Estimation of a Forward-Looking Monetary Policy Rule:
A Time-Varying Parameter Model using Ex-Post Data

by

Chang-Jin Kim
Korea University and University of Washington

Charles R. Nelson [1]
University of Washington

June, 2004

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) 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 1970's can in general be divided into three sub periods: the 1970's, the 1980's, and 1990's. The conventional division of the sample into pre-Volcker and Volcker-Greenspan periods could mislead the empirical assessment of monetary policy.

---

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 macro economists. Based on subsample analyses, Judd and Rudebusch (1998), Clarida, Gali, and Gertler (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 (2002) 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 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 (1960) filter. Such an attempt has recently been made by Boivin (2004). Another approach, undertaken by Clarida, Gali, and Gertler (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 (2004), 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 a forward-looking monetary policy rule, that allows for time-varying parameters (TVP) by employing ex-post data. In doing so, we apply Kim's (2004) TVP-with-endogenous-regressors model in at least two 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 (2002). 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 1970's than during the 80's or 90's. 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 (2004) 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] Within the fixed-coefficients framework, this is equivalent to employing generalized least squares (GLS) in the first-step regression of two-stage least squares procedure.

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:

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

where r_t^* is the target federal funds rate; π_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 $E_t(.)$ 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:

$$r_t = (1 - \theta_t)r_t^* + \theta_t r_{t-1} + m_t, \quad 0 < \theta_t < 1. \tag{2}$$

where m_t is a random disturbance term. The model in equations (1) and (2) with fixed coefficients would be comparable to that in Clarida, Gali, and Gertler (2000).

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

$$r_t = (1 - \theta_t)(\beta_{0,t} + \beta_{1,t}\pi_{t,J} + \beta_{2,t}g_{t,J}) + \theta_t r_{t-1} + e_t, \tag{3}$$

$$\theta_t = \frac{1}{1 + exp(-\beta_{3,t})}, \tag{4}$$

$$\beta_{i,t} = \beta_{i,t-1} + \epsilon_{it}, \quad \epsilon_{it} \sim i.i.d.N(0, \sigma_{\epsilon,i}^2), \quad i = 0, 1, 2, 3 \tag{5}$$

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

As argued by Sims (2001) and Sims and Zha (2002), we note the importance of the time-varying variance of the shocks in the monetary policy rule. Thus, we approximate the distribution of e_t by the following GARCH(1,1) process:

$$e_t | \psi_{t-1} \sim N(0, \sigma_{e,t}^2) \tag{6}$$

$$\sigma_{e,t}^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \alpha_2 \sigma_{e,t-1}^2, \tag{7}$$

where ψ_{t-1} refers to information up to $t-1$.

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

$$\pi_{t,J} = z_t' \delta_{1t} + v_{1t}, \quad v_{1t} \sim i.i.d. N(0, \sigma_{v1,t}^2) \tag{8}$$

$$g_{t,J} = z_t' \delta_{2t} + v_{2t}, \quad v_{2t} \sim i.i.d. N(0, \sigma_{v2,t}^2) \tag{9}$$

with

$$\delta_{it} = \delta_{i,t-1} + u_{i,t}, \quad u_{it} \sim i.id. N(0, \Sigma_{u,i}), i = 1, 2, \tag{10}$$

$$\sigma_{vj,t}^2 = a_{0j} + a_{1j} v_{j,t-1}^2 + a_{2j} \sigma_{vj,t-1}^2, \quad j = 1, 2. \tag{11}$$

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

In the next section, we derive a general approach to dealing with the endogeneity problem in equation (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.

2.2. Derivation of a Heckman-Type (1976) Two-Step MLE Procedure

---

[3] Refer to Kim and Nelson (1989) for details.

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:

$$\begin{bmatrix} \pi_{t,1} \\ g_{t,1} \end{bmatrix} = E \begin{bmatrix} \pi_{t,1} \\ g_{t,1} \end{bmatrix} \psi_{t-1} + \begin{bmatrix} v_{1,t|t-1} \\ v_{2,t|t-1} \end{bmatrix} \tag{12}$$

$$\begin{bmatrix} v_{1,t|t-1} \\ v_{2,t|t-1} \end{bmatrix} = \Omega_{t|t-1}^{\frac{1}{2}} \begin{bmatrix} v_{1,t}^* \\ v_{2,t}^* \end{bmatrix}, \quad \begin{bmatrix} v_{1,t}^* \\ v_{2,t}^* \end{bmatrix} \sim i.i.d.N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right), \tag{13}$$

where ψ_{t-1} is information up to time $t - 1$; $\Omega_{t|t-1}$ is the time-varying conditional variance covariance matrix for a 2×1 vector of prediction errors, $v_{t|t-1} = \begin{bmatrix} v_{1,t|t-1} & v_{2,t|t-1} \end{bmatrix}'$, and both $\Omega_{t|t-1}$ and $v_{t|t-1}$ are obtained from the Kalman filter applied to model given by (8)-(11). [4]

