Learning, Macroeconomic Dynamics and the Term Structure of Interest Rates

Hans Dewachter\textsuperscript{a,}\textsuperscript{*} and Marco Lyrio\textsuperscript{b}

\textsuperscript{a}Catholic University of Leuven and Erasmus University Rotterdam
\textsuperscript{b}Warwick Business School

June 16, 2006

Abstract

We present a macroeconomic model in which agents learn about the central bank’s inflation target and the output-neutral real interest rate. We use this framework to explain the joint dynamics of the macroeconomy, and the term structures of interest rates and inflation expectations. Introducing learning in the macro model generates endogenous stochastic endpoints which act as level factors for the yield curve. These endpoints are sufficiently volatile to account for most of the variation in long-term yields and inflation expectations. As such, this paper complements the current macro-finance literature in explaining long-term movements in the term structure without reference to additional latent factors.

\textsuperscript{*}Corresponding author. Address: Center for Economic Studies, Catholic University of Leuven, Naamsestraat 69, 3000 Leuven, Belgium. Tel: (+)32(0)16-326859, e-mail: hans.dewachter@econ.kuleuven.ac.be. We are grateful for financial support from the FWO-Vlaanderen (Project No.:G.0332.01). We thank the discussant, Jordi Galí, the organizer, John Campbell, and the participants at the NBER Conference on Asset Prices and Monetary Policy for very helpful comments. We also thank Konstantijn Maes, Raf Wouters, and seminar participants at the National Bank of Belgium, 2004 Conference on Computing in Economics and Finance, Heriot-Watt University, Catholic University of Leuven, Tilburg University, University of Amsterdam, European Central Bank, and University of Warwick for useful discussions and comments on a previous version of this paper. The authors are responsible for remaining errors.
1 Introduction

Since the seminal papers by Vasicek (1977) and Cox, Ingersoll, and Ross (1985), there is a consensus in the finance literature that term structure models should respond to three requirements: absence of arbitrage opportunities and both econometric and numerical tractability. Models designed to meet these criteria can be useful, for instance, in the pricing of fixed income derivatives and in the assessment of the risks implied by fixed income portfolios. More recently, however, a number of requirements have been added to the modeling of the yield curve dynamics. Satisfactory models should also (i) be able to identify the economic forces behind movements in the yield curve, (ii) take into account the way central banks implement their monetary policies, and (iii) have a macroeconomic framework consistent with the stochastic discount factor implied by the model. In this paper, we present a model that fulfills all of the above requirements and, in addition, integrates learning dynamics within this macro-finance framework.

The model presented in this paper builds on recent developments (phases) in the affine term structure literature. The first phase is characterized by the use of latent or unobservable factors, as defined in Duffie and Kan (1996) and summarized in Dai and Singleton (2000). Although this framework excludes arbitrage opportunities and is reasonably tractable, the factors derived from such models do not have a direct economic meaning and are simply labeled according to their effect on the yield curve (i.e. as a “level”, a “slope”, and a “curvature” factor).

The second phase involves the inclusion of macroeconomic variables as factors in the standard affine term structure model. Ang and Piazzesi (2003) show that such an inclusion improves the forecasting performance of vector autoregression (VAR) models in which no-arbitrage restrictions are imposed.1 Their model, nevertheless, still includes unobservable factors without an explicit macroeconomic interpretation. Kozicki and Tinsley (2001, 2002) indicate the importance of long-run inflation expectations in modeling the yield curve and connect the level factor in affine term structure models to these long-run inflation expectations. This interpretation of the level factor is confirmed by Dewachter and Lyrio (2006), who estimate an affine term structure model based only on factors with a well-specified macroeconomic interpretation. The mentioned papers do not attempt, however, to propose a macroeconomic framework consistent with the pricing kernel implied by their models.

The third and most recent phase in this line of research is marked by the use of structural macro relations together with the standard affine term structure model. The structural macro model replaces the unrestricted VAR set-up adopted in previous research, and has commonly been based on a New-Keynesian framework. Hördahl, Tristani and Vestin (2006) find that the forecasting performance of such a model is comparable to that of standard latent factor models. They are also able to explain part of the empirical failure of the expectations hypothesis. A similar approach is adopted by Rudebusch and Wu (2003). Bekkaert, Cho and Moreno (2006) go one step further and impose consistency between the pricing kernel and the macro model.

The success of macro-finance models is remarkable given the well-documented dynamic inconsistencies between the long-run implications of the macroeconomic models and the term structure of interest rates.2 In particular, standard macroeconomic models fail to generate sufficient

---

1 Other papers using this approach include Dewachter, Lyrio and Maes (2006) and Diebold, Rudebusch and Aruoba (2006).

2 For instance, Gürkaynak, Sack and Swanson (2005) and Ellingsen and Söderström (2001) show that standard macroeconomic models cannot account for the sensitivity of long-run forward rates to standard
persistence to account for the time variation at the long end of the yield curve. The success of macro-finance models in fitting jointly the term structure and the macroeconomic dynamics in fact crucially hinges on the introduction of additional inert and latent factors with a macroeconomic interpretation. For instance, Bekaert et al. (2006), Dewachter and Lyrio (2006) and Höröahl et al. (2006), among others, introduce a time-varying (partly) exogenous implicit inflation target of the central bank and show that it accounts for the time variation in long-maturity yields.

The main goal of this paper is to build and estimate macro-finance models that generate these additional factors endogenously from a macroeconomic framework. To this end, we introduce learning into the framework of standard macro-finance models. Extending macro-finance models with learning dynamics seems a promising route to model jointly the macroeconomic and term structure dynamics for two reasons. First, learning generates endogenously additional and potentially persistent factors in the form of subjective expectations. Second, learning, especially constant gain learning, introduces sufficient persistence in the perceived macroeconomic dynamics to generate a level factor in the term structure of interest rates. Such a level factor is crucial to account for the time variation in the long end of the yield curve.

Our approach connects the macro-finance models of the term structure to the learning literature. Links between learning and the term structure of interest rates are also actively analyzed in the learning literature. For example, Cogley (2005) uses a time-varying Bayesian VAR to account for the joint dynamics of macroeconomic variables and the term structure of interest rates. Kozicki and Tinsley (2005) use a reduced form VAR in macroeconomic and term structure variables and assume agents have imperfect information with respect to the inflation target. They find that subjective long-run inflation expectations are crucial in fitting movements in long-maturity yields and inflation expectations and report a substantial difference between the central bank’s inflation target and the subjective expectations of the inflation target. Orphanides and Williams (2005a) introduce long-run inflation expectations in the structural macroeconomic models by substituting expectations by and calibrating the learning parameters on observed survey data. This paper complements this recent and rapidly growing literature. First, we do not rely on reduced form VAR dynamics. Instead, we use a standard New-Keynesian model to describe the macroeconomic dimension and impose consistency of the pricing kernel for the term structure and the macroeconomic dynamics. Second, following Sargent and Williams (2005), we generate the subjective expectations based on a learning technology that is optimal given the structural equations and the priors of the agents. Third, we estimate jointly the deep parameters of the structural equations and the learning parameters. The term structure of interest rates and surveys of inflation expectations are included as additional information variables in the measurement equation. We find that the proposed model generates sufficiently volatile subjective long-run expectations of macroeconomic variables to account for most of the time variation in long-maturity yields and surveys of inflation expectations. This is achieved without reference to additional latent factors and hence offers an alternative approach to the current macroeconomic shocks.

Milani (2005) finds that the persistence in the learning dynamics is sufficiently strong to capture much of the inertia of the macroeconomic series.

Orphanides and Williams (2005a, b) using a calibrated learning model show that learning affects the long end of the term structure.

Other papers using survey expectations as proxies for the theoretical expectations include Roberts (1997) and Rudebusch (2002).
macroe–finance literature.

The remainder of the paper is divided in four sections. In Section 2, we present the macroeconomic framework, which is based on a standard New-Keynesian macro model. We introduce imperfect information with respect to the endpoints of macroeconomic variables, discuss the respective priors, and derive the optimal learning rule. The perceived and actual laws of motion are derived together with the conditions for stability of the macroeconomic dynamics. The perceived law of motion forms the basis to generate the implied term structures of interest rates and inflation expectations. The estimation methodology is presented in Section 3. Both the yield curve and surveys of inflation expectations are used as additional information variables to identify subjective expectations. In Section 4, we present the estimation results and compare the performance of the estimated models in fitting the term structure of interest rates. We show that macro–finance models, built on structural equations and learning, explain a substantial part of the time variation of long-maturity yields and inflation expectations. We conclude in Section 5 by summarizing the main findings of the paper.

2 Macroeconomic dynamics

We use the standard monetary three-equation New-Keynesian framework as presented in, for instance, Hördahl et al. (2006), Bekaert et al. (2006) and Cho and Moreno (2006). These models can be considered as minimal versions of a fully structural model (e.g. Christiano, Eichenbaum and Evans 2005, Smets and Wouters 2003) and commonly represent the benchmark model in the literature linking macroeconomic dynamics and the term structure. We follow the standard Euler-equation procedure employed in the learning literature (Bullard and Mitra 2002, and Evans and Honkapohja 2001) and replace the rational expectations operator by a subjective expectations operator. In the model presented below, subjective expectations differ from rational expectations because we assume that agents do not observe the inflation target of the central bank nor the equilibrium output-neutral real interest rate. Finally, in Section 2.3 we solve for the macroeconomic dynamics, i.e. the actual law of motion. The solution is given in the form of a reduced VAR(I) model in an extended state space.

2.1 Structural equations

The structural model used is a parsimonious three-equation representation of the underlying macroeconomic structure, containing aggregate supply (AS) and IS equations and a monetary policy rule identifying the riskless nominal interest rate. To account for the persistence in inflation, the output gap and the policy rate, we add inflation indexation, habit formation, and interest rate smoothing to the standard model.

The design of the AS-equation is motivated by the sticky-price models based on Calvo (1983). In line with the standard Calvo price-setting theory, we assume a world where only a fraction of the firms updates prices at any given date, while the non-optimizing firms are assumed to use some rule of thumb (indexation scheme) in adjusting their prices (e.g. Galí and Gertler 1999). This setting leads to a positive relation between (transitory) inflation on the one hand and real marginal costs on the other. Specific assumptions are made with respect to the marginal costs

6 An alternative micro-founded approach to learning has been developed by Preston (2005). We leave this extension for future research.
and the indexation scheme of the non-optimizing agents. First, we assume that marginal costs are proportional to the output gap and an additional cost-push shock, $\varepsilon_t$. Second, non-optimizing firms are assumed to adjust prices according to an indexation scheme based on past inflation rates. The degree of indexation is measured by the parameter $\delta_T$ and the indexation scheme at time $t$ is given by $\pi^* + \delta_T(\pi_{t-1} - \pi^*)$ with $\pi^*$ the inflation target and $\pi_{t-1}$ the previous period inflation rate.

Following these assumptions, the standard AS curve is given by:

$$\pi_t = c_\pi + \mu_{\pi,1}E_t\pi_{t+1} + \mu_{\pi,2}\pi_{t-1} + \kappa_\pi y_t + \sigma_\pi \varepsilon_{\pi,t}$$  \hspace{1cm} (1)

$$c_\pi = (1 - \frac{\delta_T}{1 + \psi \phi})\pi^*$$  \hspace{1cm} (2)

$$\mu_{\pi,1} = \frac{\psi}{1 + \psi \phi}, \quad \mu_{\pi,2} = \frac{\delta_T}{1 + \psi \phi}$$

where $\psi$ represents the discount factor, and $\kappa_\pi$ measures the sensitivity of inflation to the output gap. Given the assumed proportionality between marginal costs and output gap, $\kappa_\pi$ is a rescaled parameter of the sensitivity of inflation to the real marginal cost. Endogenous inflation persistence, $\mu_{\pi,2} > 0$, arises as a consequence of the assumption that non-optimizing agents use past inflation in their indexation scheme. Finally, we impose long-run neutrality of output with respect to inflation. Given the set-up of the model, this amounts to setting the discount factor ($\psi$) to one. Long-run neutrality is characterized by inflation parameters in the AS equation adding up to one, implying that $\mu_{\pi,1} = (1 - \mu_{\pi,2})$.

The IS curve is recovered from the Euler equation on private consumption. Following the recent strand of literature incorporating external habit formation in the utility function (e.g. Cho and Moreno 2006), and imposing the standard market clearing condition, we obtain the following IS equation:

$$y_t = \mu_y E_t y_{t+1} + (1 - \mu_y) y_{t-1} + \phi(i_t - E_t \pi_{t+1} - r) + \sigma_y \varepsilon_{y,t}$$  \hspace{1cm} (3)

where the parameters $\mu_y$ and $\phi$ are functions of the utility parameters related to the agent’s level of relative risk aversion, $\sigma$, and (external) habit formation, $h$: 7

$$\mu_y = \frac{\sigma}{\sigma + h(\sigma - 1)}, \quad \phi = -\frac{1}{\sigma + h(\sigma - 1)}.$$  \hspace{1cm} (4)

Habit formation is introduced as a means to generate additional output gap persistence. Without habits, i.e. $h = 0$, the purely forward-looking IS curve is recovered. The demand shock $\varepsilon_{y,t}$ refers to (independent) shocks in preferences. 8 Equation (3) establishes the interpretation of $r$ as an output-neutral real interest rate. Other things equal, ex ante real interest rate levels $(i_t - E_t \pi_{t+1})$ above $r$ reduce output (and inflation), while for ex ante real interest rates below $r$ output (and inflation) increases. Although we could allow for time variation in this output-neutral real interest rate, we

7 We assume the following utility function:

$$U(C_t) = (1 - \sigma)^{-1}G_t \left( \frac{C_t}{H_t} \right)^{1-\sigma}$$

with $G_t$ an independent stochastic preference factor and an external habit level, $H_t$, specified as $H_t = C_t^h$. Note that in order to have a well-defined steady state, the habit persistence needs to be restricted, $0 \leq h \leq 1$, as explained in Fuhrer (2000).

