Estimation of a forward-looking monetary policy rule: A time-varying parameter model using ex post data

Chang-Jin Kim\textsuperscript{a,b,*}, Charles R. Nelson\textsuperscript{b}

\textsuperscript{a}Department of Economics, Korea University, Republic of Korea
\textsuperscript{b}Department of Economics, University of Washington, Seattle, WA 98195, USA

Received 7 September 2005; accepted 11 October 2005
Available online 18 October 2006

Abstract

In this paper, we consider estimation of a time-varying parameter model for a forward-looking monetary policy rule, by employing ex post data. A Heckman-type (1976. The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement 5, 475–492) two-step procedure is employed in order to deal with endogeneity in the regressors. This allows us to econometrically take into account changing degrees of uncertainty associated with the Fed’s forecasts of future inflation and GDP gap when estimating the model. Even though such uncertainty does not enter the model directly, we achieve efficiency in estimation by employing the standardized prediction errors for inflation and GDP gap as bias correction terms in the second-step regression. We note that no other empirical literature on monetary policy deals with this important issue. Our empirical results also reveal new aspects not found in the literature previously. That is, the history of the Fed’s conduct of monetary policy since the early 1970s can in general be divided into three subperiods: the
1970s, the 1980s, and the 1990s. The conventional division of the sample into pre-Volcker and Volcker–Greenspan periods could mislead the empirical assessment of monetary policy. © 2006 Elsevier B.V. All rights reserved.

**JEL classification:** E5; C32

**Keywords:** Endogeneity; Forward-looking monetary policy; Heteroscedasticity; Nonlinearity; Time-varying parameter model

1. Introduction

Since the seminal work by Taylor (1993), various versions of backward-looking and forward-looking Taylor rule for the U.S. monetary policy have been estimated by many empirical macroeconomists. Based on subsample analyses, Judd and Rudebusch (1998), Clarida et al. (2000), and Orphanides (2004) show that the Fed’s interest rate policy has changed since 1979. Cogley and Sargent (2001, 2003) and Boivin (2001) report significant time variation in the policy response to the state of the economy, within the framework of the time-varying parameter models. By applying Hamilton’s (1989) Markov-switching models, Sims (2001) and Sims and Zha (2006) argue that time-varying variance of the shocks is more important than time-varying coefficients in modeling the monetary policy rule.

Focusing on estimation of a Taylor-rule type forward-looking monetary policy rule, the literature introduces two alternative approaches depending on the data set employed. One approach, undertaken by Orphanides (2001, 2004), is to use historical real-time forecasts data by the Fed, called “Greenbook data.” If these real-time forecasts are made under the assumption that the nominal federal funds rate will remain constant within the forecasting horizon, there would be no endogeneity problem in the policy rule equation. Thus, the use of real-time forecasts data allows one to straightforwardly extend the basic model to incorporate time-varying coefficients and to employ the conventional Kalman (2006) filter. Such an attempt has recently been made by Boivin (2006). Another approach, undertaken by Clarida et al. (2000), is to use ex post data and explicitly estimate the Fed’s expectation process. An instrumental variables (IV) estimation procedure or a generalized method of moment (GMM) is applied, since the future economic variables used as regressors in the policy rule equation are correlated with the disturbance terms. However, extending the basic model to incorporate time-varying coefficients would not be as straightforward as in Boivin (2006), and no such attempts have been made so far. With the endogeneity problem that results from using the ex post data, a conventional IV estimation procedure or a conventional GMM estimation procedure cannot be readily applied to a time-varying parameter model.

In this paper, we consider estimation of a Taylor-rule type forward-looking monetary policy rule, that allows for time-varying parameters (TVPs) by employing ex post data. In doing so, we apply Kim’s (2006) TVP-with-endogenous-regressors model in at least two

---

1Sims and Zha (2006) argue that “the Taylor rule formalism, valuable as it may be as a way to characterize policy in the last 20 years, can be seriously misleading if we try to use it to interpret other historical periods, where monetary aggregate growth was an important factor in the thinking of policy-makers.” An extension of the model in this paper to include monetary aggregate growth would be worth pursuing.
directions. First, the model is extended to deal with nonlinearities, which results from the Fed’s interest rate smoothing. Second, it deals with heteroscedasticity in the disturbance terms of the monetary policy rule, as emphasized by Sims (2001) and Sims and Zha (2006). The endogeneity problem is solved by employing the Heckman-type (1976) two-step procedure, with bias correction terms in the second step. An important feature of the proposed estimation procedure is that it allows us to econometrically take into account the changing degrees of uncertainty associated with the Fed’s forecasts of future economic conditions. An inflation forecast of 5%, for example, would be associated with much higher uncertainty during the 1970s than during the 1980s or 1990s. Even though such uncertainty does not enter the model directly, we achieve efficiency in estimation by employing the standardized prediction errors for inflation and GDP gap as bias correction terms in the second-step regression.2

As argued by Orphanides (2001), estimating a forward-looking monetary policy rule using ex post data, which were not available at the time the policy was made, may distort the historical conduct of monetary policy. However, the use of real-time data as in Orphanides (2004) or Boivin (2006) also has its drawbacks. For example, if the real-time forecasts are not made under the assumption that the nominal federal funds rate will remain constant within the forecasting horizon, they would induce the endogeneity problem in the monetary policy rule equation. The main focus of this paper is not to discuss advantages or disadvantages of ex post data or real-time data. Rather, this paper focuses on taking care of the endogeneity issue that results from the use of ex post data, within the framework of the time-varying response of the Fed to future economic conditions. Incorporating the changing degree of uncertainty about future economic conditions in the estimation of the monetary policy rule is an additional important issue.

