Do Long-Horizon Expectations Matter for New Keynesian Models?

Preliminary and Incomplete

Laura Gáti

November 6, 2019

Abstract

When assessing the fit of New Keynesian (NK) models with statistical learning, common practice is to write down the log-linearized first-order conditions of the NK model and replace the objective expectations operator with the subjective one that captures learning ("Euler-equation approach"). As Preston (2005) argues, the Euler-equation approach imposes market clearing conditions that agents in the model have no knowledge of, implicitly making long-horizon expectations irrelevant for model dynamics. I take up the Preston argument to investigate numerically to what extent the Euler-equation approach delivers different dynamics than a model where long-horizon expectations are explicitly spelled out. Is the Euler-equation approach an innocent simplification that doesn't matter quantitatively, or does it distort the behavior of NK models with learning, leading to false predictions and inferences?

1 Introduction

The last two decades have seen increasing skepticism in macroeconomics about the assumption of rational expectations, so frequently made in both theoretical and applied work. This has come hand-in-hand with a renewed interest in information frictions on the hand and bounded rationality on the other. One avenue for departing from rational expectations is the adaptive learning literature that postulates that economic agents use ad-hoc forecasting rules to form expectations, but, like econometricians, update these rules as new data arrives.

Much work within the learning literature has been devoted to policy analysis using a learning version of the New Keynesian framework. In particular, the issue of optimal monetary policy has been revisited from a learning perspective. However, the learning literature has faced a conceptual divide. Many scholars have advocated the so-called Euler-equation approach that simply replaces the expectation operator in the loglinearized first-order conditions of the model with subjective expectations that are then specified according to various learning assumptions. By contrast, Preston (2005) shows that this is inconsistent with the microfoundations of the model as it implies that agents make suboptimal forecasts. Instead, he demonstrates that the conceptually correct learning representation involves what he calls "long-horizon forecasts:" infinite discounted sums of forecasts of all future periods.

In this paper I investigate whether the Euler-equation approach, if conceptually not accurate, can at least deliver dynamics of model variables that are consistent with those of Preston's long-horizon approach. If the two approaches yield dynamics are that are close enough then the issue of conceptual correctness becomes second-order because the model generates the same predictions no matter which learning framework is used. For example, conclusions about optimal monetary policy under learning would in this case be identical under Euler-equation or long-horizon learning. However, if dynamics under the two frameworks are far apart, then having the conceptually correct formulation becomes a first-order issue for any model analysis. This question is also relevant because the long-horizon formulation is mathematically more involved than the Euler-equation approach. Thus, from an applied perspective we want to know: do we get anything wrong if we take the shortcut?

1.1 Related literature

The paper belongs to the adaptive learning literature in macroeconomics. The go-to reference in this literature is Evans and Honkapohja (2001). This book not only lays out the technicalities of how to model statistical learning in macroeconomic frameworks but also implicitly defines the research agenda of the learning literature. Thus, the vast majority of work in learning analyzes when and under what conditions a macroeconomic model with a specific learning framework is expectationally stable (E-stable), that is, converges to rational expectations over time. Prominent examples are Evans and Honkapohja (2003), Marcet and Sargent (1989) or Eusepi and Preston (2018).

One branch of the learning literature investigates particular aspects of E-stability. For example, Bullard and Mitra (2002) and Preston (2005) emphasize that in New Keynesian models with learning, satisfying the Taylor principle is sufficient for E-stability. The analysis of Ferrero (2007) focuses on the speed with which nonrational expectations converge to rational ones. Ferrero (2007) finds that monetary policy can increase the speed of convergence by reacting strongly to private agents' expectations, yet may not always find it optimal to do so.

A small number of papers, such as Eusepi and Preston (2011) or Williams (2003) analyze how learning affects the responses of macro models to shocks. The main finding in this line of work is that statistical learning introduces expectations as a new state variable in the model. This in general dampens shocks on impact but increases persistence; even i.i.d. shocks take much longer to propagate through the system. Depending on how learning is modeled, the system can exhibit higher or lower levels of volatility.

A common theme throughout these strands of literature is how the assumption of adaptive learning changes the monetary policy problem. Since the New Keynesian model is the most widely used framework when analyzing monetary policy, most of the above-mentioned papers rely on a textbook NK model, extended with an adaptive learning type of expectation formation. The findings

are surprisingly diverse. While Eusepi and Preston (2018) and Molnár and Santoro (2014) conclude that optimal monetary policy in an NK model with learning needs to be more aggressive on inflation than under rational expectations, Eusepi et al. (2018a) find the opposite: optimal policy should be less aggressive.

One reason for varying conclusions about optimal monetary policy is different approaches to modeling learning. To the best of my knowledge, the only survey paper on modeling assumptions in the learning literature to date is Graham et al. (2011). This paper enumerates a set of modeling choices such as constant versus decreasing gain, including or excluding a constant in private agents' learning rule or the number of forecast horizons agents take into account. The degree to which agents are forward-looking is found to matter enormously for the speed of convergence. Other specification parameters of learning seem to matter little. Constant gains do not alter business cycle properties of the model, and adding an intercept to the learning rule only increases the effect of learning by a small amount.

