Learning About the Term Structure and Optimal Rules for Inflation Targeting

parttext
M Tesfaselassie[1], E Schaling[2] & S Eijffinger[3]

Working Paper Number 62

[1] Kiel Institute for the World Economy
[2] SARB Chair, University of Pretoria and Center for Economic Research, Tilburg University
[3] Center for Economic Research, Tilburg University

Learning About the Term Structure and Optimal Rules for Inflation Targeting

Mewael F. Tesfaselassie,[*] Eric Schaling,[†] Sylvester Eijffinger[‡]

This version 14 March 2007

Abstract

In this paper we incorporate the term structure of interest rates in a standard inflation forecast targeting framework. We find that under flexible inflation targeting and uncertainty in the degree of persistence in the economy, allowing for active learning possibilities has effects on the optimal interest rate rule followed by the central bank. For a wide range of possible initial beliefs about the unknown parameter, the dynamically optimal rule is in general more activist, in the sense of responding aggressively to the state of the economy, than the myopic rule for small to moderate deviations of the state variable from its target. On the other hand, for large deviations, the optimal policy is less activist than the myopic and the certainty equivalence policies.

JEL Classification: C53, E43, E52, F33

Keywords: Learning, Rational Expectations, Separation Principle, Term Structure of Interest Rates

[*]Corresponding author: Kiel Institute for the World Economy, Dusternbrooker Weg 120, 24105 Kiel, Germany. Tel: +49 431 8814 273, E-mail: mewael.tesfaselassie@ifw-kiel.de

[†]South African Reserve Bank Chair, University of Pretoria and Center for Economic Research, Tilburg University, The Netherlands, Pretoria 0002, South Africa, Tel: + 27 12 420 3820 (direct)/ 2423 (Sindi Magwaza), Fax: + 27 12 362 5207, E-mail: eric.schaling@up.ac.za

[‡]CentER for Economic Research, Tilburg University , RSM Erasmus University, CESifo and CEPR. P.O.Box 90153, 5000 LE Tilburg, The Netherlands. Tel: +31 13 466 2416, E-mail: s.c.w.eijffinger@uvt.nl

1

1 Introduction[1]

1.1 Overview

As pointed out by Bullard [1991], in the three decades since the publication of the seminal work on rational expectations (RE) in the early 1960s, a steely paradigm was forged in the economics profession regarding acceptable modelling procedures. Simply stated, the paradigm was that economic actors do not persist in making foolish mistakes in forecasting over time.

Since the late 1980s researchers have challenged this paradigm by examining the idea that how systematic forecast errors are eliminated may have important implications for macroeconomic policy. Researchers who have focused on this question have been studying what is called 'learning', because any method of expectations formation is known as a learning mechanism. Thus, since the late 1980s a learning literature, or learning paradigm, developed. An excellent introduction to - and survey of - this paradigm is presented in Evans and Honkapohja [2001]. Here recent papers that have applied (private sector) learning in a monetary policy context are Bullard and Mitra [2002], Bullard and Schaling [2006], and Orphanides and Williams [2002].

A different strand of literature in the economics profession has been dealing with optimal control or dynamic optimization. There is a now a growing number of papers that combines the themes of learning and (optimal) control. Examples are Wieland [2000b], Wieland [2000a], Wieland [2006], Beck and Wieland [2002] and recent work by Gaspar and Vestin [2006] and Ellison [2006]. Gaspar, Smets and Vestin (2006) and Ellison (2006).

In a pioneering paper, Wieland [2000b] analyzes the situation where a central bank has limited information concerning the transmission channel of monetary policy. Then, the central bank is faced with the difficult task of simultaneously controlling the policy target and estimating (learning) the impact of policy actions. Thus, the so-called separation principle does not

hold, and a trade-off between estimation and control arises because policy actions influence estimation (learning) and provide information that may improve future performance. Wieland analyses this trade-off in a simple model with parameter uncertainty and conducts dynamic simulations of the central bank's decision problem.

1.2   Recent related literature

Several recent studies have revived the issues of active learning and optimal monetary policy under multiplier uncertainty. The literature typically constructs the problem of learning and control around a simple regression model where the explanatory variable is also the control variable, a policy instrument such as the interest rate, whose coefficient has to be estimated and at the same time decisions have to be made about the appropriate level of the instrument that minimizes the expected loss from the variability of the dependent variable, say inflation, around a desired target level.

In particular, the central bank can affect its own learning possibilities through its current choice of the policy instrument. But in doing so it sacrifices short-term goals to carry out experiments. In this case, one has to differentiate between three policy rules: certainty equivalence, myopic, and optimal (Prescott [1972]). The first two rules ignore the dynamic link between learning and control. While the certainty equivalence rule ignores parameter uncertainty, the myopic rule allows for uncertainty surrounding the unknown parameters. On the other hand, the optimal policy incorporates active learning.

A careful reading of this strand of literature shows that most studies assume the presence of uncertainty in the policy multiplier. The policy multiplier can be modeled as the coefficient on the money supply, the rate of interest (Ellison [2006]), or inflation (Yetman [2002], Ellison and Valla [2001]). A common feature of these studies is that the linear economic process subject to central bank control is static, as in Brainard [1967]. Thus all dynamics in the economy are only due to central bank learning.

Under uncertainty in the policy multiplier, experimentation may require the central bank to be more responsive to new information on the state of the economy (thus higher variability of the policy instrument). This turns out be the case in Beck and Wieland [2002]. This conclusion undermines the basic Brainard [1967] result that calls for more cautious policy (implying a

lower variability of the policy instrument).[2] What drives these recent results (calling for a more active policy) is the possibility that the central bank can learn about the unknown parameter and thus improve performance in the future.

However, Ellison and Valla [2001] argue in favor of a more cautionary policy under active learning by appealing to strategic considerations. They show that strategic interactions between the central bank and the private sector introduce additional costs associated with activist policy. More specifically, while activist monetary policy may generate valuable information, it can lead to volatile inflation expectations that in turn hinder the central bank from holding inflation and output stable around their targets. Ellison and Valla [2001] thus restore the main message of Brainard [1967], although for reasons related to optimal learning.

1.3 Organization

Our model, although linear, differs from most of the above mentioned studies in two respects. First, the structural equations in our model are dynamic even if one assumes that there is no learning by the central bank. This is due to the presence of endogenous persistence (inertia) in the economy in the sense that future economic conditions depend in part on the current state of the economy. Second, while the literature typically studies uncertainty about a policy multiplier, the nature of information symmetry in our term structure equation implies that the persistence parameter in the linear process is unknown to the central bank.[3]

More specific, we incorporate the term structure of interest rates in a standard inflation forecast targeting framework. Learning about the transmission process of monetary policy is introduced by having heterogeneous agents - i.e. the central bank and private agents - who have different information sets about the future sequence of short-term interest rates. We analyze inflation forecast targeting in two environments. One in which the central bank has

---

[2]However, Schaling [2004] calls for a more aggressive policy response in a framework that modifies Svensson [1997] with a non-linear Phillips curve, which gives rise to uncertainty in the policy multiplier even if the model has only additive demand shocks. Under strict inflation targeting, he shows that policy is more aggressive than implied by certainty equivalence. Thus under multiplier uncertainty, Brainard's result does not always go through when there are inherent non-linearities in the economy.

[3]In a generic model, Beck and Wieland [2002] analyze uncertainty in the policy multiplier when there is endogenous persistence in the state variable. Levin and Moessner [2005] also study uncertainty about (inflation) persistence.

perfect knowledge, in the sense that it understands and observes the process by which private sector interest rate expectations are generated, and one in which the central bank has imperfect knowledge and has to learn the private sector forecasting rule for short-term interest rates. In the case of imperfect knowledge, the central bank has to learn about private sector interest rate expectations, as the latter affect the impact of monetary policy through the expectations theory of the term structure of interest rates. Here following Evans and Honkapohja [2001], the learning scheme we investigate is that of least-squares learning (recursive OLS) using the Kalman filter.

We find that that under flexible inflation targeting and uncertainty in the degree of persistence in the economy, allowing for active learning possibilities has effects on the optimal interest rate rule followed by the central bank. For a wide range of possible initial beliefs about the unknown parameter, the dynamically optimal rule is in general more activist, in the sense of responding aggressively to the state of the economy, than the myopic rule for small to moderate deviations of the state variable from its target. On the other hand, for large deviations, the optimal policy is less activist than the myopic and the certainty equivalence policies.

The remainder of this paper is organized as follows. In section 2 we present the model. Strict inflation targeting under perfect knowledge is analyzed in section 3. Flexible inflation targeting and learning is discussed in section 4. Section 5 compares certainty equivalence and myopic rules. Optimal policy– as well as a sensitivity analysis–is presented in section 6, whilst section 7 concludes.

2   The Environment

In this section we describe the inflation forecast targeting framework, with the transmission mechanism that incorporates the term structure of interest rates. First, the current period rate of inflation is determined by its lagged level and the lagged output gap.

$$\pi_{t+1} = \pi_t + \alpha_1 z_t - \eta_{t+1} \tag{1}$$

where z_t is the (log of) the output gap and $\alpha_1 > 0$ is the slope of the Phillips curve. The output gap is defined as the difference between actual output and potential output, the latter of which is conveniently normalized to zero, and η is a normally distributed white noise supply shock.

The output gap is autoregressive and is affected by the long-term real interest rate:

$$z_{t+1} = \beta_1 z_t - \beta_2 R_t + d_{t+1} \tag{2}$$

where R_t is the long-term real interest rate and d_{t+1} is a white noise demand shock that is also assumed to be normally distributed. Moreover, the two coefficients satisfy the restriction $\beta_1 > 0$ and $\beta_2 > 0$.[4]

The short-term ex-ante real interest rate, r_t, which can perfectly be controlled by the central bank,[5] and the long real rate are related by the expectations theory of the term structure

$$r_t = R_t - D(\hat{E}_t R_{t+1} - R_t) \tag{3}$$

where $\hat{E}_t R_{t+1}$ denotes private sector expectations (where the hat sign ' ^ ' denotes a possibly nonrational expectation) of next period's long real rate.[6] Here r_t represents the real yield to maturity on a one-period bond which is traded on the interbank money market. This yield must be equal to the (one-period) real holding period return on a long-term bond. The parameter D is defined such that $D + 1$ is equal to what is known as Macaulay's duration (see Eijffinger et al. [2000] for details). For our purposes it turns out to be convenient to rewrite (3) so that the current long real rate is expressed as a linear combination of r_t and $\hat{E}_t R_{t+1}$:

$$R_t = (1 - k)r_t + k\hat{E}_t R_{t+1} + \zeta_t \tag{4}$$

where $k \equiv D/(D + 1)$. We have added a normally distributed white noise term ζ_t, where $\zeta_t \sim N(0, \sigma_\zeta^2)$, to capture an unobserved term premium.