2. Model specification and a two-step MLE procedure

2.1. Model specification

A formal derivation of the empirical forward-looking monetary policy rule starts with the following specification of the Fed’s target interest rates (federal funds rates) as a function of the future expectation of macroeconomic conditions:

\[
    r_t^* = \beta_{0,t}^* + \beta_{1,t}(E_t(\pi_{t+J}) - \pi_t^*) + \beta_{2,t}E_t(g_{t+J}),
\]

where \(r_t^*\) is the target federal funds rate; \(\pi_t^*\) is the target rate for inflation; \(g_{t+J}\) is a measure of average output gap between time \(t\) and \(t+J\); \(\pi_{t+J}\) is the percent change in the price level between time \(t\) and \(t+J\); \(\beta_{0,t}^*\) is the desired nominal rate when both inflation and output are at their target levels; and \(E_t(\cdot)\) refers to the expectation formed conditional on information at time \(t\), when the target interest rate is determined. It is postulated that, each period, the Fed adjusts the federal funds rate to eliminate a fraction \((1 - \theta_t)\) of the gap between its current target level and its past level (interest rate smoothing), according to

\[
    r_t = (1 - \theta_t)r_t^* + \theta_tr_{t-1} + m_t, \quad 0 < \theta_t < 1,
\]

2Within the fixed-coefficients framework, this is equivalent to employing generalized least squares (GLS) in the first-step regression of two-stage least-squares procedure.
where \( m_t \) is a random disturbance term. The model in Eqs. (1) and (2) with fixed coefficients would be comparable to that in Clarida et al. (2000).

Combining (1) and (2), and assuming random walk dynamics for the coefficients, we have the following nonlinear TVP model of monetary policy rule to be estimated:

\[
\begin{align*}
rt &= (1 - \theta_t)(\beta_{0,t} + \beta_{1,t}\pi_{t,J} + \beta_{2,t}g_{t,J}) + \theta_tr_{t-1} + \epsilon_t, \\
\theta_t &= \frac{1}{1 + \exp(-\beta_{3,t})}, \\
\beta_{i,t} &= \beta_{i,t-1} + \epsilon_{it}, \quad \epsilon_{it} \sim i.i.d. N(0, \sigma_{\epsilon_i}^2), \quad i = 0, 1, 2, 3,
\end{align*}
\]

where \( \beta_{0,t} = \beta_{0,t-1} \) and \( \pi_{t,J}^2 \); \( \epsilon_t = (1 - \theta_t)[\beta_{1,t}(\pi_{t,J} - E_t(\pi_{t,J})) + \beta_{2,t}(g_{t,J} - E_t(g_{t,J}))] + m_t \); and the smoothing parameter \( \theta_t \) is constrained between 0 and 1 in Eq. (4). Note that the regressors \( g_{t,J} \) and \( \pi_{t,J} \) in Eq. (3) are correlated with the disturbance term \( \epsilon_t \).

As argued by Sims (2001) and Sims and Zha (2006), we note the importance of the time-varying variance of the shocks in the monetary policy rule. Thus, we approximate the distribution of \( \epsilon_t \) by the following GARCH(1,1) process:

\[
\begin{align*}
\epsilon_t|\psi_{t-1} &\sim N(0, \sigma_{\epsilon_t \mid \psi_{t-1}}), \\
\sigma_{\epsilon_t \mid \psi_{t-1}}^2 &= \alpha_0 + \alpha_1 \epsilon_{t-1}^2 + \alpha_2 \sigma_{\epsilon_{t-1} \mid \psi_{t-1}}^2,
\end{align*}
\]

where \( \psi_{t-1} \) refers to information up to \( t - 1 \).

