

Behavioral New Keynesian Models: Learning vs. Cognitive Discounting

Greta Meggiorini, Fabio Milani

Impressum:

CESifo Working Papers

ISSN 2364-1428 (electronic version)

Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH

The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute

Poschingerstr. 5, 81679 Munich, Germany

Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de

Editor: Clemens Fuest

<https://www.cesifo.org/en/wp>

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: <https://www.cesifo.org/en/wp>

Behavioral New Keynesian Models: Learning vs. Cognitive Discounting

Abstract

This paper estimates a New Keynesian model with new and old behavioral elements. Agents in the model exhibit cognitive discounting, or myopia: they discount variables far into the future at higher rates than typically implied in the benchmark model. We investigate the model under different expectational assumptions: rational expectations, subjective expectations with infinite-horizon learning, and subjective expectations with Euler-equation learning. Under rational expectations, the model necessitates of large, possibly unrealistically so, degrees of myopia. The same result persists under infinite-horizon learning, given that agents are still remarkably far-sighted. But, under Euler-equation learning, the model can fit the data with only minimal estimated degrees of myopia. The results indicate that the empirical evidence for cognitive discounting may be sensitive to the modeling of expectations, and they highlight learning as a key behavioral feature to understand macroeconomic fluctuations.

JEL-Codes: E310, E320, E520, E580, E700.

Keywords: behavioural macroeconomics, cognitive discounting, myopia, inattention, constant-gain learning, behavioural New Keynesian model.

Greta Meggiorini
Department of Economics
University of California
3151 Social Science Plaza
USA – Irvine, CA 92697-5100
gmeggior@uci.edu

<http://sites.google.com/site/gretameggiorini>

Fabio Milani
Department of Economics
University of California
3151 Social Science Plaza
USA – Irvine, CA 92697-5100
fmilani@uci.edu

<http://www.socsci.uci.edu/~fmilani>

1. INTRODUCTION

Empirical research in macroeconomics has increasingly come to rely on microfounded Dynamic Stochastic General Equilibrium (DSGE) models. In most cases, these models describe economies populated by agents who are fully rational: they have unlimited cognitive capacities, they can solve intertemporal optimization problems, and form expectations about future variables according to the rational expectations hypothesis. Recently, however, there have been growing efforts to incorporate behavioral elements, such as limitations in cognition, understanding, or information, by agents. The enriched models can successfully account for empirical aspects that would be otherwise hard to rationalize within more traditional settings.

Focusing, for example, on the leading framework for monetary policy analysis, the New Keynesian model, one problematic aspect is that, under the assumption of rational expectations, it implies only minimal discounting of the future by households and firms. This feature is responsible for puzzling predictions, such as those at the base of the “forward guidance puzzle”, uncovered by Del Negro, Giannoni, and Patterson (2015). Expected shifts in interest rates in the model have unrealistically large effects on current output, and effects that become even larger the further into the future they are set to take place.

A line of research initiated by Gabaix (2014, 2016, 2020) introduces myopic behavior by economic agents in the form of “cognitive discounting”. Cognitive discounting captures the idea that agents cannot fully understand events that will take place in the distant future: a possible interpretation is that agents mentally simulate the future paths of the economy, but when they deal with more distant simulations, they progressively shrink them toward steady-state values. The impact of expectations regarding deviations of variables from steady state, happening far into the future, is substantially mitigated compared to the fully-rational alternative. Therefore, among other implications, the behavioral model with cognitive discounting can easily avoid the forward guidance puzzle.¹

With cognitive discounting, Gabaix (2020) adds a behavioral element to the model, but otherwise retains the assumption of rational expectations. Those rational expectations about future variables are simply shrunk toward steady states more quickly for longer horizons. However, a behavioral alternative to rational expectations has been considered for a long time in the macroeconomics literature and is provided by the literature on adaptive learning (e.g., Marcat and Sargent, 1989, Sargent, 1993, 1999, Evans and Honkapohja, 2001, 2013).

¹Other mechanisms have been proposed in the literature to deal with the forward guidance puzzle: finite planning horizons (Woodford, 2018), imperfect common knowledge (Angeletos and Lian, 2018), finite lives (Del Negro et al., 2015), preferences with temptation and self-control (Airaud, 2020), and others. In many cases, they lead to similar loglinearized equations for the New Keynesian model. Meggiorini (2020) provides an empirical comparison of different approaches. Here, we focus on cognitive discounting as a mechanism that captures insights from psychological research (Gabaix, 2019) and that can be simply adapted to models with other types of expectation formation.

The models with cognitive discounting and with expectations based on learning represent two different behavioral models. In the first case, as we have discussed, expectations remain based on the rational expectations assumption, but they are more heavily discounted when they refer to longer-term horizons. Under adaptive learning, instead, expectations are formed from agents' perceived models of the economy. Often, but not always, the perceived models share the same variables that would exist in the model solution under rational expectations (that is, agents' models are correctly specified, since they include every relevant variable that would matter under rational expectations). But the agents' knowledge and degree of understanding about the economy are limited: they cannot know the values of some aggregate parameters, such as the degree of price stickiness in the economy, or the monetary policy reaction coefficients. As a result, they cannot infer the coefficients related to the reduced-form dynamics of the economy, as they would under rational expectations; for example, they cannot know how sensitive output is to interest rates, how steep the Phillips curve is, and how persistent output and inflation fluctuations are likely to be. However, they form beliefs about these interrelationships, and they update their beliefs over time based on the past time series data that they observe.

On one hand, agents' degree of understanding appears more limited under learning than under cognitive discounting: they lack full knowledge about the magnitudes of dynamic interactions among macroeconomic variables. Also, in their perceived models, agents are learning about intercept coefficients, hence they are allowed to make mistakes in inferring the steady state of the variables. On the other hand, in the cognitive discounting model, agents have a biased model of the economy (as a result of the extra discounting), and the bias is permanent: they are prevented from learning the truth even in the long run. Under learning, instead, agents have imperfect beliefs during transition phases, but they may converge to the rational expectations equilibrium (or to a distribution around it) in the long run.

The two behavioral elements can, in principle, coexist. Alternatively, it is conceivable that myopia may become redundant when rational expectations are replaced by boundedly-rational expectations with learning.

This paper studies a Behavioral New Keynesian model that incorporates cognitive discounting and alternative expectational assumptions. We examine the model under rational expectations, and two versions of subjective expectations with adaptive learning: infinite-horizon learning and Euler-equation learning. Under infinite-horizon learning, current macroeconomic variables are influenced by expectations far into the future. Under Euler-equation learning, instead, the New Keynesian model resembles the one under rational expectations, with output and inflation expressed as a function of one-period-ahead expectations.

The model is then estimated using a full-information Bayesian approach to evaluate the importance of cognitive discounting in models with rational expectations and, particularly, with learning

replacing rational expectations. In that case, the learning and cognitive discounting parameters are all jointly estimated, along with the remaining structural parameters.

Main Findings. The empirical results confirm the need for substantial degrees of cognitive discounting in the benchmark framework with rational expectations. Under learning with infinite-horizon expectations, the data still indicate that large, perhaps unrealistically so, degrees of myopia are needed. Under infinite-horizon learning, the model equations become, in fact, extremely forward-looking, with long-horizon forecasts about future output gaps, inflation, and interest rates, until the indefinite future mattering for current realizations. Such a framework fits the data as well as rational expectations, but it similarly needs a large enough degree of myopia as a counterbalance to the agents' extreme farsightedness. Under Euler-equation learning, on the other hand, the empirical importance of myopia is substantially attenuated. The estimates indicate values of the cognitive discounting factor rising much closer to one.

We compare the fit of the model under different expectation formation assumptions. The model with limited cognitive discounting and subjective expectations formed according to an Euler-equation learning framework fits macroeconomic data decisively better than the alternatives. The outperformance persists in the full sample analysis and in both pre-1979 and post-1982 subsamples. Rational expectations and infinite-horizon learning specifications with substantial myopia, instead, achieve a similar lower fit.

Under rational expectations, the data prefer a specification in which the impact of forward guidance is severely altered: the effects of monetary policy are stronger at shorter anticipation horizons, and almost nil at horizons longer than 2-3 years. But the model's effort to attenuate the otherwise implausible effects of forward guidance comes at a cost. The estimated myopia substantially reduces the effectiveness of conventional, unanticipated, monetary policy shocks. Only the model with Euler-equation learning can match closely the impulse response of the output gap to monetary policy shocks from a VAR; the responses are underestimated under rational expectations and infinite-horizon learning.

