Macroeconomic Dynamics, **9**, 2005, 651–681. Printed in the United States of America.
DOI: 10.1017.S1365100505040332

MONETARY POLICY RULES UNDER UNCERTAINTY: EMPIRICAL EVIDENCE, ADAPTIVE LEARNING, AND ROBUST CONTROL

WENLANG ZHANG
Center for Empirical Macroeconomics, University of Bielefeld

WILLI SEMMLER
Center for Empirical Macroeconomics, University of Bielefeld
and
New School University

We first explore empirical evidence of parameter and shock uncertainties in a state-space model with Markov switching. The evidence indicates that uncertainties in the U.S. economy have been too great to accurately define monetary policy rules. We then explore monetary policy rules under uncertainty with two approaches: the RLS learning algorithm and robust control. The former allows the parameters to be learned for a given model. Yet, as our results of the RLS learning in a framework of optimal control indicate, the state variables do not necessarily converge even in a nonstochastic model. The latter, by permitting uncertainty with respect to model misspecification, allows for a broader framework. Our study on robust control shows that robust optimal monetary policy rules reveal a stronger response to fluctuations in inflation and output than when no uncertainty exists, implying that uncertainty does not necessarily require caution.

Keywords: State-Space Model, Markov Switching, Recursive Least-Squares Learning, Robust Control

1. INTRODUCTION

In the profession, it has increasingly been recognized that formal modeling of monetary policy faces great challenges because of many kinds of uncertainties such as model uncertainty, shock uncertainty, and data uncertainty. Recent studies dealing with these uncertainties include Isard et al. (1999), Martin and Salmon (1999), Svensson (1999), Wieland (2000), Meyer et al. (2001), Tetlow and von zur Muehlen (2001a), Giannoni (2002), Orphanides and Williams (2002), and Söderström (2002).[1] Those papers explore, mostly theoretically, how certain

Address correspondence to: Prof. Willi Semmler, Department of Economics, University of Bielefeld, P.O. Box 100131, 33501 Bielefeld, Germany; e-mail: wsemmler@wiwi.uni-bielefeld.de.

© 2005 Cambridge University Press 1365-1005/05 $12.00

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

kinds of uncertainties may affect the decisions of the central banks or regulatory agencies.[2]

Empirical work on parameter uncertainty and how to capture it by modeling and estimating parameter shifts can be found in Cogley and Sargent (2001), who study the inflation dynamics of the United States after WWII by way of Bayesian VAR with time-varying parameters without stochastic volatility. Yet, Sims (2001b) 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 in estimates 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 (2003) study the drifts and volatilities of the U.S. monetary policy after WWII through a Bayesian VAR with time-varying parameters and stochastic volatility and claim to have found regime changes. Semmler et al. (2005) 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.

Thus, given such evidence on parameter and shock uncertainties, economic agents (central banks, for example) may resort to different strategies: They may either reduce uncertainty by learning or just seek a policy robust to model uncertainty without learning. As to the former, among the important research, the work by Sargent (1993, 1999) has attracted much attention. Although Sargent (1999) explores monetary policy rules with adaptive learning in an optimal control framework, he assumes that once the uncertain parameter is updated, the government pretends that the updated parameters will govern the dynamics forever. His analysis is undertaken in the traditional LQ framework. With such an approach, there may exist convergence to an ergodic distribution in a stochastic model and convergence with probability 1 in a non stochastic model. However, as mentioned by Tetlow and von zur Muehlen (2001b), the above problem in Sargent's assumption represents an inconsistency in his adaptive learning mechanism.

Because of this problem in Sargent's approach, it is a challenge to explore the learning algorithm in an appropriate manner. Therefore, in this paper, we study monetary policy rules under uncertainty with adaptive learning by using a dynamic programming algorithm recently developed by Grüne (1997) and Grüne and Semmler (2004). Our approach differs from Sargent's in that we are able to endogenize the learning of uncertain parameters and explore the problem with a dynamic-programming algorithm with adaptive grids which can deal with nonlinear constraints.

As stated above, the alternative for monetary authorities is to resort to a monetary policy rule robust to uncertainty. This is a strategy different from adaptive learning. In this approach the central bank considers the economic model only as an approximation to another model that it cannot specify. With a so-called robustness parameter, it pursues a monetary policy rule allowing for a worst-case scenario. Whereas adaptive learning considers mainly parameter uncertainty, robust control might consider more general uncertainties. The robustness parameter, as stated by

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

Gonzalez and Rodriguez (2003), can be considered as a measure of uncertainty and may affect the robust monetary policy.

The important question concerning monetary policy under uncertainty is whether uncertainty requires caution. Brainard (1967), for example, proposes that parameter uncertainty should imply a more cautious policy. This argument has been supported by the recent research of Martin and Salmon (1999), who explore the monetary policy of the United Kingdom employing a VAR model with and without parameter uncertainty. They find that the optimal rule in the presence of parameter uncertainty implies a less aggressive path for official interest rates than when no parameter uncertainty is considered. However, other researchers—Gonzalez and Rodriguez (2003) and Giannoni (2002), for example—argue that uncertainty does not necessarily require caution. We also study this problem using robust control theory, obtaining results similar to those of Giannoni (2002).

The remainder of this paper is organized as follows: In the second section, we present empirical evidence on parameter and shock uncertainties in the IS and Phillips curves by way of a state-space model with Markov switching. In Section 3, we explore monetary policy rules under model uncertainty with adaptive learning. Section 4 explores monetary policy rules with robust control and Section 5 concludes the paper.

2. EMPIRICAL EVIDENCE OF UNCERTAINTY: A STATE-SPACE MODEL WITH MARKOV SWITCHING

Consider an economic model

$$\underset{\{\boldsymbol{u}_t\}_0^\infty}{\text{Min}}\, E_0 \sum_{t=0}^{\infty} \rho^t L(\boldsymbol{x}_t, \boldsymbol{u}_t), \tag{1}$$

subject to

$$x_{t+1} = f(\boldsymbol{x}_t, \boldsymbol{u}_t, \boldsymbol{\varepsilon}_t), \tag{2}$$

where ρ is the discount factor bounded between 0 and 1, $L(x_t, u_t)$ is a loss function of an economic agent (central bank, for instance), $\boldsymbol{x}_t$ is a vector of state variables, $\boldsymbol{u}_t$ is a vector of control variables, $\boldsymbol{\varepsilon}_t$ is a vector of shocks, and E_0 denotes the mathematical expectation operator upon the initial values of the state variables. This kind of model represents the basic monetary control model employed by Clarida et al. (1999), Svensson (1997, 1999), and Beck and Wieland (2002), for example, where the constraint equations are usually IS and 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, u, to satisfy (1). However, the question arising is whether the state equation (2) can be correctly specified with time-series estimates. 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 et al. (2004) derive an optimal monetary policy rule from an optimal control problem

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

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 derived optimal monetary policy rule may not be very reliable.

One can estimate time-varying parameters of the traditional Phillips curve with the following state-space model[3]:

$$\pi_t = \boldsymbol{v_t}'\boldsymbol{\beta_t} + \varepsilon_t, \qquad \varepsilon_t \sim N\left(0, \sigma_\varepsilon^2\right), \tag{3}$$

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

where π_t denotes the inflation rate and $\boldsymbol{v_t}$ is a vector of the lags of the inflation rate and output gap. $\boldsymbol{\beta_t}$ is a vector of time-varying parameters. Note that, in this model, it is assumed that the shocks have constant variance and only $\boldsymbol{\beta_t}$ is uncertain. Cogley and Sargent (2001) study the inflation dynamics of the United States after WWII by way of a Bayesian VAR 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 macroeconomic switching of the U.S. policy and find more evidence in favor of stable dynamics with unstable disturbance variance than of clear changes in model dynamics. Therefore, Cogley and Sargent (2003) 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 (2003) assume stochastic volatility. Therefore, in the research below, we assume that $\boldsymbol{\varepsilon_t}$ has state-dependent variance. This is similar to the assumption of Cogley and Sargent (2003). But unlike Cogley and Sargent (2003), 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 by Sims and Zha (2002) who assume that there are three states of economy. With such an assumption we can explore the probability of regime switching. Another advantage of the state-space model with Markov switching is that, as we will see below, it can explore not only parameter uncertainty but also shock uncertainty.

Following Kim and Nelson (1999), we simply assume that $\boldsymbol{\varepsilon_t}$ in (3) has two states of variance with Markov property,[4] namely,

$$\boldsymbol{\varepsilon_t} \sim N\left(0, \sigma_{\varepsilon,S_t}^2\right), \tag{5}$$

with

$$\sigma_{\varepsilon,S_t}^2 = \sigma_{\varepsilon,0}^2 + \left(\sigma_{\varepsilon,1}^2 - \sigma_{\varepsilon,0}^2\right)S_t, \qquad \sigma_{\varepsilon,1}^2 > \sigma_{\varepsilon,0}^2,$$

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

and

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

where $S_t = 0$ or 1 indicates the states of the variance of $\boldsymbol{\varepsilon_t}$ 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 explore uncertainty in the traditional IS and Phillips curves which have often been taken as constraints in an optimal control model such as (1)–(2). To reduce the dimension of the model, we estimate the 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 is the output gap, R_t is 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.[5] Let r_t denote the real interest rate, namely, $r_t = R_t - \pi_t$. The model can be rewritten in a state-space form as follows:

$$\Lambda_t = X_t\phi_t + \boldsymbol{\varepsilon_t}, \tag{8}$$
$$\phi_t = \bar{\Phi}_{S_t} + F\phi_{t-1} + \eta_t, \tag{9}$$

where Λ_t denotes the dependent variables (π_t, y_t) and X_t the independent variables (π_{t-1}, y_{t-1} and r_{t-1}); ϕ_t denotes the time-varying parameters $\alpha_{n,t}$ and $\beta_{n,t}$ ($n = 1, 2, 3$); $\bar{\Phi}_{S_t}$ ($S_t = 0$ or 1) is the drift of ϕ_t; and F is a diagonal matrix with constant elements to be estimated from the model; η_t has the distribution shown in equation (4); $\boldsymbol{\varepsilon_t}$ is now assumed to have the distribution presented in equation (5).[6] The details of the state-space model with Markov switching are presented in Appendix A.[7]