In order to estimate the above model, we need IV. We assume that the relationships between the endogenous regressors (\( \pi_{t,J} \) and \( g_{t,J} \)) in Eq. (3) and the vector of IV \( z_t \) are given by

\[
\begin{align*}
\pi_{t,J} &= z_{1t}\delta_{1t} + v_{1t}, \quad v_{1t} \sim N(0, \sigma_{v_{1t}}^2),
\end{align*}
\]

\[
\begin{align*}
g_{t,J} &= z_{2t}\delta_{2t} + v_{2t}, \quad v_{2t} \sim N(0, \sigma_{v_{2t}}^2)
\end{align*}
\]

with

\[
\begin{align*}
\delta_{it} &= \delta_{i,t-1} + u_{it}, \quad u_{it} \sim i.i.d. N(0, \Sigma_{u_{it}}), \quad i = 1, 2, \\
\sigma_{v_{ij}^2}^2 &= a_{0j} + a_{1j}v_{ij,t-1}^2 + a_{2j}\sigma_{v_{ij,t-1}^2}, \quad j = 1, 2.
\end{align*}
\]

As specified above, note that the relationship between the regressors in Eq. (3) and the vector of IV \( z_t \) could be time-varying. Furthermore, shocks to Eqs. (8) and (9) could also be heteroscedastic. The specification in Eqs. (8)–(11) suggests that uncertainty associated with future inflation and output gap could be time-varying over time. With time variation in \( \delta_{it} (i = 1, 2) \), this is true even in the absence of heteroscedasticity in \( v_{1t} \) or \( v_{2t} \).

In the next section, we derive a general approach to dealing with the endogeneity problem in Eq. (3). The importance of taking into account the changing nature of uncertainty associated with future inflation and output gap will also be discussed in handling the endogeneity problem.

\(^3\)Refer to Kim and Nelson (1989) for details.
2.2. Derivation of a Heckman-type (1976) two-step MLE procedure

Given the full specification of the model, and by setting \( J = 1 \), a two-step procedure for the empirical estimation of the model is derived in this section. For this purpose, we decomposed \( g_{t,1} \) and \( \pi_{t,1} \) into two components, i.e., predicted components and prediction error components:

\[
\begin{bmatrix}
\pi_{t,1} \\
g_{t,1}
\end{bmatrix}
= \mathbf{E} \begin{bmatrix}
\pi_{t,1} \\
g_{t,1}
\end{bmatrix} | \psi_{t-1}^{-1}
+ \begin{bmatrix}
v_{1,t|t-1} \\
v_{2,t|t-1}
\end{bmatrix},
\]

(12)

\[
\begin{bmatrix}
v_{1,t|t-1} \\
v_{2,t|t-1}
\end{bmatrix}
= \Omega_{t|t-1}^{1/2} \begin{bmatrix}
v^{*}_{1,t} \\
v^{*}_{2,t}
\end{bmatrix},
\]

(13)

where \( \psi_{t-1} \) is information up to time \( t - 1 \); \( \Omega_{t|t-1} \) is the time-varying conditional variance covariance matrix for a \( 2 \times 1 \) vector of prediction errors, \( v_{t|t-1} = [v_{1,t|t-1} \ v_{2,t|t-1}]' \), and both \( \Omega_{t|t-1} \) and \( \psi_{t-1} \) are obtained from the Kalman filter applied to model given by (8)–(11).

We set a vector of \( 2 \times 1 \) standardized prediction errors \( v^{*}_{i} = [v^{*}_{1,i} \ v^{*}_{2,i}]' \), and without loss of generality, we assume the following covariance structure between \( v^{*}_{i} \) and \( e_{i} \):

\[
\begin{bmatrix}
v^{*}_{i} \\
e_{i}
\end{bmatrix}
\sim \mathcal{N}\left(\begin{bmatrix}
0 \\
0
\end{bmatrix}, \begin{bmatrix}
I_{2} & \rho \sigma_{e,c,i}

\rho' \sigma_{e,c,i} & \sigma^{2}_{e,c,i}
\end{bmatrix}\right),
\]

(14)

where \( \rho = [\rho_{1} \ \rho_{2}]' \) is a constant \( 2 \times 1 \) correlation vector. As in Kim (2006), the Cholesky decomposition of the covariance matrix in Eq. (14) results in the following representation of Eq. (14):

\[
\begin{bmatrix}
v^{*}_{i} \\
e_{i}
\end{bmatrix}
= \begin{bmatrix}
I_{2} \\
\rho' \sigma_{e,c,i}
\end{bmatrix} \sqrt{(1 - \rho' \rho) \sigma_{e,c,i}} \begin{bmatrix}
e_{i} \\
\omega_{i}
\end{bmatrix},
\]

(15)

where \( \rho \) is a \( 2 \times 1 \) vector of zeros. Then, Eq. (15) allows us to rewrite \( e_{i} \) as follows:

\[
e_{i} = \rho_{1} \sigma_{e,c,i} v^{*}_{1,i} + \rho_{2} \sigma_{e,c,i} v^{*}_{2,i} + \omega^{*}_{i}, \quad \omega^{*}_{i} \sim \mathcal{N}(0, (1 - \rho_{1}^{2} - \rho_{2}^{2}) \sigma^{2}_{e,c,i})
\]

