

ADAPTIVE LEARNING AND INFLATION PERSISTENCE

Fabio Milani*

University of California, Irvine

First draft: March, 2004

This draft: November, 2005

Abstract

What generates persistence in inflation? Is inflation persistence structural?

This paper investigates learning as a potential source of persistence in inflation. The paper focuses on the price-setting problem of firms and presents a model that nests structural sources of persistence (indexation) and learning. Indexation is typically necessary under rational expectations to match the inertia in the data and to improve the fit of estimated New Keynesian Phillips curves.

The empirical results show that when learning replaces the assumption of fully rational expectations, structural sources of persistence in inflation, such as indexation, become unsupported by the data. The results suggest learning behavior as the main source of persistence in inflation. This finding has implications for the optimal monetary policy.

The paper also shows how one's results can heavily depend on the assumed learning speed. The estimated persistence and the model's fit, in fact, vary across the whole range of constant gain values. The paper therefore provides evidence on the best-fitting constant gain in the sample, and on its variations over possible learning specifications and over time.

Keywords: adaptive learning, inflation persistence, sticky prices, best-fitting constant gain, learning speed, expectations.

JEL classification: D84, E30, E50.

*I am grateful to Michael Woodford, John Faust, Michael Kiley, Athanasios Orphanides, Giorgio Primiceri, Chris Sims, Lars Svensson, and Noah Williams for comments and helpful discussions. I would also like to thank participants at seminars at Princeton University and the Federal Reserve Board. *Address for correspondence:* Fabio Milani - Department of Economics, 3151 Social Science Plaza, University of California, Irvine, CA 92697-5100. *E-mail address:* fmilani@uci.edu; *Homepage:* www.socsci.uci.edu/~fmilani.

1. INTRODUCTION

What creates **persistence** in inflation? Is inflation persistence a **structural** characteristic of industrialized economies?

Despite a large body of literature on inflation dynamics, disagreement still exists on the main sources of inflation persistence. It is well known that sticky price models, typically used to study inflation, fail to match the observed persistence. The New Keynesian model, at least in baseline form, describes inflation as a purely forward-looking variable. All of its persistence, therefore, is inherited from the persistence of an exogenous driving variable, typically a measure of the output gap or real marginal costs, or from an exogenous autocorrelated cost-push shock.

Researchers have proposed various extensions to allow their models to endogenously generate some persistence to match actual inflation data. These extensions have the effect of introducing the dependence of current inflation on *lagged inflation*. These additional channels of inertia have been variously modeled in the literature by incorporating rule-of-thumb behavior, quadratic adjustment costs or indexation to past inflation. Galí and Gertler (1999), for example, allow for the existence of a fraction of firms that deviate from full rationality and set, instead, their prices using simple rules of thumb. Christiano, Eichenbaum and Evans (2005), Smets and Wouters (2003), Giannoni and Woodford (2003) and Woodford (2003), on the other hand, allow for partial or full indexation of prices to past inflation rates for firms not adjusting their prices optimally in a given period, as an extension to the standard Calvo (1983) pricing model. Those extensions improve the empirical fit of their models.

Together with the sources of persistence in inflation, another key issue lies in understanding whether inflation persistence is an intrinsic characteristic of industrialized economies.¹ Recent studies that aim to shed light on this issue are those by Cogley and Sargent (2005), who found evidence of a structural break in inflation dynamics, by Levin and Piger (2002), who analyzed a panel of industrial countries and concluded that high inertia is not an inherent characteristic of industrial economies, and by Erceg and Levin (2003), who similarly opposed the view of inflation persistence as structural and argued that persistence can depend on the central bank's perceived credibility. Reis and Pivetta (2003), on the other hand, found that inflation persistence has remained high and substantially unchanged since 1965.

This paper proposes a different explanation for the persistence of inflation. The paper suggests

that **adaptive learning** by economic agents may represent a crucial source of persistence, and it aims to empirically evaluate the importance of learning using time series data on inflation. One possible way to assess the empirical effects of learning might consist of simulating the economy with non-fully rational expectations to judge whether those are able to generate enough persistence compared to actual data. This paper instead proposes a different experiment: I develop an optimizing model in which I introduce learning, but I also allow for a structural characteristic that induces persistence in inflation. In fact, I allow for indexation to past inflation by non-optimizing firms. The model, therefore, *nests* two potential sources of persistence in inflation: learning and indexation. Under rational expectations, a large degree of indexation typically is needed to match the data. It becomes an empirical matter to understand whether structural sources of persistence remain essential when learning replaces the assumption of rational expectations. I discover this, by looking at the estimates for the degree of indexation to past inflation necessary to match the data in the model with learning.

Following various studies in the adaptive learning literature (see Evans and Honkapohja 2001 for example), I model private agents as econometricians, who estimate simple models of the economy and form expectations from those models. As private agents obtain more data over time, they update their parameter estimates through Constant-Gain Learning (CGL). In principle, constant-gain learning may generate persistence, not only in transitional dynamics, but perpetually.

I recognize that, when modeling learning behavior, researchers dispose of a number of degrees of freedom. An important choice is the learning speed of private agents, i.e. the value to assign to the constant gain parameter. One's results might dramatically change under different assumptions about learning speed. I, therefore, perform an additional experiment: I examine the relationship between the implied estimates of structural persistence and the possible constant gains. I find, in fact, that the estimates vary a great deal across a large range of gain values. It, therefore, becomes necessary to compare the different possible gains, and I will try to shed some light on this by comparing the fit under the various gain coefficients. In this way, I obtain the value of the best-fitting constant gain coefficient in the sample. This turns out to be close to values typically assumed (without estimation) in previous learning literature.

A related work is Ball (2000), who similarly focuses on expectations and relaxes the assumption of full rationality. Ball allows agents to use optimal univariate forecast rules as an alternative to rational or purely adaptive expectations. My paper, instead, introduces learning by agents. Also

related are the papers by Roberts (1997, 1998), and by Adam and Padula (2003), who estimate inflation equations using subjective expectations from surveys. The current paper provides in fact a way to model the formation of those subjective expectations.² Adam (2005b) similarly assumes agents that use econometric models to forecast future macroeconomic variables. He finds that deviations from rational expectations strenghten the internal propagation mechanism of a simple business cycle model.

Other explanations for inflation persistence are possible. Dotsey and King (2001), Guerrieri (2005), Holden and Driscoll (2003), Coenen and Levin (2004), for example, propose alternative adjustments in their sticky price models to generate additional persistence, but they do not focus on expectations. Mankiw and Reis (2002), Woodford (2003a), and Gumbau-Brisa (2004) try to explain inertia by agents' limited ability to update or absorb information. This paper does not aim to provide a general test among all the competing explanations. The paper proposes, instead, an explanation based on learning, and it goes on to test, using actual data, the plausibility of learning against a more structural source of persistence, i.e. indexation.

The paper also addresses the pervasive debate on the relative importance of backward- and forward-looking terms in inflation. Galí and Gertler (1999), for example, argue that inflation is mainly forward-looking: they find that roughly one third of price setters are backward-looking, from their GMM estimation. Opposite is the view of Fuhrer and Moore (1995) and Fuhrer (1997) who, instead, view inflation as a purely backward-looking process. This paper, by explicitly modeling the formation of expectations, also can help to disentangle the two components.

This paper, therefore, aims to contribute to the large literature on inflation dynamics, by proposing and evaluating a different explanation of its persistence. Moreover, the current paper also can be seen as a contribution to the growing literature on adaptive learning.

Most studies in the previous adaptive learning literature primarily have focused upon studying the convergence of models with learning to the rational expectations equilibrium. This line of research is surveyed comprehensively in Evans and Honkapohja (2001). A similar scope have the applications in monetary policy models, such as Bullard and Mitra (2002), and Evans and Honkapohja (2003).

Recently, this area of research has expanded its objectives: various studies, for example, show that learning is useful in explaining U.S. inflation in the 1970s (Sargent 1999, Orphanides and Williams 2003, Bullard and Eusepi 2005, Primiceri 2003). Williams (2003) examines, instead, the

empirical importance of adaptive learning in a business cycle model. This paper shares the interest in these new objectives and aims to contribute to the understanding of the empirical implications of learning.

As I already have mentioned, although several papers have started to employ constant-gain learning, estimates of the gain lack in the literature. This paper, therefore, contributes to the learning literature first by showing how one's results might dramatically depend on the chosen gain, and then by providing a first attempt to estimate the best-fitting constant gain in a model of inflation dynamics.

By estimating a model of inflation dynamics with deviations from rational expectations, this paper represents a simple example of what Ireland (2003) has defined as "Irrational Expectations Econometrics". Ireland pointed out the results obtained by the theoretical literature on learning and emphasized the need for an "Irrational Expectations Econometrics" that would complement those results, by assessing the empirical importance of learning. This is what the current paper tries to do. This paper focuses on a single equation estimation, while a companion paper, Milani (2004), pursues a joint estimation of a full New-Keynesian macro-model with learning by means of likelihood-based Bayesian methods.

The empirical results highlight the unimportance of forms of structural persistence in inflation. When I drop the assumption of rational expectations, by allowing instead economic agents to form their expectations through constant-gain learning, the estimates of the degree of indexation to lagged inflation fall to zero. This suggests that learning can account for a sizeable amount of persistence in inflation.

The finding that inflation persistence is not due to structural characteristics, but to learning behavior by agents carries some important policy implications. The welfare loss is different under alternative sources of persistence and so is the optimal monetary policy. A successful *management of expectations*, as emphasized by Woodford (2003b), becomes crucial under learning.

The rest of the paper is structured as follows. Section 2 presents the model, starting from the microfoundations of a dynamic optimizing general equilibrium (DGE) model under rational expectations. Section 3 presents the aggregate law of motion for inflation and describes the expectations' formation mechanism. Section 4 derives the main empirical results of the paper. Section 5 and 6 explore the relationship between learning speed and the implied estimated inflation persistence and in-sample fit. Section 7 investigates the robustness of the empirical results to alternative as-

sumptions about the learning rule, while Section 8 discusses the policy implications of the results. Section 9 concludes.

2. THE MODEL

In this section, I derive the law of motion for inflation, which will extend the New Keynesian Phillips curve along two directions. First, to induce a more realistic degree of inertia in inflation, I allow non-optimizing firms to update their prices through indexation to lagged inflation, as proposed by Christiano, Eichenbaum and Evans (2005). Then, the paper makes an important departure from the usual expectations formation mechanism. The paper relaxes the assumption of rational expectations and it assumes instead that agents behave as econometricians, by estimating an economic model and using that model to form their expectations.

I start by presenting the optimal price-setting problem for a firm under rational expectations. I shall introduce subjective (possibly non-rational) expectations and learning in the next section. The current paper focuses only on the price-setting problem by firms. A full model with consumer optimization and monetary policy is described in Milani (2004).