8 Note that only by linearly detrending output we obtain a one-to-one relation between the shock in the IS equation and preference (demand) shocks. In general, the interpretation of $\varepsilon_{y,t}$ as a demand shock is at least partially flawed, given the fact that it might also contain shocks to permanent output.
restrain from doing so in order to avoid additional complexities in the estimation arising from the fact that this variable is unobservable.

We close the model by specifying a monetary policy in terms of a Taylor rule. Following Clarida, Galí and Gertler (1999), we use a policy rule accounting for both interest rate smoothing and idiosyncratic policy shocks, \( \varepsilon_{i,t} \). The monetary policy rate equation is given by:

\[
i_t = (1 - \gamma_{i-1}) i_t^P + \gamma_{i-1} i_{t-1} + \sigma_i \varepsilon_{i,t}. \tag{5}\]

We model the central bank’s targeted interest rate, \( i_t^T \), by means of a Taylor rule in the output gap, \( y_t \), and inflation gap, \( \pi_t - \pi^* \):

\[
i_t^T = r + E_t \pi_{t+1} + \gamma_\pi (\pi_t - \pi^*) + \gamma_y y_t \tag{6}\]

where \( \pi^* \) denotes the inflation target of the central bank. This policy rule differs from the standard formulation of Taylor rules as we assign a weight of one to the expected inflation term. By imposing this condition, we model explicitly the idea that the central bank is actually targeting an ex ante real interest rate in function of the macroeconomic state, i.e. \( \pi_t - \pi^* + \gamma_y y_t \).

The model can be summarized in a standard matrix notation by defining the state space by a vector of macroeconomic variables, \( X_t = [\pi_t, y_t, i_t]^T \), and a vector of structural shocks, \( \varepsilon_t = [\varepsilon_{\pi,t}, \varepsilon_{y,t}, \varepsilon_{i,t}]^T \). Using a vector \( C \) and matrices \( A, B, D \) and \( S \) of appropriate dimensions, we write the structural equations as:

\[
AX_t = C + BE_t X_{t+1} + DX_{t-1} + S \varepsilon_t. \tag{7}\]

2.2 Perceived law of motion

The structural model (eq. (7)) is solved under two sets of expectations operators. First, we solve the model under the assumption of rational expectations. The rational expectations solution builds on perfect information of agents with respect to the structural parameters and results in a time-invariant structural VAR representation for the perceived law of motion. Next, we relax the perfect information assumption and extend the perceived law of motion by introducing uncertainty with respect to the endpoints of the macroeconomic variables. This alternative implies a perceived law of motion described in terms of a VECM. Both versions can be seen as special cases of a generic model for expectations, composed of \( i \) a set of prior beliefs of the agents, and \( ii \) an optimal learning rule. In this section, we first describe the priors of the generic model, subsequently we solve for the optimal learning rule under both rational expectations and the extended set of beliefs. Finally, we discuss the implications of the perceived law of motion for the term structure of interest rates and the term structure of inflation expectations.

2.2.1 Priors and learning. The beliefs of agents are summarized in terms of a generic model for the macroeconomic dynamics. More specifically, denoting the perceived stochastic trends by \( \zeta_t^P \) and observable macroeconomic variables by \( X_t \) the priors take the form:

\[
X_t = \xi_t^P + \Phi^P (X_{t-1} - \xi_t^P) + \Gamma^P \varepsilon_t \\
\xi_t^P = V^P \zeta_t^P \\
\zeta_t^P = \zeta_{t-1}^P + \Sigma^P v_{\zeta,t}. \tag{8}\]
Macroeconomic dynamics can be decomposed into a transitory and permanent component. The permanent component of the dynamics is generated by a set of stochastic trends $\zeta^P$. The stochastic trends are generated by a set of permanent shocks, $v_{\zeta,t}$ with standard deviation:

$$
\Sigma_\zeta = \begin{bmatrix}
\sigma_{\zeta,\pi} & 0 & 0 \\
0 & \sigma_{\zeta,y} & 0 \\
0 & 0 & \sigma_{\zeta,r}
\end{bmatrix}.
$$ (9)

The stochastic trends determine a set of perceived stochastic endpoints $\xi^P_t$, $\xi^P_t = V^P \zeta^P_t$, identifying the long-run expectations of the macroeconomic variables, $X_t$ (see Kozicki and Tinsley 2001):

$$
\xi^P_t = \lim_{s \to \infty} E_t^{P} X_{t+s}.
$$ (10)

The perceived transitory dynamics, i.e. dynamics relative to the long-run expectations, are modeled in terms of a standard VAR(1) model. More in particular, transitory dynamics are described by the matrix $\Phi^P$, modeling the inertia and interaction in the transitory dynamics, and $\Sigma^P$ the impact matrix of the transitory shocks, $\varepsilon_t$.

Finally, the priors can be restated in extended state space $\tilde{X}_t = [X^P_t, \xi^P_t]$ by defining $\eta_t = [\varepsilon^P_t, v_{\zeta,t}]'$ as:

$$
\tilde{X}_t = \Phi^P \tilde{X}_{t-1} + \Sigma^P \eta_t
$$ (11)
or equivalently:

$$
\begin{bmatrix}
X^P_t \\
\xi^P_t
\end{bmatrix} = \begin{bmatrix}
\Phi^P & (I - \Phi^P) \\
0 & I
\end{bmatrix}
\begin{bmatrix}
X^P_{t-1} \\
\xi^P_{t-1}
\end{bmatrix} + \begin{bmatrix}
\Sigma^P & (I - \Phi^P)V \Sigma_\zeta \\
0 & V \Sigma_\zeta
\end{bmatrix}
\begin{bmatrix}
\varepsilon^P_t \\
v_{\zeta,t}
\end{bmatrix}
$$ (12)

In general, we assume that the stochastic endpoints are not observed. Agents, therefore, face an inference problem for the stochastic endpoints $\xi^P_t$, which is solved by means of a mean squared error (MSE) optimal Kalman filter learning rule. Denoting the inferred values for the stochastic endpoints by $\xi^P_{t|t}$, the learning algorithm becomes:

$$
\xi^P_{t|t} = \xi^P_{t-1|t-1} + K(X_t - E_{t-1|t-1}^P X_t)
$$ (13)

where $K$ is obtained as the steady-state solution to the Kalman filtering equations:

$$
K_t = P_{t|t-1}(I - \Phi^P)'F_t^{-1}
$$

$$
F_t = (I - \Phi^P)P_{t|t-1}(I - \Phi^P)' + \Sigma^P \Sigma^P
$$

$$
P_{t+1|t} = P_{t|t-1} - P_{t|t-1}(I - \Phi^P)'F_t^{-1}(I - \Phi^P)P_{t|t-1} + \Sigma_\zeta \Sigma_\zeta'.
$$ (14)

This perceived law of motion embeds various forms of expectational assumptions. We distinguish between rational expectations and models incorporating imperfect information and/or credibility with respect to the inflation target and output-neutral real interest rate.

### 2.2.2 Rational expectations.

Rational expectations are obtained as a specific case of the perceived law of motion. More in particular, rational expectations generated by the above structural equations are recovered by two sets of informational assumptions. First, agents believe in deterministic endpoints. Within the context of the above structural model endpoints are deterministic, i.e. $\Sigma_\zeta = 0$, and are identified by solving the structural model for the steady state. Under the restriction that in the long run no trade-off exists between the output gap and the
monetary policy, i.e. \( \mu_{\pi,1} = (1 - \mu_{\pi,2}) \), the steady state of the model is determined by the level of the inflation target, \( \pi^* \), the steady state of the output gap, \( y^* \) (fixed to zero), and the output-neutral real interest rate level, \( r = r^* \):

\[
\xi_t^P = \lim_{s\to\infty} E_t \begin{bmatrix} \pi_{t+s} \\ y_{t+s} \\ i_{t+s} \end{bmatrix} = V \begin{bmatrix} \pi^* \\ y^* \\ r^* \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \pi^* \\ y^* \\ r^* \end{bmatrix},
\]

(15)

where \( E_t \) denotes the rational expectations operator. The mapping \( V \) is determined by the structural equations. Under rational expectations, the inflation target determines the long-run inflation expectations. The long-run expectations for the output gap are fixed at \( y^* = 0 \) and the long-run expectations for the nominal interest rate are determined by the Fisher hypothesis, linking the endpoint of the interest rate to the sum of the real interest rate and inflation expectations. Next, rational expectations imply that agents know the structural parameters, such that transitory dynamics, i.e. the matrices \( \Phi^P = \Phi^{re} \) and \( \Sigma^P = \Sigma^{re} \) are determined by the standard rational expectations conditions:

\[
\Phi^{re} = (A - B\Phi^{re})^{-1} D
\]

\[
\Sigma^{re} = (A - B\Phi^{re})^{-1} S.
\]

(16)

The perceived law of motion under rational expectations can be restated in an extended state space as:\(^9\)

\[
\begin{bmatrix} X_t^P \\ \xi_t^P \end{bmatrix} = \begin{bmatrix} \Phi^{re} & (I - \Phi^{re}) \\ 0 & I \end{bmatrix} \begin{bmatrix} X_{t-1}^P \\ \xi_{t-1}^P \end{bmatrix} + \begin{bmatrix} \Sigma^{re} \\ 0 \end{bmatrix} \varepsilon_t.
\]

(17)

with initial condition \( \xi_0^P \):

\[
\xi_0^P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \pi^* \\ y^* \\ r^* \end{bmatrix}.
\]

(18)

Finally, note that under rational expectations with deterministic endpoints and perfect information, agents do not face an inference problem. This perfect information assumption (i.e. \( \Sigma_\xi = 0 \)) generates a Kalman gain \( K = 0 \) (see eq.(14)). Rational expectations are, therefore, a limiting case of the learning model.

### 2.2.3 Stochastic endpoints and learning

Next to rational expectations, we introduce an alternative set of priors implying stochastic endpoints for the macroeconomic variables. Within the context of the above structural model, stochastic endpoints \( \xi_t^P \) arise as the consequence of perceived underlying stochastic trends in the economy, \( \xi_t^P = [\pi_{t+P}, y_{t+P}, r_{t+P}] \), representing the vector containing the perceived inflation target, \( \pi_t^P \), the perceived long-run output gap, \( y_t^P \) (fixed to zero), and the perceived long-run output-neutral real interest rate, \( r_t^P \). The size of the perceived time variation in the stochastic trends is measured by \( \Sigma_\xi \). Denoting the expectations operator under this set of priors by \( E_t^P \) it can be verified that:

\[
\lim_{s\to\infty} E_t^P X_{t+s} = \xi_t^P = V \xi_t^P,
\]

or equivalently

\[
\lim_{s\to\infty} E_t^P \begin{bmatrix} \pi_{t+s} \\ y_{t+s} \\ i_{t+s} \end{bmatrix} = V \begin{bmatrix} \pi_{t+P} \\ y_{t+P} \\ r_{t+P} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \pi_{t+P} \\ y_{t+P} \\ r_{t+P} \end{bmatrix}.
\]

(20)

\(^9\)Note that the solution can also be written more concisely in structural VAR form as:

\[
X_t = C^{re} + \Phi^{re} X_{t-1} + \Sigma^{re} \varepsilon_t
\]

with \( C^{re} = (I - \Phi^{re})\xi_t^{re} \).
The priors about the transitory dynamics, i.e. the dynamics relative to the stochastic endpoints, are assumed to coincide with the ones implied by the rational expectations model. This implies that the matrices \( \Phi^P \) and \( \Sigma^P \) are identical to their rational expectations equivalents: \( \Phi^P = \Phi^{re} \) and \( \Sigma^P = \Sigma^{re} \). By equating the perceived transitory dynamics to those implied by the rational expectations model, we obtain a perceived law of motion that differs from the rational expectations solution only due to the introduction of stochastic endpoints. The final PLM can be written in extended state space as:

\[
\begin{bmatrix} X_t \\ \xi_t^P \end{bmatrix} = \begin{bmatrix} \Phi^P & (I - \Phi^P) \\ 0 & I \end{bmatrix} \begin{bmatrix} X_{t-1} \\ \xi_{t-1}^P \end{bmatrix} + \begin{bmatrix} \Sigma^P & (I - \Phi^P)V \Sigma \zeta \\ 0 & V \Sigma \zeta \end{bmatrix} \begin{bmatrix} \varepsilon_t \\ \nu_{\zeta,t} \end{bmatrix}. \tag{21}
\]

Finally, under this prior, agents face an inference problem and hence resort to learning the stochastic endpoints by means of a Kalman filter. The optimal learning rule for this prior is within the class of stochastic gradient rules with the gain defined by the Kalman filter. This gain depends on the specificities of the prior, i.e. the specific values for \( \Sigma^P \) and \( \Phi^P \). As in Orphanides and Williams (2005a), we assume that agents substitute the unknown stochastic endpoints by the ones inferred by the learning rule.

