Macroeconomic Dynamics, **25**, 2021, 1635–1665. Printed in the United States of America.
doi:10.1017/S1365100519000816

AN ESTIMATED DSGE MODEL WITH LEARNING BASED ON TERM STRUCTURE INFORMATION

PABLO AGUILAR
Banco de España, Université catholique de Louvain and Universidad del País Vasco (UPV/EHU)

JESÚS VÁZQUEZ
Universidad del País Vasco (UPV/EHU)

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, S. and R. Wouters (2012a) *American Economic Journal: Macroeconomics* 4, 65–101.] by introducing the term structure of interest rates. This extension enables term structure information to fully characterize agents' expectations in real time. This feature addresses an imperfect information issue neglected in the related AL literature. The term structure of interest rates results in a strong channel of persistence driven by multi-period forecasting. Including the term structure in the AL model results in a model fit similar to that obtained in the rational expectation (RE) version of the model, but it greatly reduces the importance of other endogenous sources of aggregate persistence such as price and wage stickiness and the elasticity of the cost of adjusting capital. The model estimated also shows that term premium innovations are a major source of persistent fluctuations in nominal variables under AL. This stands in sharp contrast to the lack of transmission of term premium shocks to the macroeconomy under REs.

Keywords: Term Structure, Real-Time Adaptive Learning, Medium-Scale DSGE Model

1. INTRODUCTION

Agents can learn from financial markets to forecast macroeconomic outcomes. At the same time, learning dynamics can feed back into both the macroeconomy and

We are very grateful to Raf Wouters for his close guidance on this work. We are also thankful for helpful comments from the associate editor, two anonymous referees, Elena Mattana, Alfonso Novales, Rigas Oikonomou, Luca Pensieroso, Sergey Slobodyan, participants at the 2016 European Summer Meetings of the Econometric Society (Geneva) and seminar participants at the University of Namur, the University of Navarra, the University of the Basque Country, the University of Salamanca, and the Public University of Navarra. Some of this research was supported by the Banco de España, the Spanish Government, and the Basque Government, under grant numbers ECO2013-43773P, ECO2016-78749-P, and IT-793-13. Address correspondence to: Jesús Vázquez, Depto. FAE II, Facultad de Economía y Empresa, Av. Lehendakari Aguirre 83, 48015 Bilbao, Spain. e-mail: jesus.vazquez@ehu.es. Phone: (+34) 94 601 3779. Fax: (+34) 94 601 7123.

 1469-8056/19

financial markets. This paper introduces the term structure of interest rates into an estimated medium-scale dynamic stochastic general equilibrium (DSGE) model with adaptive learning (AL) expectations. Its aim is twofold. First, we analyze the role of the term structure of interest rates in the learning process of economic agents and show that this feature may take over other explanations of aggregate persistence. Second, we study how term structure innovations are transmitted into the macroeconomy under AL.

We build on the medium-scale AL model of Slobodyan and Wouters (2012a) based on small forecasting models by allowing agents explicitly to use data contained in the term structure of interest rates to fully characterize the formation of their expectations. This results in a major deviation from most AL models. Estimated AL models typically use *strongly revised* data, whereas actual learning dynamics must be driven by data available to agents when forming their expectations in real time.[1] Our approach overcomes this shortcoming because term structure information provides real-time information, which is not revised, about the behavior of the aggregate economy.[2] Apart from this distinctive feature, it is important to remark that our AL approach still shares an important feature with other AL approaches: The process of updating AL coefficients depends on actual values of forward-looking variables (e.g. aggregate consumption and inflation), because they are updated based on the forecast errors of these observable variables.

From a theoretical perspective, the term structure of interest rates can be viewed as summarizing information about expectations on both the macroeconomy and monetary policy. Thus, by introducing the term structure of interest rates, we consider a new channel where multi-period-ahead forecasts matter for household behavior. This feature is in line with ideas in Marcet and Sargent (1989), Preston (2005), Eusepi and Preston (2011), and Sinha (2016) of modeling optimal decisions conditional on multi-period forecasts. The rationale behind our approach is further supported by a large body of empirical literature (Fama, 1990; Mishkin, 1991; Estrella and Mishkin, 1997; Ang, Piazzesi and Wei, 2006) that evidences the ability of the term spread to predict both inflation and economic activity.

The small forecasting models assumed in this paper are very simple, but they perform well because the time-varying intercepts in them capture the long-run features of data well (e.g. the switching inflation trend in the early 1980s), whereas the time-varying coefficients associated with term spreads capture short-run movements relatively well (e.g. the growth rates in consumption and investment).

Our estimation results show that, compared with the rational expectations (REs) version of the model, including the term structure in the AL model provides a similar fit to postwar US data but results in a much lower degree of endogenous persistence associated with price and wage stickiness and the elasticity of the cost of adjusting capital. Our findings corroborate the fall in the elasticity of the cost of adjusting capital and price stickiness found in Slobodyan and Wouters (2012a) under AL.[3] The rationale behind these findings is that including the term structure in the DSGE model brings multi-period forecasting into the representative agent decision problem, which results in a much stronger persistence driven by learning dynamics than that induced by the one-period-ahead forecasts considered in their paper. Moreover, our findings are in line with those of Eusepi and Preston (2011) in a prototype real-business-cycle model, where multi-period forecasting results in much more persistent aggregate dynamics. Our empirical analysis also shows that including the term structure in the AL model provides additional support for important features found by Slobodyan and Wouters (2012a) in their estimated medium-scale DSGE model with AL. In particular, the AL model with term structure also reproduces the hump-shaped pattern of US inflation over the last 50 years. Furthermore, the AL model estimated shows that term spread innovations are a major source of fluctuations in nominal variables. This finding stands in sharp contrast to the lack of transmission of term spread shocks to the macroeconomy under RE.

The rest of the paper is organized as follows. Section 2 connects the contribution of this paper to the related literature. Section 3 introduces the term structure of interest rates into the medium-scale AL model. Section 4 discusses the main estimation results. Section 5 studies the transmission of shocks with a focus on term premium shocks. Section 6 assesses the robustness of estimation results across alternative formulations of the forecasting models. Section 7 concludes.

2. RELATED LITERATURE

Following the learning approach of Evans and Honkapohja (2001), this paper also builds on the growing literature that investigates the role of AL as an alternative to the assumption of RE in the analysis of DSGE models.[4] Recent papers (Orphanides and Williams, 2005a; Milani, 2007, 2008, 2011; Eusepi and Preston, 2011) focus on small-scale DSGE models, whereas Slobodyan and Wouters (2012a,b) introduce AL into the medium-scale DSGE model of Smets and Wouters (2007), hereinafter referred as the SW model.[5] The first group of papers considers that agents' expectations are based on a linear function of the state variables of the model whose learning coefficients are updated in each period under a gain rule (i.e. the minimum state variable approach), but Slobodyan and Wouters (2012a,b) consider an AL model with agents who form expectations using small forecasting models. Small forecasting models typically assume that agents form their expectations based on the information provided by observable endogenous variables (e.g. those showing up in Euler equations). 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 agents to know the true model economy (e.g. they know what the state variables are and they perfectly observe them).[6]

Our paper is also related to others that study the interaction between learning dynamics and term structure information (Dewachter, Iania and Lyrio, 2011; Sinha, 2015, 2016). Those papers mainly focus on how learning dynamics shape the yield curve, but our paper focuses on how term structure information influences learning dynamics to induce a strong source of aggregate persistence.

Deviating from the RE assumption by considering AL based on small forecasting models is appealing for two main reasons. First, in reality, agents face incomplete knowledge about the economy, which is at odds with the full information approach assumed under RE. Moreover, gathering and processing information is costly, so economic agents are likely to 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 that hit the economy, which provides an alternative competing hypothesis on aggregate persistence (Milani, 2007; Slobodyan and Wouters, 2012a).

Unfortunately, any form of deviation from the RE assumption studied in the literature is arguably arbitrary, and therefore requires further assessment. This is not a simple task. As suggested by Adam and Marcet (2011) and Slobodyan and Wouters (2012a), considering actual data on private sector expectations available through surveys or forward-looking variables, such as asset prices, might be very useful in disciplining expectation formation.[7] In this paper, we focus on term structure information rather than on the SPF information used in the related literature, because the former provides a much broader (market) view of agents' expectations. Moreover, we assume that the expectations hypothesis of the term structure of interest rates holds under AL and focus on the short end of the term structure. These two features can be viewed as somewhat restrictive, but we introduce them as a way of isolating the contribution of term structure information in characterizing AL. Thus, by imposing the expectation hypothesis of the term structure, we downplay the deviation of AL from RE because this hypothesis holds for the (linearized version of the) SW model under RE. Similarly, by focusing on the short end of the yield curve, we minimize the deviation of our AL model, which includes a four-quarters-ahead expectation horizon, from the AL model of Slobodyan and Wouters (2012a,b), which considers only one-quarter-ahead expectations. In a follow-up paper, Aguilar and Vázquez (2018), we relax these two assumptions, which enables us to further assess the relative importance of term structure information on AL by considering the whole yield curve.[8]

3. AN ADAPTIVE LEARNING MODEL WITH TERM STRUCTURE

This paper investigates the contribution of the term structure to the characterization of the agents' learning process. Our model builds on the SW model under AL suggested by Slobodyan and Wouters (2012a) in two important directions. First, we extend the model to account for the term structure of interest rates. Second, we assume that the information in the term structure, which is observed in real time, is the only information that enters the small forecasting models of all forward-looking variables (i.e. those involving expectations in the estimated DSGE model). This standard medium-scale DSGE model contains both nominal and real frictions that affect the choices of households and firms, and we briefly present this model below. However, our main focus is on the extensions related to both the term structure and AL. The complete log-linearized version of the SW model extended with AL and the term structure is presented in the Appendix together with a table describing parameter notation.

3.1. The SW Model

Households maximize their utility, which 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 that sets sticky nominal wages à la Calvo (1983). Households rent capital to firms and decide how much capital to accumulate depending on the capital adjustment costs that they face. Intermediate firms use capital and differentiated labor to produce differentiated goods and set prices à la Calvo. In addition, both prices and wages 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 costs and past inflation, whereas current wages are determined by past and expected future inflation and wages. Following Slobodyan and Wouters (2012a), we assume a Taylor-type rule where the short-term nominal interest rate reacts to inflation and to both the level and the growth rate of the output gap, where the latter is defined as the level of output relative to the underlying productivity process, rather than the natural output level used in the SW model.[9] In addition, we assume that the interest rate policy reacts to term structure information, which is in line with agents' learning processes. Finally, the model contains eight stochastic disturbances associated with technology shocks, price and wage markup shocks, and demand-side shocks, including policy interest rate and term premium shocks.

