Learning About the Term Structure and Optimal Rules for Inflation Targeting *

Mewael F. Tesfaselassie†, Eric Schaling‡, Sylvester Eijffinger§

Abstract

In this paper we incorporate the term structure of interest rates in the New Keynesian model and analyze optimal policy under uncertainty about private sector expectations and the degree of inflation persistence. The novel result of our paper is that, for large deviations of inflation from its target, the active learning policy is less activist—in the sense of responding less aggressively to the state of the economy—than a myopic policy, which ignores the learning channel. Moreover, for most initial beliefs the incentive for active learning increases as monetary policy’s leverage over the long-term interest rate increases.

*Eric Schaling thanks CentER for Economic Research at Tilburg University and the Research Department of the Bank of Finland for hospitality during the formative stages of the research for this paper. Earlier versions of this paper have been presented at the Bank of Finland CEPR Workshop ‘Heterogeneous Information and Modeling of Monetary Policy’, 2-3 October 2003 Helsinki, Finland, Tilburg University, the Bank of England, the University of Pretoria, the 2004 North American Summer Meeting of the Econometric Society and the 2007 SCE meeting in Montreal. We thank Seppo Honkapohja, Martin Ellison, James Bullard, Volker Wieland and seminar/conference participants for helpful comments. Finally, Eric Schaling thanks Economic Research South Africa (ERSA) for financial support.

†Corresponding author: Kiel Institute for the World Economy (E-mail: mewael.tesfaselassie@ifw-kiel.de).

‡Wits Business School and CentER for Economic Research, Tilburg University (E-mail: eric.schaling@gmail.com).

§CentER for Economic Research, Tilburg University, RSM Erasmus University, CESifo and CEPR (E-mail: s.c.w.eijffinger@uvt.nl).
1 Introduction

Inflation targeting has received considerable attention since it was first introduced in the early 1990s. The strategy involves among other things the announcement of an explicit inflation target and monetary policy decisions based on inflation forecasts. An increasing number of papers have been devoted to studying inflation targeting. Among the early contributors in modeling inflation targeting are Svensson (1997) and Svensson (1999). However, these early models sidestep the issue of learning and control, also known as active learning. In recent years economists have developed general equilibrium models based on microfoundations, the most popular being sticky price New Keynesian models. Moreover, the issue of learning under uncertainty has received renewed attention, (see, e.g., Evans and Honkapohja 2001, 2003, Svensson and Williams 2007).

Also, the past few years have seen a revival of interest in the importance of the term-structure of interest rates for the transmission of monetary policy. When short-term inflationary expectations are given, a particular level of the central bank’s key interest rate will pin down the short-term real interest rate. According to the expectations hypothesis of the term structure (e.g., Shiller 1979), the current short-term real rate and market expectations concerning future short-term real rates then determine the long real rate. This long-term real rate term will, in turn, affect the determinants of aggregate demand. Recent research on the term structure (e.g., McGough, Rudebusch and Williams 2005) has focused on analyzing indeterminacy...
of rational expectations equilibria in the context of New Keynesian models.

In this paper, we extend the New Keynesian model (e.g., Clarida, Gali and Gertler 1999, Woodford 2003) in two respects. First, we broaden the transmission mechanism of monetary policy by introducing the term structure, which measures the degree of leverage of monetary policy over the economy. Moreover, the central bank faces uncertainty about private sector expectations of the long-term real interest rate, which is then shown to give rise to uncertainty about the degree of inflation persistence in the economy. We assume that the central bank wants to learn about private sector expectations of the long-term interest rate. Second, we derive optimal policy under active learning and analyze how the active learning policy is affected by the degree of leverage of monetary policy.¹

We compare the active learning policy to a myopic policy, a benchmark case which ignores learning. The interest in the myopic policy was first motivated by Brainard (1967), where it is shown to differ from the certainty equivalence policy by allowing for parameter uncertainty. Brainard analyzes the case where a decision maker does not know the policy multiplier, that is the effect of the control variable on the state variable. Some recent studies on optimal monetary policy under parameter uncertainty commonly follow Brainard (1967) but allow for learning² (e.g., Wieland 2000, Ellison and Valla 2001, Yetman 2003, Ellison 2006, Svensson and Williams 2007).³ However, most of these studies assume a static process for

¹Our paper focuses on central bank learning. Recent papers that have applied private sector learning in a monetary policy context are Bullard and Mitra (2002), Bullard and Schaling (2009), and Orphanides and Williams (2004). An excellent introduction to—and survey of—what is called the learning paradigm is presented in Evans and Honkapohja (2001), Evans and Honkapohja (2003) and Bullard (2006).

²Schaling (2004) analyzes Brainard-type uncertainty with a non-linear Phillips curve but no possibility for learning. He shows that optimal policy is more aggressive than implied by certainty equivalence.

³For an early account of learning and control under parameter uncertainty, see Prescott (1972). The dual control literature has shown that the so-called separation principle may not hold, and a trade-off between estimation and control arises because current actions influence estimation (learning) and provide information that may improve future performance.
the target variable, with no endogenous persistence. Beck and Wieland (2002) and Wieland (2006) use a generic model with endogenous persistence but assume that the degree of persistence is known to the decision maker. Therefore, our paper provides insight on how optimal policy behaves under uncertainty about the degree of persistence, in particular in light of the recent revival of discussions on the implications of inflation persistence (for an overview see, e.g., Levin and Moessner 2005).