### 2.2.4 The term structure of interest rates.

Standard no-arbitrage conditions are used to generate bond prices consistent with the perceived law of motion. Imposing no-arbitrage under the PLM reflects the view that bond prices are set by the private sector and should, therefore, be consistent with the perceived dynamics and information set of these agents. Within the context of default-free, zero-coupon bonds, no-arbitrage implies a pricing equation of the form:

\[
P_t(\tau) = E_t^P(M_{t+1}P_{t+1}(\tau - 1)) \tag{22}
\]

where \( E^P \) denotes the subjective expectations operator generated by the PLM (see eq. (11)), \( P(\tau) \) denotes the price of a default-free, zero-coupon bond with maturity \( \tau \), and \( M_t \) denotes the pricing kernel consistent with the PLM. We follow Bekaert et al. (2006) in using the utility function implied by the macroeconomic framework to identify the prices of risk. While this approach has the advantage of guaranteeing consistency of the pricing kernel, it comes at the cost of loss of flexibility in modeling the prices of risk.\(^{10}\) The (log) pricing kernel consistent with the PLM is the homoskedastic (log) pricing kernel:

\[
m_{t+1} = -i_t - \frac{1}{2} \sigma_m^2 - \Lambda \eta_{t+1} \tag{23}
\]

where the prices of risk, \( \Lambda \), are determined by the structural parameters

\[
\Lambda = \sigma_{e_y}^{\Sigma^P} + e_x \tilde{\Sigma}^P - \sigma_y e_y \tag{24}
\]

where \( e_x \) denotes a vector selecting the elements of the \( x \)-equation, i.e. \( e_y \) selects the row of \( \tilde{\Sigma}^P \) related to the \( y \)-equation. No-arbitrage restrictions imposed on conditional Gaussian and linear state space dynamics generate exponentially affine bond pricing models (see, for instance, Ang and Piazzesi 2003):

\[
P(\tau) = \exp(a(\tau) + b(\tau)\tilde{X}_{t|t}) \tag{25}
\]

\(^{10}\)The standard approach in modeling the term structure is to assume a essentially affine term structure representation. As shown by Duffee(2002), such representations do not restrict the prices of risk to be constant.
where $\tilde{X}_{it|t}$ denotes the inferred state vector, obtained by replacing $\xi^P_t$ by its inferred value $\hat{\xi}^P_t$, $\tilde{X}_{it|t} = [X^t_t, \xi^P_t]^t$. The factor loadings $a(\tau)$ and $b(\tau)$ can be obtained by solving difference equations representing the set of non-linear restrictions imposed by the no-arbitrage conditions:

$$a(\tau) = -\delta_0 + a(\tau - 1) - (b(\tau - 1))\Sigma^P \Lambda' + \frac{1}{2} b(\tau - 1)\Sigma^P \Sigma^P b(\tau - 1)'$$ \hspace{1cm} (26)

$$b(\tau) = b(\tau - 1)\Phi^P - \delta^t_1$$

with $\delta_0 = 0$, and $\delta_1$ implicitly defined by the identity $i_s = \delta^t_1 \tilde{X}_{it|t}$. The system has a particular solution given the initial conditions $a(0) = 0$ and $b(0) = 0$.

Exponentially affine bond price models lead to affine yield curve models. Defining the yield of a bond with maturity $\tau_1$ by $y(\tau_1) = -ln(P_t(\tau_1))/\tau_1$ and the vector of yields spanning the term structure by $Y_t = [y_t(\tau_1), ..., y_t(\tau_\pi)]'$, the term structure can be written as an affine function of the extended state space variables:

$$Y_t = A_y + B_y \tilde{X}_{it|t} + v_{yt,t}$$ \hspace{1cm} (27)

where $A_y$ and $B_y$ denote matrices containing the maturity-specific factor loadings for the yield curve ($A_y = [-a(\tau_1)/\tau_1, ..., -a(\tau_\pi)/\tau_\pi]$ and $B_y = [-b(\tau_1)'/\tau_1, ..., -b(\tau_\pi)'/\tau_\pi]$), and $v_{yt,t}$ contains maturity-specific measurement errors.

2.2.5 The term structure of inflation expectations. The representation of the term structure of inflation expectations is obtained by iterating the PLM (eq. (11)) forward. It is straightforward to show that the linearity of the PLM generates an affine representation for the term structure of inflation expectations in the extended state space, $\tilde{X}_{it|t}$. The term structure of average inflation expectations is described by

$$E^P_t \pi(\tau) = \frac{1}{\tau} \sum_{i=0}^{\tau-1} E^P_t (\tau_{i+1}) = e_\pi (a_s(\tau) + b_s(\tau) \tilde{X}_{it|t})$$ \hspace{1cm} (28)

where $E^P_t \pi(\tau)$ denotes the time $t$ average inflation expectation over the horizon $\tau$, $e_\pi$ denotes a vector selecting $\pi_t$ out of the vector $\tilde{X}_{it|t}$, and $a_s(\tau)$, $b_s(\tau)$, $a_s(\tau)$ and $b_s(\tau)$ are maturity-dependent functions generated by the system:

$$a_s(\tau) = 0, \hspace{0.5cm} b_s(\tau) = b_s(\tau - 1)\Phi^P$$

$$a_s(\tau) = \frac{1}{\tau} \sum_{i=0}^{\tau-1} a_s(\tau) = 0, \hspace{0.5cm} b_s(\tau) = \frac{1}{\tau} \sum_{i=0}^{\tau-1} b_s(\tau)$$ \hspace{1cm} (29)

solved under the initial conditions $a_s(0) = 0$ and $b_s(0) = I$. Equation (28), applied over varying horizons $\tau$, forms the model-implied term structure of average inflation expectations. The term structure of inflation expectations, unlike the term structure of interest rates, is not observable. We use surveys of average inflation expectations for different maturities as a proxy for the term structure of inflation expectations. We relate these surveys, $s(\tau)$, to the model-implied average inflation expectations by allowing for idiosyncratic measurement errors, $v_{s,t}(\tau)$, in the survey responses:

$$s_t(\tau) = e_\pi a_s(\tau) + e_\pi b_s(\tau) \tilde{X}_{it|t} + v_{s,t}(\tau)$$ \hspace{1cm} (30)

where $s_t(\tau)$ denotes the time $t$ survey response concerning the average inflation expectations over the horizon $\tau$. Finally, denoting the vector containing a set of surveys of inflation
expectations for different horizons by \( S_t = [s_t(\tau_1), ..., s_t(\tau_m)]', and defining \( A_s = 0 \) and \( B_s = [(e_\pi b_s(\tau_1))', ..., (e_\pi b_s(\tau_m))']', equation (30) can be restated as:

\[
S_t = A_s + B_s \tilde{X}_{t|t} + v_{s,t}.
\]  

(31)

### 2.3 Actual law of motion

The actual law of motion (ALM), describing the observed dynamics of macroeconomic variables, is obtained by substituting the subjective expectations (eq. (11)) into the structural equations (eq. (7)). Since the subjective expectations are formed on the basis of the inferred stochastic endpoints, \( \xi_{t|t} \), and on observable macroeconomic data, the relevant space of the ALM coincides with that of the PLM, i.e. \( \tilde{X}_{t|t} \). Due to the simplicity of the learning algorithm, the ALM can be solved in closed form. In Appendix A, we show that the ALM reduces to a standard VAR(I) in the extended state space:

\[
\tilde{X}_{t|t} = \tilde{C}^A + \tilde{\Phi}^A \tilde{X}_{t-1|t-1} + \tilde{\Sigma}^A \tilde{\epsilon}_t
\]  

(32)

with

\[
\tilde{C}^A = \left[ \begin{array}{c}
(A - B(\Phi^P + K_\Phi))^{-1} C \\
K(A - B(\Phi^P + K_\Phi))^{-1} C
\end{array} \right]
\]

\[
\tilde{\Phi}^A = \left[ \begin{array}{c}
\Phi^P \\
0
\end{array} \right] \\
\tilde{\Sigma}^A = \left[ \begin{array}{c}
(A - B(\Phi^P + K_\Phi))^{-1} S \\
K(A - B(\Phi^P + K_\Phi))^{-1} S
\end{array} \right]
\]  

(33)

and \( K_\Phi = (I - \Phi^P)K_\Phi \), \( A, B \) and \( S \) and \( \Phi^P \) determined by the parameters of the structural equations, and \( K \) the constant gain matrix implied by the agents’ priors. The closed form solution can be used to highlight some of the properties of the ALM. First, subjective beliefs about the stochastic endpoints, \( \xi_{t|t} \), only affect the actual macroeconomic dynamics if an expectation channel exists, i.e. \( B \neq 0 \) in equation (7). One aspect in which macroeconomic dynamics may be affected by subjective beliefs concerns the modeling of persistence. Under rational expectations, persistence is driven by inflation indexation, habit persistence, and interest rate smoothing affecting the roots of the \( \Phi^P = \Phi^P \) matrix. Learning introduces an additional source of persistence in the form of the persistence in the subjective expectations, \( \xi_{t|t}^P \). Persistence in the beliefs follows itself from the inertia in the learning rule, i.e. the updating procedure. Milani (2005) shows in a different context that persistence due to learning is important and (partly) takes over the role of inflation indexation and habit formation. In the empirical section, we find similar results, especially for inflation persistence and interest rate smoothing.

Second, the rational expectations model is nested within the learning framework. By imposing the priors consistent with rational expectations, i.e. \( \Sigma_\zeta = 0 \) (implying \( K_\Phi = 0 \)) and \( \xi_{t|t}^P = V[\pi^*, 0, r]' \), it can be verified that the ALM simplifies to the rational expectations reduced form, equation (17). Third, the nonstationarity of the PLM does not necessarily carry over to the ALM. The eigenvalues of the matrix \( \tilde{\Phi}^A \) depend both on the structural parameters contained in \( A, B, \Phi^P \) and on the learning parameters \( K \).\(^{11}\) Finally, if the ALM is stationary, the unconditional distribution

\(^{11}\)Note that the stationarity of the ALM is inconsistent with the nonstationarity of the PLM under learning. This inconsistency arises from the fact that the ALM is solved under the assumption of a time-invariant inflation target of the central bank. In the empirical section, we allow for time variation in the
of the extended state space vector $\tilde{X}_{t|t}$ is identified. Conditional on the maintained assumption of normality of the structural shocks, $\varepsilon_t$, this distribution is given by:

$$\tilde{X}_{t|t} \sim N(E\tilde{X}_{t|t}, \Omega_{\tilde{X}})$$

(34)

with:

$$E\tilde{X}_{t|t} = \left[ \begin{array}{c}
(I - \Phi^{re})^{-1}C^{re} \\
(I - \Phi^{re})^{-1}C^{re}
\end{array} \right]$$

$$\text{vec}(\Omega_{\tilde{X}}) = (I - \tilde{\Phi}^A \otimes \tilde{\Phi}^A)^{-1} \text{vec}(\tilde{\Sigma}^A \tilde{\Sigma}^A).$$

Equation (34) represents the unconditional distribution for the extended state under learning. This distribution is characterized by two properties. First, as far as unconditional means are concerned, the ALM and the rational expectations model are observationally equivalent. The unconditional mean of the rational expectations model, i.e. $(I - \Phi^{re})^{-1}C^{re}$, coincides with the unconditional mean under the ALM for both the observable macroeconomic variables (inflation, output gap, and policy rate) and the perceived long-run expectations of the agents. The rational expectations model thus serves as a benchmark in mean for the model under learning. Second, in line with the literature on constant gain learning (e.g. Evans and Honkapohja 2001), the unconditional variance of the stochastic endpoints, $\xi^P_{t|t}$, is in general positive, implying non-convergence of the stochastic endpoints to the true values implied by the rational expectations equilibrium, $[\pi^*, 0, r + \pi^*]'.$

3 Estimation methodology

The actual law of motion for both macroeconomic variables and the inferred stochastic endpoints is used to estimate both the structural and the learning parameters. In order to identify the subjective beliefs, we use information variables directly related to the PLM, i.e. the term structure of interest rates and inflation expectations. In Section 3.1, we discuss the details of the estimation procedure. Subsequently, in Section 3.2, we explain the different versions of the model that are estimated.