3.2. The Term Structure Extension

This section introduces the term structure into 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, we obtain the standard consumption-based asset pricing equation associated with each maturity:

$$E_t\left[\beta^j \frac{U_C(C_{t+j}, L_{t+j})\left(exp(\varepsilon_t^{\{j\}})(1+R_t^{\{j\}})\right)^j}{U_C(C_t, L_t)\prod_{k=1}^{j}(1+\pi_{t+k})}\right] = 1, \; for\, j = 1,\ 2,\ ...,\ n,$$

where E_t stands for the RE or the AL operator depending on the scenarios analyzed below, β is the discount factor, U_C denotes the marginal utility consumption,

https://doi.org/10.1017/S1365100519000816 Published online by Cambridge University Press

C_t, L_t, π_t, $R_t^{\{j\}}$, and $\varepsilon_t^{\{j\}}$ denote consumption, labor, the rate of inflation, the nominal yield, and the risk premium shock associated with a j-period maturity bond, respectively. The inclusion of a risk premium shock for each maturity is in line with the view of many authors of interpreting the gap between the pure-expectations-hypothesis-implied yield, $R_t^{\{j\}}$, and the observed yield as a measure of fluctuations in the term premium (e.g. De Graeve, Emiris and Wouters, 2009).[10]

Moreover, since we focus on US Treasury bond yields in the empirical analysis carried out below, term premium shocks can be viewed as a convenience yield term (Krishnamurthy and Vissing-Jorgensen, 2012; Greenwood et al., 2015; Del Negro et al., 2017) defined as a risk premium associated with the safety and liquidity features of Treasury bonds relative to assets with the same payoff and maturity, but without such outstanding properties. Furthermore, the introduction of term premium shocks into the model captures the imperfect substitutability between bonds of different maturities observed in the data in a simple way.

Considering the utility function assumed in the SW model, after some algebra, the (linearized) consumption-based asset pricing equations can be written as

$$x_t = E_t x_{t+j} - \left(\frac{1-x_1}{\sigma_c}\right)\left[j r_t^{\{j\}} - \sum_{k=1}^{j} E_t \pi_{t+k} + j\varepsilon_t^{\{j\}}\right] + x_2\left(l_t - E_t l_{t+j}\right), \quad for\, j = 1,\ 2,\ \ldots,\ n, \tag{1}$$

where lower case variables denote the log-deviations of consumption (hours worked) from its balanced-growth (steady-state) value or, alternatively, the deviations of the nominal interest rate, nominal yields, and the rate of inflation from their respective steady-state values. In particular, $r_t^{\{j\}}$ denotes the yield (written in deviations from its steady-state value) of a bond with a j-period maturity. The following notation is also used: $x_t = c_t - x_1 c_{t-1}$, $x_1 = \frac{h}{\gamma}$, $x_2 = \frac{(\sigma_c - 1)}{\sigma_c}\frac{WL}{C}$, where h and σ_c denote the habit formation and risk aversion parameters, and γ, W, L, and C denote the balanced-growth rate and the steady-state values of the real wage, hours worked, and consumption, respectively. In particular, for $j = 1$, equation (1) can be written as equation (2) in Smets and Wouters (2007).

As in Eusepi and Preston (2011), the forward-looking behavior of both RE and AL agents can be captured by iterating the optimality condition (1), with j set to $j = 1$, j periods forward, and using the law of iterating projections, to obtain:

$$x_t = E_t x_{t+j} - \left(\frac{1-x_1}{\sigma_c}\right)\sum_{k=1}^{j} E_t\left[r_{t+k-1}^{\{1\}} - E_t \pi_{t+k} + \varepsilon_{t+k-1}^{\{1\}}\right] + x_2\left(l_t - E_t l_{t+j}\right). \tag{2}$$

Since equations (1) and (2) must hold in equilibrium, they imply the term structure of interest rates hypothesis:

$$r_t^{\{j\}} = \frac{1}{j}\sum_{k=0}^{j-1} E_t r_{t+k} + \xi_t^{\{j\}}, \tag{3}$$

where the supraindex {1} on r_{t+k} has been removed for the sake of simplicity. Equation (3) states that the nominal yield of the j-period maturity bond, $r_t^{\{j\}}$, is equal to the average of the expectations for the short-term (one-period) nominal interest rate between periods t and $t+j-1$, plus a term premium shock, $\xi_t^{\{j\}} = \left(\frac{1}{j}\sum_{k=1}^{j} E_t\varepsilon_{t+k-1}^{\{1\}}\right) - \varepsilon_t^{\{j\}}$, defined as the difference between the average of the expectations for the one-period bond risk premium shocks between periods t and $t+j-1$ and the risk premium shock associated with the j-period maturity bond. For the sake of simplicity, we assume that each term premium shock $\xi_t^{\{j\}}$ follows an AR(1) process:[11]

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

Furthermore, our empirical formulation below includes a constant term to capture the mean (steady-state value) of a yield.

The approach followed in our empirical analysis relies on the term structure hypothesis, equation (3) for $j = 4$, together with the set of log-linearized dynamic equations used in Slobodyan and Wouters (2012a) described in the Appendix, which among others includes the optimality condition (1) for $j = 1$.

According to equation (1), the term structure of interest rates can be understood as summarizing information on agents' beliefs about the future paths of inflation, consumption, and hours worked. More precisely, the nominal return associated with a j-period maturity bond must be consistent with the expected paths of inflation, consumption, and hours worked from period t until the maturity period $t+j$. Similarly, the term spread, $sp_t^{\{j\}} = r_t^{\{j\}} - r_t$ for $2 \leq j \leq n$, is clearly a forward-looking variable under both RE and AL since a (longer-term) interest rate, $r_t^{\{j\}}$, involves (according to the optimality equation (3)) the expectations of future realizations of the short-term nominal interest rate. Therefore, the term spread is also capturing the market's expectations about monetary policy. Moreover, the term spread is also driven by term premium innovations, $\eta_t^{\{j\}}$. As discussed below, the model estimated shows that term premium shocks become an important source of inflation fluctuations under AL. This finding stands in striking contrast to the absence of transmission of term premium shocks to the macroeconomy under RE. Furthermore, the consideration of term structure information in a DSGE model under AL contributes to the goal of disciplining expectations by (i) characterizing agents' expectations beyond the one-period-ahead expectations considered in most DSGE models under AL; and (ii) using term structure information as observable in the estimation procedure.

3.3. The Adaptive Learning Extension

How agents' beliefs are characterized becomes a crucial issue when one deviates from the RE hypothesis. This paper assumes small linear forecasting models that agents follow to form their expectations, the so-called "perceived law of motion" (PLM). Following Slobodyan and Wouters (2012a), the coefficients of the PLM are updated through a Kalman filter with the arrival of new information. Next, the small forecasting models are combined to form the expectation functions of the forward-looking variables of the model. The small forecasting models assumed are described below. A section in the Appendix briefly describes how learning coefficients are updated through a Kalman filter.

3.3.1. A PLM with term structure information. We adapt our extended SW model with term structure to the AL version of this model. One of the key ingredients of a model with AL is the way in which agents' expectations formation is characterized (i.e. the PLM of agents), so it is important to motivate the choice of the PLM. In an AL model with only one-period-ahead expectations, Slobodyan and Wouters (2012a) suggest an AR(2) for each expectation formed at time t. They also include a time-varying intercept coefficient, which enables expectations to track down 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 (i) the interaction between term spreads and the expectations of consumption, inflation, and hours worked implicitly determined in the equilibrium condition (1); and (ii) the empirical evidence on the ability of term spreads to predict real economic activity and inflation (Estrella and Mishkin, 1997). At first sight, considering the whole term structure of interest rates to characterize AL might be seen as useful. However, considering term spreads associated with long-horizon bonds is rather challenging because it means having to define the whole set of expectations of the short-term nominal interest rate from the one-period horizon up to a long horizon. This cannot be done without imposing further restrictions on learning dynamics because, according to term structure equation (3), the number of parameters that define the PLM associated with these expectations dramatically increases with the number of expectations of the nominal short-term interest rate defined for alternative forecasting horizons, which in practice results in a curse of dimensionality problem. Furthermore, there is evidence (Mishkin, 1991) that at maturities longer than two-quarters the term structure of interest rates does not help to anticipate future inflationary pressures.

For all these reasons, our empirical analysis focuses on the 1-year term spread (i.e. $sp_t^{\{4\}}$) because it implies a rather parsimonious AL model (thus, $j=4$ in equation (3)). Moreover, following Slobodyan and Wouters (2008, 2012a), we allow agents to combine two forecasting models at the same time, track their forecasting performance, and use a variant of the Bayesian model averaging method to generate an aggregate forecast from the two forecasting models, which is used to characterize their decisions—see Slobodyan and Wouters (2008) for further details. The two forecasting models are

$$m_1: \quad E_t y_{t+j} = \theta_{1,y,t-1}^{\{j\}} + \beta_{1,y,1,t-1}^{\{j\}} sp_t + \beta_{1,y,2,t-1}^{\{j\}} sp_{t-1},$$

$$m_2: \quad E_t y_{t+j} = \theta_{2,y,t-1}^{\{j\}} + \beta_{2,y,1,t-1}^{\{j\}} sp_t,$$

where y_{t+j} stands for any variable that is being forecasted. Notice that these forecasting models are based only on current and lagged term spreads. In contrast to the forecasting models used in the related literature (e.g. Slobodyan and Wouters, 2012a,b), these two forecasting models depend only on real-time data.[12,13] Although these PLMs are determined by actual information available to agents at the time when they formed their expectations, it is important to remark that the process of updating AL coefficients through the Kalman filter still depends on a few actual (revised) variables, such as aggregate consumption and inflation, since the AL coefficients are updated based on the forecast errors of these observables.[14] Indeed, this feature may explain the good model fit even though we consider small forecasting models based only on term structure information.

