Monetary Policy & Anchored Expectations
An Endogenous Gain Learning Model

Preliminary and Incomplete

Laura Gáti

February 14, 2020

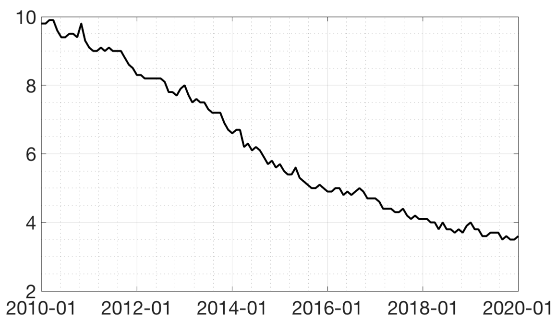
Abstract

This paper analyzes optimal monetary policy in a model where expectation formation is characterized by potential anchoring of expectations. Expectations are anchored when in an adaptive learning setting, the private sector endogenously chooses a decreasing learning gain. Within the context of an otherwise standard macro model with nominal rigidities, I find that the central bank trades off the short-run costs with the long-run benefits of anchoring expectations. Having anchored expectations reduces the volatility of observables in the long run, but getting expectations anchored is costly in terms of inducing volatility in the short run. Optimal policy is therefore conditioned on the stance of expectations.

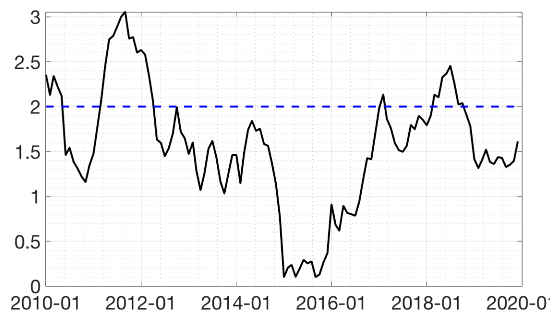
1 Introduction

The current stance of the United States business cycle is boldly defiant of mainstream macroeconomic theory. The historically low unemployment level, portrayed on panel (a) of Fig. 1, has not resulted in rising inflation. On the contrary, personal consumption expenditures (PCE) inflation has persistently undershot the Federal Reserve's 2% target, prompting the Fed to be expansionary despite the economy experiencing a boom (panels (b) and (c) of Fig. 1).

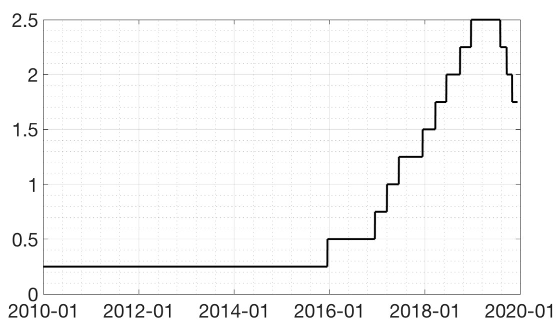
Figure 1



(a) Unemployment rate, %



(b) PCE inflation, % change from 12 months ago

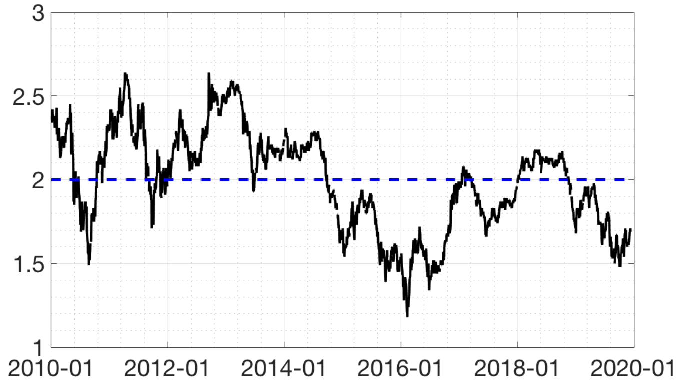


(c) Fed funds rate target, upper limit, %

In this paper I argue that the key to understanding both the puzzling behavior of inflation as well as the Fed's response to it is the time series of long-run inflation expectations. As Fig. 2 shows, long-run inflation expectations of the public, averaging a little above the 2% target prior to 2015, display a marked downward drift since 2015. This indicates that the public has doubts whether the Fed is able to restore inflation to the target. Confronted with what it sees as a changing environment, the public revises its predictions about the future course of the economy.

