CEM

Center for Empirical Macroeconomics

Working Paper No. 46

Monetary Policy Rules under Uncertainty: Empirical Evidence, Adaptive Learning and Robust Control

by

Wenlang Zhang and Willi Semmler

University of Bielefeld
Department of Economics
Center for Empirical Macroeconomics
P.O. Box 100 131

33501 Bielefeld, Germany

Monetary Policy Rules under Uncertainty: Empirical Evidence, Adaptive Learning and Robust Control

Wenlang Zhang[*] Willi Semmler[†]

March 2003

Abstract

Monetary policy faces great challenges because of many kinds of uncertainties such as model uncertainty, data uncertainty and shock uncertainty. This paper explores monetary policy rules under model and shock uncertainties. Facing such uncertainties, a central bank may resort to different strategies, it can either reduce uncertainty by learning or just choose a policy rule robust to uncertainty. Empirical evidence of model and shock uncertainties is explored in a State-Space model with Marked-switching in both linear and nonlinear Phillips curves. The evidence indicates that there has been too great uncertainty in the U.S. economy to define accurately monetary policy rules. Moreover, there seem to have been structural shifts. On the basis of this evidence, we explore monetary policy rules with the recursive least squares (RLS) learning. The simulations of the RLS learning in a framework of optimal control indicate that the state variables do not necessarily converge even in a non-stochastic model, no matter whether the linear or nonlinear Phillips curve is taken as constraints. This is different from the results of papers which discuss the RLS learning without optimal control or in the LQ framework. Finally, we explore monetary policy rules under uncertainty with robust control and find that the robustness parameter affects the economic variables greatly. A larger robustness parameter leads to a smaller variance of the state variable in a stochastic model and faster convergence of the state variable in a non-stochastic model. An evaluation of those two methods is given at the end of the paper.

JEL: E17, E19

Keywords: State-Space Model, Markov-Switching, Nonlinear Phillips Curve, Recursive Least Squares, Robust Control

[*]Center for Empirical Macroeconomics, Bielefeld University, Germany.

[†]Center for Empirical Macroeconomics, Bielefeld, and New School University, New York.

1 Introduction

In the profession it has increasingly been recognized that formal modelling of monetary policy faces great challenges because of many kinds of uncertainties such as model uncertainty, data uncertainty and shock uncertainty. Recent literature dealing with these uncertainties can be found in Isard, Laxton, and Eliasson (1999), Söderström (1999), Giannoni (2000), Meyer, Swanson and Wieland (2001), Wieland (2000), Tetlow and von zur Muehlen (2001a), Orphanides and Williams (2002), Svensson (1999), Martin and Salmon (1999), Hall, Salmon Yates and Batini (1999) and so on. These papers explore, usually theoretically, how a certain kind of uncertainty affects the decisions of the central banks or households and firms. Semmler, Greiner and Zhang (2002a) undertake a time-varying parameter estimation of the Phillips curve and Taylor rule by way of the Kalman filter and observe parameter shifts over time. Cogley and Sargent (2001) study the inflation dynamics of the U.S. after WWII by way of Bayesian Vector Autoregression with time-varying parameters without stochastic volatility. Sims (2001b), however, points out that the monetary policy behavior may not have experienced such a sharp change as shown by Cogley and Sargent (2001). Sims and Zha (2002) also study parameter shifts of the U.S. economy and find more evidence in favor of stable dynamics with unstable variance of the disturbance than of clear changes in model structure. In contrast to Sims (2001b), Cogley and Sargent (2002) study the drifts and volatilities of the U.S. monetary policies after WWII through a Bayesian Vector Autoregression with time-varying parameters and stochastic volatility and claim to have found regime changes.

Facing model and shock uncertainties, economic agents (central banks for example) may resort to different strategies: They can either reduce uncertainty by learning or just choose a policy robust to the model uncertainty without learning. The results of these two strategies may be different. By intuition we would expect all agents to improve their knowledge of the economy with all information available. But recently more and more literature is concerned with bounded-rationality and the assumption of rational expectation is being increasingly doubted. Therefore, in the research below we will consider the situation in which agents improve their knowledge of an economic model through a certain mechanism of learning. Another interesting topic in macroeconomics since the 1990s is the nonlinearity of the Phillips curve. It is argued that positive deviations of aggregate demand from potential output are more inflationary than negative deviations are dis-inflationary. The nonlinearity of the Phillips curve will also be dealt with below and this turns out to be an important difference of our paper from the others.

As stated above, central banks may also resort to a monetary policy rule robust to uncertainty. This is a completely different strategy from adaptive learning. In this approach a central bank considers the economic model only as an approximation to another model that it can not specify. With a so-called robustness parameter it pursues a monetary policy rule in the "worst case" scenario. While the adaptive learning considers mainly parameter uncertainty, the robust control might consider more general uncertainty. In spite of some criticisms, the robust control is given much attention in macroeconomics.

The remainder of this paper is organized as follows. In the second section we present empirical evidence of model and shock uncertainties in the IS and Phillips curves by way of a State-Space model with Markov-Switching. We consider both linear and nonlinear Phillips curves. In Section 3 we explore monetary policy rules under model uncertainty with adaptive learning. Section 4 explores monetary policy rules with the robust control. Section 5 briefly evaluates adaptive learning and robust control and Section 6 concludes the paper.

2 Empirical Evidence of Uncertainty: A State-Space Model with Markov-Switching

Consider an economic model

$$\underset{\{u_t\}_0^\infty}{Min} \; E_0 \sum_{t=0}^{\infty} \rho^t L(x_t, u_t), \tag{1}$$

subject to

$$x_{t+1} = f(x_t, u_t, \varepsilon_t), \tag{2}$$

where ρ is the discount factor bounded between 0 and 1, $L(x_t, u_t)$ a loss function of an economic agent (central bank for instance), x_t a vector of state variables, u_t a vector of control variables, ε_t a vector of shocks and E_0 denotes the mathematical expectation operator upon the initial values of the state variables. This kind of model can be found in many papers on monetary policy, see Svensson (1997 and 1999), Beck and Wieland (2002) and Clarida, Gali and Gertler (1999) for example, where the constraint equations are usually IS-Phillips curves. Given the loss function $L(x, u)$ and the state equation (2), the problem is to derive a path of the control variable to satisfy (1). The question arising is, however, whether the state equation can be correctly specified in reality. The uncertainty of the state equation can be caused by the uncertainty in the shock ε_t and uncertainty in parameters and data. Svensson (1999) and Semmler, Greiner and Zhang (2002b) derive an optimal monetary policy rule from an optimal control problem similar to the model above and find that the optimal monetary policy rule is greatly affected by the estimated parameters of the model. Therefore, if the parameters in the model are uncertain, the optimal monetary policy rule may also be uncertain.

Semmler, Greiner and Zhang (2002a) estimate time varying parameters of the traditional Phillips curve with the following State-Space model:

$$y_t = X_t \beta_t + \varepsilon_t, \;\; \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{3}$$

$$\beta_t = \beta_{t-1} + \eta_t, \;\; \eta_t \sim N(0, \sigma_\eta^2) \tag{4}$$

with β_t being a vector of time-varying parameters. Note that in this model it is assumed that the shocks have constant variance and only β_t is uncertain. Cogley and Sargent (2001) study the inflation dynamics of the U.S. after WWII by way of Bayesian Vector Autoregression with time-varying parameters without stochastic volatility. Sims (2001b), however, claims that the monetary policy behavior may not have experienced such a sharp change as demonstrated by Cogley and Sargent (2001). Sims and Zha (2002) also study the macroeconomic switching of the U.S. and find more evidence in favor of stable dynamics with unstable disturbance variance than of clear changes in model dynamics. Therefore, Cogley and Sargent (2002) modify the model by considering both time-varying parameters and stochastic volatility and claim to have found regime switching. A drawback of the traditional State-Space model such as (3) and (4) is that the changes of the time-varying parameters may be exaggerated, because the shocks are assumed to have constant variance. This is the reason why Cogley and Sargent (2002) assume stochastic volatility. Therefore in the research below we assume that ε_t has state-dependent variance. This is similar to the assumption of Cogley and Sargent (2002). But unlike Cogley and Sargent (2002), who assume the variances of the shocks to change from period to period, we assume that there are only two states of disturbance variance with Markov property. This is to some extent similar to the assumption of Sims and Zha (2002) who assume that there are three states of economy. With such an assumption we can figure out the probability of regime switching. One more advantage of the State-Space model with Markov-switching is that, as we will see afterwards, it can explore not only parameter uncertainty but also shock uncertainty.

Following Kim (1993) and Kim and Nelson (1999), we simply assume that ε_t in (3) has two states of variance with Markov property, namely,

$$\varepsilon_t \sim N(0, \sigma^2_{\varepsilon,S_t}), \tag{5}$$

with

$$\sigma^2_{\varepsilon,S_t} = \sigma^2_{\varepsilon,0} + (\sigma^2_{\varepsilon,1} - \sigma^2_{\varepsilon,0})S_t, \ \ \sigma^2_{\varepsilon,1} > \sigma^2_{\varepsilon,0},$$

and

$$Pr[S_t = 1|S_{t-1} = 1] = p,$$
$$Pr[S_t = 0|S_{t-1} = 0] = q,$$

where $S_t = 0$ or 1 indicates the states of the variance of ε_t and Pr stands for probability. In the research below we explore uncertainty in the IS and Phillips curves, since these two curves form the core of a monetary policy model. We will consider both linear and nonlinear Phillips curves.

2.1 Linear IS-Phillips Curves

In this subsection we will explore uncertainty in the traditional linear IS-Phillips curves which have often been taken as constraints in an optimal control model such as (1) and (2). In order to reduce the dimension of the model, we estimate simple Phillips and IS curves with only one lag of the inflation rate and output gap:

$$\pi_t = \alpha_{1t} + \alpha_{2t}\pi_{t-1} + \alpha_{3t}y_{t-1} + \varepsilon_{\pi,t}, \tag{6}$$
$$y_t = \beta_{1t} + \beta_{2t}y_{t-1} + \beta_{3t}(r_{t-1} - \pi_{t-1}) + \varepsilon_{y,t}, \tag{7}$$

where π_t is the inflation rate, y_t the output gap, r_t the short-term nominal interest rate, and $\varepsilon_{\pi,t}$ and $\varepsilon_{y,t}$ are shocks subject to Gaussian distributions with zero mean and Markov-Switching variances.[1] Let i_t denote the real interest rate, namely, $i_t = r_t - \pi_t$, the model can be rewritten in a State-Space form as follows:

$$Y_t = X_t \phi_t + \varepsilon_t, \tag{8}$$

$$\phi_t = \bar{\Phi}_{S_t} + F\phi_{t-1} + \eta_t, \tag{9}$$

where $\bar{\Phi}_{S_t}$ (S_t=0 or 1) is the drift of ϕ_t and F a diagonal matrix with constant elements to be estimated from the model. η_t has the distribution shown in eq. (4). ε_t is now assumed to have the distribution presented in eq. (5).[2]

Let ψ_{t-1} denote the vector of observations available as of time t$-$1. In the usual derivation of the Kalman filter in a State-Space model without Markov-Switching, the forecast of ϕ_t based on ψ_{t-1} can be denoted by $\phi_{t|t-1}$. Similarly, the matrix denoting the mean squared error of the forecast can be written as

$$P_{t|t-1} = E[(\phi_t - \phi_{t|t-1})(\phi_t - \phi_{t|t-1})'|\psi_{t-1}],$$

where E is the expectation operator.

In the State-Space model with Markov-Switching, the goal is to form a forecast of ϕ_t based not only on ψ_{t-1} but also conditional on the random variable S_t taking on the value j and on S_{t-1} taking on the value i (i and j equal 0 or 1):

$$\phi_{t|t-1}^{(i,j)} = E[\phi_t|\psi_{t-1}, S_t = j, S_{t-1} = i],$$

and correspondingly the mean squared error of the forecast is

$$P_{t|t-1}^{(i,j)} = E[(\phi_t - \phi_{t|t-1})(\phi_t - \phi_{t|t-1})'|\psi_{t-1}, S_t = j, S_{t-1} = i].$$

Conditional on $S_{t-1} = i$ and $S_t = j$ ($i, j = 0, 1$), the Kalman filter algorithm for our model is as follows:

[1] Forward-looking behaviors have been frequently taken into account in the Phillips curve. A survey of this problem can be found in Clarida, Gali and Gertler (1999). Because it is quite difficult to estimate a State-Space model with forward-looking behaviors, we just consider backward-looking behaviors in this section. In fact a justification of the backward-looking model can be found in Rudebusch and Svensson (1999).