While the aforementioned paper by Preston (2005) studies optimal monetary policy, it also carries a similar modeling message. Like Graham et al. (2011), Preston argues that the assumption about how many periods-ahead forecasts agents consider when making economic decisions matters. In particular, he demonstrates that modeling learning using the Euler-equation approach is inconsistent with optimal household behavior because it implies households ignoring their lifetime wealth.

The objective of this paper is related to that of Graham et al. (2011) and Preston (2005) in focusing on the role of a particular modeling choice: long-horizon forecasts versus Euler-equation learning. The model setup will thus echo that of Preston (2005).

The paper is organized as follows. Section 2 describes the main body of the model while Section 3 describes the learning framework. Section 4 presents the numerical results. Section 5 concludes.

2 The model

Following Preston (2005), the model is a textbook New Keynesian model where rational expectations (RE) have been replaced by nonrational ones, formed using statistical learning. In outlining the model, I will thus skip derivations and refer the reader to Woodford (2011), Chapter 3.

2.1 Households

A continuum of infinitely-lived households maximize expected lifetime utility

$$\hat{\mathbb{E}}_t \sum_{T=t}^{\infty} \beta^{T-t} \left[U(C_T^i) - \int_0^1 v(h_T^i(j))dj \right] \tag{1}$$

where $U(\cdot)$ and $v(\cdot)$ are the utility of consumption and disutility of labor, respectively. $\hat{\mathbb{E}}$ is the subjective expectation operator, defined in Section 3. C_t^i is the Dixit-Stiglitz aggregate of consumption goods $c_t^i(j)$

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

with $\theta > 1$ representing the elasticity of substitution between the varieties of consumption goods. Note that each good j requires its own type of labor, $h_T^i(j)$, so that the household supplies labor in the production of all goods j, leading to the integral in Equation (1). Since each good j trades at price $p_t(j)$, the aggregate price level is given by

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

Households' budget constraint is given by

$$B_t^i \leq (1+i_{t-1})B_{t-1}^i + \int_0^1 w_t(j)h_t^i(j) + \Pi_t^i(j)dj - T_t - P_t C_t^i \tag{4}$$

where the term in the integrals captures the income households receive from supplying labor and from remitted profits from each firm j. B_t^i signifies the riskless bond purchases at time t.[1]

2.2 Firms

Firms are producers of the varieties $y_t(j)$ in a monopolistically competitive environment. They set the price of the variety they produce, $p_t(j)$ to maximize the present discounted value of profit streams

$$\hat{\mathbb{E}}_t \sum_{T=t}^{\infty} \alpha^{T-t} Q_{t,T} \left[\Pi_t^j(p_t(j)) \right] \tag{5}$$

where

$$Q_{t,T} = \beta^{T-t} \frac{P_t U_c(C_T)}{P_T U_c(C_t)} \tag{6}$$

is the stochastic discount factor from households and firm profits Π_t^j are given by

$$\Pi_t^j = p_t(j)y_t(j) - w_t(j)f^{-1}(y_t(j)/A_t) \tag{7}$$

Firms produce using the production technology $y_t(j) = A_t f(h_t(j))$, whose inverse, $f^{-1}(\cdot)$, signifies the amount of labor input. Moreover, since they are monopolistically competitive, firms face a downward-sloping demand curve

$$y_t(j) = Y_t \left(\frac{p_t(j)}{P_t} \right)^{-\theta} \tag{8}$$

The parameter α in Equation (5) is the Calvo probability of not being able to adjust one's price in a given period. As is well known, this gives rise to nominal rigidities and makes the New Keynesian model suitable for the analysis of inflation dynamics.

2.3 Three-equation New Keynesian model

Again referring to textbook treatments for the model derivations (such as Woodford (2011)), I turn to the lognlinearized first-order conditions of the model. Adding a specification for monetary policy, the NK models boils down to three equations. Denoting the current output gap with x_t, inflation with π_t and the nominal interest rate with i_t, the model can be summarized as follows.

$$x_t = \mathbb{E}_t\, x_{t+1} - \sigma(i_t - \mathbb{E}_t\, \pi_{t+1}) + \sigma r_t^n \tag{9}$$

$$\pi_t = \kappa x_t + \beta\, \mathbb{E}_t\, \pi_{t+1} + u_t \tag{10}$$

$$i_t = \psi_\pi \pi_t + \psi_x x_t + \bar{i}_t \tag{11}$$

Equation (9) is the IS curve, linking today's output gap to expectations of tomorrow's inflation and output gap, as well as to today's interest rate and a shock, r_t^n which can be shown to have the interpretation of the natural real rate of interest, $r_t^n = \mathbb{E}_t\, y^n - y_t^n$, where y_t^n is the natural level of output. Equation (10) is the Phillips curve, establishing current inflation as the result of firms' price setting behavior. The variable u_t is a cost-push shock. The model is closed with a specification of monetary policy, Equation (11). Here this takes the simplest Taylor-rule form, letting monetary policy react to current inflation and current output gap with the coefficients ψ_π and ψ_x, and allowing for innovations to the nominal interest rate as $\bar{i}_t$.