Note that equation (4) can be rewritten as $R_t = (1-k)\sum_{\tau=t}^{\infty} k^{\tau-t}\hat{E}_t r_\tau + \zeta_t$; or alternatively, as

$$R_t = (1-k)r_t + (1-k)\sum_{\tau=t+1}^{\infty} k^{\tau-t}\hat{E}_t r_\tau + \zeta_t \tag{5}$$

---

[4]This relationship is similar to the one used by Rudebusch and Svensson [2002]. The differences are that here the output gap depends on the long-term real interest rate rather than the short-term real interest rate, and that they consider an additional lagged z term.

[5]The central bank has perfect control over the real short rate, r_t, because $r_t = i_t - E_t\pi_{t+1}$ where $E_t\pi_{t+1}$ is predetermined at time t.

[6]This will be relevant when discussing central bank learning under imperfect knowledge.

Thus the long term real interest rate is a weighted average of the current ex ante real short rate and the expected future sequence of future short real rates over the $t + 1$–infinity horizon.[7]

The current short-term real interest rate will be equal to:

$$r_t = i_t + E_t \pi_{t+1} \tag{6}$$

Here i_t is the nominal interest rate on the inter-bank money market and represents the expected rate of inflation in period t+1 conditional on the information set in period t.

3 Strict Inflation Targeting under Perfect Knowledge

It remains to specify the preferences of the central bank. The central bank chooses a sequence of current and future short-term nominal interest rates to meet its objectives. Since $E_t \pi_{t+1}$ is predetermined, we write the optimization problem in terms of choosing $\{r_\tau\}_{\tau=t}^\infty$.

$$\min_{\{r_\tau\}_{\tau=t}^\infty} E_t \sum_{\tau=t}^\infty \delta^{\tau-t} \frac{1}{2}(\pi_\tau - \pi^*)^2 \tag{7}$$

subject to (1), (2), (4) and (6). Here, π^* is the central bank's inflation target and δ is the discount rate with $(0 < \delta < 1)$. The expectations operator E_t refers to the central bank's expectations conditional on the information set in period t. It is obvious that the derivation of the optimal short rate depends on the assumed information structure, including model and data uncertainty faced by the central bank.

To get some straightforward results, this section assumes that the central bank can observe and respond directly to private sector expectations and moreover that the private sector and the central bank have rational expectations.[8] In this section, we also set $\zeta_t = 0$ for all t without loss of generality.

---

[7]It follows that $\hat{E}_t R_{t+1} = \left(\frac{1-k}{k} \right) \sum_{\tau=t+1}^\infty k^{\tau-t} \hat{E}_t r_\tau$.

[8]So, the central bank knows how much policy is 'in the pipeline' according to financial markets.

The timing of events is such that the central bank chooses its interest rate policy after private sector expectations are set. In the terminology of game theory, the private sector is a Stackelberg leader and the central bank is a Stackelberg follower. In each period, the sequence of events is summarized as follows:

Sequence of events in period t		
z_t, π_t realize $\rightarrow$	private sector sets $\rightarrow$	central bank chooses
	$\hat{E}_t R_{t+1}$	$r_t = r(z_t, \pi_t, \hat{E}_t R_{t+1})$

The first-order condition in terms of expected inflation is thus (see Appendix A of Eijffinger et al. [2004] for details)[9]

$$E_t \pi_{t+2} = \pi^* \tag{8}$$

Thus, optimal policy for a strict inflation targeting regime defined by (7) implies inflation forecast targeting in the sense of Svensson [1997]. That is, the best the central bank can do is to set the short rate such that it expects inflation two-period ahead to be on target. The first-order condition, together with the structural constraints, leads to the following closed form solution for the short-term ex ante real interest rate:

$$r_t = \frac{1}{\alpha_1 \beta_2 (1-k)} (\pi_t - \pi^*) + \frac{1 + \beta_1}{\beta_2 (1-k)} z_t - \frac{k}{1-k} \hat{E}_t R_{t+1} \tag{9}$$

Note that rule (9) is in the form of the Evans and Honkapohja [2003b] (hereafter EH) *expectations-based optimal rule*. By construction, it implements what EH label "optimal discretionary policy" in every period and for all values of private sector expectations.

Rule (9) differs from simple (non-optimal) Taylor-type rules, which are widely used for policy analysis, as policy responds directly to $\hat{E}_t R_{t+1}$, the private sector's forecast of the long real rate. This modification is a result of the forward-looking term structure equation that has been embedded in an otherwise backward-looking model. It is easy to check that if $k = 0$ and the term structure equation vanishes, the policy rule collapses to a version of

---

[9]See also Svensson [1997]. Even though the first-order condition in terms of $E_t \pi_{t+2}$ is equivalent to that in Svensson [1997], our optimal level of the short rate responds to three state variables– π_t, z_t, and $\hat{E}_t R_{t+1}$, which are part of the central bank's information set. In Svensson [1997] the state variables are π_t and z_t.

the Taylor rule analyzed by Svensson [1997] (hereafter the Svensson-Taylor rule).[10]

An interesting characteristic of this solution is that the central bank's optimal choice of r_t is inversely related to private sector expectations about its future short rates. In the words of Mervyn King, Governor of the Bank of England, such a rule would imply "letting the market do the work for us."

For example, if the private sector expects rates to go up (down) in the future, as a consequence current real short rates are lowered (raised) today. The property that if the private sector expects future short rates to go down the central bank raises the current short real rate (or talks about raising it) reminds us of the old joke about the Bundesbank (BuBa): "The BuBa is just like cream, the more you stir it, the thicker it gets."[11] The reason for this inverse relationship is that the central bank's inflation forecast– given the current period inflation rate and the output gap– depends on the present level of the real long-term interest rate. So, an optimal inflation forecast implies an optimal level of the current long real rate. Since the optimal long real rate (consistent with strict inflation targeting) is a weighted average of the current ex ante optimal short rate r_t^* and $\hat{E}_t R_{t+1}$, i.e. $R_t^* = (1 - k)r_t^* + k\hat{E}_t R_{t+1}$, a higher value of $\hat{E}_t R_{t+1}$ necessitates a lower value of r_t^*, and vice versa.[12]

3.1 The Rational Expectations Equilibrium

Next, we study the rational expectations equilibrium. Inserting the reaction function (9) in the term structure (4) implies that R_t has the following reduced form solution:

$$R_t = \frac{1}{\alpha_1 \beta_2}(\pi_t - \pi^*) + \frac{1 + \beta_1}{\beta_2} z_t \tag{10}$$

Importantly, this solution shows that the effects of private sector expectations of the long rate are completely offset when the central bank observes and responds optimally to those expectations. We can of course derive the

---

[10]Taylor rules are often written in terms of the i_t. Given the definition of r_t one can easily derive the optimal level of i_t from (9).

[11]In addition, the Bundesbank always considered the long-term interest rate as a reflection of its credibility.

[12]In other words, if the private sector expects a looser monetary policy in the future, this leads to a tighter policy stance today to compensate.

exact form of those expectations. Inserting (10) in the output equation gives the reduced form solution:

$$z_{t+1} = -\frac{1}{\alpha_1}(\pi_t - \pi^*) - z_t + d_{t+1} \tag{11}$$

Next, leading (10) one period and using equations (11) and (1) we get:

$$R_{t+1} = -\frac{\beta_1}{\alpha_1\beta_2}(\pi_t - \pi^*) + \frac{\beta_1}{\beta_2}z_t + u_{t+1} \tag{12}$$

where $u_{t+1} \equiv \frac{1+\beta_1}{\beta_2}d_{t+1} - \frac{1}{\beta_2}\eta_{t+1}$ is a composite white noise shock. Taking private sector expectations as of time t gives:

$$\hat{E}_t R_{t+1} = -\frac{\beta_1}{\alpha_1\beta_2}(\pi_t - \pi^*) - \frac{\beta_1}{\beta_2}z_t \tag{13}$$

This result represents the rational expectations solution under symmetric information, where private sector expectations are consistent with the solution for the long-term real interest rate implied by strict inflation targeting. Alternatively, we can express r_t in reduced form by plugging (13) into the reaction function (9):

$$r_t = \frac{1+k\beta_1}{\alpha_1\beta_2(1-k)}(\pi_t - \pi^*) + \frac{1+\beta_1(k+1)}{\beta_2(1-k)}z_t \tag{14}$$

This is what Evans and Honkapohja [2003b] call the *fundamentals-based form* of the optimal policy rule. Note that this rule is more specific than the expectations-based rule, as it is obtained by substituting the rational expectations solution for the future long real rate into the expectations-based rule.

As can be seen from (14) the ex ante real rate, r_t, is now expressed in terms of current inflation and output. In this way we can compare the behavior of r_t in (14) with the Svensson-Taylor rule that results by setting $k = 0$. We can easily check that when $0 < k < 1$ and the term structure is relevant, r_t becomes more sensitive to changes in the state variables π_t and z_t. Moreover the higher the value of k, the more volatile r_t. The intuition for this result is as follows. If k increases the term structure says that private sector expectations become more important in determining R_t, a variable that directly affects aggregate demand. To offset the decrease in the policy leverage over R_t, the central bank needs to be more aggressive in responding to changes in π_t and z_t.

10

4 Imperfect Knowledge, Beliefs and Learning

The previous section looked at a benchmark case of *perfect knowledge* where the central bank perfectly observes and responds to private sector expectations of the long real rate. Although the model extends Svensson [1997], endowing the central bank with perfect knowledge of private sector expectations is hardly realistic. Central banks typically observe private sector expectations with error, and these observation errors can be large [Evans and Honkapohja, 2003a].[13]

Suppose that the private sector's forecasting function for the long real rate takes the same form as the rational expectations solution under full information, namely, equation (13). That is, expectations respond to information on the current period state variables–inflation and output. Specifically,

$$\hat{E}_t R_{t+1} = \gamma w_t \tag{15}$$

where we have defined a new variable $w_t \equiv -(\pi_t + z_t)$ and for the purpose of tractability we have set $\alpha_1 = 1$ and $\pi^* = 0$. The nature of imperfect information is such that, the central bank knows the private sector's forecasting rule, but not the actual value of γ.[14] Let c_t denote the best forecast of γ and p_t the variance of c_t, (the degree of confidence placed upon c_t), of which more will be said later.[15]

4.1 Learning

Since the central bank is assumed to have full knowledge of the functional form (15), it knows that the term structure equation under imperfect knowledge obeys

$$R_t = (1-k)r_t + k\gamma w_t + \zeta_t \tag{16}$$

---

[13]We continue to assume that the central bank knows the structural model of the economy, including the parameters.

