the likelihood value

The likelihood function is derived as follows

$$
\begin{aligned}
p(I_T;\varpi) &= \prod_{t=1}^{T} p(i_t|I_{t-1};\varpi) \\
&= \prod_{t=1}^{T} \left(\int p(i_t|\mathbf{x}_t;\varpi) p(\mathbf{x}_t|I_{t-1};\varpi) d\mathbf{x}_t \right) \\
&= \prod_{t=1}^{T} \left(\int \int p(i_t|\mathbf{x}_t;\varpi) p(\mathbf{x}_t|X_{t-1};\varpi) p(X_{t-1}|I_{t-1};\varpi) dX_{t-1} d\mathbf{x}_t \right) \\
&= \prod_{t=1}^{T} \left(\int \int p(i_t|\mathbf{x}_t;\varpi) p(\mathbf{x}_t|\mathbf{x}_{t-1};\varpi) p(X_{t-1}|I_{t-1};\varpi) dX_{t-1} d\mathbf{x}_t \right) \\
&= \prod_{t=1}^{T} \left(\int \frac{p(i_t|\mathbf{x}_t;\varpi) p(\mathbf{x}_t|\mathbf{x}_{t-1};\varpi)}{\pi(\mathbf{x}_t|X_{t-1}, I_t)} \pi(\mathbf{x}_t|X_{t-1}, I_t) p(X_{t-1}|I_{t-1};\varpi) dX_t \right) \\
&\approx \prod_{t=1}^{T} \left(\frac{1}{N} \sum_{k=1}^{N} \frac{p(i_t|\mathbf{x}_t^k;\varpi) p(\mathbf{x}_t^k|\mathbf{x}_{t-1}^k;\varpi)}{\pi(\mathbf{x}_t^k|X_{t-1}^k, I_t)} \right) \\
&= \prod_{t=1}^{T} \left(\sum_{k=1}^{N} \omega_t^{k-1} \frac{p(i_t|\mathbf{x}_t^k;\varpi) p(\mathbf{x}_t^k|\mathbf{x}_{t-1}^k;\varpi)}{\pi(\mathbf{x}_t^k|X_{t-1}^k, I_t)} \right) \\
&= \prod_{t=1}^{T} \left(\sum_{k=1}^{N} \omega_t^k \right).
\end{aligned}
$$

In the above derivation, we used an assumption that $\pi(X_{t-1}|I_{t-1}) = \pi(X_{t-1}|I_t)$.

4 Constructing the Contemporaneous Real-Time HP Output Gap Series

In order to construct the contemporaneous real-time HP output gap from 1965Q4 to 2007Q4, we use two data sets: the Greenbook projections and the real-time data set for macroeconomists. Both can be downloaded from the website of the Philadelphia Fed. For more explanations about real-time data, we refer the readers to Orphanides (2001).

The procedure to construct the contemporaneous real-time HP output gap for a given quarter i is described as follows:

Step 1: Collect the entire time-series history perceived at the quarter i vintage (the vintage is shown at the column header of the real-time data set) which includes the real output up to the previous quarter. Note that the real output of quarter i is not available to observe in that quarter. Denote this series by $X_{j:i-1|i} = [x_{j|i}, x_{j+1|i}, x_{j+2|i}, ..., x_{i-1|i}]$ where j is the first quarter with data recorded in the vintage i data set and $x_{h|i}$ is the data of real output of the quarter h perceived at the vintage i.

Step 2: Use the Greenbook forecasts for the quarter-to-quarter growth in real GDP (with quarterized percentage points) to calculate the expected value of real GDP for the contemporaneous quarter $x_{i|i}$ from $x_{i-1|i}$. In order to reduce the end-of-sample issue of the HP filter, we also compute the expected value of real GDP in the following quarters when the forecasts of the growth rate for those quarters are available at that vintage.

Step 3: Combine these expected values with the historical series to generate the new series: $X_{j:i+k|i} = [x_{j|i}, x_{j+1|i}, x_{j+2|i}, ..., x_{i-1|i}, x_{i|i}, ..., x_{i+k|i}]$ where $0 \leq k \leq 4$.

