A Bayesian DSGE Model with Infinite-Horizon Learning: Do "Mechanical" Sources of Persistence Become Superfluous?

FABIO MILANI

University of California, Irvine

Abstract. This paper estimates a monetary DSGE model with learning introduced from the primitive assumptions. The model nests infinite-horizon learning and features, such as habit formation in consumption and inflation indexation, that are essential for the model fit under rational expectations.

I estimate the DSGE model by Bayesian methods, obtaining estimates of the main learning parameter, the constant gain, jointly with the deep parameters of the economy.

The results show that relaxing the assumption of rational expectations in favor of learning may render mechanical sources of persistence superfluous. In particular, learning appears a crucial determinant of inflation inertia.

Keywords: Infinite-Horizon Learning, DSGE model, Bayesian Estimation, Non-Rational Expectations, Inflation Persistence, Habit Formation.

JEL classification: C11, D84, E30, E50, E52.

Date: December, 2005.
Department of Economics, 3151 Social Science Plaza, University of California, Irvine, CA 92697-5100.
E-Mail: fmilani@uci.edu. *Homepage:* www.socsci.uci.edu/~fmilani.

1. Introduction

Recent DSGE models have proved successful in describing macroeconomic data. Smets and Wouters (2003, 2004, 2005) have provided the first example of a structural model that can compete in fit with unrestricted Bayesian VARs. Christiano, Eichenbaum, and Evans (2005), Giannoni and Woodford (2003), and Boivin and Giannoni (2005), have similarly developed models that approximate the impulse responses derived from VARs. The success of these papers stems from extending the simplest DSGE setup to include several features that help in generating endogenous persistence in their models. Modern DSGE models have increasingly followed their example. They typically incorporate habit formation in consumption, inflation and wage indexation, capital adjustment costs, and several autocorrelated disturbances. These additional sources of persistence, which we may view as "mechanical", together with persistent structural shocks are essential for the empirical success of the models.

Milani (2004b), however, shows that allowing for a minimal deviation from the conventional assumption of rational expectations might lead to reconsider the role of "mechanical" sources of persistence. In a model with subjective expectations and learning, in fact, the estimated degrees of habit formation in consumption and inflation indexation become negligible. Learning also improves the fit of a monetary DSGE model: the model with learning alone is preferred to the corresponding model with rational expectations, habits, and indexation.

Milani (2004b), following most of the adaptive learning literature (Evans and Honkapohja 2001, Bullard and Mitra 2002, among others), derives the model under rational expectations, and then he introduces subjective expectations and learning only on the linearized equations found under rational expectations. But Preston (2005a) argues that introducing learning directly from the primitives of the model would lead to different law of motions for inflation and output gap. The derived aggregate dynamics of the economy implies, in fact, that long-horizon expectations also matter. Preston (2005b) explains that decision rules that depend only on one-period-ahead expectations will generally not provide optimal decision rules under adaptive learning for the corresponding infinite horizon decision problems. The problem arises from the use of a different conditional distribution with respect to which expectations are taken. For example, he shows that the Euler equation under one-period-ahead learning would not satisfy the intertemporal budget constraint and, therefore, will lead to suboptimal decisions.

In this paper, I follow Preston's approach and build the model assuming subjective expectations from the primitives. I generalize Preston's framework to allow for habit formation in consumption and inflation indexation in price setting. Since Milani (2004b) shows that inserting learning in an optimizing DSGE model may make typical sources of persistence redundant, it is therefore important to verify if the results hold also when more attention is paid to the microfoundations of the model under learning.

I therefore derive a simple monetary DSGE model that incorporates infinite-horizon learning and mechanical sources of persistence, such as habit formation and inflation indexation. I then estimate the model using Bayesian methods. The paper provides the first estimation in the literature of a DSGE model with infinite-horizon learning. The main learning parameter, the constant gain, is jointly estimated with the 'deep' parameters of the economy. Estimation of the constant gain is crucial, for the empirical results often depend on the assumed gain, as shown in Milani (2004a), for example.

I find that infinite-horizon learning can generate substantial persistence in the model. When agents form subjective expectations and learn the relevant parameters, I find that the role of habit formation and indexation becomes smaller. Inflation indexation is superfluous. The persistence in inflation appears to be driven more by learning than by structural features such as indexation. Learning, in fact, substitutes for both indexation and a strong serial correlation in the exogenous cost-push shock. The results are less sharp for habit formation. In this case, the results depend on the assumed persistence in the aggregate demand disturbance. With a large autoregressive coefficient, habit formation becomes redundant. But with a small autoregressive coefficient, the model still needs a sizeable coefficient on habits to match the data.

2. A Microfounded Model with Adaptive Learning

I derive the aggregate dynamics of the economy introducing learning directly from the primitives of the model, as in Preston (2005a). This section generalizes Preston (2005a) by incorporating also habit formation in consumption and inflation indexation in price setting. In the model, agents know: 1) their own preferences; 2) the constraints they face; 3) how to solve their optimization problems. But they do not have any knowledge of other agents' preferences. Therefore, they are not able to infer the aggregate probability laws of the variables of interest, as they would be, instead, under rational expectations. To derive optimal decisions, agents need to form expectations about future macroeconomic variables. Here, I depart from the strong informational assumptions required by rational expectations, and I allow agents to form arbitrary subjective expectations.