[14]which is equal to $\gamma \equiv \beta_1/\beta_2$.

[15]By positing a simple forecasting function for the private sector, we abstract from the interaction of optimal monetary policy and rational expectations on the part of the private sector. However, even though expectations are non-rational in the short-run, they turn out to be rational in the limit since the central bank learns the unknown parameter with probability one. The solution of the model is then identical to the full information rational expectations equilibrium.

Observe that the unobserved term structure shock, ζ_t, injects additional uncertainty and prevents the central bank from inferring, in any period, the value of γ from (16).[16] Next, substituting equation (16) in the IS and Phillips curves gives,

$$w_{t+1} = \tilde{\gamma} w_t - \beta_1 z_t + \tilde{\beta} r_t + \nu_{t+1} \tag{17}$$

where $\tilde{\beta} \equiv \beta_2(1-k)$ and $\nu_{t+1} \equiv \beta_2 \zeta_t - d_{t+1} + \eta_{t+1}$ is a composite white noise shock. The persistence parameter $\tilde{\gamma} \equiv 1 + \beta_2 k \gamma > 0$ is unknown to the central bank because of the unknown parameter γ coming from equation (15).

When active learning by the central bank is involved, the optimization problem usually gets complicated as the number of state variables increases (due to what is known as the "curse of dimensionality"). In our case, besides the state variable w_t, the current output gap, z_t, also appears as a state variable in equation (17). Together with the beliefs about the mean and variance of the persistence parameter, there would be four state variables in the model. The dual control literature usually uses a constraint similar to (17) but with $\beta_1 = 0$ so that next period's state variable, w_{t+1}, depends on the policy instrument, r_t, and possibly lagged values of the state variable; see Beck and Wieland [2002]. In our case, setting $\beta_1 = 0$ reduces the number of state variables to three. This makes the dynamic equation comparable to that considered by Beck and Wieland [2002]:[17]

$$w_{t+1} = \tilde{\gamma} w_t + \tilde{\beta} r_t + \nu_{t+1} \tag{18}$$

Let $\tilde{c}_t$ denote an estimate of $\tilde{\gamma}$. Now, since $\tilde{\gamma} = 1 + \beta_2 k \gamma$, after $\tilde{c}_t$ has been estimated, c_t can be inferred indirectly from $c_t = -(1 - \tilde{c}_t)/k\beta_2$. when period $t+1$ arrives, the central bank updates its estimate by including the latest available data (w_{t+1}, w_t, r_t) in the regression equation (18). Using the widely used method of recursive least squares we have the following updating

---

[16]This is true even if the central bank knows the parameter k, and has data on the yield curve (r_t and R_t) and the current state of the economy, w_t.

[17]Note that if $\beta_1 = 0$, the perfect knowledge, rational expectations solution–where in addition we have set $\pi = 0$ and $\alpha_1 = 1$–gives $\hat{E}_t R_{t+1} = (\beta_1/\beta_2)w_t = 0$. This is inconsequential to the way learning is modeled under imperfect knowledge. The motivation for central bank learning comes from the knowledge of the functional form of (15), where in the case $\beta_1 = 0$, the central bank would be assumed to lack information about the true γ being actually zero.

equations for $\tilde{c}_t$ and its variance, denoted by $\tilde{p}_t$ [see for e.g. Beck and Wieland, 2002, Pollock, 2002].[18]

$$\begin{aligned} \tilde{c}_{t+1} &= \tilde{c}_t + \kappa_{t+1}(w_{t+1} - \tilde{c}_t w_t - \tilde{\beta} r_t) \\ \tilde{p}_{t+1} &= \tilde{p}_t - \kappa_{t+1} w_t \tilde{p}_t \end{aligned} \qquad (19)$$

where $\kappa_{t+1} \equiv w_t \tilde{p}_t F_{t+1}^{-1}$ is commonly referred to as the Kalman gain, which is the weight assigned to new information coming from the forecast error $w_{t+1} - \tilde{c}_t w_t - \tilde{\beta} r_t$. The Kalman gain in turn depends on the conditional variance of w_{t+1}, (based on information in period t), given by $F_{t+1} \equiv w_t^2 \tilde{p}_t + \sigma_\nu^2$. As can be seen from (19), the current state of the economy, w_t is part of the Kalman gain, and affects the path of the conditional variance of the parameter estimate, $\tilde{p}_{t+1}$. Due to the presence of autoregressive behavior in w_{t+1}, changes in the current state of the economy, w_t, have direct effects on the variability of w_{t+1}, and consequently on the variability of $w_{t+2}, w_{t+3}, ...$[19]

4.2 Flexible Inflation Targeting under Imperfect Knowledge

Under *flexible inflation targeting* the central bank chooses $\{r_\tau\}_{\tau=t}^{\infty}$ to minimize the discounted sum of expected current and future losses, subject to the constraints (18) and (19).

$$\min_{\{r_\tau\}_{\tau=t}^{\infty}} E_t \sum_{\tau=t}^{\infty} \delta^{\tau-t} L_\tau \qquad (20)$$

where $L_t = \frac{1}{2}(w_t^2 + \lambda r_t^2)$ and $\lambda > 0$ is the relative weight assigned to the loss from the variability in r_t, the control variable.[20] As before, the discount rate is denoted by δ, $(0 < \delta < 1)$. The expectations operator E_t refers to the

---

[18]If the unknown parameter is time-varying, the updating equations can be modified to allow for this variability via the Kalman filter; see for instance Sargent [1999] and Beck and Wieland [2002].

[19]The assumption that the shock is normally distributed with known variance is standard in the learning literature. If the prior belief also have a normal distribution, then the posterior belief is a normal distribution. This property of the posterior belief is convenient when dealing with numerical computations (see the appendix to chapter 6 of Tesfaselassie [2005]).

[20]This follows the learning and control literature. See Beck and Wieland [2002] for a detailed discussion.

central bank's expectations conditional on information set in period t. Note that when $\lambda \to \infty$, the central bank engages in full stabilization of r_t, i.e., $r_t = 0$ for all t.

The updating equations in (19) capture the idea that central bank learning about the unknown parameter is influenced by policy decisions made in period t, r_t. The channel work as follows: r_t affects the state variable in period t+1, w_{t+1}, and consequently, beliefs about the unknown parameter (i.e. $\tilde{c}_{t+1} = c(w_{t+1})$ and $\tilde{p}_{t+2} = p(w_{t+1})$). The link between current and future policy choices is established because expected future interest rate decisions by the central bank depend on the expected future state of the economy. From the principle of least squares estimation the precision of the estimate $\tilde{c}_{t+2}$ depends positively on the variance of w_{t+1}. One gets a more precise estimate (in other words, a smaller value of $\tilde{p}_{t+2}$) when the variance of w_{t+1} increases, and vice versa. Since we recognize that the current choice of monetary policy r_t affects $E_t w_{t+1}$, and given that F_{t+1} is predetermined, the coefficient of variation, defined by the ratio $\sqrt{F_{t+1}}/E_t w_{t+1}$ is also a function of r_t. This relationship between r_t, w_{t+1} and $\tilde{p}_{t+2}$ raises a potential tension between the urge to minimize current period loss from variability in r_t (the *control* part) and the need to get a more precise estimate of the degree of persistence in the economy that would help improve future outcomes (the *learning* part).

When the central bank conducts policy based on *passive learning*, by construction it disregards the potential tradeoff between estimation and control. In other words, the central bank simply ignores the effect of current policy actions on the degree of precision of future estimates of the unknown parameter, thereby treating control and estimation separately. Formally, the passive learning policy first calculates c_t and p_t and then takes these parameters to be fixed at the stage of optimization, which means that when choosing policy, the dynamic process of these estimates (the updating equations) are ignored. Likewise, before choosing r_{t+1} in period $t + 1$ the central bank updates its belief about γ to c_{t+1}, ignoring the fact that it will have to update this estimate in the future.[21] In this way the central bank fails to internalize the effect of current actions on future beliefs.

Certainty equivalence is a special case of a passive learning policy since it ignores p_t by assumption. Thus, the policy maker does not incorporate p_t even if the effect of p_t on the loss function could in principle be reduced by

---

[21] In the sense of Sargent [1999], passive learning implies that in any period t the central bank pretends that its current estimate c_t will apply forever, as if it is the true parameter. But the central bank's updating of its estimate in period t+1 falsifies this pretense.

an appropriate choice of r_t.[22]

In a strict inflation targeting regime with passive learning, the central bank minimizes the inter-temporal loss function (20) setting $\lambda = 0$ and taking into account the linear dynamic equation (18). The optimization problem would give solutions for r_t analogous to the case of perfect knowledge (section 3), as the problem is solved period-by-period and the first-order condition sets the conditional expectation of inflation two-period ahead equal to the desired target. As we will see below, this is no longer true with flexible inflation targeting ($\lambda > 0$) as future losses and the discount rate play a role in the choice of the current policy rate.

5   Passive Learning: Certainty Equivalence vs. Myopic

Before considering the role of active learning in optimal policy, this section solves for optimal policy under passive learning. There are two subcases under passive learning– one that ignores parameter uncertainty (certainty equivalence rule) and the one that incorporates parameter uncertainty (myopic rule). Both rules are passive in nature because they ignore the link between policy choices today and future learning that is apparent from the updating equations. From the vantage point of the current period the central bank's belief is not expected to be updated in the future, implying that when choosing current policy, it anticipates the initial belief $(\tilde{c}_t, \tilde{p}_t)$ to remain fixed for all future periods. Consequently, in both cases, the non-linear updating equations drop out of the optimization problem.[23]

The certainty equivalence rule is a special case of the myopic policy rule since under the former $\tilde{p}_t = 0$. The implication of this can be seen by decomposing $E_t w_{t+1}^2$ as follows[24]

$$E_t w_{t+1}^2 = (E_t w_{t+1})^2 + F_{t+1} \tag{21}$$

where $E_t w_{t+1} = \tilde{c}_t w_t + \tilde{\beta} r_t$ and under the myopic policy $F_{t+1} = w_t^2 \tilde{p}_t + \sigma_\nu^2$. On the other hand, under certainty equivalence, $\tilde{p}_t = 0 \Rightarrow w_t^2 \tilde{p}_t = 0$, and so

---

[22]Note that having full confidence in c_t is equivalent to assuming $p_t = 0$.