Step 4: Apply the HP filter to the series $X_{j:i+k|i}$ to achieve the HP output gap series $X^*_{j:i+k|i} = [x^*_{j|i}, x^*_{j+1|i}, x^*_{j+2|i}, ..., x^*_{i-1|i}, x^*_{i|i}, ..., x^*_{i+k|i}]$. We then record $x^*_{i|i}$ as the contemporaneous real-time HP output gap at the quarter i.

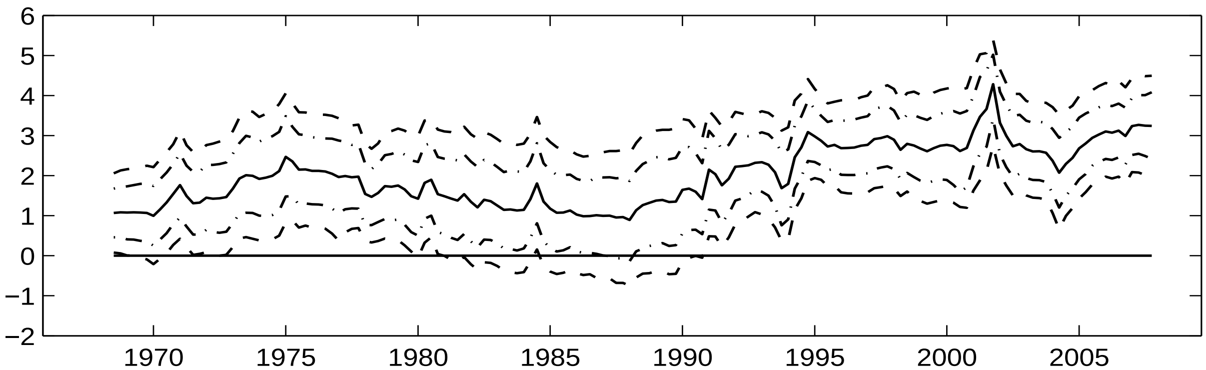
5 Robustness checks with different real activity measures

This section presents the robustness checks with three different measures of real activity. The first two use alternative measures of the natural rate of unemployment: a historical average and a 3-year moving average, leading to the two different measures of unemployment gap. Meanwhile, the third exercise uses the real-time HP output gap, which is constructed as described in section 4. The results of these robustness checks are shown in Figures 1, 2, and 3. As discussed in the baseline, our main findings remain unchanged with respect to the output gap measures.

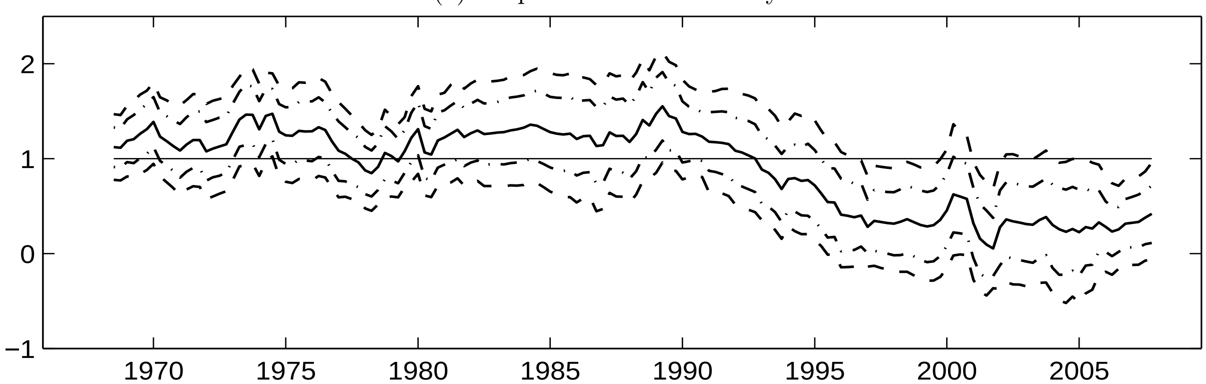
6 Optimal Policy Rule in the Case of Asymmetric Preferences to Both Inflation and Output Gap