[2] Theoretically, the elements of F and the variance of η_t may also have Markov property, but since there are already many parameters to estimate, we just ignore this possibility to improve the efficiency of estimation. Note that if the elements of F are larger than 1 in absolute value, that is, if the time-varying parameters are non-stationary, the transition equation should be altered to be the form of eq. (4).

$$\phi^{(i,j)}_{t|t-1} = \bar{\Phi}_j + F\phi^{i}_{t-1|t-1}, \quad (10)$$

$$P^{(i,j)}_{t|t-1} = FP^{i}_{t-1|t-1}F' + \sigma^2_{\eta}, \quad (11)$$

$$\xi^{(i,j)}_{t|t-1} = Y_t - X_t\phi^{(i,j)}_{t|t-1}, \quad (12)$$

$$\nu^{(i,j)}_{t|t-1} = X_t P^{(i,j)}_{t|t-1} X'_t + \sigma^2_{\varepsilon,j}, \quad (13)$$

$$\phi^{(i,j)}_{t|t} = \phi^{(i,j)}_{t|t-1} + P^{(i,j)}_{t|t-1}X'_t[\nu^{(i,j)}_{t|t-1}]^{-1}\xi^{(i,j)}_{t|t-1}, \quad (14)$$

$$P^{(i,j)}_{t|t} = (I - P^{(i,j)}_{t|t-1}X'_t[\nu^{(i,j)}_{t|t-1}]^{-1}X_t)P^{(i,j)}_{t|t-1}, \quad (15)$$

where $\xi^{(i,j)}_{t|t-1}$ is the conditional forecast error of Y_t based on information up to time t−1 and $\nu^{(i,j)}_{t|t-1}$ is the conditional variance of the forecast error $\xi^{(i,j)}_{t|t-1}$. In order to make the above Kalman filter algorithm operable, Kim and Nelson (1999) develop some approximations and manage to collapse $\phi^{(i,j)}_{t|t}$ and $P^{(i,j)}_{t|t}$ into $\phi^{j}_{t|t}$ and $P^{j}_{t|t}$ respectively.[3]

For the Phillips curve in our model, we have

$$Y_t = \pi_t, \;\; X_t = (1 \;\; \pi_{t-1} \;\; y_{t-1}), \;\; \phi_t = (\alpha_{1t} \;\; \alpha_{2t} \;\; \alpha_{3t})', \;\; \varepsilon_t = \varepsilon_{\pi t},$$

with

$$\varepsilon_{\pi t} \sim N(0, \sigma^2_{\varepsilon\pi,S_t}),$$
$$\sigma^2_{\varepsilon\pi,S_t} = \sigma^2_{\varepsilon\pi,0} + (\sigma^2_{\varepsilon\pi,1} - \sigma^2_{\varepsilon\pi,0})S_t, \;\; \sigma^2_{\varepsilon\pi,1} > \sigma^2_{\varepsilon\pi,0},$$

and

$$\eta_t = (\eta_{\alpha 1t} \;\; \eta_{\alpha 2t} \;\; \eta_{\alpha 3t})',$$
$$\sigma^2_{\eta} = (\sigma^2_{\eta_{\alpha 1}} \;\; \sigma^2_{\eta_{\alpha 2}} \;\; \sigma^2_{\eta_{\alpha 3}})',$$
$$\bar{\Phi}_{S_t} = (\bar{\Phi}_{\alpha_1,S_t} \;\; \bar{\Phi}_{\alpha_2,S_t} \;\; \bar{\Phi}_{\alpha_3,S_t})',$$
$$F = \begin{pmatrix} f_{\alpha_1} & 0 & 0 \\ 0 & f_{\alpha_2} & 0 \\ 0 & 0 & f_{\alpha_3} \end{pmatrix},$$

and similarly for the IS curve, we have

$$Y_t = y_t, \;\; X_t = (1 \;\; y_{t-1} \;\; i_{t-1}), \;\; \phi_t = (\beta_{1t} \;\; \beta_{2t} \;\; \beta_{3t})', \;\; \varepsilon_t = \varepsilon_{yt},$$

with

$$\varepsilon_{yt} \sim N(0, \sigma^2_{\varepsilon y,S_t}),$$
$$\sigma^2_{\varepsilon y,S_t} = \sigma^2_{\varepsilon y,0} + (\sigma^2_{\varepsilon y,1} - \sigma^2_{\varepsilon y,0})S_t, \;\; \sigma^2_{\varepsilon y,1} > \sigma^2_{\varepsilon y,0},$$

[3] As for the details of the State-Space model with Markov-Switching, the reader is referred to Kim and Nelson (1999, ch. 5). The program applied below is based on the Gauss Programs developed by Kim and Nelson (1999).

and

$$\eta_t = (\eta_{\beta 1t} \;\; \eta_{\beta 2t} \;\; \eta_{\beta 3t})',$$
$$\sigma_\eta^2 = (\sigma_{\eta\beta 1}^2 \;\; \sigma_{\eta\beta 2}^2 \;\; \sigma_{\eta\beta 3}^2)',$$
$$\bar{\Phi}_{S_t} = (\bar{\Phi}_{\beta_1,S_t} \;\; \bar{\Phi}_{\beta_2,S_t} \;\; \bar{\Phi}_{\beta_3,S_t})'$$
$$F = \begin{pmatrix} f_{\beta_1} & 0 & 0 \\ 0 & f_{\beta_2} & 0 \\ 0 & 0 & f_{\beta_3} \end{pmatrix}.$$

Next we use the U.S. quarterly data 1961.1-1999.4 for the estimation. The inflation rate is measured by changes in the Consumer Price Index, the output gap is measured by the percentage deviation of the log value of the Industrial Production Index (IPI) from its HP filtered trend and r_t is the Federal Funds rate.[4] The data source is the International statistics Yearbook 2000. The estimates of the hyper-parameters are shown in Table 1. We find significant differences between $\sigma_{\varepsilon\pi,0}$ (0.0021) and $\sigma_{\varepsilon\pi,1}$ (0.0053), and $\sigma_{\varepsilon y,0}$ (1.11×10^{-7}) and $\sigma_{\varepsilon y,1}$ (0.0205). The differences between $\bar{\Phi}_{\beta_2,0}$ (0.3972) and $\bar{\Phi}_{\beta_2,1}$ (0.8460), $\bar{\Phi}_{\beta_3,0}$ (0.0273) and $\bar{\Phi}_{\beta_3,1}$ (-0.4893), and $\bar{\Phi}_{\alpha_3,0}$ (0.0046)and $\bar{\Phi}_{\alpha_3,1}$ (0.0132) are also significant. The fact that all the elements of F are smaller than 1 indicates that the time-varying parameters are stationary and therefore justifies the adoption of eq. (9).

The paths of α_{2t} are shown in Figure 1A.[5] We leave aside the paths of the intercepts in the IS- and Philips curves. In Figure 1A, "Alpha_2t,0" denotes the path of α_{2t} when $[S_t = 0|Y_t]$, namely $\alpha_{2t,0}$. "Alpha_2t,1" denotes the path of α_{2t} when $[S_t = 1|Y_t]$, namely $\alpha_{2t,1}$. "Alpha_2t" denotes the weighted average of $\alpha_{2t,0}$ and $\alpha_{2t,1}$, α_{2t}. That is,

$$\alpha_{2t} = Pr[S_t = 0|Y_t]\alpha_{2t,0} + Pr[S_t = 1|Y_t]\alpha_{2t,1}.$$

The paths of α_{3t} are shown in Figure 1B. In Figure 1B, "Alpha_3t,0" denotes the path of α_{3t} when $[S_t = 0|Y_t]$ ($\alpha_{3t,0}$). "Alpha_3t,1" denotes the path of α_{3t} when $[S_t = 1|Y_t]$ ($\alpha_{3t,1}$). Similarly, "Alpha_3t" denotes the weighted average of $\alpha_{3t,0}$ and $\alpha_{3t,1}$ (α_{3t}). Figure 1C represents the weighted average of the forecast errors $\xi_{t|t-1}$. In (13) we find that the conditional variance of the forecast errors consists

[4] The IPI has also been used by Clarida, Gali and Gertler (1998) to measure the output for Germany, France, the U.S., the U.K., Japan and Italy. As surveyed by Orphanides and van Norden (2002), there are many methods to measure the output gap. We find that filtering the IPI using the Band-Pass filter developed by Baxter and King (1995) leaves the measure of the output gap essentially unchanged from the measure with the HP-filter. The Band-Pass filter has also been used by Sargent (1999).

[5] In order to eliminate the effects of the initial startup idiosyncracies of the Kalman filter algorithm, we present the paths of the variables concerned from t=12 on, namely from 1964.3 to 1999.4.

