*American Economic Journal: Macroeconomics 2012, 4(2): 65–101*
*http://dx.doi.org/10.1257/mac.4.2.65*

# Learning in a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models†

*By* Sergey Slobodyan and Raf Wouters*

*This paper evaluates the empirical performance of a medium-scale DSGE model with agents forming expectations using small forecasting models updated by the Kalman filter. The adaptive learning model fits the data better than the rational expectations (RE) model. Beliefs about the inflation persistence explain the observed decline in the mean and the volatility of inflation as well as Phillips curve flattening. Learning about inflation results in lower estimates for the persistence of the exogenous shocks that drive price and wage dynamics in the RE version of the model. Expectations based on small forecasting models are closely related to the survey evidence on inflation expectations.* (*JEL* C53, D83, D84, E13, E17, E31)

In modern macroeconomics, the behavior of economic agents is defined by a series of intertemporal optimization problems. The solution to these problems depends crucially on how agents form their expectations about future events. From a theoretical perspective, RE appear as a logical assumption: expectations are formed consistently with the underlying model and the policy environment, and all available information is used efficiently. This hypothesis is extremely useful for a macroeconomist as it tightens the link between theory and estimation, and it allows for an efficient estimation of the deep parameters of the model by exploiting all the cross-equation restrictions that are imposed through the model-consistent expectations hypothesis. However, RE do not provide a description of the information problem that agents have to solve to discover these systematic relations. As far as this information processing or learning process affects and changes economic decisions of various agents over time, RE will provide an incorrect and implausibly restrictive

* Slobodyan: CERGE-EI, Politickych veznu 7, 111 21 Prague 1, Czech Republic (e-mail: sergey.slobodyan@cerge-ei.cz); Wouters: National Bank of Belgium/Université catholique de Louvain, Boulevard de Berlaimont, B-1000 Brussels, Belgium (e-mail: rafael.wouters@nbb.be). The first author expresses sincere gratitude to NBB for hospitality and support which made this project possible. We thank participants at the San Francisco Fed Conference on "Macroeconomic Models for Monetary Policy," the EABCN Workshop on "Uncertainty over the Business Cycle," the Magyar Nemzeti Bank/CEPR Macroeconomic Policy Research Workshop on "DSGE Models: A Closer Look at the Workhorse of Macroeconomics," 2009 SED Conference, "Adaptive Learning and Macroeconomic Policy" conference at Cambridge University, Joint Lunchtime Seminar of the European Central Bank (ECB), Center for Financial Studies (CFS), and Deutsche Bundesbank, and 2008 "Learning Week" conference at St. Louis Fed. We particularly thank our discussants George Evans, Marco Del Negro, Kaushik Mitra, and Krisztina Molnár for very useful comments. The views expressed are solely our own and do not necessarily reflect those of the National Bank of Belgium.

† To comment on this article in the online discussion forum, or to view additional materials, visit the article page at http://dx.doi.org/10.1257/mac.4.2.65.

specification of the economy.[1] In reality, households and firms have limited knowledge and diffused information about the correct form of the underlying model, about the exact value of the model parameters or the state vector of variables, and especially about the exogenous and latent disturbances that hit the economy. Agents, like econometricians, need to find out the dynamic structure of the economy using the data available in real time. As processing information is costly, it is more realistic to assume that they will concentrate on a limited amount of information, and that they update their beliefs about the underlying economic relations as new data becomes available in order to capture possible changes in the stochastic structure or in the policy environment. If expectations are allowed to deviate from the RE solution, the model dynamics will change as well, and expectations become, potentially, an important additional source of business cycle fluctuations.

Therefore, in this paper, we assume that expectations are based on a limited information set, meaning that agents use small forecasting models in forming their beliefs about future realizations, i.e., by using simple autoregressive models, and that they adapt the coefficients of these forecasting models by a simple Kalman filter updating procedure. This approach implies only a modest departure from the RE model as we leave the decision problems of the agents intact, and we only replace the model-consistent expectations in their optimizing behavioral rules by a more realistic adaptive expectations mechanism. By allowing agents to use under-parameterized expectation models, the macroeconomist receives some additional degree of freedom as he has to make a decision on the relevant forecasting model or, in other words, on the information set that agents are using in their expectations model. However, this does not appear as an important empirical issue for several reasons. First, it turns out that the data are informative on this decision, so that the additional flexibility can be exploited optimally to fit the overall model on the data without any danger of ending up in an unidentified system. Moreover, larger forecasting models, incorporating more information, are characterized by higher parameter uncertainty and more instability in the updating process. This overfitting problem introduces a bias toward smaller forecasting models. Second, the estimates of structural parameters and the identification of the shocks do not seem to be very sensitive to the specific choice of the information set included in the forecasting model. Finally, the use of explicit data on private sector expectations available through surveys or forward-looking variables, like asset prices, might be very helpful in the selection of the relevant forecasting model.

Recent empirical DSGE models systematically retain the hypothesis of rational expectations. Smets and Wouters (2003, 2007) have shown that these models, when equipped with a rich set of frictions and a general stochastic structure, explain the data relatively well. However, it is still somewhat problematic that these models

[1] These limitations of the RE approach are widely recognized in the macroeconomic literature and generally acknowledged as a serious weakness and challenge for the new generation of DSGE models that are now popular in both academic and applied research. Over the last decades, alternative approaches to model expectations have been suggested in the literature: the Bounded Rationality approach of Sargent (1993), the Rational Inattention approach of Sims (2003), the Epidemiology of Macroeconomic Expectations by Carroll (2001), the Sticky Information model of Mankiw and Reis (2002), the Partial Information model of Pearlman, Currie, and Levine (1986) and Svensson and Woodford (2003), and the learning approach of Evans and Honkapohja (2001). Our paper is situated in the learning tradition and neglects the imperfect information issue in the estimation process.

require highly persistent exogenous shocks to explain the observed persistence in the data (Chari, Kehoe, and McGrattan 2009). Milani (2006, 2007) and Orphanides and Williams (2005a) claim that learning can significantly influence the macroeconomic dynamics and increase the persistence in the responses to shocks. Milani (2006, 2007) estimates a small-scale model, both under RE and learning, and shows that the learning reduces the scale of structural frictions and results in an improved marginal likelihood relative to the RE model. Orphanides and Williams (2005a) illustrate how adaptive learning can lead to inflation persistence. Slobodyan and Wouters (2012a) analyze the learning dynamics in the SW model and find that learning hardly influences model dynamics if the information set used in the learning process is the same as under rational expectations. Restricting information available to the agents improves the model fit and produces IRFs that match those from the best-fitting DSGE-VAR models.

In this paper, we show that the empirical fit of the standard medium-scale DSGE model is significantly improved when private agents are assumed to form their expectations on the basis of recursively updated small forecasting models. This empirical success depends crucially on the fact that private sector expectations about future inflation are allowed to adjust to the changing persistence in the observed inflation dynamics over the last 40 years. This perceived inflation persistence is also crucial for the appreciation of the nominal volatility and the variance of the inflation forecast errors in particular. Therefore, with adaptive learning about inflation dynamics, the Kalman filter is capable of tracking the time-varying covariance structure of the observable variables much better than under RE, and this explains most of the improvement in the model likelihood. The resulting time variation in the propagation mechanism of the various shocks explains the upward and downward trend in inflation volatility during the great inflation and the subsequent great moderation period, and is also consistent with the observed flattening of the Phillips curve over the last 20 years. These results are obtained with homoscedastic innovations for the exogenous processes and with constant parameters in the monetary policy rule. Recursive updating of the private sector's perceived law of motion for inflation persistence and for the long-run inflation target drives the results in our model. Accordingly, the high volatility and close to the random walk persistence of macroeconomic variables in the 1970s is mainly the result of a particular realization of positive markup and negative productivity shocks during that period. This interpretation differentiates our results from that of Cogley, Primiceri, and Sargent (2010) who identify the reduced volatility in the central bank's inflation target as the exogenous source for the change in the inflation dynamics. In contrast, during the great moderation, the low perceived inflation persistence reduced the sensitivity of inflation expectations and actual inflation to marginal cost fluctuations and monetary policy shocks. The important role of the inflation expectations dynamics in our model, and in particular the possible divergence between private sector expectations and the complete information-based rational expectation forecast, also justifies why central bankers are so concerned about these expectations.

Adaptive learning in our model has an important effect on the estimated stochastic structure and the implied decomposition of the business cycle in terms of the exogenous shocks. In particular, the highly persistent component in the exogenous disturbances to price and wage dynamics disappears, reducing these processes to independently and identically distributed innovations. In other words, the learning process substitutes for the exogenous persistence in the shocks that is typically required in the RE model. This also mitigates the long-run impact of these mark-up shocks on the real economy, so that productivity again becomes the dominant source of long-run growth fluctuations. This result is noteworthy because the dominant role of the mark-up shocks and the high persistence in these exogenous processes has been an important source of criticism of the NK-DSGE models. The impact of learning on the estimated structural parameters that determine the behavioral policy rules of the various agents is relatively minor, with the exception of the wage indexation parameter in the Calvo model, which is estimated to be significantly lower than in the RE setting. This result confirms the importance of adaptive expectations for generating persistence in the nominal responses.

Besides Slobodyan and Wouters (2011), the closest study to our paper is Milani (2007). Our findings confirm Milani's results by showing that adaptive learning is a feasible and empirically preferable alternative to RE when estimating DSGE models. By estimating the learning dynamics in a larger DSGE model, we were obliged to consider the particular problem of restricting the information set that is taken into account in the belief forecasting models, instead of using the Minimum State Variable information set as in Milani (2007). We observe that Kalman filter learning is more efficient and adjusts more quickly than the constant gain learning that Milani uses, a finding that is in line with Sargent and Williams (2005). We make an extra effort to provide some intuition for the time variation in the belief coefficients, and we illustrate that the implied expectations under learning are more consistent with the survey evidence measured by the Survey of Professional Forecasters (SPF), than the expectations observed under RE. We document more extensively the macroeconomic implication of the learning dynamics. The impact is mainly concentrated in the inflation dynamics, and contrary to Milani, we do not observe an important effect on the role of real frictions in households' and firms' decision problems.

Assuming that agents use only a limited information set in forming expectations may be criticized for being largely arbitrary. Therefore, we conduct an extensive robustness exercise to underline that our results do not depend on a specific choice of the small forecasting model or of the initial beliefs. We also document that the out-of-sample forecast performance of the DSGE model with adaptively learning agents using small forecasting models is competitive with a RE DSGE model in which expectations are formed using a much larger forecasting model. The use of small forecasting models is important for the learning dynamics to adjust in a flexible, fast, and stable way. The empirical performance of the learning model depends on three properties: the specification of the forecasting model, the initial beliefs, and the efficient updating procedure. All three aspects are contributing to a successful fit, but, nevertheless, the results are robust for relatively minor changes on each of the three properties.

In the next section, we explain how we change the expectation assumption in the medium-scale DSGE model of Smets and Wouters (2007, henceforth SW) and how the learning mechanism is specified. Section II presents the estimation results under learning with limited information and compares these results with the RE version. In Section III, we interpret the learning dynamics and provide additional evidence in favor of our time-varying expectations mechanism. Section IV discusses how the learning dynamics affect the propagation mechanism of the shocks in the model and explain the changes in volatility and the slope of the Phillips curve in the US data over the last four decades. Finally, in Section V, an extensive sensitivity analysis is performed to show the robustness of the results.

## I. Model and Learning Dynamics

In this paper, we evaluate the potential role of adaptive learning (AL) in an estimated medium-scale DSGE model. The model that we consider in this application is the one estimated in SW for the US economy, updated with the most recent data covering 1966–2008.[2] This DSGE model, following the work of Christiano, Eichenbaum, and Evans (2005), contains many frictions that affect both nominal and real decisions of households and firms. Households maximize expected utility over an infinite horizon. Consumption appears in this utility function relative to a time-varying external habit variable. Their labor services are further differentiated by a union that sets the nominal wage according to a Calvo model. Households decide on how much capital to accumulate, given the investment adjustment cost function, and rent this capital to firms. Depending on the rental rate, the capital stock will be used more or less intensively. Firms produce differentiated goods, decide on labor and capital inputs, and set prices again according to the Calvo model. Their marginal cost depends on wages, the rental rate of capital, and the exogenous productivity process. The Calvo models for price- and wage-setting assume partial indexation to lagged inflation, so that the inflation dynamics have both a forward- and a backward-looking component. The standard Dixit-Stiglitz aggregator in both the goods and labor market is replaced by a more general aggregator, which allows for time-varying demand elasticity.

As in SW, monetary policy is described by a generalized Taylor rule with inertia in the policy reaction to inflation and the output gap. We deviate, however, from SW by defining the output gap simply as the deviation of output from its underlying neutral productivity process and not as the natural output gap. In doing so, we avoid the modelling of the flexible economy, which considerably reduces the number of forward variables on which agents have to form expectations.

The model contains 14 endogenous variables summarized by the vector $y_t$. In addition, the stochastic structure of the model is determined by seven exogenous disturbances and their innovations. Neutral and investment-specific technological progress, risk premiums, exogenous spending, and nonsystematic monetary policy actions are represented by a first-order autoregressive process, while the price and wage mark-up disturbances are modelled as an ARMA$(1,1)$ process. The vector $w_t$ represents both the seven exogenous variables and the lagged innovations $\epsilon_{t-1}$ for

[2] We refer to SW for the formal presentation of the model. The equations and corresponding coefficients are repeated in the Model Appendix.

the mark-up shocks. After linearization around the deterministic steady state, the model can be represented as follows:[3]

$$\mathbf{A}_0\begin{bmatrix}\mathbf{y}_{t-1}\\ \mathbf{w}_{t-1}\end{bmatrix} + \mathbf{A}_1\begin{bmatrix}\mathbf{y}_t\\ \mathbf{w}_t\end{bmatrix} + \mathbf{A}_2 E_t \mathbf{y}_{t+1} + \mathbf{B}_0 \boldsymbol{\epsilon}_t = const. \tag{1}$$

Under rational expectations, the solution of the model is provided by

$$\begin{bmatrix}\mathbf{y}_t\\ \mathbf{w}_t\end{bmatrix} = \boldsymbol{\mu} + \mathbf{T}\begin{bmatrix}\mathbf{y}_{t-1}\\ \mathbf{w}_{t-1}\end{bmatrix} + \mathbf{R}\boldsymbol{\epsilon}_t, \tag{2}$$

where the matrices $\mathbf{T}$ and $\mathbf{R}$ are nonlinear functions of the structural parameters of the model, $\Theta$. The intercept, $\boldsymbol{\mu}$, can be a nonzero vector under RE only for observable variables that are not demeaned. The vector $\mathbf{y}$ can be further decomposed into state variables $\mathbf{y}^s$ (those appearing with a lag), forward variables $\mathbf{y}^f$ (appearing with a lead in the model), and the so-called static variables. More specifically, in the SW model, agents have to form expectations on seven forward variables: consumption, investment, hours worked, wages, inflation, price, and return of existing capital.