Contribution to the Literature. Our paper aims to contribute to the growing literature that can be broadly referred to as behavioral macroeconomics. Various papers have proposed more or less behavioral adjustments to the benchmark New Keynesian model, often with the goal to increase discounting in order to attenuate the forward guidance puzzle. Del Negro et al. (2015) are the first to document the puzzle and show that a perpetual-youth specification can dampen the effects of forward guidance. Different behavioral features can be added into the model to induce stronger discounting of future variables: besides Gabaix (2020), who introduces microfoundations for cognitive discounting, Woodford (2018) assumes that agents optimize only with finite planning horizons, but are boundedly-rational beyond such horizons, Angeletos and Lian (2018) model imperfect coordination and higher-order uncertainty, Airaudo (2020) inserts temptation in the utility

function, and Garcia-Schmidt and Woodford (2019) and Farhi and Werning (2019) introduce level- k thinking as a mechanism to dampen future fluctuations.² While many of these approaches lead to comparable loglinearized equations for the benchmark New Keynesian model, Gabaix's behavioral modification is possibly easier to generalize to alternative models and expectational assumptions.

In this area, our paper aims to provide further empirical evidence on the importance of cognitive discounting. Behavioral New Keynesian models with cognitive discounting are estimated in Ilabaca, Meggiorini, and Milani (2020), Meggiorini (2020), and Andrade, Cordeiro and Lambais (2019). This paper shows that the importance of the newly added behavioral feature is sensitive to the modeling of expectations. The evidence for myopia may be only minimal when rational expectations are replaced by its leading alternative, subjective expectations with adaptive learning.

At a broader level, our work contributes to the research denoted in Woodford (2013) as macroeconomics without the rational expectations hypothesis. Cole and Milani (2019) pinpoint the modeling of expectations as a key source of misspecification in New Keynesian models. The assumption of rational expectations is responsible for the empirical failure of the model to capture the dynamic interrelationships among macroeconomic variables, monetary policy decisions, and observed expectations. More backward-looking expectations, instead, help reconcile the model with the data. The evidence from experimental work (e.g., Hommes, 2011, 2021) is also supportive of similarly adaptive behavior in the formation of macroeconomic expectations.

More specifically, our paper replaces rational expectations with adaptive learning. In this respect, the paper adds to a long literature that considers learning as a natural refinement for rational expectations (DeCanio, 1979, Bray and Savin, 1986, Sargent, 1993, 1998, Evans and Honkapohja, 2001). Some papers propose learning as a single behavioral feature that can improve the model's ability to fit time series data (Milani, 2007, Slobodyan and Wouters, 2012). Learning helps generating endogenous persistence in the model (Orphanides and Williams, 2003, Milani, 2007, 2017), it leads to business cycle amplification (Eusepi and Preston, 2011, Milani, 2011), endogenous time-varying volatility (Branch and Evans, 2007, Milani, 2014), and it accounts for changing responses to shocks over time (e.g., Milani, 2009). Cole (2020, 2021) highlights private sector's learning as a mechanism that can lessen the impact of forward guidance policies.

Here, we merge new (cognitive discounting) and old (boundedly-rational expectations and adaptive learning) behavioral elements in the New Keynesian model to show that learning is still an important feature that needs to be accounted for to explain macroeconomic observations.

²McKay, Nakamura and Steinsson (2016), instead, assume limited risk sharing and incomplete markets to introduce additional discounting. These modifications yield a precautionary saving motive that reduces the response of consumption to future interest rates.

2. BEHAVIORAL NEW KEYNESIAN MODELS

We assume that the economy is summarized by a behavioral version of the New Keynesian model, as derived in Gabaix (2020). The behavioral feature that is added into the model is *cognitive discounting*: when thinking about economic developments far into the future, agents shrink their perceptions toward steady-state values of the variables. Deviations around the steady state that are expected to happen k periods in the future are more heavily discounted by a factor \bar{m}^k , with $\bar{m} \in [0, 1]$ measuring the strength of cognitive discounting. The additional discounting captures agents' partial myopia toward the future, or a 'term structure' of inattention, with agents being more inattentive with respect to more distant events.³ Gabaix (2020) also outlines possible deeper microfoundations for cognitive discounting based on iterated mental simulations under noisy signals.

Let's assume that the state vector of the economy S_t , once demeaned and loglinearized, evolves as:

$$S_{t+1} = FS_t + \epsilon_{t+1} \quad (1)$$

where F is a matrix of coefficients and ϵ_{t+1} denotes a vector of mean-zero innovations. Behavioral agents perceive instead

$$S_{t+1} = \bar{m} (FS_t + \epsilon_{t+1}) \quad (2)$$

with \bar{m}^k representing the previous cognitive discounting parameter. Behavioral agents thus form expectations $E_t^{BR} [S_{t+k}] = \bar{m}^k E_t [S_{t+k}] = \bar{m}^k F^k S_{t+k}$, where the operators E_t^{BR} and E_t denote behavioral and rational expectations, respectively. Perceived fluctuations in the future are, therefore, discounted by a factor \bar{m}^k .

We present the behavioral model under alternative expectational formation assumptions: rational expectations, subjective expectations with infinite-horizon learning, and subjective expectations with Euler-equation (i.e., one-period-ahead) learning.

2.1. Rational Expectations. When agents are subject to cognitive discounting, but still form expectations about macroeconomic variables according to the rational expectations hypothesis as in Gabaix (2020), the model is represented by the following loglinearized equations:

$$x_t = ME_t x_{t+1} - \sigma(i_t - E_t \pi_{t+1} - r_t^n) \quad (3)$$

$$\pi_t = \beta M^f E_t \pi_{t+1} + \kappa x_t + u_t \quad (4)$$

$$i_t = \rho i_{t-1} + (1 - \rho) [\chi_\pi \pi_{t-1} + \chi_x x_{t-1}] + \varepsilon_t. \quad (5)$$

³Therefore, we will refer to this behavioral feature in the paper interchangeably as cognitive discounting, or myopia, or inattention.

where x_t denotes the output gap, π_t inflation, i_t the nominal interest rate, r_t^n and u_t are demand (natural-rate) and supply (cost-push) disturbances, which evolve as AR(1) processes as $r_t^n = \rho_r r_{t-1}^n + \varepsilon_t^r$ and $u_t = \rho_u u_{t-1} + \varepsilon_t^u$, and ε_t is an *i.i.d.* monetary policy shock. The parameter σ measures the elasticity of intertemporal substitution, β denotes the household discount factor, $\kappa \equiv \frac{(1-\alpha\beta)(1-\beta)}{\alpha} \frac{\omega+\sigma^{-1}}{1+\omega\theta}$ denotes the slope of the Phillips curve, which inversely depends on the Calvo price stickiness parameter α ,⁴ and ρ , χ_π , χ_x , are monetary policy coefficients. Equation (3) represents a loglinearized Euler equation, equation (4) is a New Keynesian Phillips curve, and (5) is a Taylor rule, which characterizes monetary policy in the economy.⁵ The model equations remain similar to those in the canonical New Keynesian model, assuming that conventional expectations are replaced by behavioral expectations $E_t^{BR} = M^k E_t$. In this case, E_t still denotes rational expectations, which imply that the mathematical conditional expectation is based on the true model of the economy. Agents know the model, the model parameters, and the stochastic properties of exogenous disturbances, and condition their expectations on this information. The behavioral addition, however, is that agents are subject to cognitive discounting and, therefore, they discount future output gaps and inflation (in deviation from their steady states) at higher rates. The parameters $M, M^f \in [0, 1]$ are given by $M = \bar{m}$ and $M^f = \bar{m} \left(\alpha + \frac{1-\beta\alpha}{1-\beta\alpha\bar{m}} (1 - \bar{m}) \right)$, which depend on \bar{m} , the main behavioral coefficient of interest. Forcing the coefficient \bar{m} to equal one would imply $M = M^f = 1$ and would yield the standard formulation of the New Keynesian model.

Under rational expectations, the model can be written as

$$\Gamma_0 \xi_t = \Gamma_1 \xi_{t-1} + \Psi \nu_t + \Pi \eta_t \quad (6)$$

where $\xi_t = [x_t, \pi_t, i_t, r_t^n, u_t, E_t x_{t+1}, E_t \pi_{t+1}]'$ is a vector including the model's endogenous variables, the structural disturbances, and the expectation terms, $\nu_t = [\varepsilon_t^r, \varepsilon_t^u, \varepsilon_t]$ is a vector of exogenous demand, supply, and monetary policy, innovations, and $\eta_t = [\eta_t^x, \eta_t^\pi]'$ collects the expectational errors, which are defined as $\eta_t^x = x_t - E_{t-1} x_t$ and $\eta_t^\pi = \pi_t - E_{t-1} \pi_t$, such that $E_t \eta_{t+1}^x = E_t \eta_{t+1}^\pi = 0$, $\forall t$. If the equilibrium exists and is unique, a unique mapping can be found between expectational errors and structural innovations. The law of motion for the economy is then given by the model solution under rational expectations (e.g., Sims, 2002):

$$\xi_t = F \xi_{t-1} + G \nu_t. \quad (7)$$

⁴The remaining parameters ω and θ denote the inverse of the Frisch elasticity of labor supply and the elasticity of substitution across differentiated goods.

⁵The monetary policy rule is 'operational' in the sense of McCallum (1999), i.e., it is assumed to respond to variables dated in (t-1) that the central bank can observe in real-time.