We also analyze two additional forecasting models.[15] The first model considers a third forecasting model given by an AR(2) as in Slobodyan and Wouters (2012a) in addition to models m_1 and m_2:

$$m_3: \quad E_t y_{t+j} = \theta^{\{j\}}_{3,y,t-1} + \beta^{\{j\}}_{3,y,1,t-1} y_t + \beta^{\{j\}}_{3,y,2,t-1} y_{t-1},$$

whereas the second model includes the level of the 1-year rate, $r_t^{\{4\}}$, in m_3 instead of the AR(2) process, since its level is associated with inflation. That is,

$$m_3': \quad E_t y_{t+j} = \theta^{\{j\}}_{3,y,t-1} + \beta^{\{j\}}_{3,y,1,t-1} r_t^{\{4\}}.$$

As shown below, the inclusion of lagged variables or the level of the 1-year yield in the agents' forecasting model does not improve model fit once the term spreads are included.

4. ESTIMATION RESULTS

In this section, we first describe the data and the estimation approach. We then show the results of a few regression models for the observable forward-looking variables of the model to illustrate the relative predictive power of the 1-year term spread. 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 (2012a) (referred to hereafter as SlW); and (iv) the SW model with AL and term structure (hereafter SlWTS). Models (i) and (iii) are discussed and compared at length in Slobodyan and Wouters (2012a), whereas De Graeve, Emiris and Wouters (2009) discuss model variants of (i) and (ii). Therefore, our discussion focuses mainly on the interaction of AL expectations formation and the term structure of interest rates. This section also discusses the evolution of learning coefficients over time.

4.1. Data and the Estimation Approach

To facilitate comparison with the estimation results of Slobodyan and Wouters (2012a), the alternative models are estimated using US data for a 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 federal 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

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

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 that feature a long-run trend. $\overline{l}$, $\overline{\pi}$, $\overline{r}$, and $\overline{r}^{\{4\}}$ are the steady-state levels of hours worked, inflation, the federal funds rate, and the 1-year (four-quarters) constant maturity Treasury yield, respectively.[16]

The procedure for estimation under AL is essentially the same as in Slobodyan and Wouters (2012a), except that it is modified to take into account of the fact that agents are using both multi-period forecasting and a set of alternative forecasting models based only on term spread information. The estimation approach follows a two-step Bayesian estimation procedure. First, the log posterior function is maximized 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 (2012a). Moreover, we consider rather loose priors for the parameters that characterize 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.[17]

4.2. The Predictive Power of the Term Spread

Before discussing the estimates found in the Bayesian estimation of the DSGE model with AL, we illustrate the relative predictive power of the 1-year term spread by running five regression models for seven observable forward-looking variables of the model.

The first regression model includes the first lag of the dependent variable. The second model includes the first lags of the growth rates of consumption, investment, and the real wage together with the first lags of hours worked, inflation, and the short-term rate. The third model includes the first lag of the term spread. The

TABLE 1. Adjusted R^2 of alternative regression models

	Regression models				
Dependent variables	(1)	(2)	(3)	(4)	(5)
Consumption growth	0.03	0.18	0.21	0.14	0.08
Investment growth	0.27	0.39	0.23	0.22	0.15
Real wage growth	0.02	0.03	0.01	0.01	0.01
Hours worked	0.95	0.97	0.11	0.13	0.14
Inflation	0.75	0.76	0.47	0.48	0.44
Short-term interest rate	0.90	0.92	0.55	0.90	0.88
Short rate four-quarters ahead	0.51	0.58	0.38	0.51	0.53

Note: All regression models include a constant, and linear and quadratic trends to capture fluctuations in low-frequency components. In addition, model (1) includes the first lag of the dependent variable; model (2) includes the first lags of the growth rates of consumption, investment and the real wage together with the first lags of hours worked, inflation and the short-term rate; model (3) includes the first lag of the term spread; model (4) includes the first lag of the short-term interest rate; and model (5) includes the first lag of the 1-year yield.

fourth model includes the first lag of the short-term interest rate. Finally, the fifth model includes the first lag of the 1-year yield. In addition, all regression models include a constant along with linear and quadratic trends to capture fluctuations in low-frequency components. The first two models include revised data in general, while the other three use either the term spread or interest rates, which are observed in real time.[18]

Table 1 shows the (adjusted) R^2 for the alternative regression models. The term spread regression model does at least as well as the other models in predicting the growth rates of consumption, investment, and the real wage one-quarter ahead. However, it does much worse than the first two (and similarly to the last two) in predicting one-quarter-ahead inflation and hours worked, and the four-quarters-ahead short-term interest rate. Finally, the term spread does a reasonable job in predicting short-term interest rates. Overall, this regression analysis shows that the 1-year term spread can be helpful in forecasting many of the forward-looking variables of the model in real time.

4.3. Main Empirical Findings

Table 2 shows the estimation results associated with the four alternative DSGE models estimated divided into two panels. Panel A shows the marginal likelihood and the structural parameter estimates, while Panel B shows the estimates of the parameters that describe shock processes. In line with the findings of Slobodyan and Wouters (2012a), the marginal likelihood of the AL model without term structure using their PLM featured by an AR(2), SlW, is larger (−960.22) than that associated with the RE model, SW model (−973.76). However, when term structure information is included in the two models, the RE version performs similarly (−864.89) to our AL version (−867.97), which features small forecasting models

TABLE 2. Panel A: Priors and estimated posteriors of the structural parameters of the four alternative models

	Priors			Posteriors											
				SW			SW spread			SlW			SlWTS		
Log-				−973.76			−864.89			−960.22			−867.97		
likelihood.	Distr	Mean	SD	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
φ	N	4.00	1.50	5.96	4.21	7.63	6.06	4.33	7.89	3.34	1.88	3.87	1.72	1.14	2.32
h	B	0.70	0.10	0.79	0.74	0.86	0.77	0.69	0.83	0.68	0.63	0.75	0.77	0.72	0.83
σ_c	N	1.50	0.37	1.22	1.04	1.34	1.22	1.02	1.45	1.53	1.19	1.63	0.56	0.46	0.63
σ_l	N	2.00	0.75	1.50	0.64	2.32	1.82	0.89	2.53	1.74	1.02	2.60	2.12	1.37	2.78
ξ_p	B	0.50	0.10	0.70	0.62	0.80	0.74	0.57	0.88	0.64	0.59	0.69	0.42	0.35	0.49
ξ_w	B	0.50	0.10	0.71	0.62	0.78	0.86	0.81	0.91	0.82	0.76	0.85	0.66	0.58	0.72
ι_w	B	0.50	0.15	0.51	0.30	0.72	0.24	0.10	0.42	0.18	0.07	0.26	0.45	0.25	0.64
ι_p	B	0.50	0.15	0.25	0.10	0.38	0.27	0.10	0.39	0.27	0.11	0.39	0.17	0.07	0.27
ψ	B	0.50	0.15	0.55	0.36	0.72	0.53	0.34	0.74	0.50	0.31	0.71	0.58	0.38	0.78
Φ	N	1.25	0.12	1.62	1.48	1.73	1.57	1.41	1.68	1.58	1.45	1.73	1.38	1.26	1.49
r_π	N	1.50	0.25	1.98	1.71	2.25	1.80	1.15	2.09	1.74	1.38	2.04	1.69	1.45	1.99
ρ_r	B	0.75	0.10	0.84	0.80	0.87	0.87	0.83	0.89	0.88	0.85	0.91	0.80	0.76	0.85
r_y	N	0.12	0.05	0.10	0.05	0.13	0.11	0.06	0.15	0.13	0.07	0.18	0.16	0.10	0.22
$r_{\Delta y}$	N	0.12	0.05	0.15	0.12	0.18	0.17	0.14	0.20	0.13	0.10	0.16	0.20	0.16	0.24
r_{sp}	N	0.12	0.05	–	–	–	–	–	–	–	–	–	0.14	0.06	0.21
π	G	0.62	0.10	0.67	0.51	0.82	0.71	0.54	0.88	0.63	0.53	0.74	0.69	0.53	0.86
β	G	0.25	0.10	0.20	0.08	0.28	0.20	0.10	0.30	0.18	0.09	0.27	0.24	0.10	0.37
l	N	0.00	2.00	1.05	−0.5	2.14	0.59	−0.72	2.03	1.10	−0.76	1.96	0.01	−1.33	1.21
γ	N	0.40	0.10	0.40	0.38	0.42	0.41	0.39	0.42	0.40	0.36	0.43	0.43	0.40	0.45
$\bar{r}^{\{4\}}$	N	1.00	0.50	–	–	–	1.38	1.19	1.58				1.23	0.77	1.70
α	N	0.30	0.05	0.19	0.16	0.22	0.18	0.15	0.20	0.17	0.13	0.19	0.17	0.13	0.19
ρ	B	0.50	0.28	–	–	–	–	–	–	0.97	0.96	0.98	0.84	0.79	0.98

TABLE 2. Panel B: Priors and estimated posteriors of the structural parameters of the four alternative models

	Priors			Posteriors											
				SW			SW spread			SlW			SlWTS		
	Distr	Mean	SD	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
σ_a	IG	0.10	2.00	0.44	0.39	0.48	0.45	0.40	0.49	0.46	0.41	0.49	0.49	0.45	0.55
σ_b	IG	0.10	2.00	0.25	0.21	0.29	0.24	0.20	0.28	0.15	0.14	0.20	0.89	0.78	0.99
σ_g	IG	0.10	2.00	0.52	0.47	0.57	0.51	0.48	0.56	0.50	0.46	0.56	0.50	0.45	0.55
σ_i	IG	0.10	2.00	0.44	0.37	0.52	0.44	0.36	0.52	0.46	0.42	0.52	1.57	1.41	1.72
σ_R	IG	0.10	2.00	0.22	0.20	0.24	0.22	0.20	0.24	0.21	0.19	0.23	0.24	0.21	0.26
σ_p	IG	0.10	2.00	0.14	0.11	0.17	0.17	0.14	0.19	0.15	0.13	0.16	0.30	0.27	0.33
σ_w	IG	0.10	2.00	0.23	0.20	0.27	0.30	0.27	0.33	0.21	0.18	0.22	0.56	0.51	0.61
$\sigma^{\{4\}}$	IG	0.10	2.00	–	–	–	0.10	0.09	0.11				0.19	0.17	0.22
ρ_a	B	0.50	0.20	0.93	0.89	0.96	0.93	0.90	0.97	0.98	0.98	0.99	0.96	0.92	0.99
ρ_b	B	0.50	0.20	0.17	0.04	0.26	0.26	0.15	0.40	0.43	0.29	0.55	0.91	0.87	0.95
ρ_g	B	0.50	0.20	0.98	0.97	0.99	0.98	0.97	0.99	0.96	0.95	0.98	0.98	0.96	0.99
ρ_i	B	0.50	0.20	0.70	0.59	0.79	0.69	0.59	0.79	0.50	0.37	0.57	0.96	0.93	0.98
ρ_R	B	0.50	0.20	0.08	0.01	0.14	0.06	0.01	0.11	0.09	0.01	0.16	0.15	0.05	0.27
ρ_p	B	0.50	0.20	0.84	0.74	0.95	0.85	0.75	0.99	0.32	0.06	0.57	0.87	0.79	0.94
ρ_w	B	0.50	0.20	0.97	0.96	0.99	0.95	0.93	0.98	0.54	0.32	0.80	0.97	0.96	0.98
$\rho^{\{4\}}$	B	0.50	0.20	–	–	–	0.84	0.77	0.91	–	–	–	0.94	0.90	0.99
μ_p	B	0.50	0.20	0.67	0.49	0.85	0.75	0.61	0.90	0.47	0.29	0.67	0.41	0.26	0.53
μ_w	B	0.50	0.20	0.87	0.82	0.94	0.98	0.86	0.99	0.43	0.11	0.70	0.26	0.15	0.36
ρ_{ga}	B	0.50	0.20	0.51	0.37	0.65	0.52	0.39	0.67	0.17	0.13	0.19	0.54	0.41	0.66

Note: The labels N, B, G, and IG denote normal, beta, gamma, and inverse gamma prior distribution, respectively.