Macroeconomic theory that seeks to understand this phenomenon thus needs to account for expectation formation that features a notion of stability of forecasting behavior. I propose a model in the adaptive learning tradition where the public sector's choice of learning gain is endogenous, as in Carvalho et al. (2019). This captures the idea that in normal times, when firms and households observe economic data that confirms their previous predictions, agents choose a decreasing gain and thus do not change their forecasting rules by much. By contrast, when incoming data suggests that the current forecasting rule is incorrect, agents switch to a constant gain, updating their forecasting rule strongly.

Figure 2



(a) Market-based inflation expectations, 10 year, average, %

I refer to the former case as *anchored* and to the latter as *unanchored expectations*.

The contribution of this paper is to investigate how this affects the optimal conduct of monetary policy. I embed the anchoring mechanism in an otherwise standard New Keynesian model of the type widely used for monetary policy analysis.[1] This allows a crisp comparison between optimal monetary policy in the standard rational expectations model and the model with the expectation-anchoring mechanism. It turns out that optimal monetary policy under anchoring takes the stance of expectations explicitly into account. In particular, the central bank finds it optimal to anchor expectations whenever it is not too costly to do so. This is because having anchored expectations reduces the volatility of expectations. Since expectations affect outcomes via feedback effects, this lowers the volatility of observables as well. However, anchoring expectations comes at a short-run cost of heightened volatility as agents internalize future policy responses causing fluctuations in long-run expectations. This introduces tradeoffs in the conduct of monetary policy, presenting a novel case for violations of the divine coincidence. This insight allows us to interpret the Fed's fall 2019 decision to lower interest rates despite a strong economy as an attempt to anchor expectations or to keep them from becoming unanchored.

1.1 Related literature

My work draws on two strands of macroeconomic research. The first is the extensive literature on optimal monetary policy in the New Keynesian model. Most of this literature, such as Clarida et al. (1999) or Woodford (2011), relies on the rational expectations and thus serves as a natural benchmark of comparison.

The second branch of related work is the adaptive learning literature. Following the book by Evans and Honkapohja (2001), this literature replaces the rational expectations assumption by postulating an ad-hoc forecasting rule, the perceived law of motion (PLM), as the expectation-formation process. Agents use the PLM to form expectations and update it in every period using recursive estimation techniques.

[1] For comparability, I follow the exposition of the New Keynesian model in Woodford (2011).

The early learning literature concentrated on the question of under what conditions learning converges to a rational expectations equilibrium (this is what Evans and Honkapohja (2001) term "expectational stability" or "E-stability"). E-stability and related notions still form the core of work in learning, as in Ferrero (2007), Eusepi and Preston (2018), or, in the context of the New Keynesian model, Bullard and Mitra (2002) and Preston (2005).

Of more concern for monetary policy analysis is the quantitative dynamics of models with learning. Williams (2003) and Eusepi and Preston (2011) investigate real business cycle (RBC) and NK models with learning and find dampened impact effects and increased persistence in response to shocks. A prevalent yet not extensively discussed finding in the quantitative learning literature is that learning models with a sufficiently high constant gain have a tendency to exhibit oscillatory impulse responses. This becomes relevant for my work because it highlights both why anchoring expectations is beneficial for lowering economic volatility in the long run, and at the same time induces short-run volatility.

My work is closely related to studies that reevaluate optimal monetary policy from the lens of a learning model (Orphanides and Williams (2005), Gaspar et al. (2006), Evans and Honkapohja (2006), Ferrero (2007), Preston (2008), Molnár and Santoro (2014), Eusepi and Preston (2018), Eusepi et al. (2018a)). There is no consensus on how learning affects optimal monetary policy. In the case of the New Keynesian model, for example, Eusepi and Preston (2018) and Molnár and Santoro (2014) conclude that optimal monetary policy is more aggressive on inflation than under rational expectations, yet Eusepi et al. (2018a) find the exact opposite. My finding that policy trades off short-run with long-run volatility finds elements of truth in both arguments, and echoes the insight of Lubik and Matthes (2016) that the intertemporal tradeoff induces monetary policy to be time-inconsistent.

To the best of my knowledge, only few papers study models with an endogenous gain: Marcet and Nicolini (2003), Milani (2014), and Carvalho et al. (2019). Having an endogenous gain allows the modeler to capture an essential element of learning: the public's confidence that it has found the correct model of the economy. This is precisely the notion Carvalho et al. (2019) use to model anchored expectations and to estimate the endogenous gain. I embed an anchoring mechanism related to that of Carvalho et al. (2019) in a general equilibrium New Keynesian model and investigate the interaction between monetary policy and anchored expectations.