2.1. **Households' Optimal Consumption Decisions.** The economy is populated by a continuum of households indexed by $i \in [0,1]$. Each household i maximizes the expected discounted utility

$$\widehat{E}_t^i \left\{ \sum_{T=t}^{\infty} \beta^{T-t} \left[U\left(C_T^i - \eta C_{T-1}^i ; \zeta_T\right) - \int_0^1 v(h_T^i(j); \zeta_T) dj \right] \right\} \tag{2.1}$$

where $\widehat{E}_t^i$ indicates subjective expectations for household i. Households derive utility from the deviation of current consumption C_T^i from a stock of internal habits in consumption ηC_{T-1}^i, and they derive disutility from the hours of labor supplied $h_T^i(j)$. An aggregate shock ζ_T may affect the consumption-leisure decision in each period. The coefficient $0 < \beta < 1$ denotes the usual discount factor, while η measures the degree of habit formation in consumption. The consumption index C_T^i is the Dixit-Stiglitz CES aggregator of different goods, so that $C_t^i \equiv \left[\int_0^1 c_t^i(j)^{\frac{\theta-1}{\theta}} dj\right]^{\frac{\theta}{\theta-1}}$, and the associated price index is $P_t \equiv \left[\int_0^1 p_t(j)^{1-\theta} dj\right]^{\frac{1}{1-\theta}}$, where θ is the elasticity of substitution between differentiated goods. For simplicity, I assume homogeneous beliefs across agents (although this is *not known* to agents, who do not have any information about other agents' beliefs). As standard in the adaptive learning literature, the subjective expectations of individual agents obey the law of iterated expectations, $\widehat{E}_t^i \widehat{E}_{t+s}^i z = \widehat{E}_t^i z$ for any variable z.

I follow Preston (2005a) in assuming *incomplete* asset markets.[1] Agents can use a single one-period riskless asset to transfer wealth intertemporally. The flow budget constraint is given by:

$$M_t^i + B_t^i \leq (1 + i_{t-1}^m) M_{t-1}^i + (1 + i_{t-1}) B_{t-1}^i + P_t Y_t^i - T_t - P_t C_t^i \qquad (2.2)$$

where M_t^i denotes end-of-period money holdings, B_t^i end-of-period riskless bond holdings, i_t^m and i_t denote nominal interest rates on money and bonds, and T_t are lump sum taxes and transfers. Y_t^i is household's real income in period t, given by $\int_0^1 \left[w_t(j) h_t^i(j) + \Pi_t(j)\right] dj$, where $w_t(j)$ represents the wage received by the household for labor supplied in the production of good j and $\Pi_t(j)$ the share of profits received from the sale of each firm's good j (households own an equal share of all the firms).

The intertemporal budget constraint (IBC) is

$$\widehat{E}_t^i \sum_{T=t}^{\infty} \beta^{T-t} C_T^i = \omega_t^i + \widehat{E}_t^i \sum_{T=t}^{\infty} \beta^{T-t} Y_T^i \qquad (2.3)$$

where $\omega_t^i \equiv \frac{W_t^i}{P_t \bar{Y}}$ is the share of nominal wealth ($W_t^i \equiv (1 + i_{t-1}) B_{t-1}^i$) as a fraction of nominal steady-state income. With habit formation, the first

[1] This assumption limits the extent of information revelation from prices.

order conditions become

$$\begin{aligned} \lambda_t^i &= U_c\left(C_T^i - \eta C_{T-1}^i; \zeta_T\right) - \beta\eta \widehat{E}_t^i\left[U_c\left(C_{T+1}^i - \eta C_T^i; \zeta_{T+1}\right)\right] && (2.4) \\ \lambda_t^i &= \beta \widehat{E}_t^i\left[\lambda_{t+1}^i (1+i_t) P_t / P_{t+1}\right] && (2.5) \end{aligned}$$

where λ_t^i is the marginal utility of real income in period t. Substituting (2.4) into (2.5), and taking a log-linear approximation of the implied Euler equation, I obtain

$$\widetilde{C}_t^i = \widehat{E}_t^i \widetilde{C}_{t+1}^i - (1-\beta\eta)\,\sigma\left(i_t - \widehat{E}_t^i \pi_{t+1}\right) + g_t - \widehat{E}_t^i g_{t+1} \tag{2.6}$$

where

$$\widetilde{C}_t^i = C_t^i - \eta C_{t-1}^i - \beta\eta \widehat{E}_t^i\left[C_{t+1}^i - \eta C_t^i\right] \tag{2.7}$$

and where $\sigma \equiv \frac{-U_c}{(\bar{C} U_{cc})} > 0$ represents the elasticity of intertemporal substitution of consumption in the absence of habit formation, and $g_t \equiv \frac{\sigma U_{c\zeta}\zeta_t}{U_c}$ is a preference shock. Solving (2.6) backwards, taking expectations, substituting into the modified IBC,[2] using $C_t = Y_t$, and expressing everything in terms of the output gap $x_t \equiv Y_t - Y_t^n$ yields the aggregate demand equation[3]

$$\widetilde{x}_t = \widehat{E}_t \sum_{T=t}^{\infty} \beta^{T-t}\left[(1-\beta)\widetilde{x}_{T+1} - (1-\eta\beta)\sigma\left(i_T - \pi_{T+1} - r_T^n\right)\right] \tag{2.8}$$

where

$$\widetilde{x}_t \equiv (x_t - \eta x_{t-1}) - \beta\eta \widehat{E}_t\left(x_{t+1} - \eta x_t\right)$$

and where Y_t^n is the natural rate of output (the equilibrium level of output under flexible prices) and $r_T^n \equiv [(1-\eta\beta)\sigma]^{-1}\left[\left(Y_{t+1}^n - g_{t+1}\right) - \left(Y_t^n - g_t\right)\right]$ is the flexible-price equilibrium real interest rate. Current output gap, therefore, depends on lagged and expected one-period ahead output gap, on the ex-ante real interest rate, plus on long-horizon forecasts of future output gaps, real interest rates, and disturbances until the indefinite future.

[2]Found by substituting $C_t^i = \widetilde{C}_t^i + \eta C_{t-1}^i + \beta\eta \widehat{E}_t^i\left[C_{t+1}^i - \eta C_t^i\right]$ into the IBC.

