The role of term structure in an estimated DSGE model with learning*

Pablo Aguilar[a,b] and Jesús Vázquez[b]

March, 2015

ABSTRACT: Agents can learn from financial markets to predict macroeconomic outcomes and learning dynamics can feed back into both the macroeconomy and financial markets. This paper builds on the adaptive learning (AL) model of Slobodyan and Wouters (2012b) by introducing the term structure of interest rates. This feature results in more stable learning coefficients over the whole sample period. Our estimation results show that the inclusion of the term spread in the AL model results in an increase of the parameters characterizing endogenous persistence whereas the persistence of the exogenous shocks driving price and wage dynamics decreases. Moreover, the estimated model shows that the term spread innovations are an important source of persistent fluctuations under AL. This finding stands in sharp contrast to the lack of transmission of term premium shocks to the macroeconomy under rational expectations. Furthermore, our empirical results show that our extended model with term structure does an overall better job when reproducing U.S. business cycle features.

JEL classification: C53, D84, E30, E43
Keywords: Term spread, adaptive learning, learning coefficients variability, medium-scale DSGE model

a IRES, Université catholique de Louvain
b Universidad del País Vasco (UPV/EHU)

---

*We are very grateful to Raf Wouters for his close guidance on this work. We are also thankful for helpful comments from Elena Mattana, Alfonso Novales, Rigas Oikonomou, Luca Pensieroso, and seminar participants at the University of Namur, the University of Navarra and the University of the Basque Country. Some of this research was supported by the Spanish Ministry of Economy and Competition and the Basque Government (Spain), grant numbers ECO2013-43773P and IT-793-13, respectively.

1 Introduction

Agents can learn from financial markets to predict macroeconomic outcomes. At the same time learning dynamics can feed back into both the macroeconomy and financial markets. This paper introduces the term structure of interest rates in an estimated medium-scale dynamic stochastic general equilibrium (DSGE) model with adaptive learning (AL) expectations. The aim of this paper is twofold. First, we analyze the role of term structure of interest rates in the learning process of economic agents. Second, we study how term structure innovations are transmitted into the macroeconomy under AL.

We build on the AL model of Slobodyan and Wouters (2012b) allowing the agents to use explicit data contained in the term structure of interest rates to characterize private expectations formation. The term structure provides an important additional source of information, beyond the one provided by macroeconomic variables used in previous AL models, which is observable by private market participants when pricing government bonds. The rationale of this approach is further motivated by a large empirical literature (among others, Fama, 1990; Mishkin, 1991; Estrella and Mishkin, 1997; Ang, Piazzesi and Wei, 2006) showing evidence of the ability of the term spread to predict the future evolution of both inflation and economic activity.[1]

More generally, this paper also builds on a recent growing literature investigating the role of AL, as an alternative to the assumption of rational expectations (RE), in the analysis of DSGE models. Recent papers (for instance, Orphanides and Williams, 2005a; Milani, 2007, 2008, 2011; and Eusepi and Preston 2011) focused their attention on small-scale DSGE models whereas Slobodyan and Wouters (2012a, 2012b) introduced AL in a medium-scale DSGE

[1]This paper is also related with another fast-growing strand of the literature (see for instance, Hördahl, Tristani and Vestin, 2006; Rudebusch and and Wu, 2008; Bekaert, Cho and Moreno, 2010) aiming to link the small-scale new Keynesian monetary model dynamics with the term structure of interest rates. Moreover, De Graeve, Emiris and Wouters (2009) show evidence on the importance of considering medium-scale rational expectations DSGE models, as Smets and Wouters (2007) model, to understand the links between the term structure and the aggregate economy. Thus, our paper builds also on De Graeve, Emiris and Wouters (2009) by considering AL instead of rational expectations.

model.[2] While the first group of papers considers that agent's expectations are based on a linear function of the state variables of the model whose learning coefficients are updated every period under a gain rule (i.e. the minimum state variable approach), Slobodyan and Wouters (2012b) consider an AL model with agents forming expectations using small forecasting models updated by the Kalman filter. Small forecasting models typically assume that agents form their expectations based on the information provided by observable endogenous variables, such as those showing up in the Euler equations of a DSGE model.[3]

Deviating from the RE assumption by considering AL based on small forecasting models is largely appealing for three main reasons. First, in reality agents face limited information about the economy which is at odds with the full information approach assumed under RE. Moreover, gathering and processing information is costly. So, it is likely that economic agents rely on a small set of variables when trying to figure out the relevant economic environment in their decision processes. Second, AL typically features a sluggish reaction to exogenous and latent shocks hitting the economy, which provides an additional source for explaining aggregate persistence. Third, as in Eusepi and Preston (2011) and Milani (2011), AL may add a potential important source of fluctuations associated with expectational shifts (from a certain degree of optimism to pessimism and vice versa) driving the learning process. Unfortunately, any form of deviation from the RE assumption studied in the literature is also largely arbitrary, which requires further assessment. As suggested by Adam and Marcet (2011) and Slobodyan and Wouters (2012b), considering actual data on private sector expectations available through surveys or forward-looking variables, like asset prices, might be

[2]There is also a large macroeconomic literature analyzing deviations from the RE assumption in the context of small-scale models, where the assumption of perfect information assumed under RE is somehow relaxed. This literature includes, among others, the rational inattention approach (Sims, 2003; Adam, 2007; Mackowiach and Wiederholt, 2009), the sticky information approach (Mankiw and Reis, 2002; Reis, 2009) and the imperfect information approach (Svensson and Woodford, 2004; Coenen, Levin and Wieland, 2005; Levine, Pearlman, Perendia and Yang, 2012; Pruitt, 2012).

[3]Considering small forecasting models based only on observable variables is arguably a more appealing approach to AL than the minimum state variable approach since the latter requires that agents perfectly observe the realizations of all relevant shocks. Other papers (Adam, 2005; Orphanides and Williams, 2005b; Branch and Evans, 2006; Eusepi and Preston 2011; Ormeño and Molnár, 2014) have also provided support for the use of small forecasting models on several grounds such as their relative forecast performance and their ability to approximate well the Survey of Professional Forecasters.

very useful in disciplining expectation formation.[4]

The introduction of the term structure in small forecasting models is further motivated because it helps to overcome two type of shortcomings associated with the analysis of AL models. On the one hand, AL models are arguably based on an extremely backward-looking structure by relying only on lagged values of the forecasted variable. On the other hand, estimated AL models typically use revised data whereas actual learning dynamics are likely to be driven by real-time data available to agents when forming their expectations.[5] Considering the term structure of interest rates in small forecasting models partially alleviates these limitations since term spreads are forward-looking variables and, in addition, they provide real-time information about the behavior of the aggregate economy beyond the (potentially inaccurate) information provided by real-time data on production and inflation.[6]

As a robustness check, we study two additional sensible ways of disciplining expectations in an estimated DSGE model with AL. First, a comparison of the U.S. pre-1984 and the Great Moderation periods provides a good environment for disciplining AL expectations and discriminating across alternative small forecasting models. Thus, reasonable small forecasting models should feature more stable learning coefficients in low volatility regimes such as the Great Moderation period than in high volatility regimes as the pre-1984 period. Second, we look for small forecasting models characterized by rather stable learning coefficients. The rationale for this requirement in disciplining expectations can be explained as follows: while we believe in the potential role of learning as a source of business cycle fluctuations, we also think is appropriate to play conservative by disregarding forecasting models characterized by excessive volatility of the learning coefficients since this feature would eventually result in a

[4]In this vein, recent papers introducing AL in DSGE models typically use the Survey of Professional Forecasters (SPF) to include additional observables in order to disciplining agents' predictions as in Milani (2011) and Ormeño and Molnár (2014), or to assess the performance of AL expectations as in Slobodyan and Wouters (2012b).

[5]An exception is Orphanides and Wei (2012). They used real time data, but their focus was on VAR models rather than a DSGE model. As an alternative to the use of real-time data, Ormeño and Molnár (2014) used the one-quarter-ahead expectations of the GDP deflator from the SPF recorded in real-time as a way of disciplining inflation expectations.

[6]In a parallel paper, we analyze the interactions of introducing together AL, the term structure and real-time data in the Smets and Wouters (2007) model.

likely overestimation of the importance of learning in explaining actual business cycles to the detriment of structural shocks.

Our empirical analysis shows that the inclusion of the term structure in the AL model provides additional support to important features found by Slobodyan and Wouters (2012b) in their estimated medium-scale DSGE model with AL. In particular, the AL model with term spread also reproduces the upward trend of perceived inflation persistence in the 1960s and 1970s, followed by a sharp decline until the mid-1980s, which explains the hump-shaped pattern of U.S. inflation in the last fifty years. Moreover, the AL model with term spread also results in lower estimates for the persistence of the exogenous shocks driving price and wage dynamics than those obtained in an estimated RE version of the model. The intuition is rather simple. The inclusion of AL introduces a source of persistence through the learning dynamics that reduces the role of exogenous shock persistence. However, when a forward-looking variable as the term spread is included in the AL model the sluggishness of the learning process decreases, which results in an increase of the parameters characterizing endogenous persistence in order to mimic the persistence of actual aggregate data. More precisely, the new version of the AL model with term spread exhibits the following three important features. First, the estimated persistence in the belief coefficients is lower than the one estimated for the AL model without the term structure. Second, the estimated values of a few structural parameters characterizing endogenous persistence become higher when the term spread is incorporated in the AL expectation formation mechanism. Thus, the estimated values of parameters characterizing the elasticity of the cost of adjusting capital, price and wage stickiness and both wage and price indexation schemes increase. Last but not least, the dynamics of the small forecasting models are affected by the information content of the term spread. Thus, including the term spread largely reduces the variability of the coefficients featuring the learning processes of the forward-looking variables, which allows us to be conservative when assessing the importance of learning dynamics in explaining business cycle fluctuations.

In regards to the assessment of the extended AL model with term structure, our empirical results show that our extended model does a good job when reproducing most U.S. business cycle features. For instance, the AL model with term structure outperforms the other models when replicating the size of fluctuations of many aggregate variables as well as their first-order autocorrelation and their comovement both with output growth and inflation. These results suggest that the mixture of the forward-looking dynamics incorporated through the term spread and the backward-looking dynamics of the standard AL process helps to enhance the overall fit of the model. Moreover, the estimated time-varying impulse-response functions show that the term spread innovations are an important source of persistent fluctuations under AL, especially during the 1960s and 1970s. This finding stands in sharp contrast to the absence of responses of aggregate variables to term premium shocks under RE.

The rest of the paper is organized as follows. Section 2 introduces the term structure in the medium-scale AL model. Section 3 discusses the main estimation results and analyzes the evolution of the estimated learning process coefficients over time. Section 4 provides a model assessment based on a measure of in-sample fit and business cycle features. Section 5 assesses the robustness of estimation results across alternative formulations of small forecasting models that include the term spread as a predictor. Finally, Section 6 concludes.

2 An adaptive learning model with term structure

This paper investigates the potential contribution of the term spread in the characterization of the agent's learning process. Our model builds on the estimated AL medium-scale DSGE model of Slobodyan and Wouters (2012b) by first extending the model to account for the term structure of interest rates. Second, a term spread is used in the small forecasting models of a few forward-looking variables (i.e. those involving expectations in the estimated DSGE model of Smets and Wouters, 2007). This standard medium-scale estimated DSGE model contains both nominal and real frictions affecting the choices of households and firms, we briefly present this model next. However, our main focus is on the extensions related to both the term structure and AL. The complete log linearized model of the Smets and Wouters (2007) model extended with AL and the term structure is presented in the Appendix together with a table describing parameter notation.

2.1 The DSGE model

Our model is based on the standard DSGE model of Smets and Wouters, hereinafter referred as the SW model. Households maximize their utility that depends on their levels of consumption relative to an external habit component and leisure. Labor supplied by households is differentiated by a union with monopoly power setting sticky nominal wages à la Calvo (1983). Households rent capital to firms and decide how much capital accumulate depending on the capital adjustment costs they face. Intermediate firms use capital and differentiated labor to produce differentiated goods and set prices à la Calvo. In addition, both wages and prices are partially indexed to lagged inflation when they are not re-optimized, introducing another source of nominal rigidity. As a result, current prices depend on current and expected marginal cost and past inflation whereas current wages are determined by past and expected future inflation and wages. The monetary authorities follow a Taylor-type rule reacting to inflation and output gap defined like output relative to the underlying productivity process, rather than the natural output level used in the SW model. This assumption avoids the modeling of the flexible economy which includes many additional forward-looking variables.[7] Finally, the model contains seven stochastic disturbances associated with technology, demand-side, and price and wage markup shocks.