The paper is structured as follows. Section 2 introduces the model and the monetary policy problem. Section 3 describes the learning framework and spells out the anchoring mechanism. Section 4 presents results and provides intuition. Section 5 concludes.

2 The model

Apart from expectation formation, the model is a standard New Keynesian model where the rational expectations (RE) assumption is replaced by the expectation-anchoring mechanism. The advantage of having a standard NK backbone to the model is that one can neatly isolate the way the anchoring mechanism alters the behavior of the model. Since the mechanics of the rational expectations version of this model are well understood, I only lay out the model briefly and pinpoint the places where the assumption of nonrational expectations matters.[2]

2.1 Households

The representative household is infinitely-lived and maximizes expected discounted lifetime utility from consumption net of the disutility of supplying labor hours:

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

$U(\cdot)$ and $v(\cdot)$ denote the utility of consumption and disutility of labor respectively and β is the discount factor of the household. $h_t^i(j)$ denotes the supply of labor hours of household i at time t to the production of good j and the household participates in the production of all goods j. Similarly, household i's consumption bundle at time t, C_t^i, is a Dixit-Stiglitz composite of all goods in the economy:

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

$\theta > 1$ is the elasticity of substitution between the varieties of consumption goods. Denoting by $p_t(j)$ the time-t price of good j, the aggregate price level in the economy can then be written as

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

The budget constraint of household i 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 $\Pi_t^i(j)$ denotes profits from firm j remitted to household i, T_t taxes, and B_t^i the riskless bond purchases at time t.[3]

The only difference to the standard New Keynesian model thus far is the expectation operator, $\hat{\mathbb{E}}$. This is the subjective expectation operator that differs from its rational expectations counterpart, $\mathbb{E}$, in that it does not encompass knowledge of the model. In particular, households have no knowledge of the fact that they are identical and by extension they also do not internalize that they hold identical beliefs about the evolution of the economy. As we will see in Section 2.3, this has implications for their forecasting behavior and will result in decision rules that differ from those of the rational expectations version of the model.

[2]For the specifics of the NK model the reader is referred to Woodford (2011).

[3]For ease of exposition I have suppressed potential money assets here. This has no bearing on the model implications since it represents the cashless limit of an economy with explicit money balances.

2.2 Firms

Firms are monopolistically competitive producers of the differentiated varieties $y_t(j)$. The production technology of firm j is $y_t(j) = A_t f(h_t(j))$, whose inverse, $f^{-1}(\cdot)$, signifies the amount of labor input. Noting that A_t is the level of technology and that $w_t(j)$ is the wage per labor hour, firm j profits at time t can be written as

$$\Pi_t^j = p_t(j)y_t(j) - w_t(j)f^{-1}(y_t(j)/A_t) \tag{5}$$

Firm j's problem then is to set the price of the variety it produces, $p_t(j)$, to maximize the present discounted value of profit streams

$$\hat{\mathbb{E}}_t \sum_{T=t}^{\infty} \alpha^{T-t} Q_{t,T} \left[\Pi_t^j(p_t(j)) \right] \tag{6}$$

subject to the downward-sloping demand curve

$$y_t(j) = Y_t \left(\frac{p_t(j)}{P_t} \right)^{-\theta} \tag{7}$$

where

$$Q_{t,T} = \beta^{T-t} \frac{P_t U_c(C_T)}{P_T U_c(C_t)} \tag{8}$$

is the stochastic discount factor from households. Nominal frictions enter the model through the parameter α in Equation (6). This is the Calvo probability that firm j is not able to adjust its price in a given period.

Analogously to households, the setup of the production side of the economy is standard up to the expectation operator. Also here the rational expectations operator $\mathbb{E}$ has been replaced by the subjective expectations operator $\hat{\mathbb{E}}$. This implies that firms, like households, do not know the model equations and fail to internalize that they are identical. Thus their decision rules, just like those of the households, will be distinct from their rational expectations counterparts.

2.3 Aggregate laws of motion

The model solution procedure entails deriving first order conditions, taking a loglinear approximation around the nonstochastic steady state and imposing market clearing conditions to reduce the system to two equations, the New Keynesian Phillips curve (NKPC) and IS curve (NKIS). The presence of subjective expectations, however, implies that firms and households are not aware of the fact that they are identical. Thus, as Preston (2005) takes pains to point out, imposing market clearing conditions in the expectations of agents is inconsistent with the assumed information structure.[4]

[4] The target of Preston (2005)'s critique is the Euler-equation approach as exemplified for example by Bullard and Mitra (2002). This approach involves writing down the loglinearized first order conditions of the model, and simply replacing the rational expectations operators with subjective ones. In a separate paper, I demonstrate that the Euler-equation approach is not only inconsistent on conceptual grounds as Preston (2005) shows, but also delivers substantially different quantitative dynamics in a simulated New Keynesian model. Thus relying on the Euler-equation approach when investigating the role of learning is not only incorrect in terms of microfoundations, but also leads to mistaken quantitative inferences.