We set a vector of 2×1 standardized prediction errors $v_t^* = \begin{bmatrix} v_{1,t}^* & v_{2,t}^* \end{bmatrix}'$, and without loss of generality, we assume the following covariance structure between v_t^* and e_t:

$$\begin{bmatrix} v_t^* \\ e_t \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_2 & \rho \sigma_{e,t} \\ \rho' \sigma_{e,t} & \sigma_{e,t}^2 \end{bmatrix} \right), \tag{14}$$

where $\rho = \begin{bmatrix} \rho_1 & \rho_2 \end{bmatrix}'$ is a constant 2×1 correlation vector. As in Kim (2004), the Cholesky decomposition of the covariance matrix in equation (14) results in the following representation of equation (14):

$$\begin{bmatrix} v_t^* \\ e_t \end{bmatrix} = \begin{bmatrix} I_2 & 0_2 \\ \rho' \sigma_{e,t} & \sqrt{(1 - \rho' \rho)} \sigma_{e,t} \end{bmatrix} \begin{bmatrix} \epsilon_t \\ \omega_t \end{bmatrix}, \quad \begin{bmatrix} \epsilon_t \\ \omega_t \end{bmatrix} \sim i.i.d.N \left(\begin{bmatrix} 0_2 \\ 0 \end{bmatrix}, \begin{bmatrix} I_2 & 0_2 \\ 0_2' & 1 \end{bmatrix} \right), \tag{15}$$

where 0_2 is a 2×1 vector of zeros. Then, equation (15) allows us to rewrite e_t as follows:

$$e_t = \rho_1 \sigma_{e,t} v_{1t}^* + \rho_2 \sigma_{e,t} v_{2t}^* + \omega_t, \quad \omega_t \sim N(0, (1 - \rho_1^2 - \rho_2^2) \sigma_{e,t}^2) \tag{16}$$

where ω_t is uncorrelated with either $v_{1,t}^*$ or $v_{2,t}^*$. That is, the role of equation (16) is to decompose e_t in equation (3) into two components: the components ($v_{1,t}^*$ or $v_{2,t}^*$) which are correlated with $\pi_{t,1}$ and $g_{t,1}$ and the component (ω_t) which are uncorrelated with them. Substituting equation (16) into equation (3), we get:

---

[4] In order to consider heteroscedasticity in the disturbance terms in equations (8) and (9), we adopt Harvey et. al's (1992) approach.

$$r_t = (1 - \theta_t)(\beta_{0,t} + \beta_{1,t}\pi_{t,1} + \beta_{2,t}g_{t,1}) + \theta_t r_{t-1} + \rho_1 \sigma_{e,t} v_{1t}^* + \rho_2 \sigma_{e,t} v_{2t}^* + \omega_t \qquad (3')$$

In the transformed model in equation $(3')$, as the new disturbance term ω_t is uncorrelated with $\pi_{t,1}$, $g_{t,1}$, $v_{1,t}^*$, or $v_{1,t}^*$, the following two-step MLE follows:

Step 1: Estimate equations (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 $\hat{v}_{2,t|t-1}^*$.

Step 2: Using a maximum likelihood method via the Kalman Filter, estimate the following equation along with equations (4)-(7):

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

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

The "standardized" prediction errors $\hat{v}_{1,t}^*$ and $\hat{v}_{2,t}^*$ enters in equation $(3'')$ 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 equations (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 equation (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 equations (4)-(7):

$$r_t = (1 - \theta_t)(\beta_{0,t} + \beta_{1,t}E(g_{t,1}|\psi_{t-1}) + \beta_{2,t}E(\pi_{t,1}\psi_{t-1})) + \theta_t r_{t-1} + e_t^*, \qquad (17)$$

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 se-

rious 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. However, this is very difficult, if not impossible, and a failure to correct for the standard errors at each iteration of the Kalman filter creates a serious inference problem, as documented by Kim (2004) based on a Monte Carlo experiment.

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 equation (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 equations (3$''$) and (4)-(7), we have:

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

$$\beta_t = \beta_{t-1} + \epsilon_t, \quad \epsilon_t \sim i.i.d. N(0, \Sigma_\epsilon) \tag{19}$$

$$\sigma_{e,t}^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \alpha_2 \sigma_{e,t-1}^2 \tag{20}$$

where $\beta_t = \begin{bmatrix} \beta_{0,t} & \beta_{1,t} & \beta_{2,t} & \beta_{3,t} \end{bmatrix}'$; Σ_ϵ is a diagonal matrix; and $e_t = \rho_1 \sigma_{e,t} v_{1t}^* + \rho_2 \sigma_{e,t} v_{2t}^* + \omega_t$; and

$$f(x_t; \beta_t) = (1 - \frac{1}{1 + exp(-\beta_{3t})})(\beta_{0,t} + \beta_{1,t} g_{t,1} + \beta_{2,t} \pi_{t,1}) + \frac{1}{1 + exp(-\beta_{3t})} r_{t-1} \tag{21}$$

and where $x_t' = \begin{bmatrix} 1 & g_{t,1} & \pi_{t,1} & r_{t-1} \end{bmatrix}$ and ϵ_t and ω_t are independent.