2.1. The Household Problem

I consider a standard economy populated by a continuum of households indexed by i , maximizing a discounted sum of future utilities. The generic household maximizes the following intertemporal utility function:

$$E_t^i \sum_{T=t}^{\infty} \beta^{T-t} \left[U(C_T^i, ; \xi_T) - \int_0^1 v(h_T^i(j); \xi_T) dj \right] \quad (1)$$

where consumer's utility depends positively on an index of consumption C_T^i , and negatively on the amount of labor supplied for the production of good j , $h_T^i(j)$; ξ_T is an aggregate preference shock, whereas $\beta \in (0, 1)$ is the usual household's discount factor. E_t^i here denotes rational (model-consistent) expectations. The consumption index is of the Dixit-Stiglitz CES form

$$C_t^i \equiv \left[\int_0^1 c_t^i(j)^{\frac{\theta-1}{\theta}} dj \right]^{\frac{\theta}{\theta-1}} \quad (2)$$

and the associated aggregate price index is expressed by

$$P_t \equiv \left[\int_0^1 p_t(j)^{1-\theta} dj \right]^{\frac{1}{1-\theta}} \quad (3)$$

where $\theta > 1$ represents the elasticity of substitution between differentiated goods. From the household's problem, I only need to derive the marginal utility of real income (here equal to the marginal utility of consumption, i.e. $\lambda_t = U_C(C_T^i; \xi_T)$), which will enter the aggregate supply relationship. The other details of consumer optimization can therefore remain implicit in the background.³

2.2. Optimal Price Setting

I assume Calvo price-setting, so that a fraction $0 < 1 - \alpha < 1$ of prices are allowed to change in a given period and are optimally set. The price of the remaining fraction α , which is not optimally fixed in the period, is adjusted according to the indexation rule

$$\log p_t(i) = \log p_{t-1}(i) + \gamma \pi_{t-1} \quad (4)$$

Following Christiano, Eichenbaum and Evans (2005), Smets and Wouters (2003) and Giannoni and Woodford (2003), therefore I allow firms to index their prices to past inflation when they cannot set their prices optimally. This extension typically improves the fit of the model and it generates realistic impulse response functions. $0 \leq \gamma \leq 1$ represents the degree of indexation to past inflation (Christiano, Eichenbaum and Evans (2005) assumed $\gamma = 1$, meaning full indexation). An alternative solution to introduce dependence on lagged inflation would consist of assuming a fraction of rule-of-thumb firms. An example is Galí and Gertler (1999), who define the derived equation 'New Hybrid Phillips Curve'.

The demand curve for product i takes the form:

$$y_t(i) = Y_t \left(\frac{p_t(i)}{P_t} \right)^{-\theta} \quad (5)$$

where $Y_t = \left[\int_0^1 y_t(i)^{\frac{\theta-1}{\theta}} di \right]^{\frac{\theta}{\theta-1}}$ is the aggregate output and P_t is the aggregate price as in (3). Each firm i has a production technology $y_t(i) = A_t f(h_t(i))$, where A_t is an exogenous technology shock, $h_t(i)$ is labor input and the function $f(\cdot)$ satisfies the usual Inada conditions.

Since each firm faces the same demand function (5), all firms allowed to change their price in period t will set the same price p_t^* that maximizes the expected present discounted value of future profits:

$$E_t^i \left\{ \sum_{T=t}^{\infty} \alpha^{T-t} Q_{t,T} \left[\Pi_T^i \left(p_t(i) \left(\frac{P_{T-1}}{P_{t-1}} \right)^\gamma \right) \right] \right\} \quad (6)$$

where $Q_{t,T} = \beta^{T-t} \frac{P_t}{P_T} \frac{U_c(Y_T, \xi_T)}{U_c(Y_t, \xi_t)}$ is the stochastic discount factor (U_c is the marginal utility of an additional unity of income), and $\Pi_T^i(\cdot)$ denotes profits. Firms discount future profits at rate α , since they can expect the optimal price chosen at date t to apply in period T with probability α^{T-t} . The firm chooses $\{p_t(i)\}$ to maximize the flux of profits (6), for given $\{Y_T, P_T, w_T(j), A_T, Q_{t,T}\}$ for $T \geq t$ and $j \in [0, 1]$.

The firm's problem results in the first-order condition

$$E_t^i \left\{ \begin{array}{l} \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} U_c(Y_T; \xi_T) Y_T P_T^\theta \left(\frac{P_{T-1}}{P_{t-1}}\right)^{\gamma(1-\theta)} \\ \left[\widehat{p}_t^*(i) - \mu P_T s \left(Y_T \left(\frac{\widehat{p}_t^*(i)}{P_T}\right)^{-\theta} \left(\frac{P_{T-1}}{P_{t-1}}\right)^{-\gamma\theta}, Y_T; \xi_T \right) \right] \end{array} \right\} = 0 \quad (7)$$

where $\mu = \theta/(\theta - 1)$, $s(\cdot)$ is firm i 's real marginal cost function in period $T \geq t$, given the price $\widehat{p}_t^*(i)$, set in t .

The Dixit-Stiglitz aggregate price index evolves according to the law of motion:

$$P_t = \left[\alpha \left(P_{t-1} \left(\frac{P_{t-1}}{P_{t-2}}\right)^\gamma \right)^{1-\theta} + (1-\alpha) p_t^{*1-\theta} \right]^{\frac{1}{1-\theta}} \quad (8)$$

From a log-linear approximation of the firm's first order condition⁴ and some manipulations, I obtain:

$$\widehat{p}_t^* = E_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \left[\frac{1-\alpha\beta}{1+\omega\theta} (\omega + \sigma^{-1}) x_T + \alpha\beta (\widehat{\pi}_{T+1} - \gamma\widehat{\pi}_T) \right] \quad (9)$$

where $\omega > 0$ is the elasticity of firm i 's real marginal cost function $s(\cdot)$ with respect to its output $y_t(i)$, $\sigma \equiv -U_c/(U_{cc}\bar{C})$ is the intertemporal elasticity of substitution, and $\widehat{\cdot}$ denotes log deviations from the steady state⁵.

From a log-linear approximation of the aggregate price index, notice that $\widehat{p}_t^* = \frac{\alpha}{(1-\alpha)} (\pi_t - \gamma\pi_{t-1})$, which plugged in the previous expression gives:

$$\widetilde{\pi}_t = \delta x_t + \widehat{E}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} [\delta\alpha\beta x_{T+1} + (1-\alpha)\beta\widetilde{\pi}_{T+1}] \quad (10)$$

where

$$\widetilde{\pi}_t = \pi_t - \gamma\pi_{t-1} \quad (11)$$

$$\delta \equiv \frac{(1-\alpha)(1-\alpha\beta)(\omega + \sigma^{-1})}{\alpha(1+\omega\theta)} \equiv \frac{(1-\alpha)(1-\alpha\beta)}{\alpha} \zeta > 0 \quad (12)$$

and $\zeta \equiv \frac{(\omega + \sigma^{-1})}{1 + \omega\theta}$. Equation (10) can be quasi-differenced to obtain a relationship between current and one-period-ahead expected inflation

$$\pi_t = \frac{\gamma}{1 + \beta\gamma} \pi_{t-1} + \frac{\beta}{1 + \beta\gamma} \widehat{E}_t \pi_{t+1} + \frac{\delta}{1 + \beta\gamma} x_t + u_t \quad (13)$$

in which I follow the literature in adding an exogenous cost-push shock u_t . Galí and Gertler (1999) and Sbordone (2003) suggest using the real marginal cost in place of the output gap as the driving variable for inflation. As discussed in Woodford (2003b), the relationship between inflation and marginal costs holds under weaker assumptions. For example, when the marginal cost replaces the output gap there is no need to assume any specific theory of wage-setting.

Therefore, the relationship (13) can be re-expressed in terms of the real marginal cost s_t :

$$\pi_t = \frac{\gamma}{1 + \beta\gamma} \pi_{t-1} + \frac{\beta}{1 + \beta\gamma} \widehat{E}_t \pi_{t+1} + \frac{(1 - \beta\alpha)(1 - \alpha)}{\alpha(1 + \beta\gamma)} s_t + u_t \quad (14)$$

Notice that I could have derived similar equations for inflation dynamics, only with different restrictions on the parameters, assuming the existence of some rule-of-thumb firms (Galí and Gertler (1999), Amato and Laubach (2003)), instead of indexation. The results in the following of the paper are not dependent on this choice.

3. THE NEW KEYNESIAN PHILLIPS CURVE WITH LEARNING

The aggregate dynamics for inflation in the model is given by eq.(13). I now relax the strong informational assumptions characterizing firms' knowledge under rational expectations. In this section, I assume that firms have subjective (and possibly non-rational) expectations. Subjective expectations are denoted by \widehat{E}_t .⁶ The law of motion for inflation under subjective expectations becomes

$$\pi_t = \frac{\gamma}{1 + \beta\gamma} \pi_{t-1} + \frac{\beta}{1 + \beta\gamma} \widehat{E}_t \pi_{t+1} + \frac{\delta}{1 + \beta\gamma} x_t + u_t \quad (15)$$

Under this specification, firms need to forecast future inflation rates to determine current inflation. The next paragraph will give some details on how agents form such forecasts.

3.1. Expectations Formation: Adaptive Learning

Firms do not know the correct model of inflation dynamics. They behave as econometricians, estimating an economic model and forming expectations from that model. For simplicity, I start

by assuming that firms estimate a simple linear univariate AR(1) model to form their forecasts of inflation:

$$\pi_t = \phi_{0,t} + \phi_{1,t}\pi_{t-1} + \varepsilon_t \quad (16)$$

In the estimation, they exploit the entire history of available data up to period t , $\{1, \pi_{t-1}\}_0^{t-1}$. Eq. (16) is called the **“Perceived Law Motion”** or **PLM** of the agents. Adam (2005b) assumes similar simple forecasting rules. Notice that, although in the estimation I will use demeaned variables, I recognize that agents need to estimate an intercept as well as slope parameters. A strictly positive intercept on inflation would signal that agents expect a positive target for inflation in the sample. In the next sections, I shall evaluate the robustness of the empirical results to different PLMs. Ultimately, it will be possible to choose the forecasting rule that provides the best fit of the data.

As new data become available, agents update their estimates according to the Constant-Gain learning (CLG) formula

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \kappa R_{t-1}^{-1} X_t (\pi_t - X_t' \widehat{\phi}_{t-1}) \quad (17)$$

$$R_t = R_{t-1} + \kappa (X_{t-1} X_{t-1}' - R_{t-1}) \quad (18)$$

where the first expression describes the updating of the forecasting rule coefficients $\widehat{\phi}_t = (\phi_{0,t}, \phi_{1,t})'$ over time, and the second shows the evolution of R_t , the matrix of second moments of the stacked regressors $X_t \equiv \{1, \pi_{t-1}\}_0^{t-1}$. The constant gain is expressed by parameter κ . Constant-gain learning has also been defined as ‘perpetual’ learning, since learning will take place forever and the system will not converge to the RE solution (but at most to a stochastic distribution around it). Therefore, constant gain learning may generate persistence forever, not only transitionally. A larger κ would imply faster learning of potential structural breaks, but it would also lead to higher volatility around the steady state. An increasing number of recent papers has similarly used constant-gain learning (Orphanides and Williams 2003, 2005a, Primiceri 2003).