The estimation is undertaken with the quarterly U.S. data of 1964.1–2003.1. The inflation rate is measured by changes in the GDP deflator; the output gap is measured by the percentage deviation of the Industrial Production Index (IPI) from its fourth-order polynomial trend.[8] R_t is the Federal Funds rate. The data sources are OECD and IMF. The estimates of the hyperparameters are shown in Table 1; $\sigma_{\varepsilon\pi,1}$ (0.0021) is almost as large as twice of $\sigma_{\varepsilon\pi,0}$ (0.0011). The difference between $\sigma_{\varepsilon y,0}$ (0.0000) and $\sigma_{\varepsilon y,1}$ (0.0199) is even more significant. The difference between $\bar{\Phi}_{\alpha_2,0}$ (0.4704) and $\bar{\Phi}_{\alpha_2,1}$ (0.6406) is less significant than that between $\bar{\Phi}_{\alpha_3,0}$ (0.0502) and $\bar{\Phi}_{\alpha_3,1}$ (0.0021). The difference between $\bar{\Phi}_{\beta_3,0}$ (0.0873) and $\bar{\Phi}_{\beta_3,1}$ (-0.3364) is significant not only in magnitude and t-statistics but also in signs. Therefore, the estimation results confirm a state of economy with high volatility and a state with low volatility. The fact that all the elements of F are smaller

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

TABLE 1. Estimates of the hyperparameters in the time-varying Phillips and IS curves

Phillips curve			IS curve		
Parameter	Estimate	S.D.	Parameter	Estimate	S.D.
$\sigma_{\varepsilon\pi,0}$	0.0011	0.0021	$\sigma_{\varepsilon y,0}$	0.0000	0.0024
$\sigma_{\varepsilon\pi,1}$	0.0021	0.0018	$\sigma_{\varepsilon y,1}$	0.0199	0.0064
$\sigma_{\eta\alpha 1}$	0.0020	0.0013	$\sigma_{\eta\beta 1}$	0.0073	0.0007
$\sigma_{\eta\alpha 2}$	0.0434	0.0106	$\sigma_{\eta\beta 2}$	0.0000	0.0334
$\sigma_{\eta\alpha 3}$	0.0000	0.0069	$\sigma_{\eta\beta 3}$	0.0000	0.0196
$\bar{\Phi}_{\alpha_1,0}$	0.0080	0.0052	$\bar{\Phi}_{\beta_1,0}$	−0.0003	0.0010
$\bar{\Phi}_{\alpha_1,1}$	0.0005	0.0005	$\bar{\Phi}_{\beta_1,1}$	0.0023	0.0026
$\bar{\Phi}_{\alpha_2,0}$	0.4704	0.2183	$\bar{\Phi}_{\beta_2,0}$	0.5893	0.3372
$\bar{\Phi}_{\alpha_2,1}$	0.6406	0.2208	$\bar{\Phi}_{\beta_2,1}$	0.7272	0.3836
$\bar{\Phi}_{\alpha_3,0}$	0.0502	0.0337	$\bar{\Phi}_{\beta_3,0}$	0.0873	0.1881
$\bar{\Phi}_{\alpha_3,1}$	0.0021	0.0020	$\bar{\Phi}_{\beta_3,1}$	−0.3364	0.1809
f_{α_1}	0.3677	0.3710	f_{β_1}	0.7688	0.0810
f_{α_2}	0.3431	0.2205	f_{β_2}	0.2455	0.4093
f_{α_3}	0.8757	0.0644	f_{β_1}	−0.5817	0.2662
p	0.9875	0.0147	p	0.8451	0.1589
q	0.7586	0.2323	q	0.9591	0.0248
Likelihood	−660.6778		Likelihood	−472.5914	

than 1 in magnitude indicates that the time-varying parameters are stationary and, as a result, justifies the adoption of equation (9).

The paths of α_{2t} are shown in Figure 1A. We leave aside the paths of the intercepts in the IS and Philips curves just for simplicity. Let $\boldsymbol{Y_t}$ denote the vector of observations available as of time t. In Figure 1A, "Alpha_2t,0" and "Alpha_2t,1" denote the paths of α_{2t} when $[S_t = 0|Y_t]$ $(\alpha_{2t,0})$ and $[S_t = 1|Y_t]$ $(\alpha_{2t,1})$, respectively. "Alpha_2t" denotes the weighted sum 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" and "Alpha_3t,1" denote the paths of α_{3t} when $[S_t = 0|Y_t]$ $(\alpha_{3t,0})$ and $[S_t = 1|Y_t]$ $(\alpha_{3t,1})$ respectively. Similarly, "Alpha_3t" denotes the weighted sum of $\alpha_{3t,0}$ and $\alpha_{3t,1}$ (α_{3t}). Figure 1C represents the path of $\Pr[S_t = 1|Y_t]$. It is clear that the economy is probably in state 1 most of the time since $\Pr[S_t = 1|Y_t]$ is very high except at the beginning of the 1970s. Therefore, α_{2t} is very close to $\alpha_{2t,1}$ and α_{3t} is close to $\alpha_{3t,1}$ most of the time; α_{2t} experienced some significant changes between 1970 and 1975 and at the beginning of the 1980s; α_{3t} also experienced significant changes in the first half of the 1970s, increasing from about 0.02 to 0.20. We can

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

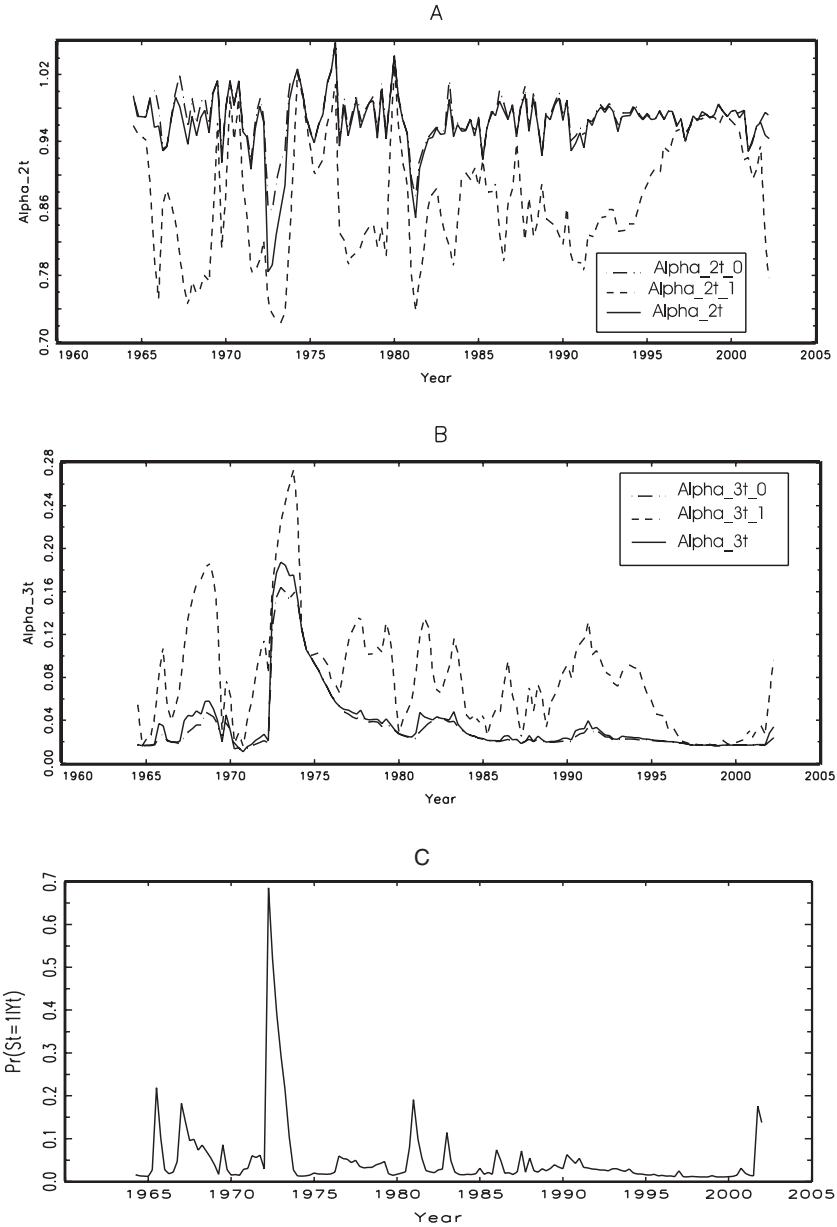


FIGURE 1. Empirical evidence of uncertainty in the Phillips curve.

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

also check these structural changes with the Chow breakpoint test of the following regression (t-statistics in parentheses):

$$\pi_t = \underset{(1.591)}{0.001} + \underset{(16.566)}{1.396}\,\pi_{t-1} - \underset{(2.197)}{0.316}\,\pi_{t-2} - \underset{(1.181)}{0.171}\,\pi_{t-3} + \underset{(0.805)}{0.066}\,\pi_{t-4} + \underset{(3.036)}{0.023}\,y_{t-1}$$

with $R^2 = 0.979$ and $DW = 2.026$. The F-statistics of the Chow breakpoint test for 1971.4 and 1981.4 are 2.325 and 4.712, respectively, larger than the critique value at the 5% level of significance, which is between 2.26 and 2.29, implying structural changes around these two points.