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

$$r_t = f(x_t; \beta_{t-1|t-1}) + \frac{\partial f(x_t; \beta_{t|t-1})}{\partial \beta_t}(\beta_t - \beta_{t|t-1}) + \rho_1 \sigma_{e,t} \hat{v}_{1t}^* + \rho_2 \sigma_{e,t} \hat{v}_{2t}^* + \omega_t \tag{22}$$

8

Evaluating the first derivatives in equation (22) and rearranging terms, we have the following linear approximation to equation (3''):

$$Y_t = X'_t \beta_t + \rho_1 \sigma_{e,t} \hat{v}^*_{1t} + \rho_2 \sigma_{e,t} \hat{v}^*_{2t} + \omega_t, \tag{23}$$

where

$$Y_t = r_t - \frac{r_{t-1}}{1 + exp(-\beta_{3,t|t-1})}$$
$$+ \frac{(r_{t-1} - \beta_{0,t|t-1} - \beta_{1,t|t-1} g_{t,t+1} - \beta_{2,t|t-1} \pi_{t,t+1}) exp(-\beta_{3,t|t-1}) \beta_{3,t|t-1}}{(1 + exp(-\beta_{3,t|t-1}))^2}, \tag{24}$$

$$X_t = \begin{bmatrix} 1 - \frac{1}{1+exp(-\beta_{3,t|t-1})} \\ g_{t,1} - \frac{g_{t,1}}{1+exp(-\beta_{3,t|t-1})} \\ \pi_{t,1} - \frac{\pi_{t,1}}{1+exp(-\beta_{3,t|t-1})} \\ \frac{(r_{t-1} - \beta_{0,t|t-1} - \beta_{1,t|t-1} g_{t,1} - \beta_{2,t|t-1} \pi_{t,1}) exp(-\beta_{3,t|t-1})}{(1+exp(-\beta_{3,t|t-1}))^2} \end{bmatrix} \tag{25}$$

3.2. Dealing with Heteroscedasticity and an Augmented Kalman Filter

In order to deal with the problem of heteroscedasticity in the disturbance term, we follow Harvey, Ruiz, and Sentena (1992) and include the ω_t term in the state equation of the following state-space representation of equations (23) and (19):

$$Y_t = \begin{bmatrix} X'_t & 1 \end{bmatrix} \begin{bmatrix} \beta_t \\ \omega_t \end{bmatrix} + \rho_1 \sigma_{e,t} \hat{v}^*_{1,t} + \rho_2 \sigma_{e,t} \hat{v}^*_{2,t} \tag{26}$$

$$(==> \quad Y_t = \tilde{X}'_t \tilde{\beta}_t + \rho_1 \sigma_{e,t} \hat{v}^*_{1,t} + \rho_2 \sigma_{e,t} \hat{v}^*_{2,t})$$

$$\begin{bmatrix} \beta_t \\ \omega_t \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \beta_{t-1} \\ \omega_{t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_t \\ \omega_t \end{bmatrix}, \quad \begin{bmatrix} \epsilon_t \\ \omega_t \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_\epsilon & 0 \\ 0 & (1 - \rho_1^2 - \rho_2^2)\sigma_{e,t}^2 \end{bmatrix} \right) \tag{27}$$

$$(==> \quad \tilde{\beta}_t = F \tilde{\beta}_{t-1} + \tilde{V}_t, \quad \tilde{V}_t \sim (0, \tilde{Q}_t))$$

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:

$$\tilde{\beta}_{t|t-1} = F\tilde{\beta}_{t-1|t-1} \tag{28}$$

$$P_{t|t-1} = FP_{t-1|t-1}F' + \tilde{Q}_t \tag{29}$$

$$\eta_{t|t-1} = Y_t - \tilde{X}'_t\tilde{\beta}_{t|t-1} - \rho_1\sigma_{e,t}\hat{v}^*_{1,t} - \rho_2\sigma_{e,t}\hat{v}^*_{2,t} \tag{30}$$

$$H_{t|t-1} = \tilde{X}'_t P_{t|t-1}\tilde{X}_t \tag{31}$$

$$\tilde{\beta}_{t|t} = \tilde{\beta}_{t|t-1} + P_{t|t-1}\tilde{X}_t H^{-1}_{t|t-1}\eta_{t|t-1} \tag{32}$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1}\tilde{X}_t f^{-1}_{t|t-1}\tilde{X}_t P_{t|t-1} \tag{33}$$