In this paper, we relax the RE assumption and, following Marcet and Sargent (1989) and Evans and Honkapohja (2001), we assume that the agents forecast the values of the forward variables as a reduced-form linear function of the state variables. This type of learning, advanced by Evans and Honkapohja (2001), is called Euler equation learning; the agents forecast only immediate future variables, which are typically present in Euler equations for firms and/or consumers. An alternative description of learning—long-horizon learning—has been suggested by Marcet and Sargent (1989) and further developed by Preston (2005). Preston (2005), considers agents forecasting economic variables (present in their budget constraint and exogenous to their decision making) infinitely many periods ahead. For a theoretical discussion on these two approaches to adaptive learning, see Preston (2005) and Honkapohja, Mitra, and Evans (2002). For a discussion of effects of the learning type on the behavior of estimated DSGE models, see Milani (2006) and references therein.

We assume that agents use only a limited information set, $\mathbf{X}_j$, in their forecasting model: $\mathbf{y}_j^f = \mathbf{X}_j^T \boldsymbol{\beta}_j$. This equation is called the Perceived Law of Motion (PLM) in the adaptive learning literature. Specifically, in the baseline specification, we assume that agents use a simple univariate AR(2) model to form expectations, and therefore, for every forward variable $\mathbf{y}_j^f$, the set $\mathbf{X}_j$ contains a constant and two lags of $\mathbf{y}_j^f$. This specification of the PLM model delivers results that are representative for the outcomes under a broad set of small forecasting models (see Section V). Note that this setup deviates from RE in three substantial ways. First, the coefficients in the forecasting models are not restricted to be consistent with the decision rules of the agents. Second, the information set that is used in the forecasting models is much smaller than the state vector that would be used under RE. Finally, the coefficients of the forecasting model are updated based on new observations. A

[3] We implemented this AL estimation approach in the Dynare program, and therefore we follow the notation used by Juillard (1996). The code is available upon request from the authors.

particularly interesting deviation from the RE is the presence of a constant in the PLM specification. This constant relaxes the restriction of a common deterministic growth rate and a well defined constant inflation objective, and it allows expectations to track recent trends in the data.

The precise learning procedure is defined as follows. Agents estimate the forecasting model at each point in time given the information set available at that time. We assume that they use an efficient Kalman filter updating mechanism.[4] They believe that the coefficients $\boldsymbol{\beta}$ (a vector obtained by stacking all $\boldsymbol{\beta}_j$) follow a vector autoregressive process around $\overline{\boldsymbol{\beta}}$ (which will be specified later): $vec(\boldsymbol{\beta}_t - \overline{\boldsymbol{\beta}}) = \mathbf{F} \cdot vec(\boldsymbol{\beta}_{t-1} - \overline{\boldsymbol{\beta}}) + \mathbf{v}_t$, where $\mathbf{F}$ is a diagonal matrix with $\rho \leq 1$ on the main diagonal.[5] Errors $\mathbf{v}_t$ are assumed to be idependently and identically distributed with variance-covariance matrix $\mathbf{V}$.

We can write the forecasting model in the following SURE format:[6]

$$(3) \qquad \begin{bmatrix} \mathbf{y}^f_{1t} \\ \mathbf{y}^f_{2t} \\ \vdots \\ \mathbf{y}^f_{mt} \end{bmatrix} = \begin{bmatrix} \mathbf{X}_{1,t-1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{X}_{2,t-1} & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{X}_{m,t-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_{1,t-1} \\ \boldsymbol{\beta}_{2,t-1} \\ \vdots \\ \boldsymbol{\beta}_{m,t-1} \end{bmatrix} + \begin{bmatrix} \mathbf{u}_{1,t} \\ \mathbf{u}_{2,t} \\ \vdots \\ \mathbf{u}_{m,t} \end{bmatrix}.$$

The errors $\mathbf{u}_{j,t}$ depend on a linear combination of the true model innovations $\boldsymbol{\epsilon}_t$, and therefore they are likely to be correlated, making the variance-covariance matrix nondiagonal: $\boldsymbol{\Sigma} = E\left[\mathbf{u}_t \cdot \mathbf{u}_t^T\right]$. With the above notation, the Kalman filter updating and transition equations for the belief coefficients and the corresponding covariance matrix are given by

$$(4a) \quad \boldsymbol{\beta}_{t|t} = \boldsymbol{\beta}_{t|t-1} + \mathbf{P}_{t|t-1}\mathbf{X}_{t-1}[\boldsymbol{\Sigma} + \mathbf{X}^T_{t-1}\mathbf{P}_{t|t-1}\mathbf{X}_{t-1}]^{-1} \times \left(\mathbf{y}^f_t - \mathbf{X}^T_{t-1}\boldsymbol{\beta}_{t|t-1}\right),$$

$$\text{with } (\boldsymbol{\beta}_{t+1|t} - \overline{\boldsymbol{\beta}}) = \mathbf{F} \cdot (\boldsymbol{\beta}_{t|t} - \overline{\boldsymbol{\beta}}).$$

$$(4b) \quad \mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \mathbf{P}_{t|t-1}\mathbf{X}_{t-1}[\boldsymbol{\Sigma} + \mathbf{X}^T_{t-1}\mathbf{P}_{t|t-1}\mathbf{X}_{t-1}]^{-1} \times \mathbf{X}^T_{t-1}\mathbf{P}_{t|t-1},$$

$$\textit{with}\ \mathbf{P}_{t+1|t} = \mathbf{F} \cdot \mathbf{P}_{t|t} \cdot \mathbf{F}^T + \mathbf{V}.$$

[4] Sargent and Williams (2005) showed that even if Kalman filter and constant gain learning are asymptotically equivalent on average, their transitory behavior may differ a lot. In particular, Kalman filter tends to result in much faster adjustment of agents' beliefs. With faster adjustment of beliefs, we are able to better understand whether the initial beliefs or time-varying coefficients matter more for the improved model fit.

[5] $\rho$ is restricted to be the same for the seven variables that are forecasted. Allowing for a variable specific autocorrelation provides some extra flexibility but also larger parameter uncertainty.

[6] The SURE format and the corresponding GLS estimator are necessary to get an efficient estimator of the complete forecasting model because the variables appearing on the RHS in each equation are not identical.

These best estimates for the beliefs $(\boldsymbol{\beta}_{t|t-1})$ are then substituted for $\boldsymbol{\beta}_t$ in (3) to generate expectations of forward-looking variables, $E_t \mathbf{y}^f_{t+1}$. Plugging these expectations into (1), we obtain a purely backward-looking representation of the model[7]:

$$\begin{bmatrix} \mathbf{y}_t \\ \mathbf{w}_t \end{bmatrix} = \boldsymbol{\mu}_t + \mathbf{T}_t \begin{bmatrix} \mathbf{y}_{t-1} \\ \mathbf{w}_{t-1} \end{bmatrix} + \mathbf{R}_t \boldsymbol{\epsilon}_t. \tag{5}$$

The resultant time-dependent matrices $\boldsymbol{\mu}_t$, $\mathbf{T}_t$, and $\mathbf{R}_t$ replace the constant equivalents in the RE solution (2). These matrices depend now on both the parameters of the decision problem $(\Theta)$ and on the best estimates of the forecasting model $(\boldsymbol{\beta}_{t|t-1})$, and contain all necessary information to describe the dynamics and the propagation of the shocks in the model under learning. In terms of adaptive learning literature, equation (5) represents the Actual Law of Motion (ALM) of the model.

In order to initialize this Kalman filter for the belief coefficients, we need to specify $\boldsymbol{\beta}_{1|0} = \overline{\boldsymbol{\beta}}$, $\mathbf{P}_{1|0}$, $\boldsymbol{\Sigma}$, and $\mathbf{V}$. In our baseline approach, all these expressions are derived from the correlations between the model variables implied by the RE Equilibrium evaluated for the corresponding structural parameter vector $\Theta$. In other words, the initial beliefs are assumed to be model consistent.[8]

Using the fact that $\hat{\boldsymbol{\beta}}_{OLS} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$ is unbiased, we use the theoretical moment matrices $E[\mathbf{X}^T\mathbf{X}]$ and $E[\mathbf{X}^T\mathbf{y}]$ from the RE solution, and set $\boldsymbol{\beta}_{1|0} = (E[\mathbf{X}^T\mathbf{X}])^{-1} \cdot E[\mathbf{X}^T\mathbf{y}]$. Given $\boldsymbol{\beta}_{1|0}$, we calculate $\boldsymbol{\Sigma}$ as $\boldsymbol{\Sigma} = E[(\mathbf{y}^f_t - \mathbf{X}^T_{t-1}\boldsymbol{\beta}_{1|0}) \times (\mathbf{y}^f_t - \mathbf{X}^T_{t-1}\boldsymbol{\beta}_{1|0})^T]$, again using the RE theoretical moments. Finally $\mathbf{P}_{1|0}$, the initial guess about the mean square forecast error of the belief coefficients, and $\mathbf{V}$, the variance–covariance matrix of shocks $\mathbf{v}_t$ to these coefficients, are both taken to be proportional to $(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$:[9] $\mathbf{P}_{1|0} = \sigma_0 \cdot (\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$, and $\mathbf{V} = \sigma_v \cdot (\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$. This initialization leaves just three parameters, $\sigma_0$, $\sigma_v$, and $\rho$, to fully describe the learning dynamics.

## II. Estimation Results

In this section, we document the estimation approach, and we present the estimation results in terms of the posterior distribution of the estimated parameters and the marginal likelihood of the model. We compare the results under RE and under our baseline learning model that assumes a simple extrapolative AR(2) forecasting model.

[7] Note that we expand the state vector $y$ in this representation with additional lags that occur in the forecasting models.

[8] An alternative approach would be to derive the initial beliefs and the underlying moment matrixes from the restricted expectations equilibrium. Given our underparameterized beliefs, this equilibrium deviates from the REE and requires the solution of the underlying ODE. Computationally, this procedure was not feasible in the estimation context.

[9] $(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$ is equal to $\mathrm{Var}[\hat{\boldsymbol{\beta}}_{GLS}]$, where $\hat{\boldsymbol{\beta}}_{GLS} = (\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{y}$, which gives an efficient estimator for the SURE model. Given knowledge of theoretical moments and of $\boldsymbol{\Sigma}$, the matrix $(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$ could be readily calculated.

### A. *Estimation Approach*

As in SW (2007), the model is estimated using seven key macroeconomic US time series as observable variables: the log difference of real GDP, real consumption, real investment and the real wage, log hours worked, the log difference of the GDP deflator, and the federal funds rate. The corresponding measurement equation is:

$$O_t = \begin{bmatrix} dlGDP_t \\ dlCons_t \\ dlINV_t \\ dlWag_t \\ lHOURS_t \\ dlP_t \\ FEDFUNDS_t \end{bmatrix} = \begin{bmatrix} \bar{\gamma} \\ \bar{\gamma} \\ \bar{\gamma} \\ \bar{\gamma} \\ \bar{l} \\ \bar{\pi} \\ \bar{r} \end{bmatrix} + \begin{bmatrix} \hat{y}_t - \hat{y}_{t-1} \\ \hat{c}_t - \hat{c}_{t-1} \\ \hat{i}_t - \hat{i}_{t-1} \\ \hat{w}_t - \hat{w}_{t-1} \\ \hat{l}_t \\ \hat{\pi}_t \\ \hat{R}_t \end{bmatrix}, \tag{6}$$

where $l$ and $dl$ stand for log and log difference respectively; $\bar{\gamma} = 100(\gamma - 1)$ is the common quarterly trend growth rate to real GDP, consumption, investment, and wages; $\bar{\pi} = 100(\Pi_* - 1)$ is the quarterly steady-state inflation rate; and $\bar{r} = 100(\gamma^{\sigma_c}\Pi_*/\beta - 1)$ is the steady-state nominal interest rate. Given the estimates of the trend growth rate and the steady-state inflation rate, the latter will be determined by the estimated discount rate. Finally, $\bar{l}$ is steady-state hours worked. The model is estimated over the sample period from 1966:I until 2008:IV.

As is typical in the DSGE estimation literature, we are using revised data, taken from the latest available data release. There could be large discrepancies between this data and the data available to the agents at the time their decisions were made. For an extensive introduction to the real-time data issue, see Croushore (2011). Real-time data would be the natural environment for analyzing learning dynamics, also because survey evidence on expectations is by definition real-time information. While believing that estimating our model using the real-time data would be appropriate, we note that implementation of such a project requires careful modeling of the data revision process. We are currently working on the implementation of such a fully integrated approach, and present in this paper only some preliminary evidence when comparing the estimated belief dynamics with survey data.

The Bayesian estimation method proceeds in two steps. First, we estimate the mode of the posterior distribution by maximizing the log posterior function, which combines the prior information on the parameters with the likelihood of the data. In a second step, the Metropolis-Hastings algorithm is used to get a complete picture of the posterior distribution and to evaluate the marginal likelihood of the model. The prior assumptions on the structural parameters are exactly the same as in Smets and Wouters (2007). For the learning parameters, it turned out that the three parameters are not simultaneously identified by the estimation procedure. In the baseline version, we estimate $\rho$ using a uniform prior over the interval $[0, 1]$, and we fix the

Table 1—Posterior Estimates for RE and AL Models

| | RE | AL | AL (baseline) | AL |
|---|---|---|---|---|
| | | ARMA markup | i.i.d. markup | No updating |
| Selection of parameter estimates (mean and 90 percent interval of the posterior distribution) | | | | |
| Wage markup AR(1) | 0.96 | 0.53 | — | — |
| | [0.93–0.99] | [0.28–0.79] | | |
| Wage markup MA(1) | 0.88 | 0.43 | — | — |
| | [0.81–0.95] | [0.13–0.73] | | |
| Price markup AR(1) | 0.85 | 0.28 | — | — |
| | [0.75–0.94] | [0.06–0.49] | | |
| Price markup MA(1) | 0.70 | 0.48 | — | — |
| | [0.55–0.86] | [0.30–0.66] | | |
| Wage indexation | 0.51 | 0.18 | 0.21 | 0.37 |
| | [0.30–0.71] | [0.07–0.29] | [0.09–0.32] | [0.19–0.54] |
| Price indexation | 0.24 | 0.29 | 0.19 | 0.34 |
| | [0.10–0.38] | [0.11–0.45] | [0.08–0.29] | [0.15–0.53] |
| Wage stickiness | 0.77 | 0.83 | 0.84 | 0.84 |
| | [0.68–0.85] | [0.78–0.88] | [0.80–0.88] | [0.79–0.89] |
| Price stickiness | 0.72 | 0.66 | 0.65 | 0.68 |
| | [0.65–0.80] | [0.61–0.72] | [0.59–0.71] | [0.61–0.75] |
| Marg.lik. | −1,005.22 | −982.23 | −983.86 | −999.76 |

scale of the initial variance-covariance matrix of the belief coefficients to some large number ($\sigma_0 = 0.03$) relative to the variance–covariance of the shocks to the beliefs ($\sigma_v = 0.003$).