2.2. Subjective Expectations & Infinite-Horizon Learning. We now relax the assumption of rational expectations and allow agents to form subjective, boundedly-rational, expectations. The model can be re-derived from its primitives under the assumption of subjective expectations, as shown in Preston (2005). Households' decisions are, therefore, optimal conditional on their beliefs and satisfy at every point in time their intertemporal budget constraint.

The economy, under subjective expectations, is summarized by the following equations:

$$x_t = \widehat{E}_t \sum_{T=t}^{\infty} (\beta M)^{T-t} [(1 - \beta) x_{T+1} - \sigma(i_T - \pi_{T+1} - r_T^n)] \quad (8)$$

$$\pi_t = \widehat{E}_t \sum_{T=t}^{\infty} (\alpha \beta M^f)^{T-t} [\kappa x_T + (1 - \alpha) \beta \pi_{T+1} + u_T] \quad (9)$$

$$i_t = \rho i_{t-1} + (1 - \rho) [\chi_\pi \pi_{t-1} + \chi_x x_{t-1}] + \varepsilon_t. \quad (10)$$

Preston (2005) shows that, with more general (subjective) expectational assumptions, long-horizon expectations explicitly enter the model. Contemporaneous values of the output gap and of the economy's inflation rate are, therefore, influenced by expectations until the entire infinite future. Models that follow the same infinite-horizon approach have also been adopted in Preston (2006, 2008), Eusepi and Preston (2010), Milani (2006), among others.

Conventional expectations are again replaced using the behavioral expectation operator $\widehat{E}_t^{BR} Y_{t+k} = M^k \widehat{E}_t Y_{t+k}$. But now the formation of expectations allows for subjective, not necessarily rational, expectations, denoted by \widehat{E}_t . The model hence includes two behavioral elements: infinite-horizon learning and cognitive discounting.

Expectations \widehat{E}_t about future macroeconomic variables are formed by agents according to their perceived law of motion (PLM)

$$Y_t = \phi_{0,t} + \phi_{1,t} Y_{t-1} + \varepsilon_t \quad (11)$$

where $Y_t = [x_t, \pi_t, i_t, r_t^n, u_t]'$ and ε_t denotes the error term in the regressions. The perceived model is assumed to be correctly-specified, as it includes the same variables that would appear in the Minimum State Variable (MSV) solution under rational expectations. Agents' beliefs regarding intercept terms are collected in the vector $\phi_{0,t}$, and those regarding the interrelationships among

variables are in the matrix $\phi_{1,t}$:⁶

$$\phi_{0,t} = \begin{bmatrix} \phi_{0,t}^x \\ \phi_{0,t}^\pi \\ \phi_{0,t}^i \\ 0 \\ 0 \end{bmatrix}, \quad \phi_{1,t} = \begin{bmatrix} \phi_{1,t}^{xx} & \phi_{1,t}^{x\pi} & \phi_{1,t}^{xi} & \phi_{1,t}^{xr} & \phi_{1,t}^{xu} \\ \phi_{1,t}^{\pi x} & \phi_{1,t}^{\pi\pi} & \phi_{1,t}^{\pi i} & \phi_{1,t}^{\pi r} & \phi_{1,t}^{\pi u} \\ \phi_{1,t}^{ix} & \phi_{1,t}^{i\pi} & \phi_{1,t}^{ii} & \phi_{1,t}^{ir} & \phi_{1,t}^{iu} \\ 0 & 0 & 0 & \rho_r & 0 \\ 0 & 0 & 0 & 0 & \rho_u \end{bmatrix},$$

The beliefs, collected in $\widehat{\phi}_t = (\phi_{0,t}, \phi_{1,t})'$, evolve according to a constant-gain updating algorithm as

$$\widehat{\phi}_t = \widehat{\phi}_{t-1} + \bar{\mathbf{g}} R_t^{-1} X_t (Y_t - X_t' \widehat{\phi}_{t-1})' \quad (12)$$

$$R_t = R_{t-1} + \bar{\mathbf{g}} (X_t X_t' - R_{t-1}), \quad (13)$$

where $X_t \equiv [1, Y_{t-1}]'$ and where the second expression characterizes the updating of the coefficients' precision matrix R_t . The parameter $\bar{\mathbf{g}}$ denotes the constant-gain coefficient and it measures a different type of possible cognitive bias: "recency bias". With a constant gain, agents weigh recent observations more heavily than those in the more distant past; recency bias is more prevalent the higher the value of $\bar{\mathbf{g}}$. An alternative interpretation is also possible: learning with a constant gain offers agents an updating algorithm that is successful in tracking systems with parameters that may vary at unknown times. In such case, a higher gain would indicate agents who are more worried about structural breaks and, therefore, more responsive to incoming information.

Expectations are formed at each horizon T , based on (11) and with beliefs updated according to (12)-(13), to yield

$$\widehat{E}_{t-1} Y_{T+1} = (I - \phi_{1,t-1})^{-1} (I - \phi_{1,t-1}^{T-t+2}) \phi_{0,t-1} + \phi_{1,t-1}^{T-t+2} Y_{t-1}. \quad (14)$$

Agents form expectations based on an information set up to $t-1$, which is assumed to include the exogenous disturbances. Expectations formed for each horizon T as in (14) are substituted into the original system (8)-(10), to derive the Actual Law of Motion of the economy, or ALM.

⁶It is typical in these models to assume that agents know the values of ρ_r and ρ_u : this is irrelevant for the results, since agents would learn very quickly the parameters for exogenous processes.

The infinite sums, which we denote by $\Sigma_{\beta M}$ and $\Sigma_{\alpha\beta M}$, become:

$$\begin{aligned}\Sigma_{\beta M} &= \widehat{E}_{t-1} \sum_{T=t}^{\infty} (\beta M)^{T-t} Y_{T+1} = (I - \phi_{1,t-1})^{-1} \left[(1 - \beta M)^{-1} I - \phi_{1,t-1}^2 (I - \beta M \phi_{1,t-1})^{-1} \right] \phi_{0,t-1} + \\ &\quad + (I - \beta M \phi_{1,t-1})^{-1} \phi_{1,t-1}^2 Y_{t-1} \\ &= B_0(\phi_{0,t-1}, \phi_{1,t-1}) + B_1(\phi_{0,t-1}, \phi_{1,t-1}) Y_{t-1}\end{aligned}\tag{15}$$

$$\begin{aligned}\Sigma_{\alpha\beta M} &= \widehat{E}_t \sum_{T=t}^{\infty} (\alpha\beta M^f)^{T-t} Y_{T+1} = (I - \phi_{1,t-1})^{-1} \left[(1 - \alpha\beta M^f)^{-1} I - \phi_{1,t-1}^2 (I - \alpha\beta M^f \phi_{1,t-1})^{-1} \right] \phi_{0,t-1} + \\ &\quad + (I - \alpha\beta M^f \phi_{1,t-1})^{-1} \phi_{1,t-1}^2 Y_{t-1} \\ &= C_0(\phi_{0,t-1}, \phi_{1,t-1}) + C_1(\phi_{0,t-1}, \phi_{1,t-1}) Y_{t-1}.\end{aligned}\tag{16}$$

The system can then be written as

$$A_0 Y_t = A_1 Y_{t-1} + A_2 \Sigma_{\beta M} + A_3 \Sigma_{\alpha\beta M} + A_4 \nu_t\tag{17}$$

$$\begin{aligned}&= A_1 Y_{t-1} + A_2 (B_0(\phi_{0,t-1}, \phi_{1,t-1}) + B_1(\phi_{0,t-1}, \phi_{1,t-1}) Y_{t-1}) + \\ &\quad A_3 (C_0(\phi_{0,t-1}, \phi_{1,t-1}) + C_1(\phi_{0,t-1}, \phi_{1,t-1}) Y_{t-1}) + A_4 \nu_t\end{aligned}\tag{18}$$

The ALM, more compactly, is given by

$$Y_t = A(\widehat{\phi}_{t-1}, \Theta, \bar{m}) + F(\widehat{\phi}_{t-1}, \Theta, \bar{m}) Y_{t-1} + G(\Theta, \bar{m}) \nu_t\tag{19}$$

where $A(\cdot) = A_0^{-1} [A_2 B_0 + A_3 C_0]$, $F(\cdot) = A_0^{-1} [A_1 + A_2 B_1 + A_3 C_1]$, and $G = A_0^{-1} A_4$. The expression makes clear that the vector and matrices of parameters in the ALM depend on the state of beliefs $\widehat{\phi}_{t-1}$, on the structural parameters of the New Keynesian model collected in Θ , and on the degree of myopia or cognitive discounting \bar{m} .