In order to process the above Kalman filter, we need the e^2_{t-1} term in order to calculate $\sigma^2_{e,t} = \alpha_0 + \alpha_1 e^2_{t-1} + \alpha_2\sigma^2_{e,t-1}$ to be employed in the $\tilde{Q}_t$ matrix of equations (29) as well as in equations (30). As in Harvey *et al.* (1992), the term e^2_{t-1} is approximated by $E(e^2_{t-1}|\psi_{t-1})$, where ψ_{t-1} is information up to time $t-1$. Since we know

$$e_{t-1} = \hat{v}^*_{1,t-1}\rho_1\sigma_{e,t-1} + \hat{v}^*_{2,t-1}\rho_2\sigma_{e,t-1} + \omega_{t-1} \tag{34}$$

and

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

we have:

$$
\begin{aligned}
E(e^2_{t-1}|\psi_{t-1}) \\
= E(e_{t-1}|\psi_{t-1})^2 + E((e_{t-1} - E(e_{t-1}|\psi_{t-1}))^2) \\
= (v^*_{1,t-1}\rho_1\sigma_{e,t-1} + v^*_{2,t-1}\rho_2\sigma_{e,t-1} + E(\omega_{t-1}|\psi_{t-1}))^2 + E((\omega_{t-1} - E(\omega_{t-1}|\psi_{t-1}))^2).
\end{aligned}
\tag{36}
$$

Here, $E(\omega_{t-1}|\psi_{t-1})$ is obtained from the last element of $\tilde{\beta}_{t-1|t-1}$ 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 β_t, the $P_{t|t-1}$ and $P_{t|t}$ terms are incorrect measures of the conditional variances of β_t. In order to correct for the endogeneity bias, inferences on β_t should be made conditional on bias correction terms $v^*_{1,t}$ and $v^*_{2,t}$. Equation (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

β_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 β_t, should not be made conditional on the bias correction term. As suggested by Kim (2004), we augment the above Kalman filter with the following equations for correct inferences of the conditional variances of β_t:

$$H_{t|t-1}^* = \tilde{X}_t' P_{t|t-1} \tilde{X}_t + \rho_1^2 \sigma_{e,t}^2 + \rho_2^2 \sigma_{e,t}^2 \tag{37}$$

$$P_{t|t}^* = P_{t|t-1} - P_{t|t-1} \tilde{X}_t H_{t|t-1}^{*-1} \tilde{X}_t' P_{t|t-1} \tag{38}$$

$$P_{t+1|t}^* = F P_{t|t}^* F' + \tilde{Q}_{t+1} \tag{39}$$

4. Empirical Results

The data we employ are quarterly data covering the period 1960.I - 2001:II. [5] As in Clarida, Gali, and Gertler (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 instrumental variables include 4 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 equations (3)-(7), with a focus on equation (3″) with bias correction terms. While the coefficient estimate for the bias correction term for inflation, ρ_1, is negative and statistically significant, the coefficient estimate for the bias correction term for GDP gap, ρ_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 equation (3) would result in serious bias in

---

[5] Even 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.

the estimation of the time-varying coefficients. When the likelihood ratio test statistic was calculated for the null hypothesis of constant regression coefficients, it was 31.30 and the null hypothesis is rejected at a 1% significance level. Considering that the likelihood ratio test 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 augmented Kalman filter. These are depicted in Figures 1 through 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 Figures 2 and 3.

In Figure 2, we observe that the Fed's response to inflation during the 1970's was 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, Gali, and Gertler (2000), who suggest that the great inflation during the 1970's is a result of the Fed's policy that accommodated inflation. Figure 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 rise in inflation. What we observe in Figure 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 1980's, however, the Fed's response to expected inflation increased sharply and stayed at the high level throughout the entire 1980's and 1990's. In particular, this response is statistically greater than 1 during the 1980's, suggesting that the Fed increased the real interest rate with an increase in inflation. In the 1990's, confidence bands are wider than in the 1980's, and the Fed's response to inflation is longer significantly differently from 1, even though the point estimate is as high as in the 1980's. We don't interpret this as the Fed having paid less attention to the inflation rate during the 1990's. As the volatility of inflation has decreased since the early 1980's, the federal funds rate carry less information about the response to inflation, resulting in wider confidence bands. [6]

---

[6] For simplicity, consider the following simple regression equation: $Y_t = \beta_0 + \beta_1 X_t + e_t$, $e_t \sim i.i.d.N(0, \sigma^2)$. The variance of $\hat{\beta}$ is given by $Var(\hat{\beta}) = \frac{\sigma^2}{\Sigma(X_t - X)^2}$. Note that, as the

Figure 3 depicts the response of the federal funds rate to expected real GDP gap. Between 1974 and 1979, this response is positive and statistically different from zero. However, throughout the entire 1980's this response is not statistically different from 0. Combining this result with that for response to inflation, we can conclude that the 1980's 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 1990's, 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 Figure 4, the degree of interest rate smoothing is shown. It's been continuously increasing since the mid 1970's. Finally, in Figure 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 (2002).