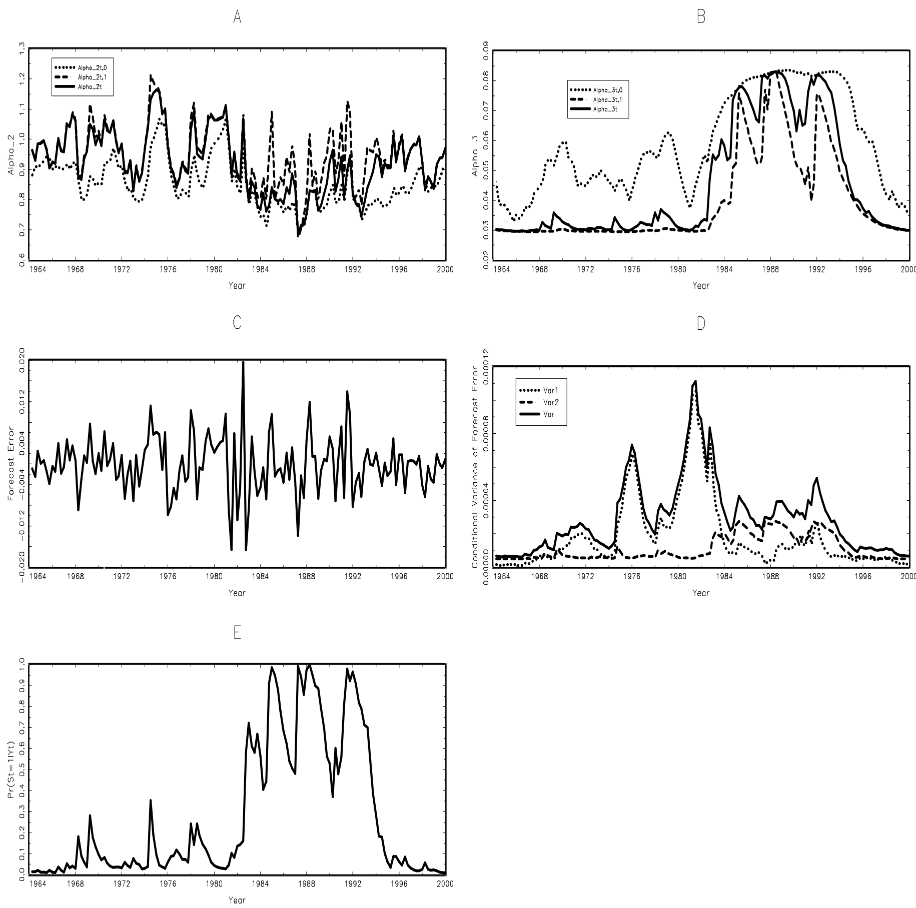

Figure 1: Results of the Linear Time-Varying Phillips Curve

Phillips curve			IS curve		
Parameter	Estimate	S.D.	Parameter	Estimate	S.D.
$\sigma_{\varepsilon\pi,0}$	0.0021	0.0004	$\sigma_{\varepsilon y,0}$	1.11×10^{-7}	0.0125
$\sigma_{\varepsilon\pi,1}$	0.0053	0.0009	$\sigma_{\varepsilon y,1}$	0.0205	0.0032
$\sigma_{\eta_{\alpha 1}}$	$2.53{\times}10^{-8}$	0.0051	$\sigma_{\eta_{\beta 1}}$	0.0061	0.0005
$\sigma_{\eta_{\alpha 2}}$	0.0698	0.0113	$\sigma_{\eta_{\beta 2}}$	0.0106	0.1951
$\sigma_{\eta_{\alpha 3}}$	$2.63{\times}10^{-9}$	0.0146	$\sigma_{\eta_{\beta 3}}$	0.0420	0.0374
$\bar{\Phi}_{\alpha_1,0}$	0.0009	0.0015	$\bar{\Phi}_{\beta_1,0}$	0.0011	0.0009
$\bar{\Phi}_{\alpha_1,1}$	0.0057	0.0073	$\bar{\Phi}_{\beta_1,1}$	0.0011	0.0020
$\bar{\Phi}_{\alpha_2,0}$	0.3401	0.1133	$\bar{\Phi}_{\beta_2,0}$	0.3972	0.1112
$\bar{\Phi}_{\alpha_2,1}$	0.2733	0.1014	$\bar{\Phi}_{\beta_2,1}$	0.8460	0.1496
$\bar{\Phi}_{\alpha_3,0}$	0.0046	0.0096	$\bar{\Phi}_{\beta_3,0}$	0.0273	0.0518
$\bar{\Phi}_{\alpha_3,1}$	0.0132	0.0249	$\bar{\Phi}_{\beta_3,1}$	-0.4893	0.0762
f_{α_1}	0.2584	0.9793	f_{β_1}	0.6177	0.0890
f_{α_2}	0.6536	0.1095	f_{β_2}	0.3645	0.1004
f_{α_3}	0.8431	0.2936	f_{β_1}	0.2129	0.1429
p	0.9687	0.0339	p	0.8187	0.1012
q	0.9867	0.0139	q	0.9442	0.0269
Likelihood	-567.50		Likelihood	-458.24	

Table 1: Estimates of the Hyperparameters in the Linear Time-Varying IS- and Phillips Curves

of two distinct terms: The conditional variance due to changing coefficients $X_t P_{t|t-1}^{(i,j)} X_t'$ and the conditional variance due to the switching of $\sigma_{\varepsilon,j}^2$. In Figure 1D "Var1" denotes $X_t P_{t|t-1}^{(i,j)} X_t'$, "Var2" denotes $\sigma_{\varepsilon,j}^2$ and "Var" is the sum of the two terms, $\nu_{t|t-1}^{(i,j)}$. When there is no switching in the variance of the forecast errors, $\sigma_{\varepsilon,j}^2$ is constant. Figure 1E represents the path of $Pr[S_t = 1|Y_t]$. The probability that there is regime switching in the Phillips curve around 1982-83, 1992 and 1994-96 is very high. 1983 seems to be a break point for α_{3t}: Before 1983 it has been quite smooth in state 1 and experienced small changes in state 0, but increased suddenly to a much higher value in 1984 in both state 1 and state 0. α_{2t} has also experienced some changes in 1983, though not so obviously as α_{3t}.

The result of the IS curve estimation is demonstrated in Figure 2, which has a similar interpretation as Figure 1. The paths of β_{2t} and β_{3t} are represented in Figure 2A and 2B. The forecast errors are represented in Figure 2C. The conditional variance of the forecast errors are represented in Figure 2D and Figure 2E is the path of $Pr[S_t = 1|Y_t]$. From Figure 2E we find that the probability that there is regime switching in the IS curve around 1970, 1983 and 1992 is very high. From Figure 2A and 2B we find similar evidence. β_{2t} evolves between 0 and 1.4, with β_{3t} between -0.7 and 0.1.

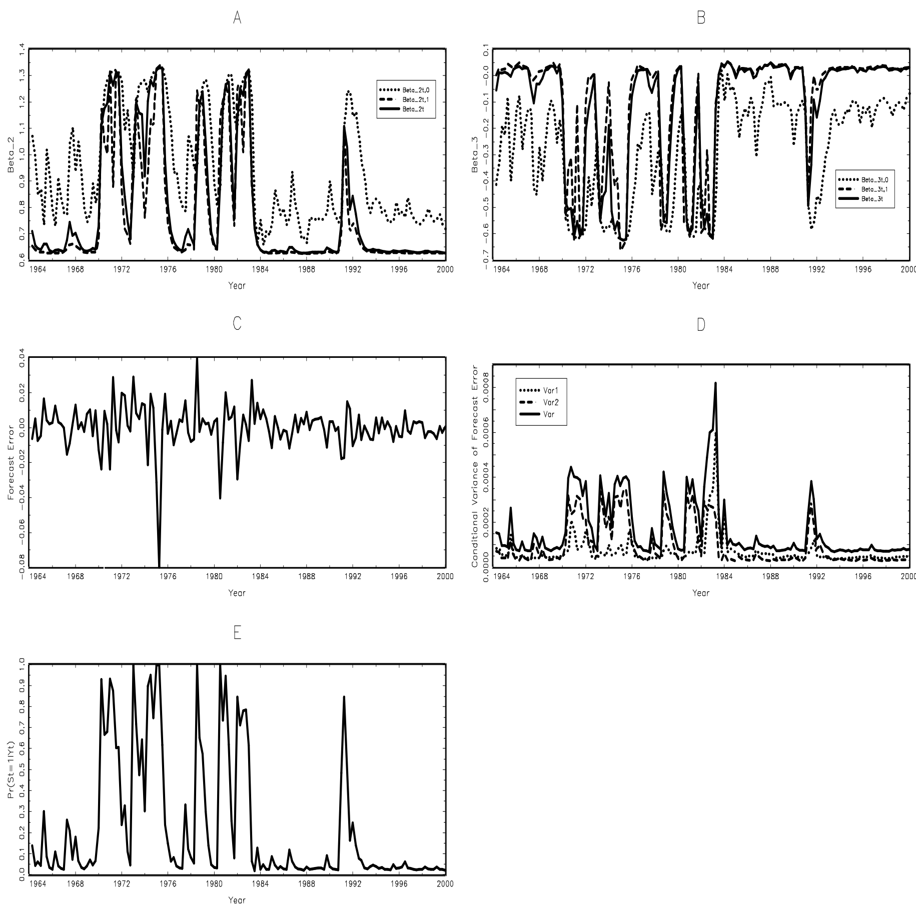

Figure 2: The Results of the Time-Varying IS Curve

2.2 Nonlinear Phillips Curve

In the previous subsection we have explored uncertainty in the simple IS and Phillips curves. The 1990s, however, has seen the development of the literature on the so-called nonlinear Phillips curve. More specifically, according to this literature, positive deviations of aggregate demand from potential are more inflationary than negative deviations are dis-inflationary.[6] Dupasquier and Ricketts (1998a) survey several models of the nonlinearity in the Phillips curve. The five models surveyed are the *capacity constraint model*, the *mis-perception* or *signal extraction model*, the *costly adjustment model*, the *downward nominal wage rigidity model* and the *monopolistically competitive model*. As mentioned by Akerlof (2002), the nonlinearity of the Phillips curve has been an important issue of macroeconomics. Aguiar and Martins (2002), for example, test three kinds of nonlinearities (quadratic, hyperbole and exponential) in the Phillips curve and Okun's law with the aggregate EURO-area macroeconomic data and find that the Phillips curve turns out to be linear, but the Okun's law nonlinear. Many empirical studies have been undertaken to explore the Phillips-curve nonlinearity. Dupasquier and Ricketts (1998a) explore nonlinearity in the Phillips curve for Canada and the U.S. and conclude that there is stronger evidence in favor of nonlinearity for the U.S than for Canada. Other studies on the nonlinearity of the Phillips curve include Knoppik (2001), Razzak (1997), Gómez and Julio (2000), Clements and Sensier (2002), Dupasquier and Ricketts (1998b), Chadha, Masson and Meredith (1992), Laxton, Meredith and Rose (1995) and Bean (2000). Monetary policy with a nonlinear Phillips curve has also been explored, see Schaling (1999), Tambakis (1998) and Flaschel, Gong and Semmler (2001) for example. Since monetary policy with a linear Phillips curve can be different from that with a nonlinear Phillips curve, we will explore uncertainty in a nonlinear Phillips curve below. In the following section we will also take into account the nonlinearity of the Phillips curve in an optimal control model.

As discussed by Aguiar and Martins (2002), there may be different forms of nonlinearity in the Phillips curve. In the research below we just follow Schaling (1999) and assume that the nonlinear form of the output gap in the Phillips curve reads as[7]

$$f(y_t) = \frac{\alpha y_t}{1 - \alpha\beta y_t}, \ \alpha > 0, 1 > \beta \geq 0, \tag{16}$$

where y_t denotes the output gap and the parameter β indexes the curvature of the curve. When β is very small, the curve approaches a linear relationship. Assuming α=10 and β=0.99, we present $f(y_t)$ with the U.S. quarterly data in Figure 3. It is obvious that when the actual output is lower than the potential output, the curve of $f(y_t)$ is flatter. From the figure we see this function

[6] There is, of course, also the other issue, that the central bank may react with interest rate changes more to inflationary than to deflationary pressures, for the Euro-area case, see Semmler, Greiner and Zhang (2002b).

[7] Note that this function is not continuous with a breaking point at $y_t = \frac{1}{\alpha\beta}$. When $y_t < \frac{1}{\alpha\beta}$, $f''(y_t) > 0$ and if $y_t > \frac{1}{\alpha\beta}$, $f''(y_t) < 0$. In the research below we choose appropriate values of α and β so that with the U.S. output gap data we have $f''(y_t) > 0$.

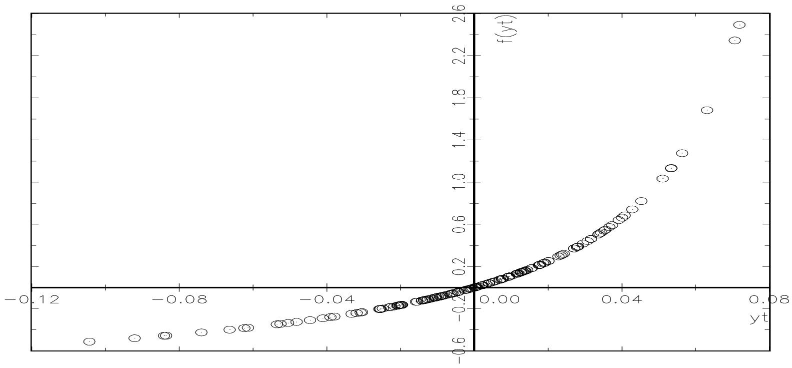

Figure 3: An Example of $f(y_t)$

describes very well the idea that positive deviations of aggregate demand from potential are more inflationary than negative deviations are dis-inflationary.

Substituting $f(y_t)$ for y_t in the Phillips curve, we have now

$$\pi_t = \alpha_{1t} + \alpha_{2t}\pi_{t-1} + \alpha_{3t}f(y_{t-1}) + \varepsilon_{\pi,t}. \tag{17}$$

Following the same procedure in the previous subsection, we present the results of the State-Space form of eq. (17) in Table 2 and Figure 4.

In Figure 4 we also observe some structural changes in the coefficients. But the difference between Figure 4 and Figure 1 is obvious. The structural changes of the coefficients show up mainly between the second half of the 1970s and the beginning of the 1990s in Figure 4, while they show up between the second half of the 1980s and the first half of the 1990s in Figure 1.

Above we have explored model and shock uncertainties in the IS-Phillips curves with the U.S data. We also have explored whether regime changes have occurred in the U.S. economy since the 1960s. The results are, to some extent, consistent with the line of research that maintains that there were regime changes in the U.S. economy.[8] Overall, the uncertainty of parameters and shocks, and their impact on monetary policy rules suggest exploring monetary policy rules with learning and robust control.