Instead, I follow Preston (2005) in preventing firms and households from internalizing market clearing conditions. As Preston (2005) demonstrates, this leads to long-horizon forecasts showing up in firms' and households' first order conditions. As a consequence, instead of the familiar expressions, the NKIS and NKPC take the following form:

$$x_t = -\sigma i_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} \beta^{T-t}\big((1-\beta)x_{T+1} - \sigma(\beta i_{T+1} - \pi_{T+1}) + \sigma r_T^n\big) \tag{9}$$

$$\pi_t = \kappa x_t + \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t}\big(\kappa\alpha\beta x_{T+1} + (1-\alpha)\beta\pi_{T+1} + u_T\big) \tag{10}$$

Here x_t, π_t and i_t are the log-deviations of the output gap, inflation and the nominal interest rate from their steady state values, and σ is the intertemporal elasticity of substitution. The variables r_t^n and u_t are exogenous disturbances representing a natural rate shock and a cost-push shock respectively.

The laws of motion (9) and (10) are obtained by deriving individual firms' and households' decision rules, which involve long-horizon expectations, and aggregating across the cross-section. Importantly, agents in the economy have no knowledge of these relations since they don't know that they are identical and thus are not able to impose market clearing conditions required to arrive at (9) and (10). Thus, although the evolution of the observables (π, x) is governed by the exogenous state variables (r^n, u) and long-horizon expectations via these two equations, agents in the economy are unaware of this. As I will spell out more formally in Section 3, it is indeed the equilibrium mapping between states and jump variables the agents are attempting to learn.[5]

The model is closed by the standard specification of monetary policy as a Taylor rule:

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

where ψ_π and ψ_x represent the responsiveness of monetary policy to inflation and the output gap respectively, $\bar{\pi}$ and $\bar{x}$ are the central bank's targets. Lastly, $\bar{i}_t$ is a monetary policy shock. I assume that the central bank publicly announces the Taylor rule. Thus Equation (11) is common knowledge and is therefore not the object of learning.[6]

Next, to simplify notation, I gather the exogenous state variables in the vector s_t and jump variables in the vector z_t as

$$s_t = \begin{bmatrix} r_t^n \\ \bar{i}_t \\ u_t \end{bmatrix} \qquad z_t = \begin{bmatrix} \pi_t \\ x_t \\ i_t \end{bmatrix} \tag{12}$$

[5]The learning of (9) and (10) is complicated by the fact that the current stance of expectations figures into the equations, resulting in the well-known positive feedback effects of learning.

[6]In an extension I consider the case where the Taylor rule is not known (or not believed) by the public and therefore is learned together with the relations (9) and (10). This dampens intertemporal expectation effects as long as the Taylor rule is not learned; afterwards, the model dynamics are identical to those of the baseline.

Then, denoting long-horizon expectations as

$$f_a(t) \equiv \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} z_{T+1} \qquad\qquad f_b(t) \equiv \hat{\mathbb{E}}_t \sum_{T=t}^{\infty} (\beta)^{T-t} z_{T+1} \tag{13}$$

I write the laws of motion of jump variables (Equations (9), (10) and (11)) compactly as

$$z_t = A_a f_a(t) + A_b f_b(t) + A_s s_t \tag{14}$$

where the matrices $A_i,\ i = \{a, b, s\}$ gather coefficients and are given in App. A. Assuming that exogenous variables evolve according to independent AR(1) processes, I write the state transition matrix equation as

$$s_t = h s_{t-1} + \epsilon_t \qquad \epsilon_t \sim \mathcal{N}(\mathbf{0}, \Sigma) \tag{15}$$

where h gathers the autoregressive coefficients ρ_j, ϵ_t the Gaussian innovations ε_t^j, and η the standard deviations σ_t^j, for $j = \{r, i, u\}$. $\Sigma = \eta\eta'$ is the variance-covariance matrix of disturbances.