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 uncertainty 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 engodeneity 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 sub periods: the 1970's, the 1980's, and 1990's. Notice that the usual practice is to divide the whole sample into two: pre-Volcker (pre-1979) period and Volcker-Greenspan period (post-1979).

volatility of the explanatory variable X_t decreases, the variance of $\hat{\beta}$ would increase.

However, dividing the sample in this way could mislead the Fed's historical performance of the monetary policy.

The 1970's 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 1970's. During the 1980's, however, the Fed mainly focused on stabilizing inflation. The response to inflation was significantly larger than one and the response to GDP gap was not significantly different from zero. 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 1990's. 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 1990's.

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. In order to overcome such drawback, Boivin (2004) employed real-time data to estimate the Fed's time-varying policy reaction to future expected macroeconomic conditions. However, unlike our approach in this paper, Bovin (2004) completely ignores the changing nature of uncertainty about the Fed's forecasts of future macroeconomic conditions. Thus, it would be worth while 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

[1] Boivin, Jean, 2001, "The Fed's Conduct of Monetary Policy: Has it Changed and Does it Matter?" Working Paper, Graduate School of Business, Columbia University.

[2] Boivin, Jean, 2004, "Has U.S. Monetary Policy Changed? Evidence from Drifting Coefficients and Real-Time Data," Working Paper, Graduate School of Business, Columbia University.

[3] Clarida, Gali, and Gertler, 2000, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," Quarterly Journal of Economics, 2000, February, 147-180.

[4] Cogley, T. and T. Sargent, 2001, "Evolving Post World War II U.S. Inflation," NBER Macroeconomics Annual, 16.

[5] Cogley, T. and T. Sargent, 2003, "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S., Working Paper, New York University.

[6] Hamilton, James D., 1989, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, Econometrica, Vol. 57, No. 2, 357-384.

[7] Harvey, Andrew C., 1989, *forecasting, Structural Time Series Models, and the Kalman Filter," Cambridge, U.K.: Cambridge University Press.*

[8] Harvey, Andrew C., Esther Ruiz and Enrique Sentana (1992), "Unobserved Component Time Series Models with ARCH Disturbances," *Journal of Econometrics,* 52, 129-157.

[9] Judd and Rudebusch, 1998, "Taylor's Rule and the Fed: 1970:1997," FRB of San Francisco, *Economic Review* 3-16.

[10] 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.

[11] Kim, Chang-Jin, 2004, Time-Varying-Parameter Models with Endogenous Regressors: A Two-Step MLE Approach and an Augmented Kalman Filter," Working paper, Korea

University and University of Washington.

[12] Kim, Chang-Jin and Charles R. Nelson, 1989, "The Time-Varying-Parameter Model for Modeling Changing Conditional Variance: The Case of the Lucas Hypothesis," *Journal of Business and Economic Statistics*, 7, 433–440.

[13] Orphanides, A., 2001, "Monetary Policy Rules Based on Real-Time Data," *American Economic Review,* 91 (4), 964-985.

[14] Orphanides, A., 2004, "Monetary Policy Rules, Macroeconomic Stability and Inflation: A View from the Trenches," 2004, forthcoming, Journal of Money, Credit, and Banking.

[15] Sims, Christopher A. and Tao Zha, 2002, "Macroeconomic Switching," Working paper, Princeton University.

[16] Sims, Christopher A., 2001, "Stability and Instability in US Monetary Policy Behavior," Working Paper, Princeton University.

Table 1. Estimation of the Hyper-parameters for a Forward-Looking Monetary Policy Rule

$$r_t = (1 - \theta_t)(\beta_{0,t} + \beta_{1,t} g_{t,1} + \beta_{2,t} \pi_{t,1}) + \theta_t r_{t-1} + e_t, \quad e_t \sim N(0, \sigma^2_{e,t})$$

$$e_t = \rho_1 \sigma_{e,t} \hat{v}^*_{1,t} + \rho_2 \sigma_{e,t} \hat{v}^*_{2,t} + \omega_t, \quad \omega_t \sim N(0, (1 - \rho_1^2 - \rho_2^2)\sigma^2_{e,t})$$

$$\sigma^2_{e,t} = \alpha_0 + \alpha_1 e^2_{t-1} + \alpha_2 \sigma^2_{e,t-1}$$

$$\theta_t = \frac{1}{1 + exp(-\beta_{3,t})}$$

$$\beta_{i,t} = \beta_{i,t-1} + \epsilon_{it}, \quad \epsilon_{it} \sim i.i.d.N(0, \sigma^2_{\epsilon,i}), \quad i = 0, 1, 2, 3$$