[23]Of course, when next period arrives, the bank updates its belief but then expect it to remain fixed from that period on.

[24]Remember, by assumption the additive shock ν_{t+1} is i.i.d. and thus uncorrelated with period t estimation error, $\tilde{\gamma} - \tilde{c}_t$.

F_{t+1} is completely exogenous, depending only on the variance of the additive shock, σ_ν^2.

$$E_t w_{t+1}^2 = (E_t w_{t+1})^2 + \sigma_\nu^2 \qquad (22)$$

Thus the difference in the way the conditional variance, F_{t+1}, is treated is also reflected in the solution of the dynamic control problem under each case (see below).

5.1 The Certainty Equivalence Policy (CER)

Under *certainty equivalence* the central bank ignores parameter uncertainty. In other words, the central bank is fully confident about its estimate $\tilde{c}_t$ so that its current belief is given by $(\tilde{c}_t, \tilde{p}_t) = (\tilde{c}_t, 0)$. We can also think of this situation as the limiting case of the updating equations, where $\tilde{p}_t = 0$ implies that $\tilde{c}_{t+1} = \tilde{c}_t$. Thus, the state $\tilde{c}_t$ does not change and is independent from the policy instrument. The minimization problem is

$$\min_{\{r_\tau\}_{\tau=t}^\infty} E_t \Big[\sum_{\tau=t}^\infty \delta^{\tau-t} L_\tau | (w_t, \tilde{c}_t) \Big] \qquad (23)$$

subject to the linear constraint (18). Importantly, with the non linear updating equations ignored, the problem is linear-quadratic, and the derivation of the optimal level of r_t is similar to that under perfect knowledge. To get the certainty equivalence rule, one usually proceeds in two steps. First, solve for the optimal rule assuming perfect knowledge of $\tilde{\gamma}$ and second, simply replace $\tilde{\gamma}$ by its recent estimate $\tilde{c}_t$. Alternatively, first replace $\tilde{\gamma}$ by its recent estimate $\tilde{c}_t$ in the linear constraint (18) and then solve the optimization problem taking the estimate as a fixed parameter. In any case, we can rewrite the above minimization problem using recursive dynamic programming and then use the 'guess and verify' method on the value function as in Schaling [2001].[25]

With this in mind, one can write the Bellman equation associated with the minimization of (23).[26]

$$V(w_t) = \min_{r_t} [L(w_t, r_t) + \delta E_t V(w_{t+1})] \qquad (24)$$

---

[25] For a discussion of the relative merits of dynamic programming and Lagrange method see Schaling [2001]. For applications of the latter to a non-linear optimization problem, and a regime switching model see Schaling [2004] and Bullard and Schaling [2001], respectively.

[26] Note that the value function in the Bellman equation does not have time subscript. This is because in infinite horizon problems, we are interested only in the unique time

subject to (18) with $\tilde{\gamma}$ replaced by its certainty equivalence estimate $\tilde{c}_t$, which at the stage of optimization is understood by the central bank to be a fixed parameter (and not a state variable). Because of the resulting linear-quadratic structure of the minimization problem, the value function will be quadratic in the state w_t.

$$V(w_t) = \mu_0 + \frac{1}{2}\mu w_t^2 \tag{25}$$

where the two coefficients remain to be determined. If (25) is correct, it follows that:

$$E_t V(w_{t+1}) = \mu_0 + \frac{1}{2}\mu E_t w_{t+1}^2 \tag{26}$$

where $E_t w_{t+1}^2$ follows from equation (22). Using (26) in (24) and taking the first order condition, we get

$$\lambda r_t + \mu\tilde{\beta}\delta(\tilde{c}_t w_t + \tilde{\beta} r_t) = 0 \tag{27}$$

which can easily be solved for r_t in the fundamentals-based form:

$$r_t = -\frac{\mu\delta\tilde{c}_t\tilde{\beta}}{\lambda + \mu\delta\tilde{\beta}^2} w_t \tag{28}$$

It is important to observe that, even though we have placed a time subscript on $\tilde{c}_t$, as far as passive learning is concerned, $\tilde{c}_t$ should be thought of as a fixed parameter (not a state variable) and it is not expected to be affected by the current policy choice. This is the sense in which forecasting and control are separated by construction.[27]

In order to identify μ, first differentiate (25) with respect to w_t:

$$V_w(w_t) = \mu w_t \tag{29}$$

---

invariant value function, V, and associated unique, stationary policy rule, that result from repeated iterations on the Bellman equation starting from any bounded continuous V_0 (e.g. $V_0 = 0$). Convergence of the value function is guaranteed due to the *contraction mapping theorem* [see Sargent, 1987]. For linear-quadratic control problems, convergence is achieved in a single iteration if V_0 is quadratic.

[27]Thus, due to the sequential nature of decision making, the updating of the parameter estimate is kept in the background.

Next, invoking the envelope theorem on the Bellman equation (24), and taking note of (28):

$$V_w(w_t) = w_t + \delta\mu\tilde{c}_t(\tilde{c}_t w_t + \tilde{\beta} r_t) = f(\mu)w_t \qquad (30)$$

where $f(\mu) \equiv 1 + \lambda\delta\tilde{c}_t^2\mu/(\lambda + \delta\tilde{\beta}^2\mu)$. For the conjectured value function (25) to be correct, it is required that the coefficients of (29) and (30) have to be equalized:[28]

$$\mu = f(\mu) \qquad (31)$$

Rearrange (31) to get the following quadratic equation for μ:

$$\delta\tilde{\beta}^2\mu^2 + [\lambda - \delta(\tilde{\beta}^2 + \tilde{c}_t^2\lambda)]\mu - \lambda = 0 \qquad (32)$$

It is straightforward to check that $\lim_{\mu\to 0} f(\mu) = 1$ and $\lim_{\mu\to\infty} f(\mu) = 1 + \lambda\tilde{c}_t^2/\tilde{\beta}^2$. Thus a unique solution for μ, such that $\mu \geq 1$, is :

$$\mu^{ce} = \frac{1}{2}\left(1 - \frac{\lambda(1 - \delta\tilde{c}_t^2)}{\delta\tilde{\beta}^2} + \sqrt{\left(1 + \frac{\lambda(1 - \delta\tilde{c}_t^2)}{\delta\tilde{\beta}^2}\right)^2 + \frac{4\lambda}{\tilde{\beta}^2}}\right) \qquad (33)$$

where superscript 'ce' stands for certainty equivalence. We see that the estimate of the persistence parameter, $\tilde{c}_t$, which from the perspective of period t is not expected to be updated in the future (as per passive learning), affects the value of the coefficient μ^{ce}. For future reference, note also that $\mu^{ce} \to 1$ as $\lambda \to 0$.

The solution (33) is similar to equation (B.6) of Svensson [1997], except that in Svensson [1997] the persistence parameter is known with certainty (and simply set equal to 1) and flexible inflation targeting is defined in terms of inflation and output stabilization. To see the effect of λ on policy responsiveness to w_t, rewrite (28), (bearing (33) in mind) so that $r_t = -\tilde{c}_t w_t/(\phi+\tilde{\beta})$, where $\phi = \lambda/(\mu^{ce}\delta\tilde{\beta})$; $\partial\phi/\partial\lambda > 0$.[29] Thus given $\tilde{c}_t$, r_t is less responsive to

---

[28]Alternatively, one can work directly with the value function. Substitute (28) in (24) and match the resulting coefficients with those in the conjectured value function (25).

[29]Upon simplifying

$$\phi = 1/2\left(\delta\tilde{\beta}^2 - (1 - \delta\tilde{c}_t^2) + \sqrt{\left(\frac{\delta\tilde{\beta}^2}{\lambda} + 1 - \delta\tilde{c}_t^2\right)^2 + \frac{4\delta^2}{\lambda\tilde{\beta}^2}}\right)$$

from which it is easily seen that $\partial\phi/\partial\lambda > 0$.

w_t as λ increases. Moreover, as $\lambda \to 0$, $\phi \to 0$ and $r_t = -(\tilde{c}_t/\tilde{\beta})w_t$ (strict inflation targeting). On the other hand, as $\lambda \to \infty$, $\phi \to \infty$ implying $r_t = 0$ (full stabilization of r_t).

5.2 The Myopic Policy (MR)

A *myopic policy rule* differs from certainty equivalence only because the myopic policy takes account of the current degree of uncertainty in the current estimate of the persistence parameter (thus $\tilde{p}_t > 0$). The central bank continues to ignore the fact that current policy can affect future beliefs and so treats $\tilde{c}_t$ and $\tilde{p}_t$ as *fixed* parameters, implying that the only state variable from the central bank's point of view is w_t.[30] The conjecture for the value function is then the same as (25) and the first order condition with respect to r_t will take the same form as (27). The difference is that now the coefficient μ will be a function of $\tilde{p}_t$ as well as $\tilde{c}_t$.