[7]This simplifying assumption was suggested by Slobodyan and Wouters (2012b).

2.2 The term spread extension

This section introduces the term spread in the SW model. Following De Graeve, Emiris and Wouters (2009) and Vázquez, María-Dolores and Londoño (2013), we extend the DSGE model by explicitly considering the interest rates associated with alternative bond maturities indexed by j (i.e. $j = 1, 2, ..., n$). From the first-order conditions characterizing the optimal decisions of the representative consumer, one can obtain the standard consumption-based asset pricing equations associated with each maturity:

$$E_t \left[\beta^t \frac{U_C(C_{t+j}, N_{t+j}) R_t^{\{j\}}}{U_C(C_t, N_t)} \right] = 1, \; for \; j = \; 1, \; 2, \; ..., \; n,$$

where E_t stands for the RE or the AL operator depending on the scenarios analyzed below, U_C denotes the marginal utility consumption, and C_t, N_t and $R_t^{\{j\}}$denote consumption, labor and gross real return of a j-period maturity bond, respectively. After some algebra, the (linearized) model implies the expectations hypothesis model, i.e. the nominal interest rate associated with the j-period (long-term) maturity bond is the average of the expected successive interest rates associated with rolling a 1-period bond over the maturity horizon. Formally,

$$r_t^{\{j\}} = \frac{1}{j} \sum_{k=0}^{j-1} E_t r_{t+k},$$

where the interest rates are written in deviations from their respective steady-state values. Therefore, the nominal yield of the j-period maturity bond, $r_t^{\{j\}}$, is equal to the average of the expectations of the short-term (1-period) nominal interest rate, r_{t+k} for $k = 0, 1, 2, ..., j - 1$.

As is standard in the related literature, we allow for deviations of the model-implied yields from actual yields by adding a term premium shock, $\xi_t^{\{j\}}$:

$$r_t^{\{j\}} = \frac{1}{j} \sum_{k=0}^{j-1} E_t r_{t+k} + \xi_t^{\{j\}}, \tag{1}$$

We further assume that $\xi_t^{\{j\}}$ follows an AR(1) process:

$$\xi_t^{\{j\}} = \rho^{\{j\}} \xi_{t-1}^{\{j\}} + \eta_t^{\{j\}}. \tag{2}$$

This autoregressive structure allows for term premium shock persistence, measured by $\rho^{\{j\}}$, whereas $\eta_t^{\{j\}}$ is the white noise innovation of the term premium shock associated with the j-period maturity bond.[8] Furthermore, our empirical formulation below includes a constant to capture the mean of a yield.

The term spread, $sp_t^{\{j\}} = r_t^{\{j\}} - r_t$ for $2 \leq j \leq n$, is clearly a forward-looking variable under RE since a (longer term) interest rate, $r_t^{\{j\}}$, involves by definition expectations of future realizations of the short-term nominal interest rate. However, under AL, the forward-looking behavior of the term spread is entirely driven by the term premium innovations, $\eta_t^{\{j\}}$, since expectations of one-period interest rate to alternative forecast horizons are purely backward-looking. As discussed below, the estimated model shows that the term premium shocks become an important source of aggregate fluctuations under AL. This finding stands in striking contrast to the absence of transmission of term premium shocks to the macroeconomy under RE. Moreover, the consideration of the term spread in a DSGE model under AL contributes to the goal of disciplining expectations by characterizing agents' expectations beyond the one-period ahead expectations considered in standard DSGE models. Our analysis will focus on the 1-year term spread (i.e. $sp_t^{\{4\}}$) from now on because it implies a more parsimonious AL model as explained below.

[8] This structure differs from the one considered by De Graeve, Emiris and Wouters (2009) in two aspects. First, they consider a measurement error in the term spread instead of a term premium shock. Second, they consider a time-varying inflation target in the monetary policy rule. Whereas the first difference is mainly a matter of semantics the second difference may introduce an additional source of exogenous persistence. We choose to ignore this potential source of exogenous persistence for two main reasons. First, our empirical analysis shows that a time-varying inflation target is no longer needed to reproduce the actual aggregate persistence when the output gap is defined as in Slobodyan and Wouters (2012b). That is, the output gap is defined as the deviation of output from its underlying neutral productivity process and not as the natural output gap. Second, it allows a more straightforward comparison with the Slobodyan and Wouters (2012b) model that helps to assess the importance of the AL expectations formation and the role of the term spread.

2.3 The adaptive learning extension

When a researcher decides to deviate from the standard RE hypothesis, the way agents' beliefs are characterized becomes a crucial issue. This paper assumes a small linear forecasting model (for instance, an autoregressive process) that agents follow to update their expectations, the so-called "perceived law of motion" (PLM). The coefficients of the PLM are updated through a Kalman filter with the arrival of new information. Next, the small linear forecasting models are combined to form the expectation functions for the different forward-looking variables of the model. Consequently, the AL model does not impose a perfect knowledge of the model structure and shock realizations. Moreover, the AL approach allows us to investigate the ability of the term spread to restrict the set of observed relevant variables taken into account when agents form their forecast as well as the variables entering in the updating expectation processes.

We now proceed to a brief explanation of how AL expectation formation works.[9] A DSGE model can be represented in matrix form as follows:

$$A_0 \begin{bmatrix} y_{t-1} \\ w_{t-1} \end{bmatrix} + A_1 \begin{bmatrix} y_t \\ w_t \end{bmatrix} + A_2 E_t y_{t+1} + B_0 \epsilon_t = 0,$$

where y_t is the vector of endogenous variables at time t and w_t is the exogenous driving force following an AR(1):

$$w_t = \Gamma w_{t-1} + \Pi \epsilon_t,$$

where ϵ_t is the vector of innovations.

Under AL, the adaptive expectations of the forward-looking variables,$E_t y_{t+1}$, are defined as linear functions of lagged values of the variables, whose time-varying (learning) coefficients are updated as explained in the subsection below. Once the expectations of the forward-looking variables,$E_t y_{t+1}$, are computed they are plugged into the matrix representation of

[9]For a detailed explanation see Slobodyan and Wouters (2012b).

the DSGE model to obtain a backward-looking representation of the model as follows

$$\begin{bmatrix} y_t \\ w_t \end{bmatrix} = \mu_t + T_t \begin{bmatrix} y_{t-1} \\ w_{t-1} \end{bmatrix} + R_t \epsilon_t,$$

where the time-varying matrices μ_t, T_t and R_t are nonlinear functions of structural parameters (entering in matricesA_0, A_1,A_2 andB_0) together with learning coefficients discussed below.

Forming and updating expectations

Agents are assumed to have a rather limited view of the economy under AL. More precisely, their PLM process is generally defined as follows:

$$y_{t+1} = X_{t-1}\beta_{t-1} + u_{t+1},$$

where y is the vector containing the k forward-looking variables of the model, X is the matrix of the kxn regressors, β is the vector of the n updating parameters, which includes an intercept, and u is a vector of errors. These errors are linear combinations of the true model innovations, ϵ_{t+1}. So, the variance-covariance matrix, $\Sigma = E[u_{t+1}u_{t+1}^T]$, is non-diagonal.

Agents are further assumed to behave as econometricians under AL. In particular, it is assumed that they use a linear projection scheme in which the parameters are updated to form their expectations for each forward variable:

$$E_t y_{t+1} = X_{t-1}\beta_{t-1}.$$

The updating parameter vectorβ is further assumed to follow an autoregressive process where agents' beliefs are updated through a Kalman filter. This updating process can be represented as in Slobodyan and Wouters (2012b) by the following equation:

$$\beta_t - \bar{\beta} = F(\beta_{t-1} - \bar{\beta}) + v_t,$$

where F is a diagonal matrix with the learning parameter $\mid \rho \mid \leq 1$ on the main diagonal and v_t are i.i.d. errors with variance-covariance matrix V. The Kalman filter updating and transition equations for the belief coefficients and the corresponding covariance matrix are given by

$$\beta_{t|t} = \beta_{t|t-1} + R_{t|t-1}X_{t-1}\left[\Sigma + X_{t-1}^T R_{t|t-1}^{-1} X_{t-1}\right]^{-1}\left(y_t - X_{t-1}\beta_{t|t-1}\right), \tag{3}$$

with $(\beta_{t+1|t} - \bar{\beta}) = F(\beta_{t|t} - \bar{\beta})$. $\beta_{t|t-1}$ is the estimate of β using the information up to time $t-1$ (but further considering the autoregressive process followed by β), $R_{t|t}$ is the variance-covariance matrix of X, $R_{t|t-1}$ is the estimate of matrix R based on the information at time $t-1$. Therefore, the updated learning vector $\beta_{t|t}$ is equal to the previous one, $\beta_{t|t-1}$, plus a correction term that depends on the forecast error, $\left(y_t - X_{t-1}\beta_{t|t-1}\right)$. Moreover, the mean-square error, $R_{t|t}$, associated with this updated estimate is given by

$$R_{t|t} = R_{t|t-1} - R_{t|t-1}X_{t-1}\left[\Sigma + X_{t-1}^T R_{t|t-1}^{-1} X_{t-1}\right]^{-1} X_{t-1}^T R_{t|t-1}^{-1}, \tag{4}$$

with $R_{t+1|t} = F R_{t|t} F^T + V$.

A PLM with term spread

We adapt our extended SW model with the term spread to the AL version of this model. As mentioned above, one of the key ingredients of a model with AL is the way agents' expectations formation are characterized (i.e. the PLM of agents). Therefore, it is important to motivate the choice of the PLM. Slobodyan and Wouters (2012b) suggests the following form for the PLM:

$$E_t y_{t+1} = \theta_{y,t-1} + \beta_{1,y,t-1} y_{t-1} + \beta_{2,y,t-1} y_{t-2}.$$

That is, each expectation formed at time t using the information up to time $t-1$ depends on an intercept and its first two lagged values (i.e. an AR(2) model). The presence of this intercept relaxes the RE assumption of agents having perfect knowledge about a common deterministic growth rate and a constant inflation target assumed in the SW model. Thus, the consideration of a time-varying intercept coefficient allows expectations to trace trend shifts in the data and changes in the inflation target. In our DSGE model with term structure, we alternatively suggest a PLM motivated by the ability of term spreads to predict real economic activity and inflation (Estrella and Mishkin, 1997). More precise, we adopt the following PLM

$$E_t y_{t+1} = \theta_{y,t-1} + \beta_{1,y,t-1} y_{t-1} + \sum_{j=1}^{n} \beta_{sp,y,t-1}^{\{k\}} sp_{t-1}^{\{k\}}.$$

At first sight, one might think that considering the whole term structure of interest rates to characterize AL would be useful. However, considering term spreads associated with long-horizons bonds implies the need of defining the whole set of expectations of the short-term nominal interest rate from the1-period horizon up to the n-period horizon. This task cannot be accomplished because, according to the term structure equation (1), the number of parameters defining the PLM associated with these expectations dramatically increases with the number of expectations of the nominal short-term interest rate defined for alternative forecast horizons, which in practice results in severe identification problems of the PLM parameters.[10] Furthermore, there is evidence (Mishkin, 1991) showing that at longer maturities than two quarters, the term structure of interest rates helps to anticipate future inflationary pressures. For these reasons, we focus our attention on the role of the 1-year term spread to characterize the PLM of forward-looking variables:

$$E_t y_{t+1} = \theta_{y,t-1} + \beta_{1,y,t-1} y_{t-1} + \beta_{sp,y,t-1} sp_{t-1},$$

[10] For instance, if we consider the 2-year maturity yield in addition to the 1-year yield we have only an additional observable time series, but we have to estimate the parameters associated with the additional expectations of the nominal short-term interest rate from the 4-quarter ahead up to the 7-quarter ahead expectations. In the simplest case where each of these expectations involve only two parameters, considering the 2-year maturity rate requires the estimation of 8 ($= 2(7-3)$) additional learning coefficients. If we want to consider a long term rate as the 10-year maturity rate, it would then require the estimation of 72 ($= 2(39-3)$) additional learning coefficients taking into account only five additional observables (the 2-year, 3-year, 5-year, 7-year and 10-year maturity rates).

where the superscript {4} has been removed from both the spread and the associated coefficient in order to simplify notation.