Parameters	Estimates (SE)
$\sigma_{\epsilon,0}$	0.3812 (0.1777)
$\sigma_{\epsilon,1}$	0.0431 (0.0568)
$\sigma_{\epsilon,2}$	0.0763 (0.0535)
$\sigma_{\epsilon,3}$	0.1289 (0.0646)
α_0	0.0676 (0.0513)
α_1	0.4099 (0.2056)
α_2	0.5574 (0.2089)
ρ_1	-0.3260 (0.0849)
ρ_2	-0.1144 (0.0891)

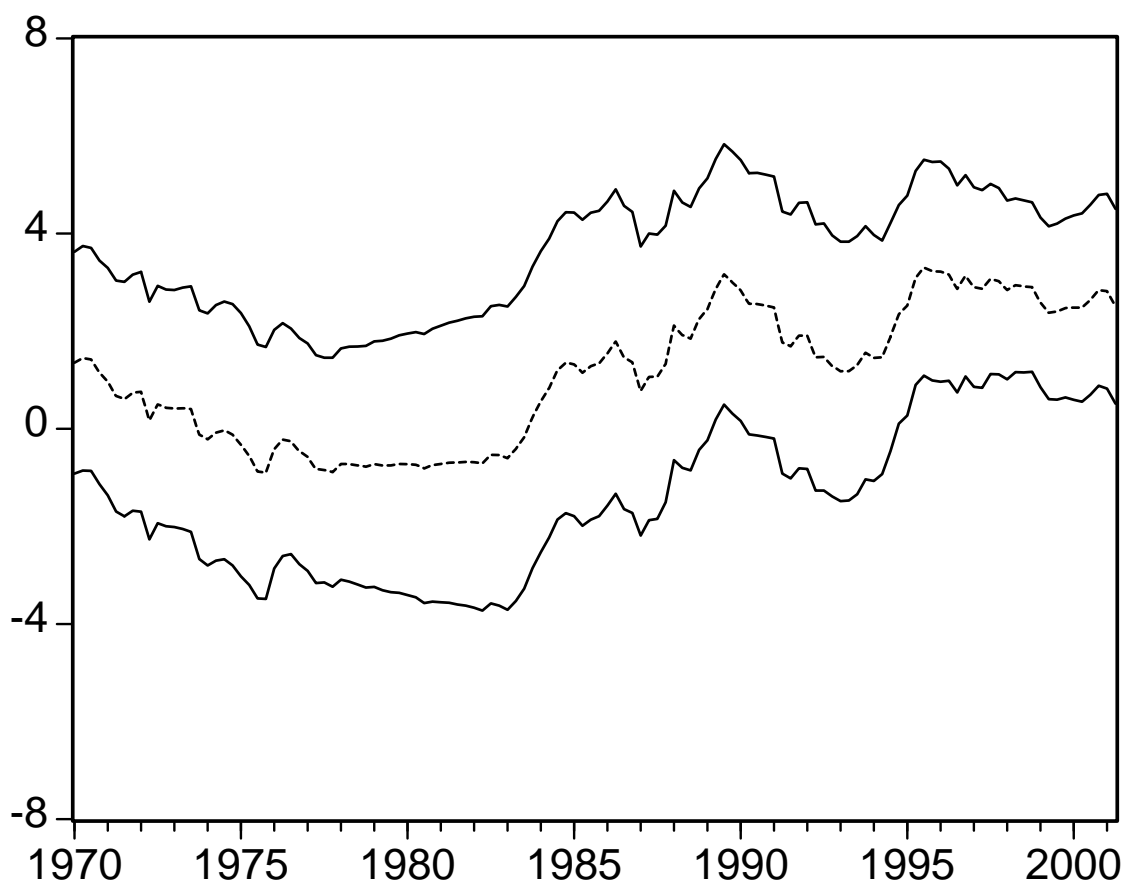Figure 1. Time-Varying Intercept Term and 90% Confidence Bands