TABLE 3. Model's fit for alternative PLMs

	Baseline	Adding AR(2)	Adding 1-year yield	Only intercept
Log-likelihood	−865.20	−865.41	−866.42	−910.83

Notes: Log-likelihood values evaluated at the mode (Laplace approximation).

based only on term structure information.[19] This finding might come as a surprise because the forecasting models assumed can be viewed on the one hand as rather realistic given that they ignore information on revised macroeconomic data which are not available to economic agents on forming their expectations; but on the other hand they look rather restrictive in that they take into account only term structure information.

In order to motivate the use of PLMs based only on term spreads, we estimated two additional models as discussed above. Thus, we firstly analyzed a forecasting model in which lagged (revised) endogenous variables are included in the PLMs (model m_3) in addition to the forecasting models based on term spreads considered in the baseline case. Second, we considered a forecasting model in which the level of the 1-year yield is included in the PLMs (model m_3') in addition to the forecasting models based on term spreads. Table 3 shows that including either lagged endogenous variables or the level of the 1-year rate does not improve model fit once the term spread is included in the PLMs (models m_1 and m_2). As argued above, the good performance of the PLM including term spreads is due to the way in which AL coefficients are updated through the Kalman filter. Namely, the learning coefficients are updated based on the forecast errors and those forecast errors depend on actual values of forward-looking variables, such as aggregate consumption and inflation. In this way, a few learning coefficients are updated using actual data. In particular, the time-varying intercept is able to capture the low-frequency movement of inflation as shown in Figure 1 below. Moreover, in reality, term spreads might be capturing other information that agents have about the course of the economy not otherwise present in the model. As a consequence, the reasonable performance of the AL model with term structure may be due to the information conveyed by term spreads over and above that captured in the model.

As also shown in Table 3, a simple PLM which includes only a time-varying intercept fits the data much worse than the PLMs based on term structure information. Nonetheless, such a simple PLM based only on a time-varying intercept fits the data better than PLMs that include many more variables analyzed in Section 6 below. In particular, we show that the marginal likelihood decreases rather dramatically if the small forecasting models follow the AR(2) structure assumed in Slobodyan and Wouters (2012a). The reason is rather simple: Since many variables in the model largely comove, including a few of those variables in the same forecasting model may induce (multi-) collinearity problems. Why might multicollinearity be a problem? The chances of converging to a local maximum in this scenario increase because the likelihood function becomes flat in some regions of the parameter space under multicollinearity. These considerations certainly make a compelling case for the use of a combination—i.e. model averaging—of small forecasting models instead of a single forecasting model with many explanatory variables. The PLMs assumed in this paper are very simplistic, but the bottom line is that the time-varying intercepts of the PLMs capture the long-run features of data (e.g. inflation) well, whereas the time-varying coefficients associated with term spreads seem to capture short-run movements (growth rates of consumption and investment) relatively well as shown in Table 1.

Our estimation results also suggest that the fitting of the short-term yield curve summarized by the 1-year yield becomes somewhat more challenging for the AL model than for the RE model. That is, the increase in the log marginal likelihood when the 1-year yield is considered as observable in the estimation procedure is much larger under RE, at 108.87 $(= -864.89 - (-973.76))$, than under AL, at 92.25 $(= -867.97 - (-960.22))$.[20]

Table 2 shows that in general many parameter estimates are rather robust across models, with a few important differences. These differences are discussed 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 determine the contribution of the term spread when interacting with AL.[21]

4.3.1. SW model versus SlW. As found by Slobodyan and Wouters (2012a), considering AL instead of the RE assumption in the SW model estimated greatly reduces the importance of both exogenous persistence due to price markup and wage markup shocks and endogenous persistence induced by habit formation, the elasticity of the cost of adjusting capital, and wage indexation. The intuition of these findings is rather simple: AL dynamics introduce a major channel of endogenous persistence that is ignored when the RE assumption is considered. As a consequence, a few sources of persistence under the RE assumption are less significant under AL in reproducing the observed persistence in most macroeconomic variables.

4.3.2. SlW model versus SlWTS. The introduction of the term spread into the PLM results in many more changes than those introduced by the single-step extension of the SW model with AL analyzed above. On the one hand, a few sources of endogenous persistence are less important. Thus, the estimates of the inverse of consumption intertemporal elasticity, σ_c, and the elasticity of the cost of adjusting capital, φ, are much smaller under AL with term structure (0.56 and 1.72, respectively) than without (1.53 and 3.34, respectively).[22] Similarly, the price and wage probabilities (ξ_p and ξ_w) and the price indexation parameter, ι_p, are much smaller under AL with term structure than without term structure, whereas the opposite occurs for the wage indexation parameter, ι_w. We find contrasting results as regards exogenous sources of price and wage markup persistence. Thus, the autoregressive parameters of the ARMA processes that characterize price and wage markup shocks (ρ_p and ρ_w) are higher in the AL model with term structure (and similar to those estimated in the RE models) than in the model without term structure. However, the moving-average parameters of the ARMA processes that characterize price and wage markup shocks (μ_p and μ_w) are lower in the AL model with term structure than in the AL model without term structure. In short, we observe that AL under the multi-period forecasting hypothesis associated with the inclusion of term structure information results in lower estimates of the parameters that characterize important sources of endogenous persistence than those implied by either the AL model without term structure information or the RE model. These results are in line with those found in Eusepi and Preston (2011) in the sense that AL with multi-period forecasting takes over in explaining aggregate persistence, while other sources become less important.

On the other hand, it is important to recall that the term spread is a forward-looking variable. This means that learning dynamics endowed with term spread information are less sluggish. Thus, the estimated persistence of belief coefficients, ρ, is lower when the term spread in the SlWTS model (0.84) is considered than in the SlW model (0.97). As a consequence of the much faster adjustment of belief coefficients, the estimates of most parameters that capture exogenous persistent dynamics in the SlWTS (i.e. ρ_b, ρ_i, ρ_p, and ρ_w) are much higher, so they mimic actual data persistence. Moreover, a lower value of ρ means not only less sluggish learning dynamics but also that the belief coefficients themselves are much less volatile. When the belief coefficients are less volatile, the forecasting models have less chance of hitting the projection facility boundary (i.e. the beliefs react less to the same forecasting errors). Hitting the projection facility boundary is costly in terms of likelihood, so this is a valuable feature which enables the initial values of the beliefs to be initialized closer to the unit root boundary. Since we follow Slobodyan and Wouters (2012a,b) and initialize the beliefs by calculating them at the RE equilibrium implied by the estimated parameters, larger values of the estimated parameters that feature exogenous persistence are also consistent with the smaller estimated value of ρ.[23]

Our results also show that the risk premium and investment disturbances—and to a lesser extent markup disturbances—are more volatile under AL based on term structure information. The sensitivity of the estimated shock process parameters to the specific PLM considered may suggest that these disturbances can be viewed as simply wedges (i.e. reduced-form shocks) rather than truly structural shocks as argued in Chari, Kehoe and McGrattan (2009).[24]

4.4. Analysis of the PLM

Figure 1 shows the trend in the PLM coefficients for inflation and consumption. We focus on these two forward-looking variables because they have observable counterparts. The time-varying intercept of the PLM of inflation (consumption)

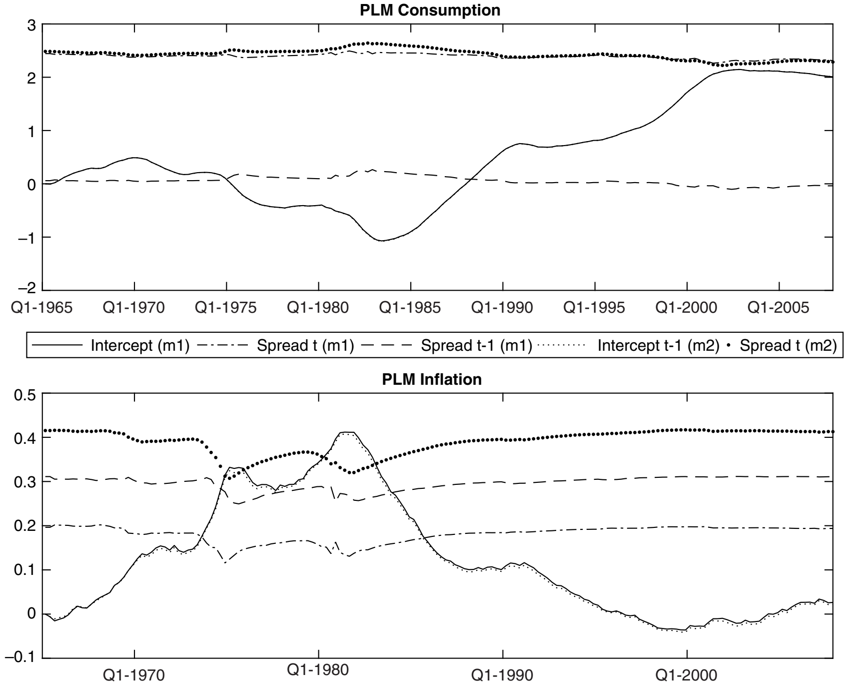

FIGURE 1. PLM coefficients of inflation and consumption.

shows how agents' perceptions of steady-state inflation (balanced-growth consumption) vary over time. Thus, the intercept of inflation expectations captures the rise of inflation expectations during the 60s and 70s and the rapid fall of inflation expectations from the early 80s onward. The intercept of consumption captures the fall in long-term consumption expectations from the early 70s to the mid-80s, capturing the economic recessions around the oil price shocks, followed by a steady rise in long-term consumption expectations up to the early 2000s.