2.3. Subjective Expectations & Euler-Equation Learning. Under some conditions, derived and discussed in Honkapohia, Mitra, and Evans (2012),⁷ the equations in (8)-(10), can be simplified to arrive at a system where only one-period-ahead expectations matter. Agents' behavior is based on the Euler equation, but obtained under subjective expectations. The same approach, which

⁷In the Euler-equation approach, agents need to forecast their consumption next period. In the typical New Keynesian economy with market clearing condition $c_t = y_t$, agents, indexed by i , would need to recognize that $c_t^i = y_t$, $\forall i, t$. Agents should be able to observe this condition from past data or to quickly learn that the equality always holds true. As a result, they would understand and be able to forecast that next-period consumption will equal next-period output. This condition, in addition to the law of iterated expectations holding at the individual level, allows the microfoundations under subjective expectations to produce the same system that is obtained under rational expectations.

can be denoted as Euler-equation learning, is used in the majority of works that assume adaptive learning (e.g., Bullard and Mitra, 2002, Evans and Honkapohja, 2001, 2003, Sargent, 1993, 1999, among countless others).

The model, therefore, shares the same loglinearized equations as the rational expectations version, only with a different expectation operator.

$$x_t = M\widehat{E}_t x_{t+1} - \sigma(i_t - \widehat{E}_t \pi_{t+1} - r_t^n) \quad (20)$$

$$\pi_t = \beta M^f \widehat{E}_t \pi_{t+1} + \kappa x_t + u_t \quad (21)$$

$$i_t = \rho i_{t-1} + (1 - \rho) [\chi_\pi \pi_{t-1} + \chi_x x_{t-1}] + \varepsilon_t. \quad (22)$$

Expectations, denoted by \widehat{E}_t , are formed according to the same PLM given by expression (11), and with beliefs updated over time according to the constant-gain learning algorithm in (12)-(13), as they were under infinite-horizon learning. Expectations regarding variables at $(t+1)$ are formed as

$$\begin{aligned} \widehat{E}_{t-1} Y_{t+1} &= (I + \phi_{1,t-1}) \phi_{0,t-1} + \phi_{1,t-1}^2 Y_{t-1} \\ &= B_0(\phi_{0,t-1}, \phi_{1,t-1}) + B_1(\phi_{0,t-1}, \phi_{1,t-1}) Y_{t-1}. \end{aligned} \quad (23)$$

Such expectations can be substituted into the original system (20)-(22), which can be written as

$$A_0 Y_t = A_1 Y_{t-1} + A_2 \widehat{E}_{t-1} Y_{t+1} + A_3 \nu_t, \quad (24)$$

to obtain the ALM:

$$Y_t = A(\widehat{\phi}_{t-1}, \Theta, \bar{m}) + F(\widehat{\phi}_{t-1}, \Theta, \bar{m}) Y_{t-1} + G(\Theta, \bar{m}) \nu_t \quad (25)$$

where $A(\cdot) = A_0^{-1}(A_2 B_0)$, $F(\cdot) = A_0^{-1}(A_1 + A_2 B_1)$, and $G = A_0^{-1} A_3$. As before, the model incorporates two main behavioral features: subjective expectations from a constant-gain learning model and cognitive discounting.

3. STRUCTURAL BAYESIAN ESTIMATION

We estimate the vector of structural and behavioral parameters jointly. We use quarterly data on the output gap, inflation, and the Federal Funds rate, as observables for x_t , π_t , and i_t . The output gap is calculated as the log-difference between Real GDP and Real Potential GDP, as estimated by the Congressional Budget Office. Inflation is the log-first difference of the Implicit GDP Price Deflator. The short-term nominal interest rate is given by the Federal Funds Rate, transformed from yearly into the corresponding quarterly rate. All series are obtained from FRED, the Federal Reserve Database, maintained by the Federal Reserve Bank of St. Louis.

The estimation sample spans the period from 1960:Q1 to 2007:Q2.⁸ As a robustness check, we will later present the results obtained when splitting the sample in pre- and post-Volcker periods.

Table 1 summarizes our choices of priors, which are conventional in the literature. The elasticity of intertemporal substitution σ is assumed to follow a prior Gamma distribution with mean 1.5 and standard deviation 0.75. The Calvo price-setting parameter has a Beta distribution prior that implies prices that are reset, on average, every three quarters, in line with microeconomic evidence. The policy reaction coefficients to inflation and the output gap are assumed to have Normal priors with mean 1.5 and standard deviation 0.25, and mean 0.1 and standard deviation 0.05, respectively. For the cognitive discounting parameter \bar{m} , we follow previous papers in selecting a Beta prior with mean 0.8 and standard deviation 0.15; later, we will consider a Uniform prior and a Gamma prior that weighs against discounting by favoring a value of 1, for robustness. We assume a Beta prior also for the constant gain coefficient, with a 0.025 mean and 0.01 standard deviation, as in Milani (2017). Finally, we choose Beta priors for the remaining autoregressive coefficients and inverse Gamma priors for the standard deviations of the exogenous innovations.⁹

The structural model is estimated using a full-information Bayesian approach. Draws are generated using a Metropolis-Hastings algorithm; we run a total number of 1,500,000 MCMC draws for each specification. We discard the initial 40% of draws as burn-in and thin the chain by saving only one every hundred draws to reduce the autocorrelation of the draws.

4. RESULTS

The posterior estimates for the full-sample estimation are shown in Table 1. Under rational expectations, substantial levels of inattention are needed to fit the data. The posterior mean estimate for the cognitive discounting parameter \bar{m} is equal to 0.416, with the entire 95% posterior density interval falling well below one. The data, therefore, require substantial discounting of future expected variables, in contrast with the implications of benchmark New Keynesian models that yield discounting factors equal to one for output and to β for inflation.

With subjective expectations and under infinite-horizon learning, contemporaneous output gaps, inflation, and interest rates, are expressed as a function of long-horizon expectations until the indefinite future. As a consequence of such long horizons, the model still requires extreme degrees

⁸In the estimation, we use data from 1954:Q3, the first quarter of availability of Federal Funds rate data, but we start computing the likelihood function only from 1960:Q1. Observations before 1960 are simply used for initialization of the learning process. We conclude the estimation before the beginning of the 2007-2009 Great Recession, to avoid having to deal with the binding effective lower bound constraint in policy rates.

⁹The only parameters that are fixed in the estimation are the discount factor $\beta = 0.99$, the inverse of the Frisch elasticity $\omega = 2$, and the elasticity of substitution across differentiated goods $\theta = 11$, which implies a steady-state markup of prices over marginal costs equal to 10%.

of myopia to fit the data. The posterior mean for \bar{m} falls around 0.2, which indicates even stronger discounting than under rational expectations.

The importance of cognitive discounting, however, is considerably reduced when rational expectations are replaced by subjective expectations formed from a learning model with the Euler-equation approach. In this case, the model specification resembles the one under rational expectations: only one-period-ahead expectations enter the Euler equation and the New Keynesian Phillips curve. Under this type of learning, the model can fit macroeconomic data without resorting to extreme degrees of inattention. The estimate for \bar{m} rises much closer to 1, with a posterior mean equal to 0.9366.

Therefore, while the Behavioral New Keynesian model that has been proposed in the literature retains the assumption of rational expectations, we show that the addition of another behavioral feature, such as deviations from rational expectations and learning by the agents, can substantially reduce the empirical need for cognitive discounting.

The specification with expectations formed according to the Euler-equation learning model also provides the best overall fit of the data. We calculate the models' log marginal likelihoods (shown in the last row of Table 1). Rational expectations and infinite-horizon learning achieve similar fit, with the former slightly outperforming. But Euler-equation learning leads to sizable improvements in model fit: the resulting Bayes factor would suggest 'decisive' evidence in favor of Euler-equation learning versus the alternatives, based on Jeffreys' (1961) widely-used interpretative scale.

Cognitive discounting, however, is not redundant. When we re-estimate the model with Euler-equation learning and fix \bar{m} at 1, the resulting marginal likelihood falls to -270.4209 (the full set of posterior estimates for this case are shown in Table 4). The specification with both Euler-equation learning and (moderate) myopia is, therefore, preferred by the data.

Turning to the remaining estimated parameters, the posterior results show that the monetary policy estimates remain similar across model specifications. The degree of price stickiness also remains in the same range, around 0.8, while the estimate for σ is lower under rational expectations. Another major difference relates to the properties of structural disturbances: whereas rational expectations and infinite-horizon learning require large autocorrelation coefficients in the natural rate and cost-push disturbance processes, the required persistence is reduced under Euler-equation learning. The estimates for ρ_r and ρ_u fall from values around 0.9 under rational expectations to 0.2278 and 0.0654 under Euler-Equation learning. As already discussed in Milani (2007, 2017), learning helps the models match the observed persistence in macroeconomic data.

The estimates for the constant-gain coefficients differ between the two types of learning models: the discounting of observations in the more distant past is considerably larger under infinite-horizon learning (as implied by a posterior mean for the gain equal to 0.0454) than under Euler-equation learning (with a lower estimated gain equal to 0.0049).