$$h \equiv \begin{pmatrix} \rho_r & 0 & 0 \\ 0 & \rho_i & 0 \\ 0 & 0 & \rho_u \end{pmatrix} \quad \epsilon_t \equiv \begin{pmatrix} \varepsilon_t^r \\ \varepsilon_t^i \\ \varepsilon_t^u \end{pmatrix} \quad \text{and} \quad \eta \equiv \begin{pmatrix} \sigma_r & 0 & 0 \\ 0 & \sigma_i & 0 \\ 0 & 0 & \sigma_u \end{pmatrix} \tag{16}$$

Thus, while the state-space form of the solution for the rational expectations version of the model takes the form

$$s_t = h s_{t-1} + \epsilon_t \tag{17}$$

$$z_t = g^{RE} s_t \tag{18}$$

the model with learning leaves the state transition equation (17) unchanged, but replaces (18) with the law of motion for observables (14). Once I have specified expectation formation, I will revisit this formulation to highlight more formally the difference between the rational expectations and learning versions of the model.

2.4 The monetary policy problem

I assume the monetary authority seeks to maximize welfare of the representative household under commitment. As shown in Woodford (2011), a second-oder Taylor approximation of household utility delivers a central bank loss function of the form

$$L^{CB} = \mathbb{E}_t \sum_{T=t}^{\infty} \{\pi_T^2 + \lambda_x (x_T - x^*)^2 + \lambda_i (i_T - i^*)\} \tag{19}$$

where $\lambda_j\ j = \{x, i\}$ is the weight the central bank assigns to stabilizing variable j and j^* is its target value. The central bank's problem, then, is to determine paths for inflation, the output gap and the interest rate that minimize the loss in Equation (19), subject to the model equations (9) and (10), as well as the evolution of long-horizon expectations, spelled out in Section 3. A second question is to find a policy rule, that is a response function for the policy instrument i_t, that implements the optimal allocation.

For the optimal plan, I consider the time-zero and timelessly optimal solutions to the problem, as advocated by Woodford (2011). The comparison between the two is informative because of the finding that anchoring introduces an intertemporal tradeoff for monetary policy. This implies that optimizing at time t_1 yields a different optimal policy than the one obtained when optimizing at time t_2; as under rational expectations, time-zero optimal commitment is time-inconsistent. In the learning model, however, this happens for different reasons than under RE.

This is related and yet distinct from the observation of Lubik and Matthes (2016) that learning results in time-varying monetary policy coefficients. In Lubik and Matthes (2016), learning behavior on the part of the central bank combined with data misperceptions drives the time-variability of Taylor-rule coefficients. In other words, central bank beliefs become a state variable that affects the choice of policy coefficients. As in learning models in general, it is true also here that beliefs enter the model as an endogenous state variable. Thus Lubik and Matthes (2016)'s point is valid also in this context.

However, there is an additional channel coming from the anchoring mechanism. As I explain in detail in Section 4.1, the same monetary policy coefficients have different volatility implications depending on whether expectations are anchored or not. At the same time, getting expectations anchored also results in heightened volatility in the short run. This intertemporal tradeoff provides an additional reason for optimal time-zero commitment to be time-inconsistent.

3 Learning with anchored expectations

The main informational assumption of the model is that agents have no knowledge of the equilibrium mapping between states and jumps in the model. Therefore they are not able to form rational expectations forecasts. To see this, observe that an agent with rational expectations would internalize the rational expectations state-space system (17) - (18) and would therefore forecast future jumps as $\mathbb{E}_t\, z_{t+h} = g^{RE}\,\mathbb{E}_t\, s_{t+h} = g^{RE} h^h s_t$. Agents in the learning model however don't know g^{RE} and are thus indeed unable to form the rational expectations forecast. Instead, agents postulate an ad-hoc forecasting relationship between states and jumps and seek to refine it in light of incoming data.

3.1 Perceived law of motion

I assume agents consider a forecasting model for jumps of the form

$$\hat{\mathbb{E}}_t z_{t+1} = a_{t-1} + b_{t-1} s_t \tag{20}$$

where a and b are estimated coefficients of dimensions 3×1 and 3×3 respectively. This perceived law of motion (PLM) reflects the assumption that agents forecast jumps using a linear function of current states and a constant, with last period's estimated coefficients. I also assume that

$$\hat{\mathbb{E}}_t \phi_{t+h} = \phi_t \quad \forall\, h \geq 0 \tag{21}$$

This assumption, known in the learning literature as anticipated utility, means that agents fail to internalize that they will update the forecasting rule in the future.[7] Assuming that agents know the evolution of states, that is they have knowledge of Equation (17)[8], the PLM together with anticipated utility implies that h-period ahead forecasts are constructed as

$$\hat{\mathbb{E}}_t z_{t+h} = a_{t-1} + b_{t-1} h^{h-1} s_t \quad \forall h \geq 1 \tag{22}$$