[3]In the derivation, I also use $\int_i \omega_t^i di = 0$ from bond's market clearing, and I integrate over the i households, using $C_t = \int_i C_t^i di$, $Y_t = \int_i Y_t^i di$, and $\widehat{E}_t[\cdot] \equiv \int_i \widehat{E}_t^i[\cdot]\, di$, which denotes average private-sector expectations.

2.2. **Firms' Problem.** I assume Calvo price-setting. A fraction $0 < 1-\alpha < 1$ of firms can set prices optimally in a given period t. The remaining α firms that are not allowed to optimize in t can still adjust their prices following the indexation rule proposed by Christiano, Eichenbaum and Evans (2005):

$$\log p_t(i) = \log p_{t-1}(i) + \gamma\pi_{t-1} \tag{2.9}$$

where the parameter $0 \leq \gamma \leq 1$ measures the degree of indexation to past inflation. The aggregate price index P_t evolves according to

$$P_t = \left[\alpha\left(P_{t-1}\left(\frac{P_{t-1}}{P_{t-2}}\right)^{\gamma}\right)^{1-\theta} + (1-\alpha)p_t^{*1-\theta}\right]^{\frac{1}{1-\theta}}. \tag{2.10}$$

Each firm i maximizes the expected present discounted value of future profits $\Pi_T^i(\cdot)$

$$\widehat{E}_t^i\left\{\sum_{T=t}^{\infty}\alpha^{T-t}Q_{t,T}\left[\Pi_T^i\left(p_t^*(i)\left(\frac{P_{T-1}}{P_{t-1}}\right)^{\gamma}\right)\right]\right\} \tag{2.11}$$

where a unit of income in date T is valued by the stochastic discount factor $Q_{t,T} = \beta^{T-t}\frac{P_t}{P_T}\frac{\lambda_T}{\lambda_t}$.

The first-order conditions for the problem are

$$\widehat{E}_t^i\left\{\begin{array}{c}\sum_{T=t}^{\infty}(\alpha\beta)^{T-t}\lambda_T Y_T P_T^{\theta}\left(\frac{P_{T-1}}{P_{t-1}}\right)^{\gamma(1-\theta)}\cdot \\ \left[p_t^*(i) - \mu P_T s\left(Y_T\left(\frac{\widehat{p}_t^*(i)}{P_T}\right)^{-\theta}\left(\frac{P_{T-1}}{P_{t-1}}\right)^{-\gamma\theta}, Y_T; \widetilde{\zeta}_T\right)\right]\end{array}\right\} = 0 \tag{2.12}$$

where $\mu = \theta/(\theta-1) > 1$, $\widetilde{\zeta}_t$ is a vector of exogenous real disturbances incorporating both preference shocks ζ_t and technology shocks A_t, and where $s(\cdot)$ is firm i's real marginal cost function.

Log-linearization of the first-order condition yields

$$\widehat{p}_t^*(i) = \widehat{E}_t^i\sum_{T=t}^{\infty}(\alpha\beta)^{T-t}\left[\frac{1-\alpha\beta}{1+\omega\theta}\left(\omega Y_T - \lambda_T^i + \frac{v_{y\zeta}}{v_y}\zeta_T\right) + \alpha\beta\left(\pi_{T+1} - \gamma\pi_T\right)\right] \tag{2.13}$$

where $\widehat{p}_t^* \equiv \log(p_t^*/P_t)$ and $\omega \equiv v_{yy}\overline{Y}/v_y$ is the elasticity of the marginal disutility of producing output with respect to an increase in output.

From a log-linear approximation of the aggregate price index and integrating over the i firms, I can derive the aggregate supply relation

$$\begin{aligned}\widetilde{\pi}_t &= \xi_p\left(\omega x_t + \left[(1-\eta\beta)\sigma\right]^{-1}\widetilde{x}_t\right) + \widehat{E}_t\sum_{T=t}^{\infty}(\alpha\beta)^{T-t} \\ &\quad \left[\alpha\beta\xi_p\left(\omega x_{T+1} + \left[(1-\eta\beta)\sigma\right]^{-1}\widetilde{x}_{T+1}\right) + (1-\alpha)\beta\widetilde{\pi}_{T+1} + u_T\right]\end{aligned} \tag{2.14}$$

where

$$\begin{aligned}\widetilde{\pi}_t &\equiv \pi_t - \gamma\pi_{t-1} \\ \widetilde{x}_t &\equiv (x_t - \eta x_{t-1}) - \beta\eta\widehat{E}\left(x_{t+1} - \eta x_t\right) \\ \xi_p &= \frac{(1-\alpha)(1-\alpha\beta)}{\alpha(1+\omega\theta)}\end{aligned}$$

Current inflation therefore depends on lagged inflation, current, lagged and one-period ahead output gaps, and on the long-horizon forecasts of future output gaps, inflation rates, and supply shocks. Deviations of the empirical output gap from the theoretically relevant gap will show up in the supply shock u_t.

2.3. **Monetary Authority.** I assume that the following Taylor rule with partial adjustment describes monetary policy in this economy

$$i_t = \rho i_{t-1} + (1-\rho)\left[\chi_\pi\pi_t + \chi_x x_t\right] + \varepsilon_t$$

where ρ denotes the degree of interest-rate smoothing, ψ_π and ψ_x are feedback coefficients, and ε_t accounts for unanticipated deviations from systematic monetary policy.