The estimation result of the IS curve is demonstrated in Figure 2. The paths of β_{2t} and β_{3t} are represented in Figure 2A and 2B, and Figure 2C, is the path of $\Pr[S_t = 1|Y_t]$. From Figure 2C, we find that the IS curve was probably in state 1 before the beginning of the 1980s and in state 0 afterward. Therefore, β_{2t} was close to $\beta_{2t,1}$ before 1982 and close to $\beta_{2t,0}$ afterward. The same is true of β_{3t}. Both β_{2t} and β_{3t} experienced some significant structural changes in the 1970s and small changes around 1990. It is also obvious that the time-varying parameters show some structural changes at the beginning of the 1980s, which coincides with the beginning of the post-Volcker period. Then, after that, for the period from 1984 to 1990, the time-varying parameters seemed to be relatively stable.[9]

In the preceding, we have explored parameter and shock uncertainties in the IS and Phillips curves with the U.S. data. The results are consistent with the line of research that maintains that there were regime changes in the U.S. economy.[10] 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

The question arising is: What is the optimal monetary policy rule in case some parameters or shocks in an economic model as presented by equations (1)–(2) are uncertain? Recently, numerous papers have been contributed to this problem. Svensson (1999), Tetlow and von zur Muehlen (2001a), Beck and Wieland (2002), Orphanides and Williams (2002), Söderström (2002), and McGough (2003), 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 approach is that the economic agents may learn the parameters using the Kalman filter. This approach has been pursued by Tucci (1997) and Beck and Wieland (2002). Another learning mechanism that is also applied frequently is the recursive least-squares (RLS) algorithm. 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 over time, using all information available.

Beck and Wieland (2002) and Orphanides and Williams (2002) have explored monetary policy with adaptive learning. Besides the difference in the learning

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

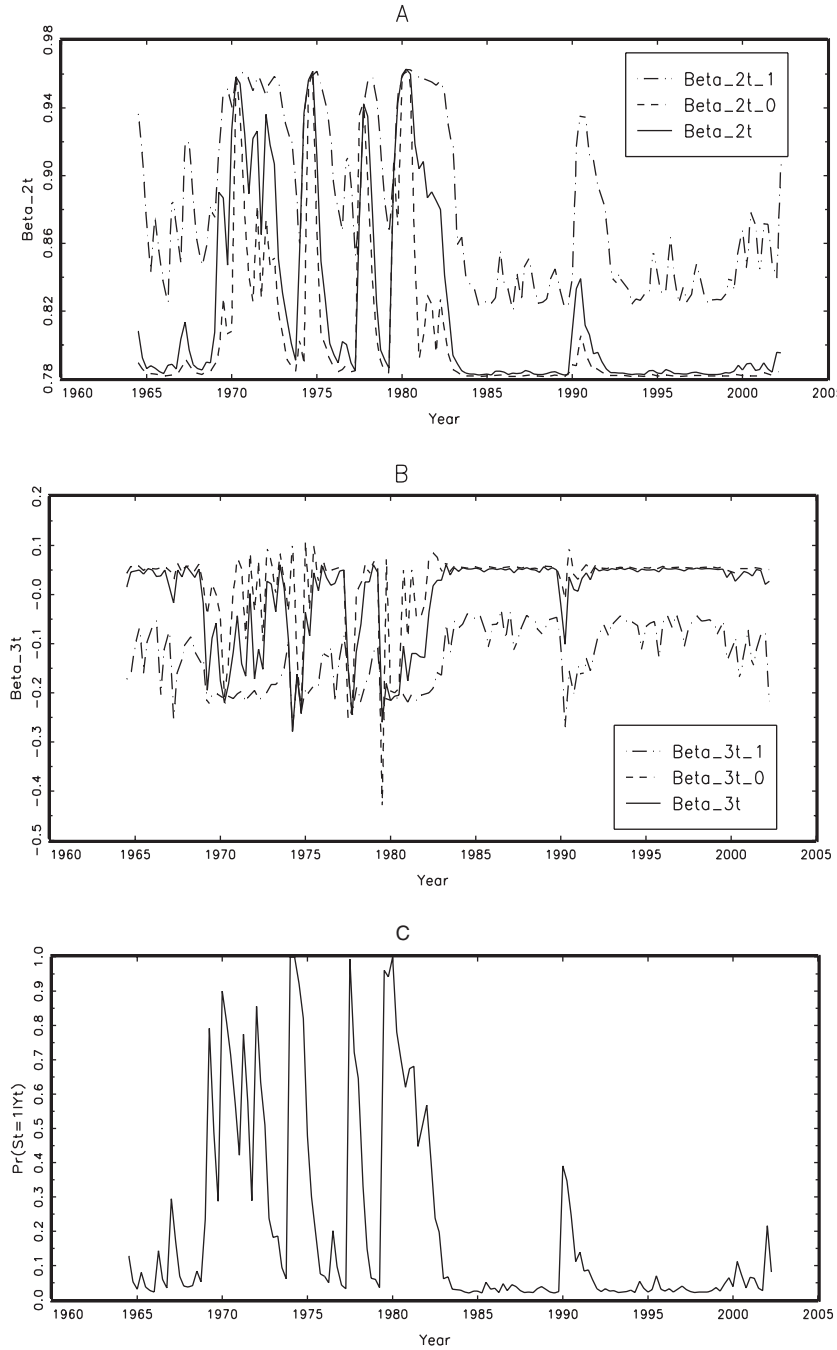


FIGURE 2. Empirical evidence of uncertainty in the IS curve.

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

algorithm they use, another difference between Beck and Wieland (2002) and Orphanides and Williams (2002) is that the former do not consider the role of expectations in the model, whereas 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.

Among the research of monetary policy with adaptive learning, the work of Sargent (1999) has attracted much attention. Sargent (1999) employs a learning algorithm as well as a discounted loss function, but in an LQ model. Yet, he constructs his results in two steps. First, assuming the RLS learning algorithm with a decreasing or constant gain, the agents estimate a model of the economy using the latest available data and update parameter estimates from period to period. Second, once the unknown parameter is updated, an optimal policy is derived from an LQ control model under the assumption that the updated parameter will govern the dynamics forever. As remarked by Tetlow and von zur Muehlen (2001b), however, the two steps are inconsistent with each other.

Because of this problem in Sargent's approach, it is required to explore such models by employing appropriate solution techniques. Therefore, in this section we will explore monetary policy with adaptive learning by using a recently developed dynamic programming algorithm.[11] To overcome the problem of Sargent (1999), we endogenize the changing parameters in a nonlinear optimal control problem and explore how our conclusion may be different from that of Sargent (1999).

Thus, 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 (1981) and Sargent (1999) prove that RLS is a specific form of the Kalman filter.

3.1. RLS Learning in a 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:

$$\pi_t = \gamma_1 \pi_{t-1} + \gamma_2 \pi_t^e + \gamma_3 y_t + \varepsilon_t, \qquad \varepsilon_t \sim i.i.d.\left(0, \sigma_\varepsilon^2\right), \tag{10}$$

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

where π_t^e denotes the agents' (including the central bank) expected inflation rate based on the time $t-1$ information, $\gamma_1, \gamma_2 \in (0, 1)$, $\gamma_3 > 0$; ε_t is a serially uncorrelated innovation. We further assume that the IS equation reads as:[12]

$$y_t = -\theta r_{t-1}, \qquad \theta > 0, \tag{11}$$

where r_t denotes the deviation of the short-term real interest rate from its equilibrium level. Substituting equation (11) into (10), we have

$$\pi_t = \gamma_1 \pi_{t-1} + \gamma_2 \pi_t^e - \gamma_3 \theta r_{t-1} + \varepsilon_t. \tag{12}$$

In the 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

$$\bar{a} = \frac{\gamma_1}{1-\gamma_2}, \qquad \bar{b} = -\frac{\gamma_3 \theta}{1-\gamma_2}.$$

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

$$\pi_t = \bar{a}\pi_{t-1} + \bar{b} r_{t-1} + \varepsilon_t. \tag{13}$$

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 $i = 0, \ldots, t-1$. 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

$$\pi_t^e = a_{t-1}\pi_{t-1} + b_{t-1} r_{t-1}. \tag{14}$$

The standard least-squares formula gives the equations

$$\begin{pmatrix} a_t \\ b_t \end{pmatrix} = \left(\sum_{i=1}^{t} z_i' z_i \right)^{-1} \left(\sum_{i=1}^{t} z_i' \pi_i \right), \tag{15}$$

where $z_i = (\pi_{i-1} \quad r_{i-1})'$.

Defining

$$c_t = \begin{pmatrix} a_t \\ b_t \end{pmatrix},$$

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

we can also compute equation (15) using the stochastic approximation of the recursive least-squares equations

$$c_t = c_{t-1} + \kappa_t V_t^{-1} z_t (\pi_t - z_t' c_{t-1}), \tag{16}$$

$$V_t = V_{t-1} + \kappa_t (z_t z_t' - V_{t-1}), \tag{17}$$

where c_t and V_t denote respectively, the coefficient vector and the moment matrix for z_t using data $i = 1, \ldots, t$; κ_t is the gain. To generate the least-squares values, one must set the initial value for the recursion.[13] 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 $\bar{c} = \left(\frac{\bar{a}}{\bar{b}}\right)$ 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 nonstochastic and κ_t sufficiently small, π_t^e can converge 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 implies that the agents believe the Phillips curve might exhibit structural changes and they allocate larger weights to the recent observations of the inflation rate than to the earlier ones. Orphanides and Williams (2002) denote the case of $\kappa_t = 1/t$ as infinite memory and the case of constant κ_t as finite memory. Following Svensson (1997, 1999) we assume that the central bank pursues a monetary policy by minimizing a quadratic loss function. The problem reads as

$$\underset{\{r_t\}_0^\infty}{\text{Min}} E_0 \sum_{t=0}^{\infty} \rho^t L(\pi_t), \qquad L(\pi_t) \equiv (\pi_t - \pi^*)^2, \tag{18}$$

subject to equations (12), (14), (16), and (17); π^* is the target inflation rate, which is assumed to be zero.[14]

As mentioned above, if the unknown parameters are adaptively estimated with RLS learning with a small and constant gain, they will converge in distributions in a stochastic model and converge with probability 1 in a nonstochastic model. However, an optimal control problem such as (18) with nonlinear state equations, embedded in equations (12), (14), (16), and (17), is difficult to solve and, using appropriate solution techniques, indicates that the model will not necessarily converge even if the state equations are nonstochastic.