The introduction of the term spread seems like a natural step forward when defining the PLM of the alternative forward-looking variables because a few of them can be potentially better predicted using the information contained in the term spread. However, the estimation of the SW model under AL incorporating this generalization of the PLM results in large parameter uncertainty leading to large standard deviations and confidence intervals for many structural parameter estimates. Moreover, as discussed above we look for small forecasting models satisfying two important criteria. First, we focus on small forecasting models featuring more stable learning coefficients in low volatility regimes such as the Great Moderation period than in high volatility regimes as the pre-1984 period. Second, we look for small forecasting models characterized by rather stable learning coefficients. Avoiding an excessive volatility of the learning coefficients is a reasonable feature for a forecasting model in order to overcome the potential issue of overestimating the importance of learning in explaining actual business cycles. The following set of parsimonious PLM for the alternative forward-looking variables of the model satisfy these criteria:[11]

$$\begin{cases} E_t y_{t+1} = \theta_{y,t-1} + \beta_{1,y,t-1} y_{t-1} + \beta_{2,y,t-1} y_{t-2}, \quad for \;\; y = l, \; q, \; i, \; w \\ E_t y_{t+1} = \theta_{y,t-1} + \beta_{1,y,t-1} y_{t-1} + \beta_{sp,y,t-1} sp_{t-1}, \quad for \; y = \; c, \; \pi, \; r^k \\ E_t r_{t+j} = \theta_{j,t-1} + \beta_{j,t-1} r_{t-1}, \quad for \; j = \; 1, \; 2, \; 3 \end{cases}$$

where l, q, i, w, c, π and r^k stand for (in deviation from their respective steady-state values or detrended by its balanced growth rate) hours worked, Tobin's q, investment, real wage, consumption, inflation and the rental rate of capital, respectively. In simple words, it is found that the SW model with term spread under an AL scheme performs better when the term spread is used instead of the second lag of the corresponding variable in the PLM of consumption, inflation and the rental rate of capital. This result suggests that the informative

[11] As discussed in Section 5 below, other PLM specifications also performed relatively well.

content of the term spread is richer than their own second lagged value for these three expectations. Moreover, the introduction of the 1-year term spread in the AL model requires the introduction of the PLM functions associated with the 1-period, 2-period and 3-period ahead forecasts of the short-term nominal interest rate at time t (i.e. $E_t r_{t+1}$, $E_t r_{t+2}$ and $E_t r_{t+3}$, respectively). In these three cases, a parsimonious PLM including only the intercept and the lagged short-term nominal interest rate, r_{t-1}, works well. Furthermore, the inclusion of a time-varying intercept coefficient in the PLM of interest rates allows the AL expectations to track term premium swings in the data due to shifts in aggregate uncertainty (e.g. the pre-Volcker period versus the Great Moderation period).

3 Estimation results

This section starts describing the data and the estimation approach. Subsequently, the estimation results for the following four alternative DSGE models are discussed: (i) the SW model, (ii) the SW model with term spread, (iii) the SW model with AL suggested by Slobodyan and Wouters (2012b), from now on SlW, and (iv) the SW model with AL and term spread, hereafter SlWTS. Models (i) and (iii) were discussed and compared at length by Slobodyan and Wouters (2012b) whereas De Graeve, Emiris and Wouters (2009) discussed model variants of (i) and (ii). Therefore, our discussion is mainly focusing on the interaction of AL expectations formation and the term spread. The section also discusses the evolution of learning coefficients over time and the associated time-varying impulse response functions.

3.1 Data and the estimation approach

To facilitate the comparison with Slobodyan and Wouters (2012b) estimation results, the alternative models are estimated using U.S. data for the sample period running from 1966:1 until 2007:4. The set of observable variables is identical to theirs (i.e. the quarterly series of the inflation rate, the Fed funds rate, the log of hours worked, and the quarterly log differences of real consumption, real investment, real wages and real GDP) with the addition of the 1-year Treasury constant maturity yield. GDP, consumption, investment and hours worked are measured in per-working age population terms. The measurement equation is

$$X_t = \begin{bmatrix} dlGDP_t \\ dlCONS_t \\ dlINV_t \\ dlWAG_t \\ lHours_t \\ dlP_t \\ FEDFUNDS_t \\ One\ year\ TB\ yield_t \end{bmatrix} = \begin{bmatrix} \overline{\gamma} \\ \overline{\gamma} \\ \overline{\gamma} \\ \overline{\gamma} \\ \bar{l} \\ \overline{\pi} \\ \overline{r} \\ \overline{r}^{\{4\}} \end{bmatrix} + \begin{bmatrix} y_t - y_{t-1} \\ c_t - c_{t-1} \\ i_t - i_{t-1} \\ w_t - w_{t-1} \\ l_t \\ \pi_t \\ r_t \\ r_t^{\{4\}} \end{bmatrix},$$

where l and dl denote the log and the log difference, respectively. $\overline{\gamma} = 100(\gamma - 1)$ is the common quarterly trend growth rate for real GDP, real consumption, real investment and real wages, which are the variables featuring a long-run trend. $\bar{l}$, $\overline{\pi}$, $\overline{r}$ and $\overline{r}^{\{4\}}$ are the steady-state levels of hours worked, inflation, the Fed funds rate and the 1-year (four-quarter) constant maturity Treasury yield, respectively.

The estimation approach follows a standard two-step Bayesian estimation procedure. First, a maximization of the log posterior function is carried out by combining prior information on the parameters with the likelihood of the data. The prior assumptions are exactly the same as in Slobodyan and Wouters (2012b). Moreover, we consider rather loose priors for the parameters characterizing the 1-year yield dynamics. The second step implements the Metropolis-Hastings algorithm, which runs a massive sequence of draws of all the possible realizations for each parameter in order to draw a picture of the posterior distribution.[12]

[12]The RE versions of the DSGE models are estimated using standard Dynare routines whereas the AL versions of the models used the codes gently provided by Slobodyan and Wouters with a few minor modifications to accommodate the presence of the term spread in the PLM.

3.2 Estimation results

Table 1 shows the estimation results associated with the four alternative estimated DSGE models divided in two panels. Panel A shows the structural parameter estimates whereas Panel B shows the estimates of the parameters describing shock processes. In general, we observe that many parameter estimates are rather robust across models with a few important differences. The discussion of these differences is organized in two parts. First, the differences between the SW model under RE and the SW model under AL (i.e. SW model versus SlW model) are discussed to assess the contribution of AL. Second, the differences between the SW model and the SW model with term spread, both under AL (i.e. SlW model versus SlWTS) are studied to understand the contribution of the term spread when interacting with AL. A similar structure is followed below when assessing the differences found across models by considering second moment statistics.[13]

SW model versus SlW

As found by Slobodyan and Wouters (2012b) the consideration of AL instead of the RE assumption in the estimated SW model largely reduces, on the one hand, the sources of exogenous persistence due to price mark-up and wage mark-up shocks. Thus, the parameter estimates of ρ_p and ρ_w decrease from 0.84 and 0.97 to 0.32 and 0.54, respectively, when considering AL instead of RE. On the other hand, AL reduces the importance of endogenous persistence induced by habit formation, h (from 0.79 to 0.68), the elasticity of the cost of adjusting capital φ (from 5.96 to 3.34) and wage indexation, ι_w (from 0.51 to 0.18). The intuition of these findings is rather simple: AL dynamics introduce an important channel

[13]We also considered the 1-year term spread (i.e. the difference between the 1-year Treasury constant maturity yield and the Fed funds rate) to estimate the SW model under RE with term spread. Introducing the term spread in the SW model barely changes parameter estimates with a few exceptions. The wage Calvo probability estimate, ξ_w, increases from 0.71 to 0.86 when the term spread is considered. The opposite occurs for the wage indexation parameter, ι_w, that goes from 0.51 to 0.24. These results suggest that the relative importance of the endogenous sources (versus the exogenous sources) in explaining price and wage persistence increases when the model is extended with the term spread dynamics. Moreover, the inverse of the Frisch elasticity of labor supply, σ_l, the volatility of the innovation and the moving average coefficient associated with the wage shock, σ_w and μ_w, and the persistence of the risk premium shock, ρ_b, increase slightly.

of endogenous persistence ignored when considering the RE assumption. As a consequence, a few sources of persistence under the RE assumption are of less importance under AL in order to reproduce the observed persistence in most macroeconomic variables. An exception is that the persistence of risk premium shocks, ρ_b, increases with AL (from 0.17 to 0.43).

SlW model versus SlWTS

The introduction of the term spread in the SlW model results in much more changes than the ones introduced by the single-step extension of the SW model with AL analyzed above. On the one hand, the new version of the AL model with term spread, reinforcing the findings of Slobodyan and Wouters (2012b), results in lower estimates for the persistence of the exogenous shocks that drive price and wage dynamics than those obtained in an estimated RE version of the model. In particular, the estimated persistence of wage mark-up shocks, $\rho_w = 0.25$, is roughly twice lower than the estimate found for the SlW model (0.54) and four times lower than the estimate (0.97) associated with the RE version of the model.

On the other hand, it is important to recall that the term spread is a forward-looking variable.[14] This feature implies that learning dynamics endowed with term spread information are less sluggish. Thus, the estimated persistence of belief coefficients, ρ, is much lower when considering the term spread in the SlWTS model (0.82) than in the SlW model (0.97). As a consequence of the much faster adjustment of belief coefficients, the estimates of most parameters capturing persistent dynamics in the SlWTS are higher in order to mimic actual data persistence. Thus, the estimates of price and wage stickiness parameters, $\xi_p = 0.71$ and $\xi_w = 0.85$, are higher than the corresponding estimates in the SlW model (0.64 and 0.82, respectively) and in the SW model (0.70 and 0.71, respectively). In the same line, the estimates of price and wage indexation parameters, $\iota_p = 0.56$ and $\iota_w = 0.53$, are also higher than the ones in the SlW model (0.27 and 0.18, respectively) and in the SW model (0.25 and 0.51, respectively). Similarly, the estimate of the elasticity of the cost of adjusting

[14]Recall that under AL, the forward-looking dynamics associated with the term spread come exclusively through the term premium innovations.

capital, $\varphi = 6.69$, is higher than the ones estimated for the SlW model (3.34) and the SW model (5.96).[15] The only exception to this pattern is that the estimate of the habit formation parameter, $h = 0.61$, is lower than the value estimated for the SW model (0.79), but closer to the one in the SlW (0.68).

Four additional differences among parameter estimates are also found. First, the estimate of the elasticity of capital utilization adjustment cost for the SlWTS model, $\psi = 0.14$, is much lower than the estimated values obtained under AL (0.50) and RE (0.55). Second, the estimate of the Frisch elasticity parameter, $\sigma_l = 2.22$, is higher than the estimated values obtained under AL (1.74) and RE (1.50). Third, the estimate of the risk aversion parameter, $\sigma_c = 1.64$, is close to the estimated value under AL (1.53), but higher than under RE (1.22). Notice that these three differences obtained by considering the term structure in the PLM further reinforces the direction of the estimate changes obtained by considering AL instead of RE. Finally, the estimated persistence of risk premium shocks, $\rho_b = 0.21$, is close to the one obtained under RE (0.17), but twice lower than the estimated value under AL (0.43). A possible explanation for this finding is that considering the term structure of interest rates, and thus longer-term expectations, helps to identify risk premium shock process parameters.

3.3 Analysis of the PLM

Figures 1A, 1B and 1C show the evolution over time of the PLM coefficients for inflation, consumption and the rental rate of capital, respectively. Each one of them contains two graphs. The graphs on the left correspond to the SlW model whereas those on the right correspond to the SlWTS model. Focusing on the SlW model graphs, we observe a strong negative correlation between the coefficients associated with the first two lags of the corresponding forward-looking variable. Indeed, these correlation coefficients are very large as discussed below. These findings suggest that the information provided by the second lag of the variable is mostly redundant in these PLM.

[15] As emphasized by Smets and Wouters (2007), a higher elasticity of the cost of adjusting capital reduces the sensitivity of investment to the real value of the existing capital stock, q.