[1] I have suppressed the money asset in this expression to simplify the exposition. This does not change the implication of the model in any way since it simply represents the cashless limit of an economy with explicit money balances.

The Euler-equation approach consists of adopting equations (9)-(11) and replacing the rational expectation (RE) operator $\mathbb{E}$ with a subjective expectation operator, $\hat{\mathbb{E}}$. This gives rise to the following system:

$$x_t = \hat{\mathbb{E}}_t x_{t+1} - \sigma(i_t - \hat{\mathbb{E}}_t \pi_{t+1}) + \sigma r_t^n \tag{12}$$

$$\pi_t = \kappa x_t + \beta \hat{\mathbb{E}}_t \pi_{t+1} + u_t \tag{13}$$

$$i_t = \bar{i}_t + \psi_\pi \pi_t + \psi_x x_t \tag{14}$$

The subjective expectation $\hat{\mathbb{E}}$ is then specified separately, usually as a reduced-form forecasting rule with coefficients updated using recursive least squares. As I spell out explicitly in Section 3, I will use the most general formulation here, namely decreasing gain learning where agents learn both the slope and the constant of their forecasting rule.

Preston (2005) makes a case for a different way of translating the RE NK model to a learning setting. His argument is that the derivation of system (9)-(10) relies on imposing information on individual agents' forecasts that are only available to firms and households under rational expectations. For example, only under rational expectations are firms and households aware of the fact that they are identical in equilibrium and can thus impute market clearing conditions. Agents unaware of the structure of the model, as is assumed in the adaptive learning literature, do not have access to this information. Therefore the representation (12)-(13) is not a valid learning representation of the NK model. In particular it ignores agents' need to rely on forecasts of infinite-period ahead variables to take optimal economic decisions.

Thus, as Preston (2005) demonstrates, the conceptually valid learning representation of the NK model according consists of the following equations:

$$x_t = -\sigma i_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} \beta^{T-t} \big((1-\beta) x_{T+1} - \sigma(\beta i_{T+1} - \pi_{T+1}) + \sigma r_T^n\big) \tag{15}$$

$$\pi_t = \kappa x_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} \big(\kappa\alpha\beta x_{T+1} + (1-\alpha)\beta \pi_{T+1} + u_T\big) \tag{16}$$

$$i_t = \psi_\pi \pi_t + \psi_x x_t + \bar{i}_t \tag{17}$$

Clearly, the difference between systems (12)-(13) and (15)-(16) is that in the latter, long-horizon expectations are present. A more subtle issue is that the long-horizon expectations in the latter system also embody a different informational assumption, namely that agents in the model do not know the model and thus do not rely on market clearing conditions when forecasting.

It is clear that from an informational consistency standpoint that the system (15)-(16) is conceptually the correct one. It is also visibly more complicated, and more difficult to work with. Therefore the question arises: is the difference between the two approaches merely of conceptual nature, or does it affect model dynamics?

If the Euler-equation learning approach of (12)-(13) yields dynamics that are close enough to the more complicated long-horizon learning approach in (15)-(16), then the Preston argument is not a first-order issue because taking the shortcut does not impact the model's predictions and thus its policy usefulness.

2.4 Compact notation

For expositional purposes, let me now introduce the following notation. Let exogenous states be summarized as:

$$s_t \equiv \begin{pmatrix} r_t^n \\ \bar{i}_t \\ u_t \end{pmatrix} \quad P \equiv \begin{pmatrix} \rho_r & 0 & 0 \\ 0 & \rho_i & 0 \\ 0 & 0 & \rho_u \end{pmatrix} \quad \epsilon_t \equiv \begin{pmatrix} \varepsilon_t^r \\ \varepsilon_t^i \\ \varepsilon_t^u \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_r & 0 & 0 \\ 0 & \sigma_i & 0 \\ 0 & 0 & \sigma_u \end{pmatrix} \tag{18}$$

where P and Σ are the matrix of shock persistence and standard deviation respectively. The law of motion of the exogenous states can then be written in the compact form

$$s_t = Ps_{t-1} + \epsilon_t \qquad \epsilon_t \sim \mathcal{N}(\mathbf{0}, \Sigma) \tag{19}$$