(16)

where \( \omega^{*}_{i} \) is uncorrelated with either \( v^{*}_{1,i} \) or \( v^{*}_{2,i} \). That is, the role of Eq. (16) is to decompose \( e_{i} \) in Eq. (3) into two components: the components \( (v^{*}_{1,i} \text{ or } v^{*}_{2,i}) \) which are correlated with \( \pi_{t,1} \) and \( g_{t,1} \), and the component \( (\omega^{*}_{i}) \) which are uncorrelated with them. Substituting Eq. (16) into Eq. (3), we get

\[
r_{i} = (1 - \theta_{i})(\beta_{0,i} + \beta_{1,i} \pi_{t,1} + \beta_{2,i} g_{t,1}) + \theta_{i} r_{t-1} + \rho_{1} \sigma_{e,c,i} v^{*}_{1,i} + \rho_{2} \sigma_{e,c,i} v^{*}_{2,i} + \omega^{*}_{i}.
\]

(3’)

In the transformed model in Eq. (3’), as the new disturbance term \( \omega^{*}_{i} \) is uncorrelated with \( \pi_{t,1}, \ g_{t,1}, \ v^{*}_{1,i}, \text{ or } v^{*}_{2,i} \), the following two-step MLE follows:

**Step 1**: Estimate Eqs. (8) and (9) via the maximum likelihood estimation procedure based on Harvey et al.’s (1992) modified Kalman filter, and obtain “standardized” one-step-ahead forecast errors \( (J = 1) \), \( \hat{v}^{*}_{1,t|t-1} \) and \( \hat{v}^{*}_{2,t|t-1} \).

---

^In order to consider heteroscedasticity in the disturbance terms in Eqs. (8) and (9), we adopt Harvey et al.’s (1992) approach.
Step 2: Using a maximum likelihood method via the Kalman filter, estimate the following equation along with Eqs. (4)–(7):

\[ r_t = (1 - \theta_t)(\beta_{0,t} + \beta_{1,t}x_{t,1} + \beta_{2,t}g_{t,1}) + \theta_t r_{t-1} + \rho_1 \sigma_{e,t}\hat{v}_{1,t}^* + \rho_2 \sigma_{e,t}\hat{v}_{2,t}^* + \omega_t^*, \quad (3') \]

\[ \omega_t^* \sim N(0, (1 - \rho_1^2 - \rho_2^2)\sigma_{e,t}^2). \]

The “standardized” prediction errors \( \hat{v}_{1,t}^* \) and \( \hat{v}_{2,t}^* \) enter in Eq. (3”) as bias correction terms in the spirit of Heckman’s (1976) two-step procedure for a sample selection model. That we use standardized prediction errors for inflation and GDP gap as bias correction terms suggests that the changing degrees of uncertainty associated with future inflation and GDP gap are taken into account in the estimation of a forward-looking monetary policy rule.

One might be tempted to employ the following alternative two step MLE procedure:

Step 1”: Estimate Eqs. (8) and (9) via the maximum likelihood estimation procedure based on Harvey et al.’s (1992) modified Kalman filter, and obtain predicted components of the regressors in Eq. (3), \( E(\pi_{t,1}|\psi_{t-1}) \) and \( E(g_{t,1}|\psi_{t-1}) \).

Step 2”: Using a maximum likelihood method via the Kalman filter, estimate the following equation along with Eqs. (4)–(7):

\[ r_t = (1 - \theta_t)(\beta_{0,t} + \beta_{1,t}E(\pi_{t,1}|\psi_{t-1})) + \beta_{2,t}E(g_{t,1}|\psi_{t-1}) + \theta_t r_{t-1} + e_t. \quad (17) \]

One problem with the above two-step procedure is that it fails to take into account the changing degrees of uncertainty associated with future inflation and GDP gap. A more serious problem with this approach is that, as we are using generated regressors, the standard errors of the time-varying coefficients should be corrected at each iteration of the Kalman filter.

Given the general framework for correcting the endogeneity problem in the presence of time-varying coefficients in this section, the next section discusses procedures for explicitly handling the issues of nonlinearity and heteroscedasticity in the disturbance terms involved in Eq. (3).

3. Estimation of the model: dealing with nonlinearity and heteroscedasticity in the monetary policy rule

3.1. Dealing with nonlinearity: a linear approximation

Rewriting Eqs. (3”) and (4)–(7), we have

\[ r_t = f(x_t; \beta_t) + \rho_1 \sigma_{e,t}\hat{v}_{1,t}^* + \rho_2 \sigma_{e,t}\hat{v}_{2,t}^* + \omega_t^*, \quad \omega_t^* \sim N(0, (1 - \rho_1^2 - \rho_2^2)\sigma_{e,t}^2), \quad (18) \]

\[ \beta_t = \beta_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim i.i.d. N(0, \Sigma_\varepsilon), \quad (19) \]

\[ \sigma_{e,t}^2 = \sigma_0 + \sigma_1 e_{t-1}^2 + \sigma_2 \sigma_{e,t-1}^2, \quad (20) \]
where $\beta_t = [\beta_{0,t} \beta_{1,t} \beta_{2,t} \beta_{3,t}]'$; $\Sigma_e$ is a diagonal matrix; and $e_t = \rho_1 \sigma_{e,t} v_{1t}^* + \rho_2 \sigma_{e,t} v_{2t}^* + \omega_t^*$; and

$$f(x_t; \beta_t) = \left( 1 - \frac{1}{1 + \exp(-\beta_{3,t})} \right) (\beta_{0,t} + \beta_{1,t} \pi_{t,1} + \beta_{2,t} g_{t,1}) + \frac{1}{1 + \exp(-\beta_{3,t})} r_{t-1}$$

(21)

and where $X_t^i = [1 \ g_{t,1} \ \pi_{t,1} \ r_{t-1}]$ and $e_t$ and $\omega_t^*$ are independent.