The Lagrangian of the policy problem in this case is written as follows

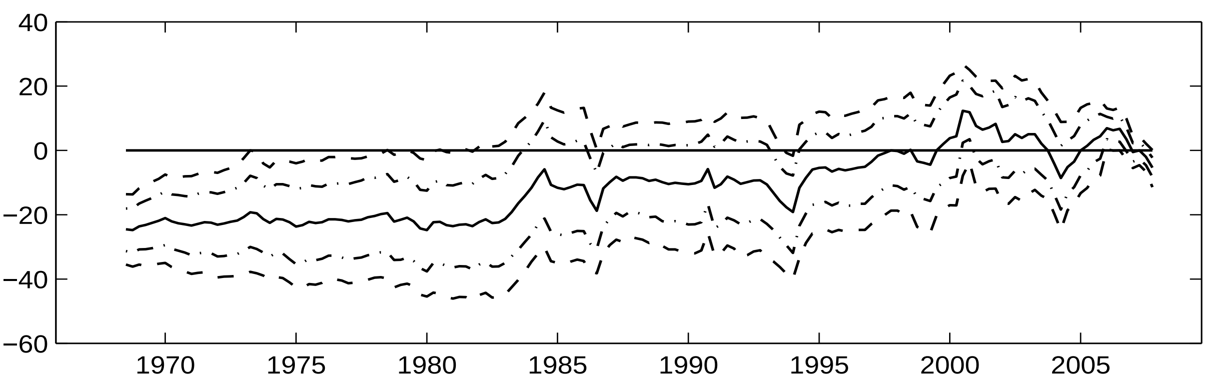
$$\begin{aligned} \underset{\pi_t, y_t, i_t}{Min} E_t \Big\{ & \frac{e^{\alpha(\pi_t - \pi^*)} - \alpha(\pi_t - \pi^*) - 1}{\alpha^2} + \mu \left[\frac{e^{(\lambda y_t)} - \lambda y_t - 1}{\lambda^2} \right] + \frac{\gamma}{2}(i_t - i^*)^2 \\ & - \phi^\pi_t(\pi_t - \kappa y_t - \varepsilon^s_t) - \phi^y_t(y_t + \varphi i_t - \varepsilon^d_t) \Big\}, \end{aligned} \tag{6.1}$$