3. Infinite-Horizon Learning

With learning introduced as in Preston (2005a,b), long-horizon expectations also matter. In the previous section, I have generalized Preston's framework to include habit formation and indexation. The model economy can be summarized as

$$\widetilde{x}_t = \widehat{E}_t \sum_{T=t}^{\infty} \beta^{T-t} \left[(1-\beta)\, \widetilde{x}_{T+1} - (1-\eta\beta)\sigma(i_T - \pi_{T+1} - r_T^n)\right] \tag{3.1}$$

$$\widetilde{\pi}_t = \widehat{E}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \left[\xi_p \left(\omega x_T + [(1-\eta\beta)\sigma]^{-1} \widetilde{x}_T\right) + \right. \tag{3.2}$$
$$\left. + (1-\alpha)\, \beta \widetilde{\pi}_{T+1} + u_T\right]$$

$$i_t = \rho i_{t-1} + (1-\rho)\left[\chi_\pi \pi_t + \chi_x x_t\right] + \varepsilon_t \tag{3.3}$$

$$r_t^n = \phi_r r_{t-1}^n + \nu_t^r \tag{3.4}$$

$$u_t = \phi_u u_{t-1} + \nu_t^u \tag{3.5}$$

where $\widetilde{x}_t$ and $\widetilde{\pi}_t$ have the usual meaning. I have assumed that the disturbances r_t^n and u_t follow autoregressive processes. The shocks ε_t, ν_t^r, ν_t^u are *i.i.d. Normal* with mean 0 and variance-covariance matrix Q.[4]

From (3.1) and (3.2), it is clear that economic agents need to form forecasts of macroeconomic variables until the indefinite future. I follow a number of papers in the adaptive learning literature (see Evans and Honkapohja 2001 for a comprehensive treatment) and assume that agents use simple linear economic models to form expectations. The agents have the following "Perceived Law of Motion" (PLM)

$$Z_t = a_t + b_t Z_{t-1} + c_t r_t^n + d_t u_t + \varepsilon_t, \tag{3.6}$$

where $Z_t \equiv [\pi_t, x_t, i_t]'$ and a_t, b_t, c_t, d_t are coefficient vectors and matrices of appropriate dimensions. The PLM has the same structural form of the rational expectations solution of the system, i.e. it includes the same regressors that appear in the Minimum State Variable (MSV) solution under rational expectations. The agents, however, lack knowledge about the parameters of the model. Therefore, they use historical data to learn the parameters over

[4] When learning is introduced on the linearized equations found under RE, the aggregate demand and supply equations become $\widetilde{x}_t = \widehat{E}_t \widetilde{x}_{t+1} - (1-\beta\eta)\sigma\left[i_t - \widehat{E}_t \pi_{t+1} - r_t^n\right]$ and $\widetilde{\pi}_t = \xi_p\left[\omega x_t + [(1-\eta\beta)\sigma]^{-1}\widetilde{x}_t\right] + \beta \widehat{E}_t \widetilde{\pi}_{t+1} + u_t$. I refer the reader to Preston (2005a,b) and Honkapohja, Mitra, and Evans (2003), for a discussion of the different approaches.

time. As soon as they observe additional data, agents update their estimates of the parameter vector (a_t, b_t, c_t, d_t) through constant-gain learning, as described by the following formulas

$$\widehat{\boldsymbol{\phi}}_t = \widehat{\boldsymbol{\phi}}_{t-1} + \overline{\mathbf{g}} R_{t-1}^{-1} X_t (Z_t - X_t' \widehat{\boldsymbol{\phi}}_{t-1}) \tag{3.7}$$

$$R_t = R_{t-1} + \overline{\mathbf{g}} (X_{t-1} X_{t-1}' - R_{t-1}) \tag{3.8}$$

where (3.7) describes the updating of the learning rule coefficients $\widehat{\boldsymbol{\phi}}_t = (a_t', vec(b_t, c_t, d_t)')'$, and (3.8) describes the updating of the matrix of second moments R_t of the stacked regressors $X_t \equiv \{1, Z_{t-1}, u_t, r_t^n\}_0^{t-1}$. The parameter $\overline{\mathbf{g}}$ denotes the constant gain, which indicates the speed at which agents update their beliefs. From their PLM, and using the updated parameters through (3.7) and (3.8), agents can form expectations for any future horizon $T > t$ as

$$\begin{aligned} \widehat{E}_t Z_T = {} & (I_5 - b_{t-1})^{-1} (I_5 - b_{t-1}^{T-t}) a_{t-1} + b_{t-1}^{T-t} E_t Z_t + \\ & + \phi_r r_t^n (\phi_r I_5 - b_{t-1})^{-1} (\phi_r^{T-t} I_5 - b_{t-1}^{T-t}) c_{t-1} + \\ & + \phi_u u_t (\phi_u I_5 - b_{t-1})^{-1} (\phi_u^{T-t} I_5 - b_{t-1}^{T-t}) d_{t-1} \end{aligned} \tag{3.9}$$

where I_5 is a 5×5 identity matrix.

4. Bayesian Estimation

The paper provides the first empirical analysis of a model with Infinite-Horizon learning. I estimate the system using Bayesian methods to fit the series for output gap, inflation and the nominal interest rate. I use quarterly U.S. data for the period 1960:I to 2004:II. Inflation is defined as the annualized quarterly rate of change of the GDP Implicit Price Deflator, output gap as the log difference between GDP and Potential GDP (CBO estimate), and I use the federal funds rate as the nominal interest rate.

The main learning parameter, the constant gain, is estimated *jointly* with the deep parameters of the economy. I can substitute the expectations formed as in (3.9) into (3.1) and (3.2) and re-write the model in state-space form

$$\xi_t = A_t + F_t \xi_{t-1} + G_t w_t \tag{4.1}$$
$$Y_t = H \xi_t$$

where $\xi_t = [x_t, \pi_t, i_t, u_t, r_t^n]$, $w_t \sim N(0, Q)$, H is a matrix of zeros and ones selecting observables from ξ_t, and A_t, F_t, G_t are time-varying matrices of coefficients, which are convolutions of structural parameters of the economy and agents' beliefs. Expression (4.1) represents the "Actual Law of Motion" (ALM) of the economy: the ALM has the same structural form as the PLM, but possibly different parameter values. Having expressed the model as a linear Gaussian system, I can evaluate the likelihood function using the Kalman Filter. To derive the parameter estimates, I use a Random-Walk Metropolis-Hastings algorithm to generate draws from the posterior distribution.[5] I generate $300,000$ draws with an initial burn-in of $60,000$ draws. A similar estimation procedure has been used by several recent papers that focus on DSGE models under rational expectations (see An and Schorfheide (2006) for a first survey of this literature). This paper, instead, exploits similar techniques to provide the first estimation of a DSGE model with infinite-horizon learning.