The novel result of our paper is that, unlike the case where the central bank is learning about the policy multiplier (see, e.g., Beck and Wieland 2002), for large deviations of the state from the target, the active learning policy is less activist—in the sense of responding less aggressively to the state of the economy—than the myopic policy. This result holds for a wide range of possible initial beliefs about the degree of persistence. For small to moderate deviations of the state variable from its target, the optimal policy under active learning is more activist than the myopic rule. Moreover, it turns out that for most initial beliefs experimentation increases as monetary policy’s leverage over the long-term interest rate increases (from the term structure, this means that the long-term interest rate depends less on private sector expectations of future short real rates).

By analyzing optimal policy under central bank learning, our paper complements existing studies which analyze private sector learning but abstract from optimal policy (e.g., Bullard and Mitra 2002) or abstract from central bank learning (e.g., Gaspar, Smets and Vestin 2006).

The remainder of this paper is organized as follows. Section 2 presents the details of the model. Section 3 analyzes inflation targeting under rational expectations.

---

4The Gaspar, Smets and Vestin paper is concerned with private sector learning instead of central bank learning. This means that they analyze optimal monetary policy under what they call “sophisticated central banking”: that is, the central bank perfectly observes private sector expectations, even when those expectations are formed using a recursive (learning) algorithm.
while section 4 extends the problem by allowing for uncertainty and learning, discussing two policy scenarios—the myopic rule and the dynamically optimal rule. It also conducts sensitivity analysis with respect to important parameters. Section 5 concludes.

2 The Environment

In this section we layout the building blocks of the model—the New Keynesian Phillips and IS curves, as well as the term structure equation.

The New Keynesian model allows for a central role for forward-looking inflation and output expectations in the transmission mechanism. We work with the hybrid New Keynesian Phillips curve (see, e.g., Clarida, Gali and Gertler 1999, Walsh 2003, Woodford 2003).\footnote{For a discussion of the empirical fit of the hybrid Phillips curve under learning, see Milani (2007).}

\[
\pi_t = \phi \pi_{t-1} + (1 - \phi) \delta \hat{E}_t \pi_{t+1} + \alpha_1 z_t + \eta_t \tag{1}
\]

where the hat sign ‘\(^\hat{\cdot}\)’ denotes subjective (possibly nonrational) expectations. As pointed out by Clarida, Gali and Gertler (1999), in the New Keynesian model, the IS equation is obtained by log-linearizing the consumption Euler equation that arises from households’ optimal saving decisions. Translated into this paper’s notation the New Keynesian IS equation is given by

\[
z_t = \hat{E}_t z_{t+1} - \beta_2 r_t + d_t \tag{2}
\]

After recursive forward substitution, and assuming \(d_t\) is i.i.d., we get

\[
z_t = -\beta_2 \sum_{\tau=t}^{\infty} \hat{E}_t r_\tau + d_t \tag{3}
\]
(Woodford 2003, p. 244) and (Clarida, Gali and Gertler 1999, p. 1689) interpret the term \( \sum_{\tau=t}^{\infty} \hat{E}_t r_\tau \) in (3) as the long real rate. In order to relate the output gap to an \( n \)-period real bond rate, \( R_t \), Kozicki and Tinsley (2008) approximate the term \( \sum_{\tau=t}^{\infty} \hat{E}_t r_\tau \) by a simple average \( R_t = (1/n) \sum_{\tau=t}^{t+n-1} \hat{E}_t r_\tau \) so that

\[
  z_t = -\beta_2 R_t + d_t
\]

We follow this approximation but we use a weighted average representation, which is commonly used in the term structure literature (e.g., Shiller 1979, Fuhrer and Moore 1995, McGough, Rudebusch and Williams 2005).\(^6\) The weighted average representation is derived from an arbitrage equation that equalizes the real yield to maturity on a one-period bond to the one-period holding return on a long-term bond.

\[
r_t = R_t - D(\hat{E}_t R_{t+1} - R_t)
\]

where \( \hat{E}_t R_{t+1} \) denotes private sector expectations of next period’s long real rate.\(^7\) The parameter \( D \) is defined such that \( D+1 \) is equal to what is known as Macaulay’s duration (for details see Eijffinger, Schaling and Verhagen 2000). For our purposes it turns out to be convenient to rewrite equation (5) so that the current long real rate is expressed as a linear combination of \( r_t \) and \( \hat{E}_t R_{t+1} \)

\[
  R_t = (1-k)r_t + k\hat{E}_t R_{t+1} + \zeta_t
\]

where \( k \equiv D/(D+1) \). We have added a normally distributed white noise term \( \zeta_t \), where \( \zeta_t \sim N(0, \sigma^2_\zeta) \), to capture an unobserved term premium.\(^8\)

\(^6\)Thus long rates are directly relevant to spending decisions. This is a special case and not a prediction of the microfoundations underlying the New Keynesian model.

\(^7\)This will be relevant for the case of optimal policy under learning.

\(^8\)Note that equation (6) can be rewritten as \( R_t = (1-k)r_t + (1-k) \sum_{\tau=t+1}^{\infty} k^{\tau-t} \hat{E}_t r_\tau + \zeta_t \). 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. It follows that \( \hat{E}_t R_{t+1} = ((1-k)/k) \sum_{\tau=t+1}^{\infty} k^{\tau-t} \hat{E}_t r_\tau \).