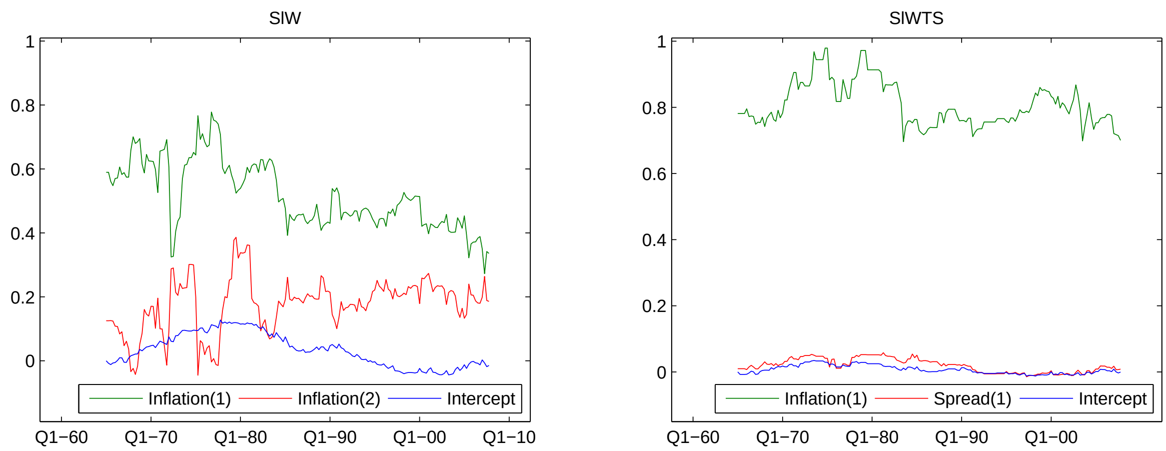


Figure 1A. PLM coefficients of inflation

In regards to the inflation learning process, Figure 1A shows that the introduction of the lagged term spread instead of the second lag of inflation in the PLM of inflation results in much more stable learning coefficients. Moreover, the intercept coefficient showing the perceived trend based on the past observed inflation rate (more precisely, the deviations of perceived long-term inflation from the constant inflation target set by the central banker) takes values closer to zero when the lagged term spread is considered. This finding suggests that the expected mean of inflation does not deviate much from the constant inflation objective of the central bank when allowing private agents to form their inflation expectations by taking into account the term spread information. Similarly, the lagged term spread coefficient always takes values close to zero but both the intercept and the term spread learning coefficients show a positive correlation with observed inflation. Thus, they increased from the start of the sample period until the early eighties when U.S. inflation was increasing and then, during the rest of the eighties and early nineties, they decreased approaching zero when inflation went down. In particular, the value of the correlation coefficient between actual U.S. inflation and the estimated term spread learning coefficient is 0.76.

The coefficient associated with lagged inflation measures perceived inflation persistence in the SlWTS model. Similarly, inflation persistence is measured by the sum of the coefficients associated with the first two lags of inflation in the SlW model. The two models implied rather similar perceived inflation persistence patterns. Thus, as found by Slobodyan and Wouters (2012b), perceived inflation persistence show an upward trend in the 1960s and early 1970s with a peak around the mid-1970 and another in the late 1970s, followed by a sharp decline until the mid-1980s. Since then, perceived inflation persistence has exhibited milder fluctuations. Thus, it moderately increases in the late 1990s followed by a mild downward trend in the 2000s.

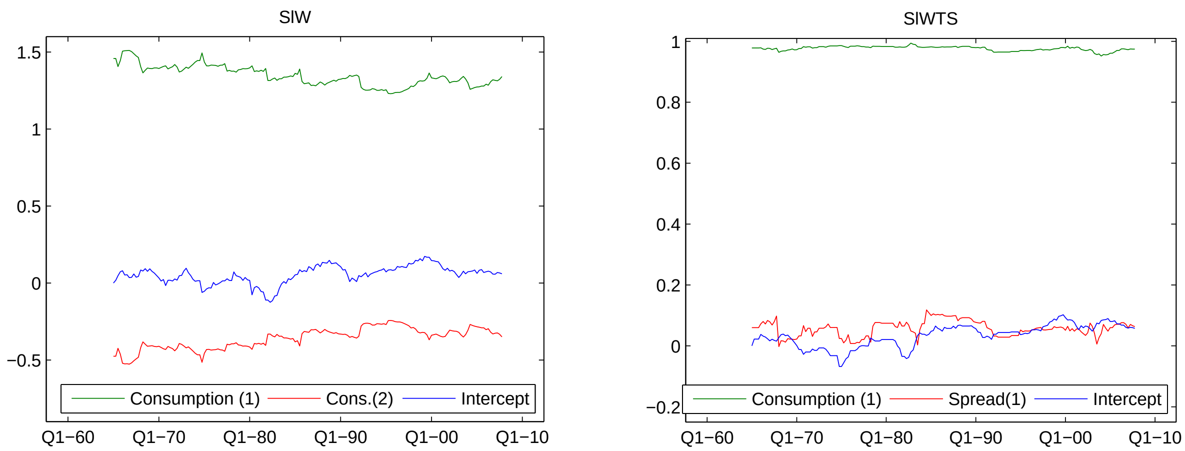


Figure 1B. PLM coefficients of consumption

The effects of introducing the lagged term spread on the PLM of consumption and the rental rate of capital are somehow qualitatively similar to the ones associated with the PLM of inflation. Thus, the inclusion of the term spread instead of the second lag of these two variables results in more stable estimated values of both the intercept and the first lag coefficient of the two PLM as shown in Figures 1B and 1C. Moreover, the term spread learning coefficients associated with the PLM of consumption and the rental rate of capital always take values close to zero, showing more pronounced fluctuations in the case of the PLM of consumption before 1984 than after as discussed below.

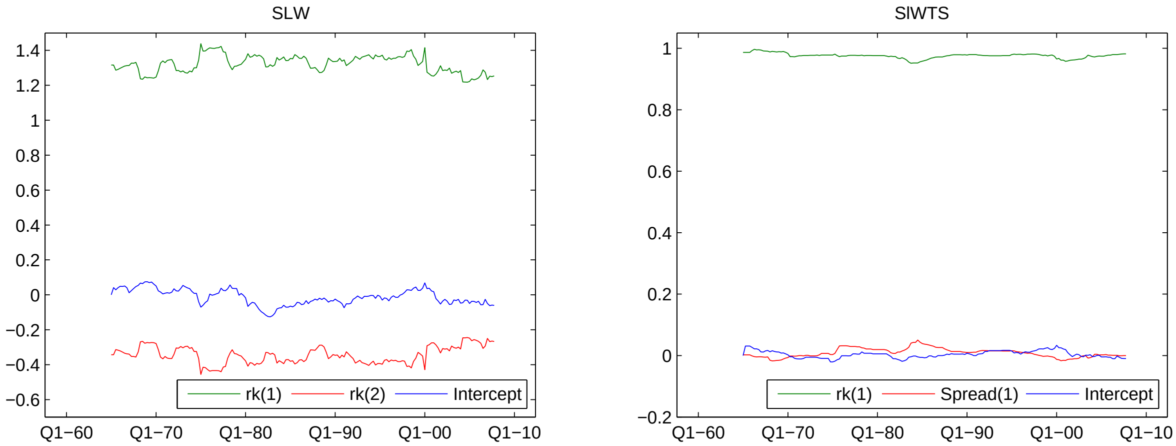


Figure 1C. PLM coefficients of the rental rate of capital

The values of the term spread coefficients close to zero shown in the PLM of inflation, consumption and the rate rate of capital may induce the reader to think that the introduction of the lagged term spread has a relative minor impact. However, the inclusion of the lagged term spread largely reduces the short-run fluctuations of the coefficients associated with the first lag of the variable and the intercept of the three PLM. In particular, we believe that a rather stable intercept coefficient is a desirable property of a forecasting model because the shifts of this coefficient mainly captures changes in the long-run expectations (i.e. expectations of the steady-state value or the balanced growth path) of the corresponding variable. It is then reasonable to think that the intercept mainly features low frequency fluctuations rather than high frequency fluctuations. Moreover, the consideration of the lagged term spread in these PLM has four additional effects. First, the intercept coefficient of the PLM of inflation falls close to zero, which is consistent with the fact that inflation is measured in deviation with respect to its steady-state value. This result is important because, as pointed out by Slobodyan and Wouters (2012b), the intercept of the PLM of inflation captures agents' belief deviations from a constant inflation target as determined by the monetary policy rule. Therefore, the inclusion of the lagged term spread in the PLM of inflation reduces the importance of this source of inflation bias. Second, the correlation between the coefficients of the first two lags of the variable decreases for all those PLM where the lagged term spread is not included reducing the importance of the redundancy issues as discussed below. Third, the introduction of the lagged term spread in the PLM of inflation, consumption and the rental rate of capital results in an overall improvement on the ability of the AL model with term spread to reproduce U.S. business cycle features as shown below. Finally, as shown in the next subsection, term spread innovations become an important source of fluctuations under AL.

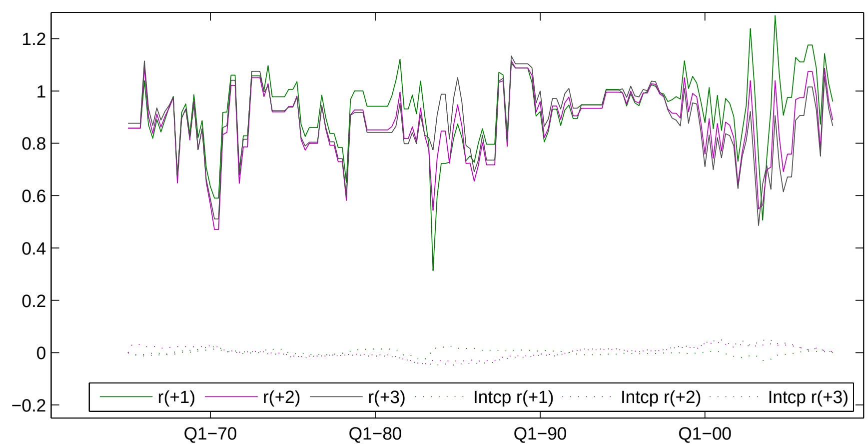


Figure 1D. PLM coefficients of future short term interest rates

Figure 1D shows the evolution of the coefficients associated with the PLM of the 1-period, 2-period and 3-period ahead forecasts of the short-term nominal interest rate (i.e. $E_t r_{t+1}$, $E_t r_{t+2}$ and $E_t r_{t+3}$, respectively). Notice that the coefficient of the lagged interest rate fluctuates around one in all cases. Moreover, the size of these fluctuations becomes smaller as the forecasting horizon increases, which means that the informational content of the lagged interest rate is larger for short-term forecasting than for long-term forecasting as expected.

Volatility of learning coefficients

As discussed above, a sound criteria for disciplining expectations is to select the PLM by disregarding forecasting models characterized by excessive volatility of the learning coefficients as a way of avoiding the overestimation of the importance of learning in explaining actual business cycles. Table 2 shows the standard deviations of the different learning coefficients across PLM, AL models (SlW versus SlWTS) and sample periods. When comparing alternative sample periods, we split the full sample period in two sub-samples, where the first (pre-1984) sub-sample is characterized by high volatility of most aggregate variables and the second (post-1984) covers the Great Moderation period (1984-2007) characterized by low volatility.[16] The comparison of the pre-1984 and the Great Moderation periods provides a good environment for conducting a field experiment useful for disciplining AL expectations and discriminating across alternative small forecasting models. In particular, one should expect that the learning coefficients should be more stable in low volatility regimes as the Great Moderation period than in high volatility regimes as the pre-1984 period.[17]

Comparing the learning coefficients of each PLM across models, we observe that the forecasting models with term spread (i.e. the ones associated with the SlWTS) result in more stable learning coefficients than the corresponding ones associated with the SlW. This result holds for the two sub-samples and the whole sample as well. Moreover, the two forecasting models exhibit, in general, a lower volatility of learning coefficients during the Great Moderation period than during the pre-1984 period. Interestingly, the real wage growth rate is more volatile during the Great Moderation than in the previous period. Accordingly, the three coefficients of the PLM of the real wage associated with the SlWTS model exhibit higher volatility after 1984 than before. However, only the intercept coefficient of the PLM of the real wage associated with the SlW model shows this feature.

Table 2 also shows the statistic *corr(*$\beta_{1,y,t-1}$*,*$\beta_{j,y,t-1}$*)* denoting the correlation between the coefficient associated with the first lag of the variable y, $\beta_{1,y,t-1}$, and the coefficient associated with its second lag, $\beta_{2,y,t-1}$, or alternatively the coefficient associated with the lagged term spread, $\beta_{sp,y,t-1}$, depending on whether the lagged term spread enters or not in the PLM. A value of *corr(*$\beta_{1,y,t-1}$*,*$\beta_{2,y,t-1}$*)* close to one warns us about the possibility of a redundancy issue and motivates the exercise of exploring the effects of substituting the second lag of

[16]Stock and Watson (2002) and Smets and Wouters (2007), among others, mark 1984:1 as the start of the Great Moderation.