I collect the structural parameters in the vector Ψ:

$$\Psi = \left\{\eta, \beta, \alpha, \sigma, \gamma, \xi_p, \omega, \rho, \chi_\pi, \chi_x, \phi_r, \phi_u, \sigma_\varepsilon, \sigma_r, \sigma_u, \sigma_{\varepsilon,r}, \sigma_{\varepsilon,u}, \sigma_{r,u}, \overline{\mathbf{g}}\right\}$$

I fix some of the parameters: $\beta = 0.99$, $\xi_p = 0.0015$, and $\omega = 0.8975$ (ξ_p and ω are fixed at the values estimated in Giannoni and Woodford 2003 for the flexible wages case). I fix the autoregressive parameters ϕ_r and ϕ_u to 0.9 (I will also consider the case $\phi_r = \phi_u = 0.1$).

Table 1 presents information about the priors. The habit and indexation parameters η and γ are assumed to follow Uniform distributions in the interval $[0, 1]$. The intertemporal elasticity of substitution coefficient

[5] More details about the estimation method can be found in Milani (2004b).

σ follows a Gamma distribution with mean 0.125 and standard deviation 0.09. I choose inverse gamma distributions for the standard deviations of the shocks. The constant-gain coefficient follows a Gamma distribution with prior mean 0.031 and prior standard deviation 0.022.

I estimate the initial conditions for the learning algorithm using pre-sample data for the 1954:III-1959:IV period. The evolution of agents' beliefs is shown in figure 1 and 2, together with the 95% probability bands. For example, we see that agents perceive inflation as more persistent starting in the second half of the 1970s until the first half of the 1980s (parameter b_{22}), and they perceive a smaller sensitivity of output to interest rates after 1980 (parameter b_{13}).

Table 1 presents the estimation results. First, I assume that the autoregressive coefficients regarding the disturbances r_t^n and u_t equal 0.9. I find very weak evidence of habit formation in consumption and no evidence of indexation in inflation. I estimate, in fact, η, the habit parameter, equal to 0.113, while I estimate γ, the inflation indexation parameter, equal to 0.009. The two parameters are tightly estimated: the 95% posterior probability intervals also remain close to zero. Therefore, infinite-horizon learning appears to account for the persistence in the data. Additional "mechanical" sources of persistence, which are essential under rational expectations, become superfluous under learning. Under infinite-horizon learning, however, the estimate of α, the Calvo price-stickiness parameter, is unrealistic: I find α equal to 0.992, which implies an extreme degree of rigidity in prices. I obtain a value of 0.067 for the intertemporal elasticity of substitution parameter σ. The estimates for the monetary policy rule ballpark most estimates in the literature ($\rho = 0.91$, $\chi_\pi = 1.52$, and $\chi_x = 0.68$). A crucial parameter in the estimation is represented by the constant-gain parameter. The paper estimates the constant gain jointly with the deep parameters of the economy. I estimate the gain equal to 0.006. To get some intuition about this value, it may be useful to think about the gain as a rough indication of how many observations agents use to form their expectations. A gain equal to 0.006, therefore, mimics the situation of an econometrician running rolling-window regressions using a window with 1/0.006 observations (corresponding to 166.67 quarters of data, or 41.668 years). The estimated value implies a substantially slower learning than what found in Milani (2004a,b, 2005a) assuming only one-period ahead expectations.

Infinite-horizon learning, therefore, weakens the role of habits and indexation in a model where the disturbances are highly persistent. But can learning also substitute for the typically strongly autocorrelated structural disturbances? Here, I re-estimate the model by fixing the autoregressive parameters ϕ_r and ϕ_u to 0.1. I obtain different results for habits and indexation. In the case of habits, the results seem to depend on the assumed persistence of the disturbances. When the assumed autocorrelation is low, a large degree of habit formation in consumption is still needed to fit the data (I find $\eta = 0.87$). The results are more favorable for inflation indexation. Even assuming a low autocorrelation of the disturbances, the estimated indexation is small (I estimate $\gamma = 0.21$). The results suggest that learning matters for inflation dynamics. A minimal deviation from rational expectations is sufficient to account for the persistence in inflation, so that both indexation and a strongly autocorrelated cost-push shock become redundant.

The estimated Calvo parameter α is now small: I find α equal to 0.19 in this case, suggesting a much smaller price rigidity. The results about α are therefore strongly dependent on the assumed autocorrelation and suggest difficulties in robustly identifying this parameter. I obtain different results also for the constant gain. Now, the gain coefficient equals 0.017. This estimate implies faster learning than in the previous case and is more similar to what found by Milani (2004a,b, 2005a). In general, various recent papers (Milani 2004a,b, 2005a, Orphanides and Williams 2005a,b, Branch and Evans 2005) are starting to accumulate evidence that the most realistic values of the gain lie in the $0.01 - 0.05$ range, with the majority of estimates around 0.02.