In identifying μ, we remember that in (26), $E_t w_{t+1}^2$ is a function of $\tilde{p}_t$ via (21). So

$$
\begin{aligned}
V_w(w_t) &= w_t + \delta\mu\tilde{c}_t(\tilde{c}_t w_t + \tilde{\beta} r_t) + \delta\mu p_t w_t \\
&= \left(1 + \frac{\lambda\delta c_t^2\mu}{\lambda + \delta\tilde{\beta}^2\mu} + \delta\tilde{p}_t\mu\right)w_t
\end{aligned}
\tag{34}
$$

Thus following the steps analogous to the previous section, we match the coefficients of (29) and (34),

$$
\mu = 1 + \frac{\lambda\delta c_t^2\mu}{\lambda + \delta\tilde{\beta}^2\mu} + \delta\tilde{p}_t\mu
\tag{35}
$$

or

$$
(1 - \delta\tilde{p}_t)\mu = 1 + \frac{\lambda\delta c_t^2\mu}{\lambda + \delta\tilde{\beta}^2\mu}
\tag{36}
$$

where the certainty equivalence case arises if $\tilde{p}_t = 0$, that is if one disregards parameter uncertainty. In finding the solution for μ that satisfies (36),

---

[30]One can also think of $\tilde{c}_t$ and $\tilde{p}_t$ as state variables. However, these states do not change and are independent from the policy instrument. This means that, when optimizing, the terms $\kappa_{t+1}(w_{t+1} - \tilde{c}_t w_t - \tilde{\beta} r_t)$ and $\kappa_{t+1} w_t \tilde{p}_t$ on the right hand side of the updating equations drop out.

we note that the expression on the right side of the equation is identical to the corresponding term under certainty equivalence, but on the left hand side of the equation, the coefficient on μ is $0 \leq 1 - \delta \tilde{p}_t \leq 1$ if $\tilde{p}_t$ is not too large. This holds if, say, the initial parameter uncertainty is such that $\tilde{p}_0 \leq 1$, which is actually not that restrictive if $\tilde{c}_0$ is also not too large. In this case, the prior belief $(\tilde{c}_0, \tilde{p}_0)$ such that $\tilde{p}_0 = \tilde{c}_0/2$ implies high initial uncertainty (see Beck and Wieland [2002]). This means that if we restrict the central bank's belief such that $\tilde{c}_0 \leq 2$, we can reasonably assume as well that $\tilde{p}_0 \leq 1$.[31] With this in mind, we can easily observe that when parameter uncertainty is taken into account, the fixed point for μ, denoted by μ^m, will be larger or equal to its fixed point under certainty equivalence, μ^{ce}. Thus the myopic policy is more aggressive (in the sense of larger policy response to new information about the state of the economy) than the certainty equivalent rule.

As before, to get the solution for μ, rewrite (36):

$$A\mu^2 + B\mu - \lambda = 0 \tag{37}$$

where $A \equiv (1 - \tilde{p}_t \delta) \delta \tilde{\beta}^2$, $B \equiv (1 - \tilde{p}_t \delta)\lambda - \delta(\tilde{\beta}^2 + \tilde{c}_t^2 \lambda)$. We can check that if $\tilde{p}_t \to 0$, then $A \to \delta \tilde{\beta}^2$ and $B \to \lambda - \delta(\tilde{\beta}^2 + \tilde{c}_t^2 \lambda)$, which are the coefficients under the certainty equivalence rule.

The solution for equation (37) depends, among other things, on $\tilde{p}_t$.

$$\mu^m = \frac{1}{2}\left(-\frac{B}{A} + \sqrt{\left(\frac{B}{A}\right)^2 + \frac{4\lambda}{A}} \right) \tag{38}$$

$\mu^m \geq 1$ as long as the initial value of $\tilde{p}_0$ is not too large. We know that $\tilde{p}_t$ goes down in magnitude over time as more data about w and r arrive. This is true even for policy under passive learning.[32]

Unlike the case of certainty equivalence, $\lim_{\lambda \to 0} \mu^m = 1/(1 - \delta \tilde{p}_t)$. The myopic case collapses to the certainty equivalence only if $\tilde{p}_t \to 0$, which also implies that $\mu^m = \mu^{ce} \to 1$.

Remember that the central bank knows the policy multiplier, $\tilde{\beta}$, with certainty. On the other hand, when choosing its interest rate under the myopic policy, the central bank behaves as if its initial belief, including $\tilde{p}_0 >$

---

[31]For e.g., in Beck and Wieland [2002] the prior belief about the unknown parameter is characterized by a mean of 0.5 and variance 0.25, so that the central bank faces considerable uncertainty.

[32]See chapter 5 of Tesfaselassie [2005] for a discussion of convergence under learning.

0, will not be updated. In other words, the policy maker currently thinks that he will live with an uncertain estimate now and in the future. This perception somehow exaggerates actual future uncertainty because it neglects the fact that as time goes by, the precision of $\tilde{c}$ will increase ($\tilde{p}$ will decline) with the arrival of new economic data.

What is then the intuition behind a more aggressive rule under the myopic policy? The control problem is dynamic, so the bank expects to incur losses from variability in w_{t+1} (via its effect on $F_{t+2} = \tilde{p}_{t+1}w_{t+1}^2 + \sigma_\nu^2$). Since as of period t the central bank does not internalize the effect of policy on future beliefs, we have $\tilde{p}_{t+1} = \tilde{p}_t$ and $E_t F_{t+2} = \tilde{p}_t E_t w_{t+1}^2 + \sigma_\nu^2 = \tilde{p}_t(E_t w_{t+1})^2 + \tilde{p}_t F_{t+1} + \sigma_\nu^2$, which shows the benefits from a policy that sets $E_t w_{t+1}$ closer to zero. It follows that, given $\lambda > 0$, from the perspective of the myopic policy r_t responds more strongly to deviations of w_t so that $E_t w_{t+1}$ is closer to the target w^* (zero) than implied under the certainty equivalence rule.[33]

It is important to note that policy at time t can affect only $E_t w_{t+1}$, which corresponds with the expected loss under certainty equivalence. On the other hand, the conditional variance, F_{t+1}, is independent of r_t since p_t and w_t are predetermined as of time t. In the case where $\lambda = 0$, the myopic rule collapses to the certainty equivalence rule. This result is different from the classic study by Brainard [1967] and other related papers, where the policy multiplier is assumed to be unknown, so that policy can affect the conditional variance component, and the certainty equivalence principle breaks down even when the control variable (policy instrument) does not enter the loss function. What we have shown in this section is that, as long as $\lambda > 0$ the myopic rule differs from the certainty equivalence for the case where the persistence parameter is unknown while the policy multiplier is known with certainty.

6   Optimal Monetary Policy under Active Learning

We now formalize the active learning problem, in which the central bank does not separate estimation and control, as future beliefs about the persistence parameter depend on the whole history of the state variables and the interest

---

[33]Note that p_t scales up the component of the loss function associated with the variability of w_{t+1}. The presence of p_t does not matter under strict inflation targeting since the policy instrument is set such that $E_t w_{t+1} = 0$. By contrast, with flexible inflation targeting, $\lambda > 0$ implying that $E_t w_{t+1} \neq 0$. The presence of p_t then requires a more aggressive policy that drives $E_t w_{t+1}$ closer to zero.

rate choices of the central bank, including the current one. Under optimal policy, the policy maker can take actions now such that w_{t+1} is more informative and contributes to a more precise future estimate of the persistence parameter.

More precisely, the benefits from experimenting with the policy rate are in terms of reduced variability of the economy in the future from more precise estimates and improved control associated with less uncertainty in the unknown parameter. Under flexible inflation targeting, optimal policy exploits the tradeoff between those future benefits and current costs from large movements in r_t.

As before, the central bank chooses a sequence of current and future short-term nominal interest rates (note that $E_t\pi_{t+1}$ is predetermined) to minimize the intertemporal loss function

$$\min_{\{r_\tau\}_{\tau=t}^{\infty}} E_t\Big[\sum_{\tau=t}^{\infty} \delta^{\tau-t} L_\tau | (w_t, \tilde{p}_t, \tilde{c}_t)\Big] \tag{39}$$

But now, the constraint is not just the linear Phillips curve (18) but also (and importantly) the non linear updating equations (19). The effect of current policy on future beliefs becomes visible from the Bellman equation associated with the dynamic programming problem (39) [see Beck and Wieland, 2002]:

$$V(w_t, \tilde{c}_t, \tilde{p}_t) = \min_{r_t} \left[L(w_t, r_t) + \delta E_t V(w_{t+1}, \tilde{c}_{t+1}, \tilde{p}_{t+1})\right] \tag{40}$$

subject to (18) and (19).

The following table summarizes the results under CER, MR and the fully optimal policy.

Table 1: Summary of Key Results

	Strict IT	Flexible IT
Perfect Knowledge*	$r_t = -\frac{\tilde{\gamma}}{\tilde{\beta}} w_t$	$r_t = -\frac{\mu \delta \tilde{\gamma} \tilde{\beta}}{\lambda + \mu \delta \tilde{\beta}^2} w_t$
Imperfect Knowledge–CER	$r_t = -\frac{\tilde{c}_t}{\tilde{\beta}} w_t$	$r_t = -\frac{\mu^{CE} \delta \tilde{c}_t \tilde{\beta}}{\lambda + \mu^{CE} \delta \tilde{\beta}^2} w_t$
Imperfect Knowledge–MR	$r_t = -\frac{\tilde{c}_t}{\tilde{\beta}} w_t$	$r_t = -\frac{\mu^M \delta \tilde{c}_t \tilde{\beta}}{\lambda + \mu^M \delta \tilde{\beta}^2} w_t$
Imperfect Knowledge– optimal policy	$r_t = -\frac{\tilde{c}_t}{\tilde{\beta}} w_t$	$V(w_t, \tilde{c}_t, \tilde{p}_t) = \min_{r_t} [L(w_t, r_t) + \delta E_t V(w_{t+1}, \tilde{c}_{t+1}, \tilde{p}_{t+1})]$

*μ is the same as μ^{CE} except that $\tilde{c}_t$ is replaced by $\tilde{\gamma}$.

The second term on the right side of (40) (also reproduced in the last row of Table 1) is the expected discounted loss from period $t + 1$ onwards, with time $t + 1$ state vector $(w_{t+1}, \tilde{c}_{t+1}, \tilde{p}_{t+1})$ depending on current period policy actions and state vector via the updating equations. This term thus captures the value of information and is given by (see also the appendix to chapter 6 of Tesfaselassie [2005])

$$ E_t V(.) = \int V\Big(w_{t+1}, c(w_{t+1}, \tilde{c}_t, \tilde{p}_t, w_t, r_t), p(\tilde{p}_t, w_t)\Big) f(w_{t+1}|.) dw \qquad (41) $$

where $f(w|.)$ is the conditional density function of w_{t+1}. Even if it has to control a linear process (18) and its loss function is quadratic in the two arguments, the policy maker faces a non-linear constraint because of the updating equations. Thus the dynamic programming problem falls outside the linear-quadratic formulation that is usually assumed in many economic applications. Fortunately, using a standard contraction mapping argument, Kiefer and Nyarko [1989] have shown the existence of a stationary policy and a unique value function that solve the dynamic programming problem. It is thus possible to use numerical dynamic programming methods to approximate the value function and associated policy rule. Judd [1998] describes extensively the numerical methods for solving Bellman equations, while Wieland [2006], and Beck and Wieland [2002], among others, apply these methods to optimal policy under parameter uncertainty.

As pointed out by Wieland [2006], a drawback of this procedure is the so-called curse of dimensionality, which sets in as the number of state variables becomes large. The reason is that the number of computations increases geometrically with the number of state variables in the optimization problem, which in turn undermines the precision of the numerical approximation. This will not pose a problem for us since we only have three state variables, $\tilde{c}_t$, $\tilde{p}_t$ and w_t.