Summarizing the estimated coefficients as $\phi_{t-1} \equiv \begin{bmatrix} a_{t-1} & b_{t-1} \end{bmatrix}$, here 3×4, I can rewrite Equation (20) as

$$\hat{\mathbb{E}}_t z_{t+1} = \phi_{t-1} \begin{bmatrix} 1 \\ s_t \end{bmatrix} \tag{23}$$

The timing assumptions of the model are as follows. In the beginning of period t, the current state s_t is realized. Agents then form expectations according to (20) using last period's estimate ϕ_{t-1} and the current state s_t. Given exogenous states and expectations, today's jump vector z_t is realized. This allows agents to evaluate the most recent forecast error $f_{t-1} \equiv z_t - \phi_{t-1} \begin{bmatrix} 1 \\ s_{t-1} \end{bmatrix}$ to update their forecasting rule. The estimate is updated according to the following recursive least-squares algorithm:

$$\phi_t = \left(\phi_{t-1}' + k_t^{-1} R_t^{-1} \begin{bmatrix} 1 \\ s_{t-1} \end{bmatrix} \left(z_t - \phi_{t-1} \begin{bmatrix} 1 \\ s_{t-1} \end{bmatrix} \right)' \right)' \tag{24}$$

$$R_t = R_{t-1} + k_t^{-1} \left(\begin{bmatrix} 1 \\ s_{t-1} \end{bmatrix} \begin{bmatrix} 1 & s_{t-1} \end{bmatrix} - R_{t-1} \right) \tag{25}$$

where R_t is the 4×4 variance-covariance matrix of the regressors and k_t is the learning gain, specifying to what extent the updated estimate loads on the forecast error. Clearly, a high gain implies high loadings and thus strong changes in the estimated coefficients ϕ. A low gain, by contrast, means that the current forecast error only has a small effect on ϕ_t.

The vast majority of the learning literature specifies the gain either as a constant, $\bar{g}$, or decreasing with time so that $k_t^{-1} = (k_{t-1} + 1)^{-1}$. Instead, in the spirit of Carvalho et al. (2019), I allow firms and households in the model to choose whether to use a constant or a decreasing gain. I use the following endogenous gain specification: let ω_t denote agents' time t estimate of the forecast error variance and

[7] This is a conventional assumption in the learning literature and serves to simplify the algebra. As Sargent (1999) shows, similar results obtain upon relaxing anticipated utility.

[8] Allowing agents to know the state transition equation is also a common simplifying assumption in the learning literature. In an extension, I relax this assumption and find that it has similar implications as having agents learn the Taylor rule: initial responses to shocks lack intertemporal expectation effects, but these reemerge as the evolution of state variables is learned.

θ_t be a statistic evaluated by agents in every period as

$$\omega_t = \omega_{t-1} + \tilde{\kappa} k_{t-1}^{-1}(f_{t-1}f_{t-1}' - \omega_{t-1}) \tag{26}$$

$$\theta_t = \theta_{t-1} + \tilde{\kappa} k_{t-1}^{-1}(f_{t-1}'\omega_t^{-1} f_{t-1} - \theta_{t-1}) \tag{27}$$

where $\tilde{\kappa}$ is a parameter that allows agents to scale the gain compared to the previous estimation and f_{t-1} is the most recent forecast error, realized at time t. Then for a specified threshold $\tilde{\theta}$, the gain is determined endogenously as

$$k_t = \begin{cases} k_{t-1} + 1 & \text{if} \quad \theta_t < \tilde{\theta} \\ \bar{g}^{-1} & \text{otherwise.} \end{cases} \tag{28}$$

In other words, agents choose a decreasing gain when the criterion θ_t is lower than the threshold $\tilde{\theta}$; otherwise they choose a constant gain. This framework, which I refer to as an anchoring mechanism, captures the intuition that when current squared forecast errors are large compared to agents' estimated forecast error variance, agents conclude that the forecasting performance of their current PLM ϕ_{t-1} is poor. Since ϕ_{t-1} appears to provide an inaccurate description of the evolution of observables, agents choose a constant gain, reflecting their desire to update ϕ strongly using the most recent data. By contrast, if $\theta_t < \tilde{\theta}$, current squared forecast errors are not sizable compared the estimated forecast error variance; it seems that ϕ_{t-1} is close to the data-generating process and so agents see no need to change it, thus opting for a decreasing gain. I therefore refer to the situation when $\theta_t < \tilde{\theta}$ as anchored expectations.