In contrast to the time-varying intercept coefficients, the term spread coefficients in the two PLMs are fairly stable. Moreover, the term spread coefficients associated with these two PLMs are positive, indicating that a higher 1-year bond yield today anticipates tighter financial conditions in the future (say four-quarters ahead), which results in higher one-quarter-ahead expectations of inflation and consumption because agents substitute current consumption for future consumption. Notice that the sum of the term spread coefficients associated with the PLM of consumption is roughly five times larger than that associated with the PLM of inflation, which results in much larger swings in consumption expectations than in inflation expectations due to term spread changes. This greater sensitivity of consumption expectations (relative to inflation expectations) to term spread variations is consistent with the relative size of the impulse responses of consumption and inflation to a term spread innovation shown below. Also,

https://doi.org/10.1017/S1365100519000816 Published online by Cambridge University Press

notice that the small forecasting model m_1 characterizes inflation expectations differently because both contemporaneous and lagged spreads have nontrivial coefficients in the agents' forecasting function, while consumption expectations are described by the simpler m_2 forecasting model. This empirical finding further motivates the combination of two alternative small forecasting models to describe all forward-looking variables of the model.

5. THE TRANSMISSION OF SHOCKS

As shown by Slobodyan and Wouters (2012a,b), the transmission of structural shocks is crucially determined by the way in which agents form their expectations. Therefore, it is important to analyze how impulse response functions (IRFs) shift over time driven by changes in the updating belief coefficients. Our IRF analysis focuses on the estimated time-varying responses of a selected group of real (i.e. output and consumption) and nominal (inflation and the short-term interest rate) variables to a term premium innovation. As emphasized above, the term spread is a forward-looking variable. Thus, the responses of output and inflation to term premium innovations illustrate how term structure innovations anticipate movements in these variables.[25]

5.1. Impulse Responses to a Term Premium Shock

The introduction of AL extended with term structure information permits a feedback from the term structure to the macroeconomy through the learning dynamics that is missing under RE. Figure 2 shows the time-varying IRF of output, consumption, inflation, and the nominal interest rate to a term premium innovation under AL. The stability of term spread learning coefficients associated with the PLM featured only by the term spread shown above means that the IRF change modestly over time, with more prominent changes in the responses during the 70s and early 80s. A positive innovation in the term premium shock increases the 1-year yield relative to the (short-term) federal funds rate, which brings forward consumption and investment decisions, resulting in higher current economic activity (output and consumption), inflation, and (short-term) nominal interest rate. The impulse responses of all variables are hump-shaped, capturing a different transmission mechanism (relative to RE). This hump-shaped feature is more pronounced in the nominal variables (inflation and nominal interest rate) than in the real variables (output and consumption).

5.2. Variance Decomposition Analysis

Table 4 shows the variance decomposition of output, consumption, inflation, and the short-term nominal interest rate for two alternative forecasting horizons. Risk premium shocks explain a large proportion of output and consumption

TABLE 4. Variance decomposition

	1-year horizon				10-year horizon			
	y	c	π	r	y	c	π	r
Productivity	22.82	7.54	6.85	7.74	31.61	7.87	10.33	11.48
Risk premium	51.40	74.70	0.11	18.61	35.60	59.74	0.09	9.37
Exogenous spending	13.25	10.26	0.08	4.42	23.07	23.26	0.06	3.47
Investment technology	3.34	2.02	1.77	44.93	1.42	0.99	1.45	21.85
Monetary policy	1.46	0.74	76.80	15.52	1.47	0.96	51.36	18.19
Price markup	5.06	3.56	0.06	1.61	5.09	5.21	0.90	1.64
Wage markup	2.62	1.08	2.45	3.55	1.45	0.73	2.96	5.15
Term premium	0.06	0.08	11.87	3.62	0.28	1.25	32.84	28.85

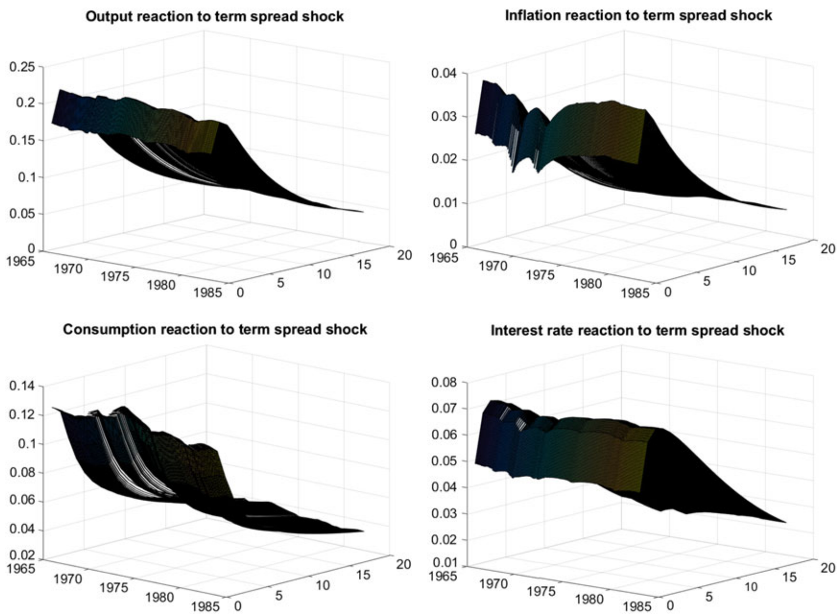

FIGURE 2. IRF to a term premium shock.

variability at short-term forecasting horizons, whereas both productivity and exogenous spending significantly increase their proportions as the forecasting horizon increases.

Monetary policy shocks explain a large proportion of inflation variability at short-term forecasting horizons, but term premium shocks account for an increasing proportion at long-term forecasting horizons (32.84%). This last feature also shows up when the variability of the short-term nominal interest rate is analyzed.

https://doi.org/10.1017/S1365100519000816 Published online by Cambridge University Press

Thus, 28.85% of the short rate long-run variability is explained by the term premium shock. Interestingly, price and wage markup shocks play a small role in explaining inflation variability. This is in contrast to the variance decompositions reported in Smets and Wouters (2007) and Slobodyan and Wouters (2012a), where these two shocks explain a sizable share of inflation variability.

In short, our results suggest that term premium innovations together with monetary innovations take over markup shocks in explaining nominal variability. However, their role in real variables is actually small, with the risk premium shock playing a prominent role, especially in explaining consumption variability.

6. ROBUSTNESS ANALYSIS

This section studies the robustness of the estimation results of the SlWTS model using four alternative specifications for the PLM across two dimensions: in-sample fit and parameter estimates. The first alternative specification (labeled "I") uses the PLM suggested by Slobodyan and Wouters (2012a) (i.e. an AR(2)). The second alternative formulation (labeled "II") considers the minimum set of state variables. The third specification (labeled "III") includes only the lagged term spread in all PLMs. Finally, the fourth specification (labeled "IV") considers the baseline learning specification, but with the belief coefficient, ρ, restricted to zero. The baseline formulation and the last two use only term spread information observed in real time, while the other two use revised data. Table 5 shows the estimation results for the five alternative specifications of the PLM studied.[26] The marginal likelihood evaluated at the mode clearly favors the baseline specification. In particular, notice that the (posterior) marginal likelihood for the Slobodyan and Wouters estimation is -960.22 in Table 2, whereas the marginal likelihood computed at the mode is -958.62. When the 1-year yield is included in the set of observables, the marginal likelihood for the Slobodyan and Wouters estimation is -952.45 at the mode as reported in Table 5. This small improvement in model fit when adding the 1-year yield as observable in the Slobodyan and Wouters estimation is in contrast to the large fit improvement found in our model: 6.17 $(= -952.45 - (-958.62))$ versus 101.27 $(= -865.20 - (-966.47))$, where all these values are computed at the mode. This finding suggests that the fitting of the 1-year yield becomes much more challenging for the forecasting models of Slobodyan and Wouters than for the forecasting models based on term spreads. Moreover, the use of the lagged term spread instead of the current term spread results in a large deterioration in the fit of the model. The conclusions are similar for the other three formulations. The deterioration becomes much larger when the learning coefficient, ρ, is restricted to zero.