The likelihood of the model is evaluated with a Kalman filter that uses equation (5) as the state equation and (6) as the measurement equation. At each moment in time, the beliefs of the agents on their forecasting model are updated using (4), which delivers the necessary input to solve for next period $\mu_t$, $\mathbf{T}_t$, and $\mathbf{R}_t$. The forecasting model is evaluated and updated in terms of the filtered states of the overall model.[10]

## B. *Posterior Estimates*

Table 1 summarizes the marginal likelihood and the posterior estimates for a selection of the parameters under RE and under AL with small forecasting models. The first observation is the noticeable improvement in the marginal likelihood for the model under AL (−982.23) versus RE (−1005.22), which implies that the posterior odds ratio is definitely in favor of the learning approach. We thus confirm results of Milani (2007), who also finds marginal likelihood to be in favor of a model with adaptive learning. This result is confirmed if we allow agents to combine the two forecasting models, AL and RE, explicitly. The posterior distribution of this

[10] Note that not all forward variables are observed, while others are observed in first differences but appear in the forecasting models in levels or, more precisely, in log deviations from the steady-state growth path.

Table 2—Variance Decomposition for RE and AL Models

| | RE | | | AL (baseline), average | | |
|---|---|---|---|---|---|---|
| Variance decomposition | Inflation | Output growth | Output | Inflation | Output growth | Output |
| Productivity | 8.20 | 9.26 | 18.54 | 18.71 | 12.32 | 62.40 |
| Risk premium | 0.27 | 34.58 | 4.13 | 1.44 | 59.14 | 23.26 |
| Exogenous demand | 2.41 | 25.42 | 9.86 | 0.07 | 17.50 | 2.15 |
| Investment specific technology | 0.67 | 14.57 | 12.00 | 0.71 | 4.01 | 0.80 |
| Monetary policy | 2.45 | 7.72 | 5.84 | 0.72 | 5.35 | 8.38 |
| Price markup | 29.76 | 3.97 | 7.03 | 59.96 | 1.11 | 0.77 |
| Wage markup | 56.24 | 4.47 | 42.61 | 18.39 | 0.58 | 2.23 |

encompassing model gives a dominant weight to the AL forecasts (90 percent interval between 0.86 and 0.97).

Secondly, there are some important changes in the estimated stochastic structure depending on the retained assumption about the expectations. Especially the process of the price and wage mark-up shocks changes quite dramatically. While these shocks were estimated as an ARMA$(1,1)$ process with a very persistent component under RE, these shocks follow basically an independently and identically distributed process under learning. With both the autoregressive and the moving average coefficient close to the mean of the prior distribution centered around 0.5, the implied dynamics for these processes are equivalent to a white noise process. This interpretation is confirmed by re-estimating a version of the learning model in which the two mark-up shocks are explicitly defined as independently and identically distributed processes. The marginal likelihood of the model hardly changes ($-983.86$). It appears that the propagation of the mark-up shocks under learning is completely captured by the expectations mechanism and by the internal dynamics of the decision rules, while it was dependent on the persistence in the exogenous dynamics of the mark-up process under RE.[11] This learning model with independent and identically distributed price and wage mark-up shocks is considered as the baseline AL model in the rest of the paper.

The change in the persistence of the shocks has some interesting consequences for the implied variance decomposition for both inflation and output. Table 2 illustrates these effects. A first observation is that the relative contribution of the price mark-up shock to inflation volatility increases at the cost of the wage mark-up shock. Rational agents understand that wage mark-up shocks have a very persistent effect on the marginal cost, while price mark-up shocks are expected to reverse relatively quickly. But under learning and our specification of the belief equations, agents do not distinguish between the persistence of the two shocks, and therefore they tend to underestimate the impact of the wage shocks and overestimate the price mark-up shocks. On average over the sample, the two mark-up shocks together generate a similar amount of inflation volatility under learning and under RE, but their impact varies strongly over time as we will discuss in the next section. On the other hand, the impact of the mark-up shocks, and in particular the wage mark-up shock, on the

[11] Note that these results are not driven by the fact that two lags are used for forecasting. Section V describes an example where inflation and price of capital are just AR$(1)$, delivering an essentially identical outcome.

real economy is substantially lower under learning. The general equilibrium effects of a negative supply shock, operating through expectations of future inflation and demand, are realized much more slowly and incompletely. In particular, aggregate demand resists the downward adjustment because monetary policy needs to respond less aggressively as inflation reacts less as well.

The same mechanism of slowly adjusting expectations works for TFP and investment technology shocks. Under learning, the TFP shock comes out as more persistent, which explains why its contribution to the output variance increases. The investment shock, however, becomes less important in the variance decomposition because it gets less persistence under learning, which is probably related to the fact that under learning it generates a negative correlation between investment and consumption, consumption not being supported by the wealth effect as under RE. In addition, consumption under learning is suppressed through longer interest rate responses. Similarly to TFP, the risk premium shock, that is the shock to the interest rate at which households can access the financial market, is estimated to be more persistent, and in addition its impact under learning is stronger because learning agents do not anticipate future restrictive policy reactions and the resulting real rate increases.

We note that the impact of adaptive learning on the estimated exogenous processes is stonger for those shocks that represent deviations from intertemporal conditions, and which are therefore directly related to forward-looking variables. For example, the risk premium shock is present in the consumption Euler equation and in the cost of capital equation, while price and wage mark-up shocks affect, respectively, the NK Phillips curve for price and wage setting. Expectations play a crucial role in these equations.

A third observation that follows from the estimation outcomes is that the structural parameters that govern the decision rules of the agents remain relatively robust under alternative hypotheses on the expectations. The estimated mode of several parameters changes, but the posterior intervals under RE and under AL do overlap considerably in most cases. The most significant change is observed for the wage indexation parameter which drops from 0.51 under RE to 0.21 under learning. Note that the indexation for price-setting is estimated consistently low at a level of 0.20 in both cases. This result points in the same direction as the change in the mark-up shocks, namely that the learning dynamics provide a strong propagation mechanism for the mark-up shocks so that alternative frictions, such as the indexation mechanism, are less crucial to explain the observed persistence in the inflation process. The estimated price and wage stickiness, however, are not significantly different across models.

These results can be considered as an interesting achievement because empirical DSGE models are often criticized for depending too much on the dynamics incorporated in the exogenous disturbances, which leaves only a minor role for the endogenous propagation mechanism. In this context, the reduced role of the wage mark-up shock in the model with learning is especially welcome given that this shock has been questioned as being too large and having a dubious interpretation (Chari, Kehoe, and McGrattan 2009). Learning dynamics seem to substitute for the exogenous persistence but leave the structural parameters of the model unaffected. It is also important to note that the identified historical innovations to the exogenous processes are very similar under RE and under AL. The correlation between the estimates of these independently and identically distributed innovation series varies from 0.83 for the price mark-up shock to 0.98 for the productivity shock. Notice that the productivity shock does not affect, at least directly, any model equation in which an expectation is present, which leads to almost identical identification of its innovation under both AL and RE.

Finally, note that the estimated persistence $(\rho)$ in the belief coefficients is estimated to be quite high $(0.96/0.97)$. In the sensitivity analysis, we will consider alternative assumptions about the parameters driving the learning dynamics.

As mentioned before, our learning model deviates from RE by considering a smaller information set and by allowing for time-variation in the beliefs based on the latest observations. To identify which of these two deviations is more important for explaining the improved fit of the model, we consider a model in which agents use the same small forecasting model but with constant belief coefficients fixed at the RE-implied initial coefficients $(\boldsymbol{\beta}_{1|0})$. This model produces a marginal likelihood of $-999.76$, which is still an improvement on the RE-model. The model with constant expectations based on a limited information set improves the model fit compared to the fully model consistent expectations case. But updating of the expectations through the KF-learning process is more important for improving the fit. Learning explains roughly three-quarters of the overall improvement in the marginal likelihood. This important contribution of the learning dynamics is independent of the specific choice of the initial beliefs. Note that we do not optimize over the initial beliefs, but instead derive them from the RE-consistent version of the model. Optimizing over initial beliefs further improves the overall fit of the model as discussed in the robustness exercise in Section V.

### C. *Improved Forecasting Performance of Inflation Dynamics*

The good fit of the AL model is also reflected in the out-of-sample prediction performance. Table 3 compares the root mean squared forecast error at different forecast horizons under RE and under the baseline AL model. The overall forecast performance of the AL model is better at a one-quarter forecast horizon, which is consistent with the higher marginal likelihood. The biggest gain is realized in the forecast of inflation which improves by 9 percent relative to the RE model. Given that the learning dynamics are most active for the inflation expectations, this is a promising result for the AL approach. On the other hand, the forecasts for consumption and real wages are slightly worse under AL even for one-period-ahead forecasts. The forecast performance of the AL model generally deteriorates at longer horizons for all variables but wages. This result applies independently of the exact procedure that is used to construct the long-term forecasts.[12] Longer-horizon forecasts for consumption and the interest rate deteriorate the most. For these variables,

[12] The results reported in Table 2 are produced under the hypothesis that the ALM process, exemplified by (5), remains constant over the forecast horizon. Minor improvements are obtained if the ALM is updated over the forecast horizon in line with the perceived mean reversion in the belief coefficients as captured by the estimated $\rho$ parameter.

Table 3—Out-of-Sample Prediction Performance of RE and AL Models

| | GDP | $\pi$ | $R$ | Hours | Wage | Cons. | Inv. | Overall |
|---|---|---|---|---|---|---|---|---|
| RMSE of the RE-model at different forecast horizons | | | | | | | | |
| 1q | 0.65 | 0.27 | 0.13 | 0.52 | 0.63 | 0.60 | 1.66 | −12.13 |
| 2q | 1.01 | 0.27 | 0.23 | 0.85 | 0.99 | 0.88 | 2.85 | −7.15 |
| 4q | 1.63 | 0.23 | 0.40 | 1.35 | 1.51 | 1.15 | 5.19 | −3.06 |
| 8q | 2.62 | 0.23 | 0.54 | 2.00 | 2.63 | 1.73 | 8.53 | 1.13 |
| 12q | 3.11 | 0.23 | 0.56 | 2.26 | 3.73 | 2.06 | 9.49 | 2.21 |
| Percentage gains (+) or losses (−) of AL relative to RE model | | | | | | | | |
| 1q | 5.08 | 8.93 | 3.96 | 2.65 | −2.13 | −2.60 | 6.20 | 1.56 |
| 2q | 2.08 | 1.79 | −3.24 | −3.51 | −2.19 | −11.44 | 5.68 | −1.17 |
| 4q | 3.95 | −14.62 | −8.09 | −5.97 | −4.17 | −19.08 | 4.64 | −3.42 |
| 8q | 6.83 | −4.38 | −23.45 | −8.46 | 1.63 | −25.00 | 0.27 | −6.45 |
| 12q | 0.25 | −5.14 | −38.53 | −12.03 | 4.95 | −42.22 | −8.09 | −10.9 |

*Notes:* The forecast period is 1990:I–2008:IV. Models are re-estimated each year, starting in 1966:I and ending in the fourth quarter of the year before the forecast starts. The overall measure of forecast performance is the log determinant of the uncentered forecast error covariance matrix. Gains and losses in the overall measure are expressed as the difference in the overall measure divided by the number of variables and divided by two to convert the variance to standard errors (times 100).

the restrictions imposed by the RE assumption appear to be very useful in longer-horizon forecasts. This suggests that there is a trade-off between the flexibility of the learning dynamics in capturing the short-run behavior on the one hand, and the RE restrictions that are useful for the long-run forecasts on the other hand. In this context, it would be very interesting to test whether Preston's infinite-horizon learning approach (2005) can resolve this trade-off more efficiently by using the restrictions from the budget constraints in the forecasts.

## III. Expectations Implied by the Learning Model

In this section, we discuss in detail the behavior of the expectations implied by the time-varying small forecasting or PLM model. First, the time variation in the belief coefficients for the various forward variables is illustrated graphically. This time variation or updating of the coefficients is driven by the forecasting errors and has a simple intuitive interpretation. In particular, we illustrate how the updating of the inflation belief coefficients responds to the innovations in the exogenous disturbances. Secondly, we evaluate the quality of the forecasts implied by these small forecasting models by comparing the RMSE with alternative forecasting models, and by comparing the implied forecasts with empirical evidence on these expectations as measured by the Survey of Professional Forecasters. This evidence confirms that the small forecasting models provide a plausible description of the way expectations are formed by agents in the real world.

### A. *Time Variation in the Beliefs*

KF learning leads to important time variation in the coefficients of the forecasting model. Figure 1 illustrates the time variation in the coefficients of the AR(2) forecasting model for four of the seven forward variables in the model. To facilitate the

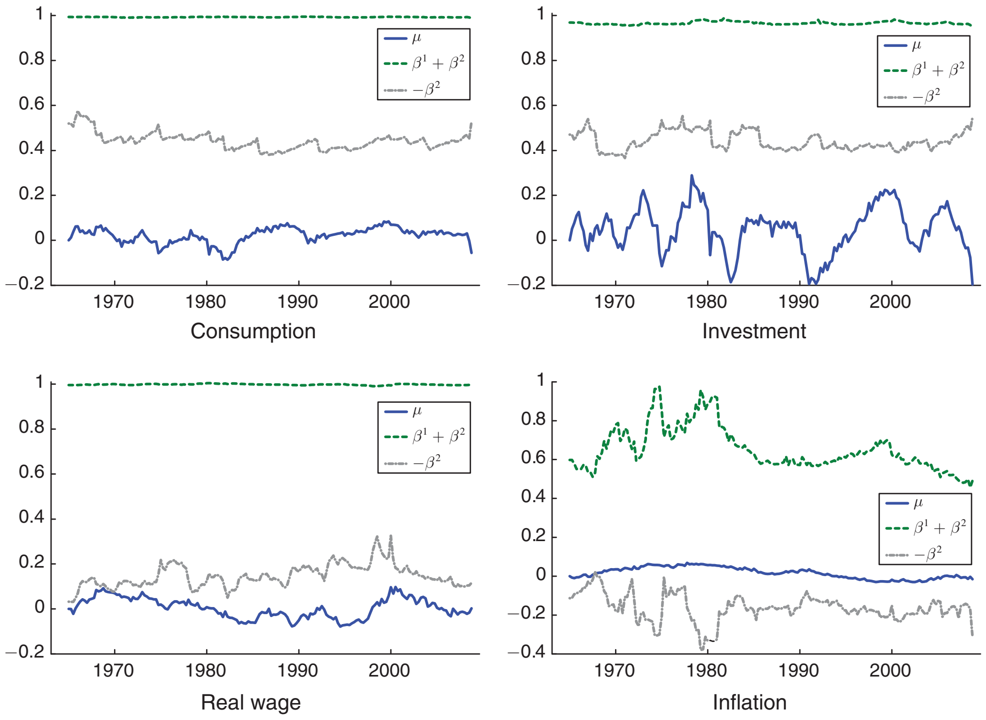


Figure 1. Time Variation in the Belief Coefficients of the AR(2) Forecasting Model

interpretation of the time variation, we can rewrite the AR(2) forecasting equation in the following format:

$$\mathbf{y}^f_{j,t} = \mu_{j,t} + (\beta^1_{j,t} + \beta^2_{j,t}) \times \mathbf{y}^f_{j,t-1} - \beta^2_{j,t} \times \Delta \mathbf{y}^f_{j,t-1}.$$

Figure 1 provides the evidence on the constant $(\mu_{j,t})$, the persistence in the expectations $(\beta^1_{j,t} + \beta^2_{j,t})$, and $(-\beta^2_{j,t})$, which we refer to as the persistence in the growth rate of the expectations.