3.1 Maximum likelihood estimation

The model is estimated by means of loglikelihood in the extended state space with the ALM dynamics serving as the transition equation:

$$\tilde{X}_{t|t} = \tilde{C}^A + \tilde{\Phi}^A \tilde{X}_{t-1|t-1} + \tilde{\Sigma}^A \varepsilon_t$$

(35)

and a measurement equation, relating the extended state to observable economic variables. The observable variables included in the measurement equation consist of macroeconomic variables, $X_t$ (inflation, output gap, and policy rate), a sample of yields spanning the term structure of interest rates, $Y_t$ (1, 2, 3, 4, 5 and 10 year yields), and a sample of the term structure of inflation expectations, proxied by survey data on inflation expectations, $S_t$ (1 and 10 year average inflation expectations).12

---

12Survey expectations are increasingly used in the empirical literature. Roberts (1997) shows that models including survey expectations can account for some of the (unexplained) inflation inertia. Survey expectations are also starting to be used in the bond pricing literature. Kim and Orphanides (2005) use survey expectations on short-term interest rate movements as an additional input in a otherwise standard Vasicek model. Also, Chunn (2005) uses several survey expectations as additional inputs in a two-factor term structure model. Finally, Bekaert et al. (2006) show the empirical relevance of surveys on inflation by showing that surveys help to forecast inflation better than any rational expectations model.
The observable variables are collected in the vector \( Z_t = [X_t', Y_t', S_t']' \). Using the affine representation of each of these variables in the extended state space, as discussed in Section 2.2, the measurement equation becomes:

\[
Z_t = A_m + B_m \tilde{X}_{t|t} + v_{z,t}
\]  

(36)

where \( v_{z,t} \) denotes idiosyncratic measurement errors with variance-covariance matrix \( \Psi \), and \( A_m \) and \( B_m \) represent the derived affine representations of the respective subsets of observable variables \( X_t, Y_t \) and \( S_t \) (\( B_X \) is defined as: \( X_t = B_X \tilde{X}_{t|t} \), i.e. \( B_X = [I_{3 \times 3}, 0_{3 \times 3}] \), \( A_y, B_y \) and \( B_s \) are defined in equations (27) and (31), respectively):

\[
A_m = \begin{bmatrix} 0 \\ A_y \\ A_s \end{bmatrix}, B_m = \begin{bmatrix} B_X \\ B_y \\ B_s \end{bmatrix} \quad \text{and} \quad \Psi = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Psi_y & 0 \\ 0 & 0 & \Psi_s \end{bmatrix}.
\]

Prediction errors, \( Z_t - E_{t-1}^A Z_t \), and their corresponding loglikelihood value \( l(Z_t - E_{t-1}^A Z_t; \theta) \), where \( E_{t-1}^A \) denotes the expectations operator based on the ALM, are functions of both the structural macroeconomic shocks and the measurement errors:

\[
l(Z_t - E_{t-1}^A Z_t; \theta) = -\frac{1}{2} \Omega_Z - \frac{1}{2} (Z_t - E_{t-1}^A Z_t) \Omega_Z^{-1} (Z_t - E_{t-1}^A Z_t)
\]  

(37)

\[
\Omega_Z = B_m \sqrt{A} \Sigma A^\top B_m^\top + \Psi.
\]

One contribution of this paper is that the deep parameters of the structural equations and the parameters of the learning procedure are estimated jointly based on a wide variety of information variables, i.e. macroeconomic variables, term structure variables, and surveys of inflation expectations.\(^{13}\) The parameters to be estimated are collected in the parameter vector \( \theta \), containing the deep parameters of the structural equations \( (\delta_x, \kappa_x, \sigma, h, r, \pi^*, \gamma_{x\gamma}, \gamma_{i-1}, \sigma_x, \sigma_y, \sigma_i, \sigma_{z,\pi}, \sigma_{z,r}, \sigma_{z,\gamma}, \sigma_{z,\gamma}, \sigma_{z,\gamma}, \sigma_{z,\gamma}) \), the learning parameters (priors on the volatility of the stochastic trends \( \sigma_{z,\pi}, \sigma_{z,r}, \) and initial values \( \zeta_{0|0} \)), and the variances of the measurement errors \( (\text{diag}(\Psi)) \):

\[
\theta = \left\{ \delta_x, \kappa_x, \sigma, h, r, \pi^*, \gamma_{x\gamma}, \gamma_{i-1}, \sigma_x, \sigma_y, \sigma_i, \sigma_{z,\pi}, \sigma_{z,r}, \sigma_{z,\gamma}, \sigma_{z,\gamma}, \sigma_{z,\gamma}, \sigma_{z,\gamma}, \sigma_{z,\gamma}, \sigma_{z,\gamma}, \sigma_{z,\gamma}, \zeta_{0|0}, \text{diag}(\Psi) \right\}.
\]  

(38)

Not all deep parameters and learning parameters are estimated. We follow Hördahl et al. (2006) and Bekoert et al. (2006) by fixing the discount factor to one, \( \psi = 1 \). Also, throughout the estimation the prior on the uncertainty of the long-run value for the output gap is restricted to zero, \( \sigma_{z,y} = 0 \). This restriction guarantees that the long-run expected output gap is fixed to zero under the PLM. Furthermore, we impose the theoretical constraints \( \sigma_{z,\pi}, \sigma_{z,r} \geq 0 \) and \( 0 \leq h \leq 1 \). Finally, parameter estimates are constrained to satisfy two conditions. First, parameter estimates must be consistent with the existence of a unique rational expectations solution. Second, under learning, parameter estimates should imply eigenvalues of \( \Phi^A \) strictly smaller than one in absolute value in order to guarantee stability of the ALM. The model is estimated using a Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm.

\(^{13}\)Other research estimating learning parameters include Orphanides and Williams (2005a) and Milani (2005). Orphanides and Williams (2005a) estimate the constant gain by minimizing the distance between the model-implied inflation expectations and those reported in the survey of professional forecasters. Milani (2005) estimates jointly, using Bayesian methods, the constant gain and the deep parameters of a structural macroeconomic model. We complement their analyses by including more information in the measurement equation, notably term structure of interest rates.
3.2 Estimated versions of the model

We estimate a total of eight models. Model versions differ depending on (i) the type of information included in the measurement equation, (ii) the assumptions concerning the learning procedure, (iii) the time variation in the inflation target, and (iv) the prices of risk. Four versions are based on the baseline model presented in the previous section, two versions are extensions allowing for heterogeneity in the monetary policy, and the final two versions allow for more flexibility in the prices of risk.

Regarding the information included in the measurement equation, we distinguish between the Macro and the general versions of the model. In the Macro version, we restrict the measurement equation to incorporate only macroeconomic information, while in the general version we include all available information. The Macro version of the model is motivated by the concern that including term structure and survey information in the measurement equation may bias the estimates of the deep and learning parameters in order to fit the term structure and the survey expectations. To avoid this problem, a two-step procedure is employed. In the first step, the deep and learning parameters are estimated while restricting the measurement equation to contain only macroeconomic variables. In the second step, we fix the parameter estimates for the deep and learning parameters obtained in the first step, and optimize the likelihood based on the full measurement equation over the remaining parameters, \( \text{diag}(\Psi) \). In the general versions of the model, the estimation of all parameters is performed in one step on the basis of the most general measurement equation.

We estimate both rational expectations and learning versions of the model. The learning versions of the model include four additional parameters describing the priors of the agents, \( \sigma_{\zeta, \pi}, \sigma_{\zeta, r} \) and the starting values for the stochastic trends, \( \zeta_{0|0} \). The distinction between rational expectations and learning models identifies the contribution of learning to the overall fit of the respective series. The four baseline models can be summarized as follows:

- **Rational Expectations Macro**: the rational expectations version is estimated using a two-step procedure ensuring that the deep parameters are based only on macroeconomic information.
- **Rational Expectations I**: the rational expectations version is estimated using a one-step procedure based on the general measurement equation.
- **Learning Macro**: the learning version is estimated using a two-step procedure ensuring that the deep parameters are based only on macroeconomic information.
- **Learning I**: the learning version is estimated using a one-step procedure based on the general measurement equation.

In addition to the four baseline models, we estimate two extensions to allow for heterogeneity in the monetary policy rule and in the agents’ priors. The heterogeneity is modeled by means of chairman-specific policy rules and priors.\(^{14}\) Specifically, the time-invariant policy rule parameters \( \pi^* \),

\(^{14}\)This procedure differs from other research that allows for time variation in the inflation target. For instance, Dewachter and Lyrio (2006), Kozicki and Tinsley (2005), and Hördahl et al. (2006) allow for variation in the inflation target of the central bank by modeling the inflation target as an inert autoregressive process. This approach results in quite variable inflation target dynamics. In contrast, this paper allows for discrete jumps in the inflation target at pre-specified dates. Beyond these dates, the inflation target is constant.
\(\gamma_{\pi}, \gamma_{\y},\) and \(\gamma_{i-1}\) of the baseline models are replaced by chairman-specific parameters \(\pi_{j}, \gamma_{j}, \gamma_{y,j}\) \text{ and } \gamma_{i-1,j}, \text{ where } j \text{ denotes the presiding chairman.}^{15} \text{ The heterogeneity in priors is modeled analogously by replacing the learning parameters } \sigma_{\zeta,\pi} \text{ and } \sigma_{\zeta,r} \text{ by their chairman-specific equivalents, } \sigma_{\zeta,\pi,j} \text{ and } \sigma_{\zeta,r,j}. \text{ We estimate both the rational expectations version of this model, labeled Rational Expectations II, and the learning version of the model, labeled Learning II. The model versions Rational Expectations I and Learning I, implying time-invariant policy rules and beliefs, are nested in the respective extensions and hence identify the contribution of allowing for policy heterogeneity in the overall fit.}^{16} \text{ Finally, the last two models, Rational Expectations III and Learning III extend the models Rational Expectations II and Learning II by allowing for time variation in the prices of risk. In these versions, we disregard consistency of the pricing kernel with the IS curve and posit an affine function for the prices of risk: } \Lambda_{t} = \Lambda_{0} + \Lambda_{1}\tilde{X}_{t|t}.^{17}

4 Estimation results

4.1 Data

We estimate the proposed models using quarterly data for the USA. The data covers the period from 1963:Q4 until 2003:Q4 (161 quarterly observations). The data set contains three series of macroeconomic observations: quarter-by-quarter inflation (based on the GDP deflator and collected from the National income and Product Accounts), the output gap (constructed as the log of GDP minus the log of the natural output level, based on Congressional Budget Office data), and the Federal funds rate, representing the policy rate. Next to the macroeconomic variables, the data set includes six yields with maturities of 1, 2, 3, 4, 5 and 10 years. The data for yields up to five years are from the CRSP database.\(^{18} \text{ The ten-year yields were obtained from the Federal Reserve. Finally, we also use survey data on short- and long-run inflation expectations. More specifically, we include the one- and ten-year average inflation forecast, as reported by the Federal Reserve Bank of Philadelphia in the Survey of Professional Forecasters.} \n
Table 1 presents descriptive statistics on the data set described above, which are depicted in Figure 1. These statistics point to the usual observations: the average term structure is upward sloping; the volatility of yields decreases with maturity; normality is rejected for all series (based on JB statistics); and all variables display significant inertia, with a first-order autocorrelation coefficient typically higher than 0.90. Inflation displays a somewhat lower inertia, i.e., an autocorrelation coefficient of 0.76.

Insert Table 1 and Figure 1

Table 1 also presents the correlation structure of the data. Three data features can be highlighted. First, the yields are extremely correlated across the maturity spectrum. This points

---

\(^{15}\)The chairmen included in the analysis are Martin (1951-1970), Burns (1970-1978), Miller (1978-1979), Volcker (1979-1987) and Greenspan (1987-2006). We divide the Volcker period in two sub-periods in order to account for the well-documented change in monetary policy that took place during this term, i.e., the change from monetary targeting to a more conventional monetary policy. The first Volcker period ends in 1982Q3.

\(^{16}\)For an analysis of regime changes on monetary policy, see Schorfheide (2005) or Sims and Zha (2004). Both papers make use of Markov switching techniques identifying the regime breaks endogenously. We, in contrast, fix the dates of the breaks to the moments of a change in the Fed chairman.

\(^{17}\)Note that allowing for time-varying prices of risk adds 42 parameters to be estimated. In order to keep the estimation tractable, we restrict \(\Lambda_{1}\) to be diagonal.

\(^{18}\)We thank Geert Bekaert, Seonghoon Cho and Antonio Moreno for sharing the data set.
to the well-known fact that a limited number of factors account for the comovement of the yields. Second, there is a strong correlation between the term structure and the macroeconomic variables, with significant positive correlations between inflation and the term structure and significant negative correlations between the term structure and the output gap. These correlation patterns are an indication of common factors driving macroeconomic and yield curve dynamics. Finally, we observe a substantial and positive correlation between the surveys of inflation expectations and both the macroeconomic variables (especially inflation and the Federal funds rate) and the yield curve. Again, this suggests that the factors affecting the yield curve and macroeconomic variables also drive movements in the surveys of inflation expectations.

4.2 Parameter estimates

Tables 2, 3 and 4 report the estimation results for the rational expectations versions of the model. Our estimates for the Macro model (Table 2) are broadly in line with the literature. We observe a mild domination of the forward looking terms for both the AS and IS curves ($\mu_{\pi,1} = 0.524$ and $\mu_y = 0.509$, respectively). The deviation from the purely forward-looking model ($\mu_{\pi,1} = 1$ and $\mu_y = 1$) is explained by the relatively high values for the inflation indexation parameter, $\delta_\pi$, and the habit persistence, $h$, estimated at 0.908 and 1, respectively. Both estimates for the inflation sensitivity to the output gap, $\kappa_\pi$, and the output gap sensitivity to the real interest rate, $\phi$, are small, 0.00055 and −0.019, respectively. Although these values are smaller than the ones typically used in calibration-based studies, they are commonly found in empirical studies using GMM or FIML methods. Our estimates imply an active monetary policy rule. The ex ante real interest rate reacts positively to both inflation and the output gap, $\gamma_\pi = 0.674$ and $\gamma_y = 0.569$. Significant interest rate smoothing is also observed in the policy rule ($\gamma_{i-1} = 0.862$). As often found in the literature, some of the estimated parameters are not statistically significant. Similar results have been reported, for instance, by Cho and Moreno (2006).

Extending the standard macro model by (i) including yield curve and inflation survey data in the measurement equation (Rational Expectations I, II and III), (ii) allowing for chairman-specific monetary policy (Rational Expectations II and III), and (iii) introducing time variation in the prices of risk (Rational Expectations III) affects the parameter estimates significantly. First, the estimated persistence decreases, as shown by the decrease in the indexation parameter $\delta_\pi$, which takes a value of 0.67 and 0.57 in the Rational Expectations II and III models, respectively, and by the decrease in the habit persistence, $h$, for the Rational Expectations I, II and III models to 0.738, 0.721, and 0.750, respectively. As a result of the drop in the indexation and/or the habit persistence, the forward-looking components ($\mu_{\pi,1}$ and $\mu_y$) in the AS and IS equation increase. The estimates of monetary policy rule indicate for all versions of the model that (i) monetary policy is relatively inert, and (ii) the Taylor principle is satisfied since the ex ante real interest rate tends to increase with both the inflation gap and the output gap. Nevertheless, the estimated inflation and output gap responses vary across the alternative versions of the model. Based on the results in Table 3 and 4, we find, as Clarida, Galí and Gertler (2000), a strong increase in the responsiveness to the inflation gap during the Volcker and Greenspan periods.