In regard to parameter estimates, a few structural parameters (e.g. habit formation parameter, h, Frisch elasticity, σ_l, and steady-state markup, Φ), and all

TABLE 5. Panel A: Robustness analysis of alternative PLM: structural parameters

Log-likelihood	Baseline −865.20		I −952.45		II −952.01		III −933.63		IV −1155.90	
	Mode	SD	Mode	SD	Mode	SD	Mode	SD	Mode	SD
φ	1.60	0.38	4.46	0.18	4.74	0.24	1.34	0.38	1.02	0.02
h	0.78	0.04	0.65	0.11	0.81	0.07	0.73	0.07	0.62	0.02
σ_c	0.50	0.11	1.37	0.15	1.20	0.14	0.53	0.07	0.70	0.02
σ_l	2.16	0.48	1.55	0.10	2.05	0.19	2.00	0.51	1.41	0.01
ξ_p	0.44	0.09	0.48	0.13	0.70	0.14	0.44	0.10	0.43	0.02
ξ_w	0.63	0.04	0.58	0.12	0.57	0.14	0.62	0.16	0.91	0.02
ι_w	0.38	0.17	0.18	0.13	0.43	0.21	0.37	0.24	0.31	0.02
ι_p	0.16	0.26	0.39	0.14	0.55	0.17	0.32	0.39	0.23	0.02
ψ	0.49	0.24	0.04	0.10	0.11	0.17	0.34	0.49	0.87	0.03
Φ	1.36	0.08	1.51	0.13	1.57	0.12	1.44	0.11	1.35	0.03
r_π	1.73	0.17	1.84	0.11	1.86	0.19	1.72	0.35	2.05	0.02
ρ_r	0.81	0.04	0.86	0.08	0.89	0.05	0.63	0.04	0.88	0.02
r_y	0.13	0.20	0.03	0.14	0.14	0.20	–	–	0.11	0.02
$r_{\Delta y}$	0.20	0.08	0.15	0.11	0.15	0.15	–	–	0.35	0.02
r_{sp}	0.14	0.09	0.07	0.11	0.15	0.17	0.14	0.47	0.11	0.02
π	0.65	0.10	0.84	0.08	0.61	0.19	0.63	0.16	0.79	0.02
β	0.22	0.08	0.27	0.15	0.17	0.16	0.34	0.22	0.35	0.02
l	−0.32	0.34	−0.08	0.18	−1.02	0.21	−1.43	0.52	−1.02	0.02
γ	0.43	0.02	0.39	0.09	0.40	0.05	0.41	0.04	0.44	0.03
$\bar{r}^{\{4\}}$	1.30	0.16	1.99	0.18	2.44	0.24	1.49	0.35	1.43	0.02
α	0.15	0.05	0.18	0.11	0.19	0.16	0.15	0.06	0.23	0.03
ρ	0.95	0.03	0.99	0.03	0.99	0.03	0.97	0.01	0.0	–

policy parameter estimates are observed to be fairly robust across the alternative specifications of the PLM studied. However, there are also some important differences. In particular, the estimates of most parameters that measure endogenous persistence (ξ_p, φ, and ψ) are somewhat sensitive to the specification used. Nevertheless, these parameter estimates are rather similar for the two PLM formulations that use the term spread (Baseline and "III"), indicating that the weakness of these endogenous sources of persistence is robust when the term spread is introduced into the PLM.

The parameter estimates of the shock processes are also observed to be quite stable across PLM formulations. However, there are a few noticeable differences. Thus, the persistence of risk premium and investment shocks and the size of their innovations are much greater for the PLM specifications that include the term spread (Baseline, "III", and "IV") than for the others. This high sensitivity of shock process estimates reinforces the suggestion discussed above that these disturbances may capture reduced-form shocks rather than truly structural shocks (Chari, Kehoe and McGrattan, 2009).

Table 5. Panel B: Robustness analysis of alternative PLM: shock process parameters

	Baseline		I		II		III		IV	
	Mode	SD	Mode	SD	Mode	SD	Mode	SD	Mode	SD
σ_a	0.50	0.05	0.44	0.11	0.47	0.12	0.48	0.06	0.50	0.02
σ_b	0.90	0.08	0.22	0.13	0.31	0.12	0.79	0.08	4.69	0.02
σ_g	0.49	0.03	0.50	0.14	0.51	0.10	0.50	0.07	0.50	0.02
σ_i	1.52	0.10	0.46	0.13	0.94	0.14	1.54	0.09	1.75	0.02
σ_R	0.23	0.02	0.23	0.14	0.22	0.14	0.24	0.07	0.24	0.02
σ_p	0.29	0.05	0.16	0.10	0.16	0.16	0.27	0.12	0.32	0.02
σ_w	0.55	0.04	0.61	0.14	0.44	0.09	0.55	0.06	0.56	0.01
$\sigma^{\{4\}}$	0.19	0.02	0.15	0.10	0.19	0.09	0.15	0.02	0.16	0.02
ρ_a	0.95	0.02	0.88	0.12	0.97	0.03	0.94	0.04	0.98	0.01
ρ_b	0.90	0.02	0.56	0.14	0.23	0.17	0.96	0.02	0.76	0.02
ρ_g	0.98	0.01	0.98	0.15	0.99	0.01	0.98	0.01	0.97	0.03
ρ_i	0.96	0.02	0.58	0.14	0.59	0.21	0.95	0.03	0.93	0.02
ρ_R	0.12	0.23	0.23	0.16	0.09	0.18	0.59	0.18	0.12	0.02
ρ_p	0.85	0.06	0.84	0.14	0.59	0.18	0.87	0.10	0.76	0.03
ρ_w	0.97	0.10	0.88	0.09	0.91	0.09	0.98	0.03	0.98	0.02
$\rho^{\{4\}}$	0.92	0.04	0.95	0.07	0.96	0.03	0.93	0.06	0.97	0.02
μ_p	0.41	0.11	0.55	0.13	0.59	0.14	0.57	0.16	0.39	0.02
μ_w	0.29	0.16	0.24	0.15	0.44	0.16	0.28	0.26	0.07	0.01
ρ_{ga}	0.54	0.09	0.62	0.12	0.57	0.15	0.52	0.25	0.53	0.02

Notes: Log-likelihood values evaluated at the mode (Laplace approximation). The column labeled "Baseline" shows the parameter estimates for the baseline model discussed in the previous section. Column "I" shows the estimates for the model using the PLM suggested by Slobodyan and Wouters (2012a) (i.e. an AR(2)). Column "II" shows the estimates for the model that considers the minimum set of state variable approach to describe the PLM. Column "III" reports the estimates for the model that includes only the lagged term spread in all PLMs. Finally, Column "IV" shows the estimates when the learning ρ is restricted to zero.

7. CONCLUSIONS

In this paper, we extend the AL model of Slobodyan and Wouters (2012a) by introducing the term structure of interest rates. Our extension enables the term spread of interest rates to fully characterize the expectations of all forward-looking variables with term structure information, which is observed in real time, while retaining the features of AL based on small forecasting models and the process of updating learning coefficients based on actual values of forward-looking variables. This extension overcomes, to some extent, a general shortcoming of estimated AL versions of DSGE models based only on final revised data when in reality agents only have access to real-time data when updating their expectations over time.

Our estimation results show that including the term structure in the AL model results in decreases in the importance of a few endogenous sources of aggregate persistence (such as price and wage stickiness and the elasticity of the cost of

https://doi.org/10.1017/S1365100519000816 Published online by Cambridge University Press

adjusting capital). The rationale for this finding is rather simple. The extension of the estimated DSGE model introduces a multi-period forecasting environment, which means that learning dynamics are much more persistent than in one-period-ahead forecasting models (e.g. Slobodyan and Wouters, 2012a,b). The importance of multi-period forecasting is in line with results found in Eusepi and Preston (2011) for a real-business-cycle model. Moreover, our estimated DSGE model with term structure information shows that term premium innovations are an important source of persistent fluctuations in the nominal variables (inflation and the short-term interest rate). This finding stands in sharp contrast to the lack of transmission of term premium shocks to the macroeconomy under REs.

This paper should be viewed as a first step toward understanding the importance of relying on multi-period AL based on real-time data to characterize aggregate persistence. In a follow-up paper, Aguilar and Vázquez (2018), we investigate an extended version of our model in which AL and the term structure features from the whole yield curve are combined with real-time macroeconomic information to determine agents' expectations and their decisions.

NOTES

1. There are a few exceptions that address this important shortcoming in the AL literature. Milani (2011) focuses on real-time data on output and inflation and forecasts of them from the Survey of Professional Forecasters (SPF) recorded in real time when estimating a small-scale DSGE model. However, he ignores revised data on macroeconomic variables, which more accurately describe the actual economy, when estimating and assessing the model's fit. Orphanides and Wei (2012) also use real-time data, but focus on a vector autoregression (VAR) model rather than a DSGE model. Similarly, Slobodyan and Wouters (2013) use both real-time data and survey expectations data in a univariate model of inflation.

2. The idea of using only term structure information to predict business cycle conditions is not new (McCallum, 1994). The use of only term structure information to characterize agents' expectations can certainly be seen as somewhat restrictive because agents indeed observe additional relevant information. Below we show that considering additional (revised) data does not improve model fit once term structure information is included in small forecasting models. Moreover, in a follow-up paper, Aguilar and Vázquez (2018) show that adding real-time data into small forecasting models does not improve the overall performance of the DSGE model estimated.

3. Milani (2007) reached qualitatively similar conclusions, but he studied a small-scale DSGE model, where many sources of aggregate persistence were neglected.

4. This paper is also related to another strand of the literature (Hördahl, Tristani and Vestin, 2006; Rudebusch and Wu, 2008; Bekaert, Cho and Moreno, 2010) that seeks to link small-scale new Keynesian monetary model dynamics with the term structure of interest rates under RE. De Graeve, Emiris and Wouters (2009) show evidence of the importance of considering medium-scale RE-DSGE models in order to understand the links between the term structure and the aggregate economy.

5. There is also a large body of macroeconomic literature that analyzes deviations from the RE assumption in the context of small-scale models. This literature includes, among others, the rational inattention approach (Sims, 2003; Adam, 2007; Mackowiak and Wiederholt, 2009), the sticky information approach (Mankiw and Reis, 2002; Reis, 2009), and several approaches that deal with imperfect information issues (Svensson and Woodford, 2004; Coenen, Levin and Wieland, 2005; Levine, Pearlman, Perendia and Yang, 2012; Pruitt, 2012; Casares and Vázquez, 2016).

6. Other papers (Adam, 2005; Orphanides and Williams, 2005b; Branch and Evans, 2006; Eusepi and Preston, 2011; Hommes and Zhu, 2014; Ormeño and Molnár, 2015) have also provided support

https://doi.org/10.1017/S1365100519000816 Published online by Cambridge University Press

for the use of small forecasting models on several grounds such as their relative forecast performance and their ability to approximate the SPF well.

7. In this vein, recent papers that introduce AL into DSGE models typically use the SPF to include additional observables in order to discipline agents' beliefs (e.g. Milani, 2011, and Ormeño and Molnár, 2015, Cole and Milani, 2019) or to assess the performance of AL expectations as in Eusepi and Preston, 2011 and Slobodyan and Wouters (2012a).

8. Furthermore, Aguilar and Vázquez (2018) consider real-time inflation as well as SPF data on consumption growth and inflation forecasts to discipline agents' expectations. Nonetheless, the use of consumption growth forecasts from the SPF force us to restrict our attention to the Great Moderation period (1984–2007) in that paper because this time series starts in 1981:3. The sample period is much shorter than the one considered (1966–2007) in this paper, which is similar to that studied in Slobodyan and Wouters, 2012a,b for the purposes of comparison. Other important differences between these two papers that show alternative AL modeling choices are highlighted below.