3 Monetary Policy Rules with Adaptive Learning

Svensson (1999) and Semmler, Greiner and Zhang (2002b) derive an optimal monetary policy rule in an optimal control model with a quadratic loss function

[8]See Cogley and Sargent (2001, 2002) for example.

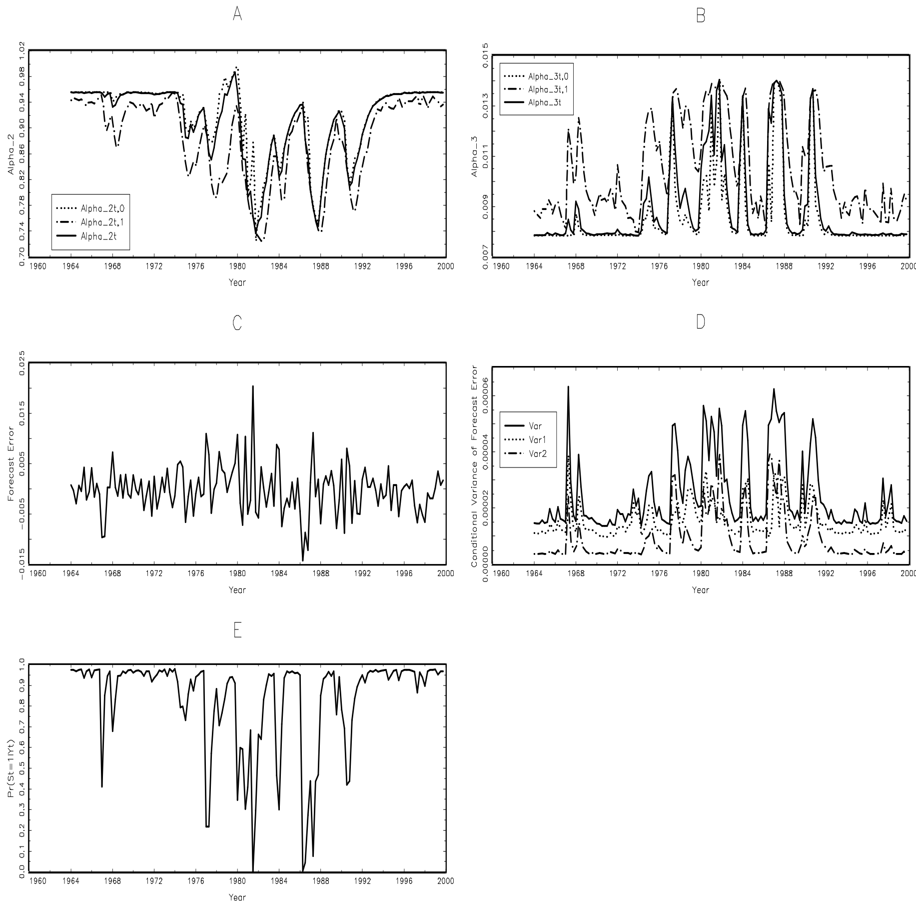

Figure 4: Results of the Time-Varying nonlinear Phillips Curve

Parameter	Estimate	S.D.
$\sigma_{\varepsilon\pi,0}$	0.0073	0.0013
$\sigma_{\varepsilon\pi,1}$	1.470 $\times 10^{-9}$	0.0024
$\sigma_{\eta_{\alpha 1}}$	0.003	0.0002
$\sigma_{\eta_{\alpha 2}}$	1.087 $\times 10^{-8}$	0.0104
$\sigma_{\eta_{\alpha 3}}$	3.258 $\times 10^{-9}$	0.0013
$\bar{\Phi}_{\alpha_1,0}$	0.0124	0.0018
$\bar{\Phi}_{\alpha_1,1}$	0.0015	0.0007
$\bar{\Phi}_{\alpha_2,0}$	0.152	0.051
$\bar{\Phi}_{\alpha_2,1}$	0.213	0.069
$\bar{\Phi}_{\alpha_3,0}$	0.012	0.010
$\bar{\Phi}_{\alpha_3,1}$	0.0065	0.005
f_{α_1}	0.169	0.089
f_{α_2}	0.778	0.071
f_{α_3}	0.174	0.639
p	0.949	0.021
q	0.746	0.101
Likelihood	-582.60	

Table 2: Estimates of the Hyperparameters in the Nonlinear Time-Varying Phillips Curve

and the IS- and Phillips curves as constraints. This optimal monetary policy rule is similar to the Taylor rule (Taylor 1993 and 1999). Yet, it is found that the optimal monetary policy rule can be greatly influenced by the parameters in the state equations. The question arising is, therefore, what is the optimal monetary policy rule in case some parameters or shocks in an economic model such as eq. (2) are uncertain? Recently numerous papers have been contributed to this problem. Svensson (1999), Orphanides and Williams (2002), Tetlow and von zur Muehlen (2001a), Söderström (1999), and Beck and Wieland (2002), for example, explore optimal monetary policy rules under the assumption that the economic agents learn the parameters in the model in a certain manner. One assumption is that the economic agents may learn the parameters using the Kalman filter. This assumption has been taken by Tucci (1997) and Beck and Wieland (2002). Another learning mechanism which is also applied frequently is recursive least squares (RLS). This kind of learning mechanism has been applied by Sargent (1999) and Orphanides and Williams (2002). By intuition we would expect that economic agents reduce uncertainty and therefore improve economic models by learning with all information available. Of course, there is the possibility that economic agents do not improve model specification but seek a monetary policy rule robust to uncertainty. This is what robust control theory aims at.

In this section we will explore monetary policy rules under uncertainty under the assumption that the central banks reduce uncertainty by way of learning. As mentioned above, some researchers, Beck and Wieland (2002) and Orphanides and Williams (2002) for example, have explored this problem. Besides the difference in the learning algorithm, another difference between Beck and Wieland (2002) and Orphanides and Williams (2002) is that the former do not consider role of expectations in the model, while the latter take into account expectations in the Phillips curve. Unlike Beck and Wieland (2002), Orphanides and Williams (2002) do not employ an intertemporal framework. They provide a learning algorithm with constant gain but do not use a discounted loss function. Moreover, Orphanides and Williams (2002) assume that the government knows the true model, but the private agents do not know the true model and have to learn the parameters with the RLS algorithm. In their case the government and the private agents are treated differently. Sargent (1999) employs both a learning algorithm as well as a discounted loss function but in a linear-quadratic (LQ) model. This implies that after one step of learning the learned coefficient is presumed to remain forever when the LQ problem is solved. In our model, however, both the central bank and the private agents are learning the parameters, that is, they are not treated differently.

The difference of our model from that of Beck and Wieland (2002) can be summarized in three points: First, we consider both linear and nonlinear Phillips curves. Second, we take into account expectations. This is consistent with the model of Orphanides and Williams (2002). An important characteristic of New-Keynesian economics is that current economic behavior depends not only on the current and past policy but also on the expectations of agents. Third, we employ the RLS learning algorithm instead of the Kalman filter. In fact, Harvey (1989) and Sargent (1999) prove that RLS is a specific form of the Kalman filter. Evans and Honkapohja (2001) analyze expectations and learning mechanisms in macroeconomics in detail. The difference to Sargent (1999) is that we in fact can allow for both coefficient drift through learning by RLS and solve a nonlinear optimal control model using a dynamic programming algorithm.

3.1 RLS Learning in Linear Phillips Curve

Orphanides and Williams (2002) assume that the current inflation rate is not only affected by the inflation lag but also by inflation expectations. Following Orphanides and Williams (2002), we assume that the linear Phillips curve takes the following form:

$$\pi_t = \gamma_1 \pi_{t-1} + \gamma_2 \pi_t^e + \gamma_3 y_t + \varepsilon_t, \ \ \varepsilon \sim iid(0, \sigma_\varepsilon^2), \tag{18}$$

where π_t^e denotes the agents' (including the central bank) expected inflation rate based on the time t information, γ_1, $\gamma_2 \in$ (0,1), $\gamma_3 > 0$ and ε is a serially uncorrelated innovation. In order to simplify the analysis, we further assume the IS equation to be deterministic taking the following form:[9]

$$y_t = -\theta r_{t-1}, \ \theta > 0, \tag{19}$$

[9]This is the same as Orphanides and Williams (2002), except that they include a noise in the equation.

where

$$r_t = rr_t - r^*,$$

with rr_t denoting the real interest rate and r^* the equilibrium real rate. Substituting eq. (19) into (18), we have

$$\pi_t = \gamma_1 \pi_{t-1} + \gamma_2 \pi_t^e - \gamma_3 \theta r_{t-1} + \varepsilon_t, \ \ \varepsilon \sim iid(0, \sigma_\varepsilon^2). \tag{20}$$

In case of rational expectations, namely, $\pi_t^e = E_{t-1}\pi_t$, we get

$$E_{t-1}\pi_t = \gamma_1 \pi_{t-1} + \gamma_2 E_{t-1}\pi_t - \gamma_3 \theta r_{t-1},$$

that is,

$$E_{t-1}\pi_t = \bar{a}\pi_{t-1} + \bar{b} r_{t-1},$$

with

$$\bar{a} = \frac{\gamma_1}{1-\gamma_2} \tag{21}$$

$$\bar{b} = -\frac{\gamma_3 \theta}{1-\gamma_2}. \tag{22}$$

With these results we get the rational expectations equilibrium (REE)

$$\pi_t = \bar{a}\pi_{t-1} + \bar{b} r_{t-1} + \varepsilon_t. \tag{23}$$

Now suppose that the agents believe the inflation rate follows the process

$$\pi_t = a\pi_{t-1} + b r_{t-1} + \varepsilon_t,$$

corresponding to the REE, but that a and b are unknown and have to be learned. Suppose that the agents have data on the economy from periods $i = 0, ..., t-1$. Thus the time-(t-1) information set is $\{\pi_i, r_i\}_{i=0}^{t-1}$. Further suppose that agents estimate a and b by a least squares regression of π_i on π_{i-1} and r_{i-1}. The estimates will be updated over time as more information is collected. Let (a_{t-1}, b_{t-1}) denote the estimates through time t-1, the forecast of the inflation rate is then given by

$$\pi_t^e = a_{t-1}\pi_{t-1} + b_{t-1} r_{t-1}. \tag{24}$$

The standard least squares formula gives the equations

$$\begin{pmatrix} a_t \\ b_t \end{pmatrix} = \left(\sum_{i=1}^{t} z_i' z_i \right)^{-1} \left(\sum_{i=1}^{t} z_i' \pi_i \right), \tag{25}$$

where $z_i = \begin{pmatrix} \pi_{i-1} & r_{i-1} \end{pmatrix}'$.

Defining $c_t = \begin{pmatrix} a_t \\ b_t \end{pmatrix}$, we can also compute eq. (25) using the stochastic approximation of the recursive least squares equations

$$c_t = c_{t-1} + \kappa_t V_t^{-1} z_t (\pi_t - z_t' c_{t-1}), \tag{26}$$

$$V_t = V_{t-1} + \kappa_t (z_t z_t' - V_{t-1}), \tag{27}$$

where c_t and V_t denote the coefficient vector and the moment matrix for z_t using data $i = 1, ..., t$. κ_t is the gain. To generate the least squares values, the initial value for the recursion must be set appropriately.[10] The gain κ_t is an important variable. According to Evans and Honkapohja (2001), the assumption that $\kappa_t = t^{-1}$ (decreasing gain) together with the condition $\gamma_2 < 1$ ensures the convergence of c_t as $t \to \infty$. That is, as $t \to \infty$, $c_t \to \bar{c}$ with probability 1, with $\bar{c} = \begin{pmatrix} \bar{a} \\ \bar{b} \end{pmatrix}$ and therefore $\pi_t^e \to$ REE.