Similarly, letting z_t summarize the endogenous variables as

$$z_t \equiv \begin{pmatrix} \pi_t \\ x_t \\ i_t \end{pmatrix} \tag{20}$$

allows me to write the models (9)-(10), (12)-(13) and (15)-(16) compactly as

$$z_t = A_p^{RE}\, \mathbb{E}_t\, z_{t+1} + A_s^{RE} s_t \tag{21}$$

$$z_t = A_p^{RE} \hat{\mathbb{E}}_t z_{t+1} + A_s^{RE} s_t \tag{22}$$

$$z_t = A_a^{LR} f_a(t) + A_b^{LR} f_b(t) + A_s^{LR} s_t \tag{23}$$

$$s_t = Ps_{t-1} + \epsilon_t \tag{24}$$

where the A matrices gather coefficients and are specified below. Here $f_a(t)$ and $f_b(t)$ capture period t long-horizon expectations of the endogenous states z:

$$f_a(t) \equiv \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} z_{T+1} \qquad\qquad f_b(t) \equiv \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\beta)^{T-t} z_{T+1} \tag{25}$$

The coefficient matrices are given by:

$$A_p^{RE} = \begin{pmatrix} \beta + \frac{\kappa\sigma}{w}(1-\psi_\pi\beta) & \frac{\kappa}{w} & 0 \\ \frac{\sigma}{w}(1-\psi_\pi\beta) & \frac{1}{w} & 0 \\ \psi_\pi\big(\beta + \frac{\kappa\sigma}{w}(1-\psi_\pi\beta)\big) + \psi_x \frac{\sigma}{w}(1-\psi_\pi\beta) & \psi_x(\frac{1}{w}) + \psi_\pi(\frac{\kappa}{w}) & 0 \end{pmatrix} \tag{26}$$

$$A_s^{RE} = \begin{pmatrix} \frac{\kappa\sigma}{w} & -\frac{\kappa\sigma}{w} & 1 - \frac{\kappa\sigma\psi_\pi}{w} \\ \frac{\sigma}{w} & -\frac{\sigma}{w} & -\frac{\sigma\psi_\pi}{w} \\ \psi_x(\frac{\sigma}{w}) + \psi_\pi(\frac{\kappa\sigma}{w}) & \psi_x(-\frac{\sigma}{w}) + \psi_\pi(-\frac{\kappa\sigma}{w}) + 1 & \psi_x(-\frac{\sigma\psi_\pi}{w}) + \psi_\pi(1 - \frac{\kappa\sigma\psi_\pi}{w}) \end{pmatrix} \tag{27}$$

$$A_a^{LR} = \begin{pmatrix} g_{\pi a} \\ g_{xa} \\ \psi_\pi g_{\pi a} + \psi_x g_{xa} \end{pmatrix} \quad A_b^{LR} = \begin{pmatrix} g_{\pi b} \\ g_{xb} \\ \psi_\pi g_{\pi b} + \psi_x g_{xb} \end{pmatrix} \quad A_s^{LR} = \begin{pmatrix} g_{\pi s} \\ g_{xs} \\ \psi_\pi g_{\pi s} + \psi_x g_{xs} + \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \end{pmatrix} \tag{28}$$

$$g_{\pi a} = (1 - \frac{\kappa\sigma\psi_\pi}{w}) \left[(1-\alpha)\beta, \kappa\alpha\beta, 0\right] \tag{29}$$

$$g_{xa} = \frac{-\sigma\psi_\pi}{w} \left[(1-\alpha)\beta, \kappa\alpha\beta, 0\right] \tag{30}$$

$$g_{\pi b} = \frac{\kappa}{w} \left[\sigma(1-\beta\psi_\pi), (1-\beta-\beta\sigma\psi_x, 0\right] \tag{31}$$

$$g_{xb} = \frac{1}{w} \left[\sigma(1-\beta\psi_\pi), (1-\beta-\beta\sigma\psi_x, 0\right] \tag{32}$$

$$g_{\pi s} = (1 - \frac{\kappa\sigma\psi_\pi}{w}) \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} (I_3 - \alpha\beta P)^{-1} - \frac{\kappa\sigma}{w} \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} (I_3 - \beta P)^{-1} \tag{33}$$

$$g_{xs} = \frac{-\sigma\psi_\pi}{w} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} (I_3 - \alpha\beta P)^{-1} - \frac{\sigma}{w} \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} (I_3 - \beta P)^{-1} \tag{34}$$

$$w = 1 + \sigma\psi_x + \kappa\sigma\psi_\pi \tag{35}$$

Before turning to the specification of learning, I also note that the solution of the RE model can

be written as the state-space representation

$$s_t = hs_{t-1} + \epsilon_t \tag{36}$$
$$z_t = gs_t \tag{37}$$

where g and h are matrices of coefficients. Using this, one can write the law of motion of the observables z_t as

$$z_t = ghs_{t-1} + ge_t \tag{38}$$

Note that a RE solution to the NK model implies that current observables depend on past states and current innovations with a loading matrix gh and g respectively.