It is worthwhile to compare my anchoring criterion θ to the one in Carvalho et al. (2019). The criterion employed by Carvalho et al. (2019) is computed as the absolute difference between subjective and model-consistent expectations, scaled by the variance of shocks. On the one hand, their specification requires the private sector to evaluate model-consistent expectations, which runs counter to the maintained informational assumptions. It is more consistent with the present model, then, to assume that firms and households employ a statistical test of structural change. This motivates my choice of a statistic for θ as a multivariate time series version of the squared CUSUM test.[9]

On the other hand, simulation of the model using Carvalho et al. (2019)'s criterion reveals that their criterion leads to the opposite comparative statics of anchoring with respect to monetary policy aggressiveness. In particular, while policy that is more aggressive on inflation (a higher ψ_π) leads to more anchoring in a model with the CUSUM-inspired criterion, if one uses Carvalho et al. (2019)'s criterion, the same comparative static involves *less* anchoring. This comes from the fact that Carvalho et al. (2019)'s criterion endows the public sector with capabilities to disentangle volatility due to the learning mechanism from that owing to exogenous disturbances. Thus agents in the Carvalho et al. (2019) model are able to make more advanced inferences about the performance of their forecasting rule and understand that a higher ψ_π causes more learning-induced volatility. This is however not possible for agents who process data in real time without knowledge of the model. Therefore the

[9]See Brown et al. (1975) and Lütkepohl (2013) for details.

CUSUM-inspired criterion is preferable both on conceptual and quantitative grounds.

Further inspection of the anchoring mechanism foreshadows how anchoring and monetary policy interact in the model. Recall that $\phi = [a, b]$, where a is the estimate of the constant and b the estimate of the slope in the law of motion of jumps. Believing the estimate b to be correct means that agents think they know how observables respond to shocks. Analogously, thinking that the estimate a is correct has the interpretation of agents being confident about the long-run average values of the observables. But that is equivalent to agents trusting that the central bank is able to implement its proclaimed targets in the long run. In this way, anchored expectations has a natural interpretation as trust in the central bank's ability to achieve the long-run target.[10]

Having thus established anchored expectations as a metric of trust in the central bank implementing the announced target, it becomes intuitive why a monetary authority would want to make sure that expectations do not become unanchored. Clearly, unanchored expectations have the opposite interpretation to anchored ones: they reflect that the public has doubts whether the bank is committed or able to achieve the long-run target. And due to the feedback from expectations to observables (recall Equation (14)), unanchored expectations can have self-confirming effects in that they cause observables to drift away from the announced target. Thus a central bank that doesn't come across as committed to the target may fail to anchor expectations and will thus face additional difficulty in implementing the target.

3.2 Actual law of motion

Having laid out the expectation formation, I can now characterize the evolution of the jump variables under learning. Using the PLM from Equation (20), I write the long-horizon expectations in (13) as

$$f_a(t) = \frac{1}{1-\alpha\beta}a_{t-1} + b_{t-1}(I_3 - \alpha\beta h)^{-1}s_t \qquad f_b(t) = \frac{1}{1-\beta}a_{t-1} + b_{t-1}(I_3 - \beta h)^{-1}s_t \tag{29}$$

Substituting these into the law of motion of observables (Equation (14)) yields the actual law of motion (ALM):

$$z_t = g^l_{t-1}\begin{bmatrix}1\\ s_t\end{bmatrix} \tag{30}$$

where g^l is a 3×4 matrix given in App. B. Thus, instead of the state-space solution of the RE version of the model (Equations (17) and (18)), the state-space solution for the learning model is characterized by the pair of equations (17) and (30).

[10]For this reason, a learning specification in which only the constant is learned ($b = b^{RE}$) is sufficient to analyze anchored expectations qualitatively. Quantitatively, however, learning the slope makes a big difference. For this reason, figures that aim to provide intuition on the workings of anchoring, such as the figures of Section 4.1, are, unless stated otherwise, computed using the constant-only specification.

4 Monetary policy and anchoring

In this section I use the model developed above to analyze the interaction between monetary policy and the anchoring mechanism. Within the class of policy rules such as Equation (11), I ask whether optimal monetary policy is more aggressive on inflation under learning than under RE. I also investigate how the choice of Taylor-rule coefficients affects the anchoring mechanism. Subsequently, I explore whether the monetary authority could do better than with a Taylor rule. Does an alternative rule exist that can implement the optimal plan for the observables? And can such an alternative rule be implemented in practice?