7   Numerical Results

Having described the main elements of the policy problem under active learning, we now present some numerical results. We first present our results for benchmark values for the weight on interest rate stabilization, λ, the policy multiplier, $\tilde{\beta}$, the discount factor, δ, and the variance of the composite exogenous shock, σ_ν^2. Then we compare these results with those derived for alternative parameter values.

Table 2: Parameter configuration

Parameter	Benchmark value	New value
$\tilde{\beta}$	0.5	1
λ	0.5	0.1
δ	0.95	0.75
σ_ν^2	1	0.5

Possible initial beliefs for $\tilde{c}$ are as high as 1.4 and as low as 0.2, while the beliefs about the variance $\tilde{p}$ range from 0.1 to 0.7. The relative degree of confidence in an estimate measured by the coefficient of variation, $\sqrt{\tilde{p}}/\tilde{c}$, takes its lowest value at $(\tilde{c}_0, \tilde{p}_0) = (1.4, 0.1)$ implying $\sqrt{\tilde{p}}/\tilde{c} < 0.23$. In this case, the uncertainty associated with $\tilde{c}$ is quite small. As will be shown below, given the parameter configuration, the role of parameter uncertainty in inducing an active learning policy decreases with the coefficient of variation. Moreover, to get an idea of how initial beliefs about the degree of uncertainty surrounding the parameter estimate matters for policy, 16 alternative pairs of $(\tilde{c}_0, \tilde{p}_0)$ are considered from the sets $\tilde{c}_0 \in \{0.2, 0.6, 1, 1.4\}$ and $\tilde{p}_0 \in \{0.1, 0.3, 0.5, 0.7\}$. All the figures shown below have $(\pi_t + z_t)$ on the horizontal axis. This follows from the definition $\pi_t + z_t = -w_t$. Using the original variables π_t and z_t helps to interpret the results in terms of inflation and output gap.

7.1   Results under Baseline Parameters

In this section, we compare the three decision rules- certainty-equivalent, myopic and dynamically optimal rules- given the benchmark parameter values and the central bank's initial beliefs about the unobserved persistence parameter. Figure 1 shows the response of interest rate r_t to deviations of the state w_t from its target level of zero, for a specific belief characterized by the mean $\tilde{c}_0 = 1$ and variance $\tilde{p}_0 = 0.5$.

Since monetary policy is conducted under a flexible inflation targeting regime, $(\lambda > 0)$, there is a tradeoff between stabilizing inflation and stabilizing the short real rate. In this case, all three decision rules do not completely offset the predictable impact of w_t on w_{t+1}. With the model featuring autoregressive behavior, any random shock to w_t will then have a long lasting effect under the three policy rules. However the degree of gradualism associated with a given level of λ differs from one decision rule to another.

As we saw in the preceding section, the *myopic rule* is more aggressive than the *certainty equivalent rule* for all levels of the state variable. The reason is that, under the myopic policy rule, the central bank recognizes
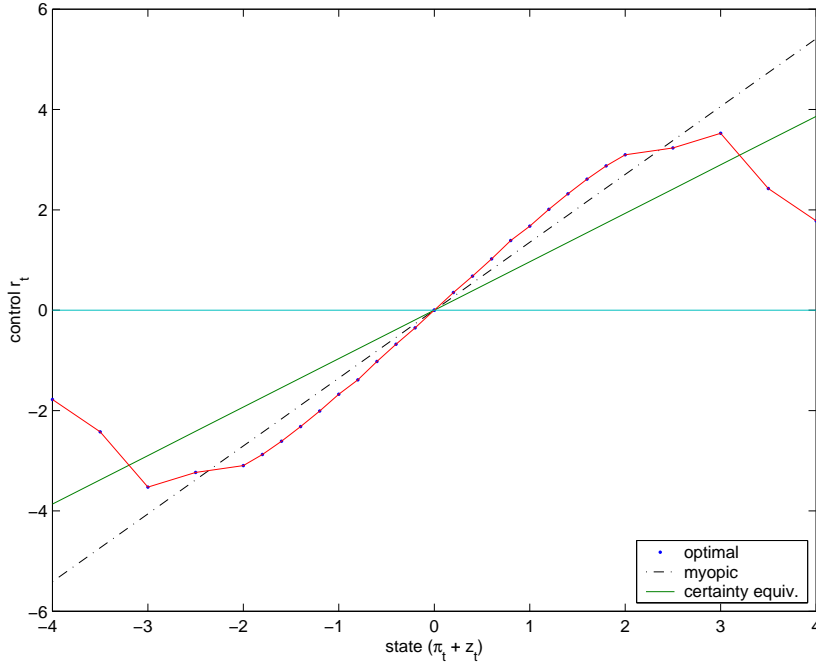
Figure 1: The three decision rules for initial beliefs $\tilde{c}_0 = 1$ and $\tilde{p}_0 = 0.5$

that, provided that there is uncertainty surrounding the belief $\tilde{c}_0$, in other words, $\tilde{p}_0 > 0$, the contribution of this source of uncertainty to the variability of w_{t+1} (and subsequently w_{t+2}, w_{t+3}, ...) increases with w_t. In a dynamic setting, stabilizing w_{t+1} also helps minimize the negative impact of this source of uncertainty. Thus the rate of interest has to move more in response to w_t compared to a policy that disregards parameter uncertainty.

Now, as can be seen from Figure 1, for low to moderate deviations of the state from the target, the *dynamically optimal policy* is even more aggressive than the myopic one. But if the deviations are large, the optimal policy responds less aggressively, even compared to the certainty equivalence policy. The intuition for this is the following. From the updating equations, the larger the deviation of w_t from zero (due to say an exogenous shock), the smaller $\tilde{p}_{t+1}$ (implying that $\tilde{c}_{t+1}$ is a more precise estimate) leading to a smaller control error when setting r_{t+1}. Thus, in contrast to the myopic policy, the actively learning central bank anticipates future improvements in policy performance as $|w_t|$ increases. This shows that when realized exogenous shocks, ν_t, that ultimately drive data generation for w_t, are large there is less of a role for deliberate probing by the central bank, more so the larger

25

the deviation of w_t from the target.[34]

7.2 Sensitivity Analysis

The qualitative results shown in Figure 1 carry over to other possible initial beliefs about the persistence parameter. In Figure 2, 16 alternative plots are shown, each plot corresponding to a specific configuration of the initial belief, $\tilde{c}_0$ and $\tilde{p}_0$.[35] Perhaps not surprisingly, the three rules diverge with the magnitude of $\tilde{p}_0$ and $\tilde{c}_0$. For example, when parameter uncertainty is small (say $\tilde{p}_0 = 0.1$) and there is a small degree of expected persistence in the economy ($\tilde{c}_0 = 0.2$), the three decision rules tend to be identical. At the other extreme, when both $\tilde{c}_0$ and $\tilde{p}_0$ are large, there are clear divergences between the decision rules.

Next we examine how the incentives for the central bank to deviate from the certainty equivalence and myopic rules may depend on other parameters of interest. As can be expected, the three decision rules are affected by changes in $\tilde{\beta}, \lambda$ and δ. Thus when analyzing the effect of changes in these parameters, it is not proper to compare directly the optimal policies arising from each parameter setting. Rather, one has to take the difference between the optimal and the myopic policies. On the other hand, a change in σ_ν^2 affects only the dynamically optimal policy (more about this will be said below). In this case, we compare directly the dynamically optimal rules associated with each value of σ_ν^2.

7.2.1 Policy under Less Volatile Shocks

First, we look at the effect of a decrease in the variance of the exogenous random shock, σ_ν^2. In the benchmark case, the variance was set to 1, while now the variance is reduced by half, standing at 0.5, which implies that, for a given policy path, the economy is inherently more stable as it is subject to less volatile shocks. Note here that unlike the *optimal policy*, the *certainty equivalence* and *myopic rules* are not affected by σ_ν^2, as the shock enters the model additively. So only the optimal decision rules for the two alternative values of σ_ν^2 are shown in Figure 3 below.

---

[34]At the same time, of course, F_{t+1}, the conditional variance of w_{t+1}, increases with w_t^2 but this component of $E_t w_{t+1}^2$ is independent of r_t.

[35]Because the rate of interest responds symmetrically to positive and negative deviations of the state variable from the target (i.e. zero), only positive deviations are shown in the plots.
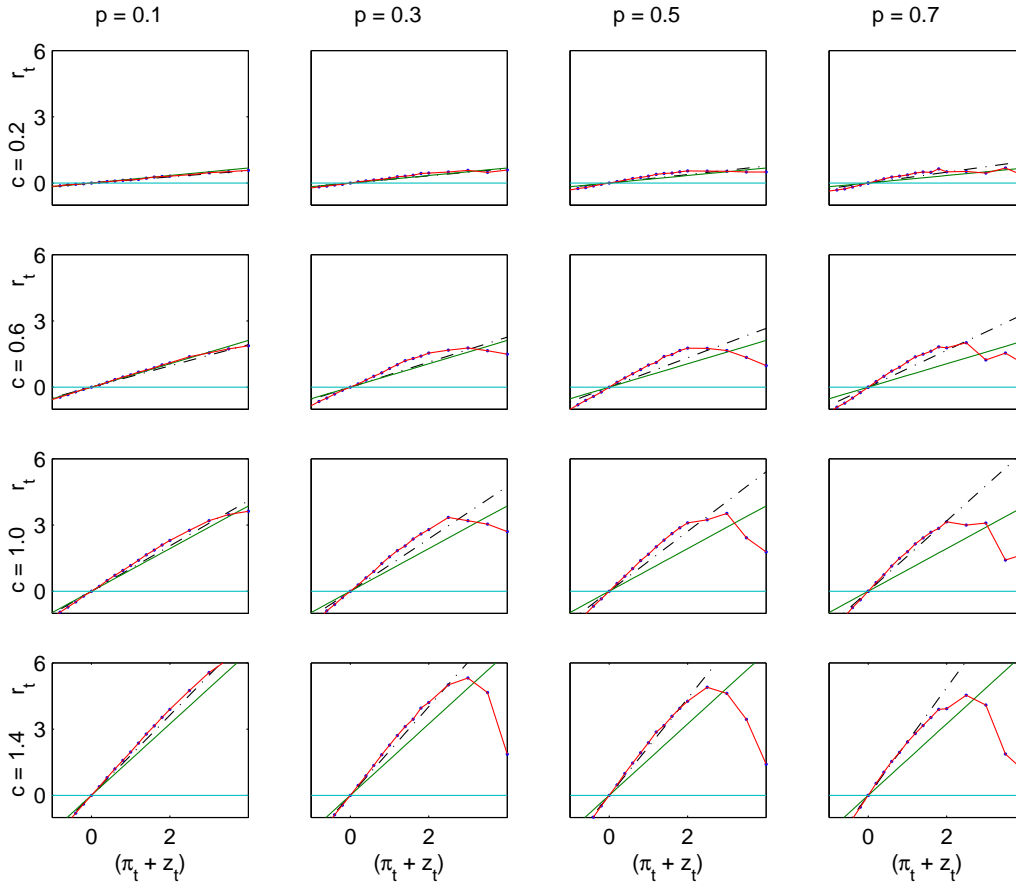
Figure 2: The three decision rules for alternative initial beliefs