(a) Response to Real Activity

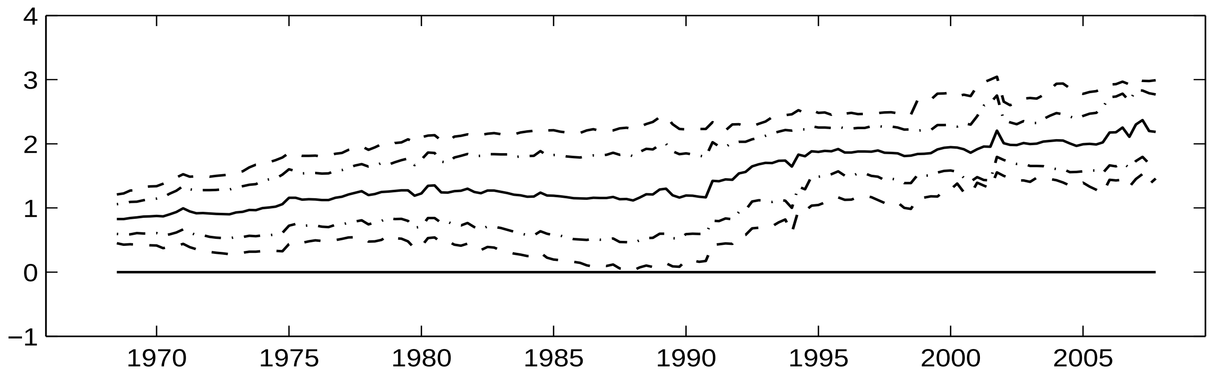


(b) Response to Inflation

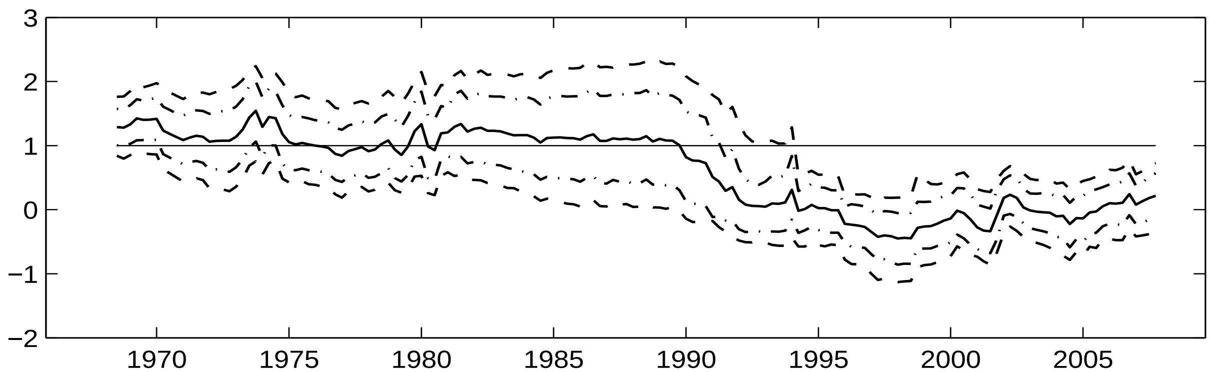


(c) Response to Inflation Variance

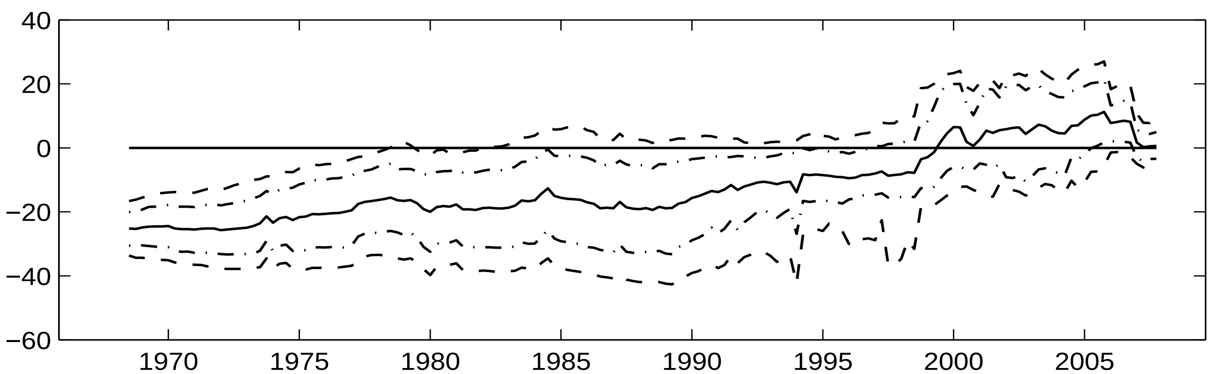
Figure 1: Robustness Check with Three-Year Moving Average Unemployment Gap
Note: Dashed lines are 68% and 90% percentile intervals.