I assume that agents, when forming expectations in period t , have access to information only up to $t - 1$: I therefore replace \widehat{E}_t with \widehat{E}_{t-1} . Using their PLM and the updated parameter estimates $\widehat{\phi}_t$, agents form expectations for $t + 1$ as

$$\widehat{E}_{t-1}\pi_{t+1} = \phi_{0,t-1} (1 + \phi_{1,t-1}) + \phi_{1,t-1}^2 \pi_{t-1} \quad (19)$$

To summarize, the model of the economy is composed by the inflation dynamics equation (15) (Phillips curve), agents' beliefs (16), updating equations (17), (18), and the forecasting rule (19).

Substituting agents' expectations formed from their PLM as in (19) into the inflation dynamics equation (15), I derive the **“Actual Law of Motion”** or **ALM** of the economy, i.e. the law of motion of π_t for a *given* PLM:

$$\pi_t = \frac{\beta\phi_{0,t}(1 + \phi_{1,t})}{1 + \beta\gamma} + \frac{\gamma + \beta\phi_{1,t}^2}{1 + \beta\gamma}\pi_{t-1} + \frac{\delta}{1 + \beta\gamma}x_t + u_t \quad (20)$$

in which the reduced-form coefficients are time-varying and are convolutions of the structural parameters describing inflation dynamics and of the coefficients representing agents' beliefs. In models with adaptive learning, it is commonly assumed that, in each period t , agents use an econometric model to form their expectations about future inflation, but they do not take into account their subsequent updating in periods $T > t$. Therefore, they act as adaptive decision-makers, in accordance with what Kreps (1998) defines as an *anticipated utility model*.⁷

Obviously, several alternative ways to model learning are possible. The paper considers a simple learning rule (assuming that agents use univariate autoregressive models). But it is worth pointing out that such a simple mechanism of expectations formation fits quite well with inflation expectations from surveys, such as the expected inflation series from the Survey of Professional Forecasters, for example (see Branch and Evans 2005).

3.1.1. Is the Expectations Formation Realistic?

As described in Brayton et al. (1997), the most recent Federal Reserve model, the FRB/WORLD, also employs non-fully rational expectations.

Recognizing the uncertainty surrounding expectations formation, the main Fed model not only uses model consistent rational expectations, but it also models expectations as derived from a VAR for five ‘core’ macro variables (federal funds rate, CPI inflation, output gap, long-run inflation expectations and long-run interest rate expectations). The underlying justification is that the agents understand the main features of the economy, as represented by small economic models, and use the information to form their expectations. The Fed model, however, does not currently incorporate time-variation in the parameters.

The realism of forming expectations by adaptive learning can be gauged by looking at survey-based expectations. The expectations derived according to the adaptive learning algorithm do a

good job in tracking inflation expectations from the Survey of Professional Forecasters. In particular, they replicate the underestimation of inflation in the two peaks in the 1970s and the overestimation of inflation during most of the 1990s. Less successful in tracking survey-based forecasts are purely adaptive (naïve) expectations and RE (since the data display large and persistent forecast errors, which are less consistent with RE). A substantial literature, in fact, emphasizes how survey expectations reflect an intermediate degree of rationality, rejecting full rationality as well completely naïveté (Roberts 1998 is an example).

Adam (2005a) provides some experimental evidence on the formation of inflation expectations, showing that forecast rules in which agents condition on lagged inflation successfully mimic the inflation expectations of the subjects in his experiment. The AR(1) forecasting rule assumed in this paper is consistent with such experimental evidence.⁸

Moreover, constant gain learning seems a plausible choice to model the behavior of professional forecasters. Branch and Evans (2005), in fact, show that constant gain models of learning fit forecasts from surveys better than other methods for both inflation and output growth. They find that constant-gain learning models dominate models with optimal constant gain (obtained by minimizing the forecasts' Mean Square Error), with Kalman Filter, and with Recursive Least Squares learning. Their results therefore support constant-gain learning as a model of actual expectations formation.

3.2. Real Marginal Cost as the Driving Variable

Following recent research (Galí and Gertler 1999, Sbordone 2003), I also work with a version of the Phillips curve in which the real marginal cost is the relevant driving variable for inflation

$$\pi_t = \frac{\gamma}{1 + \beta\gamma} \pi_{t-1} + \frac{\beta}{1 + \beta\gamma} \widehat{E}_t \pi_{t+1} + \frac{(1 - \beta\alpha)(1 - \alpha)}{\alpha(1 + \beta\gamma)} s_t + u_t \quad (21)$$

Again I assume that firms form their expectations by estimating an AR(1) specification for inflation as in (19).

The ALM for inflation becomes

$$\pi_t = \frac{\beta\phi_{0,t}(1 + \phi_{1,t})}{1 + \beta\gamma} + \frac{\gamma + \beta\phi_{1,t}^2}{1 + \beta\gamma} \pi_{t-1} + \frac{(1 - \beta\alpha)(1 - \alpha)}{\alpha(1 + \beta\gamma)} s_t + u_t \quad (22)$$

4. SOME SIMPLE IRRATIONAL EXPECTATIONS ECONOMETRICS: EMPIRICAL RESULTS

I estimate the inflation equation assuming that firms form expectations and update their beliefs through constant-gain learning as described. I use quarterly U.S. data on inflation, output, and real marginal costs from 1960:01 to 2003:04. Inflation is defined as the annualized quarterly rate of change of the GDP Implicit Price Deflator, the output gap as detrended GDP after removing a quadratic trend, the real marginal cost as the unit labor cost, which is empirically proxied by the log labor income share in deviation from the steady state.⁹ I allow agents to initialize their estimates of coefficients, variances, and covariances, using pre-sample data from 1951 to 1959.¹⁰

The empirical exercise proceeds as follows. First, I estimate the agents' PLM in which agents learn the forecasting coefficients over time through constant-gain learning, as described by expressions (17) and (18). I can then substitute the resulting forecasts into the original inflation equations, (15) and (21). I initially fix the constant gain κ at the value of 0.015. This is consistent with values derived by minimizing the deviation of the constructed series from the expected inflation series from the Survey of Professional Forecasters, as found by Orphanides and Williams (2005a).

I can then simply estimate the ALM for inflation by NLLS (Nonlinear Least Squares). In this way, I am able to *disentangle* the effects of learning from the effects due to structural sources of persistence in inflation, such as indexation. Notice that having explicitly modeled the formation of agents expectations, it is not necessary to estimate the inflation equation by GMM as under RE: I can therefore avoid the typical problems of GMM in the estimation of Phillips curves (Lindé 2002, Jondeau and Le Bihan 2003, Fuhrer and Olivei 2004).

This paper therefore focuses on a single equation estimation of inflation dynamics. This approach avoids infecting the inflation equation by potential misspecifications in other parts of the model. In a companion paper (Milani 2004) I jointly estimate, instead, a full New-Keynesian model with learning by likelihood-based Bayesian methods. The two papers represent different thought experiments: this paper is implicitly assuming that agents' learning is correctly modeled and it focuses on estimating the structural parameters *given* the assumed learning specification. Milani (2004) pushes the experiment a step further and aims to *jointly* extrapolate from the data the learning rule coefficients together with the structural coefficients. In that way, agents' beliefs and their learning speed are jointly estimated together with the rest of the model. Structural estimates

will be affected by the uncertainty concerning agents' learning specification. This allows for a better account of total uncertainty in the system. The drawback is that misspecifications in the learning equation will bias the rest of the model coefficients. The current paper instead separately estimates the learning equation; when the results are inserted in the ALM, they are treated as certain. The standard errors of the structural coefficients then underestimate the true underlying uncertainty.

These papers provide a first example of the “Irrational Expectations Econometrics” that Ireland (2003) has judged needed to complement the mainly theoretical results of the previous adaptive learning literature.

4.1. *What Creates Persistence in Inflation?*

The inflation equation nests indexation by firms and non-rational expectations. Whether the persistence in inflation is structural or due to learning behavior becomes then an empirical question. Understanding the role of different sources of persistence is crucial from both a positive and a normative point of view: the recommendations for optimal monetary policy are likely to differ in the cases that indexation or learning drive the persistence in inflation.

4.1.1. *Agents' Beliefs*

Figure 1 shows the evolution of the estimated coefficients in the agents' forecasting equation (agents' PLM). Economic agents are updating the coefficients through constant gain learning. The reported evolution of beliefs is obtained by assuming a constant gain equal to 0.015.

In the beginning of the sample, agents were coming from periods of low and volatile inflation (the 1950s and 1960s), and, consequently, they estimated low autoregressive coefficients for inflation (around 0.15 at the beginning of the sample). In the 1970s, inflation rose substantially and also became more persistent. The agents recognized the shift and in the 1970s they started estimating larger autoregressive coefficients (with a pick around 1975, when the estimated $\phi_{1,t}$ jumped to 0.958). The estimates of perceived inflation persistence declined in the last part of the sample, though remaining above 0.8.

The evolution of the intercept in the inflation equation (recalling that the true value should always be 0, since I work with demeaned variables) indicates that the perceived inflation target was low in the 1960s, it increased and remained high through the second half of the 1970s, and it continuously decreased after Volcker's disinflation.

4.1.2. Structural Estimates

Table 1 shows the estimation results. The table reports the estimates for the alternative specifications with the output gap and the real marginal cost as main driving variables for inflation, and for different values of the constant-gain coefficient ($\kappa = 0.015, 0.02, \text{ and } 0.025$).

I obtain coefficients on indexation to lagged inflation, γ , equal to 0.139 and 0.047 in the output gap and real marginal cost equations, for the case in which $\kappa = 0.015$. The estimates are small and not significantly different from 0. I estimate δ , the sensitivity of inflation to changes in the output gap, equal to 0.22. In the equation with real marginal costs as the driving variable, I also estimate α , the Calvo parameter. I find $\alpha = 0.671$, which suggests prices that remain fixed for 3.04 quarters. The estimate is in line with most estimates in the literature (on the lower end). This suggests that although learning reduces the relevance of indexation, it does not significantly alter the estimated magnitude of price stickiness. With other gain coefficients, the results still indicate that inflation indexation is not supported by the data (I obtain γ equal to -0.001 and 0.045 when $\kappa = 0.02$, and γ equal to -0.10 and -0.006 , both not significantly different from 0, when $\kappa = 0.025$). The results about indexation contrast with the estimates typically computed in the literature.

Likewise, I can estimate a reduced form equation for inflation, given by

$$\pi_t = \omega_b \pi_{t-1} + \omega_f \widehat{E}_t \pi_{t+1} + \delta y_t + \varepsilon_t \quad (23)$$

where $y_t = x_t, s_t$. The specification is similar to reduced forms usually estimated in the empirical literature that aims to evaluate the relative importance of forward-looking and backward-looking behavior in inflation (the New Hybrid Phillips curve for example). The only difference comes from the use of subjective expectations instead of RE. The estimation yields

$$\pi_t = \underset{(0.15)}{0.105} \pi_{t-1} + \underset{(0.19)}{0.992} \widehat{E}_t \pi_{t+1} + \underset{(0.037)}{0.201} x_t + \widehat{\varepsilon}_t \quad (24)$$

when the output gap is used and

$$\pi_t = \underset{(0.166)}{0.045} \pi_{t-1} + \underset{(0.197)}{0.972} \widehat{E}_t \pi_{t+1} + \underset{(0.029)}{0.147} s_t + \widehat{\varepsilon}_t \quad (25)$$

when the real marginal cost is used. The results therefore suggest that when learning replaces rational expectations, inflation becomes strongly forward-looking.