First, it is clear that the constants vary a lot for all four forward variables. Under RE, these constants would be zero, reflecting the fact that all real variables are modelled as deviations from a constant deterministic growth rate and that inflation and the nominal rate fluctuate around a constant inflation objective and a corresponding nominal rate as determined by the monetary policy reaction function. By allowing for a time-varying constant in the belief models, these restrictions are relaxed. The perceived trend growth rate and inflation objective can vary now over time and across variables as a function of the observed trends in the recent history. The constants in the real variables fluctuate over the cycle reflecting the past growth rates observed in each of the individual variables. Clearly, the constant for the expected investment rate is the most cyclical, while the constants for consumption and real wages more closely reflect the long-term growth rates in these variables which deviate quite persistently from the imposed common productivity growth rate in the model. For inflation, the constant also reflects the trend in the past observed inflation rate. The constant term in the inflation beliefs rose during the 1970s and started to decline only slowly after the disinflation of the early 1980s. The coefficient has stabilized around zero since the mid-1990s, meaning that the expected mean inflation of the private agents has converged to the constant inflation objective of the central bank since then. Note that the perceived long-run inflation rate that is implied by these belief equations does not only depend on the constant term, but also on the perceived persistence.

Second, the perceived persistence, as measured by the sum of the AR(2) coefficients, is stable and close to one for the real variables. These coefficients suggest that the true data-generating processes for these expectations are close to an AR(1) in first differences. The more interesting coefficient is therefore the persistence in the growth rate (measured here by $-\beta^2_{j,t}$), which is clearly positive for the real variables, and, in the case of consumption, it was slightly higher during the 1970s, but declining later on.

The most important updating dynamics seem to be taking place in the perceived inflation persistence. This process followed a clear upward trend in the 1960s and 1970s with a peak around the mid-1970s and again around 1980, followed by a quick decline toward a level of 0.6 since the mid-1980s, and a further downward shift in the most recent period.[13] Movements in the inflation rate were perceived as much more persistent in the 1970s then they were during the 1960s or the more recent period. These updating dynamics for the perceived inflation persistence correspond with the statistical properties of the observed inflation process over this period. For instance, Cogley, Primiceri, and Sargent (2010) obtain a very similar pattern for the persistence in the inflation gap. In the following sections, we discuss, in detail, how this perceived persistence affects the impulse responses of various shocks, and how they can be helpful in understanding the great inflation in the 1970s and the moderation afterward. These estimates confirm the general observation that monetary policy and the inflation target of the central bank have become much more credible over the last decades.

This time variation in the beliefs of private agents has a relatively simple and intuitive interpretation. The updating expression in (4) states that the updating in the beliefs is determined by the forecast errors multiplied by a Kalman gain matrix. These expressions are very general and complex because the forecasting model takes the form of a SURE model that treats all forward variables jointly, and because the Kalman gain matrix itself depends on updates of the second-moment matrices. The outcome of this mechanism nevertheless has a straightforward interpretation. For the constant terms, this simply means that higher (lower) than expected realizations of a forward variable result in upward (downward) revisions in the constant of the forecasting equation. The updating of the persistence, which is especially relevant for

[13] The extremely high perceived inflation persistence in the mid- and late 1970s also explains why the updating in the beliefs during these years sometimes leads to explosive AR(2) beliefs. As is standard in the learning literature, the projection facility in our estimation process eliminates updates in the beliefs that would result in unstable dynamics for the model's ALM.

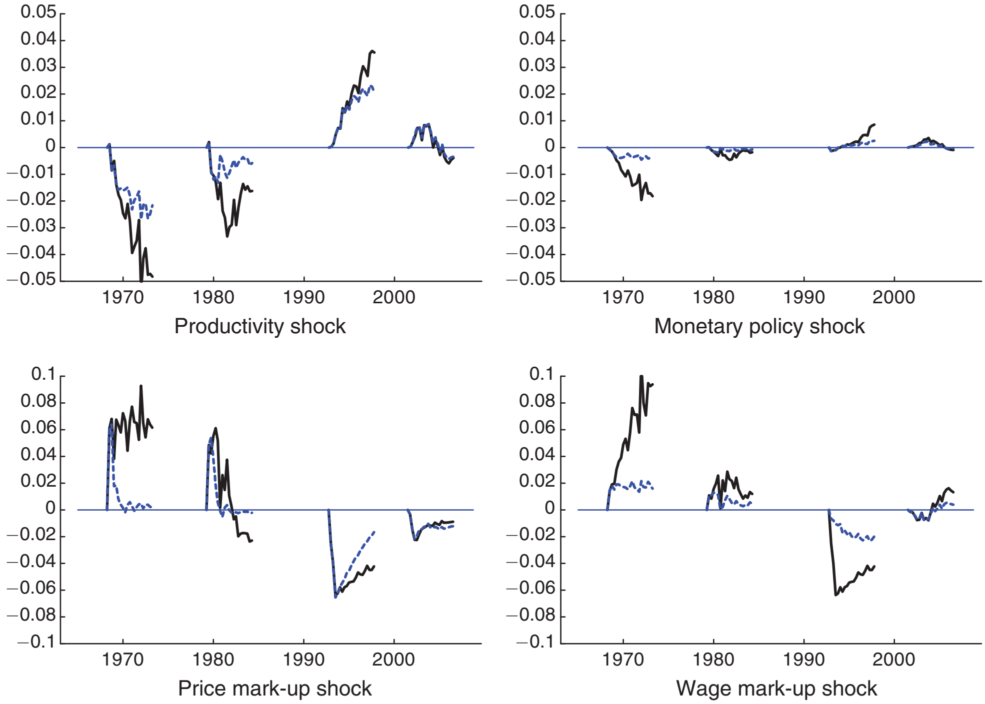


Figure 2. Impulse Response Function of Perceived Inflation Persistence

the inflation beliefs, follows a slightly more complicated logic because of its state-dependent nature. When inflation is high relative to its long-run mean value, agents will expect inflation to decline in the future, with the speed of convergence depending on the perceived persistence. With a new realization of inflation that is higher than expected, agents will tend to revise their perceived persistence upward in order to avoid underestimation in the future. With a new realization that is lower than expected, the perceived persistence will be adjusted downward. So, in periods when the inflation rate is high, revisions in persistence are positively correlated with the inflation innovations. This interpretation explains why the perceived inflation persistence was rising in the seventies and quickly declining in the early 1980s. But the opposite relation applies when inflation is at a relatively low level: when a positive inflation innovation is realized in a low inflation state, the perceived inflation persistence will decline as inflation seems to rise towards its long-run target faster than previously expected. This negative relationship between inflation realizations and revisions in the persistence is relevant in the more recent periods with low inflation rates.

To further illustrate this updating process, we show in Figure 2 the reaction of the perceived inflation persistence to various structural innovations. The figure shows how the perceived inflation persistence reacts to four types of shocks (one standard error innovation to productivity, monetary policy, price markup, and wage markup shock) during four different moments in our sample (1968:I, 1979:III, 1993:I, and 2001:IV). For each observation, we plot the IRF of the perceived inflation persistence during five years following the shock. Two types of IRFs are calculated: the dotted line represents the updates in the belief coefficients that are implied by the new realizations of the forward variables, but disregarding the feedback effect of these belief updates on the actual law of motion of the economy, while the full line takes into account this feedback effects of the beliefs on $\boldsymbol{\mu}_t$, $\mathbf{T}_t$, and $\mathbf{R}_t$. Several observations can be noted. First, the sign of the belief updates changes over time in line with the above discussion. Second, beliefs are very sensitive to innovations in the shocks especially in the first part of our sample, but this sensitivity decreases during the more recent period when actual inflation is more in line with long-term expectations. Third, perceived inflation persistence is mainly driven by innovations to the price markup, which generates mainly short-run volatility in the beliefs, and by innovations to the wage markup and productivity process which have a more gradual but also more persistent effect on the beliefs. Finally, monetary policy shocks have almost no direct impact on the perceived inflation persistence, represented by the dotted line, but they had some indirect effects through the feedback effects at least during the first subperiod. Monetary policy affects inflation only through the marginal costs and aggregate economic activity, and expectations and beliefs underlying these variables contribute to the overall reaction of inflation and inflation expectations.

A complete decomposition of the belief dynamics in terms of historical shocks is not possible because of the highly nonlinear nature of the updating process. Nevertheless, the impulse response exercise, as presented in Figure 2, suggests that the upward trend in perceived inflation in the 1970s was mostly driven by positive price and wage markup and negative productivity shocks. The series of accommodating monetary policy shocks during the great inflation period also contributed significantly to the destabilization of inflation beliefs, but it is not possible to quantify this contribution exactly. The role of the restrictive monetary policy innovations on the inflation beliefs during the Volker disinflation process is even harder to assess. Our learning model implies a relatively quick decline in the perceived inflation persistence during this period, following the sharp decline in actual inflation realizations. Monetary policy shocks alone are not able to generate such a quick response in our updating mechanism, which suggests that other shocks, and price mark-up shocks in particular, were active during this period as well.

### B. *More Evidence on the Expectations Model*

In order to further justify the choice of the small forecasting models as an approximation for the expectations of the private sector, it is useful to analyze in detail the quality and properties of the implied forecasts. Two types of evidence are presented in this section. First, we evaluate the forecast performance of the small PLM model relative to RE model and to the actual law of motion (ALM) of the learning model. It would be difficult to maintain the hypothesis that our small forecasting model provides a reasonable approximation for actual historical expectations if the forecasting performance of this model is much worse compared to alternative forecasting models. Secondly, we compare the forecasts implied by the small PLM models and the RE model with survey evidence on actual historical forecasts of economic agents. In principle, this type of evidence enables the historical relevance of alternative expectation models to be tested directly.

Table 4—RMSE Comparison of the PLM Forecasts

| | $\pi$ | Hours | Wage | Cons. | Inv. |
|---|---|---|---|---|---|
| PLM | 0.29 | 0.60 | 0.55 | 0.72 | 1.90 |
| | (0.22) | (0.48) | (0.63) | (0.60) | (1.57) |
| RE | 0.28 | 0.55 | 0.53 | 0.62 | 1.85 |
| | (0.24) | (0.48) | (0.62) | (0.57) | (1.57) |
| ALM | 0.28 | 0.55 | 0.54 | 0.72 | 1.84 |
| | (0.22) | (0.49) | (0.63) | (0.61) | (1.53) |
| PLM–no update | 0.30 | 0.60 | 0.55 | 0.77 | 1.86 |
| | (0.25) | (0.48) | (0.63) | (0.63) | (1.51) |

*Notes:* The first figure stands for the RMSE over the complete estimation sample 1966:I–2008:IV, the figure in parentheses applies for the period 1990:I–2008:IV, which is the same period considered in the out-of-sample forecasts in Table 3.

Table 4 summarizes the RMSE statistics from the PLM forecasts for the forward variables. We report results for the five forward variables, out of seven, that are also included in the list of observed variables. To assess this performance, similar statistics are provided for the model under RE, in which expectations are model-consistent and the PLM and the ALM yield the same forecast. We also report the RMSE for the ALM under learning, and for the PLM of the AR(2) model without updating, the same model that was considered in Table 1, with the AR(2) coefficients fixed at the initial beliefs. The statistics are all based on in-sample forecasts, in contrast to Table 3 where out-of-sample forecasts were considered, because here we want to evaluate the historical forecasts of the PLM model that are implicit in the learning model when estimated over the full sample. Only one-quarter-ahead forecasts are considered because these are the relevant ones in the Euler equation learning approach.

The RE and ALM in-sample forecast errors are minimized in the estimation procedure, and therefore they provide a minimum bound against which the PLM performance can be evaluated. Overall, the AR(2)-based PLM does seem to perform reasonably well compared to these optimized predictors. The maximum loss for the small model forecasts relative to RE model is realized for consumption, with a deterioration of 15 percent in the RMSE statistic. For inflation, the small-model forecasts are only slightly worse than the benchmark. Note also that the PLM forecasts perform almost as well as the RE forecasts for the more recent sample. Updating of the belief coefficients in the AR(2) forecasting model has only a minor impact on the actual forecast performance.

A reasonable performance of the expectation models in terms of RMSE is preferable for obtaining a good overall fit of the model, but provides only indirect evidence on the empirical validity of these expectations. To test this last objective, it is more useful to compare the forecast errors in our expectation models with direct evidence on these forecast errors as collected in empirical surveys. The Survey of Professional Forecasters (SPF) contains information about private sector expectations on future outcomes for inflation (GDP deflator), consumption and investment, which are three variables that are also present in our list of forward variables. This survey evidence is therefore directly relevant to evaluate our expectation model.

Table 5—Correlation of Forecast Errors with SPF Data

| | Inflation | | Consumption | | Investment | |
|---|---|---|---|---|---|---|
| | PLM | RE | PLM | RE | PLM | RE |
| Complete sample | 0.53 | 0.48 | 0.56 | 0.52 | 0.77 | 0.71 |
| 1969:I–1980:IV | 0.47 | 0.54 | — | — | — | — |
| 1981:IV–1990:IV | 0.44 | 0.34 | 0.66 | 0.59 | 0.71 | 0.62 |
| 1991:I–2000:IV | 0.33 | 0.04 | 0.38 | 0.41 | 0.79 | 0.75 |
| 2001:IV–2008:IV | 0.71 | 0.61 | 0.56 | 0.52 | 0.81 | 0.76 |

A problem that complicates any direct comparison of the forecasts is related to the fact that the surveys are based on real-time data, while our PLM forecasts are based on ex-post revised data. To overcome this problem, we concentrate on the forecast errors rather than on the forecast levels, and we focus on systematic bias during prolonged periods rather than on the period-by-period forecast performance. To construct the forecast errors, we use the survey forecast at time $t$ for changes in the variable between $t + 1$ and $t$, and calculate the prediction error as the difference between these expected changes and changes realized in the next quarter (based on the first data release as reported in the SPF database). Table 5 compares the correlation between these forecast errors in the surveys and the forecast errors realized by our PLM and RE model. Clearly, the correlation of the PLM forecast errors with the survey errors is higher than for the RE forecast errors. The difference in correlation is most pronounced for inflation and especially during the 1990s.[14]

To further illustrate the relevance of this difference, Figure 3 plots the cumulative forecast errors for inflation observed in the SPF data and in the PLM and RE model. All three forecasts display similar systematic bias over different subperiods in our sample: forecasters were surprised by the general increase in inflation in the 1970s and the disinflation in the early 1980s resulting in cumulated negative forecast errors first and a reversion in the trend later on. However, it appears very clearly from this picture that expectations in the survey and in the small PLM models tend to overestimate inflation systematically during the 1990s, whereas the RE model underestimates inflation systematically during this period. This evidence suggests that the expectations in the PLM are more in line with the survey evidence during the 1990s, which is exactly the period during which the learning model outperforms the RE model in predicting inflation (see Table 3) and in terms of overall likelihood. We consider these results as favorable evidence for the superiority of the small model PLM hypothesis relative to the RE-based expectations. This result also suggests that

[14] There are a number of other papers that considered survey expectations to be an outcome of some adaptive learning process. In Orphanides and Williams (2005b) agents use a three-variable VAR with constant gain Recursive Least Squares to form expectations of inflation, the unemployment rate, and the federal funds rate. They fit the model to the SPF expectations and conclude that geometrically discounting past data with a rate of 1–2 percent per quarter produces forecasts close to those found in SPF. Branch and Evans (2006) consider both constant gain and Kalman filter adaptive learning and find that constant gain learning with gain 0.0345 fits the SPF expectations on GDP growth and GDP deflator inflation the best. Nunes (2009) considers inflation expectations to be a weighted average of rational expectations and SPF forecasts that could be represented as an outcome of a simple constant gain learning process.