As in Harvey (1989), we first consider a linear approximation to Eq. (18) obtained by taking a Taylor series expansion of the nonlinear function $f(x_t; \beta_t)$ around $\beta_t = \beta_{tt-1}$, where $\beta_{tt-1} = E(\beta_t | \psi_{t-1})$ 5:

$$r_t = f(x_t; \beta_{tt-1}) + \frac{\partial f(x_t; \beta_{tt-1})}{\partial \beta_t} (\beta_t - \beta_{tt-1}) + \rho_1 \sigma_{e,t} \hat{\nu}_{1t}^* + \rho_2 \sigma_{e,t} \hat{\nu}_{2t}^* + \omega_t^*.$$ 

(22)

Evaluating the first derivatives in Eq. (22) and rearranging terms, we have the following linear approximation to Eq. (30):

$$Y_t = X_t' \beta_t + \rho_1 \sigma_{e,t} \hat{\nu}_{1t}^* + \rho_2 \sigma_{e,t} \hat{\nu}_{2t}^* + \omega_t^*,$$

(23)

where

$$Y_t = \begin{bmatrix} r_{t-1} \\ \frac{r_{t-1} - \beta_{0,t|t-1} - \beta_{1,t|t-1} \pi_{t,1} - \beta_{2,t|t-1} g_{t,1}}{1 + \exp(-\beta_{3,t|t-1})} \end{bmatrix}$$

(24)

$$X_t = \begin{bmatrix} 1 - \frac{1}{1 + \exp(-\beta_{3,t|t-1})} \\ \frac{\pi_{t,1}}{1 + \exp(-\beta_{3,t|t-1})} \\ \frac{g_{t,1}}{1 + \exp(-\beta_{3,t|t-1})} \\ \frac{(r_{t-1} - \beta_{0,t|t-1} - \beta_{1,t|t-1} \pi_{t,1} - \beta_{2,t|t-1} g_{t,1}) \exp(-\beta_{3,t|t-1})}{(1 + \exp(-\beta_{3,t|t-1}))^2} \end{bmatrix}.$$ 

(25)

3.2. Dealing with heteroscedasticity and the Kalman filter

In order to deal with the problem of heteroscedasticity in the disturbance term, we follow Harvey et al. (1992) and include the $\omega_t^*$ term in the state equation of the following state-space representation of Eqs. (23) and (19):

$$Y_t = [X_t' \ 1] \begin{bmatrix} \beta_t \\ \omega_t^* \end{bmatrix} + \rho_1 \sigma_{e,t} \hat{\nu}_{1t}^* + \rho_2 \sigma_{e,t} \hat{\nu}_{2t}^*$$

$$\Rightarrow Y_t = X_t^i \hat{\beta}_t + \rho_1 \sigma_{e,t} \hat{\nu}_{1t}^* + \rho_2 \sigma_{e,t} \hat{\nu}_{2t}^*,$$

(26)

---

5This procedure would work well in case innovations to $\beta_t$ are small so that $\beta_t$ is well predicted one step in advance. As an alternative, one could consider nonlinear filtering via the Bayesian MCMC, as in Cogley and Sargent (2005).
\[
\begin{bmatrix}
\beta_t \\
\omega_t^*
\end{bmatrix} = \begin{bmatrix}
1 & 0 \\
0 & 0
\end{bmatrix}\begin{bmatrix}
\beta_{t-1} \\
\omega_{t-1}^*
\end{bmatrix} + \begin{bmatrix}
\varepsilon_t \\
\omega_t^*
\end{bmatrix},
\begin{bmatrix}
\varepsilon_t \\
\omega_t^*
\end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix}
0 \\
0
\end{bmatrix}, \begin{bmatrix}
\Sigma_e & 0 \\
0 & (1 - \rho_1^2 - \rho_2^2)\sigma^2_{e, t}
\end{bmatrix}\right)
\]

\[\Rightarrow \hat{\beta}_t = F\hat{\beta}_{t-1} + \tilde{v}_t, \quad \tilde{v}_t \sim (0, \tilde{Q}_t). \quad (27)\]