Next, we undertake an appropriate numerical study of the model. Though the return function is quadratic and the Phillips curve is linear, the problem falls outside the scope of LQ optimal control problems since some parameters in the Phillips curve are time varying and follow nonlinear paths. The problem cannot be solved analytically, and numerical solutions have to be employed. In the numerical

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

study below, we resort to the algorithm developed by Grüne (1997), who applies adaptive rather than uniform grids.[15] The numerical study is undertaken for the deterministic case. To simplify the numerical study, we assume that $a_t = \bar{a}$ and 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, p. 1361), the reason for focusing on the unknown parameter b is that "this parameter is multiplicative to the decision variable r_t and therefore central to the trade-off between current control and estimation."

3.1.1. Numerical Study. In the numerical study, undertaken by the above-mentioned dynamic programming algorithm, 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 Figures 3A–D, respectively. Figure 3E is the phase diagram of π_t and r_t. Neither the state variables nor the control variable converges. In fact, they fluctuate cyclically. We explore solution paths with many different initial values of the state variables and smaller κ_t (0.01, for example) and find that in no cases 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 of the stochastic version is

$$\pi_t = \pi_{t-1} - 0.33 r_{t-1} + \varepsilon_t. \tag{19}$$

In the case of RLS learning, however, we have

$$\pi_t = \pi_{t-1} + \tilde{b}_t r_{t-1} + \varepsilon_t, \tag{20}$$

with

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

If there is perfect knowledge with 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 nonstochastic model. We define this case as the benchmark model. The results of the benchmark model are shown in Figure 3A and 3D (dashed line). Note that the benchmark model contains only one state variable, namely π_t, with dynamics denoted by (19). In this benchmark case, it is clear that the reaction function should be a linear function of the state variables. In the nonstochastic benchmark model, we find that the optimal monetary policy rule turns out to be $r_t = 3.00\pi_t$ and, substituting this into the deterministic version of equation (19) generates the optimal trajectory of π_t, $\pi_t = 0.01\pi_{t-1}$. It is obvious that in the steady state, we have $\pi_t = \pi_{t-1} = 0$. From Figure 3A and 3D, we observe that π_t and r_t converge to zero over time in the benchmark model.

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

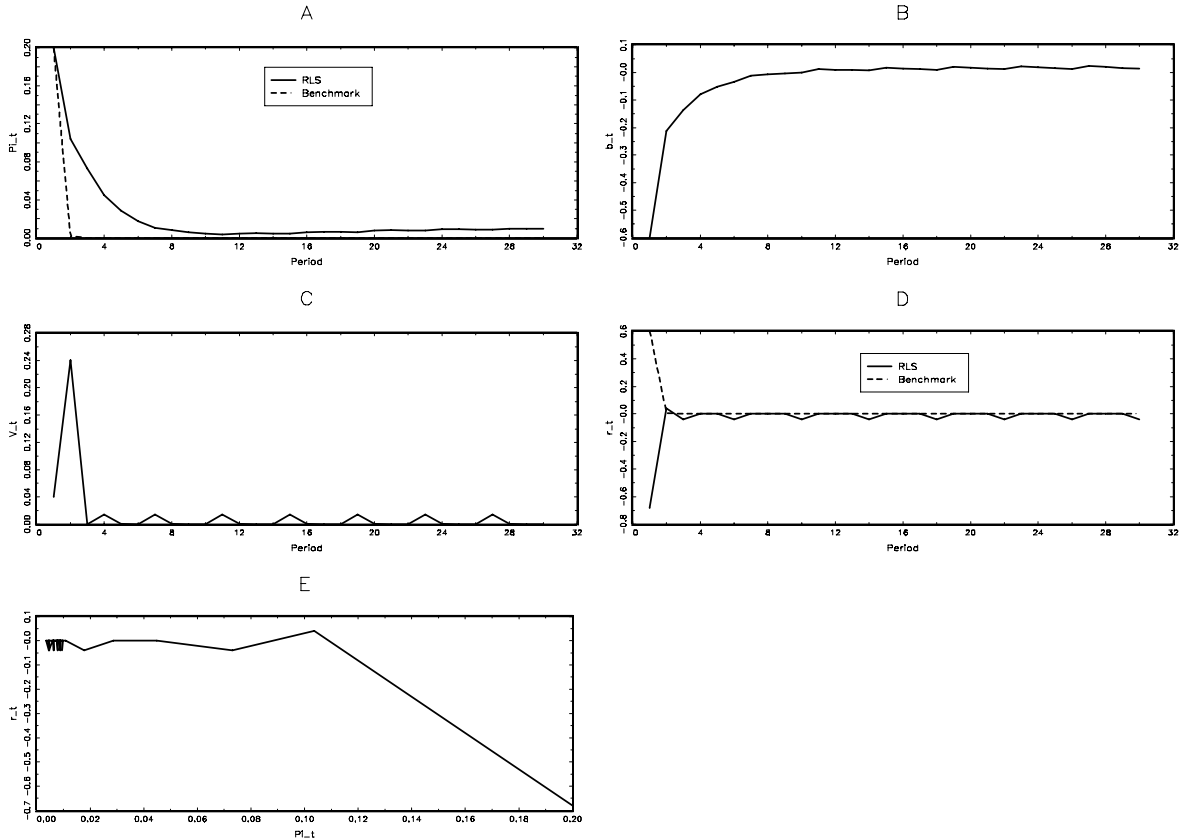


FIGURE 3. Simulations of RLS learning (solid) and benchmark model (dashed) with linear Phillips curve.

3.2. RLS Learning in Nonlinear Phillips Curve

We have explored adaptive learning in the context of a linear Phillips curve. The recent literature, however, argues that the Phillips curve might be nonlinear. Therefore, we next explore adaptive learning with a nonlinear Phillips curve.

As discussed by Aguiar and Martins (2002), there may be different forms of nonlinearity in the Phillips curve. Laxton et al. (1998) explore alternative shapes (concave, linear, and convex) of the U.S. Phillips curve and argue that the Fed should assume the convex form. More specifically, according to this literature, negative deviations of aggregate demand from potential are less disinflationary than positive deviations are inflationary. Therefore, we just follow Schaling (2002) and assume that the nonlinear form of the output gap in the Phillips curve reads as[16]

$$f(y_t) = \frac{\alpha y_t}{1 - \alpha\beta y_t}, \qquad \alpha > 0, 1 > \beta \geq 0, \tag{21}$$

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. Given the above nonlinear Phillips curve, equation (10) reads as

$$\pi_t = \gamma_1 \pi_{t-1} + \gamma_2 \pi_t^e + \gamma_3 f(y_t) + \varepsilon_t, \tag{22}$$

with $f(y_t)$ given by equation (21). Substituting equation (11) into equation (21), and then (21) into (22), we get the following nonlinear Phillips curve

$$\pi_t = \gamma_1 \pi_{t-1} + \gamma_2 \pi_t^e - \gamma_3 g(r_{t-1}) + \varepsilon_t, \tag{23}$$

where

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

The REE turns out to be

$$\pi_t = \bar{a}\pi_{t-1} + \bar{b} g(r_{t-1}) + \varepsilon_t, \tag{24}$$

where $\bar{a}$ is the same as in the preceding subsection, but $\bar{b}$ is changed to be $-\gamma_3/(1 - \gamma_2)$. The forecast of the inflation rate is now given by

$$\pi_t^e = a_{t-1}\pi_{t-1} + b_{t-1} g(r_{t-1}). \tag{25}$$

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

$$z_i = [\pi_{i-1} \quad g(r_{i-1})]'.$$

The optimal control problem (18) now turns out to have constraints (23), (25), (16), and (17).

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

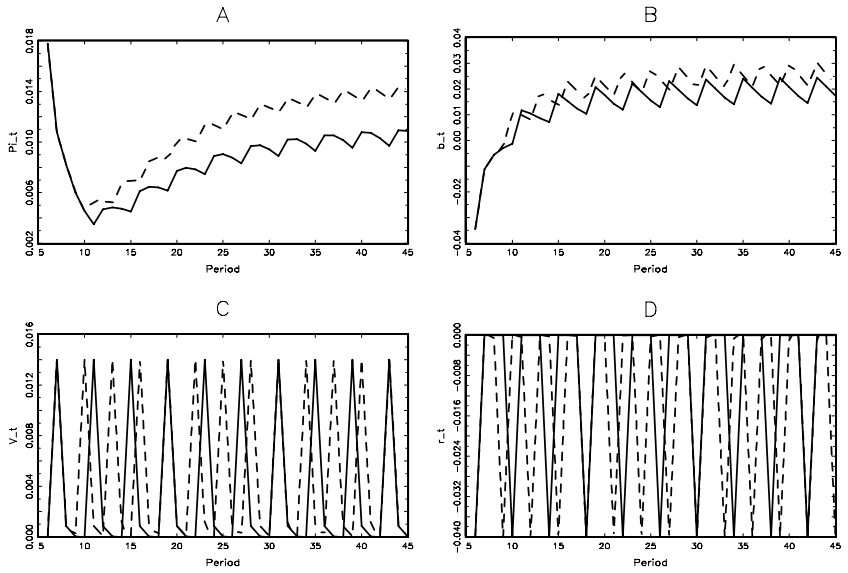


FIGURE 4. RLS learning with linear (solid) and nonlinear Phillips curves (dashed).

3.2.1. Numerical Study. In this version, again using the aforementioned dynamic programming algorithm, we assume $\alpha = 10$ and $\beta = 0.99$. The solution paths with the same starting values of the state variables as in the preceding subsection are presented in Figure 4A–D. Figure 4A represents the path of π_t, 4B the path of b_t, 4C the path of V_t, and 4D the path of r_t. The results of this subsection (nonlinear Phillips curve) are presented by dashed lines, whereas the results from the preceding subsection (linear Phillips curve) are indicated by solid lines.[17]

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. Experiments with many different initial values of state variables were undertaken and in no cases are the state variables found to converge. However, the difference between the simulations with linear and nonlinear Phillips curves cannot be ignored. Figure 4 indicates that both π_t (Figure 4A) and b_t (Figure 4B) 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 experiments are shown in Table 2. The standard deviation (S.D.) and magnitudes of the mean of these variables are larger in the case of the nonlinear Phillips curve than in the case of the linear one. As in the preceding subsection, the solutions are explored with different γ_1 and γ_2 and the results are similar.