In each panel of the figure the benchmark case ($\sigma_\nu^2 = 1$) is shown by the dotted line while the solid line arises from the smaller variance of the shock ($\sigma_\nu^2 = 0.5$). As is apparent from most of the panels, *optimal policy* is more aggressive when the variance σ_ν^2 decreases, that is when the shocks driving the w_t process are less volatile. This is especially the case with large values of $\tilde{c}$ and $\tilde{p}$ shown in the lower right corner of the figure. The intuition is that, with low variability in the shocks, optimal policy needs to actively manage data generation by increasing the (conditional) coefficient of variation in w_{t+1}, defined here as $CV = \sqrt{F_{t+1}}/E_t w_{t+1}$. Since F_{t+1} is predetermined, the CV can be increased only by a lower value of $E_t w_{t+1}$, which in turn requires a more aggressive response of r_t to w_t.[36]

---

[36]Quantitatively, this difference is not large when $\tilde{c}$ or $\tilde{p}$ is small.
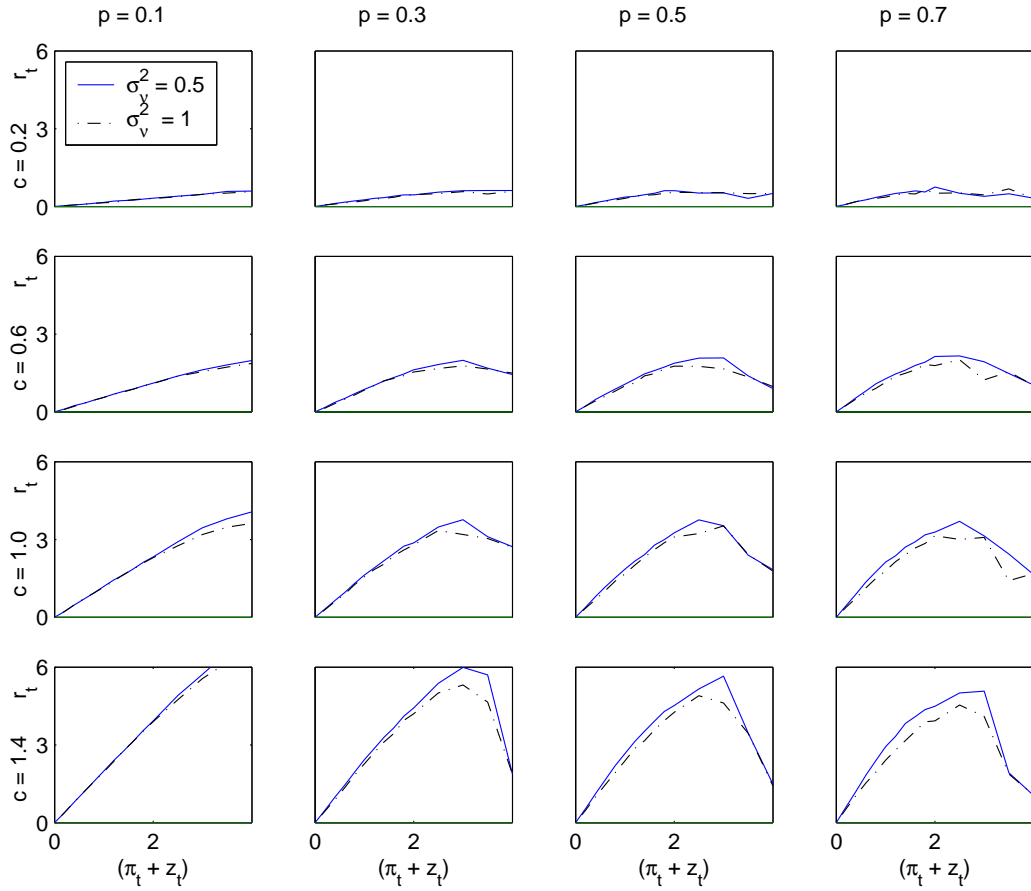
Figure 3: Performance of optimal policy ($\sigma_\nu^2 = 1$ vs. $\sigma_\nu^2 = 0.5$)

7.2.2 Policy Cares More about Inflation Stabilization

With a very small weight on interest rate stabilization, the dynamically optimal policy is still more aggressive than the myopic for small to moderate deviations of the state variable, though to a lesser extent than when interest rate stabilization is more important. The reason is that the myopic policy is already aggressive enough when inflation stabilization receives high priority. The incentive to deviate from the certainty equivalence policy diminishes for a central bank that gives more attention to inflation stabilization. In the limit of strict inflation targeting optimal policy is not affected by uncertainty in the persistence parameter.

7.2.3 The Effect of a Larger Policy Multiplier

Figure 4 shows how the degree of probing is affected by the size of the policy multiplier, $\tilde{\beta} \in \{0.5, 1\}$. The solid line is associated with a larger multiplier, $\tilde{\beta} = 1$, while the dotted line is associated with a smaller multiplier, $\tilde{\beta} = 0.5$. From (17), we know that $\tilde{\beta} \equiv \beta_2(1 - k)$; thus a smaller value of $\tilde{\beta}$ is associated with a weak leverage of monetary policy on the current long real rate, (from the term structure equation, the weight on the short real rate, $(1-k)$, decreases). In other words, as k gets closer to one, the model becomes more forward-looking since the long term interest rates will be determined mainly by movements in private sector expectations. Due to this fact, the importance of learning increases as well.

Using the difference between the optimal policy and the myopic policy as a measure of experimentation, Figure 4 shows that, given small to moderate deviations of $(\pi_t + z_t)$, experimentation decreases as the value of $\tilde{\beta}$ gets larger. For most of the alternative initial beliefs, experimentation is maintained over large deviations of the state from the target when $\tilde{\beta} = 1$. The effect on the relative response of the optimal policy of a change in $\tilde{\beta}$ diminishes if $\tilde{c}_0$ or $\tilde{p}_0$ is sufficiently small.

7.2.4 Policy Cares Less about the Future

As said earlier, the benefits from experimenting with the policy rate are in terms of reduced variability of the economy in the future from more precise estimates and improved control associated with less uncertainty in the unknown parameter. Under flexible inflation targeting, optimal policy exploits the tradeoff between these future benefits and current costs from large movements in r_t. In this regard, one would expect that the incentive to experiment increases as the central bank gives more weight to expected future losses from variability in $w_{t+\tau}$ and $r_{t+\tau}$, for $\tau = 1, 2, 3....$ Thus there is more tendency to probe for a relatively large discount factor (alternatively, for a relatively small discount rate). Figure 5 confirms this intuition about the effect of changes in the discount factor, δ, especially for large initial beliefs about the persistence parameter.

In Figure 5, each plot shows the difference between the optimal and myopic policies for two sets of values of the discount factor, $\delta \in \{0.95, 0.75\}$. The baseline value of 0.95 is shown by the solid line and is compared to the smaller value 0.75 given by the dotted line. With a smaller discount factor, the central bank cares less about the future expected losses and thus policy
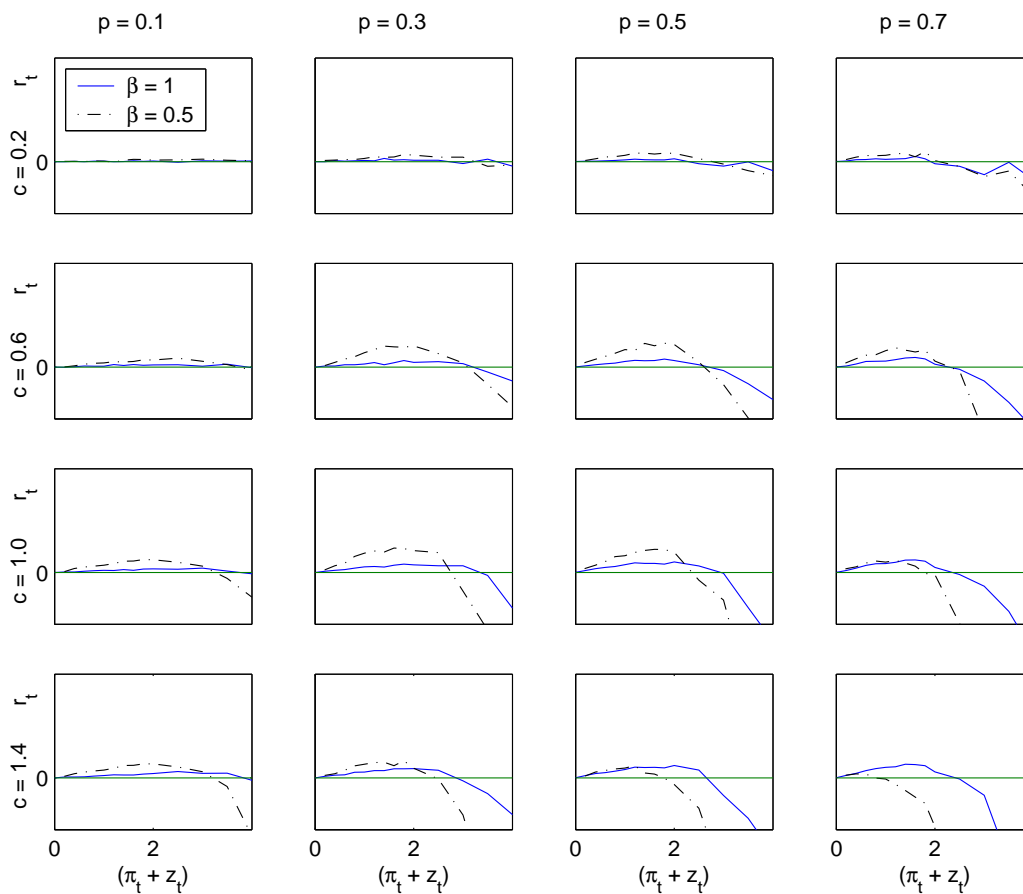
Figure 4: Relative performance of optimal policy ($\tilde{\beta} = 1$ vs. $\tilde{\beta} = 0.5$)

tilts towards stabilizing current expected loss from variability in r_t.