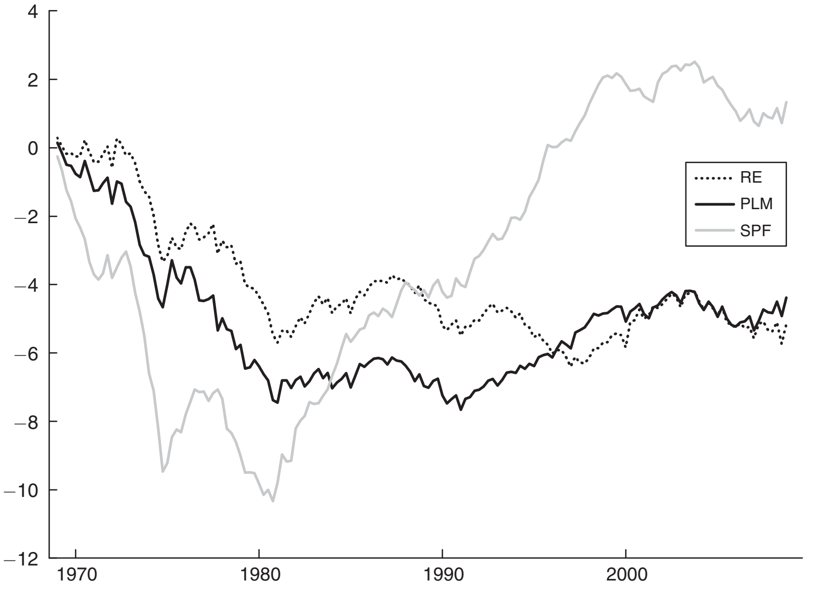


Figure 3. Cumulative Forecast Errors for Inflation

the overall improvement in the marginal likelihood of the learning model is based on a more realistic modelling of private sector expectations.

To understand the divergence of PLM and RE inflation forecasts during the 1990s, we note that inflation persistence implied by PLM beliefs was very low and close to its steady state value in this interval. Adjustment of beliefs in this situation takes place mostly through the constant $\mu_t$. The constant was indeed falling rapidly, following a relatively fast disinflation observed during the 1990s. Naturally, adaptive expectations are lagging actual inflation experience, which leads to overprediction during this time interval. Analyzing reasons for systematic underprediction of RE expectations is harder because they depend on a large number of model variables. One notices, however, that while inflation was generally falling, marginal costs were dropping even faster until the mid-1990s, and started to grow at about the same time RE expectations changed the direction of the bias. Falling marginal cost together with the relative high elasticity of inflation to marginal costs under RE explains the underprediction of inflation during this period.

Finally, we run an exercise in which we combined the estimated AR(2) belief coefficients with the real-time data in order to evaluate the implied expectations. Therefore, we apply the AL-PLM belief coefficients (given in the lower right panel of Figure 1) on real-time inflation data, using first available data releases, and generate the one-period ahead forecasts using this time-varying AR(2) model. This forecast is compared to the actual real-time inflation realisation and the SPF one-quarter ahead inflation forecast, using the same subintervals as in the Table 5. For comparison, we also repeat this exercise using the fixed-coefficient AR(2) process for inflation dynamics that is implied by the rational expectations model (column 1 of Table 1). As is clear from Table 6, the "real-time" PLM inflation forecasts of the

Table 6—Correlation of Real-Time Inflation and SPF Inflation Expectations with Univariate Inflation Forecasts over Subsamples

| | Real-time inflation (first release) | | SPF inflation expectations | |
|---|---|---|---|---|
| Correlation with | PLM | RE | PLM | RE |
| Complete sample | 0.84 | 0.83 | 0.91 | 0.89 |
| 1969:I–1980:IV | 0.72 | 0.73 | 0.84 | 0.81 |
| 1981:IV–1990:IV | 0.70 | 0.65 | 0.93 | 0.87 |
| 1991:I–2000:IV | 0.50 | 0.48 | 0.68 | 0.56 |
| 2001:IV–2008:IV | 0.09 | –0.01 | 0.50 | 0.35 |

AL model display a higher correlation with SPF expectations than “AR(2)-PLM-like” expectations implied by the RE model. The difference is especially pronounced after 1980, consistent with the results for inflation given in the Table 5. The adaptive learning-based expectations are also marginally better than rational expectations-derived beliefs in forecasting the actual real-time inflation realizations.

## IV. Macrodynamics Implied by the Learning Process

In this section, we discuss the implications of the time-varying beliefs for the overall dynamics of the model. All results presented in this section take the time-varying PLM ($\beta_t$) and ALM processes ($\boldsymbol{\mu}_t$, $\mathbf{T}_t$, and $\mathbf{R}_t$) as given and their consequences for various properties of the overall model are assessed.

### A. *Time Variation in the Impulse Response Functions*

The mechanism for the transmission of structural shocks depends crucially on the way private agents form their expectations. Therefore, it is interesting to illustrate how the IR functions depend and vary over time depending on the belief coefficients. In Figure 4, we plot the time-varying IRF for the productivity shock, the risk premium shock, the monetary policy shock, and the wage mark-up shock for the baseline learning model with beliefs based on the AR(2) model. Only the effects on output and inflation are shown. The IRFs reported here are calculated for fixed belief coefficients (and corresponding ALM matrices) at each point in time and disregard the updating of these beliefs that might be caused by the shock, as discussed in the Section IVA. In doing so, these pseudo-IRFs might underestimate the persistence and the magnitude of the actual responses. We add the corresponding IRF in the RE model at the end of the sample for comparison.

For all shocks, the reaction of inflation depends crucially on the perceived persistence of inflation by the private agents. Inflation reacted much more strongly and persistently to the shocks in the 1970s when inflation was perceived as very persistent. During the periods when perceived inflation persistence was more moderate, the reaction of inflation was also much smaller and more gradual. This profile of the inflation response under learning with small PLM models contrasts sharply with the typical response under RE. The impact of the shocks on output displays less time

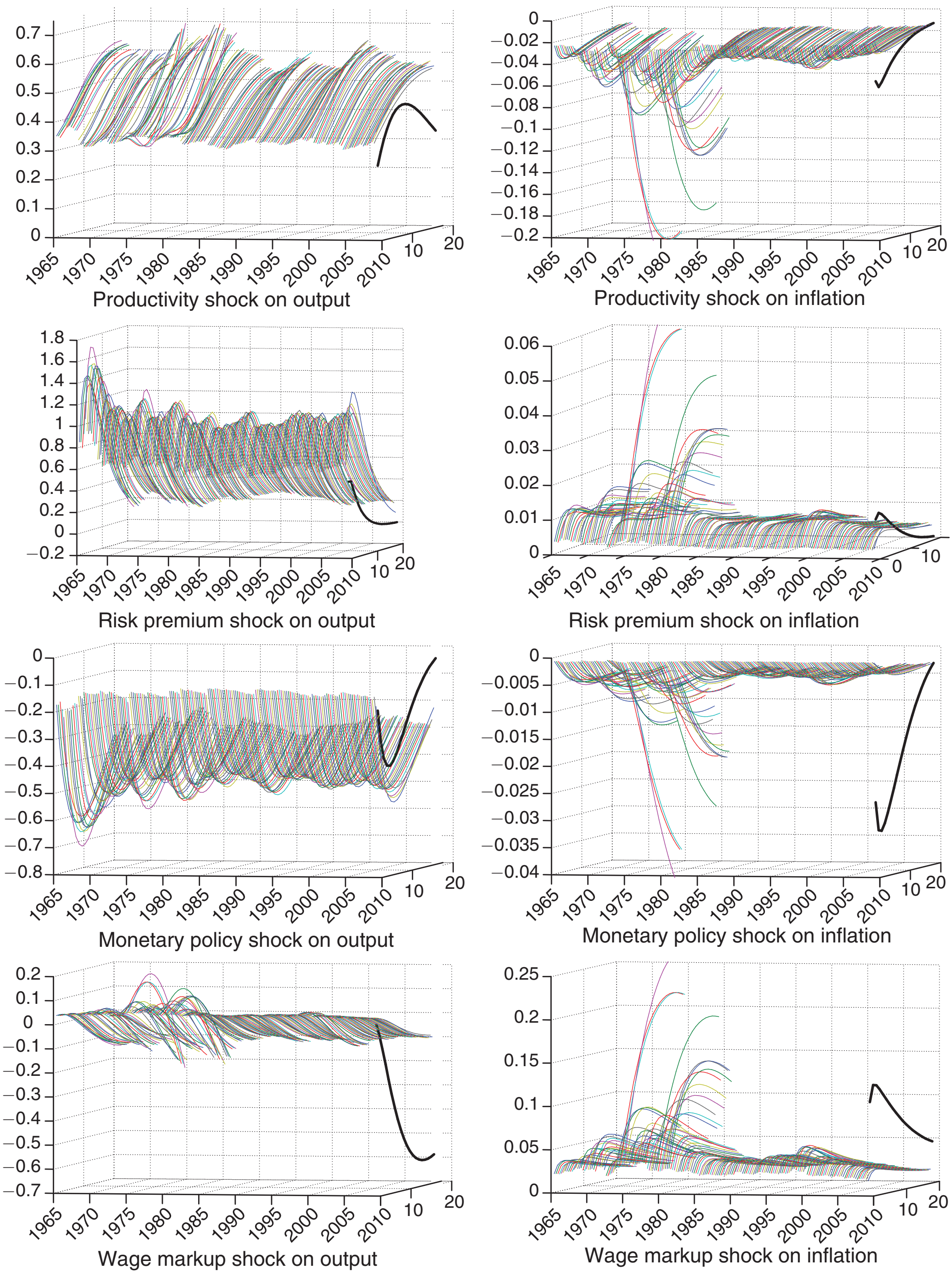


FIGURE 4. IRFUNCTIONS IN THE AL AND RE MODEL

variation. This is especially the case for the impact effects, while the transmission of the shocks in the subsequent quarters tends to vary somewhat more with larger and more persistent effects in the 1970s as well. The profile of the output reaction is more in line with the reaction under RE for most shocks but not so for the mark-up shocks, which have a much smaller impact on output under learning.

Table 7—Mean and Volatility of Inflation and Output over Subsamples

| | DATA | | AL Model | |
|---|---|---|---|---|
| | Before 1984 | After 1984 | Before 1984 | After 1984 |
| Inflation | | | | |
| Mean | 1.43 | 0.62 | 1.03 | 0.63 |
| SD | 0.60 | 0.25 | 0.55 | 0.32 |
| Output growth | | | | |
| Mean | 0.36 | 0.43 | 0.36 | 0.45 |
| SD | 1.12 | 0.59 | 1.06 | 0.96 |
| HP-output | | | | |
| SD | 2.02 | 0.96 | 1.87 | 1.66 |

Changes in the IRFs, especially impact responses, also illustrate the impact on the transmission mechanism brought about by the small forecasting model assumption. For example, when the agents are using only lags of inflation to form inflation expectations instead of a full RE information set, all other state variables, including shocks, tend to become less important for determining inflation. As a result, impact effect decreases. As exogenous shocks get incorporated into inflation over time, the agents' expectations adjust as well, producing a more persistent inflation IRF. In Slobodyan and Wouters (2012a), we observed explicitly the improved match in terms of persistence between IRFs under learning and IRFs under more data-driven DSGE-VAR identification schemes.

## B. *Time Variation in the Volatility*

Given the time variation in the way agents formed expectations in our PLM model and the effect of this on the transmission mechanism of the different shocks, it is interesting to evaluate how this contributed to the overall volatility in the economy. The results are most outspoken for inflation. The model produces both a higher mean inflation and a higher inflation volatility in the 1970s than in the period since 1984, see Table 7.[15] The outcome for the mean inflation is mainly related to the time varying constant in the belief equations. The higher volatility is explained by the higher perceived inflation persistence and the stronger and more persistent reaction of inflation to all the shocks in the 1970s. Averaging over the subperiods before and after 1984, the model explains a drop in inflation volatility from 0.55 to 0.32, a 42 percent drop in volatility $\big((0.55 - 0.32)/0.55 \times 100$ percent$\big)$, while in the historical data, the observed decline is 58 percent. Thus, the model explains 0.67 of the historical moderation in inflation volatility. These results clearly illustrate the crucial role of inflation expectations to explain the great inflation experience of the 1970s. The series of upward inflation shocks that

[15] To produce these numbers, 500 draws from the MCMC were randomly selected. At every parameter draw, the time-varying $\mu$, $\mathbf{T}$, and $\mathbf{R}$ implied by the changing beliefs, were saved. Then this time-varying ALM model was simulated 500 times to produce 500 hypothetical alternative histories for our sample period. Before and after 1984 means and standard deviations were then averaged over all histories, and then over all parameter draws.

arose in the mid-1970s led to an upward revision in the mean expected inflation rate by private agents and, at the same time, they also revised their perceived inflation persistence, which reinforced the impact of the unfavorable shocks on inflation even further. This revision in the inflation expectations of the private sector happened independently of the systematic monetary policy behavior, as the policy rule in our model is assumed to be constant over the complete estimation period. At the beginning of the 1980s, a combination of restrictive monetary policy and negative inflation shocks caused agents to revise downward their expectations about future mean inflation and the perceived inflation persistence, so that inflation gradually converged toward the inflation objective of the central bank. The crucial mechanism in this explanation of the great inflation is the interaction between the way inflation expectations are formed and the specific series of historical shocks that appear over time. This interpretation suggests that monetary policymakers should continuously be careful about inflation expectations and how these anticipations react to positive inflation shocks.

For output, the model is able to replicate the increase in the average growth rate over the two subperiods, but it explains only a small fraction of the great moderation in the volatility of the real variables. The baseline AL model explains, on average, 18 percent of the decline in volatility in the growth rate of output, or 15 percent in terms of HP-filtered output gap volatility. Most of the real moderation generated by our model seems to be related to consumption behavior.