3 Learning

Let me now turn to the specification of the subjective expectation operator, $\hat{\mathbb{E}}$. In the adaptive learning literature, it is common practice to postulate the following forecasting rule:

$$\hat{\mathbb{E}}_t z_{t+1} = a_{t-1} + b_{t-1}s_t \tag{39}$$

where a and b are the learning coefficients that agents update as time goes by. It is worth contrasting this perceived law of motion (PLM) with the law of motion of z_t under rational expectations, equation (38). In particular, note that (38) implies $a = 0$ and $b = gh$, that is, a particular slope coefficient and an intercept of zero. The interpretation of this is that the crucial assumption of the learning literature is that agents do not know the model. In particular, since they do not know the model, they do not know equation (38). Instead, they form expectations using (39) and evaluate the performance of this forecast in each period in order to update (a, b). Their goal is thus to learn the correct, RE value of (a, b).

Introducing the notation $\phi_{t-1} = (a_{t-1}, b_{t-1})$, here 3×4, I can rewrite the PLM of equation (39) as $\hat{\mathbb{E}}_t z_{t+1} = \phi_t \begin{bmatrix} 1 \\ s_t \end{bmatrix}$. Also in line with standard practice in the learning literature, I impose the anticipated utility assumption:

$$\hat{\mathbb{E}}_{t-1}\phi_{t+h} = \phi_{t-1} \quad \forall\, h \geq 0 \tag{40}$$

This embodies two things. On the one hand, it means that agents today mistakenly believe that they will not update the forecasting rule in the future. This is a technical assumption that simplifies the solution of learning models.[2] On the other hand, anticipated utility also carries a timing assumption: it implies that the belief about ϕ_t was formed at $t-1$. In other words, in the evening of period $t-1$, agents update ϕ_{t-2} to get the ϕ_{t-1} they will use to form expectations in period t. Adding the assumption that agents know the law of motion of the exogenous process (also standard in the learning literature), h-horizon forecasts can be written using the PLM of equation (39) as:

$$\hat{\mathbb{E}}_t z_{t+h} = a_t + b_t P^{h-1} s_t \quad \forall h \geq 1 \tag{41}$$

To close the learning framework one needs to specify how the regression coefficients are updated. I once again follow the standard assumption in the learning literature and let agents update ϕ using a decreasing gain recursive least squares (RLS) algorithm:

$$\phi_t = \left(\phi'_{t-1} + t^{-1}R_t^{-1}\begin{bmatrix} 1 \\ s_{t-1} \end{bmatrix}\left(z_t - \phi_{t-1}\begin{bmatrix} 1 \\ s_{t-1} \end{bmatrix}\right)'\right)' \tag{42}$$
$$R_t = R_{t-1} + t^{-1}\left(\begin{bmatrix} 1 \\ s_{t-1} \end{bmatrix}\begin{bmatrix} 1 & s_{t-1} \end{bmatrix} - R_{t-1}\right) \tag{43}$$

[2] Sargent (1999) shows that relaxing this assumption does not change the conclusions of learning models.

where R_t is the variance-covariance matrix of the regressors and is 4×4.

3.1 Actual laws of motion

Having specified the expectation formation in the learning models, I can now evaluate the laws of motion in these two models. As we saw in section (2), the rational expectation version of the NK model gives rise to a law of motion of observables z_t given by equation (38). For the learning versions of the model, we need to evaluate expectations to obtain the the actual laws of motion (ALMs) at each period t. For the Euler-equation learning model, this means using the PLM of equation (39) to evaluate one-period ahead forecasts and plug those into (22). For the long-horizon learning model, we need to evaluate the long-horizon expectations in equation (25). By plugging the expression for h-period ahead forecasts (41) into (25) and evaluating the infinite sums, I obtain:

$$f_a(t) = \frac{1}{1-\alpha\beta} a_t + b_t(I_3 - \alpha\beta P)^{-1} s_t \qquad f_b(t) = \frac{1}{1-\beta} a_t + b_t(I_3 - \beta P)^{-1} s_t \tag{44}$$

Alternatively I can evaluate each h-period ahead forecast individually using (41), and then sum H of these terms, discounting appropriately.[3]

4 Simulations

Now I turn to the simulations of the NK model under the three specifications of expectation formation: rational expectations, Euler-equation learning and long-horizon learning. The main question of interest is whether the long-horizon specification of the model yields different dynamics than the Euler-equation counterpart for identical expectation formation.