4.1. Expectations and the Effects of Monetary Policy. Under rational expectations, forward guidance announcements have implausibly large effects on the economy (Del Negro et al., 2015). In Figure 1, we show the responses of inflation to an anticipated monetary policy shock at horizons 0 to 80 quarters. The response in the left panel mirrors those identified in McKay et al. (2016) and Gabaix (2020): in the benchmark New Keynesian model with $\bar{m} = 1$, the impact of forward guidance is magnified for longer anticipation horizons, making the model potentially unstable. In the right panel, we repeat the same exercise, but now assuming the same degree of cognitive discounting that has been empirically estimated: \bar{m} is fixed at its posterior mean of 0.416. The response of inflation to forward guidance is drastically different, with shocks at the lowest horizons now having the largest effects. Anticipations with horizons above two years are largely ineffective, given that agents strongly discount them.

Therefore, to fit the data, the rational expectations specification depends upon a large enough degree of myopia, which is needed to shut down the effects of anticipations far into the future. But the addition of severely myopic agents comes at a cost: the implied responses to conventional, unanticipated, monetary policy shocks are also attenuated.

Figure 2 shows the impulse responses of the output gap to unanticipated, expansionary, monetary policy innovations, under the different expectation assumptions, and compared with the response obtained from a VAR model estimated on the same three endogenous variables (output gap, inflation, Federal funds rate). The model with Euler-equation learning generates output responses that track very closely the responses from the VAR. The specifications with rational expectations and infinite-horizon expectations, instead, in consequence of the sizable degree of myopia that they require to dampen the effects of anticipations, end up severely underestimating the effects of monetary policy.

5. ROBUSTNESS

i. Split samples. Several papers have documented time variation in the coefficients related to monetary policy and in the standard deviations of exogenous shocks before and after the early 1980s. The degree of inattention may also vary over time, and, in particular, assume very different values in samples characterized by high inflation and macroeconomic instability, such as the 1970s, compared with low-inflation low-volatility samples, as the post-1980 decades.

Therefore, to investigate the robustness of our main findings, we repeat the estimation now focusing on two sub-samples: 1954:Q3-1979:Q2 and 1982:Q1-2007:Q2. The results are reported in Table 2.

Under rational expectations, the estimated degrees of cognitive discounting shift from 0.63 to 0.409 over the samples. As before, the infinite-horizon learning case requires the largest discounting, and this remains true in both sub-samples. In the pre-1979 period, the Euler-equation learning

results suggest values of myopia almost identical to the full sample estimates. The estimated coefficient, however, is lower in the post-1982 sample, with a posterior mean estimate equal to 0.79. Inattention is thus more prevalent in periods characterized by overall macroeconomic stability, as could be expected. We need to remark, however, that for post-1982 data, the posterior distribution for \bar{m} displays a clear bimodality: the posterior mode is obtained at a value of 0.9112, suggesting more limited inattention not far off its full-sample estimate, but there is also a lower-probability mode that concentrates around lower values.

One issue with the estimation for pre-Volcker samples is the possibility of equilibrium indeterminacy. The results in Ilabaca, Meggiorini, and Milani (2020), however, show that, in a model with cognitive discounting and rational expectations, indeterminacy is rejected over each sample (the paper estimates the model allowing for both indeterminacy and determinacy). The data strongly prefer a specification where monetary policy in the 1960-70s is indeed passive, but where bounded rationality is pervasive enough to induce determinacy.¹⁰ Our estimates are similar here: they indicate a monetary policy that responds with coefficients below one toward inflation before Volcker, but agents' inattention is strong enough to restore determinacy.

The model comparison exercise paints a similar story as in the previous section: the fit of the model under rational expectations and infinite-horizon learning are comparable in the pre-1979 samples, whereas infinite-horizon learning is preferred after 1982. But, in each sample, the data are clearly best explained by the model with Euler-equation learning and only limited degrees of cognitive discounting.

ii. Taylor Rule with Anticipated Shocks. We estimate the model now allowing for explicit forward guidance by the central bank, modeled through the addition of anticipated, or 'news', shocks in the Taylor rule. The policy rule becomes

$$i_t = \rho i_{t-1} + (1 - \rho) [\chi_\pi \pi_{t-1} + \chi_x x_{t-1}] + \varepsilon_t + \varepsilon_{t-4}^4 + \varepsilon_{t-8}^8, \quad (26)$$

where ε_{t-4}^4 and ε_{t-8}^8 denote monetary shocks that are anticipated four and eight quarters in advance. We assume Gamma(0.3,0.3) priors for the standard deviations of the news shocks.¹¹

Table 3 reports the estimates. The posterior mean values for \bar{m} remain similar to those estimated in the model with only unanticipated monetary policy shocks, and the Euler-equation learning specification still outperforms the alternatives.

¹⁰In the model, and as shown in Gabaix (2020), the modified Taylor principle is given by $\chi_\pi + \frac{1-\beta M^f}{\kappa} \chi_x + \frac{(1-\beta M^f)(1-M)}{\kappa\sigma} > 1$. Under learning, instead, the estimation doesn't pose particular problems and can be performed whether the equilibrium is E-stable or E-unstable (e.g., Milani, 2008).

¹¹The prior assigns progressively lower probability to larger standard deviations. The prior choice follows Schmitt-Grohé and Uribe (2012) and Milani and Rajbhandari (2020).

iii. Uninformative priors for \bar{m} . In the benchmark estimation, we assumed an informative Beta(0.8,0.15) prior for the cognitive discounting coefficient. In the Euler-equation learning specification, we find that a limited degree of cognitive discounting still improves the model fit compared with the benchmark case, based on the marginal likelihoods. But the model comparison conclusions may be affected by the choice of an informative prior that favors a positive, albeit modest, myopia coefficient.

Therefore, we evaluate the sensitivity of the results to the use of two alternative priors: an uninformative Uniform[0,1] prior for \bar{m} and a Gamma(0.3,0.3) prior for $1 - \bar{m}$. A Gamma prior with equal values for mean and standard deviation assigns highest probability to a value of zero, and progressively lower probability to larger values. The prior is, therefore, set for $1 - \bar{m}$, to penalize more sizable deviations from the benchmark model with no inattention.

The results, presented in Table 4, are again almost identical. With a Uniform prior, the posterior mean estimate for \bar{m} is equal to 0.9382; the log marginal Likelihood is -268.7747, which still outperforms the case with fixed $\bar{m} = 1$ (for which the log marginal Likelihood equals -270.4209). With the Gamma prior, the estimate for \bar{m} is slightly higher at 0.9437, and the log marginal likelihood equals -267.9884. Therefore, independently of the prior, the data always indicate modest degrees of discounting, and favor a specification where learning and modest cognitive discounting coexist.

iv. Different gains for different variables. While it is common to assume a single constant-gain coefficient in models with adaptive learning, it is possible that agents adopt different gains for different variables. If agents perceive different probabilities of structural breaks in the dynamics of different variables, it would be natural that they would weigh incoming information differently. Therefore, we re-estimate the model, but now allowing for different constant gain coefficients for the output gap, inflation, and interest rate perceived laws of motion. The posterior estimates indicate that agents perceive more instability in the inflation equation and less in the output gap. The posterior means for the gains are 0.0059 for output, 0.0370 for inflation, and 0.0198 for the interest rate. Nonetheless, the evidence on inattention remains similar, with \bar{m} equal to 0.88, slightly lower than in the benchmark case. The marginal likelihood is -266.2732, suggesting an improved fit compared to the case of a single gain.

It is conceivable that agents similarly display different degrees of myopia when dealing with different variables. While this is theoretically plausible, the corresponding parameters will not be separately identifiable in our framework. Hence, in the paper we focus only on a general degree of myopia, rather than a variable-specific, or agent-specific, level.

6. CONCLUSIONS

The adaptive learning literature has long provided an alternative approach to modeling expectations in macroeconomic models. Compared with the rational expectations hypothesis, adaptive learning allows researchers to introduce behavioral elements into their models: economic agents can form subjective perceptions about economic relationships, they can update their beliefs based on past events, and, in doing so, they may display cognitive biases, such as the “recency bias” captured by constant-gain learning.

Recently, a different literature, which in large part arose from efforts to attenuate the forward-guidance puzzle, has proposed Behavioral New Keynesian models. Agents in these models display cognitive discounting, that is, they are myopic toward events expected to happen far into the future, which they therefore discount at a much higher rate. Aside from the extra discounting, these models still maintain the rational expectations assumption.

This paper shows that under rational expectations, the model requires extremely large degrees of myopia to fit macroeconomic data. These are necessary to avoid unrealistically strong responses of output and inflation to monetary policy shocks expected to happen far into the future.

We also estimate the model under alternative expectation formation assumptions based on the adaptive learning literature: subjective expectations with infinite-horizon learning and subjective expectations with Euler-equation learning. Under infinite-horizon learning, agents need to form expectations about macroeconomic variables until the infinite future. The estimates indicate that the degree of cognitive discounting that is needed to fit the data is even higher, in order to alleviate the agents’ farsightedness in the model. Under Euler-equation learning, instead, the model equations resemble those under rational expectations, with output gap and inflation expressed as a function of one-period-ahead expectations. In this case, the model with learning can accurately fit the data with only minimal degrees of cognitive discounting. Moreover, the Euler-equation learning specification provides the best fit across all specifications.