At each iteration of the Kalman filter, we obtain a linear approximation of the model around \( \beta_t = \beta_{t|t-1} \), and calculate \( Y_t \) and \( X_t \). Then, the following Kalman filter proceeds:

\[\hat{\beta}_{t|t-1} = F\hat{\beta}_{t-1|t-1}, \quad (28)\]
\[P_{t|t-1} = FP_{t-1|t-1}F' + Q_t, \quad (29)\]
\[\eta_{t|t-1} = Y_t - \hat{X}_t\hat{\beta}_{t|t-1} - \rho_1\sigma_{e, t}\hat{v}_{1,t}^* - \rho_2\sigma_{e, t}\hat{v}_{2,t}^*, \quad (30)\]
\[H_{t|t-1} = \hat{X}_tP_{t|t-1}\hat{X}_t', \quad (31)\]
\[\hat{\beta}_{t|t} = \hat{\beta}_{t|t-1} + P_{t|t-1}\hat{X}_tH_{t|t-1}^{-1}\eta_{t|t-1}, \quad (32)\]
\[P_{t|t} = P_{t|t-1} - P_{t|t-1}\hat{X}_tH_{t|t-1}^{-1}\hat{X}_tP_{t|t-1}. \quad (33)\]

In order to process the above Kalman filter, we need the \( \sigma^2_{e, t} \) term in order to calculate \( \sigma^2_{\psi_{t-1}} = \sigma_0^2 + \sigma_1e^2_{t-1} + \sigma_2\sigma^2_{e, t-1} \) to be employed in the \( \tilde{Q}_t \) matrix of Eq. (29) as well as in Eq. (30). As in Harvey et al. (1992), the term \( \sigma^2_{e, t} \) is approximated by \( E(e^2_{t-1} | \psi_{t-1}) \), where \( \psi_{t-1} \) is information up to time \( t - 1 \). Since we know

\[e_{t-1} = \hat{v}_{1,t-1}^*\rho_1\sigma_{e, t-1} + \hat{v}_{2,t-1}^*\rho_2\sigma_{e, t-1} + \omega_{t-1}^* \]

and

\[e_{t-1} = E[e_{t-1} | \psi_{t-1}] + (e_{t-1} - E[e_{t-1} | \psi_{t-1}]), \quad (35)\]

we have

\[
E(e^2_{t-1} | \psi_{t-1}) = E(e^2_{t-1} | \psi_{t-1})^2 + E((e_{t-1} - E(e_{t-1} | \psi_{t-1}))^2)
= (v_{1,t-1}^*\rho_1\sigma_{e, t-1} + v_{2,t-1}^*\rho_2\sigma_{e, t-1} + E(\omega_{t-1}^* | \psi_{t-1}))^2 + E((\omega_{t-1}^* - E(\omega_{t-1}^* | \psi_{t-1}))^2).
\]  

(36)

Here, \( E(\omega_{t-1}^* | \psi_{t-1}) \) is obtained from the last element of \( \hat{\beta}_{t|t-1} \) and its mean squared error \( E((\omega_{t-1}^* - E(\omega_{t-1}^* | \psi_{t-1}))^2) \) is given by the last diagonal element of \( P_{t-1|t-1} \).

Even though the above Kalman filter provides correct inferences on \( \beta_t \), the \( P_{t|t-1} \) and \( P_{t|t} \) terms are incorrect measures of the conditional variances of \( \beta_t \). In order to correct for the endogeneity bias, inferences on \( \beta_t \) should be made conditional on bias correction terms \( v_{1,t}^* \) and \( v_{2,t}^* \). Eq. (31) provides us with the variance of \( Y_t \) conditional on past information and on these bias correction terms. Thus, \( P_{t|t} \), for example, provides us with variances of \( \beta_t \) conditional on information up to time \( t \) and on the bias correction terms. However, the correct conditional variance of \( Y_t \), and thus the correct conditional variance of \( \beta_t \), should not be made conditional on the bias correction terms. We augment the above Kalman filter with the following equations for correct inferences of the conditional...
variances of $\beta_i$:

$$H^e_{t-1} = \tilde{X}_t P_{t-1} \tilde{X}_t + \rho^2 \sigma^2_{\epsilon,t} + \rho^2 \sigma^2_{\epsilon,t^*},$$  \hspace{1cm} (37)

$$P^e_{t|t} = P_{t|t-1} - H^e_{t-1} \tilde{X}_t P_{t|t-1},$$  \hspace{1cm} (38)

$$P^e_{t+1|t} = FP^e_{t|t} F^t + \hat{Q}_{t+1}.$$  \hspace{1cm} (39)

4. Empirical results

The data we employ are quarterly data covering the period 1960.I–2001:II.6 As in Clarida et al. (2000), the interest rate is the average federal funds rate in the first-month of each quarter; inflation is measured by the % change of the GDP deflator; the output gap is the series constructed by the Congressional Budget Office. The IV include four lags of each of the following variables: the federal funds rate, output gap, inflation, commodity price inflation, and M2 growth.