Insert Tables 2, 3 and 4

Tables 5, 6 and 7 report the estimation results for the versions where learning is introduced. The central parameters in the analysis, distinguishing learning models from rational expectations models,
are the standard deviations of the perceived stochastic trends $\zeta^P_t$, $\sigma_{\zeta,\pi}$ and $\sigma_{\zeta,r}$.\textsuperscript{19} Our estimates for these parameters are statistically significant, indicating a rejection of rational expectations models. This finding holds irrespective of the version of the learning model and indicates the importance of the learning specification in modeling the joint dynamics of the macroeconomic variables, the yield curve and the survey expectations. One interpretation of the parameters $\sigma_{\zeta,\pi}$ and $\sigma_{\zeta,r}$ is in terms of the perceived uncertainty with respect to the endpoints of the macroeconomic state. The estimates of the parameters $\sigma_{\zeta,\pi}$ and $\sigma_{\zeta,r}$ in the Learning II and III versions of the model indicate substantial time variation in the uncertainty with respect to the inflation and real interest rate endpoints. Interestingly, we find uncertainty for both inflation and real interest rate endpoint to be significantly lower during the Greenspan term than under previous chairmen. The introduction of learning dynamics affects significantly the estimates of the deep parameters relative to those obtained for the rational expectations counterparts. First, across learning models, we find that the forward-looking component in the AS equation ($\mu_{\pi,1}$) increases substantially and significantly relative to the rational expectations versions of the model. This increase is explained by the decrease in the inflation indexation parameter.\textsuperscript{20} The interest rate smoothing parameter drops significantly to values on average around 0.8 in the learning cases, which are more in line with Rudebusch (2002).

A second effect of learning is the increase in the inflation sensitivity to the output gap. We estimate $\kappa_{\pi}$ levels of 0.05, 0.012 and 0.009 in the Learning I, II and III models, respectively. Finally, note that one problematic feature of the estimation across learning specifications is the identification of the inflation targets, $\pi^*_j$, and the real interest rate $r$, which present large standard errors. This drop in significance can be attributed to the fact that the stochastic endpoints take over the role of these parameters in the expectation formation process.

Insert Tables 5, 6 and 7

Figures 2 to 4 plot the macro variables and their endpoints for each model. Endpoints, representing long-run (subjective) expectations, are deterministic in the rational expectations models and stochastic in the learning cases. As Figure 2 shows, in the presence of learning long-run inflation expectations are time varying and therefore different from the constant central bank’s inflation target (around 3 to 4 percent per year). These endpoints are also remarkably similar across model specifications. Allowing for chairman-specific policy rules (models II and III) leads to significantly different inflation targets across rational expectations and learning models. In the rational expectations models II and III, the estimated inflation targets show a gradual increase over the seventies until the end of the Volcker experiment, subsequently decreasing over time (around 5% in the second Volcker period and 3.2% in the Greenspan term). This gradual decline in inflation targets seems unrealistic given the strong deflationary policy conducted by Volcker.\textsuperscript{21}

\textsuperscript{19} Note that the parameter $\sigma_{\zeta,y}$ was fixed to zero for consistency with the assumption of long-run neutrality of output (see Section 2).

\textsuperscript{20} The decrease in the inflation indexation as a consequence of the introduction of learning is also found in other studies. For instance, Milani (2005), introducing constant gain learning in a New-Keynesian macroeconomic model, finds an even stronger effect, with the inflation indexation parameter close to zero after the introduction of learning.

\textsuperscript{21} One explanation for the observed time series of inflation targets is that inflation targets adapt so as to fit the surveys of inflation expectations. Since under rational expectations long-run expectations coincide with the inflation targets, inflation targets need to track the survey of inflation expectations. Some evidence in favor of this interpretation can be found in Table 8. Comparing the macro part of the likelihood, one
Under learning, the estimated chairman-specific inflation targets seem more in line with the historical record of US monetary policy. Estimates of time-varying inflation targets, in line with our results, can be found in Kozicki and Tinsley (2005) and Milani (2005).

Figure 3 shows the differences between the long-run real interest rate expectations under learning and the values implied by rational expectations models. This difference is less pronounced than in the inflation case and is also similar across learning models. Figure 4 presents equivalent graphs for the long-run expectations regarding the short-run policy rate. We observe again sizable differences between the implied rational expectations endpoints and the subjective long-run expectations under learning. The variability in the long-run expectation for the nominal interest rate is dominated by variation in the inflation endpoint.

Insert Figures 2, 3, and 4

4.3 Comparing learning and rational expectations models

4.3.1 BIC and likelihood decomposition. We use the Bayesian Information Criterion (BIC) for an overall evaluation of the performance across models. Although this criterion does not constitute a formal statistical test, it takes into account (i) the use of different procedures in the estimation of the models (i.e. Macro and general versions), and (ii) the fact that, although rational expectations and learning models are nested, standard likelihood ratio tests are not appropriate since the parameter restrictions of the rational expectations models are on the boundary of the admissible parameter space, i.e. $\sigma_{\zeta, \pi} = 0$ and $\sigma_{\zeta, r} = 0$. Next to the BIC, we also compare the performance of the different models through a likelihood decomposition, showing the contribution of the macro variables, the yield curve, and the surveys of inflation expectations.

The results are presented in Table 8. According to the BIC, learning models outperform their rational expectations counterparts. More strongly, Learning I models outperform any of the estimated rational expectations models. According to this criterion, Learning III models present the best specification, incorporating learning dynamics, heterogeneity in monetary policy rules and priors, and time variation in the prices of risk. The likelihood decomposition shows that the superior performance of learning models is accounted for in each of its components. There seems to be, however, a trade off in fitting those components. From a macro perspective, the Learning Macro model presents the best performance (12.08 as average loglikelihood). From a yield curve and inflation expectations perspective, the Learning III model provides the best fit. The inclusion of this information in the measurement equation therefore slightly biases the model towards fitting yield curve and survey data at the expense of the macroeconomic part.

Insert Table 8

Observes a drop from the Rational Expectations I to the Rational Expectations II model, indicating that allowing for chairman-specific inflation targets worsened the macroeconomic fit. This drop in likelihood is more than compensated by the increase in likelihood in the term structure of interest rates and survey parts of the likelihood.

The findings of the BIC are confirmed by approximative likelihood ratio tests. We reestimated the learning models fixing the learning parameters to small values, $\sigma_{\zeta, \pi} = \sigma_{\zeta, r} = 0.0001$, and $\zeta_{0|0} = [\pi^*, 0, r]$. Likelihood ratio (LR) tests performed using the latter models as the null hypothesis reject the proxy models at 1% significance levels. Also, note that the Rational Expectations I, II and III and Learning I, II and III are nested. Likelihood ratio tests indicate that both the Learning I and II and the Rational Expectations I and II models are rejected against the alternatives, Rational Expectations III and Learning III.
4.3.2 Prediction errors. Table 9 presents summary statistics for the prediction errors of all variables in the alternative model specifications. In all cases, we find evidence of model misspecification due to the significant means and autocorrelation coefficients of the prediction errors. Therefore, none of the models is accepted as a completely satisfactory representation of the joint dynamics of the macroeconomy, the yield curve, and surveys of inflation expectations. There is, however, a clear distinction between learning and rational expectations models. In most cases, learning models outperform their rational expectations counterpart. Introducing learning typically leads to an increase in the in-sample predictive power of all variables, except inflation, and to a decrease in the standard deviation and the autocorrelation of the prediction errors. The inclusion of chairman-specific policies seems to have a considerable positive effect in the fit of the yield curve and surveys of inflation expectations. For learning models, although the inclusion of time-varying prices of risk decreases the mean of the forecast errors, overall, it does not seem to improve the results in a significant way.

Insert Table 9

4.4 Learning dynamics, inflation expectations and bond markets

Do macroeconomic models including learning fit the term structure of interest rates and inflation expectations? To answer this question, we analyze the fitting errors of the respective models. As shown in Table 10, learning models with chairman-specific policy rules (Learning II and III) explain 95% of the variation in the yield curve and more than 85% of the variation in the surveys of inflation expectations. Furthermore, the mean fitting errors for the yield curve are low, ranging from 6 to 20 basis points for the Learning II model, and from 2 to 8 basis points for the Learning III model. These results are comparable to studies using latent factor models (e.g. de Jong 2000). This can also be seen in Figures 5 and 6, which show the fit for the one- and the ten-year yields across models. The difference in performance across models is especially pronounced for the ten-year yield. The performance across models regarding the fit of survey of ten-year average inflation expectations can be seen in Figure 7. In general terms, Learning II and III models fit both the yield curve and surveys of inflation expectations relatively well.

To identify the contribution of learning in the mentioned performance, we compare the Rational Expectations II and Learning II models (analogous results are obtained for Rational Expectations III and Learning III models). We observe an increase in fit due to learning between 4% (one-year yield) and 14% (ten-year yield). Furthermore, we observe a significant reduction in the remaining autocorrelation in the fitting errors. To identify the contribution of chairman-specific monetary policy rules and priors, we compare the Learning I and II models. Learning II models show an increase in the explained variation in the yield curve between 2 and 4 percent and in the survey of inflation expectations between 1 and 14 percent. We observe also a general decrease in the remaining autocorrelation in the fitting errors.

Insert Table 10 and Figures 5, 6, and 7

Why do learning models outperform their rational expectations counterparts? To answer this question, we analyze the affine term structure representations of rational expectations and learning

---

23The rejection of the overall model is common in the macro-finance literature (e.g. Bekaert et al. 2006 and Cho and Moreno 2006). In the pure finance literature, it has also been shown to be difficult to find affine term structure representations that are not rejected by the data.
models. More specifically, we look at the affine representations for the term structure of interest rates and inflation expectations in a transformed state space, decomposing the observed macroeconomic variables in perceived permanent and temporary components. This decomposition is achieved by the rotation matrix $T$:

$$
T = \begin{bmatrix}
I_3 & -I_3 \\
0 & I_3
\end{bmatrix}
$$

which generates the decomposition:

$$
\tilde{X}_{it} = T \begin{bmatrix}
X_t - \xi^P_{i,t} \\
\xi^P_{i,t}
\end{bmatrix} = T \begin{bmatrix}
X_t \\
\xi^P_{i,t}
\end{bmatrix}.
$$

The affine representation of the term structure of interest rates and inflation expectations can be restated in this state space as:

$$
Y_t = A_y + B_y \tilde{X}_{it} + v_{y,t} = A_y + B_y T^{-1} T \tilde{X}_{it} + v_{y,t} = A_y + B^T \tilde{X}_{it} + v_{y,t}
$$

and

$$
S_t = A_s + B_s \tilde{X}_{it} + v_{s,t} = A_s + B_s T^{-1} T \tilde{X}_{it} + v_{s,t} = A_s + B^T \tilde{X}_{it} + v_{s,t}.
$$

Figure 8 shows the transformed yield curve loadings for each of the models. We identify one slope factor driving the yield spread, represented by the perceived transitory interest rate component, and two curvature factors, i.e. the perceived output gap and the perceived inflation gap. The curvature factors affect primarily but marginally the intermediate maturity yields. We also obtain a level factor, exerting its influence equally over the entire yield curve. This factor is driven only by changes in the perceived stochastic endpoint for the policy rate. While both rational expectations and learning models share a level factor in the transformed state space, the implications of this factor differ across models. Rational expectations models imply a deterministic endpoint for the policy rate, i.e. $\xi_{i,t} = r + \pi^*$. The level factor is therefore constant and cannot explain the time variation in long-maturity yields. Learning models generate endogenous stochastic endpoints for the policy rate, which seem to be sufficiently volatile to account for the time variation in the long end of the yield curve.

5 Conclusions

In this paper we built and estimated a macroeconomic model including learning. Learning was introduced in the model by assuming that agents do not believe in time-invariant inflation targets nor in constant equilibrium real rates. Given these priors, the optimal learning rule was derived in terms of a Kalman gain updating rule. We estimated the model including in the measurement equation,

\footnote{Note that in the versions II and III of both rational expectations and learning models, yield curve and inflation expectations loadings also depend on the policy rule parameters. Given that we identify six policy regimes, we have six sets of loadings. For reasons of brevity, we only present the loadings implied by the Greenspan policy rules.}

\footnote{Note that to the extent that one allows for time-varying inflation targets within the rational expectations framework, one can generate exogenously volatility in the endpoints. This is the approach followed in the standard macro-finance literature. The Rational Expectations II and III panels in Figures 2 and 4 are examples of this approach. The main advantage of learning is that there is no need to refer to exogenous shocks (i.e. in the inflation target) to account for the time variation in the long-end of the yield curve. The stochastic endpoints are generated endogenously in the model.}
next to the standard macroeconomic variables, bond yields and surveys of inflation expectations. The structural and learning parameters were estimated jointly. The findings of the paper can be summarized as follows. First, including learning improves the fit of the model independently of the type of information included in the measurement equation. Although learning models improve on the rational expectations models, they are not fully satisfactory. Autocorrelation in the errors was found to be significant. Finally, we found that introducing learning in a standard New-Keynesian model generated sufficiently volatile stochastic endpoints to fit the variation in long-maturity yields and in surveys of inflation expectations. The learning model, therefore, complements the current macro-finance literature linking macroeconomic and term structure dynamics.
References


Appendix A: ALM dynamics

In this appendix, we derive a closed form solution for the actual law of motion (ALM). The derivation follows the standard approach in the learning literature by substituting subjective expectations, i.e. the PLM, into the structural equations. The structural equations are described in equation (7), which is repeated here as:

$$AX_t = C + BE_t X_{t+1} + DX_{t-1} + S \epsilon_t$$ (43)

while the PLM is described by means of a VECM in the inferred stochastic endpoints:

$$X_t = (I - \Phi^P)\xi^P_{t|t} + \Phi^P X_{t-1} + \Sigma^P \xi_t$$ (44)

and a learning rule based on the Kalman filter updating rule:

$$\xi^P_{t|t} = \xi^P_{t-1|t-1} + K(X_t - E^P_{t-1} X_t).$$ (45)