[17]More precisely, the Great Moderation features a lower-than-average volatility for all the observable variables considered in this paper, but for the real wage growth rate.

the forward-looking variable in the PLM for the lagged term spread. Interestingly, this correlation statistic gets smaller even in those PLM where the term spread does not enter. In sum, the introduction of the term spread instead of the second lag of the corresponding forward-looking variable not only avoids the use of redundant information from the second lag of the variable in a few PLM but also helps to overcome the redundancy issue in the rest of the PLM.

3.4 Impulse response function analysis

As shown by Slobodyan and Wouters (2012a, 2012b), the transmission of structural shocks are crucially determined by the way agents form their expectations. Therefore, it is important to show how the impulse response functions (IRF) shift over time driven by changes in the updating belief coefficients. Our IRF analysis is divided in two parts. First, we analyze the estimated time-varying responses of a selected group of real (i.e. output and consumption) and nominal (inflation and short-term interest rate) variables to term spread innovations. As emphasized above, term spread innovations are the only forward-looking components of term structure under AL. Thus, the responses of output and inflation to term spread innovations illustrate how term structure innovations anticipate movements in these variables. Second, we analyze the responses of output and inflation to a selected group of shocks (i.e. productivity, risk premium, monetary policy and wage markup shocks). By comparing these IRF with the ones reported in Slobodyan and Wouters (2012b, p. 87), we illustrate how the introduction of the term spread in the small forecasting model changes the transmission of structural shocks.[18]

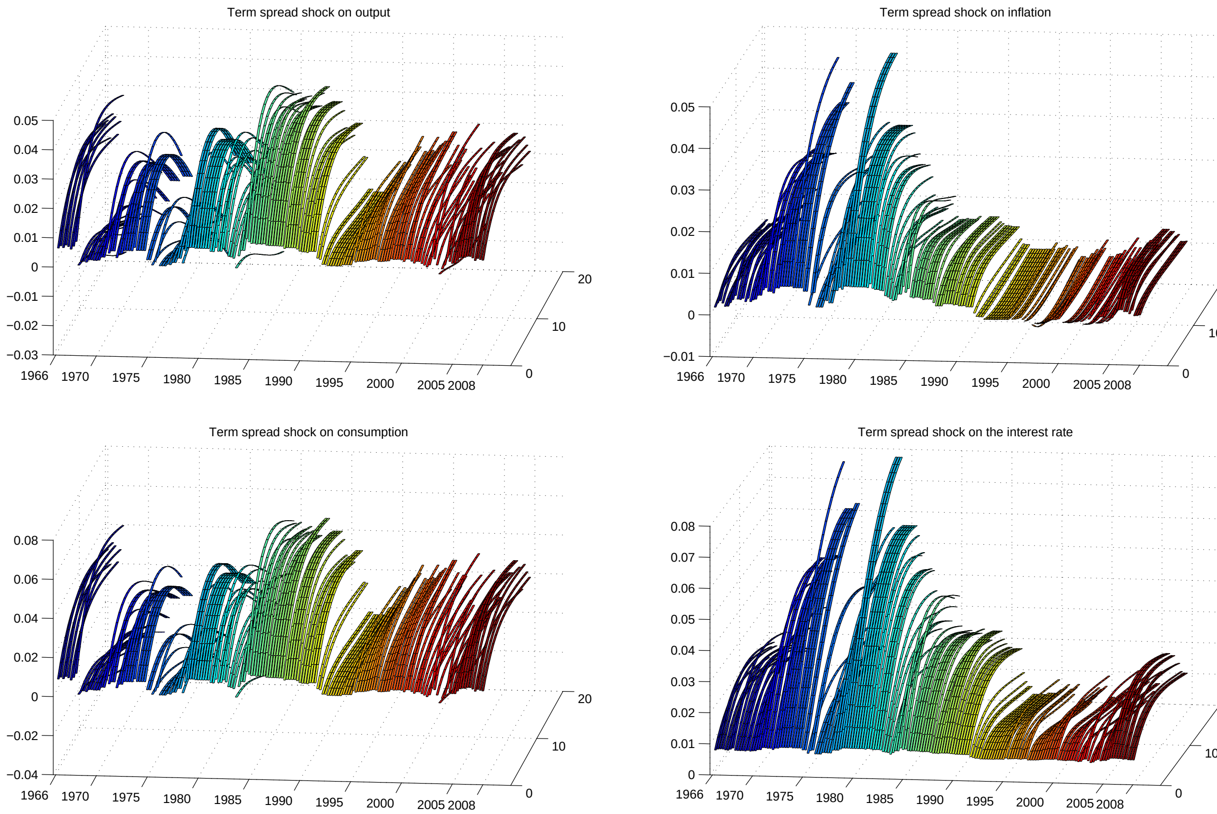
[18]Following Slobodyan and Wouters (2012b), the IRF are computed using the fixed belief coefficients obtained using the information available at each point in time, but then ignoring the updating of these beliefs driven by the shock. Therefore, these IRF might underestimate the size and persistence of actual responses.

IRF to a term premium shock

Recall that the term premium shock capture the forward-looking behavior of the term spread (see equations (1) and (2)) under AL. Figure 2 shows the time-varying IRF of output, consumption, inflation and nominal interest rate to a term premium innovation. These responses are rather sizable when compare to the responses associated with the alternative shocks below. This finding suggests that term premium shocks are an important source of aggregate fluctuations under AL. Moreover, we observe that the shapes and the timing of shifts associated with the IRF of the two real variables (output and consumption) are almost identical. Thus, the shifts of the IRF are mostly concentrated in the 1960s and 1970s, and later they only occur around 1984 and 1991. These findings stand in sharp contrast to the lack of a response of aggregate variables to term premium shocks under RE. Most papers in the literature aiming to link RE DSGE models with the term structure of interest rates (see for instance, Hördahl, Tristani and Vestin, 2006; and De Graeve, Emiris and Wouters, 2009) assume a sort of dichotomy where the DSGE model is solved first and independently from the term structure of interest rates. In contrast, the introduction of AL extended with term structure information allows for a feedback from the term structure to the macroeconomy that is missing under RE.

Regarding the IRF of the two nominal variables, we also find that the shapes and the timing of shifts are rather similar, but the initial impact and the persistence of the interest rate responses are slightly larger than the ones associated with inflation. Moreover, the persistence of inflation and interest rate responses is much larger before 1984 than afterward, which is consistent with the estimated perceived inflation persistence pattern shown in Figure 1A above.

Figure 2. Term structure shocks

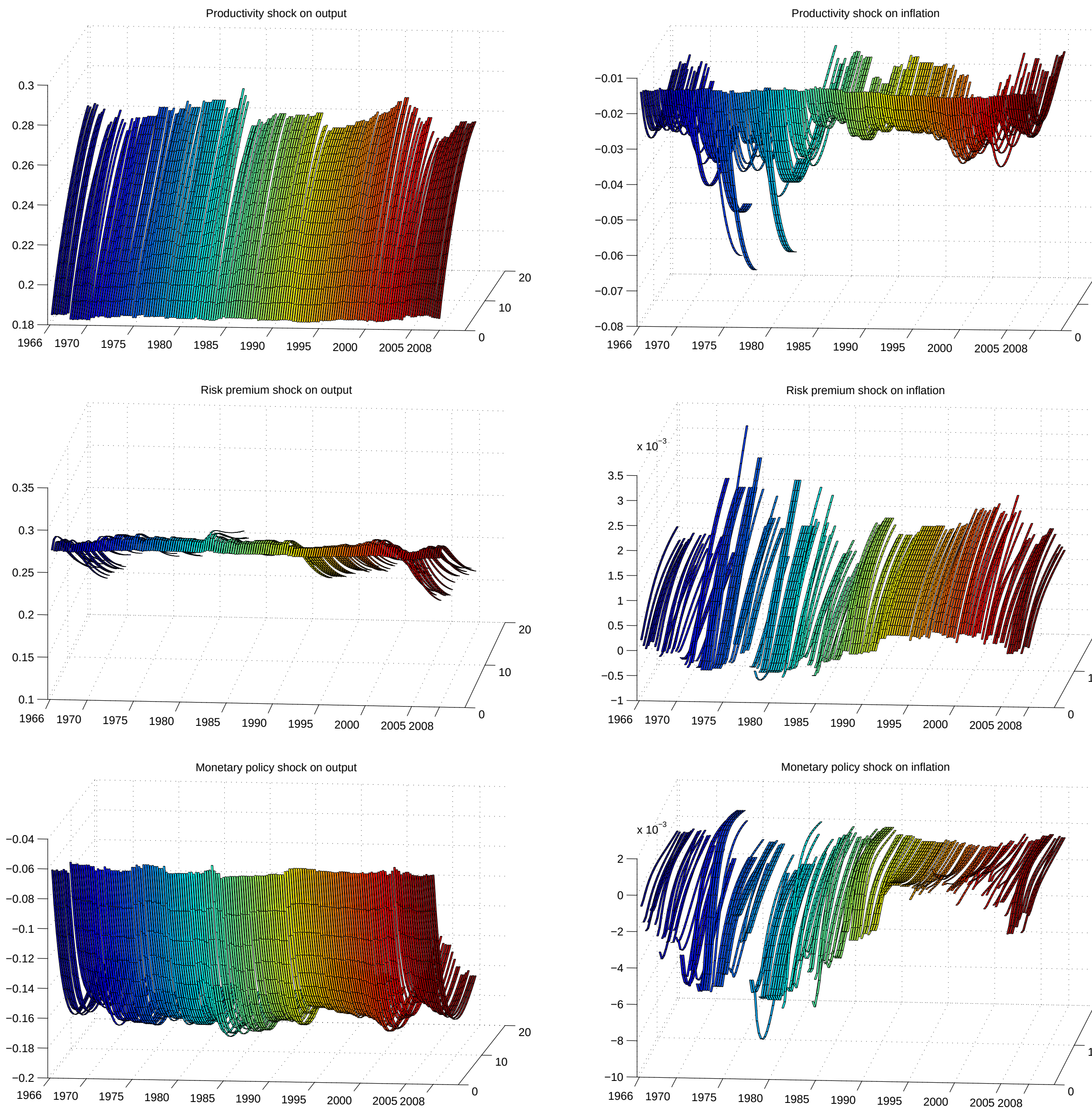


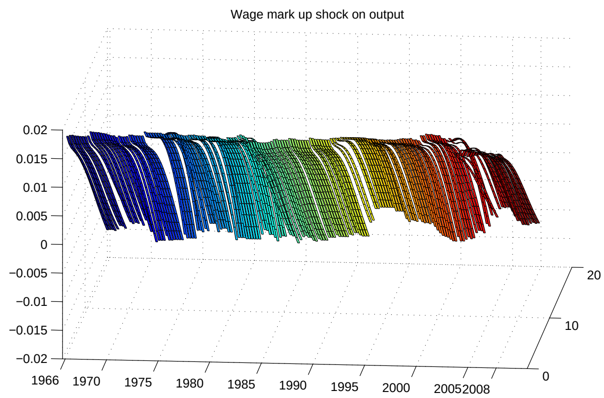
IRF of output and inflation to alternative shocks

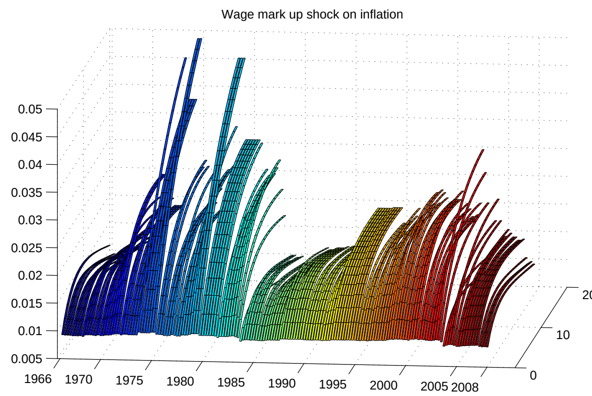
Figure 3 shows the IRF of output and inflation to productivity, risk premium, monetary policy and wage markup shocks. The shape and timing of shifts associated with the responses of inflation to these shocks derived from the SlWTS model are similar to the ones shown in Slobodyan and Wouters (2012b) based on the SlW model. However, the large responses taken place before 1984 are much more extreme in the SlW model than in the SlWTS. This result is consistent with the finding that the introduction of the term spread in the small forecasting model reduces the size of fluctuations. Moreover, as with the response of inflation to a term spread shock, the size of inflation responses to the other structural shocks decreases substantially after 1984 when inflation was perceived as less persistent.

Regarding the responses of output, we observe a few shifts but they are not so large in relative terms as those observe for inflation responses. Moreover, output responses do not display different patterns before and after 1984 as occurs with inflation responses, with the exception of the response of output to risk premium shocks which becomes less persistence since 1990.