In addition to shedding light on the behavioral features that seem to matter in macroeconomic data, the empirical results also carry implications for monetary policy. If learning matters, then price-level targeting may become optimal, as illustrated by Eusepi and Preston (2018). If cognitive discounting is an important feature of the data, instead, the analysis in Gabaix (2020) shows that price-level targeting may lead to lower-welfare outcomes. Understanding the behavioral factors that play a role at a macroeconomic level is hence important for the choice of desirable policies and is likely to be a prolific area for future research.

REFERENCES

- [1] Airaudo, M. (2020). “Temptation and forward-guidance,” *Journal of Economic Theory*, 186(C), 104989.
- [2] Andrade, J., Cordeiro, P., and G. Lambais, (2019). “Estimating a Behavioral New Keynesian Model,” Paper 1912.07601, arXiv.org.
- [3] Angeletos, G-M., and C. Lian (2018). “Forward Guidance without Common Knowledge”, *American Economic Review*, 108(9), 2477-2512.
- [4] Branch, W. A., and Evans, G. W. (2007). “Model uncertainty and endogenous volatility,” *Review of Economic Dynamics*, 10(2), 207-237.
- [5] Bray, M., and N. Savin (1986). “Rational Expectations Equilibria, Learning, and Model Specification”, *Econometrica*, 54(5), 1129-1160.
- [6] Bullard, J.B., and K. Mitra (2002). “Learning about monetary policy rules”, *Journal of Monetary Economics*, 49, 1105-1129.
- [7] Cole, S.J., (2020). “The limits of central bank forward guidance under learning,” *International Journal of Central Banking*, 16(4), 199-250.
- [8] Cole, S.J., (2021). “Learning and the effectiveness of central bank forward guidance,” *Journal of Money, Credit, and Banking*, 53(1), 157-200.
- [9] Cole, S., and F. Milani (2019). “The misspecification of expectations in New Keynesian models: a DSGE-VAR approach,” *Macroeconomic Dynamics*, 23(3), 974-1007.
- [10] DeCanio, S. (1979). “Rational Expectations and Learning from Experience”, *Quarterly Journal of Economics*, 93(1), 47-57.
- [11] Del Negro, M., Giannoni, M., and C. Patterson, (2015). “The Forward Guidance Puzzle”, Staff Reports 574, Federal Reserve Bank of New York.
- [12] Eusepi, S., and Preston, B. (2010). “Central Bank Communication and Expectations Stabilization,” *American Economic Journal: Macroeconomics*, 2(3), 235-271.
- [13] Eusepi, S., and Preston, B. (2011). “Expectations, learning, and business cycle fluctuations,” *American Economic Review*, 101(6), 2844-72.
- [14] Eusepi, S., and Preston, B. (2018). “The Science of Monetary Policy: An Imperfect Knowledge Perspective,” *Journal of Economic Literature*, 56(1), 3-59.
- [15] Evans, G. W. and Honkapohja, S. (2001). *Learning and expectations in macroeconomics*. Princeton University Press, Princeton, NJ.
- [16] Evans, G. W. and Honkapohja, S. (2003). “Expectations and the Stability Problem for Optimal Monetary Policies,” *Review of Economic Studies*, 70(4), 807-824.
- [17] Evans, G. W., and Honkapohja, S. (2013). “Learning as a rational foundation for macroeconomics and finance,” *Rethinking expectations: The way forward for macroeconomics*, Princeton University Press, Princeton, NJ.
- [18] Farhi, E., and I. Werning (2019). “Monetary Policy, Bounded Rationality, and Incomplete Markets,” *American Economic Review*, 109(11), 3887-3928.
- [19] Gabaix, X., (2014). “A Sparsity-Based Model of Bounded Rationality,” *The Quarterly Journal of Economics*, 129(4), 1661-1710.
- [20] Gabaix, X., (2016). “Behavioral Macroeconomics Via Sparse Dynamic Programming,” NBER Working Paper 21848, NBER, Inc.
- [21] Gabaix, X., (2019). “Behavioral Inattention,” *Handbook of Behavioral Economics*, 2, 261-344.

- [22] Gabaix, X., (2020). “A Behavioral New Keynesian Model,” *American Economic Review*, 110(8), 2271-2327.
- [23] Garcia-Schmidt, M., and M. Woodford (2019). “Are Low Interest Rates Deflationary? A Paradox of Perfect-Foresight Analysis,” *American Economic Review*, 109(1), 86-120.
- [24] Hommes, C. (2011). “The heterogeneous expectations hypothesis: Some evidence from the lab.” *Journal of Economic Dynamics & Control*, 35(1), 1-24.
- [25] Hommes, C. (2021). “Behavioral & Experimental Macroeconomics and Policy Analysis: a Complex Systems Approach,” *Journal of Economic Literature*, 59(1), 149-219.
- [26] Honkapohja, S., Mitra, K., and G. Evans (2011). “Notes on Agents’ Behavioral Rules Under Adaptive Learning and Studies of Monetary Policy,” CDMA Working Paper Series 201102, Centre for Dynamic Macroeconomic Analysis.
- [27] Ilabaca, F., Meggiorini, G., and F. Milani (2020). “Bounded rationality, monetary policy, and macroeconomic stability”, *Economics Letters*, Volume 186, Article 108522.
- [28] Jeffreys, H. (1961). *Theory of probability*, (3rd ed.), Oxford University Press, New York.
- [29] 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.
- [30] McCallum, B. T. (1999). “Issues in the Design of Monetary Policy Rules,” *Handbook of Macroeconomics*, 1(C), 1483-1530.
- [31] McKay, A., Nakamura, E., and J. Steinsson, (2016). “The Power of Forward Guidance Revisited”, *American Economic Review*, 106 (10), 3133-58.
- [32] Meggiorini, G., (2020). “Behavioral New Keynesian Models: A Look Through the Window”, mimeo, University of California, Irvine.
- [33] Milani, F. (2006). “A Bayesian DSGE Model with Infinite-Horizon Learning: Do “Mechanical” Sources of Persistence Become Superfluous?” *International Journal of Central Banking*, Issue 6, September.
- [34] Milani, F. (2007). “Expectations, learning and macroeconomic persistence,” *Journal of Monetary Economics*, 54(7), 2065-2082.
- [35] Milani, F. (2008). “Learning, Monetary Policy Rules, and Macroeconomic Stability”, *Journal of Economic Dynamics and Control*, 32(10), 3148-3165.
- [36] Milani, F. (2009). “Expectations, Learning, and the Changing Relationship between Oil Prices and the Macroeconomy”, *Energy Economics*, 31(6), 827-837.
- [37] Milani, F. (2011). “Expectation shocks and learning as drivers of the business cycle,” *Economic Journal*, 121(552), 379-401.
- [38] Milani, F. (2014). “Learning and time-varying macroeconomic volatility,” *Journal of Economic Dynamics and Control*, 47, 94-114.
- [39] Milani, F. (2017). “Sentiment and the US business cycle,” *Journal of Economic Dynamics and Control*, 82, 289-311.
- [40] Milani, F., and A. Rajbhandari (2020). “Observed Expectations, News Shocks, and the Business Cycle”, *Research in Economics*, 74(2), 95-118.
- [41] Orphanides, A., and J. Williams (2006). “Monetary Policy with Imperfect Knowledge”, *Journal of the European Economic Association*, 4(2/3), 366-375.
- [42] Preston, B. (2005). “Learning about Monetary Policy Rules when Long-Horizon Expectations Matter,” *International Journal of Central Banking*, 1(2), September.

- [43] Preston, B. (2006). “Adaptive learning, forecast-based instrument rules and monetary policy,” *Journal of Monetary Economics*, 53(3), 507-535.
- [44] Preston, B. (2008). “Adaptive learning and the use of forecasts in monetary policy,” *Journal of Economic Dynamics and Control*, 32(11), 3661-3681.
- [45] Sargent, T.J., (1993). *Bounded Rationality in Macroeconomics*. Oxford University Press, Oxford, U.K.
- [46] Sargent, T.J., (1999). *The Conquest of American Inflation*. Princeton University Press, Princeton, NJ.
- [47] Schmitt-Grohé, S., and M. Uribe (2012). “What’s News in Business Cycles?”, *Econometrica*, 80(6), 2733-2764.
- [48] Sims, C.A., (2002). “Solving Linear Rational Expectations Models”, *Computational Economics*, 20, 1-20.
- [49] Slobodyan, S., and Wouters, R. (2012). “Learning in a Medium-Scale DSGE Model with Expectations Based on Small Forecasting Models”, *American Economic Journal: Macroeconomics*, 4(2), 65-101.
- [50] Woodford, M., (2013). “Macroeconomic Analysis Without the Rational Expectations Hypothesis,” *Annual Review of Economics*, 5(1), 303-346.
- [51] Woodford, M. (2018). “Monetary policy analysis when planning horizons are finite”, *NBER Macroeconomics Annual*, University of Chicago Press, 33(1), 1-50.