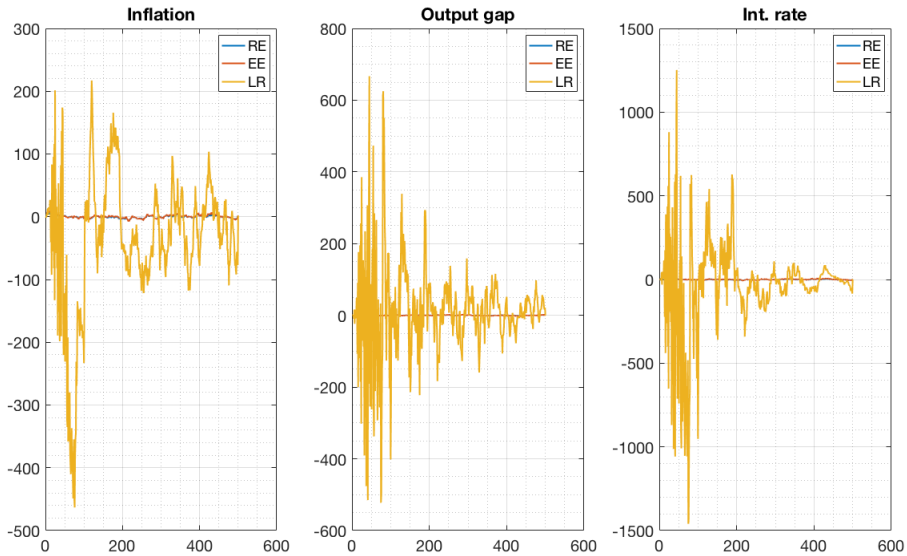
Fig. 1 presents the main results. This figure shows the evolution of the observables inflation, output gap and nominal interest rate over time conditional on the same sequence of shocks. As the top panel makes clear, the dynamics of Euler-equation learning (abbreviated as EE) is extremely different from that of long-horizon learning (abbreviated as LR). In particular, the latter produces much more volatile dynamics and takes considerably longer to converge to rational expectations than EE. As the bottom panel illustrates, after 600 periods, the EE dynamics are fairly close to rational expectations. By contrast, the long-horizon expectation dynamics remain distinct from rational expectations even after 100000 periods (Fig. 2).

One might think that the Euler-equation and long-horizon learning models only differ in terms of the horizon of forecasts that show up in the expectations of agents. To investigate this, I next plot simulations for the same sequence of shocks of the long-horizon learning model where I evaluate each h-period ahead forecast separately and add up H of these forecasts, discounting each as needed. Fig. 3 presents the results. The top panels show the observables for the long-horizon learning model. The blue lines are long-horizon model observables for the long-horizon expectations evaluated analytically (equation (44)), while the red dashed and yellow lines are for the alternative evaluation of long-horizon expectations, using $H = 10000$ and $H = 100$ respectively. It is clear that $H = 10000$ is a good approximation in the sense that observables behave exactly as if analytical expectations had been used. The opposite is true for $H = 100$; this evaluation of long-horizon expectations leads to drastically different dynamics.

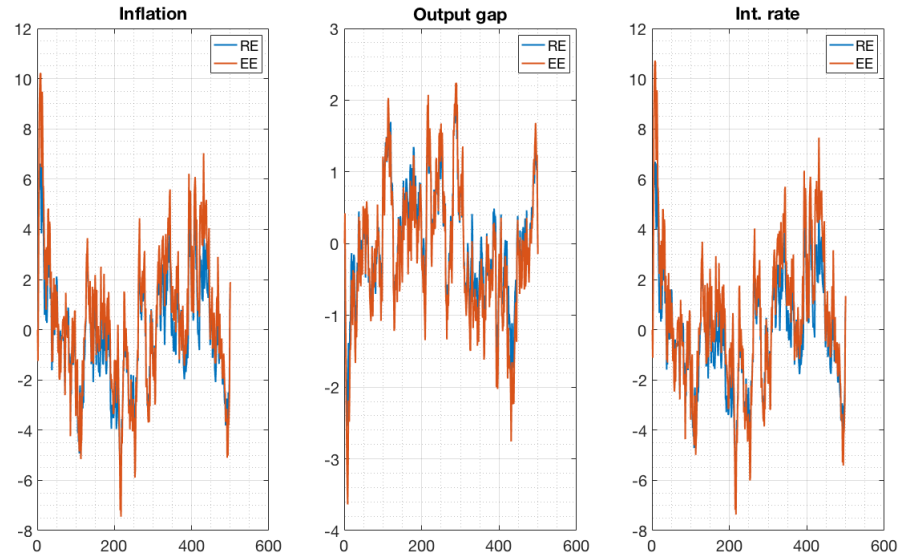
Are these the dynamics of Euler-equation learning? The bottom panel of Fig. 3 gives a stark answer: no. Even if I set $H = 1$, thus considering only one-period ahead forecasts, as in Euler-equation learning, I still do not obtain the Euler-equation learning dynamics (in red here). This is because the long-horizon forecast approach does not only differ from the Euler-equation approach in the amount and horizon of forecasts that figure into period t expectations. Instead, the presence of infinite h-period ahead forecasts reflects the fact that this approach does not endow agents with information the expectational assumptions imply they do not know.

[3] For f_a, already $H = 100$ is not a bad approximation of ∞-horizons, but for f_b to be accurate, I need at least $H = 10000$.