Figure 3. Shocks on output and inflation







4 Model evaluation

This section compares the performance of the four alternative DSGE models across a measure of in-sample fit based on the root mean squared one-period-ahead forecast error (RMSE) and a few second moment statistics. More precise, we analyze the standard deviation, the correlations with output growth, the correlations with inflation and the first-order autocorrelation of the main macroeconomic variables.

Table 3 shows the RMSE associated with each observable variable across the four alternative models. Clearly, the two AL models outperform the RE models for all variables. Moreover, the SlWTS and the SlW perform rather similar. Thus, the SlWTS performs slightly better than the SlW for inflation and the growth rates of output and consumption, whereas the opposite is true for the rest of the variables.

Table 4 shows second moment statistics obtained from the four estimated DSGE models. Regarding the standard deviations, we observe that the two AL models (SlW and SlWTS) reproduce much better the size of the fluctuations of the growth rates of investment and real wages and the long-term nominal interest rates than the two RE models (SW and SW with term spread). Moreover, the SlWTS model performs closely to the two RE models in reproducing the volatility of output growth, and to a lesser extent the volatility of the short-term interest rate and inflation, and it outperforms the other three models in reproducing the volatility of consumption growth.

Regarding the contemporaneous co-movement with output growth, the SlWTS does a reasonable job in reproducing the actual correlations: it is as good as the SW model for most correlations (with the exception of the correlation of output growth with inflation and investment growth rate) and it outperforms the SlW for all output growth correlations, but the correlations of output growth with investment growth and hours worked for which the two models do a similar job. Regarding the contemporaneous co-movement with inflation, the SlWTS is close to the SW model in reproducing the correlations of inflation with investment, real wage and output growth rates as well as the correlation between inflation and the nominal interest rate. Moreover, it outperforms the SlW for all inflation correlations.

Finally, the analysis of the first-order autocorrelation statistics shows that the SlW model generates too much persistence in the growth rates of output and consumption when compared to those observed in actual data and those generated by the other three models (SW, SW with term spread and SlWTS). For the other variables, the four models do in general a good job in matching their observed persistence. Nevertheless, there a few noticeable differences when comparing the two AL models. Thus, the SlWTS model does a better job than the SlW model when characterizing the persistence of consumption and output growth rates, inflation and the 1-year constant maturity rate. However, the opposite occurs for the real wage growth rate.

5 Robustness analysis

This section studies the robustness of estimation results of the SlWTS model using alternative specifications for the PLM of the forward-looking variables across three dimensions: in-sample fit, parameter estimates and second-moment statistics.

Table 5 shows the in-sample fit measured by the RMSE across the four alternative PLM formulations analyzed. The row labeled "Baseline" shows the RMSE for all observables variables obtained from the baseline model discussed in the previous section. That is, it includes the lagged term-spread in the PLM of inflation, the rental rate of capital and consumption instead of the second lag of the corresponding variable. The row labeled "I" shows the corresponding RMSE for the model that includes the lagged term-spread in the PLM of inflation, rental rate of capital and wages instead of the second lag of the corresponding variable. The row labeled "II" shows the corresponding RMSE for the model that includes the lagged term-spread in the PLM of inflation, rental rate of capital, consumption and wages instead of the second lag of the corresponding variable. Finally, the row labeled "III" reports the RMSE for the model that includes the lagged term-spread instead of the second lag of the corresponding variable in the PLM of inflation, rental rate of capital and consumption and where the PLM of hours worked and wages are assumed to follow an AR(1).[19] We observe that the RMSE is rather robust across the four PLM specifications. Moreover, the baseline specification is the one showing the lowest RMSE for half of the eight observable variables (i.e. for output growth, inflation, short-term nominal interest rate and 1-year yield).

Table 6 shows the estimation results for four alternative specifications of the PLM studied. We observe that parameter estimates are rather robust across the alternative specifications of the PLM studied. Nevertheless, there a few noticeable differences. First, the estimated mode of the learning coefficient, ρ, lies in the interval (0.92, 0.99) across the alternative PLM studied. Second, the estimates of a few parameters measuring endogenous persistence are somewhat sensitive depending on the specification used. Thus, the estimates of the habit formation parameter, h, and the wage stickiness parameter, ξ_w, both lie in similar intervals ([0.69,0.79] and [0.69,0.87], respectively), whereas the estimates of the price indexation parameter, ι_p, move inside the interval [0.45,0.65] and the elasticity of the cost of adjusting capital, φ, in the interval (4.27,6.82). Third, the estimates of the parameters characterizing the exogenous persistence of prices and wages, ρ_p and ρ_w, also exhibit some variability across

[19] We explored many other specifications as well, but we do not show all of them to save space and keep the discussion tractable. Recall that we focused on those PLM specifications that satisfy the selection criteria discussed above regarding PLM coefficient fluctuations. Estimation results from these additional selected specifications can be obtained from the authors upon request.

alternative PLM specifications, but they are much lower in most cases than the corresponding estimates of the RE models. These results reinforce Slobodyan and Wouters' finding that AL models substantially decrease the importance of price and wage exogenous persistence for reproducing the persistence of actual macroeconomic variables.

Table 7 shows the second moments for the alternatives specifications of the PLM analyzed. In general, we observe that the second-moment statistics are rather robust across PLM formulations. Nevertheless, it is worthwhile to highlight a few differences. Regarding the standard deviations, the baseline PLM formulation does a much better job in reproducing the size of fluctuations of consumption, output and real wage growth rates than the other PLM formulations. For the other variables, the baseline PLM formulation falls in general shorter than the other formulations when reproducing the observed standard deviations. The baseline formulation also performs much better than the alternative PLM specifications when reproducing the comovement between inflation and the short-term interest rate. Finally, the baseline formulation does a much better job in reproducing the persistence of investment growth than the other specifications, but the opposite occurs for the first autocorrelation of the real wage growth rate.

6 Conclusions

In this paper we have extended the adaptive learning model of Slobodyan and Wouters (2012b) by introducing the term structure of interest rates. While retaining the feature of AL based on small forecasting models, our extension allows the term spread of interest rates to help agents forecasting a few forward-looking variables such as inflation, consumption and the rental rate of capital. The introduction of term structure information in small forecasting models results in more stable perceived law of motions for the forward-looking variables than those obtained when ignoring the term structure. This finding is important because adaptive learning schemes are often criticized for being arbitrary (see, for instance, Adam and Marcet

(2011) and references therein), and potentially amplifying the size of fluctuations in an ad hoc manner. However, the inclusion of the term structure is useful to disciplining expectations formation by restricting in general learning coefficient dynamics and, in particular, the size of fluctuations induced by the learning process.

Our estimation results show that the inclusion of the term spread in the AL model reinforces the different features found by Slobodyan and Wouters (2012a, 2012b) between the rational expectations (RE) and AL versions of the estimated medium-scale DSGE model, such as the shifts in beliefs about inflation persistence explaining the hump-shaped pattern of U.S. inflation in the last fifty years as well as the lower estimates for the persistence of the exogenous shocks driving price and wage dynamics. Moreover, the AL model extended with the term structure of interest rates results in a strong feedback from the term structure to the macroeconomy that is missing in RE DSGE models. Furthermore, the empirical results show that our extended model with term structure does a better job than the RE model and the AL model without term structure when reproducing most U.S. business cycle features.

A general shortcoming of estimated AL versions of DSGE models is the use of only final revised data when in reality agents have only access to real-time data when updating their expectations over time. In an ongoing paper, we are investigating an extended version of our model where the elements of AL and the term spread are combined with the presence of data revisions affecting agent decisions as modeled in Casares and Vázquez (2012) under rational expectations.

Table 1. Panel A: Priors and estimated posteriors of the structural parameters of the four alternative models

	Priors			Posteriors											
				SW			SW spread			SlW			SlWTS		
	Distr	Mean	Std D.	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
φ	Normal	4.00	1.50	5.96	4.21	7.63	6.06	4.33	7.89	3.34	1.88	3.87	6.69	5.71	7.92
h	Beta	0.70	0.10	0.79	0.74	0.86	0.77	0.69	0.83	0.68	0.63	0.75	0.61	0.54	0.66
σ_c	Normal	1.50	0.37	1.22	1.04	1.34	1.22	1.02	1.45	1.53	1.19	1.63	1.64	1.41	1.84
σ_l	Normal	2.00	0.75	1.50	0.64	2.32	1.82	0.89	2.53	1.74	1.02	2.60	2.22	1.79	2.68
ξ_p	Beta	0.50	0.10	0.70	0.62	0.80	0.74	0.57	0.88	0.64	0.59	0.69	0.71	0.67	0.73
ξ_w	Beta	0.50	0.10	0.71	0.62	0.78	0.86	0.81	0.91	0.82	0.76	0.85	0.85	0.81	0.89
ι_w	Beta	0.50	0.15	0.51	0.30	0.72	0.24	0.10	0.42	0.18	0.07	0.26	0.53	0.40	0.65
ι_p	Beta	0.50	0.15	0.25	0.10	0.38	0.27	0.10	0.39	0.27	0.11	0.39	0.56	0.44	0.68
ψ	Beta	0.50	0.15	0.55	0.36	0.72	0.53	0.34	0.74	0.50	0.31	0.71	0.14	0.09	0.18
Φ	Normal	1.25	0.12	1.62	1.48	1.73	1.57	1.41	1.68	1.58	1.45	1.73	1.55	1.49	1.61
r_π	Normal	1.50	0.25	1.98	1.71	2.25	1.80	1.15	2.09	1.74	1.38	2.04	1.74	1.52	2.02
ρ_r	Beta	0.75	0.10	0.84	0.80	0.87	0.87	0.83	0.89	0.88	0.85	0.91	0.89	0.87	0.91
r_y	Normal	0.12	0.05	0.10	0.05	0.13	0.11	0.06	0.15	0.13	0.07	0.18	0.12	0.08	0.16
$r_{\Delta y}$	Normal	0.12	0.05	0.15	0.12	0.18	0.17	0.14	0.20	0.13	0.10	0.16	0.14	0.12	0.16
π	Gamma	0.62	0.10	0.67	0.51	0.82	0.71	0.54	0.88	0.63	0.53	0.74	0.83	0.73	0.95
β	Gamma	0.25	0.10	0.2	0.08	0.28	0.20	0.10	0.30	0.18	0.09	0.27	0.15	0.06	0.27
l	Normal	0.00	2.00	1.05	-0.5	2.14	0.59	-0.72	2.03	1.10	-0.76	1.96	-0.44	-1.04	0.48
γ	Normal	0.40	0.10	0.40	0.38	0.42	0.41	0.39	0.42	0.40	0.36	0.43	0.39	0.38	0.40
$\bar{r}^{\{4\}}$	Normal	1.00	0.50	-	-	-	1.38	1.19	1.58				1.28	1.08	1.45
α	Normal	0.30	0.05	0.19	0.16	0.22	0.18	0.15	0.20	0.17	0.13	0.19	0.17	0.14	0.20
ρ	Beta	0.50	0.28	-	-	-	-	-	-	0.97	0.96	0.98	0.82	0.80	0.84

Table 1. Panel B: Priors and estimated posteriors of the structural parameters of the four alternative models