Param.	Prior Distr.	Posterior Distributions		
		<u>RE</u>	<u>IH</u>	<u>EE</u>
<i>IS & PC</i>				
σ	$\Gamma(1.50,0.75)$	0.1862 [0.09,0.34]	0.3381 [0.14,0.60]	0.3258 [0.20,0.49]
α	$B(0.66,0.1)$	0.8273 [0.73,0.91]	0.7867 [0.65,0.89]	0.7989 [0.71,0.89]
<i>Taylor Rule</i>				
ρ	$B(0.7,0.2)$	0.8995 [0.85,0.94]	0.9035 [0.86,0.95]	0.9035 [0.86,0.95]
χ_π	$N(1.5,0.25)$	1.3052 [0.94,1.70]	1.3275 [0.95,1.74]	1.3182 [0.94,1.72]
χ_x	$N(0.1,0.05)$	0.1451 [0.06,0.22]	0.1471 [0.06,0.23]	0.1489 [0.07,0.23]
<i>Myopia</i>				
\bar{m}	$B(0.8,0.15)$	0.4160 [0.17,0.67]	0.1976 [0.10,0.29]	0.9366 [0.89,0.98]
<i>Learning</i>				
\bar{g}	$B(0.025,0.01)$	—	0.0454 [0.03,0.07]	0.0049 [0.002,0.01]
<i>Shocks</i>				
ρ_r	$B(0.5,0.2)$	0.9174 [0.87,0.96]	0.8939 [0.84,0.95]	0.2278 [0.09,0.37]
ρ_u	$B(0.5,0.2)$	0.8833 [0.83,0.94]	0.7796 [0.69,0.86]	0.0654 [0.01,0.16]
σ_r	$\Gamma^{-1}(0.30,1)$	3.0477 [1.64,5.23]	2.7384 [1.35,5.62]	2.5735 [1.65,4.05]
σ_u	$\Gamma^{-1}(0.30,1)$	0.1766 [0.12,0.23]	0.2689 [0.24,0.30]	0.2609 [0.24,0.29]
σ_ε	$\Gamma^{-1}(0.30,1)$	0.2338 [0.21,0.26]	0.2336 [0.21,0.26]	0.2337 [0.21,0.26]
<i>log Marg. Likelihood</i>		-276.1951	-277.0829	-267.1622

TABLE 1. Posterior Estimates: Full-Sample Estimation. *Note:* The Table reports mean posterior estimates, along with 2.5% and 97.5% percentiles. We ran 1,500,000 MH draws, discarding the first 40% as initial burn-in.

Param.	Prior Distr.	Posterior Distributions					
		<u>RE</u>		<u>IH</u>		<u>EE</u>	
		1954-1979	1982-2007	1954-1979	1982-2007	1954-1979	1982-2007
<i>IS & PC</i>							
σ	$\Gamma(1.50,0.75)$	0.3569 [0.16,0.66]	0.3029 [0.13,0.56]	0.6447 [0.30,1.05]	0.4377 [0.18,0.76]	0.7603 [0.39,1.26]	0.4195 [0.20,0.73]
α	$B(0.66,0.1)$	0.7964 [0.67,0.90]	0.7662 [0.65,0.87]	0.7082 [0.58,0.85]	0.7914 [0.67,0.90]	0.7429 [0.61,0.88]	0.7957 [0.66,0.90]
<i>Taylor Rule</i>							
ρ	$B(0.7,0.2)$	0.7711 [0.65,0.89]	0.9186 [0.88,0.96]	0.7943 [0.67,0.91]	0.9195 [0.88,0.96]	0.7987 [0.68,0.92]	0.9199 [0.88,0.96]
χ_π	$N(1,0.25)$	0.7901 [0.53,1.08]		0.7651 [0.49,1.06]		0.7705 [0.49,1.09]	
	$N(1.5,0.25)$		1.4495 [1.00,1.91]		1.4221 [0.99,1.87]		1.436 [0.99,1.89]
χ_x	$N(0.1,0.05)$	0.1666 [0.10,0.23]	0.1343 [0.05,0.23]	0.1663 [0.09,0.24]	0.1257 [0.04,0.21]	0.1654 [0.09,0.23]	0.1291 [0.04,0.21]
<i>Myopia</i>							
\bar{m}	$B(0.8,0.15)$	0.6305 [0.31,0.90]	0.4088 [0.17,0.66]	0.2746 [0.13,0.44]	0.2757 [0.16,0.39]	0.9371 [0.86,0.99]	0.7909 [0.36,0.95]
<i>Learning</i>							
\bar{g}	$B(0.025,0.01)$			0.0403 [0.02,0.06]	0.0332 [0.01,0.06]	0.0209 [0.01,0.03]	0.007 [0.002,0.02]
<i>Shocks</i>							
ρ_r	$B(0.5,0.2)$	0.8849 [0.81,0.96]	0.9316 [0.88,0.98]	0.8294 [0.73,0.93]	0.9059 [0.84,0.96]	0.2335 [0.07,0.42]	0.6089 [0.34,0.91]
ρ_u	$B(0.5,0.2)$	0.8308 [0.73,0.93]	0.8226 [0.73,0.91]	0.6970 [0.55,0.83]	0.6890 [0.57,0.81]	0.1575 [0.03,0.33]	0.2430 [0.03,0.75]
σ_r	$\Gamma^{-1}(0.30,1)$	1.4343 [0.67,2.70]	1.3964 [0.74,2.56]	1.7265 [0.97,3.293]	1.4740 [0.73,2.99]	1.3936 [0.79,2.51]	1.4655 [0.78,2.75]
σ_u	$\Gamma^{-1}(0.30,1)$	0.1864 [0.11,0.28]	0.1226 [0.09,0.16]	0.3513 [0.30,0.41]	0.1824 [0.16,0.21]	0.3357 [0.29,0.39]	0.1782 [0.15,0.21]
σ_ε	$\Gamma^{-1}(0.30,1)$	0.2065 [0.18,0.24]	0.1521 [0.13,0.17]	0.2016 [0.18,0.24]	0.1521 [0.13,0.17]	0.2070 [0.18,0.24]	0.1524 [0.13,0.17]
<i>log Marg. Likelihood</i>		-156.3274	-38.2284	-156.4760	-35.7699	-145.8191	-28.8900

TABLE 2. Posterior Estimates: pre-1979 and post-1982 Split Sample Estimation.
Note: The Table reports mean posterior estimates, along with 2.5% and 97.5% percentiles. We ran 1,500,000 MH draws, discarding the first 40% as initial burn-in.

Param.	Prior Distr.	Posterior Distributions		
		<u>RE</u>	<u>IH</u>	<u>EE</u>
<i>IS & PC</i>				
σ	$\Gamma(1.50,0.75)$	0.1843 [0.08,0.33]	0.3271 [0.16,0.59]	0.3282 [0.20,0.49]
α	$B(0.66,0.1)$	0.8337 [0.74,0.92]	0.7926 [0.67,0.89]	0.7965 [0.71,0.89]
<i>Taylor Rule</i>				
ρ	$B(0.7,0.2)$	0.8988 [0.85,0.94]	0.9034 [0.85,0.95]	0.9038 [0.85,0.95]
χ_π	$N(1.5,0.25)$	1.3144 [0.94,1.71]	1.3123 [0.95,1.74]	1.3198 [0.93,1.73]
χ_x	$N(0.1,0.05)$	0.1514 [0.07,0.23]	0.1499 [0.07,0.23]	0.1440 [0.06,0.23]
<i>Myopia</i>				
\bar{m}	$B(0.8,0.15)$	0.4655 [0.18,0.79]	0.1954 [0.10,0.29]	0.9364 [0.89,0.98]
<i>Learning</i>				
\bar{g}	$B(0.025,0.01)$	—	0.0444 [0.03,0.07]	0.0049 [0.002,0.01]
<i>Shocks</i>				
ρ_r	$B(0.5,0.2)$	0.9177 [0.87,0.96]	0.8947 [0.84,0.95]	0.2276 [0.09,0.37]
ρ_u	$B(0.5,0.2)$	0.8817 [0.82,0.94]	0.7810 [0.69,0.87]	0.0656 [0.01,0.16]
σ_r	$\Gamma^{-1}(0.30,1)$	2.8517 [1.29,5.43]	2.7602 [1.40,5.00]	2.5651 [1.62,4.02]
σ_u	$\Gamma^{-1}(0.30,1)$	0.1663 [0.09,0.23]	0.2687 [0.24,0.30]	0.2607 [0.24,0.29]
σ_ε	$\Gamma^{-1}(0.30,1)$	0.1528 [0.08,0.23]	0.1644 [0.09,0.24]	0.1668 [0.09,0.24]
σ_{ε^4}	$\Gamma^{-1}(0.30,1)$	0.1046 [0.004,0.21]	0.0987 [0.004,0.20]	0.0985 [0.004,0.21]
σ_{ε^8}	$\Gamma^{-1}(0.30,1)$	0.1091 [0.005,0.21]	0.0974 [0.003,0.21]	0.0956 [0.004,0.21]
log Marg. Likelihood		-275.8665	-278.3282	-267.7775

TABLE 3. Posterior Estimates: Taylor rule with news shocks. *Note:* The Table reports mean posterior estimates, along with 2.5% and 97.5% percentiles. We ran 1,500,000 MH draws, discarding the first 40% as initial burn-in.