(a) Response to Real Activity



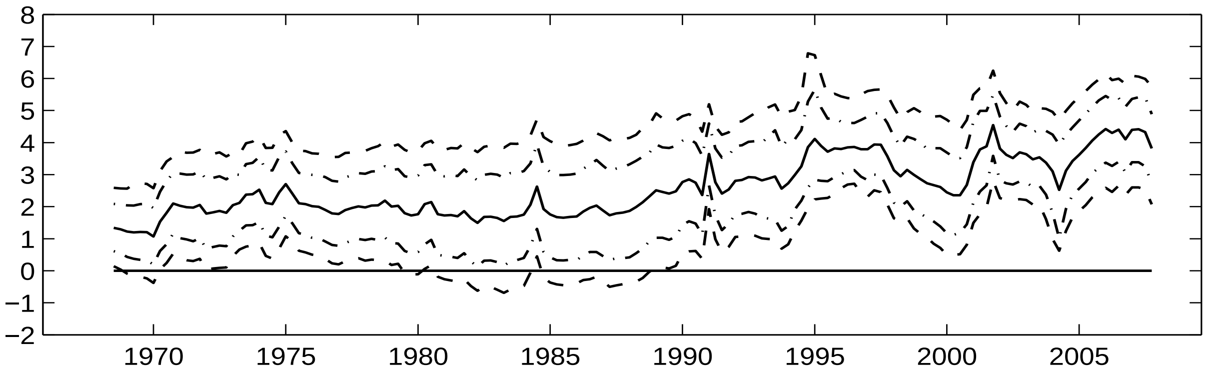
(b) Response to Inflation



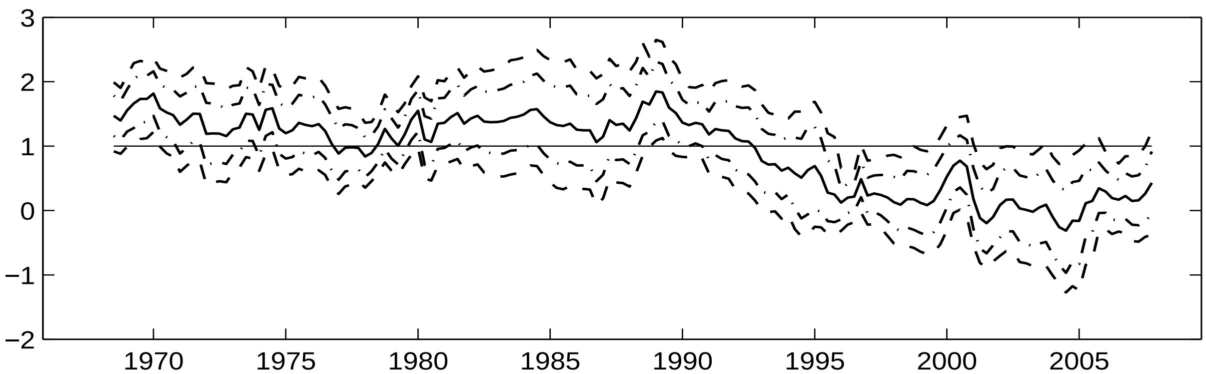
(c) Response to Inflation Variance

Figure 2: Robustness Check with Historical Average Unemployment Gap

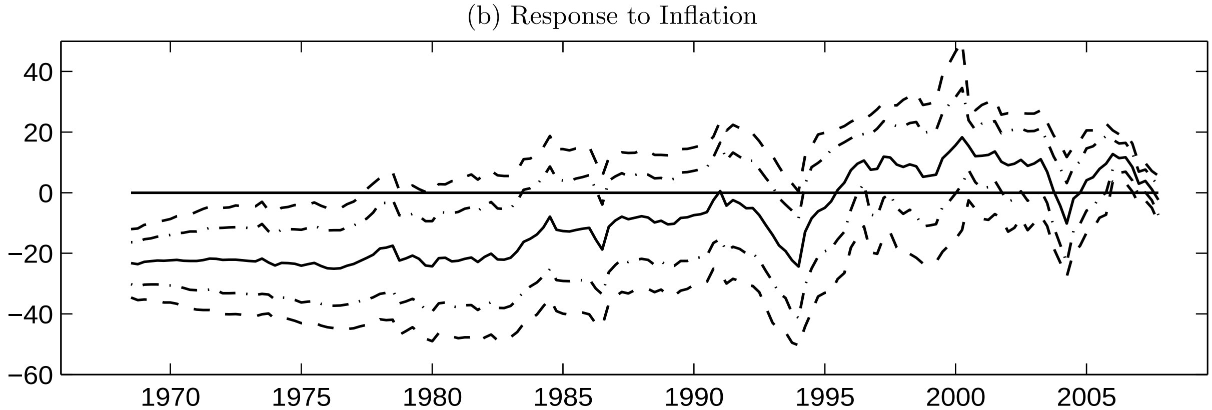
Note: Broken lines are 68% and 90% percentile intervals.