6

7

8
Finally, the current short-term real interest rate is related to the short-term nominal interest rate by the Fisher equation

\[ r_t = i_t - \hat{E}_t \pi_{t+1} \] (7)

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 \( t \).

Combining equations (1), (4), (6) and (7) we get the following consolidated constraint

\[ \pi_t = \phi \pi_{t-1} + ((1 - \phi) \delta + \beta) \hat{E}_t \pi_{t+1} - \alpha_1 \beta_2 k \hat{E}_t R_{t+1} - \beta i_t + u_t \] (8)

where \( \beta = \alpha_1 \beta_2 (1 - k) \) and \( u_t = \alpha_1 d_t + \eta_t - \alpha_1 \beta_2 \zeta_t \) is a composite shock.

### 3 Inflation Targeting under Rational Expectations

It remains to specify the preferences of the central bank. As is standard in the dual control literature, we assume that the central bank loss function includes the state variable, \( \pi_t \), and the control variable, \( i_t \). The central bank chooses a sequence of current and future short-term nominal interest rates that minimize its intertemporal loss.

\[ \min_{\{i_t\}^\infty_{t=1}} E_t \sum_{\tau=t}^\infty \delta^{\tau-t-1} \frac{1}{2} [ (\pi_\tau - \pi^*)^2 + \lambda i_\tau^2 ] \] (9)

subject to equation (8). Here, \( \pi^* = 0 \) is the central bank’s inflation target, \( \delta \) is the discount factor with \( 0 < \delta < 1 \) and \( \lambda > 0 \) is the relative weight assigned to the loss from the variability in \( i_t \), the control variable. This follows the learning and control literature (see, e.g., Beck and Wieland 2002) and will become clear in
the case of optimal policy under learning.\(^9\) 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.\(^{10}\) Under rational expectations, the central bank and the private sector have identical information sets, implying identical expectations. Thus one can use the two expectations operators \(E_t\) and \(\hat{E}_t\) interchangeably.

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.

**The Rational Expectations Solution**

In each period, the sequence of events is as follows. First, the private sector sets \(\hat{E}_tR_{t+1}\) and \(\hat{E}_t\pi_{t+1}\). Then the central bank chooses \(i_t\) determining \(R_t\) and \(z_t\). Finally the composite shock \(u_t\) realizes determining \(\pi_t\). Table 1 summarizes the sequence of events.

The timing of events implies that there is a control lag within period \(t\). While \(i_t\) is chosen at the beginning of period \(t\), \(\pi_t\) realizes at the end of period \(t\). The

---

\(^9\)Policy makers often think in terms of simple loss functions that depend on output, inflation and interest rate variation. We are abstracting from output variation on the ground of tractability. Note that interest-rate stabilization is also relevant in models with cost channels in production decisions or credit intermediation.

\(^{10}\)So, the central bank knows how much policy is ‘in the pipeline’ according to financial markets.
shocks occur in between. Under parameter uncertainty, the above timing convention prevents the central bank from backing out the unknown parameter from observable variables.\footnote{See section 5 of Clarida, Gali and Gertler (1999) for a discussion of optimal monetary policy under incomplete information.} Under rational expectations and symmetric information, the solution will take the following form

\[ \pi_t = \mu_1 \pi_{t-1} + u_t \Rightarrow \hat{E}_t \pi_{t+1} = \mu_1^2 \pi_{t-1} \]  

(10)

\[ R_t = \mu_2 \pi_{t-1} \Rightarrow \hat{E}_t R_{t+1} = \mu_2 \mu_1 \pi_{t-1} \]  

(11)

\[ i_t = \mu_3 \pi_{t-1} \]  

(12)

where \( \mu_1, \mu_2 \) and \( \mu_3 \) are functions of the model parameters. The private sector expectations are consistent with the solution for the long-term real interest rate implied by flexible inflation targeting.

4 Inflation Targeting under Learning

The previous section looked at the benchmark case of rational expectations where the central bank perfectly observes and responds to private sector expectations of the long real interest rate and of inflation. However, as noted by Evans and Honkapohja (2003), central banks typically observe private sector expectations with error, and these observation errors can be large.\footnote{We continue to assume that the central bank knows the structural model of the economy, including the parameters.} We now look at the more realistic case where the central bank has imperfect knowledge of private sector expectations.

Following the literature on learning (see, e.g., Evans and Honkapohja 2001), we suppose that the private sector’s forecasting functions for the inflation rate and the long real interest rate take the same form as the rational expectations solution
under full information, namely, equations (10) and (11). The central bank knows only that the private sector’s forecasting rules will be of the form $\hat{E}_t \pi_{t+1} = \gamma_1 \pi_{t-1}$ and $\hat{E}_t R_{t+1} = \gamma_2 \pi_{t-1}$.\(^\text{13}\)

Using these expressions in equation (8), the central bank’s perceived law of motion of inflation is

$$\pi_t = \gamma \pi_{t-1} - \beta i_t + u_t \tag{13}$$

where $\gamma = \phi + ((1 - \phi) \delta + \beta) \gamma_1 - \alpha_1 \beta_2 k \gamma_2$. As far as the central bank’s learning is concerned, the relevant unknown parameter is $\gamma$.\(^\text{14}\) Let $c_t$ denote the best forecast of $\gamma$ and $p_t$ the variance of $c_t$, (the degree of confidence placed upon $c_t$).\(^\text{15}\)