A.1 Deriving the Actual Law of Motion

A first step in obtaining the actual law of motion (ALM) consists of deriving the expectations implied by the PLM, equations (44) and (45). Under the PLM, the one-step ahead prediction, $E^P_t X_{t+1}$, is given by:

$$E^P_t X_{t+1} = (I - \Phi^P)E^P_t \xi^P_{t+1|t+1} + \Phi^P X_t.$$ (46)

Under the PLM dynamics, the stochastic endpoints $\xi^P_{t|t}$ are random walks, i.e. $E^P_{t-1}(X_t - E^P_{t-1} X_t) = 0$, such that $E^P_t \xi^P_{t+1|t+1} = \xi^P_{t|t}$. The one-step ahead expectations are given by:

$$E^P_t X_{t+1} = (I - \Phi^P)\xi^P_{t|t} + \Phi^P X_t.$$ (47)

Substituting the learning rule, equation (45), for $\xi^P_{t|t}$ we obtain a description for the expectations as:

$$E^P_t X_{t+1} = (I - \Phi^P)(\xi^P_{t-1|t-1} + K(X_t - E^P_{t-1} X_t)) + \Phi^P X_t.$$ (48)

or equivalently, by lagging equation (47) one period giving a closed form expression for $E^P_{t-1} X_t = (I - \Phi^P)\xi^P_{t-1|t-1} + \Phi^P X_{t-1}$:

$$E^P_t X_{t+1} = (I - \Phi^P)(\xi^P_{t-1|t-1} + K(X_t - (I - \Phi^P)\xi^P_{t-1|t-1} - \Phi^P X_{t-1})) + \Phi^P X_t.$$ (49)

This expression can also be written as:

$$E^P_t X_{t+1} = (I - (I - \Phi^P)K)(I - \Phi^P)\xi^P_{t-1|t-1} + (\Phi^P + (I - \Phi^P)K)X_t - (I - \Phi^P)K\Phi^P X_{t-1}.$$ (50)

Denoting the matrix $(I - \Phi^P)K$ by $K_\Phi$ we obtain the final expression for the one-step ahead expectation as:

$$E^P_t X_{t+1} = (I - K_\Phi)(I - \Phi^P)\xi^P_{t-1|t-1} + (\Phi^P + K_\Phi)X_t - K_\Phi\Phi^P X_{t-1}.$$ (51)

The second step in deriving the ALM dynamics consists of inserting the subjective expectations, equation (51), into the structural equations, i.e. equation (43):

$$AX_t = C + B((I - K_\Phi)(I - \Phi^P)\xi^P_{t-1|t-1} + (\Phi^P + K_\Phi)X_t - K_\Phi\Phi^P X_{t-1}) + DX_{t-1} + S \epsilon_t.$$ (52)
Solving for $X_t$, we obtain:

$$X_t = (A - B(\Phi^P + K_\Phi))^{-1}C + (A - B(\Phi^P + K_\Phi))^{-1}B(I - K_\Phi)(I - \Phi^P)\xi^P_{t-1|t-1}$$

$$= (A - B(\Phi^P + K_\Phi))^{-1}(D - B\Phi K_\Phi \Phi^P)X_{t-1} + (A - B(\Phi^P + K_\Phi))^{-1}S\varepsilon_t. \tag{53}$$

Note that if the rational expectations solution is unique, and if $\Phi^P = \Phi^{re}$, the expression $(A - B(\Phi^P + K_\Phi))^{-1}(D - B\Phi K_\Phi \Phi^P)$ equals $\Phi^P$ which allows us to rewrite the above dynamics as:

$$X_t = (A - B(\Phi^P + K_\Phi))^{-1}C + (A - B(\Phi^P + K_\Phi))^{-1}B(I - K_\Phi)(I - \Phi^P)\xi^P_{t-1|t-1}$$

$$\Phi^P X_{t-1} + (A - B(\Phi^P + K_\Phi))^{-1}S\varepsilon_t. \tag{54}$$

Equation (54) describes the actual law of motion for the observable macroeconomic variables as a function of the previous state, $X_{t-1}$, the inferred stochastic endpoints, $\xi^P_{t-1|t-1}$ and the structural shocks, $\varepsilon_t$. This description is only a partial description of the ALM, since the dynamics of the stochastic endpoints is not taken into account. In order to obtain a complete characterization of the ALM, we add the learning rule, i.e. equation (45). The joint dynamics of the observable macroeconomic variables, $X_t$, and the inferred stochastic endpoints, $\xi^P_{t|t}$ is given by:

$$\begin{bmatrix}
I & 0 \\
-K & I
\end{bmatrix}
\begin{bmatrix}
X_t \\
\xi^P_{t|t}
\end{bmatrix}
= \begin{bmatrix}
(A - B(\Phi^P + K_\Phi))^{-1}C \\
0
\end{bmatrix}
+ \begin{bmatrix}
\Phi^P \\
-K\Phi^P
\end{bmatrix}
\begin{bmatrix}
(A - B(\Phi^P + K_\Phi))^{-1}B(I - K_\Phi)(I - \Phi^P) \\
(I - K(I - \Phi^P))
\end{bmatrix}
\begin{bmatrix}
X_{t-1} \\
\xi^P_{t-1|t-1}
\end{bmatrix}
+ \begin{bmatrix}
(A - B(\Phi^P + K_\Phi))^{-1}S \\
0
\end{bmatrix}
\varepsilon_t,$$

where the dynamics for $\xi^P_{t|t}$ are given by equation (45). Finally, pre-multiplying by

$$\begin{bmatrix}
I & 0 \\
-K & I
\end{bmatrix}^{-1}
= \begin{bmatrix}I & 0 \\K & I\end{bmatrix} \tag{55}$$

yields a complete description of the ALM:

$$\begin{bmatrix}
X_t \\
\xi^P_{t|t}
\end{bmatrix}
= \begin{bmatrix}
(A - B(\Phi^P + K_\Phi))^{-1}C \\
K(A - B(\Phi^P + K_\Phi))^{-1}C
\end{bmatrix}
$$

$$+ \begin{bmatrix}
\Phi^P \\
-K\Phi^P
\end{bmatrix}
\begin{bmatrix}
(A - B(\Phi^P + K_\Phi))^{-1}B(I - K_\Phi)(I - \Phi^P) \\
(I - K(I - \Phi^P))
\end{bmatrix}
\begin{bmatrix}
X_{t-1} \\
\xi^P_{t-1|t-1}
\end{bmatrix}
+ \begin{bmatrix}
(A - B(\Phi^P + K_\Phi))^{-1}S \\
K(A - B(\Phi^P + K_\Phi))^{-1}S
\end{bmatrix}
\varepsilon_t.$$

This ALM is represented in extended state space, $\tilde{X}_{t|t} = [X'_t, \xi^P_{t|t}]$ by

$$\tilde{X}_{t|t} = \tilde{C}^A + \tilde{\Phi}^A \tilde{X}_{t-1|t-1} + \tilde{\Sigma}^A \varepsilon_t \tag{56}$$
Noting that we obtain that this point is a steady state if it solves:

\[
\begin{align*}
    & (A - B(\Phi^P + K_\phi))^{-1} C \\
    & K(A - B(\Phi^P + K_\phi))^{-1} C)
\end{align*}
\]

with

\[
\begin{align*}
    \hat{C} & = \begin{bmatrix} (A - B(\Phi^P + K_\phi))^{-1} C \\
    K(A - B(\Phi^P + K_\phi))^{-1} C
\end{align*}
\]

\[
\begin{align*}
    \hat{\Phi} & = \begin{bmatrix} A - B(\Phi^P + K_\phi) \quad (I - K(I - (A - B(\Phi^P + K_\phi))^{-1} B(I - K_\phi)(I - \Phi^P)) \\
    0 \quad I - K(I - (A - B(\Phi^P + K_\phi))^{-1} B(I - K_\phi)(I - \Phi^P)
\end{align*}
\]

\[
\begin{align*}
    \hat{\Sigma} & = \begin{bmatrix} (A - B(\Phi^P + K_\phi))^{-1} S \\
    K(A - B(\Phi^P + K_\phi))^{-1} S
\end{align*}
\]

(57)

A.2 Properties of the Actual Law of Motion

Based on the final representation of the ALM as stated in equation (56), some properties of the ALM can be described in more detail. A first property is that the unconditional mean of the ALM coincides with the unconditional mean of the rational expectations model. Denoting the expectations operators under rational expectations and under the ALM by respectively \( E^{re} \) and \( E^A \), the equivalence between unconditional expectations can be formalized as:

\[
E^A X_t = E^{re} X_t = (I - \Phi^{re})^{-1} C^{re}
\]

\[
E^A \xi_{tt} = E^{re} X_t = (I - \Phi^{re})^{-1} C^{re}.
\]

We show this property by showing that \( X_t = (I - \Phi^{re})^{-1} C^{re} = \xi_{tt} \) is a steady state under the ALM. In the derivation we make extensive use of the properties of the rational expectations solution. More specifically, the unconditional mean for \( X_t \) based on the rational expectations model is given by:

\[
E^{re}(X_t) = (I - \Phi^{re})^{-1} C^{re}
\]

where the values for \( \Phi^{re} \) and \( C^{re} \) satisfy the rational expectations conditions:

\[
\begin{align*}
    C^{re} & = (A - B\Phi^{re})^{-1} C + (A - B\Phi^{re})^{-1} B C^{re} \\
    \Phi^{re} & = (A - B\Phi^{re})^{-1} D \\
    \Sigma^{re} & = (A - B\Phi^{re})^{-1} S
\end{align*}
\]

(60)

We now show that the unconditional mean of \( X_t \) under the ALM, denoted by \( E^A_t X_t \), coincides with the unconditional mean of the rational expectations model:

\[
E^A_t X_t = E^{re} X_t = (I - \Phi^{re})^{-1} C^{re}.
\]

(61)

In order to show this equivalence, we show that the point \( X_t = (I - \Phi^{re})^{-1} C^{re} \) and \( \xi_{tt} = (I - \Phi^{re})^{-1} C^{re} \) are a steady state for the ALM. Substituting this particular point in the ALM, we obtain that this point is a steady state if it solves:

\[
(I - \Phi^{re})^{-1} C^{re} = (A - B(\Phi^P + K_\phi))^{-1} C + \Phi^P (I - \Phi^{re})^{-1} C^{re} +
\]

\[
(A - B(\Phi^P + K_\phi))^{-1} B(I - K_\phi)(I - \Phi^P)(I - \Phi^{re})^{-1} C^{re}.
\]

Noting that \( \Phi^{re} = \Phi^P \) we can rewrite the equation by subtracting from both sides \( \Phi^P (I - \Phi^{re})^{-1} C^{re} \), resulting in the equality:

\[
(I - \Phi^P)(I - \Phi^{re})^{-1} C^{re} = (A - B(\Phi^P + K_\phi))^{-1} C + (A - B(\Phi^P + K_\phi))^{-1} B(I - K_\phi)(I - \Phi^P)(I - \Phi^{re})^{-1} C^{re}.
\]
Pre-multiplying by \((A - B(\Phi^P + K_\Phi))^{-1}\),

\[(A - B(\Phi^P + K_\Phi))C^{re} = C + B(I - K_\Phi)C^{re}.
\]

Finally, this condition holds whenever a rational expectations equilibrium exists, i.e. adding \(BK_\Phi C^{re}\) to both sides, the above condition reduces to the rational expectations condition for \(C^{re}\)

\[(A - B\Phi^P)C^{re} = C + BC^{re}.
\]

The above derivation thus implies that if a rational expectations equilibrium exists, then the unconditional expectations of the rational expectations equilibrium coincides with the steady state of the ALM. If we assume, moreover, that all of the eigenvalues of \(\tilde{\Phi}^A\) are strictly smaller than 1 in absolute value, the steady state of the ALM is attracting and defines the unconditional mean of the observable variables \(X_t\). The second equality, i.e. \(E^A\xi^P_t = (I - \Phi^{re})^{-1}C^{re}\) can be shown analogously.