(a) Response to Real Activity



(b) Response to Inflation



(c) Response to Inflation Variance

Figure 3: Robustness Check with HP Output Gap

Note: Broken lines are 68% and 90% percentile intervals.

in which ϕ_t^π and ϕ_t^y are the Lagrange multipliers. It is straightforward to derive the first-order optimal conditions

$$E_t\Big\{\frac{e^{\alpha(\pi_t-\pi^*)}-1}{\alpha}-\phi_t^\pi\Big\}=0,$$
$$E_t(\gamma(i_t-i^*)-\phi_t^y\varphi)=0,$$
$$E_t\Big\{\mu\frac{e^{\lambda y_t}-1}{\lambda}+\phi_t^\pi\kappa-\phi_t^y\Big\}=0.$$

Combine these conditions to eliminate the Lagrange multipliers

$$E_t\Big\{\frac{e^{\alpha(\pi_t-\pi^*)}-1}{\alpha}\kappa+\frac{e^{\lambda y_t}-1}{\lambda}\mu-\frac{\gamma}{\varphi}(i_t-i^*)\Big\}=0. \tag{6.2}$$

Then the central bank sets the interest rate in order to respond to inflation and output deviations

$$i_t=i^*+E_t\Big\{\frac{e^{\alpha(\pi_t-\pi^*)}-1}{\alpha}\frac{\kappa\varphi}{\gamma}+\frac{e^{\lambda y_t}-1}{\lambda}\frac{\mu\varphi}{\gamma}\Big\}. \tag{6.3}$$

and the above expression can be approximated as

$$\begin{aligned}
i_t &= i^*+E_t\Big\{\frac{\kappa\varphi}{\gamma}(\pi_t-\pi^*)+\frac{\kappa\varphi\alpha}{2\gamma}(\pi_t-\pi^*)^2+\frac{\mu\varphi}{\gamma}y_t+\frac{\mu\varphi\lambda}{2\gamma}y_t^2\Big\}\\
&= i^*+\frac{\kappa\varphi}{\gamma}E_t(\pi_t-\pi^*)+\frac{\kappa\varphi\alpha}{2\gamma}E_t(\pi_t-\pi^*)^2+\frac{\mu\varphi}{\gamma}E_ty_t+\frac{\mu\varphi\lambda}{2\gamma}E_ty_t^2\\
&= (i^*-\frac{\kappa\varphi}{\gamma}\pi^*)+\frac{\kappa\varphi}{\gamma}\pi_{t|t}+\frac{\kappa\varphi\alpha}{2\gamma}\sigma^2_{\pi_t|t}+\frac{\mu\varphi}{\gamma}y_{t|t}+\frac{\mu\varphi\lambda}{2\gamma}\sigma^2_{y_t|t}\\
&= b_0+b_1\pi_{t|t}++b_2\sigma^2_{\pi_t|t}+b_3y_{t|t}+b_4\sigma^2_{y_t|t},
\end{aligned} \tag{6.4}$$

where $b_0=i^*-\frac{\kappa\varphi}{\gamma}\pi^*$, $b_1=\frac{\kappa\varphi}{\gamma}$, $b_2=\frac{\kappa\varphi\alpha}{2\gamma}$, $b_3=\frac{\mu\varphi}{\gamma}$, $b_4=\frac{\mu\varphi\lambda}{2\gamma}$ and $\sigma^2_{y_t|t}$ is the expected variance of unemployment gap conditional on the information available at time t. Other notations are as in the baseline model.

7 Expected Variance of Output Gap

The process of generating the expected variance of output gap series is similar to the one used to create the expected variance of inflation series. The model specification used is given by

$$y_t=c+\sum_{i=1}^{3}\beta_iy_{t-i}+\varphi i_t+\sum_{j=1}^{3}\psi_i\pi_{t-j}+\epsilon_t, \tag{7.1}$$

where the output gap is proxied by the 5-year moving average gap. This specification is derived from the IS curve equation by substituting the expectations by a linear combination of lags of inflation and the output gap.

8 Additional Discussions

As mentioned in the paper, the response of inflation fell below one since the mid-1990s. Given the fact that inflation remained at low and stable levels in the 1990s and thereafter, this finding may seem surprising. Nevertheless, the discussions of economists at the Fed provide narrative evidence regarding this approach. For instance, Alan Blinder, a former Vice Chairman of the Fed, testified before the Senate committee in 1994 that:

> If monetary policy is used to cut our losses on the inflation front when luck runs against us, and pocket the gains when good fortune runs our way, we can continue to chip away at the already low inflation rate (Blinder, 1994, p.4).