\subsection{4.1 Central Bank Beliefs and Updating}

Note that the unobserved shock, $u_t$, injects additional uncertainty and prevents the central bank from inferring, in any period, the value of $\gamma$ from equation (13). Under learning the relevant stochastic process is equation (13). Our setup resembles that of the dual control literature, which usually uses a constraint similar to equation (13) (see, e.g., Beck and Wieland 2002). However, in our case the unobservability of private sector expectations translates into uncertainty about the persistence parameter $\gamma$. Unlike Beck and Wieland (2002), we assume that the policy maker knows $\beta$.

At the end of period $t$, the central bank updates its prior estimate by including the latest available data ($\pi_t, \pi_{t-1}, i_t$) in the regression equation (13). Using

\(^{13}\)Under rational expectations, the $\gamma$s are functions of the model parameters.

\(^{14}\)It is not necessary for the central bank to separately estimate $\gamma_1$ and $\gamma_2$.

\(^{15}\)By positing a simple forecasting function for the private sector, we abstract from the interaction of optimal monetary policy under imperfect information and private sector learning. However, even though private sector 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.
the widely used method of recursive least squares we have the following updating
equations for $c_t$ and its variance, denoted by $p_t$ (see, e.g., Beck and Wieland 2002,
Pollock 1999).

\begin{align*}
  c_{t+1} &= c_t + \kappa_t (\pi_t - c_t \pi_{t-1} + \beta i_t) \\
  p_{t+1} &= p_t - \kappa_t \pi_{t-1} p_t
\end{align*}

(14)

where $\kappa_t \equiv \pi_{t-1} p_t F_t^{-1}$ is commonly referred to as the Kalman gain, which is the
weight assigned to new information coming from the forecast error $\pi_t - c_t \pi_{t-1} - 
\beta i_t$. The Kalman gain in turn depends on the conditional variance of $\pi_t$, (based
on information in period $t$), given by $F_t \equiv \pi_{t-1}^2 p_t + \sigma^2_{\nu}$. As can be seen from
equation (14), the current state of the economy, $\pi_{t-1}$ is part of the Kalman gain,
and affects the path of the conditional variance of the parameter estimate, $p_{t+1}$.
Due to the presence of autoregressive behavior in $\pi_t$, changes in the current state
of the economy, $\pi_{t-1}$, have direct effects on the variability of $\pi_t$, and consequently
on the variability of $\pi_{t+1}, \pi_{t+2}, \ldots, \pi_{t+\infty}$.

The updating equations in (14) capture the idea that central bank learning about
the unknown parameter is influenced by policy decisions made in period $t$, $i_t$. This
channel works as follows: $i_t$ affects the state variable in period $t+1$, $\pi_{t}$, and conse-
quently, beliefs about the unknown parameter (i.e. $c_{t+2} = c(\pi_{t})$ and $p_{t+2} = p(\pi_{t})$).
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

\textsuperscript{16}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).

\textsuperscript{17}The assumption that the shock is normally distributed with known variance is standard in
the learning literature. If the prior belief has a normal distribution, then the posterior belief also
follows a normal distribution. This property of the posterior belief is convenient when dealing
with numerical computations (see the appendix to chapter 6 of Tesfasellassie 2005).
of the estimate $c_{t+1}$ depends positively on the variance of $\pi_t$. One gets a more precise estimate (in other words, a smaller value of $p_{t+1}$) when the variance of $\pi_t$ increases, and vice versa. Since we recognize that the current choice of monetary policy $i_t$ affects $E_t\pi_t$, and given that $F_t$ is predetermined, the coefficient of variation, defined by the ratio $\sqrt{F_t}/E_t\pi_t$ is also a function of $r_t$. The relationship between $i_t$, $\pi_t$ and $p_{t+2}$ raises a potential tension between the urge to minimize current period loss from variability in $i_t$ (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).

Under parameter uncertainty, one has to differentiate between three policy rules: certainty equivalence, myopic, and optimal (see 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. Finally, the optimal policy incorporates active learning.

### 4.2 Derivation of Optimal Policy: Passive Learning

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 $i_{t+1}$ in period $t+1$ the central bank updates its belief about $\gamma$ to $c_{t+1}$, ignoring the fact that it will
have to update this estimate in the future.\textsuperscript{18} In this way the central bank fails to internalize the effect of current actions on future beliefs.

This section solves for optimal policy under passive learning, namely, one that incorporates parameter uncertainty but ignores learning (henceforth the \textit{myopic} rule). This rule is passive because it ignores 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 \((c_t, p_t)\) to remain fixed for all future periods. Consequently, the non-linear updating equations drop out of the optimization problem.\textsuperscript{19}

It is important to observe that, even though we have placed time subscripts on \(c_t\) and \(p_t\), as far as passive learning is concerned, \(c_t\) and \(p_t\) should be thought of as fixed parameters (not state variables) and not expected to be affected by the current policy choice.\textsuperscript{20} This is the sense in which forecasting and control are separated by construction.\textsuperscript{21} A special case of the myopic rule is certainty equivalence, where parameter uncertainty is ignored (i.e., \(p_t = 0\)).\textsuperscript{22} Therefore, we do not treat certainty equivalence in our core analysis.