Figure 1: Comparing models



(a) All models



(b) Rational expectations and Euler-equation learning only

Figure 2: Rational expectations and long-horizon learning, $T = 100000$, last 100 periods

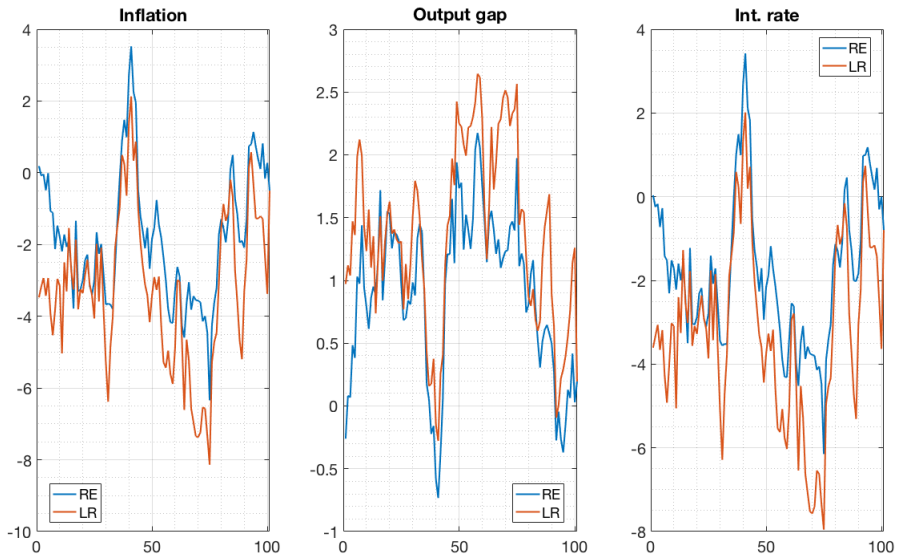
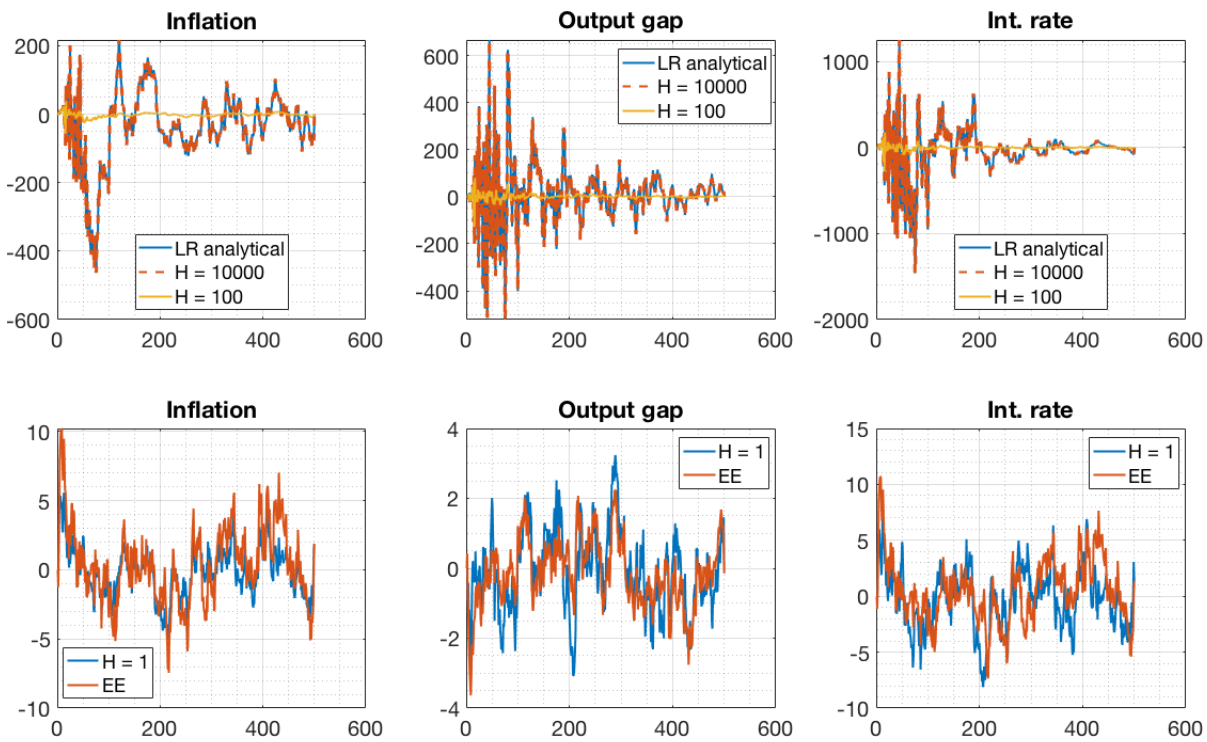


Figure 3: Comparing horizons



5 Conclusion

Many an analysis of New Keynesian models with learning has worked with the Euler-equation approach: replacing the expectation operator of the three-equation NK model with a subjective expectation operator and studying the ensuing dynamics. Here I have investigated the critique of Preston (2005) which suggests that such a formulation is not the correct learning representation of the NK model. Replacing the rational expectation operator with subjective expectations without revisiting the optimal forecasting behavior of agents in the model is at odds with the informational assumption of learning that agents do not know the model.