I also re-estimate the model under Recursive Least Squares learning: this implies a decreasing gain equal to t^{-1}. The results substantially confirm what found under constant-gain learning. When the autocorrelation of the shocks is large, I estimate $\eta = 0.059$ and $\gamma = 0.01$. The Calvo price-stickiness parameter is now slightly less extreme ($\alpha = 0.911$). When, the autocorrelation is small, instead, I find $\eta = 0.786$ and $\gamma = 0.134$. The Calvo parameter is again reduced to 0.18.[6]

5. Conclusions

DSGE models under rational expectations typically need several additional sources of persistence to match macroeconomic data. In the paper, I have developed a model in which non-rational expectations and learning enter from the primitive assumptions. As in Preston (2005a), the aggregate dynamics of the economy implies that long-horizon expectations of future macroeconomic conditions matter for the current dynamics of output, inflation, and nominal interest rates. The model, therefore, nests infinite-horizon learning and some of the "mechanical" sources of persistence, such as habit formation and inflation indexation, that are essential under rational expectations. Once the assumption of rational expectations is relaxed in favor of learning, it becomes interesting to verify whether mechanical sources of persistence remain essential for the model fit.

I estimate the model using Bayesian methods. I obtain estimates of the main learning parameter, the constant gain, jointly with the other model parameters.

[6]I do not report the results, but I have found that the model with constant-gain learning fits better than the model with RLS learning.

The results show that learning may render additional sources of persistence superfluous. Learning seems to represent the main cause of persistence in inflation: with learning, the estimated indexation is very close to zero. The results do not depend on the assumed autocorrelation of the shocks. Infinite-horizon learning generates sufficient persistence in inflation, so that it might be possible to avoid both indexation and serial correlation in the cost-push shock. The results are, instead, mixed for habit formation: learning and strongly autocorrelated shocks substantially weaken the evidence of habit formation. But the results in this case depend on the assumed autocorrelation. A low autocorrelation restores, in fact, a role to habit formation.

Overall, learning seems to provide a good description of the data. But the literature still needs to shed more light on the best way to model learning. In related research, I am comparing the estimates and fit of DSGE models under different learning mechanisms: one-period-ahead versus infinite-horizon learning, constant-gain versus recursive-least-squares learning, and different learning rule specifications. Moreover, as Preston (2004a,b) shows, monetary policy rules may have very different properties under different learning mechanisms. A priority for future research, therefore, will consist of evaluating the robustness of policy rules to different assumptions about learning.

References

[1] An, S., and F. Schorfheide, (2006). "Bayesian Analysis of DSGE Models", *Econometric Reviews*, forthcoming.

[2] Boivin, J., and M. Giannoni, (2005). "Has Monetary Policy Become More Effective?", *Review of Economics and Statistics*, forthcoming.

[3] Branch, W.A., and G.W. Evans, (2005). "A Simple Recursive Forecasting Model", forthcoming, *Economics Letters*.

[4] Bullard, J., and K. Mitra, (2002). "Learning About Monetary Policy Rules", *Journal of Monetary Economics*, 49, pp. 1105-1129.

[5] Christiano, L.J., Eichenbaum, M., and C.L. Evans, (2005). "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy", *Journal of Political Economy*, Vol. 113 (1).

[6] Giannoni, M., and M. Woodford, (2003). "Optimal Inflation Targeting Rules", Ben S. Bernanke and Michael Woodford, eds., *Inflation Targeting*, Chicago: University of Chicago Press.

[7] Honkapohja, S., Mitra, K., and G. Evans, (2003). "Notes on Agent's Behavioral Rules Under Adaptive Learning and Recent Studies of Monetary Policy", unpublished manuscript.
[8] Milani, F., (2004a). "Adaptive Learning and Inflation Persistence", mimeo, Princeton University.
[9] ——— (2004b). "Expectations, Learning and Macroeconomic Persistence", mimeo, Princeton University.
[10] ——— (2005a). "Learning, Monetary Policy Rules, and Macroeconomic Stability", mimeo, Princeton University.
[11] ——— (2005b). "Estimating and Comparing New Keynesian Models with Adaptive Learning: Euler-Equation versus Infinite-Horizon Learning", in progress.
[12] Orphanides, A., and J. Williams, (2003a). "Imperfect Knowledge, Inflation Expectations and Monetary Policy", in Ben Bernanke and Michael Woodford, eds., *Inflation Targeting*. Chicago: University of Chicago Press.
[13] ——— (2005a). "Inflation Scares and Forecast-Based Monetary Policy", *Review of Economic Dynamics*, Vol. 8 (2), pages 498-527.
[14] ——— (2005b). "The Decline of Activist Stabilization Policy: Natural Rate Misperceptions, Learning, and Expectations", *Journal of Economic Dynamics and Control*, vol. 29, issue 11, pages 1927-1950.
[15] Preston, B., (2004a). "Adaptive Learning and the Use of Forecasts in Monetary Policy", forthcoming, *Journal of Monetary Economics*.
[16] ——— (2004b). "Adaptive Learning, Forecast-Based Instrument Rules and Monetary Policy", mimeo, Columbia University.
[17] ——— (2005a). "Learning About Monetary Policy Rules When Long-Horizon Expectations Matter", *International Journal of Central Banking,* Iss. 2.
[18] ——— (2005b). "Adaptive Learning in Infinite Horizon Decisions Problems", mimeo, Columbia University.
[19] Smets, F., and R. Wouters, (2003). "Monetary Policy in an Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area", *Journal of the European Economic Association*, v. 1, iss. 5, pp. 1123-75.
[20] ——— (2004). "Forecasting with a Bayesian DSGE Model: An Application to the Euro Area", *Journal of Common Market Studies,* 42 (4), 841-867.
[21] ——— (2005). "Comparing Shocks and Frictions in US and Euro Business Cycles: A Bayesian DSGE Approach", *Journal of Applied Econometrics*, Vol. 20, No. 2, 2005, pp. 161-183.
[22] Svensson, L.E.O., (2003). "Monetary Policy and Learning", *Federal Reserve Bank of Atlanta Economic Review*, Third Quarter, 11-16.
[23] Woodford, M., (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.