\textsuperscript{18}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.

\textsuperscript{19}Of course, when next period arrives, the bank updates its belief but then expects it to remain fixed from that period on.

\textsuperscript{20}One can also think of \(c_t\) and \(p_t\) as state variables. However, here these states do not change and are independent from the policy instrument. This means that, when optimizing, the terms \(\kappa_t(\pi_t - c_t\pi_{t-1} + \beta_i)\) and \(\kappa_t\pi_t - p_t\) on the right hand side of the updating equations drop out.

\textsuperscript{21}Thus, due to the sequential nature of decision making, the updating of the parameter estimate is kept in the background.

\textsuperscript{22}The implication of this can be seen by decomposing \(E_t\pi_t^2\) as \(E_t\pi_t^2 = (E_t\pi_t)^2 + F_t\), where \(E_t\pi_t = c_t\pi_{t-1} - \beta_i\) and under the myopic policy \(F_t = \pi_t^2 + \sigma_u^2\) (assuming \(u_t\) is i.i.d., uncorrelated with the estimation error \((\gamma - c_t)\)). On the other hand, under certainty equivalence, \(p_t = 0 \Rightarrow F_t = \sigma_u^2\); \(F_t\) is thus completely exogenous.
The minimization problem is then
\[
\min_{\{i_t\}_{t=1}} \mathbb{E}\left[ \sum_{\tau=t}^{\infty} \delta^{\tau-t} \frac{1}{2} (\pi_\tau^2 + \lambda \tau_r^2) | \pi_{t-1} \right]
\] (15)
subject to the linear constraint (13). Importantly, with the non-linear updating equations ignored, the problem is linear-quadratic, and the derivation of the optimal level of \(i_t\) is similar to that under rational expectations (see the derivations in Appendix A and Appendix B)

\[
i_t = \frac{c_t \beta (1 + \delta \mu)}{\lambda + \beta^2 (1 + \delta \mu) \pi_{t-1}}
\] (16)

where

\[
\mu = -\frac{1}{2} \left( \frac{A + B - \sqrt{B^2 + C}}{G} \right)
\] (17)

with \(A = \beta^2 (1 - 2 \delta p_t), B = \lambda (1 - \delta (c_t^2 + p_t)), C = \beta^4 + 2 \beta^2 \lambda (1 + \delta (c_t^2 - p_t)),\) and \(G = \beta^2 \delta (1 - \delta p_t).\)

When choosing its interest rate under the myopic policy, the central bank behaves as if its initial belief 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 \(c\) will increase (\(p\) will decline) with the arrival of new economic data.

Under the myopic policy, \(i_t\) responds more strongly to \(\pi_{t-1}\) the larger is \(p_t\). What is the intuition? The control problem is dynamic, so the bank expects to incur losses from variability in \(\pi_t\) (via its effect on \(F_{t+1} = p_{t+1} \pi_t^2 + \sigma_u^2\)). Since as of period \(t\) the central bank does not internalize the effect of policy on future beliefs, we have \(p_{t+1} = p_t\) and \(E_t F_{t+1} = p_t E_t \pi_t^2 + \sigma_u^2 = p_t (E_t \pi_t)^2 + p_t F_t + \sigma_u^2\), which shows the benefits from a policy that sets \(E_t \pi_t\) closer to zero. It follows that, given \(\lambda > 0\), from the
perspective of the myopic policy $i_t$ should respond more strongly to deviations of $\pi_{t-1}$ so that $E_t \pi_t$ is closer to the target $\pi^*$ (zero) than implied by $p_t = 0.23$

It is important to note that $i_t$ can affect only $E_t \pi_t$. On the other hand, the conditional variance, $F_t$, is independent of $i_t$ since $p_t$ and $\pi_{t-1}$ 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, in which $\beta$ 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 $\phi$ is unknown while $\beta$ is known with certainty.

4.3 Derivation of Optimal Policy: 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 rate choices of the central bank, including the current one. Under optimal policy, the policy maker can take actions now such that $\pi_t$ 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 today 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

---

23Note that $p_t$ scales up the component of the loss function associated with the variability of $\pi_t$. The presence of $p_t$ does not matter under strict inflation targeting since the policy instrument is set such that $E_t \pi_t = 0$. By contrast, with flexible inflation targeting, $\lambda > 0$ implying that $E_t \pi_t \neq 0$. The presence of $p_t$ then requires a more aggressive policy that drives $E_t \pi_t$ closer to zero.
parameter. Under flexible inflation targeting, optimal policy exploits this tradeoff between those future benefits and current costs from large movements in $i_t$.