A second property is the unconditional normality of the extended state vector \(\tilde{X}_{t|t}\) under the ALM. Assuming a standard normal distribution for the structural shocks, \(\varepsilon_t\), it is well known that the linearity of the state space dynamics and the assumed stability of the ALM (all eigenvalues of \(\tilde{\Phi}^A\) are assumed to be strictly smaller than 1) implies that the unconditional distribution for \(\tilde{X}_{t|t}\) is:

\[\tilde{X}_{t|t} \sim N(E^A\tilde{X}_{t|t}, \Omega_{\tilde{X}})\]

with:

\[E^A\tilde{X}_{t|t} = \iota_{2\times1} \otimes (I - \Phi^{re})^{-1}C^{re}\]

\[\text{vec}(\Omega_{\tilde{X}}) = (I - \tilde{\Phi}^A \otimes \tilde{\Phi}^A)^{-1}\text{vec}(\tilde{\Sigma}^A\tilde{\Sigma}^A)\].
Inflation ($\pi$) is expressed in annual terms and is constructed by taking the quarterly percentage change in the GDP deflator (collected from the National income and Product Accounts). The output gap ($y$) series is constructed from data provided by the Congressional Budget Office (CBO). The Fed rate is used as the short-term interest rate or the policy rate ($i$). Bond yield data concern month-end yields on zero-coupon U.S. Treasury bonds, expressed in annual terms. Data on yields up to five years are based on CRSP data and on ten-year yields based on the Federal Reserve. Mean denotes the sample arithmetic average in percentage p.a., Stdev the standard deviation, Auto the first order quarterly autocorrelation, Skew skewness, and Kurt kurtosis. JB stands for the Jarque-Bera normality test statistic with the significance level at which the null of normality may be rejected underneath it. ** indicates significance at the 5% confidence level.
Table 2
Parameter estimates - Rational Expectations Macro and Rational Expectations I

\[ \pi_t = \mu_\pi \frac{\pi_{t+1}}{E_t} + (1 - \mu_\pi)\pi_{t-1} + \kappa_\pi y_t + \sigma_\pi \varepsilon_{\pi,t} \]

\[ y_t = \mu_y \frac{y_{t+1}}{E_t} + (1 - \mu_y)y_{t-1} + \phi(i_t - E_r\pi_{t+1} - r) + \sigma_y \varepsilon_{y,t} \]

\[ i_t = (1 - \gamma_{i-1}) [r + E\pi_{t+1} + \gamma_{\pi}(\pi_t - \pi^*) + \gamma_y y_t] + \gamma_{i-1} i_{t-1} + \sigma_i \varepsilon_{i,t} \]

<table>
<thead>
<tr>
<th></th>
<th>Rat. Exp. Macro</th>
<th>Rat. Exp. 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pi )-eq.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_{\pi,1} )</td>
<td>0.524**</td>
<td>0.527**</td>
</tr>
<tr>
<td>( \kappa_\pi \times 10^2 )</td>
<td>0.055 (0.278)</td>
<td>0.582** (0.236)</td>
</tr>
<tr>
<td>( y )-eq.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \mu_y )</td>
<td>0.509**</td>
<td>0.580**</td>
</tr>
<tr>
<td>( \phi )</td>
<td>-0.019*</td>
<td>-0.012**</td>
</tr>
<tr>
<td>( i )-eq.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \gamma_{i-1} )</td>
<td>0.862**</td>
<td>0.934**</td>
</tr>
<tr>
<td>( \gamma_\pi )</td>
<td>0.674*</td>
<td>0.100 (0.165)</td>
</tr>
<tr>
<td>( \gamma_y )</td>
<td>0.569 (0.504)</td>
<td>0.010 (0.172)</td>
</tr>
<tr>
<td>( r )</td>
<td>0.025**</td>
<td>0.028**</td>
</tr>
<tr>
<td>( \pi^* )</td>
<td>0.032**</td>
<td>0.044**</td>
</tr>
<tr>
<td>Stdev</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \sigma_\pi )</td>
<td>0.0063**</td>
<td>0.0069**</td>
</tr>
<tr>
<td>( \sigma_y )</td>
<td>0.0043**</td>
<td>0.0070**</td>
</tr>
<tr>
<td>( \sigma_i )</td>
<td>0.0133**</td>
<td>0.0134**</td>
</tr>
<tr>
<td>Struct</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \delta_\pi )</td>
<td>0.908**</td>
<td>0.895**</td>
</tr>
<tr>
<td>( h )</td>
<td>1.000 —</td>
<td>0.738**</td>
</tr>
<tr>
<td>( \sigma \times 10^{-2} )</td>
<td>0.274 (0.170)</td>
<td>0.496**</td>
</tr>
</tbody>
</table>

Stdev denotes standard deviation. Struct denotes structural parameters. Maximum likelihood estimates with standard errors between brackets. ** denotes significantly different from zero at the 5% significance level, and * at the 10% level.
Table 3
Parameter estimates - Rational Expectations II

\[
\begin{align*}
\pi_t &= \mu_{\pi,1}E_t\pi_{t+1} + (1 - \mu_{\pi,1})\pi_{t-1} + \kappa_{\pi}E_t\pi_{t+1} + \sigma_{\pi}\varepsilon_{\pi,t} \\
y_t &= \mu_yE_ty_{t+1} + (1 - \mu_y)y_{t-1} + \phi(i_t - E_t\pi_{t+1} - r) + \sigma_y\varepsilon_{y,t} \\
i_t &= (1 - \gamma_{i-1})[r + E\pi_{t+1} + \gamma_y(\pi_t - \pi^*) + \gamma_yy_{t+1} + \gamma_{i-1}i_{t-1} + \sigma_i\varepsilon_{i,t}]
\end{align*}
\]

<table>
<thead>
<tr>
<th>Parameter estimates - Rational Expectations II</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\pi_{eq})</td>
</tr>
<tr>
<td>&amp; (\kappa_{\pi}(\times10^2))</td>
</tr>
<tr>
<td>(y_{eq})</td>
</tr>
<tr>
<td>&amp; (\phi)</td>
</tr>
<tr>
<td>&amp; (r)</td>
</tr>
<tr>
<td>(i_{eq})</td>
</tr>
<tr>
<td>&amp; (\gamma_{\pi})</td>
</tr>
<tr>
<td>&amp; (\gamma_y)</td>
</tr>
<tr>
<td>&amp; (\pi^*)</td>
</tr>
<tr>
<td>&amp; (\gamma_{i-1})</td>
</tr>
<tr>
<td>&amp; (\gamma_{\pi})</td>
</tr>
<tr>
<td>&amp; (\gamma_y)</td>
</tr>
<tr>
<td>&amp; (\pi^*)</td>
</tr>
<tr>
<td>(\sigma_{\pi})</td>
</tr>
<tr>
<td>(\sigma_y)</td>
</tr>
<tr>
<td>(\sigma_i)</td>
</tr>
<tr>
<td>(\delta_{\pi})</td>
</tr>
<tr>
<td>(h)</td>
</tr>
<tr>
<td>(\sigma(\times10^{-2}))</td>
</tr>
</tbody>
</table>

\(Stdev\) denotes standard deviation. \(Struct\) denotes structural parameters. Maximum likelihood estimates with standard errors between brackets. ** denotes significantly different from zero at the 5% significance level, and * at the 10% level.
Table 4
Parameter estimates - Rational Expectations III

\[ \pi_t = \mu_{\pi,1}E_t\pi_{t+1} + (1 - \mu_{\pi,1})\pi_{t-1} + \kappa_{\pi}y_t + \sigma_{\pi}\varepsilon_{\pi,t} \]
\[ y_t = \mu_yE_ty_{t+1} + (1 - \mu_y)y_{t-1} + \phi(i_t - E_t\pi_{t+1} - r) + \sigma_y\varepsilon_{y,t} \]
\[ i_t = (1 - \gamma_{i-1}) \left[ r + E\pi_{t+1} + \gamma_{\pi}(\pi_t - \pi^*) + \gamma_{y}y_t \right] + \gamma_{i-1}i_{t-1} + \sigma_i\varepsilon_{i,t} \]

<table>
<thead>
<tr>
<th>Rational Expectations III</th>
<th>Martins</th>
<th>Burns</th>
<th>Miller</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\pi)-eq. (\mu_{\pi,1})</td>
<td>0.638** (0.017)</td>
<td>(\kappa_{\pi}(\times 10^2))</td>
<td>0.925* (0.549)</td>
</tr>
<tr>
<td>(y)-eq. (\mu_y)</td>
<td>0.582** (0.013)</td>
<td>(\phi)</td>
<td>-0.025** (0.012)</td>
</tr>
<tr>
<td></td>
<td>(r)</td>
<td>0.020** (0.006)</td>
<td></td>
</tr>
<tr>
<td>(i)-eq. (\gamma_{i-1})</td>
<td>0.845** (0.029)</td>
<td>0.482** (0.041)</td>
<td>0.654** (0.289)</td>
</tr>
<tr>
<td>(\gamma_{\pi})</td>
<td>0.511 (0.681)</td>
<td>0.176 (0.170)</td>
<td>0.233 (1.391)</td>
</tr>
<tr>
<td>(\gamma_y)</td>
<td>0.363 (0.346)</td>
<td>0.516** (0.248)</td>
<td>0.846 (0.991)</td>
</tr>
<tr>
<td>(\pi)</td>
<td>0.018** (0.005)</td>
<td>0.047** (0.002)</td>
<td>0.060** (0.006)</td>
</tr>
<tr>
<td>(i)-eq. (\gamma_{i-1})</td>
<td>0.779** (0.029)</td>
<td>0.943** (0.034)</td>
<td>0.928** (0.026)</td>
</tr>
<tr>
<td>(\gamma_{\pi})</td>
<td>0.317 (0.308)</td>
<td>1.657 (1.759)</td>
<td>1.338 (0.949)</td>
</tr>
<tr>
<td>(\gamma_y)</td>
<td>0.395* (0.221)</td>
<td>-0.239 (0.396)</td>
<td>1.092 (0.692)</td>
</tr>
<tr>
<td>(\pi)</td>
<td>0.078** (0.003)</td>
<td>0.049** (0.002)</td>
<td>0.032** (0.001)</td>
</tr>
<tr>
<td>Stdev (\sigma_{\pi})</td>
<td>0.0081** (0.0006)</td>
<td>(\sigma_y)</td>
<td>0.0053** (0.0004)</td>
</tr>
<tr>
<td></td>
<td>(\sigma_i)</td>
<td>0.0136** (0.0008)</td>
<td></td>
</tr>
<tr>
<td>Struct (\delta_{\pi})</td>
<td>0.567** (0.042)</td>
<td>(h)</td>
<td>0.750** (0.044)</td>
</tr>
<tr>
<td></td>
<td>(\sigma(\times 10^{-2}))</td>
<td>0.233** (0.107)</td>
<td></td>
</tr>
<tr>
<td>Prices of risk (\Lambda_{0,\pi})</td>
<td>-0.974 (0.612)</td>
<td>(\Lambda_{0,y})</td>
<td>-0.271 (0.389)</td>
</tr>
<tr>
<td></td>
<td>(\Lambda_{0,i})</td>
<td>0.045 (0.136)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\Lambda_{1,\pi,\pi})</td>
<td>0.067 (7.131)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\Lambda_{1,y,y})</td>
<td>-0.573 (14.905)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(\Lambda_{1,i,i})</td>
<td>3.010 (2.152)</td>
<td></td>
</tr>
</tbody>
</table>

*\(Stdev\) denotes standard deviation. *\(Struct\) denotes structural parameters. Maximum likelihood estimates with standard errors between brackets. ** denotes significantly different from zero at the 5% significance level, and * at the 10% level.
Table 5
Parameter estimates- Learning Macro and Learning I

\[
\begin{align*}
\pi_t &= \mu_{\pi,1} E\pi_{t+1} + (1 - \mu_{\pi,1})\pi_{t-1} + \kappa_{\pi} y_t + \sigma_{\pi} \epsilon_{\pi,t} \\
y_t &= \mu_y E\pi_{t+1} + (1 - \mu_y)\pi_{t-1} + \phi(i_t - E\pi_{t+1} - r) + \sigma_y \epsilon_{y,t} \\
i_t &= (1 - \gamma_{i-1}) \left[ r + E\pi_{t+1} + \gamma_{\pi}(\pi_t - \pi^*) + \gamma_y y_t \right] + \gamma_{i-1} i_{t-1} + \sigma_i \epsilon_{i,t}
\end{align*}
\]

<table>
<thead>
<tr>
<th></th>
<th>Learning Macro</th>
<th>Learning I</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\pi)-eq.</td>
<td>(\mu_{\pi,1})</td>
<td>0.672** (0.056)</td>
</tr>
<tr>
<td></td>
<td>(\kappa_{\pi} \times 10^2)</td>
<td>0.431 (0.504)</td>
</tr>
<tr>
<td>(y)-eq.</td>
<td>(\mu_y)</td>
<td>0.504** (0.028)</td>
</tr>
<tr>
<td></td>
<td>(\phi)</td>
<td>-0.008 (0.023)</td>
</tr>
<tr>
<td>(i)-eq.</td>
<td>(\gamma_{i-1})</td>
<td>0.833** (0.039)</td>
</tr>
<tr>
<td></td>
<td>(\gamma_{\pi})</td>
<td>0.401 (0.300)</td>
</tr>
<tr>
<td></td>
<td>(\gamma_y)</td>
<td>0.504 (0.416)</td>
</tr>
<tr>
<td></td>
<td>(r)</td>
<td>0.028 (0.265)</td>
</tr>
<tr>
<td></td>
<td>(\pi^*)</td>
<td>0.031 (0.669)</td>
</tr>
<tr>
<td>Stdev</td>
<td>(\sigma_{\pi})</td>
<td>0.0062** (0.0005)</td>
</tr>
<tr>
<td></td>
<td>(\sigma_y)</td>
<td>0.0043** (0.0003)</td>
</tr>
<tr>
<td></td>
<td>(\sigma_{i})</td>
<td>0.0132** (0.0004)</td>
</tr>
<tr>
<td>Struct</td>
<td>(\delta_{\pi})</td>
<td>0.489** (0.123)</td>
</tr>
<tr>
<td></td>
<td>(h)</td>
<td>1.000 —</td>
</tr>
<tr>
<td></td>
<td>(\sigma(\times 10^{-2}))</td>
<td>0.618 (1.778)</td>
</tr>
<tr>
<td>Learning</td>
<td>(\sigma_{\xi,\pi})</td>
<td>0.044** (0.013)</td>
</tr>
<tr>
<td></td>
<td>(\sigma_{\xi,y})</td>
<td>0.000 —</td>
</tr>
<tr>
<td></td>
<td>(\sigma_{\xi,r})</td>
<td>0.043 (0.141)</td>
</tr>
<tr>
<td>Initial points</td>
<td>(\xi_{0,\pi})</td>
<td>0.018 (0.022)</td>
</tr>
<tr>
<td></td>
<td>(\xi_{0,y})</td>
<td>0.000 —</td>
</tr>
<tr>
<td></td>
<td>(\xi_{0,i})</td>
<td>0.014 (0.022)</td>
</tr>
</tbody>
</table>