Figure 2. Time-Varying Response of Federal Funds Rate
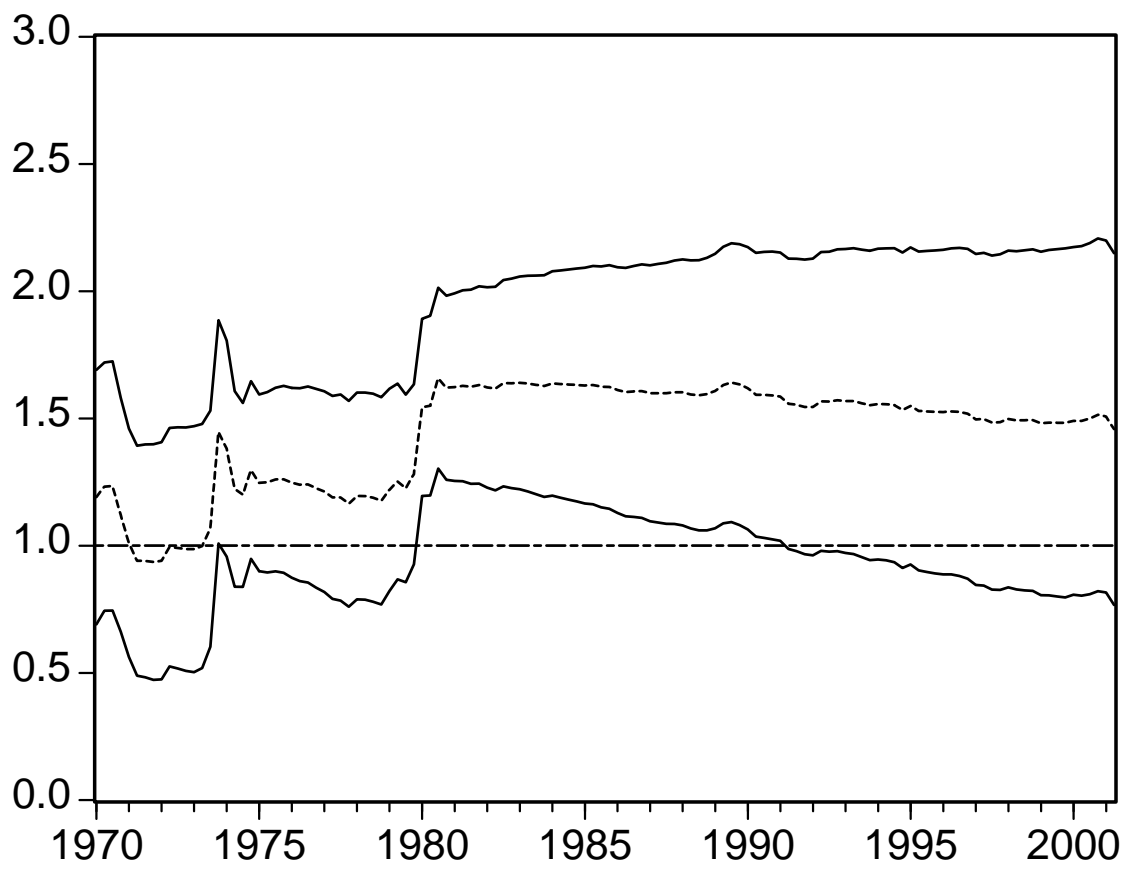to Expected Inflation and 90% Confidence Bands

Figure 3. Time-Varying Response of Federal Funds Rate
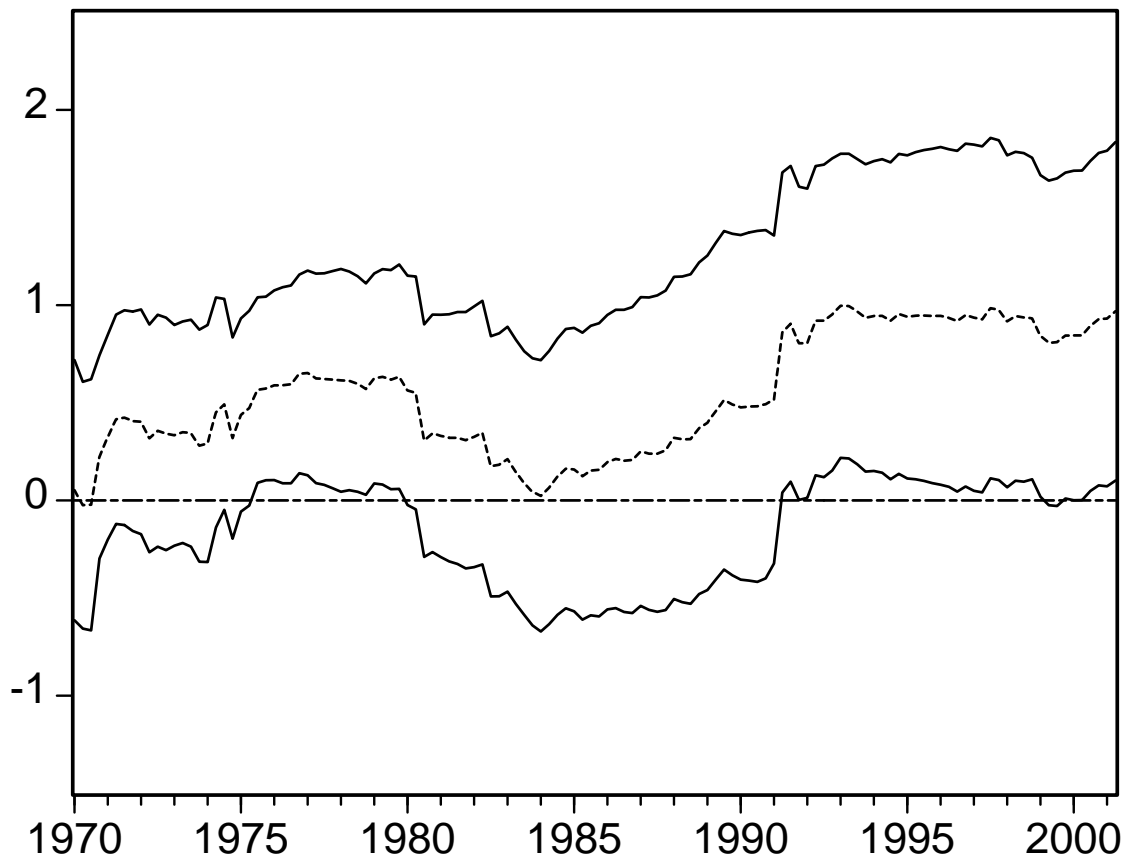to Expected Output Gap and 90% Confidence Bands