4.1.3. *Comparison with the Literature*

Christiano, Eichenbaum and Evans (2005) do not actually estimate γ , but they fix it to 1, indicating full indexation. Boivin and Giannoni (2003) and Giannoni and Woodford (2003) estimate the coefficient and find that γ equals 1. These results point towards extremely large degrees of structural persistence in inflation.

But, as we have seen, estimates of γ close to 1 hinge on the assumption of rational expectations. When the assumption is weakened, by introducing a minimal deviation from full rationality and allowing agents to learn over time, the degree of inflation persistence due to structural features of the economy drops to almost zero.

There has been a considerable debate in the literature on whether inflation is mainly a backward or forward-looking phenomenon. Galí and Gertler (1999) stress the importance of forward-looking expectations in their New Hybrid Phillips Curve, although still obtaining a positive weight on lagged inflation. Fuhrer and Moore (1995) and Fuhrer (1997), on the other hand, depicts inflation as substantially backward-looking.

The results of this paper give merit to both ideas. In fact, I obtain that inflation is mostly forward-looking (and indeed very forward-looking as seen in eq.(24) and (25)). But if expectations are formed as in this paper, the reduced form will be equivalent to a completely backward-looking specification. The effort to explicitly model subjective expectations gives a way to disentangle the inertia due to structural characteristics from those due to the sluggishness of forward-looking expectations (this is hard to identify under RE, as pointed out by Beyer and Farmer 2004). It is easy to understand, however, why many contrasting results in the literature have emerged in relation to the relative importance of backward and forward-looking terms.

A recent study by Cogley and Sbordone (2005) similarly obtain an estimated degree of indexation close to 0 by allowing for a time-varying steady-state inflation. It is beyond the paper's scope to check the role of learning in a model in which also steady-state inflation varies over time. But, taken together, the present paper and Cogley and Sbordone (2005) accumulate evidence that when more realistic features are added into the model, the role of indexation vanishes.

5. LEARNING SPEED AND INFLATION PERSISTENCE

In the estimation, I have experimented different gain coefficients between 0.015 and 0.025. These values are common in empirical studies adopting constant-gain learning, as Orphanides and Williams (2003) for example. The degree of persistence introduced in the system by learning, as well as the estimates of the indexation parameter, are likely to be strongly dependent on the choice of the gain parameter.

For this reason, it becomes essential to investigate how the estimates of structural persistence vary across a wide range of possible gain values. This experiment allows me to examine the relationship between learning speed and inflation persistence.

Figure 2 shows such a relationship. The figure illustrates how the reduced form coefficient on lagged inflation ($\frac{\gamma}{(1+\gamma\beta)}$) varies with different gain values ranging from 0 to 0.30 (the results for values above 0.30 are totally similar to those on that upper bound).

There seems to be a sort of V-shaped relationship. With a zero gain, or with very small gains, the weight given to lagged inflation in an equation like (24) is sizeable. With slightly larger gains, the implied coefficients on the backward-looking term becomes much smaller and implies coefficients below 0.2 for gains around 0.025. Inside this range, learning is successful in creating enough persistence in inflation, so that no role remains for additional sources of structural persistence. Outside that range, for lower or larger gain values, backward-looking components and indexation remain important, for learning with those gains does not generate expectations of future inflation rates that seem supported by the data.

6. LEARNING SPEED AND FIT

Having such a diverse range of results, it is important to evaluate which value of the constant gain is more supported by the data. I calculate the best-fitting constant gain by estimating the following simple New Keynesian Phillips curve with learning (and no indexation)

$$\pi_t = \delta y_t + \beta \widehat{E}_t \pi_{t+1} + u_t, \quad \text{where } y_t = x_t, s_t \quad (26)$$

over the range of all possible gain coefficients from 0 to 0.30 and evaluating how the in-sample fit changes. I assume again that agents form expectations by estimating autoregressive specifications for inflation and updating the coefficients over time through constant-gain learning. As a measure

of fit, I report the Schwartz's Bayesian information Criterion (BIC).

Figure 3 indicates how the fit (BIC) varies across the whole range of assumed gain coefficients. Again, it is possible to observe a sort of V-shaped relationship. The best-fitting specifications (lowest BIC) have gain coefficients that lie in the range typically assumed in previous learning studies (between 0.015 and 0.03). Very small gains (close to 0) and large gains (above 0.06) perform very poorly in terms of in-sample fit.

Figure 4 reports the same information, but focusing on a narrower range of gain values. It is apparent that already gains above 0.03 lead to a sizeable reduction in fit.

Table 2 reports the gain coefficients in correspondence of which the lowest BIC is derived. The best-fitting constant gain coefficient equals 0.02 for the inflation equation with the output gap as driving variable. Such a value is similar to what found by Orphanides and Williams (2005a) who minimize the deviations of their model-based expectations from data on survey-based expectations. The results are similar when real marginal costs replace the output gap as the driving variable for inflation. The best-fitting constant gain coefficient, reported in table 2, now equals 0.025. Figure 6 also illustrates a similar relationship between constant gain and fit. In ongoing work, I am finding similar values for the best-fitting constant gain using European data (Milani 2005c). The evidence that we are starting to accumulate therefore suggests that constant-gain learning with a gain around 0.02 appears to be a satisfactory description of the formation of inflation expectations.

But has the learning speed been constant over time? I can recursively re-estimate the equation over the sample to compute the evolution of the best-fitting gain. From figure 7 it is apparent that the gains preferred by the data have remained inside the narrow range 0.016-0.0205. Notice, however, that the gain has varied in the sample. In particular, it is interesting to notice how agents' learning speed has gradually increased during all the 1970s and in the early 1980s, to stabilize later on. Investigating the causes and consequences of the changes in learning speed, as well as the interaction between policy and learning speed, will be useful extensions.

7. ROBUSTNESS

I have so far assumed that economic agents use a simple autoregressive model as in (16) as their PLM. But the correct model of expectations formation is uncertain. Therefore, in this section I aim to evaluate the robustness of my results to alternative forecasting rules.

7.1. *Phillips Curve as Learning Rule*

Suppose now that private agents use the reduced form of a Phillips curve to form their inflation expectations (assuming they observe inflation, output, and real marginal costs in period $t - 1$). Again, they do not know the relevant parameters, so they gradually learn them from the data they observe. The agents' PLM is now

$$\pi_t = \phi_{0,t} + \phi_{1,t}\pi_{t-1} + \phi_{2,t}x_{t-1} + \varepsilon_t \quad (27)$$

with the output gap, and

$$\pi_t = \phi_{0,t} + \phi_{1,t}\pi_{t-1} + \phi_{3,t}s_{t-1} + \varepsilon_t \quad (28)$$

with real marginal costs entering the Phillips curve. Now agents are using the same variables that enter the actual law of motion of inflation in the model.¹¹

Under a constant gain equal to 0.02 (the best-fitting gain in the previous case), for example, agents estimate the Phillips curve coefficients reported in figure 8. The intercept again starts low and increases over the sample, decreasing only in the second half of the 1990s. The agents' belief about the persistence of inflation is low in the 1960s and early 1970s, when it jumps from below 0.4 to almost 1 (0.96) around 1975. Another period of extremely high persistence occurs in 1981-1982. Therefore private agents perceive an extremely large degree of persistence both during the run-up of inflation in the 1970s and during Volcker's disinflation in the early 1980s. The third coefficient in the graph represents the sensitivity of inflation to the output gap (slope of the Phillips curve). This sensitivity has been increasing from the late 1960s until 1975, it remained stable (around 0.2) in the second part of the 1970s, and then it declined after Volcker's disinflation. In the 1980s and 1990s, the estimated sensitivity has always been low and it has fallen even more after 1998 (stabilizing around 0.06). The dynamics is consistent with the existence of an important trade-off between inflation and output in the 1960s-1970s, which almost disappeared in the 1980s-1990s.

If I estimate the degree of indexation under this alternative learning rule, I obtain similar results as before. Figure 9 shows the estimated backward-lookingness in inflation across gain coefficients. The estimates are close to 0 for a range of gain values above 0.05. The figure also superimposes the fit under the different gains. The best-fitting gains are now larger than they were under the simpler AR(1) forecasting rule. The best-fitting gain, reported also in table 3, equals 0.068 (and in

correspondence of this gain the coefficient on the backward-looking term in the inflation equation equals 0.043).

Figure 10 instead repeats the exercise for the learning rule in which marginal costs replace the output gap. The message is similar. The best-fitting gain is now 0.0355 (larger than before) and the estimated persistence coming from structural features is small (but larger than before, 0.168). It is worth noticing from figures 9 and 10 that the models with output gap seem to provide a better fit than do those with real marginal costs (BIC is around 2.6 instead of 3).

7.2. Interest Rates in the Learning Rule

I now suppose that agents use additional information in their forecasting rule. The agents now also include nominal interest rates (the federal funds rate) in their forecasting model. The PLM becomes

$$\pi_t = \phi_{0,t} + \phi_{1,t}\pi_{t-1} + \phi_{2,t}x_{t-1} + \phi_{3,t}i_{t-1} + \varepsilon_t \quad (29)$$

This is similar to inflation equations that are commonly estimated as part of monetary VARs.

The results reported in figure 11 show that the best-fitting gains (also corresponding to the lowest levels of indexation) lie around 0.1. The best-fitting gain is estimated equal to 0.0995 (at this gain the backward-looking term in inflation equals 0.23). Although not a general rule, it appears from the cases examined so far that the larger information set, the higher the estimated speed. The constant gain increases in fact from 0.02 when only lagged inflation enters the learning rule, to 0.068 when inflation and output gap enter, to 0.0995 when inflation, output gap, and the federal funds rate enter.

The results are indicative that estimates of the gain are likely to depend on the assumed learning rule. But also the choice of the regressors entering the agents' PLM can be based on the fit of the various choices. And it is worth noticing that the learning specifications all fit better than an entirely backward-looking inflation equation, where current inflation is regressed on lagged inflation and current output gap (BIC = 3.18). In all cases, learning seems to provide a way to account for the persistence of observed inflation. Future research should shed more light on how to best model the learning process.

8. POLICY IMPLICATIONS

The scope of the paper has been so far descriptive. But shedding light on the main sources of inflation persistence is also crucial from a normative standpoint. Whether the inertia in inflation is structural or is instead due to the way agents form their expectations affects in fact the optimal monetary policy. If the mechanisms that are introduced in the models to induce persistence, such as indexation, turn out to be the wrong representation of the economy, then a welfare analysis based on such microfoundations will be erroneous as well.

8.1. Optimal Monetary Policy with Structural Inflation Persistence

Suppose first that inflation depends on its lagged values because of automatic indexation by firms. Suppose expectations are fully rational. In this case, the rigidity of prices is due to structural characteristics of the economy.