Table 1 reports estimates of the hyper-parameters of the model given by Eqs. (3)–(7), with a focus on Eq. (39), both with and without bias correction terms.7 While the coefficient estimate for the bias correction term for inflation, $\rho_1$, is negative and statistically significant, the coefficient estimate for the bias correction term for GDP gap, $\rho_2$, is not statistically significant. However, these two coefficients are jointly significant at a 1% significance level, with the likelihood ratio test statistic of 13.13. Thus, ignoring endogeneity in the regressors of the forward-looking policy rule in Eq. (3) would result in serious bias in the estimation of the time-varying coefficients. When the likelihood ratio test statistic is calculated for the null hypothesis of constant regression coefficients, it is 31.30 and the null hypothesis is rejected at a 1% significance level. Considering that the likelihood ratio test statistic for the null of constant coefficients is a very conservative one, we interpret this result as a very strong evidence in favor of time-varying reaction of the monetary policy to future macroeconomic variables. Conditional on the hyper-parameter estimates of the model, the time-varying coefficients of interest are calculated via the proposed Kalman filter. These are depicted in Figs. 1–4, along with their 90% confidence bands. Of our particular interest would be the responses of the federal funds rate to expected future inflation and expected future GDP gap in Figs. 2 and 3.

In Fig. 2, we observe that the Fed’s response to inflation during the 1970s was the lowest throughout the whole sample. During this period, this coefficient is not statistically different from 1 at the 90% significance level. Note that this result is in contrast to that in Clarida et al. (2000), who suggest that the great inflation during the 1970s is a result of the Fed’s policy that accommodated inflation. Fig. 2 shows that the Fed just did not pay enough attention to inflation, rather than accommodating inflation by lowering the real interest rate with a

---

6Even though our sample starts at the first quarter of 1960, the log likelihood function in the second-step MLE procedure was evaluated starting from the first quarter of 1970. This is because the first 20 observations were used to obtain initial values of the coefficients in the first-step MLE procedure, and the next 20 observations were used to obtain initial values of the coefficients in the second-step MLE procedure.

7In Eq. (5), the innovations to the policy rule parameters are assumed to be orthogonal. We also estimated the model by relaxing this assumption. However, the likelihood ratio tests statistic for the null hypothesis of orthogonal innovations was not significant at a 5% significance level. We thus focus our discussion on the case of orthogonal innovations.
rise in inflation. What we observe in Fig. 2 is also in contrast to Orphanides (2004), who suggests that the Fed’s policy response to expected inflation was broadly similar before and after 1979. In the early 1980s, however, the Fed’s response to expected inflation increased sharply and stayed at the high level throughout the entire 1980s and 1990s. In particular, this response is statistically greater than 1 during the 1980s, suggesting that the Fed increased the real interest rate with an increase in inflation. In the 1990s, confidence bands are wider than in the 1980s, and the Fed’s response to inflation is no longer significantly different from 1, even though the point estimate is as high as in the 1980s. We do not interpret this as the Fed having paid less attention to the inflation rate during the 1990s. As the volatility of inflation has decreased since the early 1980s, the federal funds rate carry less information about the response to inflation, resulting in wider confidence bands.

Table 1
Estimation of the hyper-parameters for a forward-looking monetary policy rule

\[
\begin{align*}
\rho_1 & = \rho_1, \\
\rho_2 & = \rho_2, \\
\rho_3 & = \rho_3,
\end{align*}
\]

Parameters Estimates (SE)

<table>
<thead>
<tr>
<th>Parameters</th>
<th>With bias correction terms</th>
<th>Without bias correction terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_{\epsilon,0}$</td>
<td>0.2812 (0.1777)</td>
<td>0.1426 (0.1973)</td>
</tr>
<tr>
<td>$\sigma_{\epsilon,1}$</td>
<td>0.0431 (0.0568)</td>
<td>0.0353 (0.1144)</td>
</tr>
<tr>
<td>$\sigma_{\epsilon,2}$</td>
<td>0.0763 (0.0535)</td>
<td>0.0760 (0.0835)</td>
</tr>
<tr>
<td>$\sigma_{\epsilon,3}$</td>
<td>0.1289 (0.0646)</td>
<td>0.2189 (0.2921)</td>
</tr>
<tr>
<td>$z_0$</td>
<td>0.0676 (0.0513)</td>
<td>0.0755 (0.0473)</td>
</tr>
<tr>
<td>$z_1$</td>
<td>0.4099 (0.2056)</td>
<td>0.4389 (0.1477)</td>
</tr>
<tr>
<td>$z_2$</td>
<td>0.5574 (0.2089)</td>
<td>0.5536 (0.1512)</td>
</tr>
<tr>
<td>$\rho_1$</td>
<td>-0.3260 (0.0849)</td>
<td>–</td>
</tr>
<tr>
<td>$\rho_2$</td>
<td>-0.1144 (0.0891)</td>
<td>–</td>
</tr>
<tr>
<td>ln((L))</td>
<td>-153.7663</td>
<td>-160.0294</td>
</tr>
</tbody>
</table>

\footnote{For simplicity, consider the following simple regression equation: \( Y_t = \beta_0 + \beta_1 X_t + e_t, e_t \sim i.i.d. \mathcal{N}(0, \sigma_e^2) \). The variance of \( \hat{\beta} \) is given by \( \text{Var}(\hat{\beta}) = \sigma^2 / \Sigma(X, - \bar{X})^2 \). Note that, as the volatility of the explanatory variable \( X \) decreases, the variance of \( \hat{\beta} \) would increase.}
throughout the entire 1980s this response is not statistically different from zero. Combining this result with that for response to inflation, we can conclude that the 1980s was a period during which the Fed paid closer attention to expected inflation than to
real economic activities, leading to the stabilization of inflation. Once inflation has been stabilized at a lower level, the Fed could have more room for actively reacting to real economic conditions such as the real GDP gap. Thus, since the early 1990s, the Fed’s
response to real GDP gap turns out to be positive and statistically significant. Actually, the Fed’s response to real GDP gap during this period is larger than ever before.