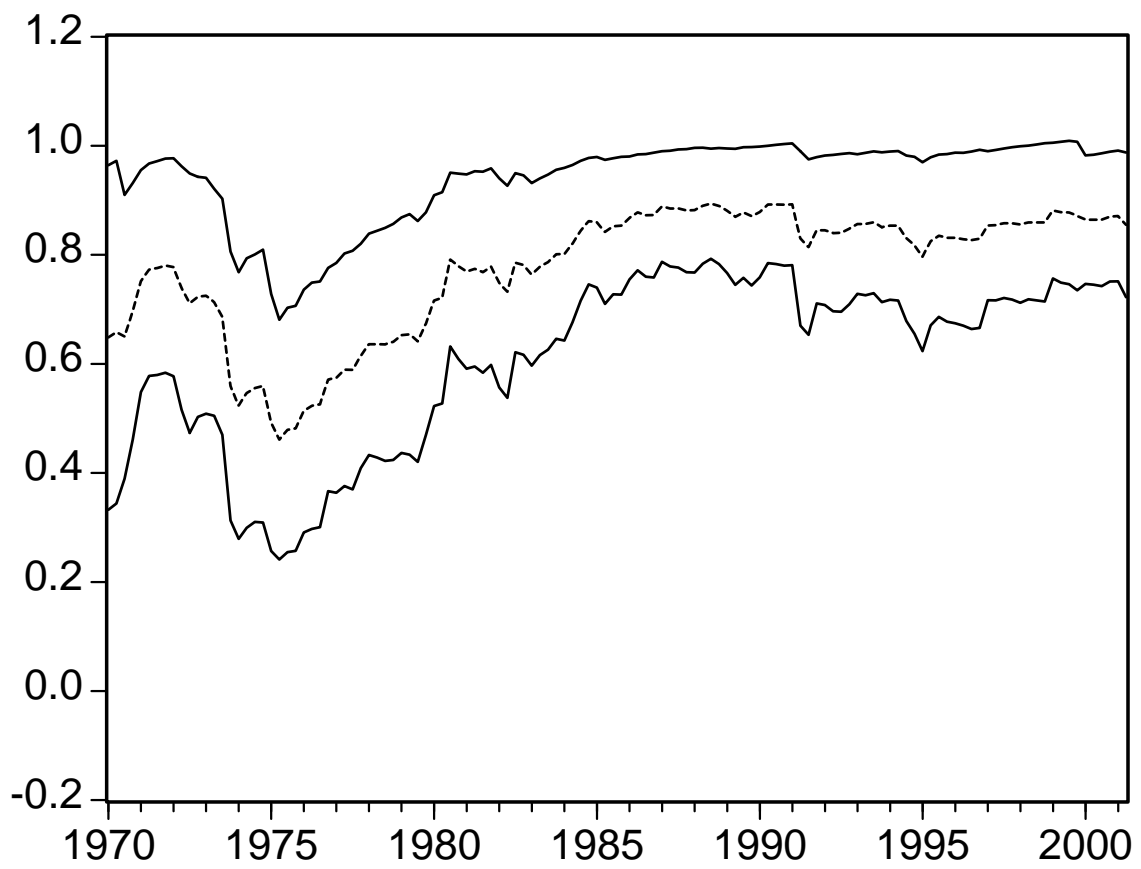Figure 4. Time-Varying Degree of Interest Rate Smoothing and 90 % Confidence Bands

Figure 5. Conditional forecast Errors and Their Variances
for the Federal Funds Rate