Such concerns may however be of a theoretical nature only. If it was the case that Euler-equation learning produces nearly identical dynamics with the long-horizon approach, macroeconomists interested only in policy applications of NK models with learning could calmly resort to the admittedly simpler Euler-equation approach. But my simulations reveal that the two approaches give rise to dramatically different dynamics. Thus relying on the Euler-equation approach can be misleading and generate false predictions.

References

Benhabib, J., S. Schmitt-Grohé, and M. Uribe (2001). The perils of Taylor rules. *Journal of Economic Theory 96*(1-2), 40–69.

Bullard, J. and K. Mitra (2002). Learning about monetary policy rules. *Journal of Monetary Economics 49*(6), 1105–1129.

Carvalho, C., S. Eusepi, E. Moench, and B. Preston (2019). Anchored inflation expectations. *Available at SSRN 3018198*.

Clarida, R., J. Gali, and M. Gertler (1999). The science of monetary policy: a New Keynesian perspective. *Journal of economic literature 37*(4), 1661–1707.

Coibion, O., Y. Gorodnichenko, S. Kumar, and M. Pedemonte (2018, June). Inflation expectations as a policy tool? Working Paper 24788, National Bureau of Economic Research.

Cukierman, A. and A. H. Meltzer (1986). A theory of ambiguity, credibility, and inflation under discretion and asymmetric information. *Econometrica 54*(5), 1099–1128.

D'Acunto, F., U. Malmendier, J. Ospina, and M. Weber (2019, September). Exposure to daily price changes and inflation expectations. Working Paper 26237, National Bureau of Economic Research.

Davig, T. and E. M. Leeper (2007). Generalizing the Taylor principle. *American Economic Review 97*(3), 607–635.

Erceg, C. J. and A. T. Levin (2003). Imperfect credibility and inflation persistence. *Journal of monetary economics 50*(4), 915–944.

Eusepi, S., M. Giannoni, and B. Preston (2018a). On the limits of monetary policy.

Eusepi, S., M. P. Giannoni, and B. Preston (2018b). Some implications of learning for price stability. *European Economic Review 106*, 1–20.

Eusepi, S. and B. Preston (2010, July). Central bank communication and expectations stabilization. *American Economic Journal: Macroeconomics 2*(3), 235–71.

Eusepi, S. and B. Preston (2011). Expectations, learning, and business cycle fluctuations. *American Economic Review 101*(6), 2844–72.

Eusepi, S. and B. Preston (2018). The science of monetary policy: An imperfect knowledge perspective. *Journal of Economic Literature 56*(1), 3–59.

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

Evans, G. W. and S. Honkapohja (2003). Expectations and the stability problem for optimal monetary policies. *The Review of Economic Studies 70*(4), 807–824.

Evans, G. W. and S. Honkapohja (2009). Learning and macroeconomics. *Annu. Rev. Econ. 1*(1), 421–449.

Evans, G. W., S. Honkapohja, et al. (2009). Expectations, learning and monetary policy: an overview of recent research. *Monetary policy under uncertainty and learning 13*, 27–76.

Ferrero, G. (2007). Monetary policy, learning and the speed of convergence. *Journal of Economic Dynamics and Control 31*(9), 3006–3041.

Goodfriend, M. (1993). Interest rate policy and the inflation scare problem: 1979-1992. *FRB Richmond Economic Quarterly 79*(1), 1–23.

Graham, L. et al. (2011). Individual rationality, model-consistent expectations and learning. *Centre for Dynamic Macroeconomic Analysis working paper 1112*.

Kozicki, S. and P. A. Tinsley (2001). Shifting endpoints in the term structure of interest rates. *Journal of Monetary Economics 47*(3), 613–652.

Marcet, A. and T. J. Sargent (1989). Convergence of least squares learning mechanisms in self-referential linear stochastic models. *Journal of Economic Theory 48*(2), 337–368.

Molnár, K. and S. Santoro (2014). Optimal monetary policy when agents are learning. *European Economic Review 66*, 39–62.

Preston, B. (2005). Learning about monetary policy rules when long-horizon expectations matter. Mpra paper, University Library of Munich, Germany.

Sargent, T. J. (1999). *The Conquest of American Inflation*. Princeton University Press.

Svensson, L. E. (2015). The possible unemployment cost of average inflation below a credible target. *American Economic Journal: Macroeconomics 7*(1), 258–96.

Williams, N. (2003). Adaptive learning and business cycles. *Manuscript, Princeton University*.

Winkler, F. (2019). The role of learning for asset prices and business cycles. *Journal of Monetary Economics*.

Woodford, M. (2011). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.