The optimal monetary policy in the case of indexation, considering a microfounded DSGE model as in Woodford (2003b), could be implemented by a central bank minimizing a welfare-based loss function that takes the following form

$$E_t \sum_{T=t}^{\infty} \beta^{T-t} L_t \quad (30)$$

$$L_t = (\pi_t - \gamma\pi_{t-1})^2 + \lambda_x (x_t - x^*)^2 \quad (31)$$

where $\lambda_x = \frac{\delta}{\theta}$ under this paper's microfoundations, and $x^* > 0$ is the optimal output gap level, which depends on microeconomic distortions such as the degree of market power and the size of tax distortions in the economy.

This loss function is derived from the microfoundations of the model. It is therefore optimal to minimize not just the deviation of inflation from target (assumed equal 0 here), but also the rate of change. The more persistent inflation, the more aggressive the optimal reaction of monetary policy. Optimal policy under commitment from the timeless perspective will satisfy the target criterion

$$\pi_t - \gamma\pi_{t-1} = -\frac{\lambda_x}{\kappa} (x_t - x_{t-1}) \quad (32)$$

The optimal rule will be given by

$$i_t^* = \psi_\pi \pi_{t-1} + \psi_x x_{t-1} + \psi_u u_t + r_t^n \quad (33)$$

where ψ_π, ψ_x, ψ_u are optimal feedback coefficients, i_t^* is the policy instrument, and r_t^n is the natural real rate of interest. Monetary policy therefore responds to past observable variables and current shocks.

8.2. *Optimal Monetary Policy with Adaptive Learning*

On the other hand, suppose that the persistence in inflation is not due to structural features (assuming $\gamma \approx 0$), but it is instead due to learning by economic agents. If the central bank recognizes this, it would not be optimal to react so aggressively to inflation as if γ was close to 1. But it would become optimal, instead, to react to private sector expectations of inflation. This would avoid fluctuations induced by mistaken expectations, not in line with the inflation target. Let's consider the following New-Keynesian model with learning, also described in Evans and Honkapohja (2003), Preston (2003), and Milani (2004)

$$\pi_t = \delta x_t + \beta \widehat{E}_t \pi_{t+1} + u_t \tag{34}$$

$$x_t = \widehat{E}_t x_{t+1} - \sigma \left(i_t - \widehat{E}_t \pi_{t+1} - r_t^n \right) \tag{35}$$

where I have added an aggregate demand equation, represented by a log-linearized Euler equation. With zero indexation, the central bank will now target the more familiar welfare-based loss function

$$L_t = \pi_t^2 + \lambda_x (x_t - x^*)^2 \tag{36}$$

which leads to an optimal target criterion similar to (32), but with $\gamma = 0$.

Evans and Honkapohja (2003) show that the optimal policy rule under commitment in this case would be

$$i_t^* = \frac{1}{\sigma} \left[\widehat{E}_t x_{t+1} - \frac{\lambda_x}{\lambda_x + \delta^2} x_{t-1} + \left(\frac{\beta \delta}{\lambda_x + \delta^2} + \sigma \right) \widehat{E}_t \pi_{t+1} + \frac{\delta}{\lambda_x + \delta^2} u_t + \sigma r_t^n \right] \tag{37}$$

The implied reaction function makes clear the need for the central bank to respond now to expectations of the relevant variables. Expectations are taken as given by the policymaker, who does not incorporate the agents' learning rule in its optimization problem. Hence, I abstract here from issues of active experimentation. In this way, the optimal target criterion is satisfied regardless of the particular expectations held by private agents.

In such a framework, a fundamental task of the central bank and optimal monetary policy becomes *the management of expectations*, as emphasized by Woodford (2003b).

In order to keep inflation expectations close to target, the importance of *transparency* and *credibility* should be emphasized. In particular, if the monetary authority lacks credibility, every attempt to reduce inflation, not believed by the public (not incorporated in $\widehat{E}_t\pi_{t+1}$), may be useless and inflation may continue to rise. A more transparent central bank is likely to facilitate private sector learning. Orphanides and Williams (2005b) provide evidence that an explicit target for inflation improves agents' learning. In this new framework, bad policy therefore can mean unsuccessful management of expectations and it can arise also in the case of a central bank following a truly optimal rule derived from dynamic optimization. Undesirable policy outcomes might be due to an important misspecification of the policymaker's model: the failure to understand and incorporate the way agents form their expectations.

Since an important concern in monetary policy points towards robustness of a chosen policy rule, it would also be necessary to examine the effects of optimal rules under learning, if the policymaker does not recognize the true agents' PLM or assumes that they form expectations rationally. Preston (2003) shows that price-level targeting is more robust than inflation targeting if the policymaker wrongly assumes that agents have RE rather than recognizing their learning rule. Similarly, it would be important to evaluate the welfare losses from optimal policy if the policymaker believes the persistence is structural while it is due to learning and vice versa. Coenen (2003) shows the potential large losses from underestimating the degree of inflation persistence and suggests aggressive policy as a safe choice. Certainly, if we are willing to believe that non-fully-rational expectations and learning behavior are important determinants of inflation dynamics, new avenues of research about optimal monetary policy open. The robustness of optimal rules not only to standard model uncertainty, but also to uncertainty in the correct specification of the agents' learning rule (dropping RE opens a wide range of different alternatives), would be an important matter. Indeed, in real world policy-making, central banks would hardly argue to target loss functions as (31), which include the rate of change besides the level of inflation; policymakers put instead a lot of effort on monitoring the evolution of private sector expectations. This is more consistent with models where learning is important than with models with structural persistence.¹²

8.3. *The Welfare Cost of Misspecifying Private Sector's Expectations*

What is the cost of a policymaker misperceiving the expectations formation of the agents? Does the optimal monetary policy obtained under the assumption of rational expectations perform

reasonably well when agents have instead imperfect knowledge? Here I briefly try to provide some preliminary evidence.

Assume that the true structure of the economy is represented by the model with learning, as described by equations (34) and (35). Assume the true degree of indexation is zero ($\gamma = 0$). The policymaker therefore minimizes the quadratic loss function (36). First, I compute the optimal policy by taking agents' learning into account: I therefore assume that the policymaker correctly understands agents' expectations formation. I fix $\beta = 1$, $\delta = 0.22$, $\sigma = 1$, $\theta = 7.69$ (as in Giannoni and Woodford 2003), so that $\lambda_x = \frac{\delta}{\theta} = \frac{0.22}{7.69} = 0.0286$, and I choose $\kappa = 0.02$. The optimal policy under learning implies a welfare loss equal to 2.28, as shown in Table 4.¹³

Suppose now that the policymaker optimizes under the mistaken assumption that agents have fully rational expectations (I still assume no indexation). If the policymaker implements the optimal policy assuming rational expectations, but the true representation of the economy is instead one with non-fully rational expectations and learning, the implied loss will be equal to 5.23. The welfare cost that arises from a misspecification of the agents' expectations formation mechanism is therefore substantial. I find in fact that under rational expectations, the policymaker does not respond as aggressively to inflation as she does under learning. This conclusion is shared by Orphanides and Williams (2003, 2005a). They similarly find that learning implies a more aggressive central bank's response to movements in inflation: they discuss, in fact, how an aggressive policy helps stabilizing inflation expectations near those implied by the rational expectations equilibrium.

9. CONCLUSIONS

Several papers have studied the determinants of inflation dynamics. Monetary policy models often include a law of motion for inflation in which current inflation depends on future expected inflation and on current output gap or marginal costs. Such a relationship, known to researchers as New Keynesian Phillips curve, can be derived from the optimizing behavior of firms in models with imperfect competition and sticky prices. Firms, in these models, have rational expectations. In their simplest form, New Keynesian Phillips curves fail to match the persistence characterizing actual data on inflation. Therefore, researchers have proposed various extensions that lead to persistence in the inflation equation. Rule-of-thumb behavior, price indexation, menu costs are popular modeling devices to account for inflation inertia.

This paper has suggested a different approach, proposing learning as an important determinant of inflation behavior. In the model, economic agents form expectations from simple economic models, not knowing the true model parameters. They have the same knowledge that econometricians would have, and, therefore, they use historical data to infer the relevant parameters and update their estimates through constant-gain learning.

The paper shows that when learning replaces the standard assumption of rational expectations, structural sources of persistence, such as indexation, are no longer essential to fit the data. Therefore, learning seems to be a major source of persistence in inflation. Disentangling the role of learning from that of structural sources of persistence carries also normative implications. If learning rather than mechanical indexation is the main source of persistence in inflation, the implied optimal monetary policy will differ.

Under constant-gain learning, one's results may heavily depend on the choice of the gain. In the paper, I have shown how the estimated backward-lookingness in inflation varies over the range of possible gain coefficients. The large differences are evidence that working with estimated, rather than arbitrarily chosen, constant gain coefficients is necessary. The paper has provided some preliminary evidence, calculating the best-fitting constant gain and showing how the fit of the inflation equation changes across the range of assumed gains. The best-fitting gain for the full sample considered in the paper seems to be around 0.02, and it is reassuringly not far from values assumed by other studies in the learning literature.

References

- ADAM, K., (2005a). "Experimental Evidence on the Persistence of Output and Inflation", ECB Working Paper No. 492.
- (2005b). "Learning to Forecast and Cyclical Behavior of Output and Inflation", *Macroeconomic Dynamics*, Vol. 9(1), 1-27.
- AND M. PADULA, (2003). "Inflation Dynamics and Subjective Expectations in the United States", mimeo.
- AMATO, J.D., AND T.LAUBACH (2003). "Rule-of-Thumb Behaviour and Monetary Policy", *European Economic Review*, Volume 47, Issue 5, pp. 791-831.
- ASCARI, G., (2004). "Staggered Prices and Trend Inflation: Some Nuisances", *Review of Economic Dynamics* 7, 642-667.
- BALL, L., (2000). "Near-Rationality and Inflation in Two Monetary Regimes", NBER WP No. 7988, October.
- BEYER, A., AND R.E.A. FARMER, (2004). "On the Indeterminacy of New-Keynesian Economics", mimeo, ECB and UCLA.
- BOIVIN, J., AND M. GIANNONI, (2003). "Has Monetary Policy Become More Effective?", NBER Working Paper no. 9459, January.
- BRANCH, W.A., AND G.W. EVANS, (2005). "A Simple Recursive Forecasting Model", forthcoming, *Economics Letters*.
- FLINT BRAYTON, ANDREW LEVIN, RALPH TRYON, AND JOHN C. WILLIAMS, (1997). "The Evolution of Macro Models at the Federal Reserve Board", Federal Reserve Board of Governors.
- BULLARD, J., AND S. EUSEPI, (2005). "Did the Great Inflation Occur Despite Policymaker Commitment to a Taylor Rule?", *Review of Economic Dynamics*, vol. 8(2), pages 324-359.
- AND K. MITRA, (2002). "Learning About Monetary Policy Rules", *Journal of Monetary Economics*, 49, pp. 1105-1129.
- CALVO, G., (1983). "Staggered Prices in a Utility-Maximizing Framework", *Journal of Monetary Economics*, 12(4), pp. 983-998.
- 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*, Issue 113(1).
- CLARIDA, R., GALI, J., AND M. GERTLER, (2000). "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory", *Quarterly Journal of Economics*, vol. CXV, issue 1, pp. 147-180.
- COENEN, G., (2003). "Inflation Persistence and Robust Monetary Policy Design", ECB Working Paper No. 290, November.
- AND A. LEVIN, (2004). "Identifying the Influences of Real and Nominal Rigidities in Aggregate Price-Setting Behavior", ECB Working Paper No. 418, November.
- COGLEY, T., AND T.J.SARGENT, (2005). "Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US", *Review of Economic Dynamics* 8, April.
- COGLEY, T., AND A.SBORDONE, (2005). "A Search for a Structural Phillips Curve", Staff Reports No. 203, Federal Reserve Bank of New York.