The fact that the inflation rate has a higher mean and experiences larger changes in the nonlinear Phillips curve than in the linear one seems to be consistent

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

TABLE 2. Mean and standard deviation of state and control variables.

	π_t		b_t		V_t		r_t	
	L[a]	NL[b]	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

[a] L = linear Phillips curve.
[b] NL = nonlinear Phillips curve.

with the research of Tambakis (1998) who analyzes the single-period Barro-Gordon optimal monetary problem with a convex Phillips curve and an asymmetric loss function. Tambakis (1998) finds that both symmetric and asymmetric loss functions with a convex Phillips curve yield a positive expected inflation bias. The reason is that, unlike the case of a quadratic loss function with a linear Phillips curve, in the case of a quadratic loss function with a convex Phillips curve, $E\pi = \pi^*$ does not necessarily satisfy the first-order condition. Therefore, inflation rate does not necessarily equal its target in equilibrium. This problem has also been analyzed by Graham and Snower (2002) and Karanassou and Snower (2003). They argue that the long-run Phillips curve might be nonvertical and the inflation rate might be different from its target in equilibrium, depending on the extent to which economic agents are forward or backward looking.

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. This is different from the work of Sargent (1999) because the state variables can converge in such a nonstochastic model with his approach. The problem in Sargent (1999), as mentioned before, is that he employs two assumptions which turn out to be inconsistent with each other. This is because he explores the problem in a traditional LQ framework, which fails to accurately endogenize the uncertain parameter. Our results are, however, consistent with the claim of Beck and Wieland (2002), who also find that the state variables do not converge even if all shocks in the stochastic model are set to zero, because economic agents never stop experimentation in each period.

4. MONETARY POLICY RULES WITH ROBUST CONTROL

A disadvantage of adaptive learning analyzed in the preceding section is that we have considered only parameter uncertainty. Monetary authorities have to guard against uncertainties such as model misspecification in a more general way. For this purpose, robust control theory has been applied. Robust control induces the economic agents to seek a strategy for the "worst case" and can deal with more general uncertainties than adaptive learning does. On the basis of some earlier papers [see Hansen and Sargent (1999, 2001a,b)], Hansen and Sargent (2002) explore robust control in macroeconomics in details. Cagetti et al. (2002)

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

also employ robust control in macroeconomics. Svensson (2000) applies the idea of robust control in a simpler framework. Giordani and Söderlind (2003) and Giannoni (2002), however, extend the use of robust control theory by considering forward-looking behaviors.

The interesting question concerning monetary policy under uncertainty is whether the central bank should show a stronger or weaker response to the fluctuations of economic variables than when no uncertainty exists. Brainard (1967), for example, proposes that parameter uncertainty should incur a more cautious policy. This argument has been supported by recent research by Martin and Salmon (1999) who explore the monetary policy of the United Kingdom employing a VAR model with and without parameter uncertainty. They find that the optimal rule in the presence of parameter uncertainty incurs a less aggressive path for official interest rates than when no parameter uncertainty is considered. Other researchers, Gonzalez and Rodriguez (2003) and Giannoni (2002), for example, however, argue that uncertainty does not necessarily require caution. Following Hansen and Sargent (2002), in this section, we explore this problem with robust control and take up the problem of the control bank's response under uncertainty with respect to model misspecification.[18]

The research undertaken below is based on the framework of Hansen and Sargent (2002). A short introduction to robust control theory is given in Appendix B. The so-called robustness parameter θ plays an important role in the problem's solution. If θ is $+\infty$, the problem collapses to the traditional optimal control without model misspecification. Gonzalez and Rodriguez (2003) explore how the robustness parameter θ affects the control variable and prove that in a one-state, one-control model, the response is characterized by a hyperbolic function with a discontinuity at $\underline{\theta}$. Such a response presents a concave shape on the right side of the discontinuity and a convex one on the left.

Related to the above is the concept of a detection-error probability. It is a statistic concept that can tell us how difficult it is to tell the approximating model apart from the distorted one. The larger the detection-error probability, the more difficult it is to tell the two models apart. The design and interpretation of the detection-error probability are shown in Appendix B.

Semmler et al. (2005) estimate the IS and Phillips curves with the U.S. quarterly data for the period 1961–1999.[19] Next, we undertake some simulations of the robust control with the parameters estimated by Semmler et al. (2005). 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 have

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

and $u_t = r_t$ with π_t, y_t, and r_t being the inflation rate, output gap, and deviation of the interest rate from its target (assumed to be zero in the simulation below),

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

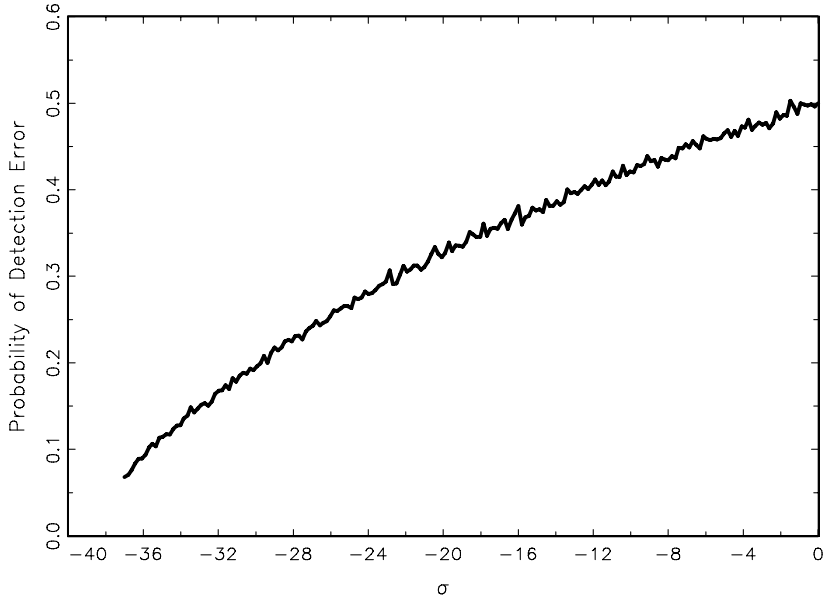


FIGURE 5. Detection error probability.

respectively. The problem then turns out to be

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

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 $C = \left(\begin{smallmatrix} 0.01 & 0 \\ 0 & 0.01 \end{smallmatrix}\right)$, the detection error probability is shown in Figure 5.[20] If we want a detection-error probability of about 0.15, $\sigma = -33$; that is, $\theta = 0.03$.

With $\theta = 0.03$, we have

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

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

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

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

TABLE 3. Standard deviations of the state and control variables with different θ.

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

and $V(\pi, y) = 11.134\pi^2 + 1.022y^2 + 0.945\pi y + 0.080$. In case $\theta = +\infty$,[21] we have $F = (6.438 \;\; 11.929)$ and $V(\pi, y) = 10.120\pi^2 + 1.020y^2 + 0.850\pi y + 0.073$. Comparing the elements in F with different θ, we find that the lower the θ is, the higher the coefficients of the inflation and output gaps in the reaction function. 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. Consistent with the conclusion of Gonzalez and Rodriguez (2003), this implies that more uncertainty with respect to model misspecification requires a stronger response of the control variable. Our numerical results are therefore consistent with the result of Giannoni (2002), who also shows that uncertainty does not necessarily require caution in a forward-looking model with robust control.

Next, we present the paths of the inflation-rate deviation, the output gap, and the interest rate with different θ in Figure 6A–C.

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 3, which indicates that the standard deviations of the state and control variables increase if θ decreases.[22]

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 nonzero shocks. That is, F and K in equations (B.15) and (B.16) do not change 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 nonzero shocks are shown in Figure 7A–C. In Figure 7, we find that the state variables converge to their equilibria zero whether the robustness parameter is small or large. However, 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 simulations indicate that the larger the robustness parameter θ, the lower π_t, y_t, and r_t and, moreover, the faster the state variables converge to their equilibria. In case $\theta = +\infty$, the state variables reach their lowest values and attain the equilibria at the highest speed.

In sum, we have shown that uncertainty with respect to model misspecification might not necessarily require caution. Though robust control can deal with problems that cannot be solved with the classical optimal control

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

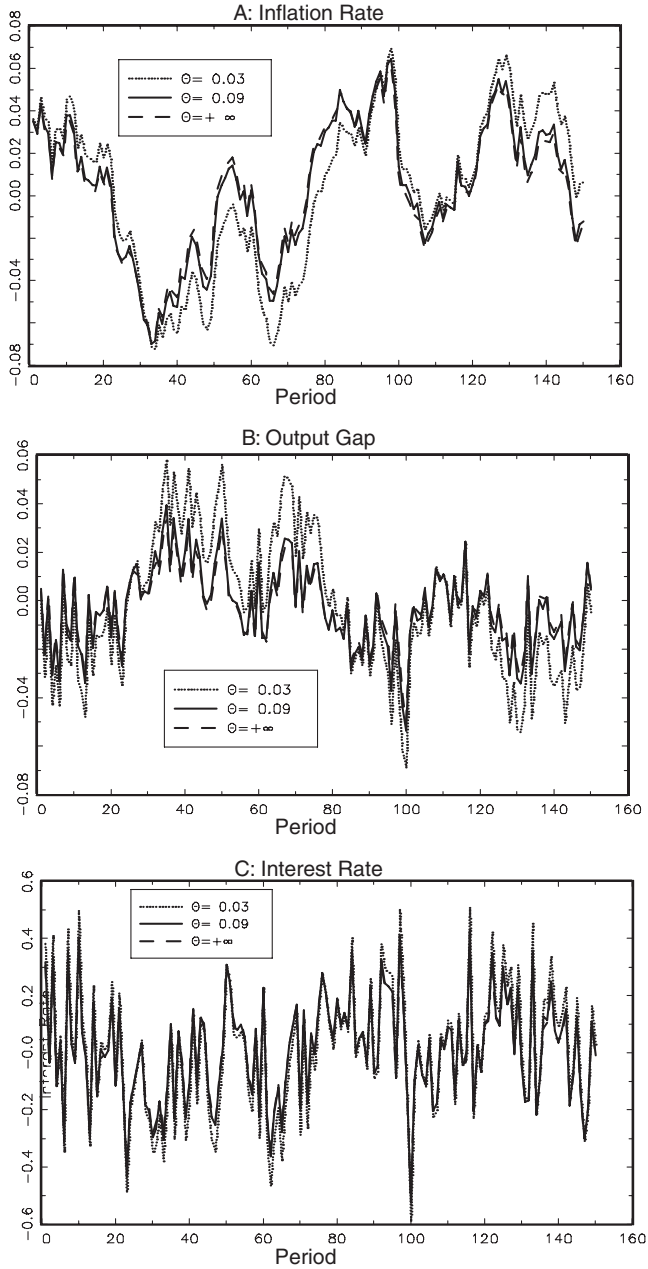


FIGURE 6. Simulation of the robust control with $\pi_0 = 0.02$ and $y_0 = 0.02$.