	Priors			Posterior											
				SW			SW spread			SlW			SlWTS		
	Distr	Mean	Std D.	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
σ_a	Invgamma	0.10	2.00	0.44	0.39	0.48	0.45	0.40	0.49	0.46	0.41	0.49	0.47	0.44	0.51
σ_b	Invgamma	0.10	2.00	0.25	0.21	0.29	0.24	0.20	0.28	0.15	0.14	0.20	0.25	0.23	0.27
σ_g	Invgamma	0.10	2.00	0.52	0.47	0.57	0.51	0.48	0.56	0.50	0.46	0.56	0.52	0.48	0.56
σ_i	Invgamma	0.10	2.00	0.44	0.37	0.52	0.44	0.36	0.52	0.46	0.42	0.52	0.36	0.30	0.40
σ_R	Invgamma	0.10	2.00	0.22	0.20	0.24	0.22	0.20	0.24	0.21	0.19	0.23	0.21	0.20	0.22
σ_p	Invgamma	0.10	2.00	0.14	0.11	0.17	0.17	0.14	0.19	0.15	0.13	0.16	0.13	0.10	0.15
σ_w	Invgamma	0.10	2.00	0.23	0.20	0.27	0.30	0.27	0.33	0.21	0.18	0.22	0.34	0.31	0.39
$\sigma^{\{4\}}$	Invgamma	0.10	2.00	-	-	-	0.10	0.09	0.11				0.17	0.16	0.19
ρ_a	Beta	0.50	0.20	0.93	0.89	0.96	0.93	0.90	0.97	0.98	0.98	0.99	0.98	0.98	0.99
ρ_b	Beta	0.50	0.20	0.17	0.04	0.26	0.26	0.15	0.40	0.43	0.29	0.55	0.21	0.13	0.27
ρ_g	Beta	0.50	0.20	0.98	0.97	0.99	0.98	0.97	0.99	0.96	0.95	0.98	0.98	0.97	0.99
ρ_i	Beta	0.50	0.20	0.70	0.59	0.79	0.69	0.59	0.79	0.50	0.37	0.57	0.49	0.44	0.55
ρ_R	Beta	0.50	0.20	0.08	0.01	0.14	0.06	0.01	0.11	0.09	0.01	0.16	0.11	0.07	0.15
ρ_p	Beta	0.50	0.20	0.84	0.74	0.95	0.85	0.75	0.99	0.32	0.06	0.57	0.37	0.21	0.50
ρ_w	Beta	0.50	0.20	0.97	0.96	0.99	0.95	0.93	0.98	0.54	0.32	0.80	0.25	0.11	0.38
$\rho^{\{4\}}$	Beta	0.50	0.20	-	-	-	0.84	0.77	0.91	-	-	-	0.96	0.94	0.98
μ_p	Beta	0.50	0.20	0.67	0.49	0.85	0.75	0.61	0.90	0.47	0.29	0.67	0.53	0.40	0.63
μ_w	Beta	0.50	0.20	0.87	0.82	0.94	0.98	0.86	0.99	0.43	0.11	0.70	0.40	0.30	0.49
ρ_{ga}	Beta	0.50	0.20	0.51	0.37	0.65	0.52	0.39	0.67	0.17	0.13	0.19	0.52	0.44	0.20

Table 2. Standard deviations of learning coefficients across learning models and sample periods

	SlW			SlWTS		
	Pre-1984	Post-1984	Full sample	Pre-1984	Post-1984	Full sample
PLM of consumption						
$\theta_{c,t-1}$	0.052	0.037	0.058	0.026	0.017	0.037
$\beta_{1,c,t-1}$	0.046	0.036	0.065	0.005	0.007	0.007
$\beta_{2,c,t-1}$	0.046	0.035	0.067	-	-	-
$\beta_{sp,c,t-1}$	-	-	-	0.025	0.023	0.025
corr($\beta_{1,c,t-1}$,$\beta_{j,c,t-1}$)	-	-	-0.997	-	-	0.316
PLM of investment						
$\theta_{i,t-1}$	0.134	0.156	0.146	0.064	0.053	0.063
$\beta_{1,i,t-1}$	0.048	0.037	0.054	0.052	0.0038	0.047
$\beta_{2,i,t-1}$	0.047	0.038	0.054	0.049	0.040	0.046
corr($\beta_{1,i,t-1}$,$\beta_{2,i,t-1}$)	-	-	-0.991	-	-	-0.980
PLM of inflation						
$\theta_{\pi,t-1}$	0.038	0.034	0.051	0.010	0.006	0.012
$\beta_{1,\pi,t-1}$	0.083	0.048	0.105	0.069	0.038	0.067
$\beta_{2,\pi,t-1}$	0.11	0.034	0.084	-	-	-
$\beta_{sp,\pi,t-1}$	-	-	-	0.014	0.016	0.020
corr($\beta_{1,\pi,t-1}$,$\beta_{j,\pi,t-1}$)	-	-	-0.590	-	-	-0.530
PLM of capital return						
$\theta_{r^k,t-1}$	0.055	0.030	0.044	0.010	0.010	0.029
$\beta_{1,r^k,t-1}$	0.051	0.049	0.050	0.008	0.007	0.008
$\beta_{2,r^k,t-1}$	0.049	0.049	0.049	-	-	-
$\beta_{spread,r^k,t-1}$	-	-	-	0.015	0.014	0.014
corr($\beta_{1,r^k,t-1}$,$\beta_{j,r^k,t-1}$)	-	-	-0.992	-	-	-0.391

Table 2. Standard deviations of learning coefficients across learning models and sample periods

(*continued*)

	SlW			SlWTS		
	Pre-1984	Post-1984	Full sample	Pre-1984	Post-1984	Full sample
PLM of hours worked						
$\theta_{l,t-1}$	0.058	0.048	0.054	0.031	0.026	0.029
$\beta_{1,l,t-1}$	0.065	0.049	0.060	0.055	0.075	0.067
$\beta_{2,l,t-1}$	0.055	0.057	0.061	0.062	0.069	0.066
corr($\beta_{1,l,t-1}$,$\beta_{2,l,t-1}$)	-	-	-0.974	-	-	-0.915
PLM of Tobin's q						
$\theta_{q,t-1}$	0.079	0.087	0.096	0.099	0.062	0.141
$\beta_{1,q,t-1}$	0.050	0.026	0.054	0.036	0.040	0.039
$\beta_{2,q,t-1}$	0.014	0.017	0.019	0.035	0.028	0.037
corr($\beta_{1,q,t-1}$,$\beta_{2,q,t-1}$)	-	-	-0.594	-	-	-0.539
PLM of real wage						
$\theta_{w,t-1}$	0.045	0.060	0.06	0.020	0.036	0.032
$\beta_{1,w,t-1}$	0.049	0.047	0.051	0.025	0.081	0.064
$\beta_{1,w,t-1}$	0.052	0.050	0.055	0.026	0.084	0.067
corr($\beta_{1,w,t-1}$,$\beta_{2,w,t-1}$)	-	-	-0.996	-	-	-0.901
PLM of interest rate						
$\theta_{1,t-1}$	-	-	-	0.011	0.010	0.010
$\beta_{1,t-1}$	-	-	-	0.140	0.120	0.135
$\theta_{2,t-1}$	-	-	-	0.018	0.020	0.020
$\beta_{2,t-1}$	-	-	-	0.129	0.120	0.129
$\theta_{3,t-1}$	-	-	-	0.014	0.020	0.023
$\beta_{3,t-1}$	-	-	-	0.119	0.130	0.128

Table 3. RMSE of the four alternative models

Standard deviation	Δc	Δinv	Δw	Δy	l	π	r	$r^{\{4\}}$
SW	0.72	1.84	0.64	0.76	1.17	0.76	1.35	-
SW spread	0.74	1.83	0.62	0.75	0.89	0.78	1.42	1.40
SlW	0.71	1.78	0.54	0.74	0.54	0.28	0.23	-
SlWTS	0.65	1.84	0.59	0.73	0.58	0.27	0.24	0.21

Table 4. Actual and synthetic second moments across models

Standard deviation	Δc	Δinv	Δw	Δy	l	π	r	$r^{\{4\}}$
Data	0.67	2.20	0.55	0.83	2.77	0.59	0.82	0.72
SW	0.49	5.7	0.32	0.84	11.6	0.45	0.52	-
SW spread	0.49	5.34	0.38	0.77	5.85	0.23	0.30	0.26
SlW	0.88	2.52	0.49	1.11	2.30	0.31	0.52	-
SlWTS	0.75	1.59	0.57	0.82	1.52	0.28	0.41	0.54
Correlation with Δy	Δc	Δinv	Δw	Δy	l	π	r	$r^{\{4\}}$
Data	0.64	0.65	0.06	1	0.09	-0.28	-0.20	-0.12
SW	0.59	0.61	0.26	1	0.08	-0.21	-0.14	-
SW spread	0.57	0.59	0.13	1	0.10	-0.14	-0.12	-0.06
SlW	0.84	0.80	0.25	1	0.14	-0.07	-0.08	-
SlWTS	0.67	0.51	0.14	1	0.03	-0.18	-0.25	-0.18
Correlation with π	Δc	Δinv	Δw	Δy	l	π	r	$r^{\{4\}}$
Data	-0.40	-0.17	-0.21	-0.28	-0.46	1	0.60	0.56
SW	-0.22	-0.14	-0.18	-0.21	-0.54	1	0.70	-
SW spread	-0.11	-0.11	-0.21	-0.14	-0.27	1	0.50	0.55
SlW	-0.04	-0.01	-0.14	-0.08	0.40	1	0.35	-
SlWTS	-0.11	-0.22	-0.16	-0.18	-0.10	1	0.65	0.23
Autocorrelation	Δc	Δinv	Δw	Δy	l	π	r	$r^{\{4\}}$
Data	0.17	0.52	0.15	0.22	0.96	0.86	0.94	0.95
SW	0.29	0.59	0.18	0.24	0.97	0.87	0.92	-
SW spread	0.28	0.58	0.09	0.21	0.96	0.77	0.88	0.89
SlW	0.46	0.51	0.15	0.37	0.94	0.67	0.88	-
SlWTS	0.23	0.53	-0.20	0.15	0.92	0.78	0.85	0.78

Table 5. RMSE of the alternative PLM

Standard deviation	Δc	Δinv	Δw	Δy	l	π	r	$r^{\{4\}}$
Baseline	0.65	1.84	0.59	0.73	0.58	0.27	0.24	0.21
I	0.63	1.91	0.62	0.77	0.58	0.28	0.24	0.22
II	0.63	1.88	0.64	0.77	0.57	0.27	0.24	0.23
III	0.65	1.83	0.57	0.74	0.58	0.28	0.24	0.21

Notes to Table 5: The row labeled "Baseline" shows the RMSE for all observables variables obtained from the baseline model discussed in the previous section. That is, it includes the lagged term-spread in the PLM of inflation, the rental rate of capital and consumption instead of the second lag of the corresponding variable. The row labeled "I" shows the corresponding RMSE for the model that includes the lagged term-spread in the PLM of inflation, rental rate of capital and wages instead of the second lag of the corresponding variable. The row labeled "II" shows the corresponding RMSE for the model that includes the lagged term-spread in the PLM of inflation, rental rate of capital, consumption and wages instead of the second lag of the corresponding variable. Finally, the row labeled "III" reports the RMSE for the model that includes the lagged term-spread instead of the second lag of the corresponding variable in the PLM of inflation, rental rate of capital and consumption and where the PLM of hours worked and wages are assumed to follow an AR(1).

Table 6. Panel A: Robustness analysis of alternative PLM: structural parameters

	Baseline		I		II		III	
Marg. Lik.	-1000		-1000		-999		-954	
	Mode	Std D.	Mode	Std D.	Mode	Std D.	Mode	Std D.
φ	6.54	0.96	4.36	0.05	4.27	0.07	6.82	1.36
h	0.69	0.12	0.79	0.03	0.77	0.05	0.69	0.07
σ_c	1.11	0.68	0.94	0.03	1.01	0.04	1.17	0.28
σ_l	2.52	0.44	2.09	0.04	2.15	0.03	1.72	0.90
ξ_p	0.67	0.06	0.69	0.04	0.70	0.04	0.72	0.05
ξ_w	0.87	0.03	0.69	0.03	0.68	0.04	0.72	0.11
ι_w	0.47	0.36	0.41	0.04	0.40	0.04	0.45	0.11
ι_p	0.45	0.22	0.63	0.04	0.60	0.05	0.65	0.35
ψ	0.30	0.06	0.23	0.05	0.22	0.05	0.08	0.10
Φ	1.55	0.18	1.53	0.03	1.55	0.05	1.56	0.12
r_π	1.86	0.22	1.68	0.04	1.66	0.06	1.51	0.42
ρ_r	0.86	0.03	0.84	0.03	0.85	0.06	0.86	0.02
r_y	0.11	0.05	0.11	0.04	0.10	0.01	0.14	0.04
$r_{\Delta y}$	0.10	0.01	0.12	0.03	0.14	0.05	0.12	0.02
π	0.71	0.12	0.86	0.04	0.88	0.04	0.68	0.08
β	0.23	0.18	0.43	0.03	0.47	0.05	0.19	0.13
l	-1.65	0.92	0.24	0.05	0.61	0.07	-0.53	0.79
γ	0.37	0.02	0.40	0.01	0.41	0.07	0.39	0.2
$\bar{r}^{\{4\}}$	1.16	0.32	1.81	0.03	1.77	0.06	1.01	0.51
α	0.11	0.04	0.18	0.03	0.18	0.04	0.17	0.02
ρ	0.92	0.04	0.98	0.04	0.99	0.01	0.92	0.06

Table 6. Panel B: Robustness analysis of alternative PLM: shock process parameters