\(Stdev\) denotes standard deviation. \(Struct\) denotes structural parameters. Maximum likelihood estimates with standard errors between brackets. ** denotes significantly different from zero at the 5% significance level, and * at the 10% level.
Table 6

Parameter estimates - Learning II

\[ \pi_t = \mu_\pi E_t \pi_{t+1} + (1 - \mu_\pi) \pi_{t-1} + \kappa_\pi y_t + \sigma_\pi \varepsilon_{\pi,t} \]

\[ y_t = \mu_y E_t y_{t+1} + (1 - \mu_y) y_{t-1} + \phi (i_t - E_t \pi_{t+1} - r) + \sigma_y \varepsilon_{y,t} \]

\[ i_t = (1 - \gamma_{i-1}) \left[ r + E \pi_{t+1} + \gamma_\pi (\pi_t - \pi^*) + \gamma_y y_t \right] + \gamma_{i-1} i_{t-1} + \sigma_i \varepsilon_{i,t} \]

<table>
<thead>
<tr>
<th></th>
<th>Learning II</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pi )-eq.</td>
<td></td>
</tr>
<tr>
<td>( \mu_\pi )</td>
<td>0.728**</td>
</tr>
<tr>
<td>( \kappa_\pi \times 10^2 )</td>
<td>1.182**</td>
</tr>
<tr>
<td>( y )-eq.</td>
<td></td>
</tr>
<tr>
<td>( \mu_y )</td>
<td>0.528**</td>
</tr>
<tr>
<td>( \phi )</td>
<td>-0.022**</td>
</tr>
<tr>
<td>( r )</td>
<td>0.026</td>
</tr>
<tr>
<td>( i )-eq.</td>
<td></td>
</tr>
<tr>
<td>( \gamma_{i-1} )</td>
<td>0.804**</td>
</tr>
<tr>
<td>( \gamma_\pi )</td>
<td>0.406</td>
</tr>
<tr>
<td>( \gamma_y )</td>
<td>0.012</td>
</tr>
<tr>
<td>( \pi^* )</td>
<td>0.028</td>
</tr>
<tr>
<td>Learning</td>
<td></td>
</tr>
<tr>
<td>( \sigma_{\zeta,\pi} )</td>
<td>0.018**</td>
</tr>
<tr>
<td>( \sigma_{\zeta,y} )</td>
<td>0.000</td>
</tr>
<tr>
<td>( \sigma_{\zeta,r} )</td>
<td>0.007</td>
</tr>
<tr>
<td>( i )-eq.</td>
<td></td>
</tr>
<tr>
<td>( \gamma_{i-1} )</td>
<td>0.185</td>
</tr>
<tr>
<td>( \gamma_\pi )</td>
<td>0.564**</td>
</tr>
<tr>
<td>( \gamma_y )</td>
<td>0.109</td>
</tr>
<tr>
<td>( \pi^* )</td>
<td>0.003</td>
</tr>
<tr>
<td>Learning</td>
<td></td>
</tr>
<tr>
<td>( \sigma_{\zeta,\pi} )</td>
<td>0.004</td>
</tr>
<tr>
<td>( \sigma_{\zeta,y} )</td>
<td>0.000</td>
</tr>
<tr>
<td>( \sigma_{\zeta,r} )</td>
<td>0.031**</td>
</tr>
<tr>
<td>Stdev</td>
<td></td>
</tr>
<tr>
<td>( \sigma_\pi )</td>
<td>0.0088**</td>
</tr>
<tr>
<td>( \sigma_y )</td>
<td>0.0048**</td>
</tr>
<tr>
<td>( \sigma_i )</td>
<td>0.0105**</td>
</tr>
<tr>
<td>Struct</td>
<td></td>
</tr>
<tr>
<td>( \delta_\pi )</td>
<td>0.374**</td>
</tr>
<tr>
<td>( h )</td>
<td>0.935**</td>
</tr>
<tr>
<td>( \sigma (\times 10^{-2}) )</td>
<td>0.235**</td>
</tr>
<tr>
<td>Initial points</td>
<td></td>
</tr>
<tr>
<td>( \xi_{0,\pi} )</td>
<td>0.009</td>
</tr>
<tr>
<td>( \xi_{0,y} )</td>
<td>0.000</td>
</tr>
<tr>
<td>( \xi_{0,i} )</td>
<td>0.033**</td>
</tr>
</tbody>
</table>

*Stdev* denotes standard deviation. *Struct* denotes structural parameters. Maximum likelihood estimates with standard errors between brackets. ** denotes significantly different from zero at the 5% significance level, and * at the 10% level.
<table>
<thead>
<tr>
<th>$\pi$-eq.</th>
<th>Learning III</th>
<th>$\mu_{\pi,1}$</th>
<th>0.702** (0.023)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_{\pi} \times 10^2$</td>
<td></td>
<td>0.883** (0.417)</td>
<td></td>
</tr>
<tr>
<td>$y$-eq.</td>
<td></td>
<td>$\mu_y$</td>
<td>0.523** (0.016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\phi$</td>
<td>-0.023** (0.010)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$r$</td>
<td>0.016 (0.094)</td>
</tr>
<tr>
<td>$i$-eq.</td>
<td></td>
<td>$\gamma_{i-1}$</td>
<td>0.735** (0.088)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\gamma_\pi$</td>
<td>0.394 (0.628)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\gamma_y$</td>
<td>0.032 (0.239)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\pi^*$</td>
<td>0.015 (0.258)</td>
</tr>
<tr>
<td>Learning</td>
<td></td>
<td>$\sigma_{\zeta,\pi}$</td>
<td>0.018** (0.004)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\sigma_{\zeta,y}$</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\sigma_{\zeta,r}$</td>
<td>0.005 (0.009)</td>
</tr>
<tr>
<td>$i$-eq.</td>
<td></td>
<td>$\gamma_{i-1}$</td>
<td>0.284** (0.047)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\gamma_\pi$</td>
<td>0.659** (0.223)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\gamma_y$</td>
<td>0.042 (0.174)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\pi^*$</td>
<td>0.002 (0.148)</td>
</tr>
<tr>
<td>Learning</td>
<td></td>
<td>$\sigma_{\zeta,\pi}$</td>
<td>0.005 (0.005)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\sigma_{\zeta,y}$</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\sigma_{\zeta,r}$</td>
<td>0.027** (0.003)</td>
</tr>
<tr>
<td>Stdev</td>
<td></td>
<td>$\sigma_\pi$</td>
<td>0.0087** (0.0007)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\sigma_y$</td>
<td>0.0045** (0.0004)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\sigma_i$</td>
<td>0.0117** (0.0010)</td>
</tr>
<tr>
<td>Struct</td>
<td></td>
<td>$\delta_\pi$</td>
<td>0.424** (0.047)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$h$</td>
<td>0.956** (0.073)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\sigma \times 10^{-2}$</td>
<td>0.223** (0.0947)</td>
</tr>
<tr>
<td>Initial points</td>
<td></td>
<td>$\xi_{0,\pi}$</td>
<td>0.009 (0.012)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\xi_{0,y}$</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\xi_{0,i}$</td>
<td>0.033** (0.007)</td>
</tr>
<tr>
<td>Prices</td>
<td></td>
<td>$\Lambda_{0,\pi}$</td>
<td>0.156 (0.256)</td>
</tr>
<tr>
<td>of risk</td>
<td></td>
<td>$\Lambda_{0,y}$</td>
<td>0.132 (0.227)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\Lambda_{0,i}$</td>
<td>-0.235** (0.104)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\Lambda_{1,\pi,\pi}$</td>
<td>0.145 (3.750)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\Lambda_{1,y,y}$</td>
<td>-2.667 (5.194)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\Lambda_{1,i,i}$</td>
<td>2.197** (0.972)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\Lambda_{1,\xi_{0,\pi},\xi_{0,\pi}}$</td>
<td>-0.937 (4.538)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\Lambda_{1,\xi_{0,y},\xi_{0,y}}$</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\Lambda_{1,\xi_{0,i},\xi_{0,i}}$</td>
<td>0.029 (1.068)</td>
</tr>
</tbody>
</table>

$Stdev$ denotes standard deviation. $Struct$ denotes structural parameters. Maximum likelihood estimates with standard errors between brackets. ** denotes significantly different from zero at the 5% significance level, and * at the 10% level.
<table>
<thead>
<tr>
<th>Components</th>
<th>Macroeconomy</th>
<th>Yield curve</th>
<th>Inflation</th>
<th>Expectation</th>
<th>Total</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat. Exp. Macro</td>
<td>12.07</td>
<td>20.98</td>
<td>5.61</td>
<td></td>
<td>38.66</td>
<td>-76.70</td>
</tr>
<tr>
<td>Rat. Exp. I</td>
<td>11.75</td>
<td>23.23</td>
<td>5.55</td>
<td></td>
<td>40.53</td>
<td>-80.42</td>
</tr>
<tr>
<td>Rat. Exp. II</td>
<td>11.57</td>
<td>25.07</td>
<td>6.21</td>
<td></td>
<td>42.85</td>
<td>-84.44</td>
</tr>
<tr>
<td>Rat. Exp. III</td>
<td>11.71</td>
<td>25.25</td>
<td>6.22</td>
<td></td>
<td>43.18</td>
<td>-85.00</td>
</tr>
<tr>
<td>Learning Macro</td>
<td>12.08</td>
<td>22.08</td>
<td>6.07</td>
<td></td>
<td>40.23</td>
<td>-79.72</td>
</tr>
<tr>
<td>Learning I</td>
<td>11.81</td>
<td>26.05</td>
<td>6.50</td>
<td></td>
<td>44.36</td>
<td>-87.97</td>
</tr>
<tr>
<td>Learning II</td>
<td>12.03</td>
<td>27.74</td>
<td>6.74</td>
<td></td>
<td>46.51</td>
<td>-91.32</td>
</tr>
<tr>
<td>Learning III</td>
<td>12.02</td>
<td>27.93</td>
<td>6.78</td>
<td></td>
<td>46.73</td>
<td>-91.50</td>
</tr>
</tbody>
</table>
Table 9
Summary statistics of prediction errors of macroeconomic variables, yield curve, and survey of inflation expectations

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>0.77</td>
<td>0.76</td>
<td>0.74</td>
<td>0.74</td>
<td>0.79</td>
<td>0.79</td>
<td>0.79</td>
<td>0.79</td>
</tr>
<tr>
<td>Mean (%)</td>
<td>0.07</td>
<td>-0.04</td>
<td>-0.05</td>
<td>-0.06</td>
<td>0.06</td>
<td>-0.07</td>
<td>-0.08</td>
<td>-0.08</td>
</tr>
<tr>
<td>Stdev (%)</td>
<td>1.20</td>
<td>1.21</td>
<td>1.27</td>
<td>1.28</td>
<td>1.15</td>
<td>1.15</td>
<td>1.15</td>
<td>1.15</td>
</tr>
<tr>
<td>Auto</td>
<td>-0.25**</td>
<td>-0.20**</td>
<td>0.09</td>
<td>0.19</td>
<td>0.09</td>
<td>0.26**</td>
<td>0.42**</td>
<td>0.38**</td>
</tr>
</tbody>
</table>

Mean denotes the sample average in percentage per year, Stdev the standard deviation in percentage per year, and Auto the first order quarterly autocorrelation. ** denotes significantly different from zero at the 5% significance level.
### Table 10
Summary statistics of fitting errors of yield curve and survey of inflation expectations

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$y_{1y}$</td>
<td>$y_{3y}$</td>
<td>$y_{5y}$</td>
<td>$y_{10y}$</td>
<td>$S_{1y}$</td>
<td>$S_{10y}$</td>
<td>$y_{1y}$</td>
</tr>
<tr>
<td>Mean (%)</td>
<td>0.84</td>
<td>0.54</td>
<td>0.38</td>
<td>0.23</td>
<td>0.70</td>
<td>0.38</td>
<td>-0.04</td>
</tr>
<tr>
<td>Stdev (%)</td>
<td>1.10</td>
<td>1.75</td>
<td>1.96</td>
<td>2.16</td>
<td>1.00</td>
<td>1.01</td>
<td>0.97</td>
</tr>
<tr>
<td>Auto</td>
<td>0.64**</td>
<td>0.85**</td>
<td>0.91**</td>
<td>0.95**</td>
<td>0.57**</td>
<td>0.97**</td>
<td>0.51**</td>
</tr>
<tr>
<td>Auto (first order quarterly</td>
<td>0.85**</td>
<td>0.91**</td>
<td>0.74**</td>
<td>0.79**</td>
<td>0.70**</td>
<td>0.85**</td>
<td>0.53**</td>
</tr>
<tr>
<td>correlation)</td>
<td>0.60</td>
<td>0.58</td>
<td>0.56</td>
<td>0.54</td>
<td>0.57</td>
<td>0.97</td>
<td>0.51</td>
</tr>
</tbody>
</table>

**Mean** denotes the sample average in percentage per year, **Stdev** the standard deviation in percentage per year, and **Auto** the first order quarterly autocorrelation. ** denotes significantly different from zero at the 5% significance level.
Figure 1: Data, USA, 1963:Q4-2003:Q4 (161 observations).
Figure 2: Inflation.
Figure 3: Real interest rate.
Figure 4: Policy interest rate.
Figure 5: Term structure fit across models, one-year yield.
Figure 6: Term structure fit across models, ten-year yield.
Figure 7: Fit of survey of ten-year average inflation expectations across models.
Figure 8: Loading - Term structure of interest rates across models.