The central bank chooses a sequence of current and future short-term nominal interest rates to minimize the intertemporal loss function

$$
\min_{\{i_t\}_{t=1}^{\infty}} E \left[ \sum_{\tau=1}^{\infty} \delta^{\tau-t} \frac{1}{2} (\pi_\tau^2 + \lambda i_\tau^2) \right](\pi_{t-1}, c_t, p_t) \right]
$$

(18)

Here, there are three constraints—the linear Phillips curve (13) and (importantly) the two non-linear updating equations (14). The effect of current policy on future beliefs becomes visible from the Bellman equation associated with the relevant dynamic programming problem (see, e.g., Beck and Wieland 2002)

$$
V(\pi_{t-1}, c_t, p_t) = \min_{i_t} E_t \left[ \frac{1}{2} (\pi_{t+1}^2 + \lambda i_{t+1}^2) + \delta V(\pi_t, c_{t+1}, p_{t+1}) \right]
$$

(19)

subject to equations (13) and (14). The second term on the right side of equation (19) is the expected discounted loss from period $t + 1$ onwards. It is a function of time $t + 1$ state vector $(\pi_t, c_{t+1}, p_{t+1})$, which in turn depends on the current policy rate $i_t$ and the current state vector $(\pi_{t-1}, c_t, p_t)$. This term thus captures the value of information and is given by (see also the appendix to chapter 6 of Tesfaselassie 2005)

$$
E_t V(.) = \int V \left( \pi_t, c(\pi_t, c_t, p_t, \pi_{t-1}), p(p_t, \pi_t) \right) f(\pi_t | .) d\pi
$$

(20)

where $f(\pi_t | .)$ is the conditional density function of $\pi_t$. Because of the non-linear updating equations 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

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 the present paper since we have only three state variables, $c_t$, $p_t$ and $\pi_{t-1}$.

### 4.4 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, $\lambda$, the long-term bond duration measure $k$ (which pins down the reduced-form parameter $\beta$ for given values of $\alpha_1$ and $\beta_2$), the discount factor, $\delta$, and the variance of the composite exogenous shock, $\sigma_u^2$. Then we compare these results with those derived for alternative parameter values (see Table 2).

We set $\delta = 0.99$, $\alpha_1 = 0.3$, and $\beta_2 = 1$. The calibrated values for $\alpha_1$ and $\beta_2$ are taken from Clarida, Gali and Gertler (1999), and the values of $k = 0.8$ and $k = 0.75$ are chosen such that the long-term bond has duration $D = 4$ and $D = 2$ (in quarters) respectively (McGough, Rudebusch and Williams (2005) and Kozicki and Tinsley (2008) consider the case $D = 2$).

Possible initial beliefs for $c$ are as high as 1.4 and as low as 0.2, while the beliefs about the variance $p$ range from 0.1 to 0.7. The relative degree of confidence in an estimate measured by the coefficient of variation, $\sqrt{p}/c$, takes its lowest value at
\((c_0, p_0) = (1.4, 0.1)\) implying \(\sqrt{p}/c < 0.23\). In this case, the uncertainty associated with \(c\) 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 \((c_0, p_0)\) are considered from the sets \(c_0 \in \{0.2, 0.6, 1, 1.4\}\) and \(p_0 \in \{0.1, 0.3, 0.5, 0.7\}\). All the figures shown below have \(\pi_{t-1}\) on the horizontal axis.

**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 \(i_t\) to deviations of the state \(\pi_{t-1}\) from its target level of zero, for a specific belief characterized by the mean \(c_0 = 1\) and variance \(p_0 = 0.5\).

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

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, optimal policy responds less aggressively
compared to the myopic policy. The intuition for this is the following. From the updating equations, the larger the deviation of $\pi_{t-1}$ from zero (due to say an exogenous shock), the smaller $p_{t+1}$ (implying that $c_{t+1}$ is a more precise estimate) leading to a smaller control error when setting $i_{t+1}$. Thus, in contrast to the myopic policy, the actively learning central bank anticipates future improvements in policy performance as $|\pi_{t-1}|$ increases. This shows that when realized exogenous shocks, $u_t$, which ultimately drive data generation for $\pi_{t-1}$, are large there is less of a role for deliberate probing by the central bank, and the more so the larger the deviation of $\pi_{t-1}$ from the target.

**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, $c_0$ and $p_0$.\(^{25}\) Perhaps not surprisingly, the three rules diverge with the magnitude of $p_0$ and $c_0$. For example, when parameter uncertainty is small (say $p_0 = 0.1$) and there is a small degree of expected persistence in the economy ($c_0 = 0.2$), the three decision rules tend to be identical. At the other extreme, when both $c_0$ and $p_0$ 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 $\beta$ and $\lambda$. Thus when analyzing the effect of changes in these parameters, it is not proper to

\(^{24}\)At the same time, of course, $F_t$, the conditional variance of $\pi_t$, increases with $\pi_{t-1}^2$ but this component of $E_t \pi_t^2$ is independent of $i_t$.

\(^{25}\)Because the interest rate 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.
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 $\sigma_u^2$ 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 $\sigma_u^2$.

**Policy under Less Volatile Shocks**

First, we look at the effect of a decrease in the variance of the exogenous random shock, $\sigma_u^2$. In the benchmark case, the variance was set to 0.2, while now the variance is reduced by half, standing at 0.1, 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 $\sigma_u^2$, as the shock enters the model additively. So only the optimal decision rules for the two alternative values of $\sigma_u^2$ are shown in Figure 3 below.