Note that the data reported in Table 7 are averages. The exercise also makes it possible to calculate the probability of the observed decline according to our posterior distribution. For inflation, we find that the volatility declines by 58 percent or more in 10 percent of the simulations, while less than 1 percent of simulations reproduce an observed decline of 47 percent for output growth volatility. Despite the relatively large uncertainty surrounding these statistics, the baseline model is hard to reconcile with the real moderation. For a successful explanation of the output moderation by the learning approach, it is necessary that the model captures a sufficiently strong decline in the perceived persistence in the growth of the real variables. The baseline model does take up some decline for consumption growth persistence, but this effect is not strong enough. Some of the alternative specifications that are discussed in the sensitivity analysis are more successful on this dimension. This suggests that the learning approach is potentially able to explain a larger fraction of the real moderation with an appropriate selection of the belief processes.[16]

### C. *Time Variation in the Inflation-Output Relation*

The time variation in the learning model also has interesting implications for the relation between inflation and output. In the literature, the flattening of the Phillips curve has been observed in various contexts, cf. Atkeson and Ohanian (2001), Stock and Watson (2006), Borio and Filardo (2007), Kuttner and Robinson (2008), and others. Using the time-varying representation of the learning model,

[16] Bullard and Singh (2009) report similar results using a regime-switching process for the underlying technology shocks and find that learning effects can explain up to 30 percent of the volatility reduction.

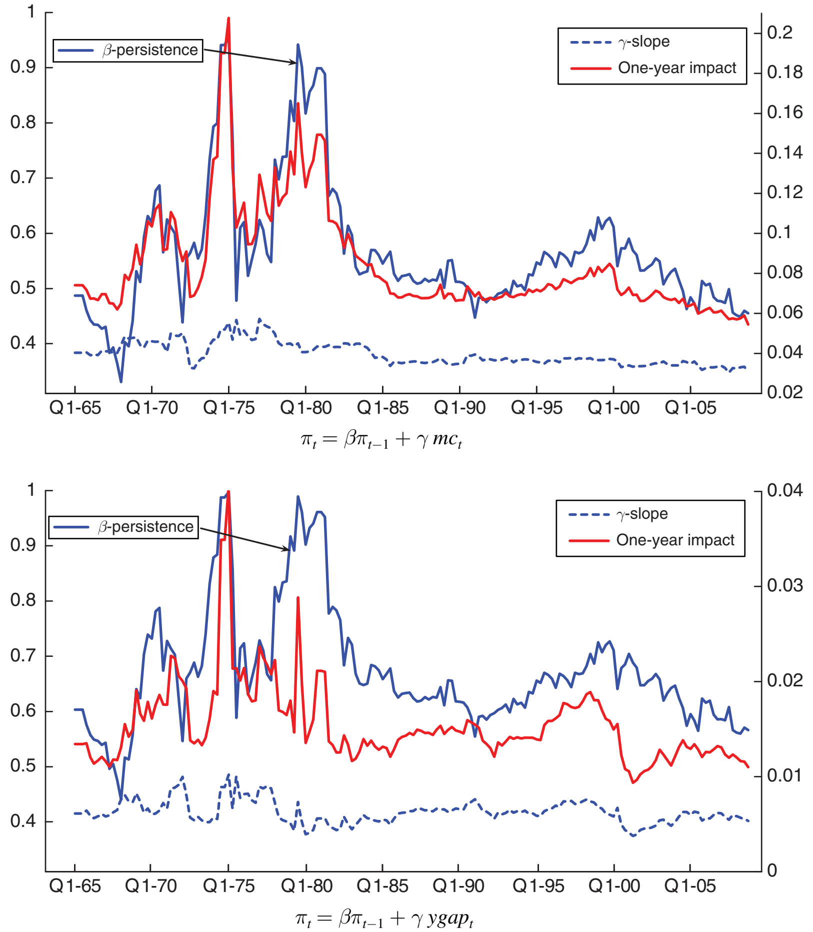


Figure 5. Time Variation in the Phillips Curve Coefficients

we can calculate for each point in time the implied coefficients for a projection of current inflation on the current measure of economic conditions (marginal cost or output gap) and past inflation. The resulting coefficients are given in Figure 5. The first panel shows that the coefficient of the real marginal cost in a simple backward-looking Phillips curve equation has indeed declined in our learning model in line with the evidence on a flattening of the Phillips curve after 1985. For the regressions using the output gap instead, the conclusion is less clear, as there are two periods of significant flattening separated by a marked increase in the slope of the Phillips curve in the early 1980s. However, if we calculate the effect of a sustained four-quarter change in economic conditions on the one-year-ahead inflation rate, the conclusion is clear—over time, the one-year-ahead impact on inflation drops significantly. In the case where the marginal cost is used as a proxy for economic conditions, the impact closely follows the perceived inflation persistence, while in the output gap case, the impact clearly depends on both the persistence and the slope coefficient.

## V. Sensitivity and Robustness Analysis

In this section, we provide evidence on the robustness of the learning dynamics. First of all, we illustrate that the results are not dependent on the choice of one specific model for the PLM. Secondly, we show that results are not sensitive to the sample period and that the learning model still outperforms the RE model when estimated over the more stable period since 1984. Finally, we also document the robustness of the estimation outcomes across alternative specifications and initialization of the learning dynamics.

### A. *Alternative Assumptions on the PLM-Model*

In the baseline learning model, the PLM model was specified as a simple AR(2) model. There are clearly no obvious arguments for choosing this particular forecasting model, and it is crucial to show that the results of the baseline model are robust across alternative choices for the PLM model. Therefore, we also considered an extended setup of the learning dynamics in which we allow agents to consider a set of alternative small forecasting models. These are five small forecasting models in which the expectations about the forward variables are modelled either as a simple AR(1) process; an AR(2) process as in the baseline model; an AR(2) plus the lagged inflation rate; an AR(2) plus the lagged inflation and interest rate; and an AR(2) plus lagged inflation, interest rate, and output. We allow agents to combine these forecasts either by simple equal weights (EW version) or by a Bayesian Model Averaging method in which the weight attached to each model varies with the past forecasting performance of each of the small forecasting models (BIC version).[17] The exercise confirms the robustness of the results in various dimensions. The AL model with a set of small forecasting models produces a marginal likelihood that is very similar to that of our baseline AR(2) specification and also outperforms the RE model on this criterion. The persistence of inflation in the average forecasting model, measured by the sum of the lagged inflation coefficients, follows a profile over the sample that is close to the one in the baseline model. Structural parameters and identified shock series do not change significantly. Some IRFs do change mainly in their short-run response. For instance, with the interest rate included in the PLM specification, the model predicts a positive inflation response to a restrictive monetary policy shock in the 1970s. Raising the policy rate created higher future inflation expectations. This response resembles the price puzzle in the SVAR literature.

We also considered alternative PLM models in which we allow agents to augment the AR(2) specification with additional RHS variables specific for each of the seven forward variables. These variables were selected based on the direct relevance of the variables suggested by the structural model: marginal cost for inflation, interest rates for the demand components, etc. Such extensions of the PLM

[17] It is interesting to note that in our estimations, a simple fixed-weight forecast combination works better than sophisticated time-varying re-weighting of the forecasts. Timmermann (2006) surveys a broad literature that reaches a similar conclusion, namely that simple forecast combinations "often dominate more refined combination schemes aimed at estimating the theoretically optimal combination weights." For demonstration of dominance of simple averaging of forecasts of quarterly GDP growth, see Stock and Watson (2004).

model do not improve the fit of the model, and in some cases they lead to significantly worse (but still better than under RE) marginal likelihoods. This finding may be related to the fact that while the model suggests a contemporaneous relationship between these variables and the forward variables, in adaptive learning we are restricting to variables' lags instead.

Moving in the other direction, however, produced better results. Using the baseline beliefs, we allowed two of the forward-looking variables (inflation and price of capital) to be AR(1) processes, leaving the remaining five forwards as AR(2). This specification produced a marginal likelihood rather close to that of the baseline model, and was more successful in matching the post-1984 moderation experience.

In still another set of exercises, we allowed agents to use a significantly broader information set for forecasting. When lags of all seven forward-looking variables are added to all forecasting equations, the model fit under adaptive learning is only marginally better than for RE DSGE but significantly worse than for the baseline learning model. When we tried to mimic the information set available to the rational agents by including either all endogenous state variables or endogenous state variables and shocks, the model fit is worse by five to seven units compared to RE. We conclude that including too many variables into the forecasting functions leads to the overfitting problem and worsens the forecasting performance.

## B. *Alternative Sample Periods*

We estimated the model under RE and under learning for the baseline AR(2)-PLM specification for two subperiods: 1966:I–1979:II and 1984:I–2008:IV. In both cases, the model with AL improves on the RE model in terms of marginal likelihood, see Table 8. Perceived inflation persistence is the main difference between the two sub-sample models. As we have already illustrated with the out-of-sample prediction exercise over the period since 1990, the model with learning dynamics based on small PLM models improves upon the RE model during the more stable environment of the great moderation period too.

## C. *Alternative Specification and Initialization of the Learning Dynamics*

The estimation results for the learning model are also robust across alternative initializations and specifications of the Kalman filter in the learning process. To illustrate this, we re-estimated the model over the same sample, but now taking into account pre-sample data ranging from 1955:I until 1965:IV to initialize the Kalman filter for the belief coefficients. The marginal likelihood of this model is very similar to the baseline version, illustrating that our results are not particularly sensitive to the initialization based on the RE-implied belief coefficients. It is interesting to note that the beliefs about inflation persistence during this pre-sample period follow an interesting pattern, with first a strong decline in the beginning of the 1960s, followed by an increase, so that the beliefs at the start of our maximization period (1966:I) are again close to the initialization under the baseline approach. After a few more years, the beliefs have completely converged to their counterparts in the baseline estimation.

Table 8—Estimation Outcomes for Different PLM Assumptions

| | Marg. Lik. | |
|---|---|---|
| Robustness w.r.t. the PLM assumption | | |
| Average over five models: Equal weights | −988.55 | |
| Average over five models: BIC weights | −992.23 | |
| | | |
| Robustness w.r.t. the estimation sample | AL | RE |
| 1966:I–1979:II | −355.34 | −363.58 |
| 1984:I–2008:IV | −411.00 | −429.49 |
| | | |
| Robustness w.r.t. the specification—initialization of the learning process | | |
| Pre-sample updating of beliefs | −981.82 | |
| | | |
| One-step optimized beliefs | −974.55 | |
| Optimized beliefs, no updating | −966.89 | |
| Optimized beliefs, **V** from RE DSGE | −957.78 | |
| Optimized beliefs, **V** from AL DSGE | −948.57 | |
| | | |
| Parameters $\gamma, \sigma$ estimated $(\rho = 1)$ | −993.41 | |
| Equation-by-equation Kalman filter learning | −984.55 | |
| Equation-by-equation constant gain learning | −991.30 | |

Secondly, we experimented with alternative initial beliefs to test their impact on the results. Ideally, one would like to estimate both the initial belief coefficients $\boldsymbol{\beta}_{1|0}$ and the assumed variance-covariance matrix $\Sigma$ (and, therefore, $\mathbf{V}$ and $\mathbf{P}_{1|0}$) of the shocks in the agents' forecasting model (3).[18] This is not feasible in our application as the number of coefficients would increase considerably. One step in that direction is to estimate a model with learning where both the $\Sigma$, $\mathbf{V}$ matrices and the initial beliefs $\boldsymbol{\beta}_{1|0}$ are kept constant and equal to their counterparts in the baseline model. This resembles a first step of an iterative optimization procedure, where initial beliefs and model parameters are optimized during alternate steps. Compared to the baseline, the marginal likelihood of this model improves by another ten units. The learning dynamics, in particular beliefs about inflation, are very similar to the baseline model.

In still another experiment with alternative initial beliefs, we used the following setup. First, we estimated the DSGE model under the assumption that all expectations were formed as univariate AR(2) processes with fixed coefficients $\boldsymbol{\beta}$; we estimated these $\boldsymbol{\beta}$ jointly with the model parameters in a resulting backward-looking model. This step provides us with beliefs $\boldsymbol{\beta}$ optimal for the case of no learning dynamics. These optimized $\boldsymbol{\beta}$ values are then used as initial beliefs $\boldsymbol{\beta}_{1|0}$ in the adaptive learning estimations. As in the previous exercise, $\Sigma$, $\mathbf{V}$, and $\mathbf{P}_{1|0}$ are left unaffected and taken from the baseline models. Here, our goal was to break the tight link between the agents' forecasts determined by $\boldsymbol{\beta}_{1|0}$ and the uncertainty about these forecasts (and, therefore, the speed with which new information is incorporated into the beliefs), which depends on $\mathbf{V}$.

[18] Both $\mathbf{P}_{1|0}$ and $V$ are assumed to be proportional to $\mathrm{Var}[\hat{\boldsymbol{\beta}}_{GLS}] = (X^T \Sigma^{-1} X)^{-1}$ as described in Section I, and, thus, the choise of $\Sigma$ determines how fast the forecasting errors are incorporated into the beliefs according to Kalman filter updating equation (4a) and (4b).

The results of this experiment are as follows. First, the purely backward-looking model fits the data very well: conditional on the estimated $\boldsymbol{\beta}$, the log marginal likelihood improves by a factor of 38 over the RE DSGE model and by 33 units over a model with fixed model-consistent beliefs (column 4 of Table 1).[19] The optimized $\boldsymbol{\beta}$ coefficients differ significantly from those that could be consistent with any REE. For example, the agents perceive inflation to be an AR(2) process with zero persistence, which is incompatible with our estimated models.[20]

Second, introduction of adaptive learning still results in large improvement over the backward-looking model: log marginal likelihood improves by a further 18 units to $-948.57$ for the best performing model. Thus, we conclude that it is not just particular initial expectations that are beneficial to the model fit. The time variation introduced by the adaptive learning matters significantly as well, even after optimizing on these beliefs.

Third, the model fit, and especially the evolution of the beliefs, depend strongly on the second moment matrices $\mathbf{V}$ and $\boldsymbol{\Sigma}$, which are used by the agents in their beliefs Kalman filtering step. For example, if the second moments are derived using the REE of the RE DSGE model (column 1 in Table 1), the log marginal likelihood equals $-957.78$. There is very little time variation in perceived inflation persistence. On the other hand, when we use REE of a model estimated under adaptive learning (column 2 in Table 1), the log marginal likelihood equals $-948.57$, and the evolution of inflation persistence exhibits a double-peak structure qualitatively similar to that in the baseline model, going up from zero to about 0.8 in 1975 and 0.6 in 1981, and then falling back to zero over the remainder of the sample, while the constant moves very little.

The REE which determines the second moment matrices matters for the behavior of beliefs for the following reason. When the RE DSGE model is used to derive the $V$ matrix, the perceived volatility of persistence coefficient in the inflation forecasting equation is an order of magnitude smaller than when adaptive learning DSGE determines $V$, while the perceived volatility of the constant is an order of magnitude larger. As a result, the persistence moves a lot in the second case but not in the first case, while the opposite holds for the constant.