		Prior Distribution				**Posterior Distribution**							
						Constant-Gain Learning				**RLS Learning**			
Description	Param.	Range	Distr.	Mean	95% Int.	Mean	95% Int.	Mean	95% Int.	Mean	95% Int.	Mean	95% Int.
Habit Formation	η	$[0,1]$	U	.5	$[.025,.975]$	.113	$[.14,.23]$	.87	$[.72,.97]$	.059	$[.002,.15]$	.786	$[.64,.92]$
Discount Rate	β	.99	–	.99	–	.99	–	.99	–	.99	–	.99	–
Calvo Parameter	α	$[0,1]$	U	.5	$[.025,.975]$	.992	$[.97,.999]$	.19	$[.02,.42]$	.911	$[.80,.98]$	.18	$[.02,.43]$
IES	σ	$\mathbb{R}^+$	G	.125	$[.015,.35]$	.067	$[.04,.10]$	.144	$[.04,.36]$	.077	$[.04,.12]$	.239	$[.09,.44]$
Infl. Indexation	γ	$[0,1]$	U	.5	$[.025,.975]$	.009	$[0,.033]$	.216	$[0,.77]$	.01	$[0,.038]$	.134	$[.005,.43]$
Fcn. of Price Stick.	ξ_p	.0015	–	.0015	–	.0015	–	.0015	–	.0015	–	.0015	–
Elasticity mc	ω	.8975	–	.8975	–	.8975	–	.8975	–	.8975	–	.8975	–
Interest-Rate Smooth	ρ	$[0,.97]$	U	.485	$[0.024,0.946]$	.91	$[.87,.95]$	.885	$[.83,.93]$	.903	$[.86,.94]$	.877	$[.83,.92]$
Feedback to Infl.	χ_π	$\mathbb{R}$	N	1.5	$[1.01,1.99]$	1.523	$[1.14,1.92]$	1.496	$[1.14,1.87]$	1.58	$[1.21,1.95]$	1.63	$[1.29,1.99]$
Feedback to Output	χ_x	$\mathbb{R}$	N	.5	$[.01,.99]$	.681	$[.30,1.08]$	.56	$[.18,.97]$	.66	$[.23,1.04]$	.54	$[.19,.95]$
Autocorr. r_t^n	ϕ_r	.9 or .1	–	.–	–	.9	–	.1	–	.9	–	.1	–
Autocorr. u_t	ϕ_u	.9 or .1	–	.–	–	.9	–	.1	–	.9	–	.1	–
Std. MP shock	σ_ε	$\mathbb{R}^+$	IG	1	$[.34,2.81]$	.889	$[.8,.99]$	.89	$[.8,.99]$	.887	$[.8,.98]$	.891	$[.8,.99]$
Std. r_t^n	σ_r	$\mathbb{R}^+$	IG	1	$[.34,2.81]$	.856	$[.77,.95]$	.82	$[.74,.92]$	.844	$[.76,.94]$	.882	$[.8,.98]$
Std. u_t	σ_u	$\mathbb{R}^+$	IG	1	$[.34,2.81]$	1.56	$[1.4,1.73]$	1.47	$[1.18,2.1]$	1.578	$[1.42,1.76]$	1.36	$[1.13,1.72]$
Cov ε_t, r_t^n	$\sigma_{\varepsilon,r}$	$[-.5,.5]$	U	0	$[-.475,.475]$	.04	$[-.09,.17]$	$-.08$	$[-.27,.06]$	.04	$[-.11,.17]$	$-.008$	$[-.15,.13]$
Cov ε_t, u_t	$\sigma_{\varepsilon,u}$	$[-.5,.5]$	U	0	$[-.475,.475]$	.315	$[.17,.46]$	.29	$[.16,.43]$	.312	$[.17,.47]$	.32	$[.17,.45]$
Cov r_t^n, u_t	$\sigma_{r,u}$	$[-.5,.5]$	U	0	$[-.475,.475]$	.03	$[-.13,.18]$	$-.06$	$[-.23,.11]$	$-.025$	$[-.19,.12]$	$-.07$	$[-.23,.09]$
Constant Gain	$\overline{\mathbf{g}}$	$\mathbb{R}^+$	G	.031	$[.0038,.087]$	.006	$[.0014,.01]$	.017	$[.007,.036]$	–	–	–	–

Table 1 - Bayesian DSGE Model with Infinite-Horizon Learning: prior distributions, posterior estimates, and 95% probability intervals.