In each panel of the figure the benchmark case ($\sigma_u^2 = 0.2$) is shown by the dotted line while the solid line arises from the smaller variance of the shock ($\sigma_u^2 = 0.1$). As is apparent from most of the panels, *optimal policy* is more aggressive when the variance $\sigma_u^2$ decreases, that is when the shocks driving the $\pi_t$ process are less volatile. This is especially the case with large values of $c$ and $p$ 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 $\pi_t$, defined here as $CV = \sqrt{F_{t+1}/E_t}\pi_t$. Since $F_t$ is predetermined, the CV can be increased only by a lower value of $E_t\pi_t$, which in turn requires a more aggressive response of $i_t$ to $\pi_{t-1}$.\(^{26}\)

\(^{26}\)Quantitatively, this difference is not large when $c$ or $p$ is small.
The Effect of a Larger Policy Leverage

Figure 4 shows how the degree of probing is affected by the size of $\beta \in \{0.06, 0.075\}$ corresponding to $k \in \{0.8, 0.75\}$. The solid line is associated with $\beta = 0.075$, while the dotted line is associated with $\beta = 0.06$. From equation (13), we know that $\beta \equiv \alpha_1 \beta_2(1 - k)$; thus a larger value of $\beta$ is associated with a stronger leverage of monetary policy on the current long real rate, (from the term structure equation, the duration of the long-term bond decreases and the weight on the short real rate increases). In other words, as $1 - k$ increases, the model becomes less forward-looking since the long-term interest rates will be determined less by movements in private sector expectations. Due to this fact, the role of experimentation 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 values of $\pi_{t-1}$, experimentation increases with $\beta$. For most of the alternative initial beliefs, experimentation is maintained even for larger values $\pi_{t-1}$ when $\beta = 0.075$. The effect on the relative response of the optimal policy of a change in $\beta$ diminishes if $c_0$ or $p_0$ is sufficiently small.

Policy Cares More about Inflation Stabilization

We also examine optimal policy when $\lambda$, the weight placed on interest rate stabilization, is reduced from 0.5 to 0.1. As Figure 5 shows, when $\lambda = 0.1$ the active learning policy is more aggressive than the myopic policy for small to moderate values of $\pi_{t-1}$, although less strongly than when $\lambda = 0.5$. Thus the incentive for the
active learning policy to deviate from the myopic policy diminishes when inflation stabilization receives more attention. In the (somewhat unrealistic) limit of strict inflation targeting optimal policy is not affected by uncertainty in the persistence parameter.

Note also that, for small to moderate values of the state variables, Figure 5 is similar to Figure 4, implying that an increase in $\beta$ has effects on the degree of experimentation similar to an increase in $\lambda$.

5 Concluding Remarks

The paper analyzes optimal monetary policy under active learning in a New Keynesian model augmented by the term structure of interest rates, which captures the degree of leverage of monetary policy over the long-term real interest rate and in turn aggregate demand. We find that under flexible inflation targeting and learning about 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. In particular, for large deviations of the state from the target, the active learning policy is less activist—in the sense of responding less aggressively to the state of the economy—than the myopic policy, which ignores learning. This result holds for a wide range of possible initial beliefs about inflation persistence. For small to moderate deviations of inflation from its target, the active learning policy is more activist than the myopic rule. Moreover, experimentation tends to increases as monetary policy’s leverage over the economy increases. The intuition is that a more precise estimate of the degree of persistence requires more variability in the data coming from the economy. It follows that, the instrument of monetary policy—the short-term interest rate—needs to move more aggressively if its leverage over
the long-term interest rate and, in turn, aggregate demand increases.

By analyzing optimal policy under central bank learning, our paper complements existing studies which analyze private sector learning but abstract from optimal policy (e.g., Bullard and Mitra 2002) or from central bank learning (e.g., Gaspar, Smets and Vestin 2006). As an extension of the paper, one could allow for two-way learning, where the private sector as well as the central bank are learning about the economy. We leave this for future research.
Appendix A  Derivation of Optimal Policy under Rational Expectations

In this appendix we derive optimal policy under rational expectations, as a basis for deriving optimal policies under parameter uncertainty.

The optimization problem is

\[
\min_{\{i_t\}_{t=1}^{\infty}} E_t \sum_{\tau=t}^{\infty} \delta^{\tau-t} \frac{1}{2} (\pi_{t-1}^2 + \lambda_i^2)
\]  

(A.1)

subject to the linear constraint (13). The problem is linear-quadratic and we can rewrite the above minimization problem using recursive dynamic programming and then use the ‘guess and verify’ method on the value function. The Bellman equation associated with the minimization of (A.1) is\(^{27}\)

\[
V(\pi_{t-1}) = \min_{i_t} [L(\pi_{t-1}, i_t) + \delta E_t V(\pi_t)]
\]  

(A.2)

subject to equation (13). Because of the linear-quadratic structure of the problem, the value function will be quadratic in the state \(\pi_t\)

\[
V(\pi_{t-1}) = \mu_0 + \frac{1}{2} \mu \pi_{t-1}^2
\]  

(A.3)

where the two coefficients remain to be determined. If equation (A.3) is correct, it follows that

\[
E_t V(\pi_t) = \mu_0 + \frac{1}{2} \mu E_t \pi_t^2
\]  

(A.4)

where \(E_t \pi_t^2\) follows from equation (13). Using equation (A.4) in equation (A.2) and taking the first order condition, we get