We also tested the robustness of the learning dynamics with respect to the three learning-specific coefficients: $\rho$, $\sigma_v$, and $\sigma_0$. The role of $\sigma_0$ is limited to the first few observations, and given our four-period initialization of the Kalman filter, this parameter is basically neutral for the results. For simple univariate cases, it can be shown that $\sigma_v$ and $\rho$ affect the Kalman gain in a similar direction, so that an exact identification of these two parameters simultaneously is most unlikely. Therefore, we estimate $\sigma_v$ and $\sigma_0$ for different values of the $\rho$ coefficient. When $\rho = 1$, $\sigma_v$ is estimated at 0.0012, while for $\rho = 0.97$, the estimate from the baseline model, the estimate for $\sigma_v$ increases to values that are close to the ones assumed in the baseline setup. However, the uncertainty around $\sigma_v$ is very large, and the marginal likelihood is hardly affected by these alternative approaches. Note that changing $\sigma_v$ does not

[19] Note that we report log marginal likelihood values conditional on $\boldsymbol{\beta}_{1|0}$; uncertainty related to the estimation of these extra 21 parameters is ignored.

[20] Under RE, inflation persistence is 0.84. Under adaptive learning, model-consistent inflation persistence is typically around 0.6.

affect the relative size of the elements of the **V** matrix, and thus relative stability of behavior of the beliefs should be expected.

Alternatively, instead of restricting $\rho$ to be the same for all forward-looking variables, we estimate the AL model with seven separate $\rho$ parameters.[21] The resulting fit is comparable to that under the RE but is significantly worse than for the baseline learning model due to extra parameter uncertainty.

Finally, we replace the SURE approach for the forecasting model with an equation-by-equation estimation approach for the seven forward variables. This is achieved by retaining only the main diagonal in the matrix $\Sigma$ derived as described in Section I. Although theoretically this estimation approach should deliver less efficient estimates, in our application it works quite well and the marginal likelihood is almost unchanged. The results show that the perceived inflation beliefs behave very similarly to the baseline model, while the beliefs about the real variables tend to display more volatility, and as a consequence, this version is marginally more successful in simulating the post-1984 real moderation.

For this simpler equation-by-equation estimation approach of the PLM, we can also consider constant-gain learning instead of Kalman filter learning. The estimated constant gain turns out to be quite low (0.006) and the updating of the PLM beliefs leads to minor revisions. Despite this low time variability, constant gain learning with AR(2) beliefs delivers a marginal likelihood that is better than under RE and under the no-update model (the last column of Table 1). The outcome is worse than under Kalman filter learning, probably reflecting slower response of constant gain learning to changes in the underlying data process. Different learning speeds of even asymptotically-equivalent constant gain and Kalman filter algorithms had previously been observed in Sargent and Williams (2005).

## VI. Concluding Remarks

The hypothesis of model-consistent expectations, especially in the context of a medium-scale DSGE model, implies that economic agents are extremely well informed both about the structure of the model and the type of shocks that are hitting the economy at each point in time. Therefore, it is not surprising that simpler, and probably more realistic, assumptions about the expectations mechanism can improve the empirical fit of these models. In addition, our results suggest that there might be an important role for learning in these expectations. Agents update their belief models in line with actual past data and, by doing so, their reactions to exogenous shocks change considerably over time. This process is particularly relevant to understanding the changing dynamics of the inflation process. Even under a constant monetary policy rule, the beliefs of the private agents about the mean and the persistence of the inflation process can vary substantially over time. The additional dynamics from the learning process substitute for the persistence in the exogenous price and wage shocks and the backward-looking indexation in wage-setting, which are both important in the rational expectations version of the model. The important

[21] This allows the agents to change the relative speed with which new information is incorporated into the beliefs in different forecasting functions.

role of learning in inflation dynamics supports the idea that private sector inflation expectations should be closely monitored by central bankers.

The specification of the small belief models may of course be criticized as being ad hoc. We have tried to take into account that problem by allowing agents to consider different small models and to weight them depending on their past forecasting performance. Still, the belief models that we consider might be too restrictive. Introducing evidence from surveys about expectations might help to pin down the relevant information set used by agents and to overcome this identification problem. There are several recent papers that use SPF data as an observable variable. Del Negro and Eusepi (2010) test an alternative hypothesis on the relevant expectation models. As in Slobodyan and Wouters (2012a), they indicate an important divergence between the implied expectations in RE-DSGE models and SPF evidence. Using SPF data combined with real-time data, Milani (2011) also observes large deviations between model forecasts and SPF evidence and suggests that these expectation "shocks" might be an important source of business cycle fluctuations. Ormeño (2010) uses SPF inflation expectations in an estimated DSGE model with adaptive learning and concludes that SPF data is useful in distinguishing between the rational expectations and adaptive learning models. In this paper, we have only tested how well the estimated belief coefficients fare against real-time data and SPF expectations, and we leave an explicit modelling of the real-time data and of SPF expectations in the context of our adaptive learning model for future research.

Surveys contain not only information about the one-period-ahead expectations, but also about longer horizon predictions. Consistent processing of this type of information would require a switch to infinite-horizon learning. This type of information and learning approach might also be necessary to overcome one of the main weaknesses of the Euler equation learning, namely the weak quality of long-term forecasting.

Two other extensions of the paper are on our research agenda. The learning dynamics can potentially also contribute to an explanation of the great moderation on the real side of the economy. At this stage, our belief models for consumption and investment do not exhibit a clearly declining persistence in the growth rates of these variables. Such beliefs might be necessary to explain the observed moderation in real volatility. Secondly, we would like to test the time variation that is generated by the learning dynamics against a more general and less restrictive time-varying VAR model.

## Data Appendix

The model is estimated using seven key macroeconomic time series: real GDP, consumption, investment, hours worked, real wages, prices, and a short-term interest rate. GDP, consumption, and investment are taken from the US Department of Commerce—Bureau of Economic Analysis database. Real Gross Domestic Product is expressed in Billions of Chained 1996 Dollars. Nominal Personal Consumption Expenditure and Fixed Private Domestic Investment are deflated by the GDP deflator. Inflation is the first difference of the log of the Implicit Price Deflator of GDP. Hours and wages come from the BLS (hours and hourly compensation for the NFB sector for all persons). Hourly compensation is divided by the GDP price deflator in order to get the real wage variable. Hours are adjusted to take into account the limited coverage of the NFB sector compared to GDP (the index of average hours for the NFB sector is multiplied by the Civilian Employment figure (16 years and over). The aggregate real variables are expressed per capita by dividing by the population over 16. All series are seasonally adjusted. The interest rate is the Federal Funds Rate. Consumption, investment, GDP, wages, and hours are expressed in 100 times log. The interest rate and inflation rate are expressed on a quarterly basis corresponding with their appearance in the model.

## Model Appendix

In this Appendix, we summarize the log-linear equations of the model. For a more detailed presentation, we refer to the discussion in SW 2007.

• Consumption Euler equation for the nonseparable utility function:

$$\begin{aligned}\hat{c}_t &= c_1 E_t[\hat{c}_{t+1}] + (1 - c_1)\hat{c}_{t-1} + c_2(\hat{L}_t - E_t[\hat{L}_{t+1}]) \\ &\quad - c_3(\hat{R}_t - E_t[\hat{\pi}_{t+1}] + \hat{\varepsilon}_t^b),\end{aligned}$$

with $c_1 = 1/(1+\overline{\eta})$, $c_2 = c_1(\sigma_c - 1)(wL/C)/\sigma_c$, $c_3 = c_1(1-\overline{\eta})/\sigma_c$, where $\overline{\eta}$ is the external habit parameter adjusted for trend growth $\overline{\eta} = (\eta/\gamma)$; $\sigma_c$ is the inverse of the intertemporal elasticity of subsitution; and $\hat{\varepsilon}_t^b$ is the exogenous AR(1) risk premium process.

• Investment Euler equation:

$$\hat{i}_t = i_1 \hat{i}_{t-1} + (1 - i_1)\hat{i}_{t+1} + i_2 \hat{Q}_t^k + \hat{\varepsilon}_t^q,$$

with $i_1 = 1/(1+\overline{\beta}\gamma)$, $i_2 = i_1/(\gamma^2\varphi)$, where $\overline{\beta}$ is the discount factor adjusted for trend growth $(\beta\gamma^{1-\sigma_c})$, and $\varphi$ is the elasticity of the capital adjustment cost function. $\hat{\varepsilon}_t^q$ is the exogenous AR(1) process for the investment specific technology.

• Value of the capital stock:

$$\hat{Q}_t^k = -(\hat{R}_t - E_t[\hat{\pi}_{t+1}] + \hat{\varepsilon}_t^b) + q_1 E_t[r_{t+1}^k] + (1 - q_1)E_t[Q_{t+1}^k],$$

with $q_1 = r_*^k/(r_*^k + (1-\delta))$, where $r_*^k$ is the steady state rental rate to capital, and $\delta$ the depreciation rate.

• Aggregate demand equals aggregate supply:

$$\begin{aligned}\hat{y}_t &= \frac{c_*}{y_*}\hat{c}_t + \frac{i_*}{y_*}\hat{i}_t + \hat{\varepsilon}_t^g + \frac{r_*^k k_*}{y_*}\hat{u}_t \\ &= \Phi_p(\alpha\hat{k}_t + (1-\alpha)\hat{L}_t + \hat{\varepsilon}_t^a),\end{aligned}$$

with $\Phi_p$ reflecting the fixed costs in production which corresponds to the price markup in steady state. $\hat{\varepsilon}_t^g$ and $\hat{\varepsilon}_t^a$ are the AR(1) processes representing exogenous demand components and the TFP process.

- Price-setting under the Calvo model with indexation:

$$\hat{\pi}_t - \iota_p \hat{\pi}_{t-1} = \pi_1\big(E_t[\hat{\pi}_{t+1}] - \iota_p \hat{\pi}_t\big) - \pi_2 \hat{\mu}_t^p + \hat{\varepsilon}_t^p,$$

with $\pi_1 = \overline{\beta}\gamma$, $\pi_2 = (1 - \xi_p \overline{\beta}\gamma)(1 - \xi_p)/[\xi_p(1 + (\Phi_p - 1)\varepsilon_p)]$; with $\xi_p$ and $\iota_p$, respectively, the probability and indexation of the Calvo model; and $\varepsilon_p$ the curvature of the aggregator function. The price markup $\hat{\mu}_t^p$ is equal to the inverse of the real marginal $\widehat{mc}_t = (1 - \alpha)\hat{w}_t + \alpha \hat{r}_t^k - \hat{A}_t$.

- Wage setting under the Calvo model with indexation:

$$\hat{\pi}_t^w - \iota_w \hat{\pi}_{t-1} = \pi_1\big(E_t[\hat{\pi}_{t+1}^w] - \iota_w \hat{\pi}_t\big) - \pi_3 \hat{\mu}_t^w + \hat{\varepsilon}_t^w,$$

with $\pi_3 = \big(1 - \xi_w \overline{\beta}\gamma\big)(1 - \xi_w)/\big[\xi_w(1 + (\phi_w - 1)\varepsilon_w)\big]$, and wage markup $\hat{\mu}_t^w = \hat{w}_t - w_1 \hat{c}_t + (1 - w_1)\hat{c}_{t-1} - \sigma_l \hat{L}_t$ with $w_1 = 1/(1 - \overline{\eta})$.

- Capital accumulation equation:

$$\hat{\overline{k}}_t = \kappa_1 \hat{\overline{k}}_{t-1} + (1 - \kappa_1)\hat{i}_t + \kappa_2 \hat{\varepsilon}_t^q,$$

with $\kappa_1 = (1 - (i_*/\overline{k}_*)$, $\kappa_2 = (i_*/\overline{k}_*)(1 + \overline{\beta}\gamma)\gamma^2 S''$. Capital services used in production is defined as $\hat{k}_t = \hat{u}_t + \hat{\overline{k}}_{t-1}$.

- Optimal capital utilization condition:

$$\hat{u}_t = (1 - \psi)/\psi \hat{r}_t^k,$$

where $\psi$ is the elasticity of the capital utilization cost function.

- Optimal capital/labor input condition:

$$\hat{k}_t = \hat{w}_t - \hat{r}_t^k + \hat{L}_t.$$

- Monetary policy rule:

$$\hat{R}_t = \rho_R \hat{R}_{t-1} + (1 - \rho_R)\big(r_\pi \hat{\pi}_t + r_y(y\widehat{gap}_t) + r_{\Delta y}\Delta(y\widehat{gap}_t)\big) + \hat{\varepsilon}_t^r,$$

with $ygap_t = (\hat{y}_t - \Phi_p \hat{\varepsilon}_t^a)$.

The following parameters are not identified by the estimation procedure and therefore calibrated: $\delta = 0.025$, $\varepsilon_p = 10$, $\varepsilon_w = 10$, $\phi_w = 1.5$.