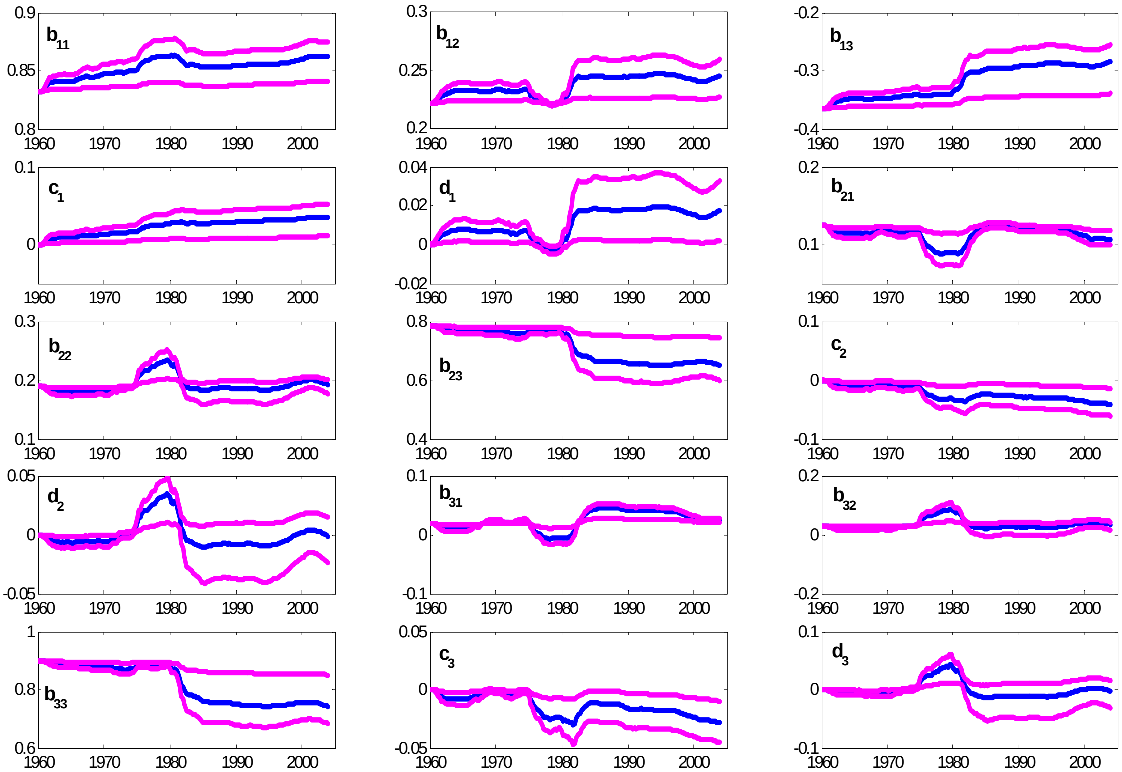

Figure 1 - Agents' time-varying beliefs 1960:I-2004:III (Autoregressive parameters = 0.9, CGL).

The learning rule is:

$$\begin{bmatrix} x_t \\ \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} b_{11,t} & b_{12,t} & b_{13,t} \\ b_{21,t} & b_{22,t} & b_{23,t} \\ b_{31,t} & b_{32,t} & b_{33,t} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ \pi_{t-1} \\ i_{t-1} \end{bmatrix} + \begin{bmatrix} c_{1,t} \\ c_{2,t} \\ c_{3,t} \end{bmatrix} r_t^n + \begin{bmatrix} d_{1,t} \\ d_{2,t} \\ d_{3,t} \end{bmatrix} u_t + \varepsilon_t.$$

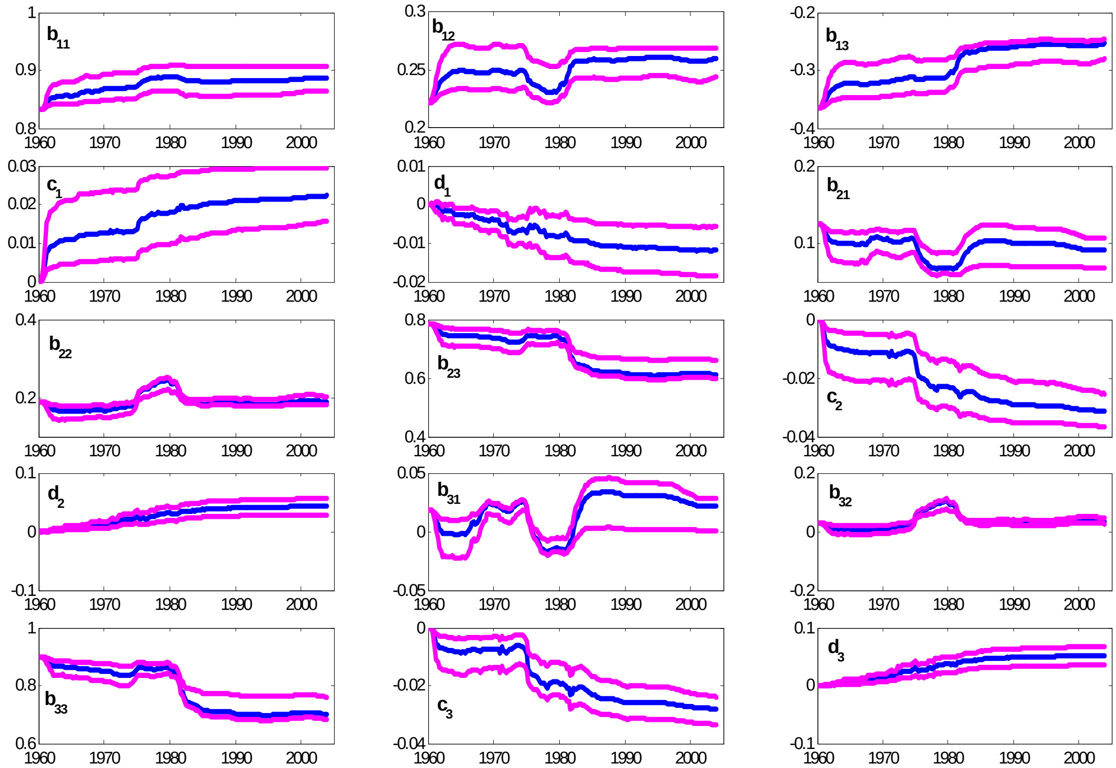

Figure 2 - Agents' time-varying beliefs 1960:I-2004:III (Autoregressive parameters = 0.1, CGL).

The learning rule is:

$$\begin{bmatrix} x_t \\ \pi_t \\ i_t \end{bmatrix} = \begin{bmatrix} b_{11,t} & b_{12,t} & b_{13,t} \\ b_{21,t} & b_{22,t} & b_{23,t} \\ b_{31,t} & b_{32,t} & b_{33,t} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ \pi_{t-1} \\ i_{t-1} \end{bmatrix} + \begin{bmatrix} c_{1,t} \\ c_{2,t} \\ c_{3,t} \end{bmatrix} r_t^n + \begin{bmatrix} d_{1,t} \\ d_{2,t} \\ d_{3,t} \end{bmatrix} u_t + \varepsilon_t.$$