- DOTSEY, M., AND R. KING, (2001). "Pricing, Production and Persistence", NBER Working Paper No. 8407, August.
- ERCEG, C., AND A.T.LEVIN, (2003). "Imperfect Credibility and Inflation Persistence", *Journal of Monetary Economics*, Volume 50, Issue 4, May 2003, Pages 915-944.
- EVANS, G.W., AND S. HONKAPOHJA, (2001). *Learning and Expectations in Macroeconomics*, Princeton University Press.
- (2003). "Adaptive Learning and Monetary Policy Design", *Journal of Money, Credit and Banking*, vol.35, pp. 1045-1072.
- FAUST, J., and L.E.O. SVENSSON, (2001). "Transparency and Credibility: Monetary Policy with Unobservable Goals", *International Economic Review* 42, pp. 369-397.
- FUHRER, J., (1997). "The (Un)Importance of Forward-Looking Behavior in Price Specifications", *Journal of Money, Credit and Banking*, vol. 29, no. 3, 338-350.
- (2000). Comment on Ball's "Near-Rationality and Inflation in Two Monetary Regimes".
- and G. MOORE, (1995). "Inflation Persistence", *The Quarterly Journal of Economics*, vol. 110(1), pages 127-59.
- and G.P. OLIVEI, (2004). "Estimating Forward-Looking Euler Equations with GMM and Maximum Likelihood Estimators: an Optimal Instruments Approach", mimeo, Federal Reserve Bank of Boston.
- GALI, J., AND M. GERTLER, (1999). "Inflation Dynamics: a Structural Econometric Analysis", *Journal of Monetary Economics*, Volume 44, Issue 2, pp. 195-222.
- GASPAR, V., SMETS, F., and D. VESTIN, (2006). "Adaptive Learning, Persistence and Optimal Monetary Policy", *Journal of the European Economic Association Papers and Proceedings*, May, forthcoming.
- GIANNONI, M., AND M.WOODFORD, (2003). "Optimal Inflation Targeting Rules", Ben S. Bernanke and Michael Woodford, eds., *Inflation Targeting*, Chicago: University of Chicago Press.
- GUERRIERI, L., (2005). "The Inflation Persistence of Staggered Contracts", *Journal of Money, Credit and Banking*, forthcoming.
- GUMBAU-BRISA, F., (2004). "Learning and Inflation Persistence in General Equilibrium", mimeo, Federal Reserve Bank of Boston.
- HOLDEN, S., AND J. DRISCOLL, (2003). "Inflation Persistence and Relative Contracting", *American Economic Review*, 93, 1369-1372.
- 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.
- IRELAND, P., (2003). "Irrational Expectations and Econometric Practice", unpublished manuscript.
- JONDEAU, E., and H. LE BIHAN, (2003). "ML vs GMM Estimates of Hybrid Macroeconomic Models (With an Application to the "New Phillips Curve")", Banque de France Working Paper, no.103.
- KILEY, M., (2004). "Is Moderate-To-High Inflation Inherently Unstable?", Federal Reserve Board of Governors, Finance and Economics Discussion Series 2004-43.

- KREPS, D., (1998). “Anticipated Utility and Dynamic Choice”, 1997 Schwartz Lecture, in *Frontiers of Research in Economic Theory*, Edited by Donald P. Jacobs, Ehud Kalai, and Morton Kamien, Cambridge University Press, Cambridge, England
- LEVIN, A., AND J. PIGER, (2003). “Is Inflation Persistence Intrinsic in Industrial Economies?”, Federal Reserve Bank of St. Louis Working Paper 2002-023E.
- LINDE, J., (2002). “Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach”, Riksbank Working Paper No. 129.
- MANKIW, N.G., AND R. REIS, (2002). “Sticky Information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve”, *Quarterly Journal of Economics*, vol. 117 (4), pp. 1295-1328.
- MILANI, F., (2004). “Expectations, Learning and Macroeconomic Persistence”, mimeo, University of California – Irvine.
- (2005a). “Learning, Monetary Policy Rules, and Macroeconomic Stability”, mimeo, University of California – Irvine.
- (2005b). “Estimating and Comparing New Keynesian Models with Adaptive Learning: Euler-Equation and Infinite-Horizon Learning”, in progress.
- (2005c). “European Inflation Dynamics and Adaptive Learning”, in progress.
- MOLNAR, K., and S.SANTORO, (2005). “Optimal Monetary Policy when Agents are Learning”, mimeo, UPF.
- ORPHANIDES, A., AND J.WILLIAMS, (2003). “Imperfect Knowledge, Inflation Expectations and Monetary Policy”, in Ben Bernanke and Michael Woodford, eds., *Inflation Targeting*. Chicago: University of Chicago Press.
- (2005a). “The Decline of Activist Stabilization Policy”, *Journal of Economic Dynamics and Control*, vol. 29(11), pages 1927-1950.
- (2005b). “Inflation Targeting under Imperfect Knowledge”, paper presented at the *Ninth Annual Conference of the Central Bank of Chile: Monetary Policy Under Inflation Targeting*.
- PRESTON, B., (2003). “Learning About Monetary Policy Rules When Long-Horizon Expectations Matter”, *International Journal of Central Banking*, forthcoming.
- (2004a). “Adaptive Learning and the Use of Forecasts in Monetary Policy”, *Journal of Monetary Economics*, forthcoming.
- (2004b). “Adaptive Learning, Forecast-Based Instrument Rules and Monetary Policy”, mimeo, Columbia University.
- PRIMICERI, G., (2003). “Why Inflation Rose and Fell: Policymakers’ Beliefs and Postwar Stabilization Policy”, mimeo, Northwestern University.
- REIS, R., AND F.PIVETTA, (2003). “The Persistence of Inflation in the United States”, mimeo, Harvard University.
- ROBERTS, J.M., (1997). “Is Inflation Sticky?”, *Journal of Monetary Economics* 39, 173-196.
- (1998). “Inflation Expectations and the Transmission of Monetary Policy”, Federal Reserve WP 1998-43.

- SANTORO, S., (2004). “Adaptive Learning and Inflation Dynamics in a Flexible Price Model”, mimeo, UPF.
- SARGENT, T.J., (1993). *Bounded Rationality in Macroeconomics*. Oxford University Press.
- (1999). *The Conquest of American Inflation*. Princeton University Press.
- SBORDONE, A., (2003). “Prices and Unit Labor Costs: a New Test of Price Stickiness”, *Journal of Monetary Economics*, Vol. 49 (2), March.
- 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.
- SVENSSON, L.E.O., (2003). “Monetary Policy and Learning”, *Federal Reserve Bank of Atlanta Economic Review*, Third Quarter, 11-16.
- WILLIAMS, N., (2003). “Adaptive Learning and Business Cycles”, mimeo, Princeton University.
- WOODFORD, M., (2003a). “Imperfect Common Knowledge and the Effects of Monetary Policy”, in P.Aghion, R.Frydman, J.Stiglitz and M.Woodford (eds.), *Knowledge, Information, and Expectations in Modern Macroeconomics: in Honor of Edmund S. Phelps*. Princeton University Press..
- (2003b). *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.

Structural Coefficients (Stand.Errors)						
$\kappa= 0.015$		$\kappa= 0.02$		$\kappa= 0.025$		
	with output gap	with marg. cost	with output gap	with marg. cost	with output gap	with marg. cost
γ	0.139 (0.21)	0.047 (0.18)	-0.001 (0.15)	0.045 (0.17)	-0.10 (0.12)	-0.006 (0.148)
β	1.125 (0.06)	1.019 (0.05)	1.114 (0.04)	1.01 (0.05)	1.10 (0.03)	1.00 (0.044)
δ	0.229 (0.04)		0.222 (0.032)		0.212 (0.028)	
α		0.671 (0.02)		0.679 (0.02)		0.687 (0.02)

Table 1 - Estimates Model with Learning. Equations estimated by NLS.

	Output gap as driving variable	Marg. cost as driving variable
	Full-sample	Full-sample
Constant-Gain κ	0.02	0.025

Table 2 - Best fitting constant-gain coefficient κ . *Note:* the estimated equation is $\pi_t = \delta x_t + \beta \hat{E}_{t-1} \pi_{t+1} + u_t$, where $\hat{E}_{t-1} \pi_{t+1} = \phi_{0,t-1} (1 + \phi_{1,t-1}) + \phi_{1,t-1}^2 \pi_{t-1}$.

	Output gap as driving variable	Marg. cost as driving variable	
	(1)	(3)	(2)
Constant-Gain κ	0.068	0.0995	0.0355
Indexation term $\gamma/(1 + \gamma\beta)$	0.0426	0.23	0.1681

Table 3 - Best-fitting gain coefficients. *Note:* learning rules

- (1) $\pi_t = \phi_{0,t} + \phi_{1,t} \pi_{t-1} + \phi_{2,t} x_{t-1} + \varepsilon_t$
- (2) $\pi_t = \phi_{0,t} + \phi_{1,t} \pi_{t-1} + \phi_{2,t} s_{t-1} + \varepsilon_t$
- (3) $\pi_t = \phi_{0,t} + \phi_{1,t} \pi_{t-1} + \phi_{2,t} x_{t-1} + \phi_{3,t} i_{t-1} + \varepsilon_t$

Welfare Loss (learning is 'true' model)	
Optimal Policy assuming Learning	2.28
Optimal Policy assuming RE	5.23

Table 8 - Welfare losses implied by optimal monetary policies assuming learning and rational expectations when learning is true model.