Table A1—Posterior Estimates for RE and AL Models—Complete List of Parameters

| | Prior distribution | | | Posterior distribution | | | |
|---|---|---|---|---|---|---|---|
| | | | | RE | | AL-Baseline | |
| | Type | Mean | SD | Mean | 5%÷95% | Mean | 5%÷95% |
| Standard deviation of the innovations[a] | | | | | | | |
| $\sigma_a$ | IG | 0.1 | 2 | 0.45 | 0.41÷0.49 | 0.46 | 0.42÷0.50 |
| $\sigma_b$ | IG | 0.1 | 2 | 0.26 | 0.21÷0.30 | 0.15 | 0.12÷0.18 |
| $\sigma_g$ | IG | 0.1 | 2 | 0.52 | 0.47÷0.56 | 0.50 | 0.46÷0.55 |
| $\sigma_q$ | IG | 0.1 | 2 | 0.42 | 0.35÷0.50 | 0.45 | 0.39÷0.50 |
| $\sigma_r$ | IG | 0.1 | 2 | 0.22 | 0.20÷0.24 | 0.22 | 0.20÷0.24 |
| $\sigma_p$ | IG | 0.1 | 2 | 0.15 | 0.13÷0.18 | 0.15 | 0.13÷0.17 |
| $\sigma_w$ | IG | 0.1 | 2 | 0.23 | 0.19÷0.26 | 0.23 | 0.21÷0.26 |
| Persistence of the exogenous processes: $\rho = \mathrm{AR}(1)$, $\theta = \mathrm{MA}(1)$ | | | | | | | |
| $\rho_a$ | B | 0.5 | 0.2 | 0.93 | 0.90÷0.97 | 0.99 | 0.99÷0.99 |
| $\rho_b$ | B | 0.5 | 0.2 | 0.25 | 0.10÷0.40 | 0.55 | 0.40÷0.72 |
| $\rho_g$ | B | 0.5 | 0.2 | 0.98 | 0.97÷0.99 | 0.97 | 0.95÷0.98 |
| $\rho_q$ | B | 0.5 | 0.2 | 0.75 | 0.66÷0.84 | 0.51 | 0.39÷0.62 |
| $\rho_r$ | B | 0.5 | 0.2 | 0.08 | 0.02÷0.15 | 0.10 | 0.02÷0.18 |
| $\rho_p$ | B | 0.5 | 0.2 | 0.85 | 0.75÷0.94 | | |
| $\rho_w$ | B | 0.5 | 0.2 | 0.96 | 0.93÷0.99 | | |
| $\theta_p$ | B | 0.5 | 0.2 | 0.70 | 0.55÷0.86 | | |
| $\theta_w$ | B | 0.5 | 0.2 | 0.88 | 0.81÷0.95 | | |
| $a\_g$[b] | N | 0.5 | 0.25 | 0.53 | 0.40÷0.68 | 0.54 | 0.41÷0.67 |
| Structural parameters | | | | | | | |
| $\varphi$ | N | 4.0 | 1.5 | 5.45 | 3.82÷7.02 | 3.23 | 1.81÷4.48 |
| $\sigma_c$ | N | 1.5 | 0.37 | 1.31 | 1.10÷1.53 | 1.58 | 1.30÷1.83 |
| $\eta$ | B | 0.7 | 0.1 | 0.77 | 0.70÷0.84 | 0.68 | 0.60÷0.76 |
| $\sigma_l$ | N | 2.0 | 0.5 | 1.48 | 0.61÷2.37 | 1.77 | 0.99÷2.47 |
| $\xi_p$ | B | 0.5 | 0.1 | 0.72 | 0.65÷0.80 | 0.65 | 0.59÷0.71 |
| $\xi_w$ | B | 0.5 | 0.1 | 0.77 | 0.68÷0.84 | 0.84 | 0.80÷0.88 |
| $\iota_p$ | B | 0.5 | 0.15 | 0.24 | 0.10÷0.38 | 0.19 | 0.08÷0.29 |
| $\iota_w$ | B | 0.5 | 0.15 | 0.51 | 0.30÷0.71 | 0.21 | 0.09÷0.32 |
| $\psi$ | B | 0.5 | 0.12 | 0.52 | 0.33÷0.70 | 0.56 | 0.35÷0.77 |
| $\Phi_p$ | N | 1.25 | 0.12 | 1.62 | 1.49÷1.74 | 1.56 | 1.43÷1.68 |
| $\rho_R$ | B | 0.75 | 0.1 | 0.85 | 0.82÷0.88 | 0.89 | 0.86÷0.92 |
| $r_\pi$ | N | 1.5 | 0.25 | 1.88 | 1.63÷2.15 | 1.75 | 1.44÷2.04 |
| $r_y$ | N | 0.12 | 0.05 | 0.10 | 0.06÷0.14 | 0.15 | 0.09÷0.20 |
| $r_{\Delta y}$ | N | 0.12 | 0.05 | 0.16 | 0.13÷0.19 | 0.14 | 0.11÷0.17 |
| $\overline{\pi}$ | G | 0.62 | 0.1 | 0.74 | 0.56÷0.91 | 0.64 | 0.55÷0.73 |
| $100(\beta^{-1} - 1)$ | G | 0.25 | 0.1 | 0.19 | 0.09÷0.29 | 0.17 | 0.06÷0.27 |
| $\bar{l}$ | N | 0.0 | 2.0 | 0.84 | −0.63÷2.34 | 0.83 | −0.72÷2.43 |
| $\gamma$ | N | 0.4 | 0.1 | 0.41 | 0.39÷0.43 | 0.41 | 0.37÷0.44 |
| $\alpha$ | N | 0.3 | 0.05 | 0.19 | 0.16÷0.22 | 0.17 | 0.14÷0.20 |

[a] The IG-distribution is defined by the degrees of freedom.
[b] The effect of TFP innovations on exogenous demand.

Table A1 provides additional information on the RE and AL estimation results.

## REFERENCES

**Atkeson, Andrew, and Lee E. Ohanian.** 2001. "Are Phillips Curves Useful for Forecasting Inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review* 25 (1): 2–11.

**Borio, Claudio, and Andrew Filardo.** 2007. "Globalisation and Inflation: New Cross-Country Evidence on the Global Determinants of Domestic Inflation." Bank for International Settlements Working Paper 227.

**Branch, William A., and George W. Evans.** 2006. "A Simple Recursive Forecasting Model." *Economics Letters* 91 (2): 158–66.

**Bullard, James, and Aarti Singh.** 2009. "Learning and the Great Moderation." Federal Reserve Bank of St. Louis Working Paper 2007-027.

**Carroll, Christopher D.** 2001. "The Epidemiology of Macroeconomic Expectations." National Bureau of Economic Research Working Paper 8695.

**Chari, V. V., Patrick J. Kehoe, and Ellen R. McGrattan.** 2009. "New Keynesian Models: Not Yet Useful for Policy Analysis." *American Economic Journal: Macroeconomics* 1 (1): 242–66.

**Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans.** 2005. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy* 113 (1): 1–45.

**Cogley, Timothy, Giorgio E. Primiceri, and Thomas J. Sargent.** 2010. "Inflation-Gap Persistence in the US." *American Economic Journal: Macroeconomics* 2 (1): 43–69.

**Croushore, Dean.** 2011. "Frontiers of Real-Time Data Analysis." *Journal of Economic Literature* 49 (1): 72–100.

**Del Negro, Marco, and Stefano Eusepi.** 2010. "Fitting Observed Inflation Expectations." Federal Reserve Bank of New York Staff Reports 476.

**Evans, George W., and Seppo Honkapohja.** 2001. *Learning and Expectations in Macroeconomics*. Princeton: Princeton University Press.

**Honkapohja, Seppo, Kaushik Mitra, and George W. Evans.** 2002. "Notes on Agents' Behavioral Rules under Adaptive Learning and Recent Studies of Monetary Policy." Unpublished. http://www.pages.uoregon.edu/gevans/learn-notes.pdf.

**Juillard, Michel.** 1996. "Dynare: A Program for the Resolution and Simulation of Dynamic Models with Forward Variables Through the Use of a Relaxation Algorithm." CEPREMAP Working Paper 9602.

**Kuttner, Ken, and Tim Robinson.** 2008. "Understanding the Flattening Phillips Curve." Reserve Bank of Australia, Research Discussion Paper 2008-05.

**Mankiw, N. Gregory, and Ricardo Reis.** 2002. "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve." *Quarterly Journal of Economics* 117 (4): 1295–1328.

**Marcet, Albert, and Thomas J. Sargent.** 1989. "Convergence of Least Squares Learning Mechanisms in Self-referential Linear Stochastic Models." *Journal of Economic Theory* 48 (2): 337–68.

**Milani, Fabio.** 2006. "A Bayesian DSGE Model with Infinite-Horizon Learning: Do 'Mechanical' Sources of Persistence Become Superfluous?" *International Journal of Central Banking* 2 (3): 87–106.

**Milani, Fabio.** 2007. "Expectations, Learning and Macroeconomic Persistence." *Journal of Monetary Economics* 54 (7): 2065–82.

**Milani, Fabio.** 2011. "Expectation Shocks and Learning as Drivers of the Business Cycle." *Economic Journal* 121 (552): 379–401.

**Nunes, Ricardo.** 2009. "Learning the Inflation Target." *Macroeconomic Dynamics* 13 (2): 167–88.

**Ormeño, Arturo.** 2010. "Disciplining Expectations: Using Survey Data in Learning Models." Unpublished. http://www.econ.upf.edu/~aormeno/ArturoOrmeno_JobMarket.pdf.

**Orphanides, Athanasios, and John C. Williams.** 2005a. "Inflation Scares and Forecast-Based Monetary Policy." *Review of Economic Dynamics* 8 (2): 498–527.

**Orphanides, Athanasios, and John C. Williams.** 2005b. "The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations." *Journal of Economic Dynamics and Control* 29 (11): 1927–50.

**Pearlman, Joseph, David Currie, and Paul Levine.** 1986. "Rational Expectations Models with Partial Information." *Economic Modelling* 3 (2): 90–105.

**Preston, Bruce.** 2005. "Learning about Monetary Policy Rules When Long-Horizon Expectations Matter." *International Journal of Central Banking* 1 (2): 81–126.

**Sargent, Thomas J.** 1993. *Bounded Rationality in Macroeconomics*. Oxford: Clarendon Press.

**Sargent, Thomas J., and Noah Williams.** 2005. "Impacts of Priors on Convergence and Escapes from Nash Inflation." *Review of Economic Dynamics* 8 (2): 360–91.

**Sims, Christopher A.** 2003. "Implications of Rational Inattention." *Journal of Monetary Economics* 50 (3): 665–90.

**Slobodyan, Sergey, and Raf Wouters.** 2012a. "Learning in an Estimated Medium-Scale DSGE Model." *Journal of Economic Dynamics and Control* 36(1): 22–46.

**Slobodyan, Sergey, and Raf Wouters.** 2012b. "Learning in a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models: Dataset." *American Economic Journal: Macroeconomics*. http://dx.doi.org/10.1257/mac.4.2.65.

**Smets, Frank, and Raf Wouters.** 2003. "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association* 1 (5): 1123–75.

**Smets, Frank, and Rafael Wouters.** 2007. "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach." *American Economic Review* 97 (3): 586–606.

**Stock, James H., and Mark W. Watson.** 2004. "Combination Forecasts of Output Growth in a Seven-Country Data Set." *Journal of Forecasting* 23 (6): 405–30.

**Stock, James H., and Mark W. Watson.** 2006. "Why Has U.S. Inflation Become Harder to Forecast?" National Bureau of Economic Research Working Paper 12324.

**Svensson, Lars E. O., and Michael Woodford.** 2003. "Indicator Variables for Optimal Policy." *Journal of Monetary Economics* 50 (3): 691–720.

**Timmermann, Allan.** 2006. "Forecast Combinations." In *Handbook of Economic Forecasting*, Vol. 1, edited by Graham Elliot, Clive W. J. Granger, and Allan Timmermann, 135–96. Amsterdam: Elsevier.

**This article has been cited by:**


1. Fabian Winkler. 2020. The role of learning for asset prices and business cycles. *Journal of Monetary Economics* **114**, 42-58. [Crossref]
2. Elton Beqiraj, Giovanni Di Bartolomeo, Marco Di Pietro, Carolina Serpieri. 2020. Bounded rationality and heterogeneous expectations: Euler versus anticipated-utility approach. *Journal of Economics* **130**:3, 249-273. [Crossref]
3. George W. Evans, Bruce McGough. 2020. Equilibrium stability in a nonlinear cobweb model. *Economics Letters* **193**, 109130. [Crossref]
4. Wei Zhao, Yi Lu, Minjuan Zhao, Peng Zhang. 2020. Fluctuations in the Open Economy of China: Evidence from the ABNK Model. *Emerging Markets Finance and Trade* **56**:9, 2073-2092. [Crossref]
5. George W. Evans, Bruce McGough. 2020. Stable near-rational sunspot equilibria. *Journal of Economic Theory* **186**, 104987. [Crossref]
6. Sacha Gelfer. 2020. The effects of professional forecast dissemination on macroeconomic volatility. *Journal of Economic Behavior & Organization* **170**, 131-156. [Crossref]
7. Jaqueson K Galimberti. 2020. Information Weighting Under Least Squares Learning. *SSRN Electronic Journal* . [Crossref]
8. Paolo Gelain, Nikolay Iskrev, Kevin J. Lansing, Caterina Mendicino. 2019. Inflation dynamics and adaptive expectations in an estimated DSGE model. *Journal of Macroeconomics* **59**, 258-277. [Crossref]
9. Michael Woodford. 2019. Monetary Policy Analysis When Planning Horizons Are Finite. *NBER Macroeconomics Annual* **33**, 1-50. [Crossref]
10. Oana Simona Hudea. 2019. The Romanian versus Eurozone economy via a DSGE model. *Economic Research-Ekonomska Istraživanja* **32**:1, 50-65. [Crossref]
11. Oana Simona Hudea. 2019. The behaviour of the Romanian economy analysed based on a dynamic stochastic general equilibrium model. *Economic Research-Ekonomska Istraživanja* **32**:1, 1645-1666. [Crossref]
12. Sebastian Poledna, Michael Gregor Miess, Cars H. Hommes. 2019. Economic Forecasting with an Agent-Based Model. *SSRN Electronic Journal* . [Crossref]
13. Daniel L Tortorice. 2018. The business cycle implications of fluctuating long run expectations. *Journal of Macroeconomics* **58**, 266-291. [Crossref]
14. Stefano Eusepi, Bruce Preston. 2018. Fiscal Foundations of Inflation: Imperfect Knowledge. *American Economic Review* **108**:9, 2551-2589. [Abstract] [View PDF article] [PDF with links]
15. Stefano Eusepi, Bruce Preston. 2018. The Science of Monetary Policy: An Imperfect Knowledge Perspective. *Journal of Economic Literature* **56**:1, 3-59. [Abstract] [View PDF article] [PDF with links]
16. Raluca-Elena Pop. Insights of Adaptive Learning Approach to Modeling Expectations: A Numerical Comparison with Adaptive Expectations and Rational Expectations 187-199. [Crossref]
17. Gavin Goy, Cars H. Hommes, Kostas Mavromatis. 2018. Forward Guidance and the Role of Central Bank Credibility under Heterogeneous Beliefs. *SSRN Electronic Journal* . [Crossref]
18. Marine Charlotte André, Meixing Dai. 2017. Is central bank conservatism desirable under learning?. *Economic Modelling* **60**, 281-296. [Crossref]

19. Stefano Eusepi, Bruce Preston. 2017. Fiscal Foundations of Inflation: Imperfect Knowledge. *SSRN Electronic Journal* . [Crossref]
20. Alessia Paccagnini. 2017. Dealing with Misspecification in DSGE Models: A Survey. *SSRN Electronic Journal* . [Crossref]
21. Seppo Honkapohja, Arja H. Turunen-Red, Alan D. Woodland. 2016. Growth, expectations and tariffs. *Canadian Journal of Economics/Revue canadienne d'économique* **49**:4, 1441-1469. [Crossref]
22. Jean-Sébastien Fontaine, René Garcia. Recent Advances in Old Fixed-Income Topics: Liquidity, Learning, and the Lower Bound 282-312. [Crossref]
23. Daniel L. Tortorice. 2016. The Business Cycles Implications of Fluctuating Long Run Expectations. *SSRN Electronic Journal* . [Crossref]
24. Federico Di Pace, Kaushik Mitra, Shoujian Zhang. 2016. Adaptive Learning and Labour Market Dynamics. *SSRN Electronic Journal* . [Crossref]
25. Thomas W.L. Norman. 2015. Learning, hypothesis testing, and rational-expectations equilibrium. *Games and Economic Behavior* **90**, 93-105. [Crossref]
26. Timothy Cogley, Christian Matthes, Argia M. Sbordone. 2014. Optimized Taylor Rules for Disinflation When Agents are Learning. *SSRN Electronic Journal* . [Crossref]
27. Stefano Eusepi, Bruce J. Preston. 2013. Fiscal Foundations of Inflation: Imperfect Knowledge. *SSRN Electronic Journal* . [Crossref]
28. Yuliya Rychalovska. 2013. The Implications of Financial Frictions and Imperfect Knowledge in the Estimated DSGE Model of the U.S. Economy. *SSRN Electronic Journal* . [Crossref]
29. Fabio Milani. The Modeling of Expectations in Empirical DSGE Models: A Survey 3-38. [Crossref]
30. Stefano Eusepi, Marc P. Giannoni, Bruce J. Preston. 2012. Long-Term Debt Pricing and Monetary Policy Transmission under Imperfect Knowledge. *SSRN Electronic Journal* . [Crossref]
31. Richard Harrison, Tim Taylor. 2012. Misperceptions, Heterogeneous Expectations and Macroeconomic Dynamics. *SSRN Electronic Journal* . [Crossref]