As indicated by Sargent (1999) and Evans and Honkapohja (2001), if κ_t is a constant, however, there might be difficulties of convergence to the REE. If the model is non-stochastic and κ_t sufficiently small , π_t^e converges to REE under the condition $\gamma_2 < 1$. However, if the model is stochastic with $\gamma_2 < 1$, the belief does not converge to REE, but to an ergodic distribution around it. Here we follow Orphanides and Williams (2002) and assume that agents are constantly learning in a changing environment. The assumption of a constant gain indicates that the agents believe the Phillips curve wanders over time and give larger weights to the recent observations of the inflation rate than to the earlier ones. Orphanides and Williams (2002) denote the case of $\kappa_t = \frac{1}{t}$ as infinite memory and the case of constant κ_t as finite memory. As many papers on monetary policy (Svensson, 1997, 1999 for example) we assume that the central bank pursues a monetary policy by minimizing a quadratic loss function. The problem reads as

$$\underset{\{r_t\}_0^\infty}{Min}\ E_0 \sum_{t=0}^{\infty} \rho^t L(\pi_t, r_t), \quad L(\pi_t, r_t) = (\pi_t - \pi^*)^2, \tag{28}$$

subject to eq. (20), (24), (26) and (27). π^* is the target inflation rate, which will be assumed to be zero just for the purpose of simplification.

Note that the difference of our model from that of Sargent (1999) is obvious, although he also applies the RLS learning algorithm and an optimal control framework with infinite horizon. Yet, Sargent (1999) constructs his results in two steps. First, following the RLS with a decreasing or constant gain, the agents estimate a model of the economy (the Phillips curve) using the latest available data, updating parameter estimates from period to period. Second, once the parameter is updated, the government pretends that the updated parameter will govern the dynamics forever and derives an optimal policy from an LQ control model. These two steps are repeated over and over. As remarked by Tetlow and von zur Muehlen (2001b), however, there is a problem in the approach of Sargent(1999). Sargent's approach is based on two assumptions: First, the economy is subject to drift in its structural parameters and second, notwithstanding this acknowledgement, the policymaker takes the estimated parameters at each date as the truth and bases policy decisions on these values. It is easy to see that the second assumption is inconsistent with the first one. Our model, however, treats the changing parameters as endogenous variables in

[10] Assuming $Z_k = (z_1, ...z_k)'$ is of full rank and letting π^k denote $\pi^k = (\pi_1, ..., \pi_k)'$, the initial value c_k is given by $c_k = Z_k^{-1}\pi^k$ and the initial value V_k is given by $V_k = k^{-1}\sum_{i=1}^{k} z_i z_i'$.

a nonlinear optimal control problem. This is similar to the methodology used by Beck and Wieland (2002).

As mentioned above, if the unknown parameters are adaptively estimated with RLS with a small and constant gain, they will converge in distributions in a stochastic model and converge to a point in a non-stochastic model. But in an optimal control problem such as (28) with nonlinear state equation the model will not necessarily converge even if the state equations are non-stochastic.

Next we undertake some simulations for the model. Though the return function is quadratic and the Phillips curve linear, the problem falls outside the scope of LQ optimal control problems, since some parameters in the Phillips curve are time-varying and follow a nonlinear path. Therefore the problem can not be solved analytically and numerical solutions have to be employed. In the simulations below we resort to the algorithm developed by Grüne (1997), who applies adaptive instead of uniform grids. A less technical description of this algorithm can be found in Grüne and Semmler (2002). The simulations are undertaken for the deterministic case. In order to simplify the simulations, we assume that a_t is known to be a constant value equal to $\bar{a}$. Therefore only b_t has to be learned in the model. In this case we have $c_t = b_t$ and $z_i = r_{i-1}$. As mentioned by Beck and Wieland (2002), the reason for focusing exclusively on incomplete information regarding b is that this parameter is multiplicative to the decision variable r_t and therefore central to the tradeoff between current control and estimation.

Simulation

In the simulations we assume $\gamma_1 = 0.6$, $\gamma_2 = 0.4$, $\gamma_3 = 0.5$, $\theta = 0.4$, $\rho = 0.985$ and $\kappa_t = 0.05$. The initial values of π_t, b_t and V_t are 0.2, -0.6 and 0.04. The paths of π_t, b_t, V_t and r_t are shown in Figure 5A-D respectively. Figure 5E is the phase diagram of π_t and r_t. Neither the state variables nor the control variable converge. In fact, they fluctuate cyclically. We try the simulations with many different initial values of the state variables and smaller κ_t (0.01 for example) and find that in no case do the variables converge. Similar results are obtained with different values for γ_1 (0.9 and 0.3 for example) and γ_2 (0.1 and 0.7 for example).

With the parameters above, we have $\bar{a} = 1$, $\bar{b} = -0.33$, therefore the REE is

$$\pi_t = \pi_{t-1} - 0.33 r_{t-1} + \varepsilon_t. \tag{29}$$

In the case of RLS learning, however, we have

$$\pi_t = \pi_{t-1} + \tilde{b}_t r_{t-1} + \varepsilon_t,$$

with

$$\tilde{b}_t = \gamma_2 b_{t-1} - \gamma_3 \theta.$$

The path of $\tilde{b}_t$ is presented in Figure 6. $\tilde{b}_t$ evolves at a higher level than $\bar{b}$. Simulations are undertaken with different initial values of the state variables and similar results for $\tilde{b}_t$ are found.

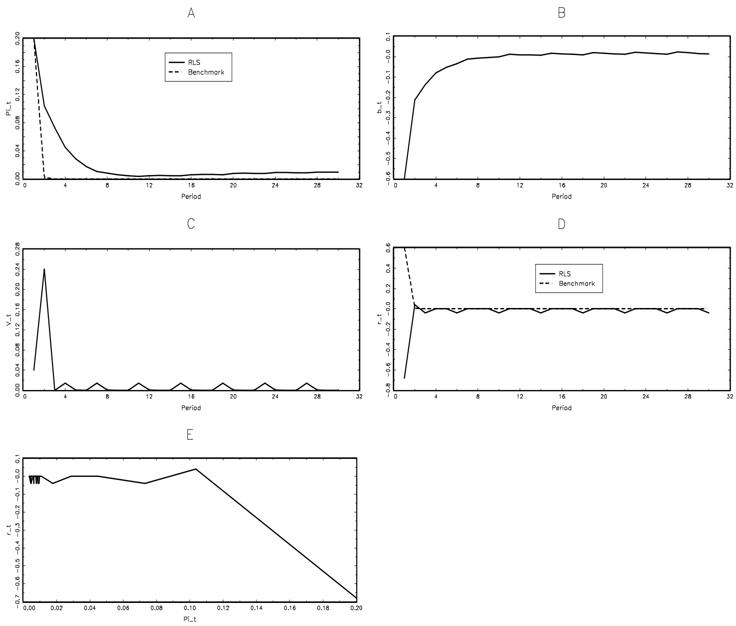

Figure 5: Simulations of RLS Learning (solid) and Benchmark Model (dashed) with Linear Phillips Curve

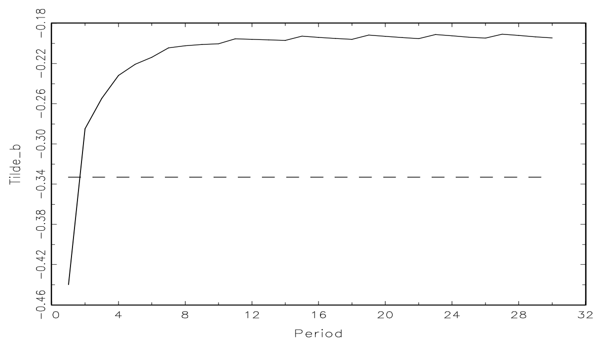

Figure 6: Path of $\tilde{b}_t$ (solid) in the Linear Phillips Curve

If there is perfect knowledge, namely, the agents have rational expectation, π_t can converge to its target value π^* (zero here), since the model then becomes a typical LQ control problem which has converging state and control variables in a non-stochastic model. We define this case as the benchmark model. The results of the benchmark model are shown in Figure 5A and 5D (dashed line). Note that in the benchmark model there is only one state variable, namely π_t with dynamics denoted by (29). In the non-stochastic benchmark model the optimal monetary policy rule turns out to be $r_t = 3.00\pi_t$ and the optimal trajectory of π_t is $\pi_t = 0.01\pi_{t-1}$. From Figure 5A and 5D we observe that as time goes on π_t and r_t converge to zero in the benchmark model.

3.2 RLS Learning in Nonlinear Phillips Curve

As mentioned in Section 2, the Phillips curve could be nonlinear. Given the nonlinearity of the Phillips curve, eq. (18) reads as,

$$\pi_t = \gamma_1 \pi_{t-1} + \gamma_2 \pi_t^e + \gamma_3 f(y_t) + \varepsilon_t, \quad \varepsilon \sim iid(0, \sigma_\varepsilon^2), \tag{30}$$

with $f(y_t)$ given by eq. (16). Substituting eq. (19) into eq. (16), and then (16) into (30), we get the following nonlinear Phillips curve

$$\pi_t = \gamma_1 \pi_{t-1} + \gamma_2 \pi_t^e - \gamma_3 g(r_{t-1}) + \varepsilon_t, \quad \varepsilon \sim iid(0, \sigma_\varepsilon^2), \tag{31}$$

where

$$g(r_t) = \frac{\alpha\theta r_t}{1 + \alpha\beta\theta r_t}.$$

The REE turns out to be

$$\pi_t = \bar{a}\pi_{t-1} + \bar{b}g(r_{t-1}) + \varepsilon_t, \tag{32}$$

where $\bar{a}$ is defined in (21) but $\bar{b}$ is changed to be $-\frac{\gamma_3}{1-\gamma_2}$. The forecast of the inflation rate is now given by

$$\pi_t^e = a_{t-1}\pi_{t-1} + b_{t-1}g(r_{t-1}). \tag{33}$$

The RLS learning mechanism is the same as the case of the linear Phillips curve, except that z_i is now modified as

$$z_i = \begin{pmatrix} \pi_{i-1} & g(r_{i-1}) \end{pmatrix}'.$$

The optimal control problem (28) now turns out to have constraints (31), (33), (26) and (27).

Simulation

In the simulations we take the same values for the parameters in the model as in the previous subsection and assume $\alpha = 10$ and $\beta = 0.99$. The simulations with the same starting values of the state variables as in the previous subsection

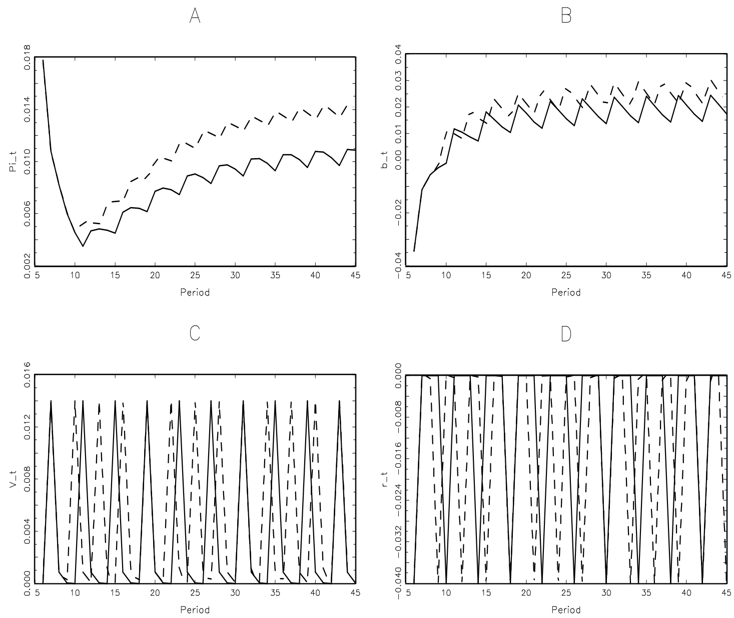

Figure 7: RLS Learning with Linear (solid) and Nonlinear Phillips Curves (dashed)

are presented in Figure 7A-D. Figure 7A represents the path of π_t, 7B the path of b_t, 7C the path of V_t and 7D the path of r_t. The results of this subsection (nonlinear Phillips curve) are presented by dashed lines, while the results from the previous subsection (linear Phillips curve) are indicated by solid lines.[11]

We find that the state variables also do not converge in the optimal control problem with the nonlinear Phillips curve. Similar to the case of the linear Phillips curve, the state and control variables fluctuate cyclically. Simulations with many different initial values of state variables were undertaken and in no case are the state variables found to converge. But the difference between the simulations with linear and nonlinear Phillips curves is not to ignore. Figure 7 indicates that both π_t (Figure 7A) and b_t (Figure 7B) evolve at a higher level in the case of a nonlinear Phillips curve than in the case of a linear one. The mean and standard deviation of π_t, b_t, V_t and r_t from the two simulations are shown in Table 3. The S.D. and absolute values of the mean of these variables are larger in the case of nonlinear Phillips curve than when the Phillips curve is linear.