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

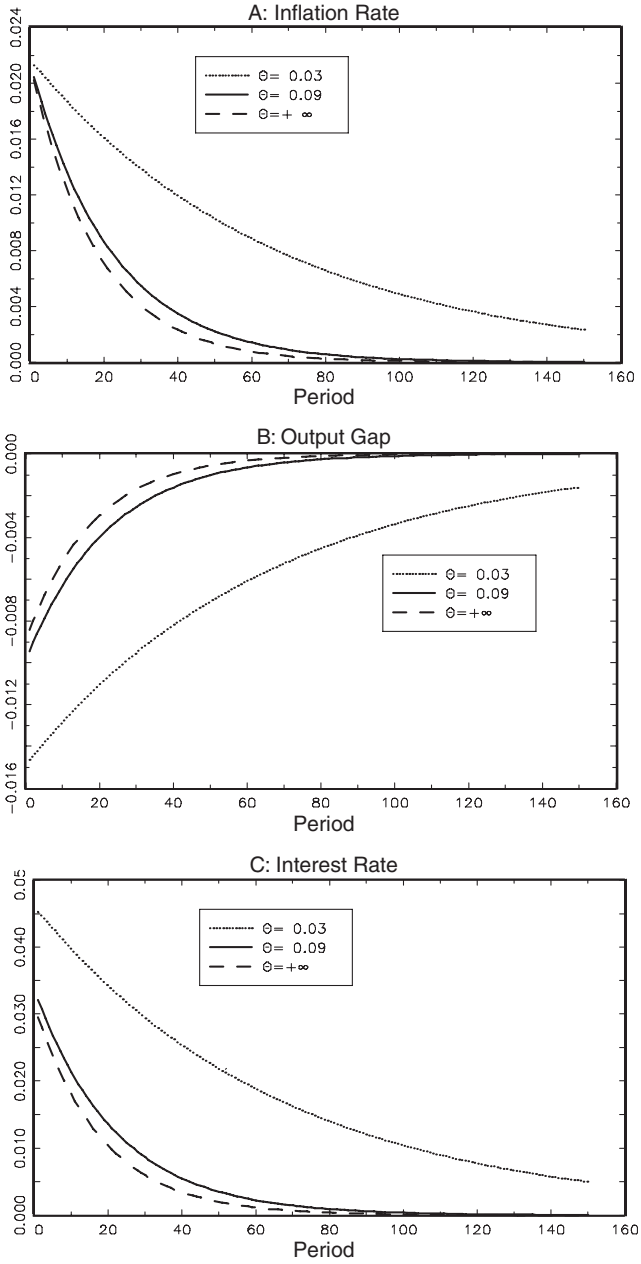


FIGURE 7. Results of the robust control with zero shocks.

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

theory, some researchers have casted doubt on 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. Therefore, Hansen and Sargent (2001b) discuss the time consistency of the alternative representations of preferences that underlie the robust control theory. An important criticism of robust control comes from Sims (2001a). He criticizes the robust control approach on conceptual grounds. Robust control also imposes neutrality properties of monetary policy implied in 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.[23] 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 nonvertical Phillips curve.[24] Yet, the robust control approach developed so far seems to follow the neutrality postulate, implying a vertical long-run Phillips curve.

5. CONCLUSION

This paper is concerned with monetary policy rules under uncertainty. We first presented some evidence of uncertainty using a state-space model with Markov switching. The empirical model using the U.S. data indicates that there have been regime changes in both parameters and shocks. On the basis of this empirical evidence, we have then explored two approaches to deal with monetary policy under uncertainty: adaptive learning and robust control. In the former case the central bank is assumed to improve its knowledge of economic models by learning from the information available. Whereas the adaptive learning mainly considers parameter uncertainty, robust control admits more general uncertainties.

As to adaptive learning, in contrast to Sargent (1999), who explores monetary policy with adaptive learning in a two-step decision process, we have endogenized the uncertain parameter and employed a dynamic programming algorithm with adaptive grids, which can deal with nonlinear state equations, so that we have solved the model appropriately. With such an approach, we were able to overcome the problem of inconsistency in the two-step decision process of Sargent (1999). Yet, different from the work of Sargent (1999), we show that, when the learning of coefficients is fully endogenized, the state variables do not necessarily converge even in a nonstochastic model with adaptive learning.

As regards robust control, which is a more general approach to guard policy against uncertainty, we have focused on model misspecification and shock uncertainty and explored whether uncertainty requires caution. Different from the common view that monetary policy under uncertainty should be more cautious as compared to no uncertainty, our research indicates that the robust policy rule may

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

respond more strongly to the economic variables in the presence of uncertainty and therefore implies that uncertainty does not necessarily require caution.

Finally, we note that there might exist other kinds of uncertainties, among which the data uncertainty has attracted much attention. The usually discussed data uncertainty is concerned with the output gap or Nairu uncertainty. A discussion of the unreliability of output gap estimated with real-time data can be found in Orphanides and van Norden (2002). The effect of data uncertainty on monetary policy may be different from the effects of model and shock uncertainties.[25] In contrast to our results on robust control, in the research on data uncertainty it is frequently found that the central bank should respond with greater caution to a variable estimated with error than it would in the absence of data uncertainty.[26] More research on data uncertainty and more succinct results are surely expected to be forthcoming.

NOTES

1. Several contributions to this problem can also be found in a recent issue of this journal [Giannoni (2002), Onatski and Stock (2002)].

2. For a study of the effect of model uncertainty in the context of ecological management problem, see Brock and Xepapadeas (2003).

3. For further details of such a model, see Semmler et al. (2005).

4. Vázquez (2003) estimates an augmented Taylor rule with a Markov-switching VAR model with the U.S. data of 1967–2002 and finds that there is no essential difference between the models with two regimes and three regimes and therefore, the model with two regimes can describe the economy well.

5. Forward-looking behaviors have been frequently taken into account in the Phillips curve. A survey of this problem can be found in Clarida et al. (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. A justification of the above type of backward-looking model can be found in Rudebusch and Svensson (1999).

6. Theoretically, the elements of F and the variance of η_t also may have a 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 nonstationary, the transition equation should be altered to be the form of eq. (4).

7. 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 subsequently is based on the Gauss programs developed by Kim and Nelson (1999).

8. The IPI has also been used by Clarida et al. (1998) to measure the output for Germany, France, the United States, the United Kingdom, Japan, and Italy. Clarida et al. (1998) use the quadratic trend of the IPI to measure the potential output for the United States of the period 1979–1993. We use the fourth-order trend of the IPI as the potential output because we cover a much longer period and the fourth-order trend fits the data better than the quadratic one. As surveyed by Orphanides and van Norden (2002), there are many approaches to measure the output gap.

9. The differences of the monetary policy rules of the United States across periods have also been explored by Clarida et al. (2000).

10. See Cogley and Sargent (2001, 2003) for example.

11. Evans and Honkapohja (2001) analyze expectations and learning mechanisms in macroeconomics in detail.

12. This is the same as Orphanides and Williams (2002), except that they assume a noise in the equation.

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

13. Evans and Honkapohja (2001, Ch. 2, p. 33) p. n. 4 explain how to set the initial values of c_t and V_t as follows. Assuming $Z_k = (z_1, \ldots z_k)'$ is of full rank and letting π^k denote $\pi^k = (\pi_1, \ldots, \pi_k)'$, the initial value c_k is given by $c_k = Z_k^{-1}\pi^k$ and the initial value V_k is given by $V_k = k^{-1}\sum_{i=1}^{k} z_i z_i'$.

14. To simplify the problem, we assume strict inflation targeting; that is, the central bank is concerned only with the inflation stabilization.

15. A less technical description and economic applications of this algorithm can be found in Grüne and Semmler (2004).

16. Note that this function is discontinuous at $y_t = 1/(\alpha\beta)$. When $y_t < 1/(\alpha\beta)$, $f''(y_t) > 0$ and if $y_t > 1/(\alpha\beta)$, $f''(y_t) < 0$. In the research below, we choose appropriate values of α and β so that with the U.S. output gap data we have $f''(y_t) > 0$. In fact, this function can be shown to be a particular *rational polynomial*, as discussed by Escribano (2004). Escribano (2004) shows that any nonlinear function can be approximated by rational polynomials by employing Padé's approximants. Escribano (2004) also discusses some new rational-polynomial nonlinear error-correction models. The reader is referred to Escribano (2004, p. 81) for the details of rational polynomials and Padé's approximants.

17. To see the differences of the simulations clearly, we present only the results from $t = 6$ onward.

18. Brock et al. (2003) and Brock and Xepapadeas (2003) also explore policy in an uncertain economic environment. Whereas they have discussed how changes of parameters in a model may affect the policy response, here we discuss how the so-called robustness parameter may influence the policy rules.

19. The data sources are OECD and IMF. The output gap is measured by the percent deviation of the IPI from its HP-filtered trend. The inflation rate is measured by the changes in the CPI, and the interest rate is the Federal Funds rate.

20. The numerical studies in this section are done with the algorithms developed by Hansen and Sargent (2002). In the computation of the detection-error probability, T (number of periods) is 150 and 5,000 simulations are undertaken here.

21. In the simulation, we just take $+\infty = 10{,}000$. In fact, the changes in the simulation results are too small to be observed if θ is a large number. Therefore, taking $\theta = 10{,}000$ can capture the simulation results of $\theta = +\infty$.