8 Concluding Remarks

In section 4.2 we set out to extend the analysis of section 3 to a general case where on the one hand monetary policy faces a tradeoff in stabilizing inflation as well as the short-term interest rate, (the policy instrument), and on the other hand, the central bank internalizes the effects of current policy choices on its learning possibilities about an unknown degree of persistence in the economy.

We find that under flexible inflation targeting and uncertainty about the degree of persistence in the economy, allowing for active learning possibilities
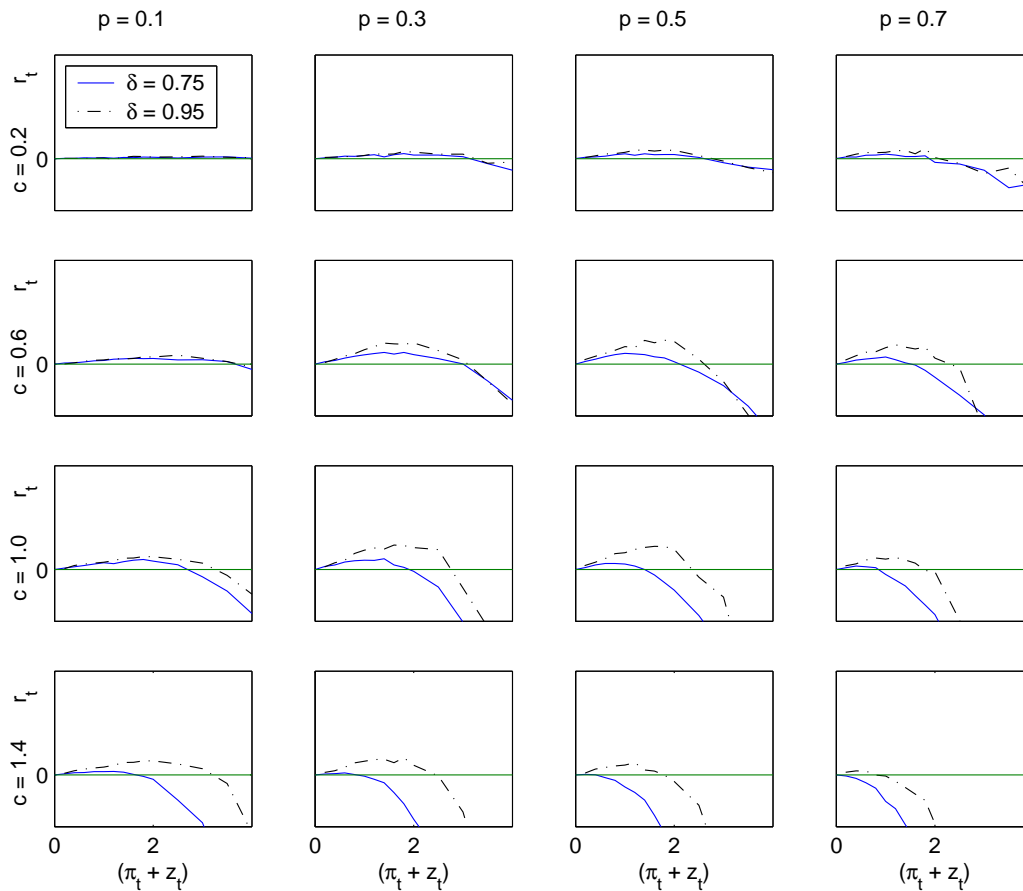
Figure 5: Relative performance of optimal policy ($\delta = 0.75$ vs. $\delta = 0.95$)

has effects on the optimal interest rate rule followed by the central bank. For a wide range of possible initial beliefs about the unknown parameter, the dynamically optimal rule is in general *more activist*, in the sense of responding aggressively to the state of the economy, than the myopic rule for small to moderate deviations of the state variable from its target. On the other hand, for large deviations, the optimal policy is less activist than the myopic and the certainty equivalence policies. This shows that the role of optimal policy in generating variability (increasing the coefficient of variation of the target variable) depends on the current state of the economy. When next period's state of the economy is expected to deviate a lot from the target due to an unpredictable shock, and thus generates data by itself, optimal policy does not need to increase the coefficient of variation of the next period's state w_{t+1}, while it does so when the economy is hit by a very small shock and the state is close to the target.

31

On the other hand, the myopic rule does not incorporate the future benefits from large deviations of the current state from its target level and thus responds linearly and more aggressively than the certainty equivalence rule. The intuition for the aggressiveness of the myopic rule lies in the fact that it takes account of the additional source of uncertainty, which is compounded by the magnitude of the state variable, as the unknown parameter is multiplicative to the current state. By acting more aggressively to stabilize the predictable impact of current period shocks on future aggregate demand and inflation, the myopic reduces the impact of uncertainty associated with estimating the persistence parameter.

This feature of the myopic rule differs from what one might find when the source of parameter uncertainty lies with the policy multiplier. In that case, policy under myopic rule tends to be *less aggressive* than policy under certainty equivalence. Uncertainty about the policy multiplier forces the central bank to be cautious about using its policy instrument freely to stabilize inflation. In our case, the analogous explanation is that, with uncertainty about the persistence parameter, the central bank would like to see less variability in next period's state variable. Under the myopic rule, this can be achieved only if the policy rate responds aggressively to new information.

References

G. Beck and V. Wieland. Learning and Control in a Changing Economic Environment. *Journal of Economic Dynamics and Control*, 26(9), 2002.

W. Brainard. Uncertainty and the Effectiveness of Policy. *American Economic Review*, 57, 1967.

J. Bullard. Learning, Rational Expectations and Policy: A Summary of Recent Research. *Federal Reserve Bank of St. Louis Review, January/February*, 1991.

J. Bullard and K. Mitra. Learning About Monetary Policy Rules. *Journal of Monetary Economics*, 49, 2002.

J. Bullard and E. Schaling. New Economy-New Policy Rules? *Federal Reserve Bank of St. Louis Review*, 83(5):57–66, 2001.

J. Bullard and E. Schaling. Monetary Policy, Determinacy, and Learnability in the Open Economy. *ECB Working Paper, No. 611*, 2006.

S. Eijffinger, E. Schaling, and M. Tesfaselassie. Heterogeneous Information about the Term Structure, Least-Squares Learning, and Optimal Rules for Inflation Targeting. *CEPR Discussion Paper 4279*, 2004.

S. Eijffinger, E. Schaling, and W. Verhagen. The Term Structure of Interest Rates and Inflation Forecast Targeting. *CEPR Discussion Paper 2375*, 2000.

M. Ellison. The Learning Cost of Interest Rate Reversals. *Journal of Monetary Economics*, 2006.

M. Ellison and N. Valla. Learning, Uncertainty and Central bank Activism in an Economy with Strategic Interactions. *Journal of Monetary Economics*, 48(1):153–171, 2001.

G.W. Evans and S. Honkapohja. *Learning and Expectations in Macroeconomics*. Princeton University Press, 2001.

G.W. Evans and S. Honkapohja. Adaptive Learning and Monetary Policy Design. *Journal of Money, Credit and Banking*, 35(6):1045–1073, 2003a.

G.W. Evans and S. Honkapohja. Expectations and the Stability Problem for Optimal Monetary Policies. *Review of Economic Studies*, 70:807–824, 2003b.

F. Smets Gaspar, V. and D. Vestin. Adaptive Learning, Persistence and Optimal Monetary Policy. *ECB Working Paper, No. 644*, 2006.

K. Judd. *Numerical Methods in Economics*. MIT Press, 1998.

N. Kiefer and Y. Nyarko. Optimal control of an unknown linear process with learning. *International Economic Review*, 30:1299–1310, 1989.

A. Levin and R. Moessner. Inflation Persistence and Monetary Policy Design - An Overview. *ECB Working Paper, No. 539*, 2005.

A. Orphanides and J. Williams. Imperfect Knowledge, Inflation Expectations, and Monetary Policy. *Finance and Economics Discussion Series, Federal Reserve Board, No. 27*, 2002.

S. Pollock. Recursive estimation in econometrics. *Queen Mary University of London, Department of Economics*, Working Paper 462, 2002.

C. Prescott. The Multi-period Control Problem with Uncertainty. *Econometrica*, 40(6):1043–1058, 1972.

G. Rudebusch and L. Svensson. Eurosystem monetary targeting: Lessons from u.s. data. *European Economic Review*, 46, 2002.

T. Sargent. *Dynamic Macroeconomic Theory*. Harvard University Press, 1987.

T. Sargent. *The Conquest of American Inflation*. Princeton University Press, New Jersy, 1999.

E. Schaling. Dynamic Programming versus Lagrange: Is There a Horse Race in Dynamic Optimization? In T. van Schaik H. Klok and S. Smulders, editors, *Economologoues– Liber Amicorum voor Theo van de Klundert*, pages 334–349. Tilburg University Press, Tilburg, 2001.

E. Schaling. The Nonlinear Phillips Curve and Inflation Forecast Targeting: Symmetric Versus Asymmetric Monetary Policy Rules. *Journal of Money, Credit and Banking*, 36(3):361–386, 2004.

L. Svensson. Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets. *European Economic Review*, 41:1111–1146, 1997.

M. F. Tesfaselassie. *Communication, Learning and Optimal Monetary Policy*. CentER Dissertation Series, No.155, Tilburg University, Tilburg, 2005.

V. Wieland. Learning by doing and the value of optimal experimentation. *Journal of Economic Dynamics and Control*, 24, 2000a.

V. Wieland. Monetary policy, parameter uncertainty and optimal learning. *Journal of Monetary Economics*, 46(1):199–228, 2000b.

V. Wieland. Monetary Policy and Uncertainty about the Natural Unemployment Rate: Brainard-Style Conservatism versus Experimental Activism. *Advances in Macroeconomics*, 6(1):1288–1288, 2006.

J. Yetman. Probing potential output: Monetary policy, credibility, and optimal learning under uncertainty. *mimeo, University of Hong Kong*, 2002.