Next we show the $\tilde{b}$ in the nonlinear Phillips curve in Figure 8. $\tilde{b}$ in the nonlinear Phillips curve equals $\gamma_2 b_{t-1} - \gamma_3$. The $\tilde{b}$ and $\bar{b}$ from the simulations with the linear Phillips curve are also shown in Figure 8, from which we find

[11] In order to see the differences of the simulations clearly, we just present the results from t=6 on.

	π_t		b_t		V_t		r_t	
	L	NL	L	NL	L	NL	L	NL
mean	0.0102	0.0135	0.0181	0.0243	0.0037	0.0049	-0.0101	-0.0135
S.D.	0.0016	0.0022	0.0069	0.0077	0.0060	0.0064	0.0174	0.0190

Table 3: Mean and S.D. of State and Control Variables. (L and NL stand for linear and nonlinear Phillips curves respectively)

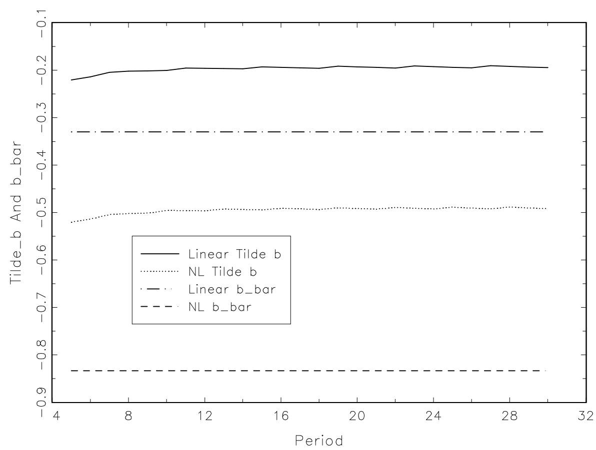

Figure 8: Paths of $\tilde{b}_t$ and $\bar{b}$ in Linear and Nonlinear Phillips Curves (NL stands for nonlinear)

that the $\tilde{b}$ evolves at a higher level than $\bar{b}$ in both linear and nonlinear Phillips curves.

Above we have explored optimal monetary policy rules with adaptive learning. The simulations indicate that the state variables do not converge no matter whether the linear or nonlinear Phillips curve is employed as constraint in the optimal control problem. But the state and control variables seem to experience larger changes in the nonlinear Phillips curve than in the linear one. The results are different from those of Sargent (1999), since in his model the state variables should converge in a non-stochastic model, as explored by Evans and Honkapohja (2001).

4 Monetary Policy Rules with the Robust Control

Facing uncertainties, economic agents can improve their knowledge of economic models by learning with all information available. This is what has been explored in Section 3. A disadvantage of the adaptive learning analyzed in the previous section is that we have considered only parameter uncertainty. Other uncertainties such as shock uncertainty explored in the first section, may also exist. Moreover, as studied in some recent literature, there is the possibility that economic agents resort to a strategy robust to uncertainty instead of learning. This problem has recently been largely explored with the robust control theory. Robust control induces the economic agents to seek a strategy for the "worst case". The robust control theory assumes that there is some model misspecfication–not only the uncertainty of the parameters like α_t and β_t estimated in the IS- and Phillips curves in Section 2, but also other kinds of uncertainties. Therefore, the robust control might deal with more general uncertainty than the adaptive learning. The robust control is now given more and more attention in the field of macroeconomics, because the classic optimal control theory can hardly deal with model misspecification. On the basis of some earlier papers (see Hansen and Sargent 1999, 2001a, 2001b and 2001c), Hansen and Sargent (2002) explore robust control in macroeconomics in details. Cagetti, Hansen, Sargent and Williams (2001) also employ the robust control in macroeconomics. Svensson (2000) analyzes the idea of robust control in a simpler framework. Giordani and Söderlind (2002), however, extend robust control by including forward-looking behavior.

In this section we will also explore monetary policy rules using the robust control. Before starting empirical research we discuss briefly the framework of robust control, following Hansen and Sargent (2002). Let the one-period loss function be L(y,u)=$-(x'Qx + u'Ru)$, with Q and R both being symmetric and positive semi-definite matrices. The optimal linear regulator problem without model misspecification is

$$\underset{\{u_t\}_{t=0}^{\infty}}{Max} \; E_0 \sum_{t=0}^{\infty} \rho^t L(x_t, u_t), \;\; 0 < \rho < 1, \tag{34}$$

subject to the so-called approximating model

$$x_{t+1} = Ax_t + Bu_t + C\check{\epsilon}_{t+1}, \;\; x_0 \; given, \tag{35}$$

where $\{\check{\epsilon}\}$ is an iid Gaussian vector process with mean zero and identity contemporaneous covariance matrix. If the decision maker thinks there is some model misspecification, he will not regard the model above as true but as a good approximation to another model that he can not specify. To represent a dynamic misspecification which can not be depicted by $\check{\epsilon}$ because of its iid nature, Hansen and Sargent (2002) take a set of models surrounding eq. (35) of the form (the so-called distorted model)

$$x_{t+1} = Ax_t + Bu_t + C(\epsilon_{t+1} + \omega_{t+1}), \tag{36}$$

where $\{\epsilon_t\}$ is another iid Gaussian process with mean zero and identity covariance matrix and ω_{t+1} a vector process that can feed back in a general way on the history of x:

$$\omega_{t+1} = g_t(x_t, x_{t-1}, ...), \tag{37}$$

where $\{g_t\}$ is a sequence of measurable functions. When eq. (36) generates the data, the errors $\check{\epsilon}$ in (35) are distributed as $\mathcal{N}(\omega_{t+1}, \mathrm{I})$ rather than as $\mathcal{N}(0,\mathrm{I})$. To express the idea that eq. (35) is a good approximation when eq. (36) generates the data, Hansen and Sargent (2002) restrain the approximation errors by

$$E_0 \sum_{t=0}^{\infty} \rho^{t+1} \omega'_{t+1} \omega_{t+1} \leq \eta_0. \tag{38}$$

In order to solve the robust control problem (34) subject to eq. (36) and (38), Hansen and Sargent (2002) consider two kinds of robust control problems, the constraint problem and the multiplier problem, which differ in how they implement the constraint (38). The constraint problem is

$$\underset{\{u_t\}_{t=0}^{\infty}}{Max} \underset{\{\omega_{t+1}\}_{t=0}^{\infty}}{Min} E_0 \sum_{t=0}^{\infty} \rho^t U(x_t, u_t), \tag{39}$$

subject to eq. (36) and (38). Given $\theta \in (\underline{\theta}, +\infty)$ with $\underline{\theta} > 0$, the multiplier problem can be presented as

$$\underset{\{u_t\}_{t=0}^{\infty}}{Max} \underset{\{\omega_{t+1}\}_{t=0}^{\infty}}{Min} E_0 \sum_{t=0}^{\infty} \rho^t \{U(x_t, u_t) + \rho\theta\omega'_{t+1}\omega_{t+1}\}, \tag{40}$$

subject to eq. (36). Hansen and Sargent (2002, ch. 6) prove that under certain conditions the two problems have the same outcomes. Therefore, solving one of the two problems is sufficient. The robustness parameter θ reflects the agents' preferences of robustness and plays an important role in the problem's solution. If θ is $+\infty$, the problem collapses to the traditional optimal control without model misspecification. In order to find a reasonable value for θ, Hansen and Sargent (2002, ch. 13) design a detection error probability function by a likelihood ratio. Consider a fixed sample of observations on the state x_t, $t = 0, ..., T-1$, and let L_{ij} be the likelihood of that sample for model j assuming that model i generates the data, the likelihood ratio is

$$r_i \equiv log \frac{L_{ii}}{L_{ij}}, \tag{41}$$

where $i \neq j$. When model i generates the data, r_i should be positive. Define

$$p_A = Prob(mistake|A) = freq(r_A \leq 0),$$
$$p_B = Prob(mistake|B) = freq(r_B \leq 0).$$

Thus p_A is the frequency of negative log likelihood ratios r_A when model A is true and p_B is the frequency of negative log likelihood ratios r_B when model

B is true. Attach equal prior weights to model A and B, the detection error probability can be defined as

$$p(\theta) = \frac{1}{2}(p_A + p_B). \tag{42}$$

When a reasonable value of $p(\theta)$ is chosen, a corresponding value of θ can be determined by inverting the probability function defined in (42). Hansen and Sargent (2002, ch. 7) find that θ can be defined as the negative inverse value of the so-called risk-sensitivity parameter σ, that is $\theta = -\frac{1}{\sigma}$.

Note the interpretation of the detection error probability. As seen above, it is a statistic concept designed to spell out how difficult it is to tell the approximating model apart from the distorted one. The larger the detection error probability, the more difficult to tell the two models apart. In the extreme case, when it is 0.5 ($\theta = +\infty$), the two models are the same. So a central bank can choose a θ according to how large a detection error probability it wants. If the detection error probability is very small, that means, if it is quite easy to tell the two models apart, it does not make much sense to design a robust rule. As stated by Anderson, Hansen and Sargent (2000), the aim of the detection error probability is to eliminate models that are easy to tell apart statistically, since it is not plausible to set the robustness parameter to be so small that we tailor decisions to be robust against alternatives that can be detected with high confidence with a limited amount of data. Note that the higher the θ, the lower the robustness, not the opposite. In the research below we can see that a larger detection error probability corresponds to a larger θ.

Next we present the solution of the multiplier problem. Define

$$\mathcal{D}(P) = P + PC(\theta I - C'PC)^{-1}C'P, \tag{43}$$

$$\mathcal{F}(\Omega) = \rho[R + \rho B'\Omega B]^{-1}B'\Omega A, \tag{44}$$

$$T(P) = Q + \rho A\left(P - \rho PB(R + \rho B'PB)^{-1}B'P\right)A. \tag{45}$$

Let P be the fixed point of iterations on $T \circ \mathcal{D}$:

$$P = T \circ \mathcal{D}(P),$$

then the solution of the multiplier problem (40) is

$$u = -Fx, \tag{46}$$

$$\omega = Kx, \tag{47}$$

with

$$F = \mathcal{F} \circ \mathcal{D}(P), \tag{48}$$

$$K = \theta^{-1}(I - \theta^{-1}C'PC)^{-1}C'P[A - BF]. \tag{49}$$

It is obvious that in case $\theta = +\infty$, $\mathcal{D}(P) = P$ and the problem collapses into the traditional LQ problem.