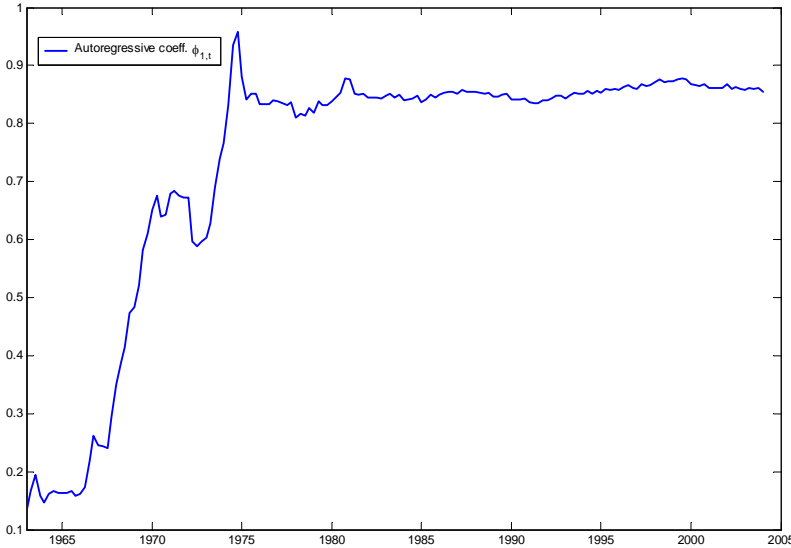
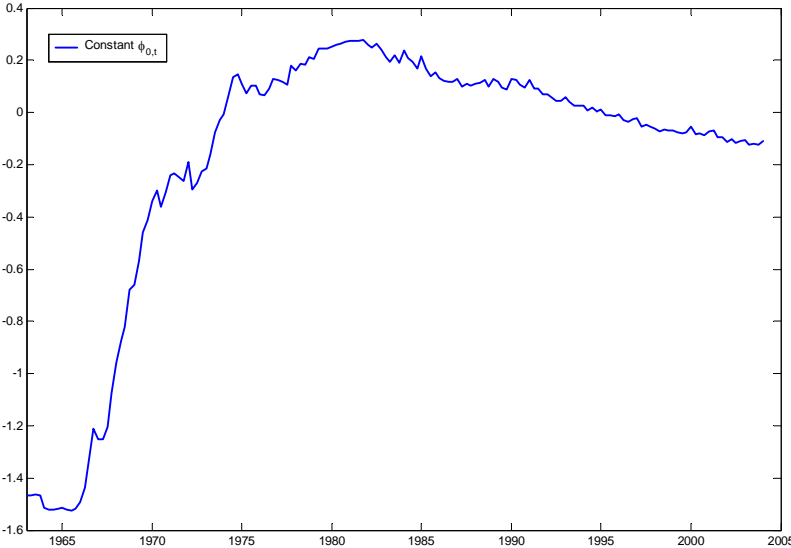


Figure 1 - Evolution of agents' beliefs over time.

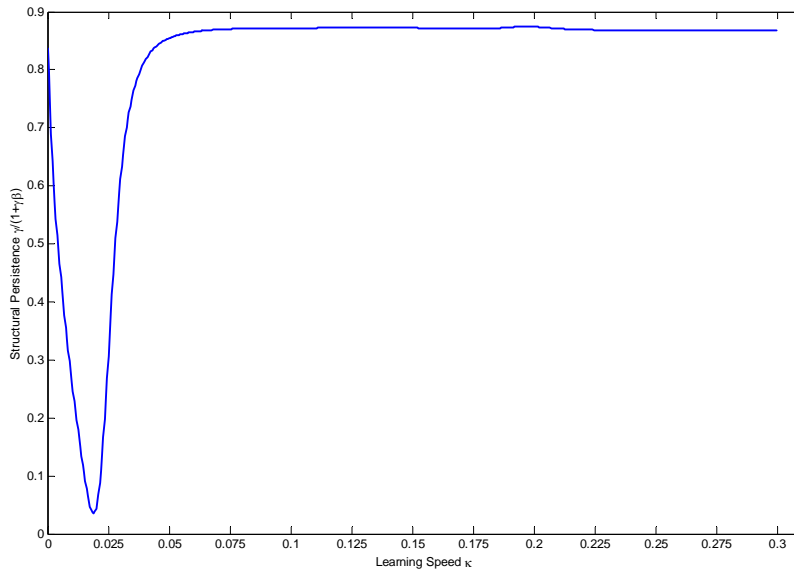


Figure 2 - Estimate of structural persistence parameter across constant gain coefficient values κ (inflation equation with output gap).

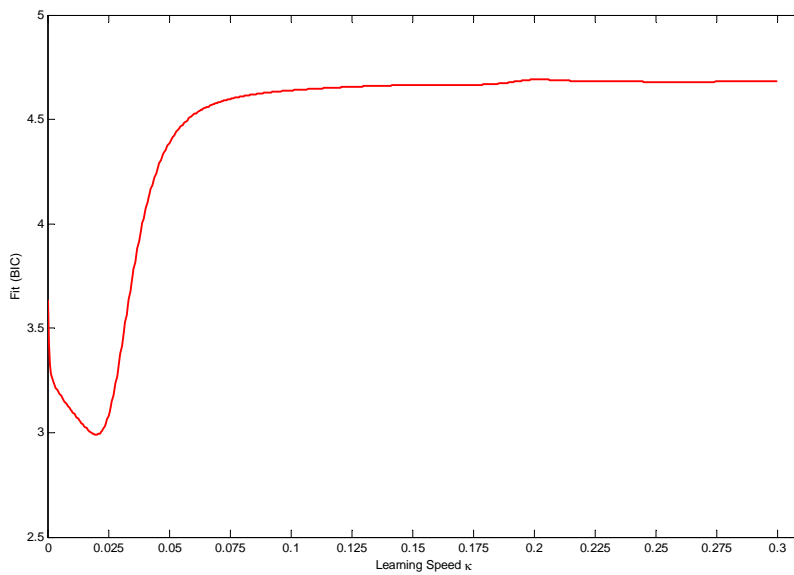


Figure 3 - Fit across constant gain coefficient values κ (inflation equation with output gap). Fit is measured by BIC.

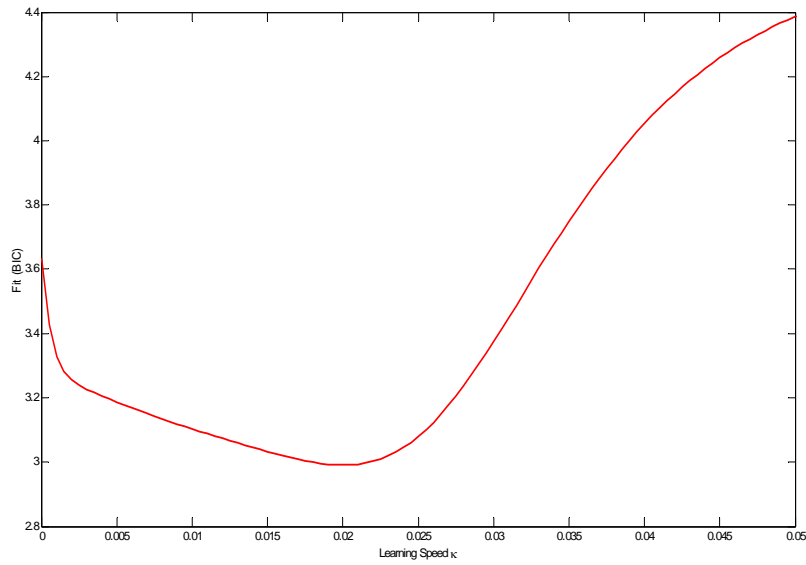


Figure 4 - Fit across constant gain coefficient values κ (inflation equation with output gap).
Narrower view.

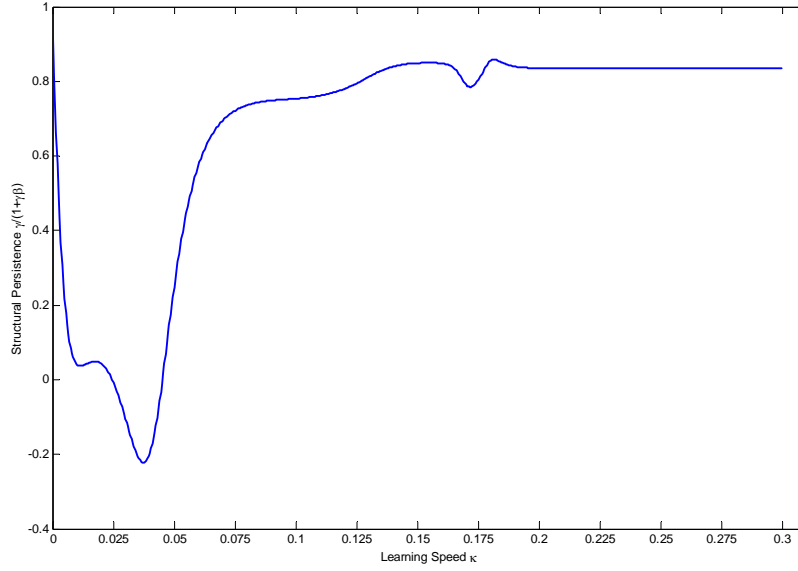


Figure 5 - Estimate of structural persistence parameter across constant gain coefficient values κ (inflation equation with marg. cost).

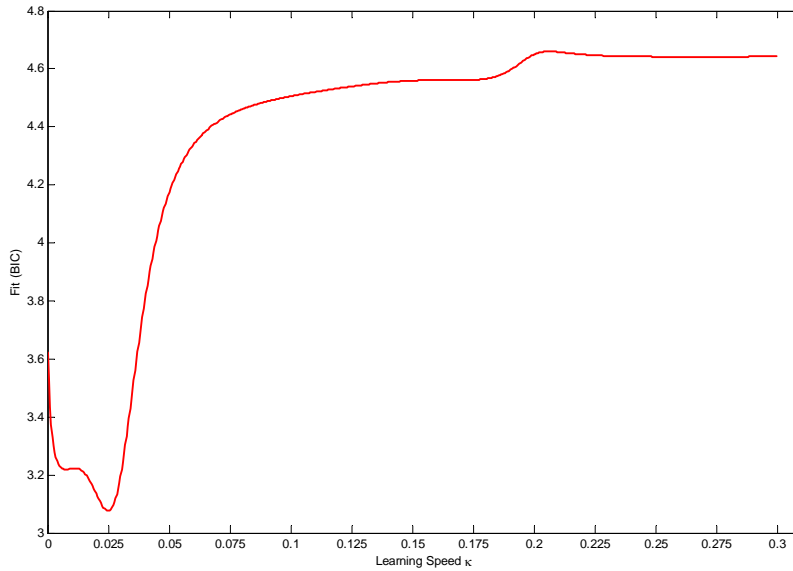


Figure 6 - Fit across constant gain coefficient values κ (inflation equation with marg. cost).

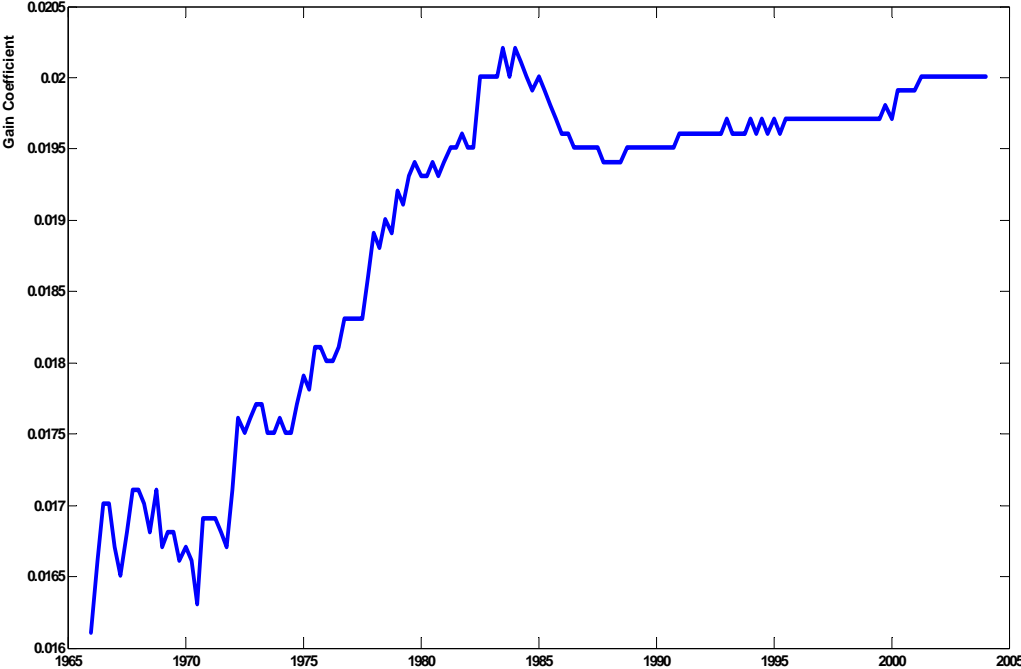


Figure 7 - The evolution of the best-fitting gain.