9. This assumption avoids the modeling of the flexible economy, which includes many additional forward-looking variables. In the rest of the paper, we continue to refer to the SW model even though we consider an alternative measure of the output gap.

10. As argued in Ireland (2004), there is a long-standing tradition of introducing additional disturbances into DSGE models in order to match the number of shocks with the number of time series used in estimation. This is because DSGE models typically introduce fewer shocks than observable variables, which means that certain combinations of endogenous variables are deterministic. If these combinations do not hold in the data, any approach that attempts to estimate the complete model will fail.

11. 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. The first difference is mainly a matter of semantics but the second 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 under AL. Second, this allows a more straightforward comparison with the Slobodyan and Wouters (2012a) model that helps to assess the importance of the AL expectations formation and the role of the term spread.

12. Readers might think that the second forecasting model adds nothing new to the first one. However, the simpler model m_2 may be useful in practice during episodes when the term spread is highly persistent. This is because the two regressors in model m_1 would be highly correlated, inducing a multicollinearity problem in those episodes which could be overcome by activating model m_2 instead.

13. This timing assumption is used in Canova and Gambetti (2010), Slobodyan and Wouters (2012a,b), and Ormeño and Molnár (2015), among others. It is shared by RE models, but it may assume a richer information set than the set observed by agents when forming their forecasts in reality. Alternatively, one can assume that at time t agents only observe lagged values, sp_{t-1} . We investigate this alternative hypothesis below, and more deeply in Aguilar and Vázquez (2018).

14. See equation (4a) in Slobodyan and Wouters (2012a). One possible way of overcoming this issue is to use additional vintage data, but this would inevitably increase the set of observables. Moreover, if the bulk of aggregate data revisions take place in the first revision, final revised data would be a good proxy of vintage revised data used in the updating procedure of AL coefficients.

15. Moreover, a preliminary version of this paper, Aguilar and Vázquez (2015), investigated a wide range of PLM combining both revised data and term structure information in a single forecasting model. The overall fit of the DSGE model deteriorates when these complicated PLMs are used.

16. Some might view the assumption of a constant inflation target, $\overline{\pi}$, as rather restrictive. However, Milani (in press) has recently shown that the inflation target is remarkably stable over time when AL is assumed instead of RE.

17. The RE versions of the DSGE models are estimated using standard Dynare routines, whereas the AL versions of the models use the codes kindly provided by Sergey Slobodyan and Raf Wouters with a few modifications to accommodate the presence of the term spread in the PLM.

18. This regression analysis should be viewed as a simple illustration of the predictive power of the term spread to predict the observable foward-looking variables of the model under AL, because these variables enter the model measured either in deviations from the balance growth path or in deviations from their respective steady-state values (neither is actually observed by the econometrician) instead of the growth rates or the levels used as dependent variables in the regression analysis. However, we consider the growth rates of consumption, investment, and the real wage instead of any alternative definition of their cyclical components, such as their linear detrended components, for two main reasons: first, the growth rates of these variables are the observables that enter into the measurement equation. Second, as argued below, the time-varying intercept of the PLMs captures the low-frequency (i.e. long-run) components of the data, whereas the remaining variables in the PLM explain mainly the high-frequency components described by the growth rates.

19. As pointed out by Del Negro and Eusepi (2011, p. 2116), a small difference of five points or less can be overturned by choosing a slightly different set of priors.

20. See Del Negro and Eusepi (2011) for a discussion of the econometric framework, based on log marginal likelihood differences, for assessing how a model estimated to fit a benchmark set of time series (in our case, the set of observables in Slobodyan and Wouters, 2012a) performs on fitting an additional time series (the 1-year yield).

21. We also consider the 1-year Treasury constant maturity yield to estimate the SW model under RE. Introducing the 1-year yield in the SW model barely changes parameter estimates with a few exceptions. The Calvo wage 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 endogenous sources (compared to exogenous sources) in explaining price and wage persistence increases when the RE model is extended with the short-term yield 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.

22. As emphasized in Smets and Wouters (2007), a lower elasticity of the cost of adjusting capital increases the sensitivity of investment to the real value of the existing capital stock, Tobin's q (see equation (A2) in the Appendix).

23. We thank Sergey Slobodyan for pointing this out.

24. This sensitivity of risk premium shock parameters is in line with the different estimates reported for these parameters in Smets and Wouters (2007) and Galí, Smets and Wouters (2012), which were obtained using a similar sample period.

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

26. All models in Table 5 are estimated using the same set of observables, which includes the seven observables in Slobodyan and Wouters (2012a) and the 1-year yield.

27. From the zero-profit condition in steady state, it should be noticed that ϕ_p also represents the value of the steady-state price markup.

28. For a detailed explanation, see Slobodyan and Wouters (2012a).

REFERENCES

Adam, K. (2005) Learning to forecast and cyclical behavior of output and inflation. *Macroeconomic Dynamics* 9, 1–27.

Adam, K. (2007) Optimal monetary policy with imperfect common knowledge. *Journal of Monetary Economics* 54, 267–301.

Adam, K. and A. Marcet (2011) Internal rationality, imperfect market knowledge and asset prices. *Journal of Economic Theory* 146, 1224–1252.

Aguilar, P. and J. Vázquez (2015) The role of term structure in an estimated DSGE model with learning. IRES, Université catholique de Louvain, Discussion paper 2015007.

Aguilar, P. and J. Vázquez (2018) Term structure and real time learning. Banco de España, Working paper 1803.

Ang, A., M. Piazzesi and M. Wei (2006) What does the yield curve tell us about GDP growth? *Journal of Econometrics* 131, 359–403.

Bekaert, G., S. Cho and A. Moreno (2010) New-Keynesian macroeconomics and the term structure. *Journal of Money, Credit and Banking* 42, 33–62.

Branch, W. A. and G. W. Evans (2006) A simple recursive forecasting model. *Economics Letters* 91, 158–166.

Canova, F. and L. Gambetti (2010) Do expectations matter? The great moderation revisited. *American Economic Journal: Macroeconomics* 2, 183–205.

Casares, M. and J. Vázquez (2016) Data revisions in the estimation of DSGE models. *Macroeconomic Dynamics* 20, 1683–1716.

Chari, V.V., P. J. Kehoe and E. R. McGrattan (2009) New Keynesian models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics* 1, 242–266.

Coenen, G., A. Levin and V. Wieland (2005) Data uncertainty and the role of money as an information variable for monetary policy. *European Economic Review* 49, 975–1006.

Cole, S. J. and F. Milani (2019) The misspecification of expectations in new Keynesian models: a DSGE-VAR approach. *Macroeconomic Dynamics* 23, 974–1007.

De Graeve, F., M. Emiris and R. Wouters (2009) A structural decomposition of the US yield curve. *Journal of Monetary Economics* 56, 545–559.

Del Negro, M. and S. Eusepi (2011) Fitting observed inflation expectations. *Journal of Economic Dynamics and Control* 35, 2105–2131.

Del Negro, M., D. Giannone, M. P. Giannoni and A. Tambalotti (2017) Safety, liquidity, and the natural rate of interest. *Brookings Papers on Economic Activity*, Spring 2017, 235–294.

Dewachter, H., L. Iania and M. Lyrio (2011) A New-Keynesian model of the yield curve with learning dynamics: A Bayesian evaluation. MPRA paper 34461.

Estrella, A. and F. 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, S. and B. Preston (2011) Expectations, learning, and business cycle fluctuations. *American Economic Review* 101, 2844–2872.

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

Fama, E. (1990) Term structure forecast of interest rates, inflation and real returns. *Journal of Monetary Economics* 25, 59–76.

Galí, J., F. R. Smets and R. Wouters (2012) Unemployment in an estimated new Keynesian model. In: D. Acemoglu and M. Woodford (eds.), *NBER Macroeconomics Annual*, vol. 26, pp. 329–360. Cambridge, MA: MIT Press.

Greenwood, R., S. G. Hanson and J. C. Stein (2015) A comparative–advantage approach to government debt maturity. *Journal of Finance* 70, 1683–1722.

Hommes, C. and M. Zhu (2014) Behavioral learning equilibria. *Journal of Economic Theory* 150, 778–814.

Hördahl, P., O. Tristani and D. Vestin (2006) A joint econometric model of macroeconomic and term structure dynamics. *Journal of Econometrics* 131, 405–444.

Ireland, P. N. (2004) A method for taking models to the data. *Journal of Economic Dynamics and Control* 28, 1205–1226.

Krishnamurthy, A. and A. Vissing-Jorgensen (2012) The aggregate demand for Treasury debt. *Journal of Political Economy* 120, 233–267.

Levine, P., J. Pearlman, G. Perendia and B. Yang (2012) Endogenous persistence in an estimated DSGE model under imperfect information. *Economic Journal* 122, 1287–1312.

Mackowiak, B. and M. Wiederholt (2009) Optimal sticky prices under rational inattention. *American Economic Review* 99, 769–803.