Param.	Prior Distr.	Posterior Distributions (EE Learning)			
		Uniform prior	Gamma Prior	Three Gains	$\bar{m} = 1$
<i>IS & PC</i>					
σ	$\Gamma(1.50,0.75)$	0.3088 [0.18,0.46]	0.3387 [0.21,0.50]	0.3274 [0.16,0.52]	0.3174 [0.20,0.47]
α	$B(0.66,0.1)$	0.8092 [0.73,0.91]	0.7991 [0.72,0.87]	0.7859 [0.69,0.88]	0.7866 [0.71,0.87]
<i>Taylor Rule</i>					
ρ	$B(0.7,0.2)$	0.9024 [0.85,0.95]	0.9044 [0.86,0.95]	0.9025 [0.85,0.95]	0.9025 [0.86,0.95]
χ_π	$N(1.5,0.25)$	1.2991 [0.89,1.78]	1.3272 [0.96,1.81]	1.3110 [0.96,1.75]	1.3096 [0.93,1.73]
χ_x	$N(0.1,0.05)$	0.1374 [0.05,0.22]	0.1550 [0.07,0.25]	0.1376 [0.05,0.21]	0.1461 [0.07,0.23]
<i>Myopia</i>					
\bar{m}	varies	0.9382 [0.89,0.98]	0.9437 [0.90,0.99]	0.8799 [0.79,0.94]	1
<i>Learning</i>					
\bar{g}	$B(0.025,0.01)$	0.0049 [0.002,0.009]	0.0050 [0.002,0.009]	—	0.0056 [0.003,0.009]
\bar{g}_x	"	—	—	0.0059 [0.003,0.011]	—
\bar{g}_π	"	—	—	0.0370 [0.024,0.049]	—
\bar{g}_i	"	—	—	0.0198 [0.007,0.038]	—
<i>Shocks</i>					
ρ_r	$B(0.5,0.2)$	0.2146 [0.09,0.37]	0.2236 [0.09,0.36]	0.2824 [0.13,0.47]	0.1776 [0.07,0.31]
ρ_u	$B(0.5,0.2)$	0.0625 [0.01,0.14]	0.0589 [0.01,0.14]	0.2218 [0.06,0.40]	0.0485 [0.09,0.12]
σ_r	$\Gamma^{-1}(0.30,1)$	2.7524 [1.71,4.37]	2.4694 [1.61,3.68]	2.6819 [1.60,5.07]	2.6773 [1.75,3.97]
σ_u	$\Gamma^{-1}(0.30,1)$	0.2606 [0.23,0.29]	0.2605 [0.23,0.29]	0.2624 [0.24,0.29]	0.2617 [0.24,0.29]
σ_ε	$\Gamma^{-1}(0.30,1)$	0.2343 [0.21,0.26]	0.2335 [0.21,0.26]	0.2346 [0.21,0.26]	0.2343 [0.21,0.26]
log Marg. Likelihood		-268.7747	-267.9884	-266.2732	-270.4209

TABLE 4. Posterior Estimates: Specification with Euler-equation learning under different priors for \bar{m} , different gain coefficients for different variables, and for no cognitive discounting ($\bar{m} = 1$).

Note: The Table reports mean posterior estimates, along with 2.5% and 97.5% percentiles.

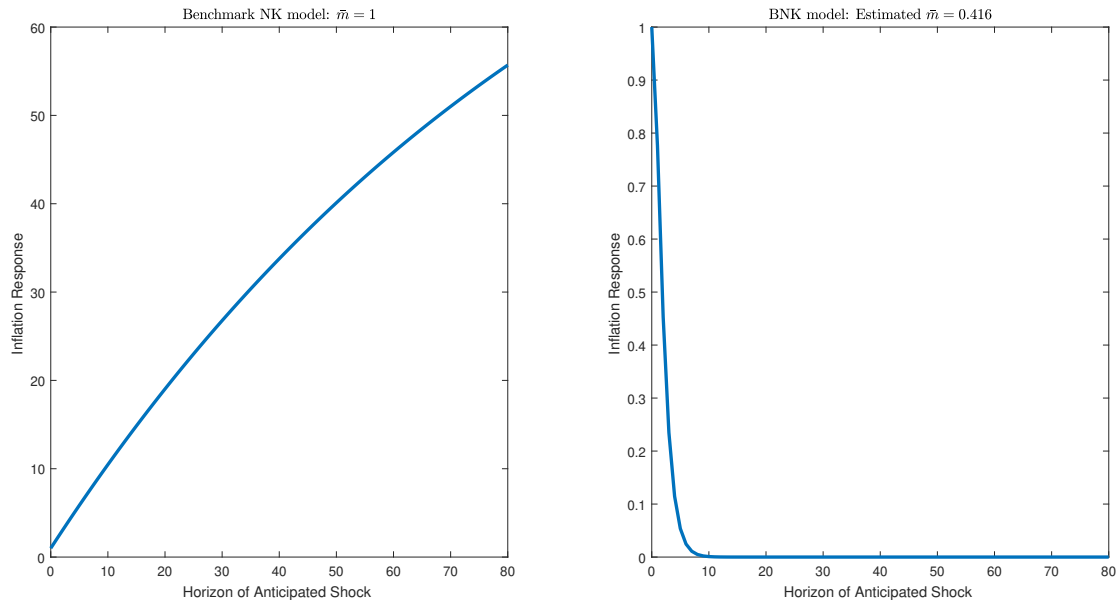


FIGURE 1. Impact of Forward Guidance shock on current inflation.

Note: The Figure shows the response of inflation at time 0 to an expansionary anticipated interest rate shock happening T periods into the future (with horizons ranging from $T = 0$ to $T = 80$). The left panel displays the response in the case of \bar{m} fixed at 1 (benchmark NK model with full rationality). The right panel shows the response when \bar{m} is equal to the estimated posterior mean ($\bar{m} = 0.416$). All the remaining parameters are set at the posterior mean values reported in Table 1.

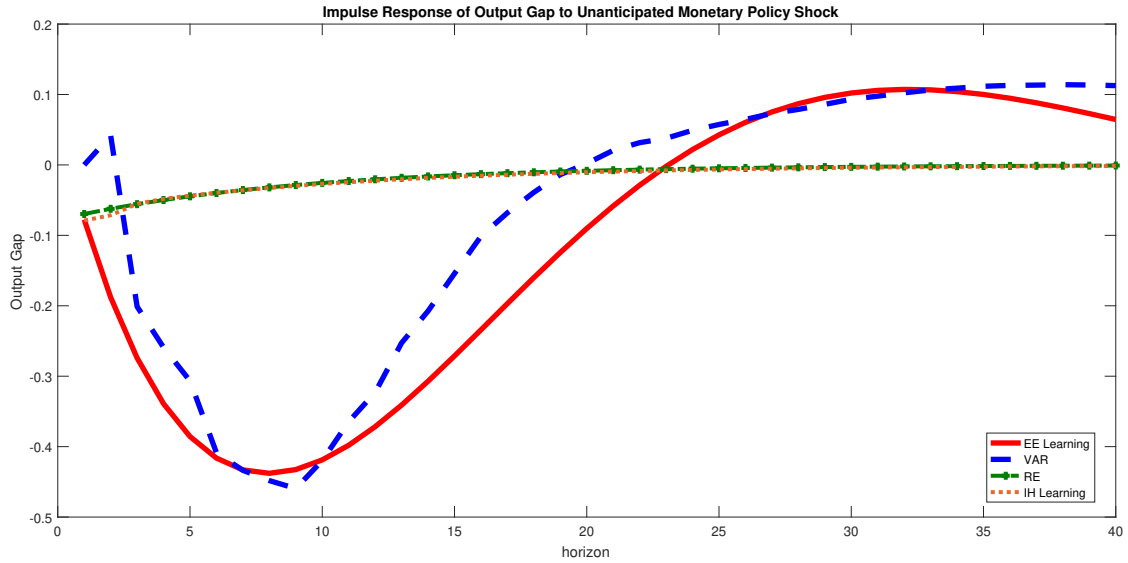


FIGURE 2. Impulse response functions of output gap to monetary policy shocks.

Note: The figure shows impulse responses for the model under the three expectation formation specifications, compared with those from a VAR. The solid red line denotes the response under EE learning; the dotted orange line denotes the response under IH learning; the dashed green line denotes the response under RE. The dashed blue line shows the corresponding impulse response in a VAR estimated on the same observable variables (output gap, inflation, Federal funds rate).