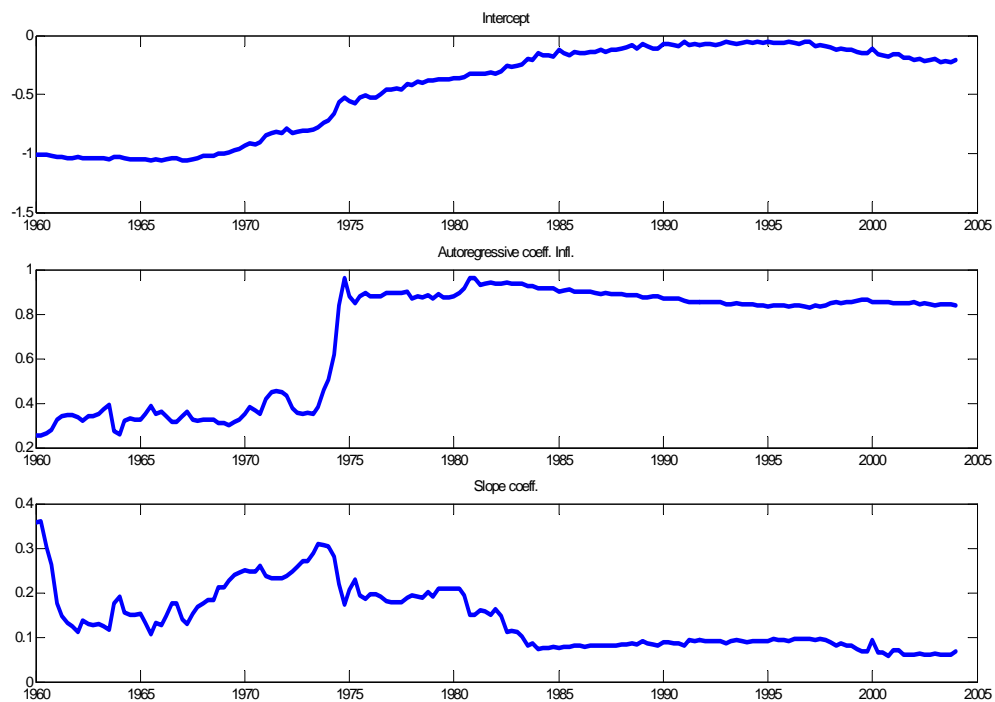


Figure 8 - Evolution of agents' beliefs (learning rule with output gap)

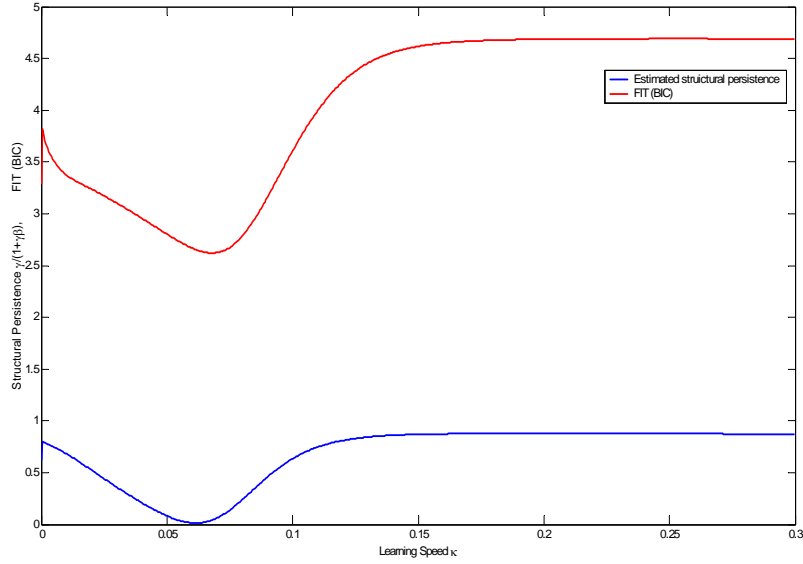


Figure 9 - Fit (dashed) and estimated persistence (solid) across constant gain coefficient values κ (inflation equation with output gap/ Phillips curve learning rule).

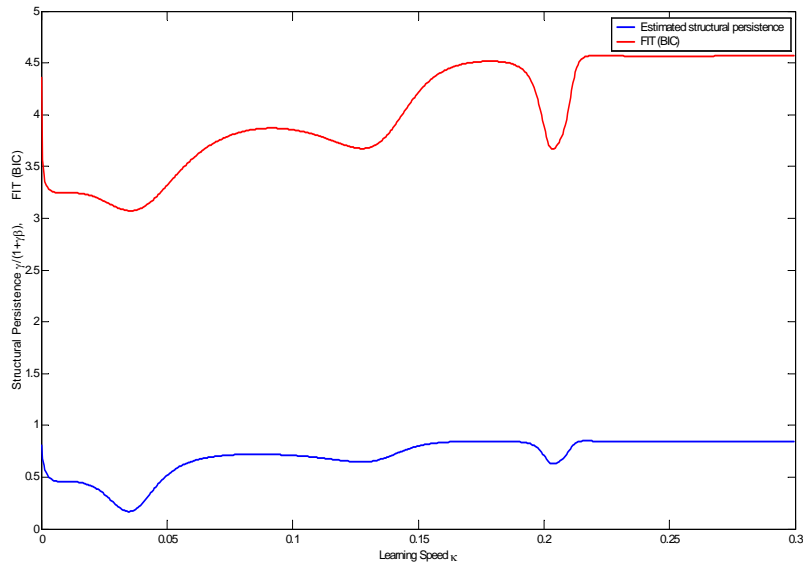


Figure 10 - Fit (dashed) and estimated persistence (solid) across constant gain coefficient values κ (inflation equation with marginal costs/ Phillips curve learning rule).

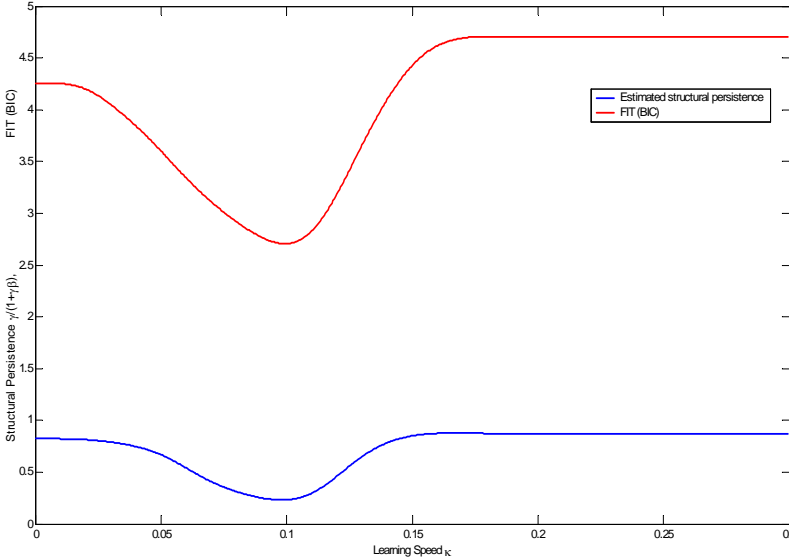


Figure 11 - Fit (dashed) and estimated persistence (solid) across constant gain coefficient values κ (learning rule with interest rates).

Notes

¹Intrinsic in the sense that inflation persistence is a ‘stylized fact’ that we should expect in most industrialized economies.

²Also related is the paper by Santoro (2004) who, independently from this work, proposes learning in a flexible price model to induce persistence in inflation.

³The full consumer maximization problem with learning is described in companion papers (Milani 2004a,b), which also examine time non-separable preferences.

⁴Consistently with most of the New Keynesian literature I log-linearize around a zero steady-state for inflation. I am therefore abstracting from the complications arising from log-linearizing around a positive inflation steady-state. For an account of the possible implications of this choice, see Kiley (2004) and Ascari (2004). Indexation permits to avoid the problems due to trend inflation, as shown in the appendix A of Ascari (2004). However, the present paper works with partial indexation, remaining partly vulnerable to such problems.

⁵I will omit ‘ $\hat{\cdot}$ ’ from the following section to save some notation. I also omit ‘ i ’ as a superscript to indicate the i -th firm, being the problem identical for every firm.

⁶For a different approach of considering learning, see Preston (2003), where learning is introduced directly from the primitive assumptions of multi-period decision problems. The derived law of motion for inflation will be equal to (10). My choice is instead similar to most papers in the adaptive learning literature (Evans and Honkapohja 2001, 2003, Bullard and Mitra 2002, Bullard and Eusepi 2003, and Williams 2003 are examples). See Milani (2004b) for estimations considering both approaches.

⁷According to an *anticipated utility model*, each period agents maximize their expected utility taking their beliefs and the model as constant, although the model is recursively estimated. But when more data become available, agents update their beliefs and use this new knowledge to maximize expected utility. Agents are therefore learning, but they are not involved in active experimentation as a fully rational behavior would imply.

⁸This particular forecasting rule would also arise as the correct functional form of rational inflation expectations in a simple New Keynesian model composed by an aggregate supply equation, given by (15), a linearized Euler equation that represents the aggregate demand equation, and a simple Taylor rule as the monetary policy rule, and where the shocks are i.i.d. Lagged inflation would be the only state variable in the model and therefore firms would learn using the Minimum State Variable solution.

⁹All the data are downloaded from FRED, the database of the Federal Reserve Bank of Saint Louis.

¹⁰The agents start with an estimate of the intercept equal to -1.00 and an estimate of the autoregressive parameter equal to 0.25. The initial variance-covariance matrix equals $\begin{bmatrix} 0.245 & 0.029 \\ 0.029 & 0.027 \end{bmatrix}$.

¹¹Using x_t and s_t instead of x_{t-1} and s_{t-1} does not alter the results.

¹²Recent papers that have started to study the optimal monetary policy under adaptive learning are those by Gaspar, Smets, and Vestin 2006, and Molnar and Santoro 2005).

¹³The optimal policy under learning is calculated by assuming that the policymaker acts as an anticipated utility maximizer. Therefore, the policymaker recursively chooses the optimal policy under the assumption that the estimated parameter will remain unchanged in the future. In this section, I am abstracting from the more complicated case of active experimentation.