Mankiw, N. G. and R. Reis (2002) Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve. *Quarterly Journal of Economics* 117, 1295–1328.
Marcet, A. and T. J. Sargent (1989) Convergence of least-squares learning in environments with hidden states variables and private information. *Journal of Political Economy* 97, 1306–1322.
McCallum, B. T. (1994) Monetary policy and the term structure of interest rates. NBER Working Paper 4938. Reprinted in the *Federal Reserve Bank of Richmond Economic Quarterly*, Volume 91/4 Fall 2005.
Milani, F. (2007) Expectations, learning and macroeconomic persistence. *Journal of Monetary Economics* 54, 2065–2082.
Milani, F. (2008) Learning, monetary policy rules, and macroeconomic stability. *Journal of Economic Dynamics and Control* 32, 3148–3165.
Milani, F. (2011) Expectation shocks and learning as drivers of the business cycle. *Economic Journal* 121, 379–401.
Milani, F. (in press) Learning and the evolution of the Fed's inflation target. *Macroeconomic Dynamics*. doi:10.1017/S136510051900004X.
Mishkin, F. S. (1991) The information in the longer maturity term structure about future inflation. *Quarterly Journal of Economics* 105, 815–828.
Ormeño, A. and K. Molnár (2015) Using survey data of inflation expectations in the estimation of learning and rational expectations models. *Journal of Money, Credit, and Banking* 47, 673–699.
Orphanides, A. and J. C. Williams (2005a) Inflation scares and forecast-based monetary policy. *Review of Economic Dynamics* 8, 498–527.
Orphanides, A. and J. 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, A. and J. C. Williams (2008) Learning, expectations formation, and the pitfalls of optimal control monetary policy. *Journal of Monetary Economics* 55 (Supplement), S80–S96.
Orphanides, A. and M. Wei (2012) Evolving macroeconomic perceptions and the term structure of interest rates. *Journal of Economic Dynamics and Control* 36, 239–254.
Preston, B. (2005) Learning about monetary policy rules when long-horizon expectations matters. *International Journal of Central Banking* 1, 81–126.
Pruitt, S. (2012) Uncertainty over models and data: The rise and fall of American inflation. *Journal of Money, Credit and Banking* 44, 341–365.
Reis, R. (2009) A sticky information general equilibrium model for policy analysis. In K. Schmidt-Heubel and C. Walsh (eds.), *Monetary Policy under Uncertainty and Learning*, Chapter 8, pp. 227–283. Central Bank of Chile.
Rudebusch, G. D. and T. Wu (2008) A macro-finance model of the term structure, monetary policy and the economy. *Economic Journal* 118, 906–926.
Sims, C. A. (2003) Implications of rational inattention. *Journal of Monetary Economics* 50, 665–690.
Sinha, A. (2015) Government debt, learning and the term structure. *Journal of Economic Dynamics and Control* 53, 268–289.
Sinha, A. (2016) Learning and the yield curve. *Journal of Money, Credit and Banking* 48, 513–547.
Slobodyan, S. and R. Wouters (2008) Estimating a medium-scale DSGE model with expectations based on small forecasting models. *Mimeo.*
Slobodyan, S. and R. Wouters (2012a) Learning in a medium-Scale DSGE model with expectations based on small forecasting models. *American Economic Journal: Macroeconomics* 4, 65–101.
Slobodyan, S. and R. Wouters (2012b) Learning in an estimated medium-scale DSGE model. *Journal of Economic Dynamics and Control* 36, 22–46.
Slobodyan, S. and R. Wouters (2013) Survey expectations and learning. *Mimeo.*
Smets, F. R. and R. Wouters (2007) Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review* 97, 586–606.
Vázquez, J., R. María-Dolores and J. 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

LOG-LINEARIZED DYNAMIC EQUATIONS

In addition to equations (1) with $j = 1$ and (3) with $j = 4$ characterizing the one-quarter rate and the 1-year bond yield, respectively, the set of the remaining log-linearized dynamic equations are the following:

- Aggregate resource constraint:

$$y_t = c_y c_t + i_y i_t + z_y z_t + \varepsilon_t^g, \tag{A1}$$

  where $c_y = \frac{C}{Y} = 1 - g_y - i_y$, $i_y = \frac{I}{Y} = (\gamma - 1 + \delta)\frac{K}{Y}$, and $z_y = r^k \frac{K}{Y}$ are steady-state ratios. As in Smets and Wouters (2007), the depreciation rate and the exogenous spending-to-GDP ratio are fixed in the estimation procedure at $\delta = 0.025$ and $g_y = 0.18$.
- Investment equation:

$$i_t = i_1 i_{t-1} + (1 - i_1) E_t i_{t+1} + i_2 q_t + \varepsilon_t^i, \tag{A2}$$

  where $i_1 = \frac{1}{1+\overline{\beta}}$, and $i_2 = \frac{1}{(1+\overline{\beta})\gamma^2\varphi}$ with $\overline{\beta} = \beta\gamma^{(1-\sigma_c)}$.
- Arbitrage condition (value of capital, q_t):

$$q_t = q_1 E_t q_{t+1} + (1 - q_1) E_t r_{t+1}^k - (r_t - E_t \pi_{t+1}) + c_3^{-1} \varepsilon_t^b, \tag{A3}$$

  where $q_1 = \overline{\beta}\gamma^{-1}(1 - \delta) = \frac{(1-\delta)}{(r^k+1-\delta)}$.
- Log-linearized aggregate production function:

$$y_t = \Phi\left(\alpha k_t^s + (1 - \alpha) l_t + \varepsilon_t^a\right), \tag{A4}$$

  where $\Phi = 1 + \frac{\phi}{Y} = 1 + \frac{\text{Steady-state fixed cost}}{Y}$ and α is the capital share in the production function.[27]
- Effective capital (with one period time-to-build):

$$k_t^s = k_{t-1} + z_t. \tag{A5}$$

- Capital utilization:

$$z_t = z_1 r_t^k, \tag{A6}$$

  where $z_1 = \frac{1-\psi}{\psi}$.
- Capital accumulation equation:

$$k_t = k_1 k_{t-1} + (1 - k_1) i_t + k_2 \varepsilon_t^i, \tag{A7}$$

  where $k_1 = \frac{1-\delta}{\gamma}$ and $k_2 = \left(1 - \frac{1-\delta}{\gamma}\right)\left(1 + \overline{\beta}\right)\gamma^2\varphi$.
- Marginal cost:

$$mc_t = (1 - \alpha) w_t + \alpha r_t^k - \varepsilon_t^a. \tag{A8}$$

- New Keynesian Phillips curve (price inflation dynamics):

$$\pi_t = \pi_1 \pi_{t-1} + \pi_2 E_t \pi_{t+1} - \pi_3 mc_t + \varepsilon_t^p, \tag{A9}$$

where $\pi_1 = \frac{\iota_p}{1+\overline{\beta}\iota_p}$, $\pi_2 = \frac{\overline{\beta}}{1+\overline{\beta}\iota_p}$, $\pi_3 = \frac{A}{1+\overline{\beta}\iota_p}\left[\frac{(1-\overline{\beta}\xi_p)(1-\xi_p)}{\xi_p}\right]$, and $A = \frac{1}{(\varphi_p-1)\varepsilon_p+1}$. The coefficient of the curvature of the Kimball goods market aggregator, included in the definition of A, is fixed in the estimation procedure at $\varepsilon_p = 10$ as in Smets and Wouters (2007).

- Optimal demand for capital by firms:

$$-\left(k_t^s - l_t\right) + w_t = r_t^k. \quad \textbf{(A10)}$$

- Wage markup equation:

$$\mu_t^w = w_t - mrs_t = w_t - \left(\sigma_l l_t + \frac{1}{1-h/\gamma}\left(c_t - (h/\gamma)\, c_{t-1}\right)\right). \quad \textbf{(A11)}$$

- Real wage dynamic equation:

$$w_t = w_1 w_{t-1} + (1-w_1)\left(E_t w_{t+1} + E_t \pi_{t+1}\right) - w_2 \pi_t + w_3 \pi_{t-1} - w_4 \mu_t^w + \varepsilon_t^w. \quad \textbf{(A12)}$$

where $w_1 = \frac{1}{1+\overline{\beta}}$, $w_2 = \frac{1+\overline{\beta}\iota_w}{1+\overline{\beta}}$, $w_3 = \frac{\iota_w}{1+\overline{\beta}}$, $w_4 = \frac{1}{1+\overline{\beta}}\left[\frac{(1-\overline{\beta}\xi_w)(1-\xi_w)}{\xi_w((\phi_w-1)\varepsilon_w+1)}\right]$ with the curvature of the Kimball labor aggregator fixed at $\varepsilon_w = 10.0$ and a steady-state wage markup fixed at $\phi_w = 1.5$ as in Smets and Wouters (2007).

- 1-year term spread:

$$sp_t^{\{4\}} = r_t^{\{4\}} - r_t. \quad \textbf{(A13)}$$

- Monetary policy rule:

$$r_t = \rho_r r_{t-1} + (1-\rho_r)\left[r_\pi \pi_t + r_y \hat{y}_t\right] + r_{\Delta y}\Delta\hat{y}_t + r_{sp} sp_t^{\{4\}} + \varepsilon_t^r, \quad \textbf{(A14)}$$

where the output gap is defined as $\hat{y}_t = y_t - \Phi\varepsilon_t^a$ (i.e. the output gap is defined as the deviation of output from its underlying neutral productivity process).

ADAPTIVE LEARNING EXPECTATION FORMATION

This part of the Appendix provides a brief explanation of how adaptive learning (AL) expectation formation works.[28] A DSGE model can be represented in matrix form as follows:

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

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 expectations of the forward-looking variables, $E_t y_{t+j}$, are defined as linear functions of variables entering in the information set of agents, whose time-varying (learning) coefficients are updated as explained below. Once the expectations of the forward-looking variables, $E_t y_{t+j}$, are computed they are plugged into the matrix representation of 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 matrices $A_0, A_1, A_2^{\{j\}}$, and B_0) together with learning coefficients discussed next.

The PLM process is generally defined as follows:

$$y_{t+j} = X_t \beta_{t-1}^{\{j\}} + u_{t+j},$$

where y is the vector containing the k forward-looking variables of the model, X is the matrix of the kxr regressors, $\beta^{\{j\}}$ is the vector of the r updating learning coefficients, which includes an intercept, and u is a vector of errors. These errors are linear combinations of the true model innovations, ϵ_t. So, the variance–covariance matrix, $\Sigma = E[u_{t+j}u_{t+j}^T]$, 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-looking variable:

$$E_t y_{t+j} = X_t \beta_{t-1}^{\{j\}}.$$

Thus, agents update their learning coefficient estimates using data up to time $t-1$, but X_t contains contemporaneous values of the regressors. The updating parameter vector, β, which results from stacking all the vectors $\beta^{\{j\}}$, 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 (2012a) 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. Notice that the learning parameter ρ plays a role very similar to the *constant gain* parameter used in the recursive least-squares learning.

The Kalman filter updating and transition equations for the belief coefficients and the corresponding covariance matrix are given by

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

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 $t-1$, and $R_{t|t-1}$ is the mean squared error associated with $\beta_{t|t-1}$. 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 squared error, $R_{t|t}$, associated with this updated estimate, $\beta_{t|t}$, is given by

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

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

https://doi.org/10.1017/S1365100519000816 Published online by Cambridge University Press

TABLE A.1. 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 markup)
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	Updating beliefs coefficient
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 markup innovation
σ_w	Standard deviation of wage markup innovation
$\sigma^{\{4\}}$	Standard deviation of 1-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 markup shock
ρ_w	Autoregressive coefficient of wage markup shock
$\rho^{\{4\}}$	Autoregressive coefficient of 1-year term premium shock
μ_p	Moving-average coefficient of price markup shock
μ_w	Moving-average coefficient of wage markup shock
ρ_{ga}	Correlation coefficient between productivity and exogenous spending shocks