\(^{27}\)Note that the value function in the Bellman equation does not have a time subscript. This is because in infinite horizon problems, we are interested only in the unique time 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). It is well known that for linear-quadratic control problems, convergence is achieved in a single iteration if \(V_0\) is quadratic.
\[-\beta(\gamma\pi_{t-1} - \beta i_t) + \lambda i_t - \beta \delta \mu (\gamma\pi_{t-1} - \beta i_t) = 0\]

which can be solved for \(i_t\)

\[i_t = -\frac{\gamma \beta (1 + \delta \mu)}{\lambda + \beta^2 (1 + \delta \mu)} \pi_{t-1}\]  
(A.5)

In order to identify \(\mu\), first differentiate equation (A.3) with respect to \(\pi_{t-1}\)

\[V_\pi(\pi_{t-1}) = \mu \pi_{t-1}\]  
(A.6)

Next, invoking the envelope theorem on the Bellman equation (A.2), and taking note of equation (A.5)

\[V_\pi(\pi_{t-1}) = \gamma (\gamma \pi_{t-1} - \beta i_t) + \delta \mu \gamma (\gamma \pi_{t-1} - \beta i_t)\]

\[= f(\mu) \pi_{t-1}\]  
(A.7)

where \(f(\mu) \equiv \frac{\gamma^2 \lambda (1 + \delta \mu)}{\lambda - \delta (\beta^2 + \gamma^2 \lambda)}\). For the conjectured value function (A.3) to be correct, the coefficients of equations (A.6) and (A.7) have to be equal. This is a fixed-point problem\(^{28}\)

\[\mu = f(\mu)\]  
(A.8)

Rearrange equation (A.8) to get the following quadratic equation for \(\mu\)

\[\delta \beta^2 \mu^2 + [\lambda - \delta (\beta^2 + \gamma^2 \lambda)] \mu - \lambda = 0\]

\(^{28}\)Alternatively, one can work directly with the value function. First substitute (A.5) in (A.2) and then match the resulting coefficients with those in the conjectured value function (A.3).
The unique solution for $\mu$ is then

$$\bar{\mu} = \frac{-A + \sqrt{A^2 + 4\beta^2\lambda\delta}}{2\beta^2\delta}$$

where $A = \beta^2 + \lambda(1 - \delta\gamma^2)$.

**Appendix B  Derivation of the Myopic Rule**

As the only state variable is $\pi_{t-1}$, the value function and the first order condition with respect to $i_t$ will take the same form as in the case of rational expectations (Appendix A), except that now $\gamma$ is replaced by its mean $c_t$. Moreover, under the myopic policy, the coefficient $\mu$ will be a function of $p_t$ as well as $c_t$. When identifying $\mu$, we remember that $E_t\pi_t^2$ is a function of $p_t$. Thus analogous to (A.7)

$$V_n(\pi_{t-1}) = c_t(c_t\pi_{t-1} - \beta i_t) + \delta \mu c_t(c_t\pi_{t-1} - \beta i_t) + (1 + \delta \mu)p_t\pi_{t-1}$$

$$= g(\mu)\pi_{t-1}$$

where $g(\mu) = \frac{c_t^2\lambda(1 + \delta \mu)}{(\lambda + \beta^2(1 + \delta \mu)) + (1 + \delta \mu)p_t} = f(\mu) + (1 + \delta \mu)p_t$. Note that $g(\mu) \geq f(\mu)$, implying that when parameter uncertainty is taken into account, the fixed point for $\mu$ is larger than the corresponding fixed point under certainty equivalence. Moreover, the myopic policy becomes more aggressive (in the sense of larger policy response to $\pi_{t-1}$) the larger is $p_t$. The solution for $\mu$ is given by equation (17) of the main text.
Table 1: Sequence of events in period $t$. 

<table>
<thead>
<tr>
<th>Event Details</th>
<th>Equation</th>
<th>Event Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Private sector sets $E_t R_{t+1}$ and $E_t \pi_{t+1}$</td>
<td>$i_t = i(\pi_{t-1}, E_t \pi_{t+1}, E_t R_{t+1})$</td>
<td>Central bank chooses $\pi_t$</td>
</tr>
<tr>
<td>$u_t$ realizes determining $\pi_t$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th>Parameter</th>
<th>Baseline value</th>
<th>Alternative value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k$</td>
<td>0.8</td>
<td>0.75</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>0.5</td>
<td>0.1</td>
</tr>
<tr>
<td>$\sigma^2_u$</td>
<td>0.2</td>
<td>0.1</td>
</tr>
</tbody>
</table>

Table 2: Parameter configuration.
Figure 1: The two decision rules under baseline parameter values and for initial beliefs $c_0 = 1$ and $p_0 = 0.5$. 
Figure 2: The two decision rules under baseline parameter values and for alternative initial beliefs
Figure 3: Performance of the active learning policy ($\sigma_u^2 = 0.2$ vs. $\sigma_u^2 = 0.1$).
Figure 4: Relative performance of the active learning policy ($\beta = 0.06$ vs. $\beta = 0.075$).
Figure 5: Relative performance of the active learning policy ($\lambda = 0.5$ vs. $\lambda = 0.1$).
References


bank Activism in an Economy with Strategic Interactions.” Journal of Monetary Economics, 48, 153-171.


McGough, Bruce, Glenn D. Rudebusch, and John C. Williams. (2005) “Using a
Long-Term Interest Rate as the Monetary Policy Instrument.”, *Journal of Monetary Economics*, 52, 855-879.