22. Some researchers, Orphanides and Williams (2002), for example, assume that the loss function of the central bank is the weighted sum of the variances of output gap and inflation. In this case, the results of the numerical computation above implies that the smaller the θ, the higher the loss function.

23. Moreover, steady states might not be optimal, if multiple steady states exist; see Greiner and Semmler (2002).

24. See Graham and Snower (2002), Blanchard (2005), and Semmler et al. (2005, Ch. 7).

25. Jenkins (2002) claims that the policy to be pursued may not be affected by pure data uncertainty in the case of additive-shock uncertainty. However, in case the central bank follows an interest-rate rule that is a function of a small number of variables, the monetary policy rule may be largely affected by data uncertainty.

26. Rudebusch (2001), for example, shows that an increase in output-gap uncertainty may reduce the coefficient on the output gap.

27. The matrices A, B, Q, and R are assumed to satisfy the assumptions stated by Hansen and Sargent (2002, Ch. 3).

REFERENCES

Aguiar, A. & M.M.F. Martins (2002) Trend, Cycle, and Nonlinear Trade-off in the Euro-area 1970–2001. Centro De Estudos Macroeconomicos E Previsão, Portugal.

Beck, G.W. & V. Wieland (2002) Learning and control in a changing economic environment. *Journal of Economic Dynamics and Control* 26, 1359–1377.

Blanchard, O. (2005) Monetary policy and unemployment. In W. Semmler (ed.), *Monetary Policy and Unemployment—U.S., Euro-Area and Japan*, pp. 9–16. London and New York: Routledge.

Brainard, W. (1967) Uncertainty and the effectiveness of policy. *American Economic Review* 52, 411–425.

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

Brock, W.A., S.N. Durlauf, & K.D. West (2003) Policy Evaluation in Uncertain Economic Environment. NBER working paper 10025.

Brock, W.A. & A. Xepapadeas (2003) Regulating nonlinear environmental systems under Knightian uncertainty. In R. Arnott, B. Greenwald, R. Kanbur, & B. Nalebuff (eds.), *Economics for an Imperfect World: Essays in Honor of Joseph Stiglitz*, pp. 127–144. Cambridge, MA: MIT Press.

Cagetti, M., L.P. Hansen, T. J. Sargent, & N. Williams (2002) Robustness and pricing with uncertain growth. *Review of Financial Studies* 15, 363–404.

Chen, Z. & L.G. Epstein (2000) Ambiguity, Risk and Asset Returns in Continuous Time. Working paper, 474, Rochester Center for Economic Research.

Clarida, R., J. Gali, & M. Gertler (1998) Monetary policy rules in practice: some international evidence. *European Economic Review* 42, 1033–1067.

Clarida, R., J. Gali, & M. Gertler (1999) The science of monetary policy: a new Keynesian perspective. *Journal of Economic Literature* XXXVII, 1661–1707.

Clarida, R., J. Gali, & M. Gertler (2000) Monetary policy rules and macroeconomic stability: evidence and some theory. *Quarterly Journal of Economics* CXV (1), 147–180.

Cogley, T. & T.J. Sargent (2001) Evolving post-World War II U.S. inflation dynamics. *NBER Macroeconomics Annual* 2001.

Cogley, T. & T.J. Sargent (2003) Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S. Working paper 2003–25, Federal Reserve Bank of Atlanta.

Epstein, L. and M. Schneider (2001) Recursive Multiple Priors. Working paper 485, Rochester Center for Economic Research.

Escribano, A. (2004) Nonlinear error correction: the case of money demand in the United Kingdom (1878–2000). *Macroeconomic Dynamics* 8, 76–116.

Evans, G.W. & S. Honkapohja (2001) *Learning and Expectations in Macroeconomics*. Princeton, NJ: Princeton University Press.

Giannoni, M.P. (2002) Does model uncertainty justify caution? Robust optimal monetary policy in a forward-looking model. *Macroeconomic Dynamics* 6, 111–144.

Giordani, P. & P. Söderlind (2003) Solution of macromodels with Hansen-Sargent robust policies: summary and some extensions. SSE/EFI Working Paper Series in Economics and Finance 499.

Gonzalez, F. & A. Rodriguez (2003) Robust control: a note on the response of the control to changes in the "free" parameter. Manuscript, University of Texas at Austin.

Graham, L. & D.J. Snower (2002) The Return of the Long-Run Phillips Curve. Discussion paper 646, Institute for the Study of Labor (IZA).

Greiner, A. & W. Semmler (2002) Monetary policy, non-uniqueness of steady states and hysteresis effects. In Reinhard Neck (ed.), *Modeling and Control of Economic Systems*. Oxford: Elsevier Science.

Grüne, L. (1997) An adaptive grid scheme for the discrete Hamilton-Jacobi-Bellman equation. *Numerische Mathematik* 75, 319–337.

Grüne, L. & W. Semmler (2004) Using dynamic programming with adaptive grid scheme for optimal control problems in economics. *Journal of Economics Dynamics and Control* 28, 2427–2456.

Hansen, L.P. & T.J. Sargent (1999) Robustness and commitment: A Monetary Policy Example. Manuscript, Department of Economics, University of Chicago, and Department of Economics, New York University.

Hansen, L.P. & T.J. Sargent (2001a) Robust Control and Model Uncertainty. Manuscript, Department of Economics, University of Chicago, and Department of Economics, New York University.

Hansen, L.P. & T.J. Sargent (2001b) Time Inconsistency of Robust Control? Manuscript, Department of Economics, University of Chicago, and Department of Economics, New York University.

Hansen, L.P. & T.J. Sargent (2002) Robust Control and Filtering for Macroeconomics. Manuscript, Department of Economics, University of Chicago, and Department of Economics, New York University.

Harvey, A.C. (1981) *Time Series Models*. Oxford: Phillip and Allan.

Isard, P., D. Laxton, & A.C. Eliasson (1999) Simple monetary policy rules under model uncertainty. *International Tax and Public Finance* 6, 537–577.

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

Jenkins, P. (2002) Monetary Policy and uncertainty. *Bank of Canada Review* (Summer), pp. 3–10.
Karanassou, M. & D. Snower (2003) An Antomy of the Phillips Curve. Manuscript, Queen Mary and Birkbeck College, University of London.
Kim, C.J. & C.R. Nelson (1999) State-space models with regime switching. Cambridge, MA: MIT Press.
Laxton, D., G. Rose, & D. Tambakis (1998) The U.S. Phillips curve: The case for asymmetry. *Journal of Economic Dynamics and Control* 23, 1459–1485.
Martin, B. & C. Salmon (1999) Should Uncertain Monetary Policy-makers Do Less. Manuscript, Bank of England.
McGough, B. (2003) Statistical learning with time-varying parameters. *Macroeconomic Dynamics* 7, 119–139.
Meyer, L.H., E.T. Swanson, & V. Wieland (2001) NAIRU Uncertainty and Nonlinear Policy Rules. Finance and Economics Discussion Series, 2001-01, Board of Governors of the Federal Reserve System.
Onatski, Alexei & James H. Stock (2002) Robust monetary policy under model uncertainty in a small model of the U.S. economy. *Macroeconomic Dynamics* 6, 85–110.
Orphanides, A. & S. van Norden (2002) The unreliability of output gap estimation in real time. *Review of Economics and Statistics* 84, 569–583.
Orphanides, A. & J.C. Williams (2002) Imperfect Knowledge, Inflation Expectations and Monetary Policy. FRBSF working paper 2002–04.
Rudebusch, G.D. (2001) Is the Fed too timid? Monetary policy in an uncertain world. *Review of Economics and Statistics* 83, 203–217.
Rudebusch, G.D. & L.E.O. Svensson (1999) Policy rules for inflation targeting. In John B. Taylor (ed.), *Monetary Policy Rules*, pp. 203–246. Chicago: University of Chicago Press.
Sargent, T.J. (1993) *Bounded Rationality in Macroeconomics*. London: Clarendon Press.
Sargent, T.J. (1999) *The Conquest of American Inflation*. Princeton: Princeton University Press.
Schaling, E. (2002) The Nonlinear Phillips Curve and Inflation Forecast Targeting—Symmetric Versus Asymmetric Monetary Policy Rules. Manuscript, http://general.rau.ac.za/economics/Ilse/nonlinphcfin[1].09.09.pdf.
Semmler, W., A. Greiner, & W. Zhang (2005) *Monetary and Fiscal Policies in the Euro-Area: Macro modeling, learning and empirics*. Elsevier Science Publisher, Amsterdam.
Sims, C.A. (2001a) Pitfalls of a minimax approach to model uncertainty. *American Economic Review* 91(2) 51–54.
Sims, C.A. (2001b) Comment on Sargent and Cogley's "Evolving U.S. Postwar Inflation Dynamics." Manuscript, Princeton University.
Sims, C.A. & T. Zha (2002) Macroeconomic Switching. Manuscript, Princeton University, www.princeton.edu/sims.
Söderström, Ulf (2002) Monetary policy with uncertain parameters. *Scandinavian Journal of Economics* 104, 125–145.
Svensson, L.E.O. (1997) Inflation forecasting targeting: implementing and monitoring inflation targets. *European Economic Review* 41, 1111–1146.
Svensson, L.E.O. (1999) Inflation targeting: Some extensions. *Scandinavian Journal of Economics* 101, 337–361.
Svensson, L.E.O. (2000) Robust Control Made Simple. Manuscript, Princeton University, http://www.princeton.edu/svensson/.
Tambakis, D.N. (1998) Monetary Policy with a Convex Phillips Curve and Asymmetric Loss. IMF working paper 98/21.
Tetlow, R.J. & P. von zur Muehlen (2001a) Simplicity versus optimality: the choice of monetary policy rules when agents must learn. *Journal of Economic Dynamics and Control* 25, 245–279.
Tetlow, R.J. & P. von zur Muehlen (2001b) Avoiding Nash Inflation: Bayesian and Robust Responses to Model Uncertainty. Manuscript, Board of Governors of the Federal Reserve System.
Tucci, M.P. (1997) Adaptive control in the presence of time-vaying parameters. *Journal of Economic Dynamics and Control* 22, 39–47.