	Baseline		I		II		III	
	Mode	Std D.	Mode	Std D.	Mode	Std D.	Mode	Std D.
σ_a	0.48	0.04	0.49	0.3	0.48	0.05	0.46	0.02
σ_b	0.27	0.05	0.30	0.04	0.29	0.05	0.28	0.02
σ_g	0.52	0.03	0.51	0.04	0.60	0.05	0.50	0.03
σ_i	0.38	0.05	0.45	0.04	0.44	0.05	0.32	0.07
σ_R	0.21	0.01	0.25	0.3	0.24	0.05	0.21	0.01
σ_p	0.12	0.04	0.15	0.04	0.15	0.05	0.12	0.05
σ_w	0.35	0.12	0.45	0.04	0.43	0.04	0.32	0.04
$\sigma^{\{4\}}$	0.18	0.01	0.21	0.03	0.20	0.05	0.18	0.01
ρ_a	0.94	0.06	0.96	0.01	0.97	0.04	0.91	0.04
ρ_b	0.19	0.29	0.28	0.04	0.28	0.04	0.19	0.11
ρ_g	0.98	0.01	0.95	0.02	0.95	0.02	0.99	0.1
ρ_i	0.41	0.09	0.34	0.04	0.36	0.0	0.45	0.17
ρ_R	0.18	0.10	0.23	0.04	0.24	0.03	0.18	0.04
ρ_p	0.53	0.15	0.44	0.03	0.42	0.05	0.27	0.79
ρ_w	0.37	0.23	0.67	0.03	0.65	0.04	0.89	0.09
$\rho^{\{4\}}$	0.90	0.03	0.92	0.03	0.92	0.03	0.97	0.01
μ_p	0.75	0.13	0.66	0.03	0.66	0.05	0.56	0.35
μ_w	0.51	0.20	0.55	0.03	0.57	0.04	0.57	0.23
ρ_{ga}	0.42	0.28	0.60	0.05	0.58	0.06	0.57	0.06

Table 7. Robustness analysis of alternative PLM: second moment statistics

Standard deviation	Δc	Δinv	Δw	Δy	l	π	r	$r^{\{4\}}$
U.S. data	0.67	2.20	0.55	0.83	2.77	0.59	0.82	0.72
Baseline	0.75	1.59	0.57	0.82	1.52	0.28	0.41	0.54
I	0.82	1.87	0.96	1.06	1.59	0.41	0.46	0.63
II	0.81	1.80	0.87	1.06	1.50	0.37	0.43	0.59
III	0.81	1.58	1.01	0.89	2.91	1.04	0.91	1.10
Correlation with Δy	Δc	Δinv	Δw	Δy	l	π	r	$r^{\{4\}}$
U.S. data	0.64	0.65	0.06	1	0.09	-0.28	-0.20	-0.12
Baseline	0.67	0.51	0.14	1	0.03	-0.18	-0.25	-0.18
I	0.75	0.65	0.28	1	0.11	-0.17	-0.22	-0.13
II	0.70	0.60	0.30	1	0.13	-0.12	-0.19	-0.11
III	0.68	0.57	0.28	1	0.08	-0.20	-0.23	-0.17
Correlation with π	Δc	Δinv	Δw	Δy	l	π	r	$r^{\{4\}}$
U.S. data	-0.40	-0.17	-0.21	-0.28	-0.46	1	0.60	0.56
Baseline	-0.11	-0.22	-0.16	-0.18	-0.1	1	0.65	0.23
I	-0.13	-0.23	-0.07	-0.17	0.19	1	0.30	0.10
II	-0.10	-0.13	-0.08	-0.12	0.41	1	0.21	0.01
III	-0.15	-0.32	0.01	-0.20	-0.62	1	0.86	0.65
First-order autocorrelation	Δc	Δinv	Δw	Δy	l	π	r	$r^{\{4\}}$
U.S. data	0.17	0.52	0.15	0.22	0.96	0.86	0.94	0.95
Baseline	0.23	0.53	-0.20	0.15	0.92	0.78	0.85	0.78
I	0.22	0.45	0.11	0.19	0.90	0.81	0.83	0.88
II	0.23	0.44	0.06	0.16	0.89	0.78	0.82	0.88
III	0.28	0.63	0.35	0.24	0.97	0.96	0.96	0.97

References

- Adam, Klaus. 2005. "Learning to forecast and cyclical behavior of output and inflation." *Macroeconomic Dynamics* 9, 1-27.

- Adam, Klaus. 2007. "Optimal Monetary Policy with Imperfect Common Knowledge." *Journal of Monetary Economics* 54, 267-301.

- Adam, Klaus, and Albert Marcet. 2011. "Internal rationality, imperfect market knowledge and asset prices." *Journal of Economic Theory* 146, 1224-1252.

- Ang, Andrew, Monica Piazzesi, and Min Wei. 2006. "What does the yield curve tell us about GDP growth?" *Journal of Econometrics* 131, 359–403.

- Bekaert, Geert, Seonghoom Cho, and Antonio Moreno. 2010. "New-Keynesian Macroeconomics and the Term Structure". *Journal of Money, Credit and Banking* 42, 33-62.

- Branch, William A., and George W. Evans. 2006. "A simple recursive forecasting model." *Economics Letters* 91, 158-66.

- Casares, Miguel, and Jesús Vázquez. 2012. "Data revisions in the estimation of DSGE models". *Macroeconomic Dynamics*, forthcoming .

- Coenen, Günter, Andrew Levin, and Volker Wieland. 2005. "Data Uncertainty and the Role of Money as an Information Variable for Monetary Policy." *European Economic Review* 49, 975-1006.

- Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan. 2009. "New Keynesian models: Not yet useful for policy analysis." *American Economic Journal: Macroeconomics* 1, 242-266.

- De Graeve, Ferre, Marina Emiris and Rafael Wouters. 2009. "A structural decomposition of the US yield curve." *Journal of Monetary Economics* 56, 545-559.

- Estrella, Arturo, and Frederic S. Mishkin. 1997. "The predictive power of the term structure of interest rates in Europe and the United States: Implications for the European Central Bank." *European Economic Review* 41,1375–1401.

- Eusepi, Stefano, and Bruce Preston. 2011. "Expectations, learning, and business cycle fluctuations". *American Economic Review* 101, 2844-2872.

- Fama, Eugene. 1990. "Term structure forecast of interest rates, inflation and real returns." *Journal of Monetary Economics* 25, 59-76.

- Hördahl, Peter, Oreste Tristani, and David Vestin. 2006. "A joint econometric model of macroeconomic and term structure dynamics". *Journal of Econometrics* 131, 405-444.

- Levine, Paul , Joseph Pearlman, George Perendia, and Bo Yang. 2012. "Endogenous persistence in an estimated DSGE model under imperfect information." *Economic Journal* 122, 1287-1312.

- Mackowiak, Bartosz, and Mirko Wiederholt. 2009. "Optimal Sticky Prices under Rational Inattention." *American Economic Review* 99, 769-803.

- Mankiw, N. Gregory, and Ricardo Reis. 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* 117, 1295-1328.

- Milani, Fabio. 2007. "Expectations, learning and macroeconomic persistence." *Journal of Monetary Economics* 54, 2065-2082.

- Milani, Fabio. 2008. "Learning, monetary policy rules, and macroeconomic stability." *Journal of Economic Dynamics and Control* 32, 3148-3165.

- Milani, Fabio. 2011. "Expectation shocks and learning as drivers of the business cycle." *Economic Journal* 121, 379-401.

- Mishkin, Frederic S. 1991. "The information in the longer maturity term structure about future inflation." *Quarterly Journal of Economics* 105, 815-828.

- Ormeño, Arturo, and Krisztina Molnár. 2014. "Using survey data of inflation expectations in the estimation of learning and rational expectations models." *Journal of Money, Credit, and Banking* (forthcoming).

- Orphanides, Athanasios, and John C. Williams. 2005a. "Inflation scares and forecast-based monetary policy." *Review of Economic Dynamics* 8, 498-527.

- Orphanides, Athanasios, and John C. Williams. 2005b. "The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations." *Journal of Economic Dynamics and Control* 29, 1927-1950.

- Orphanides, Athanasios, and John C. Williams. 2008. "Learning, expectations formation, and the pitfalls of optimal control monetary policy." *Journal of Monetary Economics* 55 (Supplement), S80-S96.

- Orphanides, Athanasios, and Min Wei. 2012. "Evolving macroeconomic perceptions and the term structure of interest rates." *Journal of Economic Dynamics and Control* 36, 239-254.

- Pruitt, Seth. 2012. "Uncertainty over Models and Data: The Rise and Fall of American Inflation." *Journal of Money, Credit and Banking* 44, 341-365.

- Reis, Ricardo. 2009 "A Sticky Information General Equilibrium Model for Policy Analysis." In Klaus Schmidt-Heubel and Carl Walsh, eds., *Monetary Policy under Uncertainty and Learning*, Central Bank of Chile: Santiago, chapter 8, 227-283.

- Rudebusch, Glenn D., and Tao Wu. 2008. "A macro-finance model of the term structure, monetary policy and the economy." *Economic Journal* 118, 906-926.

- Sims, Christopher A. 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50, 665-690.

- Smets, Frank R., and Rafael Wouters. 2003. "An estimated dynamic stochastic general equilibrium model of the Euro Area." *Journal of the European Economic Association* 1, 1123-1175.

- Smets, Frank R., and Rafael Wouters. 2007. "Shocks and frictions in US business cycles: A Bayesian DSGE approach." *American Economic Review* 97, 586-606.

- Slobodyan, Sergey, and Rafael Wouters. 2012a. "Learning in an estimated medium-Scale DSGE model." *Journal of Economic Dynamics and Control* 36, 22-46.

- Slobodyan, Sergey, and Rafael Wouters. 2012b. "Learning in a medium-Scale DSGE model with expectations based on small forecasting models." *American Economic Journal: Macroeconomics* 4, 65-101.

- Stock, James H., and Mark W. Watson. 2002. "Has the business cycle changed and why?" In Gertler, Mark, and Kenneth S. Rogoff (eds.) *NBER Macroeconomics Annual*, Cambridge, MA: MIT Press, pp. 159-218.

- Vázquez, Jesús, Ramón María-Dolores, Juan M. Londoño. 2013. "On the informational role of term structure in the US monetary policy rule." *Journal of Economic Dynamics and Control* 37, 1852-1871.

Appendix

Table A. Model parameter description

φ	Elasticity of the cost of adjusting capital
h	External habit formation
σ_c	Inverse of the elasticity of intertemporal substitution in utility function
σ_l	Inverse of the elasticity of labor supply with respect to the real wage
ξ_p	Calvo probability that measures the degree of price stickiness
ξ_w	Calvo probability that measures the degree of wage stickiness
ι_w	Degree of wage indexation to past wage inflation
ι_p	Degree of price indexation to past price inflation
ψ	Elasticity of capital utilization adjustment cost
Φ	One plus steady-state fixed cost to total cost ratio (price mark-up)
r_π	Inflation coefficient in monetary policy rule
ρ_r	Smoothing coefficient in monetary policy rule
r_Y	Output gap coefficient in monetary policy rule
$r_{\Delta Y}$	Output gap growth coefficient in monetary policy rule
π	Steady-state rate of inflation
$100(\beta^{-1}-1)$	Steady-state rate of discount
l	Steady-state labor
γ	One plus steady-state rate of output growth
$\bar{r}^{\{4\}}$	Mean of the 1-year maturity yield
α	Capital share in production function
ρ	Learning coefficient

Table A. (*Continued*)

σ_a	Standard deviation of productivity innovation
σ_b	Standard deviation of risk premium innovation
σ_g	Standard deviation of exogenous spending innovation
σ_i	Standard deviation of investment-specific innovation
σ_R	Standard deviation of monetary policy rule innovation
σ_p	Standard deviation of price mark-up innovation
σ_w	Standard deviation of wage mark-up innovation
$\sigma^{\{4\}}$	Standard deviation of one year term premium innovation
ρ_a	Autoregressive coefficient of productivity shock
ρ_b	Autoregressive coefficient of risk premium shock
ρ_g	Autoregressive coefficient of exogenous spending shock
ρ_i	Autoregressive coefficient of investment-specific shock
ρ_R	Autoregressive coefficient of policy rule shock
ρ_p	Autoregressive coefficient of price mark-up shock
ρ_w	Autoregressive coefficient of wage mark-up shock
$\rho^{\{4\}}$	Autoregressive coefficient of one year term premium shock
μ_p	Moving-average coefficient of price mark-up shock
μ_w	Moving-average coefficient of wage mark-up shock
ρ_{ga}	Correlation coefficient between productivity and exogenous spending shocks