Laurence H. Meyer, a governor of the Federal Reserve Board from 1996 to 2002, also wrote in his book "A Term at the Fed: An Insider's View" about Greenspan's behavior that:

> The staff forecast would suggest the need for a tightening, and the FOMC members would part into hawks and doves. And then the Chairman would come, as always, *speaking like a hawk and walking like a dove*. [Italics added for emphasis.]
>
> This seemed to fall into a pattern: The Chairman would ask for no change in the funds rate suggesting that the time was approaching for action, and indicate that there was a high probability of a move at the next meeting. Then at the next meeting, he would explain that the data did not yet provide a credible basis for tightening, and in any case the markets didn't expect a move. However, he would conclude that he expected the Committee would be forced to move at the next meeting (Meyer, 2004, p.83).

Another important finding is about the change in the Fed's preference before and after 1979. The question that arises is what made the Fed change behavior. According to De Long (1997), the Fed in the 1970s did not have enough autonomy to control inflation. The author provides extensive narrative evidences about the influence of Nixon's administration on the Chairmanship of Burns at the Fed. Among those is the following conversation between Richard Nixon (the speaker) and Arthur Burns (the listener), reported in Ehrlichman (1982), on October 23 1969, after Nixon had announced his intention to nominate Burns to replace Martin as chairman of the Fed:

> I know there's the myth of the autonomous Fed... [short laugh] and when you go up for confirmation some Senator may ask you about your friendship with the President. Appearances are going to be important, so you can call Ehrlichman to get messages to me, and he'll call you (Ehrlichman, 1982, p.248-249 cited in De Long, 1997, p.263).

The Fed was therefore quite sensitive to the concerns of political authorities, who were not willing to accept the possibility of recession to lower inflation given the prevailing view of a permanent negative trade-off between unemployment and inflation at the time.

Disinflation might have been thought to be more costly than inflation, so that it would be better not to reduce inflation at all or to do it very gradually (Taylor, 1997). However, given fear about the possibility of unanchored structure of expectations and the permanent double-digit inflation, fighting inflation by inducing a significant recession actually became the Fed's mandate in 1979 (De Long, 1997). This implies a greater independence of the Fed since then. In addition, Taylor (1992) argues that changes in the perceptions of how the economy works at the early 1980s, which rejects the trade-off view between unemployment and inflation, provided more impetus to curb inflation. For these reasons, the post-1980 monetary policy less likely created inflationary bias than it used to do in the 1970s.

References

Blinder, A. (1994): "Opening Statement of Alan S. Blinder at Confirmation Hearing Before the U.S Senate Committee on Banking, Housing, and Urban Affairs," mimeo, Federal Reserve Board.

De Long, J. (1997): "America's Peacetime Inflation: the 1970s," in C. Romer and D. Romer, eds., *Reducing inflation: Motivation and Strategy*, University of Chicago Press, 247–280.

Doucet, A. and A. Johansen (2009): "A Tutorial on Particle Filtering and Smoothing: Fifteen Years Later," in D. Crisan and B. Rozovskii, eds., *Handbook of Nonlinear Filtering*, Oxford University Press, 656–704.

Ehrlichman, J. (1982): *Witness to Power: The Nixon Years.*, Simon and Schuster.

Gordon, N., D. Salmond, and A. Smith (1993): "Novel Approach to Nonlinear/non-Gaussian Bayesian State Estimation," in *Radar and Signal Processing, IEE Proceedings F*, IET, 107–113.

Meyer, L. H. (2004): *A Term at the Fed: An Insider's View*, Harper Collins.

Orphanides, A. (2001): "Monetary Policy Rules Based on Real-Time Data," *American Economic Review*, 91, 964–985.

Ristic, B., S. Arulampalm, and N. Gordon (2004): *Beyond the Kalman Filter: Particle Filters for Tracking Applications*, Artech House Publishers.

Taylor, J. (1992): "The Great Inflation, the Great Disinflation, and Policies for Future Price Stability," in A. Blundell-Wignall, ed., *Inflation, Disinflation and Monetary Policy*, Sydney Ambassador Press.

Taylor, J. (1997): "Comment on 'America's Peacetime Inflation: the 1970s'," in C. Romer and D. Romer, eds., *Reducing inflation: Motivation and Strategy*, University of Chicago Press, 276–280.