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

Vázquez, J. (2003) The Role of the Term-Spread in an Augmented Taylor Rule. Manuscript, Universidad del País Vasco.
Wieland, V. (2000) Monetary policy, parameter uncertainty and optimal learning. *Journal of Monetary Economics* 46, 199–228.

APPENDIX A: BRIEF SKETCH OF STATE-SPACE MODEL WITH MARKOV SWITCHING

Below is a brief sketch of the state-space model with Markov switching explored by Kim and Nelson (1999). Let $\boldsymbol{Y}_{t-1}$ be the vector of observations available at time $t-1$. In a traditional Kalman filter in a state-space model without Markov switching, the forecast of ϕ_t based on Y_{t-1} can be written as $\phi_{t|t-1}$, and the matrix that denotes the mean squared error of the forecast reads

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

with E being the expectation operator.

In the state-space model with Markov switching, however, the forecast of ϕ_t is based on Y_{t-1}, as well as 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):

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

The mean squared error of the forecast reads

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

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

$$\phi_{t|t-1}^{(i,j)} = \bar{\Phi}_j + F\phi_{t-1|t-1}^{i}, \tag{A.1}$$

$$P_{t|t-1}^{(i,j)} = FP_{t-1|t-1}^{i}F' + \sigma_\eta^2, \tag{A.2}$$

$$\xi_{t|t-1}^{(i,j)} = Y_t - X_t\phi_{t|t-1}^{(i,j)}, \tag{A.3}$$

$$v_{t|t-1}^{(i,j)} = X_t P_{t|t-1}^{(i,j)} X_t' + \sigma_{\varepsilon,j}^2, \tag{A.4}$$

$$\phi_{t|t}^{(i,j)} = \phi_{t|t-1}^{(i,j)} + P_{t|t-1}^{(i,j)} X_t' \left[v_{t|t-1}^{(i,j)}\right]^{-1} \xi_{t|t-1}^{(i,j)}, \tag{A.5}$$

$$P_{t|t}^{(i,j)} = \left(I - P_{t|t-1}^{(i,j)} X_t' \left[v_{t|t-1}^{(i,j)}\right]^{-1} X_t\right) P_{t|t-1}^{(i,j)}, \tag{A.6}$$

where $\xi_{t|t-1}^{(i,j)}$ is the conditional forecast error of Λ_t based on information up to time $t-1$ and $v_{t|t-1}^{(i,j)}$ is the conditional variance of the forecast error $\xi_{t|t-1}^{(i,j)}$. To make the above Kalman filter algorithm operable, Kim and Nelson (1999) develop some approximations and manage to collapse $\phi_{t|t}^{(i,j)}$ and $P_{t|t}^{(i,j)}$ into $\phi_{t|t}^{j}$ and $P_{t|t}^{j}$, respectively.

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

Because the Phillips and IS curves contain only lags of variables and have uncorrelated noise, we can estimate the two equations separately. For the Phillips curve, we have the following state-space model:

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

with

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

$$\eta_t = (\eta_{\alpha 1t}\ \eta_{\alpha 2t}\ \eta_{\alpha 3t})',$$

$$\sigma^2_{\eta} = \left(\sigma^2_{\eta_{\alpha 1}}\ \sigma^2_{\eta_{\alpha 2}}\ \sigma^2_{\eta_{\alpha 3}}\right)',$$

$$\bar{\Phi}_{S_t} = \left(\bar{\Phi}_{\alpha_1, S_t}\ \bar{\Phi}_{\alpha_2, S_t}\ \bar{\Phi}_{\alpha_3, S_t}\right)',$$

$$F = \begin{pmatrix} f_{\alpha_1} & 0 & 0 \\ 0 & f_{\alpha_2} & 0 \\ 0 & 0 & f_{\alpha_3} \end{pmatrix},$$

and similarly for the IS curve, we have

$$\Lambda_t = y_t, \qquad X_t = (1\ y_{t-1}\ r_{t-1}), \qquad \phi_t = (\beta_{1t}\ \ \beta_{2t}\ \ \beta_{3t})', \qquad \boldsymbol{\varepsilon_t} = \varepsilon_{yt},$$

with

$$\varepsilon_{yt} \sim N\left(0, \sigma^2_{\varepsilon y, S_t}\right),$$

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

$$\eta_t = (\eta_{\beta 1t}\ \eta_{\beta 2t}\ \eta_{\beta 3t})',$$

$$\sigma^2_{\eta} = \left(\sigma^2_{\eta_{\beta 1}}\ \sigma^2_{\eta_{\beta 2}}\ \sigma^2_{\eta_{\beta 3}}\right)',$$

$$\bar{\Phi}_{S_t} = \left(\bar{\Phi}_{\beta_1, S_t}\ \bar{\Phi}_{\beta_2, S_t}\ \bar{\Phi}_{\beta_3, S_t}\right)',$$

$$F = \begin{pmatrix} f_{\beta_1} & 0 & 0 \\ 0 & f_{\beta_2} & 0 \\ 0 & 0 & f_{\beta_3} \end{pmatrix}.$$

APPENDIX B: BRIEF SKETCH OF ROBUST CONTROL

Below is a brief sketch of robust control theory explored by Hansen and Sargent (2002). Let the loss function be $L(y, u) \equiv -(x'Qx + u'Ru)$, with Q being positive semidefinite and R positive definite matrices, respectively. The problem without model misspecification

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

is

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

subject to the so-called approximating model[27]

$$x_{t+1} = Ax_t + Bu_t + C\check{\epsilon}_{t+1}, \qquad x_0 \text{ given}, \tag{B.2}$$

where $\{\check{\epsilon}\}$ is an i.i.d. vector process subject to normal distribution with mean zero and identity covariance matrix. If there is model misspecification that cannot be depicted by $\check{\epsilon}$, Hansen and Sargent (2002) take a set of models surrounding equation (B.2) of the form (the so-called distorted model)

$$x_{t+1} = Ax_t + Bu_t + C(\epsilon_{t+1} + \omega_{t+1}), \tag{B.3}$$

where $\{\epsilon_t\}$ is another i.i.d. process subject to normal distribution with mean zero and identity covariance matrix, and ω_{t+1} is a vector process that reads

$$\omega_{t+1} = g_t(x_t, x_{t-1}, \ldots), \tag{B.4}$$

with $\{g_t\}$ being a sequence of measurable functions. Hansen and Sargent (2002) further restrain the approximation errors by

$$E_0 \sum_{t=0}^{\infty} \rho^{t+1} \omega'_{t+1} \omega_{t+1} \leq \eta_0. \tag{B.5}$$

to express the idea that equation (B.2) is a good approximation when equation (B.3) generates the data. To solve the robust control problem (B.1) subject to equations (B.3) and (B.5), 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 (B.5). The constraint problem is

$$\underset{\{u_t\}_{t=0}^{\infty}}{\text{Max}}\ \underset{\{\omega_{t+1}\}_{t=0}^{\infty}}{\text{Min}}\ E_0 \sum_{t=0}^{\infty} \rho^t U(x_t, u_t), \tag{B.6}$$

subject to equations (B.3) and (B.5). Given $\theta \in (\underline{\theta}, +\infty)$ with $\underline{\theta} > 0$, the multiplier problem can be presented as

$$\underset{\{u_t\}_{t=0}^{\infty}}{\text{Max}}\ \underset{\{\omega_{t+1}\}_{t=0}^{\infty}}{\text{Min}}\ E_0 \sum_{t=0}^{\infty} \rho^t \{U(x_t, u_t) + \rho\theta \omega'_{t+1} \omega_{t+1}\}, \tag{B.7}$$

subject to equation (B.3). 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 plays an important role in the problem's solution. If θ is $+\infty$, the problem then becomes the usual optimal control problem without model misspecification. To find a reasonable value for θ, Hansen and Sargent (2002, Ch. 13) design a detection-error probability function by a likelihood ratio. Assume that there is a sample of observations from $t = 0$ to $t = T - 1$, and define L_{ij} as the likelihood of this

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press

sample for model j, provided that the data are generated by model i. The likelihood ratio is then

$$r_i \equiv \log \frac{L_{ii}}{L_{ij}}, \qquad i \neq j. \tag{B.8}$$

Note that r_i should be larger than zero if the data are generated by model i. Define

$$\begin{aligned} p_i &= \text{Prob}(\text{mistake} \mid i) = \text{freq}(r_i \leq 0), \\ p_j &= \text{Prob}(\text{mistake} \mid j) = \text{freq}(r_j \leq 0). \end{aligned}$$

Assuming equal prior weights to both models, the detection-error probability can be defined as

$$p(\theta) = \frac{1}{2}(p_i + p_j). \tag{B.9}$$

When a reasonable value of $p(\theta)$ is chosen, a corresponding value of θ can be determined by inverting the probability function defined in (B.9). 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, $\theta = -1/\sigma$.

Note the interpretation of the detection-error probability. The larger the detection-error probability, the more difficult it is to tell the two models apart. 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, it is not necessary to design a robust rule. Note that the higher the θ, the lower the robustness, not the opposite.

The solution of the multiplier problem is shown below. Define

$$\mathcal{D}(P) = P + PC(\theta I - C'PC)^{-1}C'P, \tag{B.10}$$

$$\mathcal{F}(\Omega) = \rho[R + \rho B'\Omega B]^{-1}B'\Omega A, \tag{B.11}$$

$$T(P) = Q + \rho A[P - \rho PB(R + \rho B'PB)^{-1}B'P]A. \tag{B.12}$$

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 (B.7) is

$$u = -Fx, \tag{B.13}$$

$$\omega = Kx, \tag{B.14}$$

with

$$F = \mathcal{F} \circ \mathcal{D}(P), \tag{B.15}$$

$$K = \theta^{-1}(I - \theta^{-1}C'PC)^{-1}C'P[A - BF]. \tag{B.16}$$

If $\theta = +\infty$, $\mathcal{D}(P) = P$, the problem above then becomes the traditional LQ problem.

https://doi.org/10.1017/S1365100505040332 Published online by Cambridge University Press