Simulations

With the same U.S. data as in Section 2 we get the following OLS estimates of the backward-looking IS- and Phillips curves (t-statistics in parentheses) for 1962.1-1999.4:

$$\pi_t = \underset{(1.961)}{0.002} + \underset{(17.408)}{1.380}\,\pi_{t-1} - \underset{(2.967)}{0.408}\,\pi_{t-2} + \underset{(1.570)}{0.214}\,\pi_{t-3} - \underset{(2.836)}{0.221}\,\pi_{t-4} + \underset{(3.024)}{0.045}\,y_{t-1}, \quad R^2 = 0.970, \tag{50}$$

$$y_t = \underset{(1.050)}{0.002} + \underset{(19.486)}{1.362}\,y_{t-1} - \underset{(7.083)}{0.498}\,y_{t-2} - \underset{(1.360)}{0.074}(R_{t-1} - \pi_{t-1}),\ R^2 = 0.843. \tag{51}$$

Let A_{11} be the sum of the coefficients of the inflation lags in the Phillips curves (0.965) and A_{22} be the sum of the coefficients of the output gap lags in the IS curve (0.864), we define

$$\mathbf{A} = \begin{pmatrix} 0.965 & 0.045 \\ 0.074 & 0.864 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 \\ -0.074 \end{pmatrix}, \quad x_t = \begin{pmatrix} \pi_t \\ y_t \end{pmatrix}.$$

The problem to solve turns out to be

$$\underset{\{R_t\}_{t=0}^{\infty}}{Max}\ \underset{\{\omega_{t+1}\}_{t=0}^{\infty}}{Min}\ E_0 \sum_{t=0}^{\infty} \rho^t \left[-(\pi_t^2 + \lambda y_t^2) + \rho\theta\omega_{t+1}'\omega_{t+1}\right]$$

subject to

$$x_{t+1} = Ax_t + BR_t + C(\epsilon_{t+1} + \omega_{t+1}).$$

With the parameters above and the starting values of π_0 and y_0 both being 0.02, $\lambda = 1$, $\rho = 0.985$ and C = $\begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}$, we present the detection error probability in Figure 9.[12] If we want a detection error probability of about 0.15, $\sigma = -33$, that is $\theta = 0.03$. With $\theta = 0.03$, we get

$$F = \begin{pmatrix} 10.462 & 12.117 \end{pmatrix}, K = \begin{pmatrix} 5.291 & 0.247 \\ 4.737 \times 10^{-7} & 5.486 \times 10^{-7} \end{pmatrix},$$

and the value function turns out to be V(π,y) = 16.240 π^2+1.033y^2+1.421πy+0.113. If we want a higher detection error probability, 0.40 for example, $\sigma = -11$ ($\theta = 0.091$) and we get

$$F = \begin{pmatrix} 7.103 & 11.960 \end{pmatrix}, K = \begin{pmatrix} 1.173 & 0.055 \\ 1.072 \times 10^{-7} & 1.805 \times 10^{-7} \end{pmatrix},$$

and V(π,y) = 11.134 π^2+1.022y^2+0.945πy+0.080. In case $\theta = +\infty$, we have F = $\begin{pmatrix} 6.438 & 11.929 \end{pmatrix}$ and V(π, y) = 10.120π^2+1.020y^2 +0.850πy+0.073. From

[12]T (number of periods) is taken as 150. 5000 simulations are undertaken here.

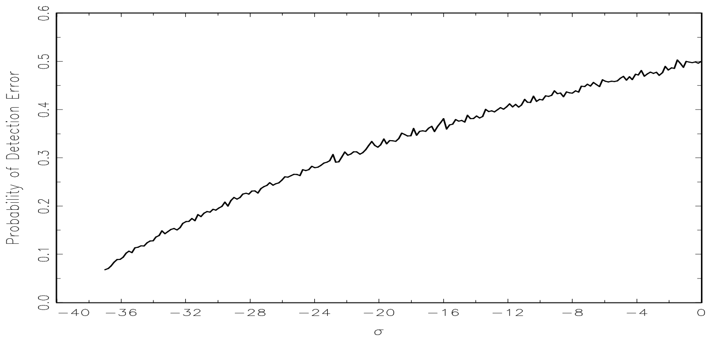

Figure 9: Detection Error Probability

θ	S.D. of π_t	S.D. of y_t	S.D. of R_t
0.03	0.038	0.028	0.223
0.09	0.032	0.017	0.186
$+\infty$	0.030	0.015	0.179

Table 4: Standard Deviations of the State and Control Variables with Different of θ

the results above we find that the lower the θ is, the higher the coefficients of the inflation and output gaps in the interest rate. That is, the farther the distorted model stays away from the approximating one, the stronger the reaction of the interest rate to the inflation and output gaps. We also find that the lower the θ is, the higher the parameters in the value function.

We present the paths of the inflation rate deviation, output gap, the nominal interest rate and the value function with different θ in Figure 10. Figure 10A presents the paths of the value function. It is obvious that the value function with $\theta = 0.03$ (namely $\sigma = -33$) evolves at a higher level than those with θ being 0.09 ($\sigma = -11$) and $+\infty$ ($\sigma = 0$). Figure 10B, 10C and 10D present the paths of the inflation deviation, output gap and interest rate. We find that the lower the θ is, the larger the volatility of the state and control variables. The standard deviations of the state and control variables are shown in Table 4, which indicates that the standard deviations of the state and control variables increase if θ decreases and therefore the value function also increases with the decrease of θ.

Next we come to a special case, namely the case of zero shocks. What do the state and control variables look like and how can the robustness parameter θ affect the state variables and the objective function? According to the certainty equivalence principle, the optimal rules of the robust control with zero shocks are the same as when there are non-zero shocks. That is, F and K in eq.

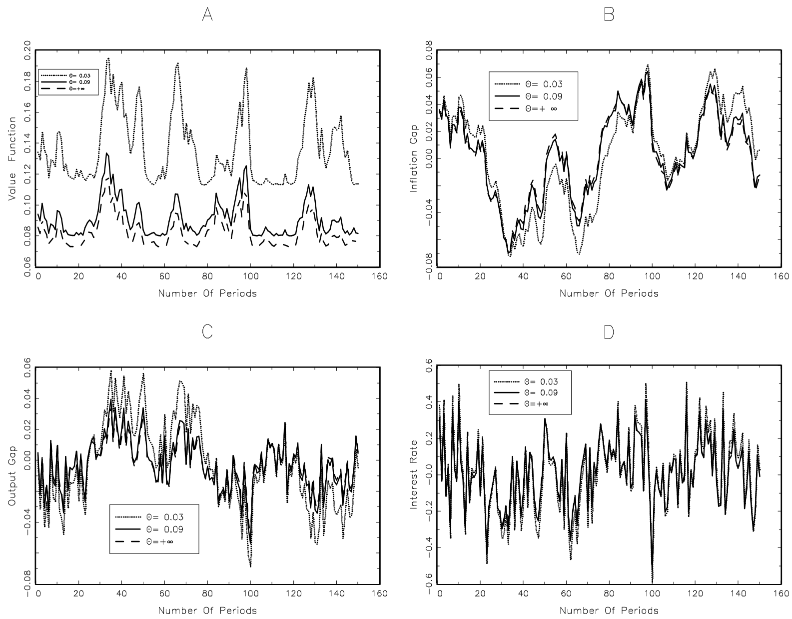

Figure 10: Simulation of the Robust Control with $\pi_0 = 0.02$ and $y_0 = 0.02$

(48) and (49) do not change no matter whether there are shocks or not. The difference lies in the value function. The simulations for zero shocks and with the same parameters as the case of non-zero shocks are shown in Figure 11. Figure 11A, 11B, 11C and 11D represent the paths of the value function, the state and control variables with different θ. In Figure 11 we find that the state variables converge to their equilibria zero as time tends to infinity no matter whether the robustness parameter is small or large. But in case the robustness parameter is small, the state variables evolve at a higher level and converge more slowly to zero than when the robustness parameter is large. The value function also evolves at a higher level in the case of small robustness parameters. It is interesting that the interest rate also converges to zero as time tends to infinity, this seems inconsistent with the fact that the nominal interest rate should be bounded by zero. This would not be surprising if we note that the interest rate turns out to be a linear function of the inflation and output gaps in the model. If we consider a long run equilibrium level of interest rate $\bar{R}$ as in the simple Taylor rule (Taylor 1993), the interest rate will converge to something around $\bar{R}$ rather than zero, though the state variables converge to zero. The simulations tell us that the larger the robustness parameter θ, the lower π_t, y_t, R_t and the value functions are, and moreover, the faster the state variables converge to their equilibria. And in case $\theta = +\infty$, the state variables reach their lowest values and attain the equilibria at the highest speed. This is consistent with the conclusion from the simulations with non-zero shocks.

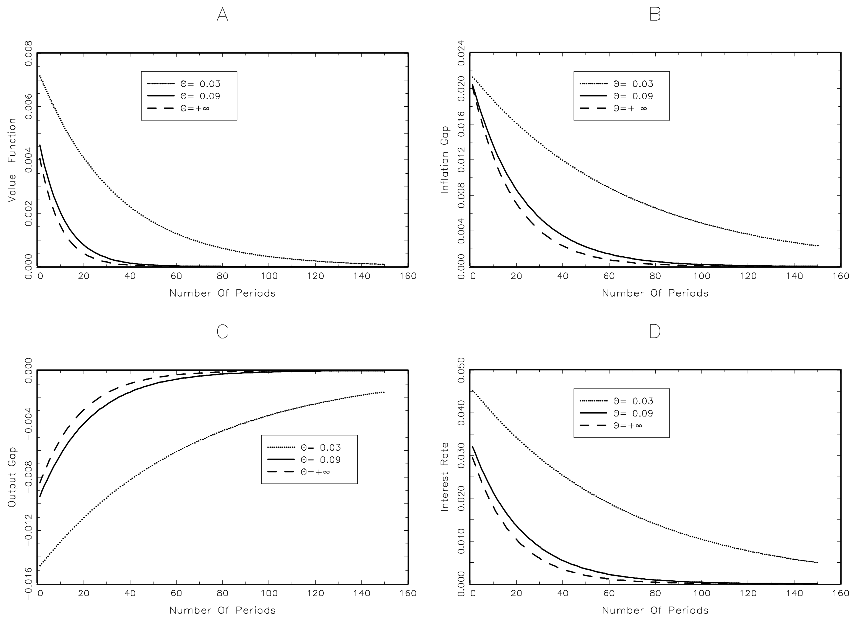

Figure 11: Results of the Robust Control with Zero Shocks

5 Monetary Policy Rules under Uncertainty: An Evaluation

We have explored two strategies of monetary policy making under uncertainty: Adaptive learning and robust control. The difference of the two strategies is clear. In the former case the central bank is assumed to improve its knowledge of economic models by learning from the information available. In the latter case, however, the central bank accepts model misspecification as a permanent state of affairs and directs its efforts to designing robust controls, rather than to using data to improve model specification over time. As mentioned before, while the adaptive learning considers mainly parameter uncertainty, the robust control considers not a specific kind of uncertainty and might deal with more general uncertainty than the adaptive learning does.

The simulations of these two strategies show much difference. In the learning strategy the state variables do not converge in both the linear and nonlinear Phillips curves even if the model is non-stochastic. With the robust control, however, the state variables may not converge in the stochastic model, but converge in the non-stochastic model.

We should, however, note two problems here. First, the robust control seeks a monetary policy rule in the so-called "worst case", which may not occur, and moreover, how to specify the "worst case" is a problem. The "worst case" and therefore the robust monetary policy rule are greatly influenced by the robustness parameter θ, as shown in our simulations. How to specify θ is a problem. As mentioned by Anderson, Hansen and Sargent (2000), if the robustness parameter is too small, it does not make sense to design a policy rule for a model which is very easy to tell apart. On the other hand, if θ is very large, the difference between the approximating model and the distorted model is very small and the robust rule may not be of much help. Although one can choose a robustness parameter with the help of the detection probability, the uncertainty is not really eliminated. In contrast to the robust control, the approach of learning, however, assumes agents reduce uncertainty through a specific algorithm of learning. The problem of this strategy is that there are many ways to specify how the parameters are learned. Different learning algorithms may lead to different policy rules. The learning algorithms discussed, for example, include the RLS, the Kalman filter, and the stochastic gradient learning. Second, in the empirical research of robust control we have assumed that the Phillips curve is linear. This leads to the convergence of the state variables in the non-stochastic model. The state variables may not converge if the nonlinear Phillips curve is used instead of the linear one.