In Fig. 4, the degree of interest rate smoothing is shown. It has been continuously increasing since the mid-1970s. Finally, in Fig. 5, the prediction errors and their conditional variances for the federal funds rate are plotted. From these plots, we can reassure the importance of incorporating heteroscedasticity in the disturbances of the monetary policy rule, as emphasized by Sims (2001) and Sims and Zha (2006).

In order to see how much the second-stage bias correction terms matter in Eq. (3), we also report the Kalman filter estimates of the time-varying policy responses for a model without bias correction terms. These are depicted in Figs. 6–9, along with their 90% confidence bands. By comparing Figs. 2 and 7, we observe that the point estimates of $\beta_{1t}$ (the response of the federal funds rates to inflation) are almost the same for both models with and without the bias correction terms. The only difference is somewhat narrower confidence bands for the model without bias correction terms. In Fig. 10, we report the probabilities at each date the $\beta_{1t}$ coefficient is greater than 1 for both models. For the model without the bias correction terms, these probabilities remain close to 1 since the early 1980s. For the model with the bias correction terms, however, these probabilities fall below 0.9 since the early 1990s.

By comparing Figs. 3 and 8, we see that the two alternative models result in quite different inferences on the $\beta_{2t}$ coefficient (response of the federal funds rates to GDP gap). For the model without the bias correction terms, the estimates of the $\beta_{2t}$ coefficient are mostly insignificant throughout most of the 1980s and the 1990s, unlike for the model with bias correction terms. In Fig. 11, we report the probabilities that $\beta_{2t}$ is greater than 0 for the two alternative models. The major difference in the results clearly shows up during the
1980s. For the model without bias correction terms, the probabilities remain high throughout the entire period since the mid-1970s. For the model with bias correction terms, however, we see a considerable drop in these probabilities during the 1980s.
Fig. 8. Time-varying response of federal funds rate to expected output gap and 90% confidence bands (model without bias correction terms).

Fig. 9. Time-varying degree of interest rate smoothing and 90% confidence bands (model without bias correction terms).
Fig. 10. Probability that the response of the federal funds rate to expected inflation is greater than 1.

Fig. 11. Probability that the response of the federal funds rate to expected GDP gap is greater than 0.

5. Summary and conclusion

This paper provides efficient estimation of a forward-looking monetary policy rule with the Fed’s time-varying responses to expected future macroeconomic conditions. Unlike
existing literature, we econometrically take into account the changing nature of uncer-
tainty associated with the Fed’s forecasts of future economic conditions, as a byproduct of applying the Heckman-type (1976) two-step procedure in dealing with endogeneity problem in the regressors of the model. Heteroscedasticity in the disturbance terms of the policy rule equation is also explicitly taken into account. Our empirical results also reveal some new aspects not found in the existing literature. Focusing on the response of the Fed to future expected inflation and GDP gap, the whole sample can be divided into three subperiods: the 1970s, the 1980s, and 1990s. Notice that the usual practice is to divide the whole sample into two: pre-Volcker (pre-1979) period and Volcker–Greenspan period (post-1979). However, dividing the sample in this way could mislead the Fed’s historical performance of the monetary policy.

The latter half of the 1970s was the period during which the Fed mainly focused on the stabilization of real economic activity. This policy, combined with the misperception of potential GDP, could have destabilized the economy during the 1970s. During the 1980s, however, the Fed mainly focused on stabilizing inflation. During the 1980s, the probability that the response of the federal funds rate to inflation is greater than one remained close to 1, even though they have somewhat decreased in the 1990s. Furthermore, during the 1980s, the Fed’s response to GDP gap decreased considerably. This policy might have stabilized inflation at a lower level. Once the inflation has been stabilized at a lower level, the Fed could pay more attention to stabilizing real economic activity since the early 1990s. This is the reason why the Fed’s response to GDP gap was higher than ever and statistically different from zero during the most of the 1990s.

One potential drawback in our approach is that the use of ex post data, which were not available at the time of policy making, could distort the empirical results on the historical conduct of the Fed’s policy. Thus, it would be worthwhile to investigate how the results in this paper would change if the approach in this paper were modified and applied to handle the real-time data. We leave this as a future research topic.

References


Heckman, J.J., 1976. The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models. Annals of Economic and Social Measurement 5, 475–492.