Some researchers have even casted doubt on the strategy of robust control. Chen and Epstein (2000) and Epstein and Schneider (2001), for example, criticize the application of the robust control for time-inconsistency in preferences. Hansen and Sargent (2001b), therefore, discuss variants in which alternative representations of the preferences that underlie robust control theory are or are not time consistent. An important criticism of the robust control comes from Sims (2001a). He criticizes the robust control approach on conceptual grounds. The robust control imposes also neutrality properties of the model which will not be removed by better local approximations. There are major sources of a more fundamental type of uncertainties that the robust control theory does not address.[13] Sims (2001a) points out that more important uncertainties are ignored in such an approach. One major uncertainty is to what extent there is a medium run trade-off between inflation and output. Sims (2001a) shows that long run effects of inflation on output may not need to be completely permanent in order to be important. On the other hand, a deflation may have strong destabilizing effects while interest rates are already very low. Thus, there may in fact be a long-run non-vertical Phillips curve.[14] Yet, the robust control approach follows the neutrality postulate, implying a vertical long-run Phillips curve.

6 Conclusion

This paper is concerned with monetary policy rules under uncertainty. We first present the evidence of uncertainty using a State-Space model with Markov-Switching. Our empirical model using the U.S. data indicates that there have been regime changes in both parameters and shocks. We have considered not only the traditional IS and linear Phillips curve, but also a nonlinear Phillips curve.

[13]Moreover, steady states might not be optimal, if multiple steady states exist, see Greiner and Semmler (2002).

[14]See Graham and Snower (2002) and Semmler, Greiner and Zhang (2002a, ch. 7).

Based on the evidence of uncertainty in monetary policy, we explore two kinds of strategies to deal with uncertainty. The first strategy is adaptive learning of unknown parameters in models. Both linear and nonlinear Phillips curves are considered. In contrast to previous models with adaptive learning (see Sargent, 1999 and Orphanides and Williams, 2002), where LQ control models have been used, our simulations with a nonlinear optimal control model indicate that the state variables do not converge in either case, but the state variables have larger means and variances in the nonlinear Phillips curve than in the linear one.

The second strategy we have considered is the robust control which may deal with more general uncertainty than the adaptive learning. Using the robust control the central bank resorts to a monetary policy rule robust to uncertainty instead of learning. The empirical research indicates that the robustness parameter plays an important role in the policy design. It influences not only the means and variances of the state and control variables but also the speed of convergence. Yet, as Sims (2001a) has argued, other major sources of uncertainty should also be considered.

References

[1] Aguiar, Alvaro and Manuel M.F. Martins (2002), "Trend, cycle, and nonlinear trade-off in the Euro-area 1970-2001", CEMPRE.

[2] Akerlof, George A. (2002), "Behavior macroeconomics and macroeconomic behavior", *American Economic Review*, Vol. 92, pp411-433.

[3] Anderson, Evan W., Lars Peter Hansen and Thomas J. Sargent (2000), "Robustness, detection and the price of risk", manuscript.

[4] Baxter, Marianne and Robert G. King (2002), "Measuring business cycles: Approximate Band-Pass filters for economic time series", NBER working paper, No. 5022.

[5] Bean, Charles (2000), "The convex Phillips curve and macroeconomic policymaking under uncertainty", manuscript, London School of Economics.

[6] Beck, Günter W. and Volker Wieland (2002), "Learning and control in a changing economic environment", *Journal of Economics Dynamics and Control*, Vol. 26, pp1359-1377.

[7] Cagetti, Hansen, Sargent and Williams (2001), "Robustness and pricing with uncertain growth", manuscript, February.

[8] Chadha, B., P. Masson and G. Meredith (1992), "Models of inflation and the costs of disinflation", *IMF Staff Papers*, Vol. 39, No. 2, pp395-431.

[9] Chen, Z. and L.G. Epstein (2000), "Ambiguity, risk and asset returns in continuous time", manuscript.

[10] Clarida, Richard, Jordi Gali and Mark Gertler (1998), "Monetary policy rules in practice: Some international evidence", *European Economic Review*, Vol. 42, 1033-1067.

[11] − − − (1999), "The science of monetary policy: A new Keynesian perspective", *Journal of Economic Literature*, XXXVII, 1661-1707.

[12] Clements, Michael P. and Marianne Sensier (2002), "Asymmetric output-gap effects in Phillips curve and mark-up pricing models: Evidence for the U.S. and the U.K.", manuscript.

[13] Cogley, Timothy and Thomas J. Sargent (2001), "Evolving post-world war II U.S. inflation dynamics", manuscript, June.

[14] − − − (2002), "Drifts and volatilities: Monetary policies and outcomes in the post WWII U.S.", manuscript, March.

[15] Dupasquier, Chantal and Nicholas Ricketts (1998a), "Nonlinearities in the output-inflation relationship", manuscript.

[16] Dupasquier, Chantal and Nicholas Ricketts (1998b), "Nonlinearities in the output-inflation relationship: Some empirical results for Canada", Bank of Canada working paper 98-14.

[17] Epstein, L. and M. Schneider (2001), "Recursive multiple priors", manuscript.

[18] Evans, George W. and Seppo Honkapohja (2001), "Learning and Expectations in Macroeconomics", The Princeton University Press.

[19] Flaschel, Peter, Gang Gong and Willi Semmler (2001), "Nonlinear Phillips curve and monetary policy in a Keynesian macroeconometric model", working paper 18, CEM, University of Bielefeld.

[20] Giannoni, Marc P. (2000), "Does model uncertainty justify caution? Robust optimal monetary policy in a forward-looking model", manuscript, September.

[21] Giordani, Paolo and Paul Söderlind (2002), "Soluation of macromodels with Hansen-Sargent robust policies: Summary and some extensions", manuscript, February.

[22] Gómez, Javier and Juan Manuel Julio (2000), "An estimation of the nonlinear Phillips curve in Colombia", manuscript.

[23] Graham, Liam and Dennis J. Snower (2002), "The return of the long-run Phillips curve", Institute for the Study of Labor (IZA), discussion paper, No. 646.

[24] Greiner, Alfred and Willi Semmler (2002), "Monetary policy, non-uniqueness of steady states and hysteresis effects", in Reinhard Neck (ed.), *Modeling and Control of Economic Systems*, Elsevier Science Ltd, Oxford.

[25] Grüne, Lars (1997), "An adaptive grid scheme for the discrete Hamilton-Jacobi-Bellman equation", *Numerische Mathematik*, Vol. 75, 319-337.

[26] Grüne, Lars and Willi Semmler (2002), "Using dynamic programming with adaptive grid scheme for optimal control problems in economics", working paper, www.wiwi.uni-bielefeld.de/∼semmler/cem.

[27] Hansen, Lars Peter and Thomas J. Sargent (1999), "Robustness and commitment: A monetary policy example", manuscript, March.

[28] − − − (2001a), "Robust control and model uncertainty", manuscript, January.

[29] − − − (2001b), "Time inconsistency of robust control?", manuscript, October.

[30] − − − (2002), "Robust Control and Filtering for Macroeconomics", book manuscript.

[31] Harvey, A. C. (1989), "Time Series Models", Phillip and Allan Publishers Limited.

[32] Heinemann, Maik (2000), "Convergence of adaptive learning and expectational stability: The case of multiple rational-expectations equilibria", *Macroeconomic Dynamics*, Vol.4, 263-288.

[33] Isard, Peter (1999), Douglas Laxton and Ann-Charlotte Eliasson, "Simple monetary policy rules under model uncertainty", *International Tax and Public Finance*, Vol.6, 537-577.

[34] Kasa, Kenneth (1999), "Will the Fed ever learn?", *Journal of Macroeconomics*, Vol. 21, No 2, 279-292.

[35] Kiefer, Nicholas and Yaw Nyarko (1989), "Optimal control of a linear process with learning", *International Economic Review*, Vol. 30, No. 3, 571-87.

[36] Kim, Chang-Jin (1993), "Sources of monetary growth uncertainty and economic activity: The time-varying-parameter model with heteroskedastic disturbances", *Review of Economics and Statistics*, Vol. 75, 483-92.

[37] Kim, Chang-Jin and Charles R. Nelson (1999), "State-Space Models with Regime Switching", The MIT Press.

[38] Knoppik, Christoph (2001), "The long-run Phillips curve of Akerlof, Dickens and Perry (1996), and the role of limits to downward nominal wage rigidity", Dept. of Economics, University of Regensburg, Germany.

[39] Laxton, D., G.Rose and D. Tambakis (1997), "The U.S. Phillips curve: The case for asymmetry", manuscript.

[40] Ljung, Lennart (2000), "Recursive least squares and accelerated convergence in stochastic approximation schemes", manuscript, October.

[41] Martin, Ben and Chris Salmon (1999), "Should uncertain monetary policymakers do less?", manuscript, Bank of England.

[42] Meyer, Laurence H., Eric T. Swanson and Volker W. Wieland (2001), "NAIRU uncertainty and nonlinear policy rules", manuscript, January.

[43] $---$ (2000), "Robust monetary policy under model uncertainty: Incorporating rational expectations", manuscript, November.

[44] Orphanides, Athanasios and Simon van Norden (2002), "The unreliability of output gap estimation in real time", *The Review of Economics and Statistics*, 84(4): 569-583.

[45] Orphanides, Athanasios and John C. Williams (2002), "Imperfect knowledge, inflation expectations and monetary policy", manuscript, November.

[46] Razzak, Weshah (1997), "The inflation-output trade-off: Is the Phillips curve symmetric? A policy lesson from New Zealand", Reserve Bank of New Zealand.

[47] Rudebusch, Glenn D. and Lars E. O. Svensson (1999), "Policy rules for inflation targeting", in John B. Taylor, ed., *Monetary Policy Rules*, The University of Chicago Press.

[48] Sargent, Thomas J. (1999), "The Conquest of American Inflation", Princeton University Press.

[49] Schaling, Eric (1999), "The nonlinear Phillips curve and inflation forecast targeting", Bank of England.

[50] Semmler, Willi, Alfred Greiner and Wenlang Zhang (2002a), "The Macroeconomy and Monetary and Fiscal Policies in the Euro-Area", book manuscript, at www.wiwi.uni-bielefeld.de/∼Semmler/cem.

[51] $---$ (2002b), "Monetary policy rules in the Euro-area: Was it too tight in the 1990s?", *Atlantic Economic Journal*, Vol. 30, 283-297.

[52] Sims, Christopher A. (2001a), "Pitfalls of a minimax approach to model uncertainty", manuscript, Princeton university.

[53] $---$ (2001b), "Comment on Sargent and Cogley's "Evolving U.S. postwar inflation dynamics", manuscript, Princeton University.

[54] Sims, Christopher A. and Tao Zha (2002), "Macroeconomic switching", manuscript, Princeton University.

[55] Söderström, Ulf (1999), “Monetary policy with uncertain parameters”, manuscript, March.

[56] Svensson, Lars E.O. (1997), “Inflation forecasting targeting: implementing and monitoring inflation targets”, *European Economic Review*, Vol. 41, 1111-46.

[57] −−− (1999), “Inflation targeting: Some extensions”, *Scandinavian Journal of Economics*, Vol. 101(3), 337-61.

[58] − − − (2000), “Robust control made simple”, manuscript, October.

[59] Tambakis, Demosthenes N. (1998), “Monetary policy with a convex Phillips curve and asymmetric loss”, working paper 98/21, IMF.

[60] Taylor, John B. (1993), “Discretion versus policy rules in practice”, *Carnegie-Rochester Conference Series on Public Policy*, Vol. 39, 1993, 195-214.

[61] − − − (1999), “Historical analysis of monetary policy rules”, in Taylor, John B., ed., *Monetary Policy Rules*, Chicago, IL: University of Chicago Press.

[62] Tetlow, Robert J. and Peter von zur Muehlen (2001a), “Simplicity versus optimality: The choice of monetary policy rules when agents must learn”, *Journal of Economic Dynamics and Control*, Vol. 25, 245-79.

[63] − − − (2001b), “Avoiding Nash inflation: Bayesian and robust responses to model uncertainty”, Board of Governors of the Federal Reserve System, manuscript.

[64] Tucci, Marco, P. (1997), “Adaptive control in the presence of time-vaying parameters”, *Journal of Economic Dynamics and Control*, Vol. 22, 39-47.

[65] Wieland, Volker (2000), “Monetary Policy, parameter uncertainty and optimal learning”, *Journal of Monetary Economics*, Vol. 46, 199-228.