I first simulate a calibrated version of the model. This serves to build intuition about the model dynamics, in particular regarding the interplay between monetary policy and the anchoring mechanism. I then proceed to the analytical characterization of optimal monetary policy.

4.1 Simulations

4.1.1 Calibration

In this section I simulate the rational expectations and learning versions of the model and compute the optimal Taylor rule coefficients numerically.[11] Table 1 summarizes the calibrated parameter values. For most of the parameters, I assign values commonly used in the macroeconomic literature. In particular, I follow Woodford (2011)'s calibration. For this section, I shut off the monetary policy parameters λ_i, λ_x and ψ_x in order to focus on the role of inflation in the central bank's problem and thus on the optimal choice of inflation aggressiveness, ψ_π.

The learning parameters $\bar{g}, \tilde{\theta}$ and $\tilde{\kappa}$ require some discussion. While the choice of $\tilde{\kappa}$ only matters for the smoothness of the endogenous gain choice and thus can be set relatively freely, the threshold $\tilde{\theta}$ has more bearing on the behavior of the model. Intuitively, the higher $\tilde{\theta}$, the more forecast error volatility agents in the economy are willing to tolerate before switching to a constant gain. Experimenting with different values reveals that once $\tilde{\theta}$ is higher than a particular threshold, expectations are anchored for any value of ψ_π. Analogously if $\tilde{\theta}$ is below a lower threshold, expectations are always unanchored regardless of the value of ψ_π. My choice of $\tilde{\theta} = 2.5$ is thus motivated by assigning a value for which the comparative static of anchoring with respect to ψ_π is meaningful.

The choice of $\bar{g}$ is far from innocent as it has considerable implications for model dynamics. In particular, as I address in the Introduction, constant gain learning models have a tendency to produce impulse responses that exhibit damped oscillations. The reason is that under an adaptive learning framework, forecast errors following an impulse are oscillatory. In fact, the higher the learning gain, the higher the amplitude of forecast error oscillations. The oscillations can even become explosive if the gain is high enough.

[11] To concentrate on the intuition, in this section I implement the learning algorithm such that only the constant is learned. The general formulation has qualitatively similar features but is more impacted by small-sample concerns prevalent in simulations.

Table 1: Calibrated parameters

β	0.99	stochastic discount factor
σ	1	intertemporal elasticity of substitution
α	0.5	Calvo probability of not adjusting prices
ψ_π	1.5	coefficient of inflation in Taylor rule
ψ_x	0	coefficient of the output gap in Taylor rule
$\bar{g}$	0.145	value of the constant gain
$\tilde{\theta}$	2.5	threshold value for criterion of endogenous gain choice
$\tilde{\kappa}$	0.2	scaling parameter of gain for forecast error variance estimation
ρ_r	0	persistence of natural rate shock
ρ_i	0.6	persistence of monetary policy shock
ρ_u	0	persistence of cost-push shock
σ_i	1	standard deviation of natural rate shock
σ_r	1	standard deviation of monetary policy shock
σ_u	1	standard deviation of cost-push shock
λ_x	0	weight on the output gap in central bank loss
λ_i	0	weight on the interest rate in central bank loss

Unfortunately, the model gives no guidance on the appropriate value for $\bar{g}$.[12] I thus turn to the admittedly thin literature on estimating learning gains. I assign the value 0.145, obtained by Carvalho et al. (2019), to my knowledge the only study to estimate the value of the constant gain for an endogenous gain model. It has to be observed, however, that this is a significantly higher value than what was found in the literature estimating gains for constant gain learning. Branch and Evans (2006)'s number of 0.062 is quite close to Eusepi et al. (2018a)'s estimate of 0.05, while Milani (2007) finds an even lower number of 0.0183. Studies that use calibrated gains such as Williams (2003) or Orphanides and Williams (2005) tend to experiment with a range of values in the $[0.01, 0.1]$ interval. The value of 0.05 seems to have attained particular prominence, but also much lower numbers have been used, such as 0.002 in Eusepi and Preston (2011). I speculate that Carvalho et al. (2019)'s estimate is so large relative to other estimates because they estimate a switching-gain model, while the rest of the estimates come from constant gain specifications. In an endogenous gain specification, data needs to assign a higher value to the constant gain parameter to rationalize the same average gain in the time series. With these caveats in mind, I adopt Carvalho et al. (2019)'s value.

4.1.2 Simulated dynamics

Having thus assigned values to the parameters, I turn to the model's behavior. Table 2 presents an overview of the optimal Taylor rule coefficient ψ_π, obtained via grid-search, for the rational expectations and anchoring models. The table also compares the baseline parameterization with several alternatives. One notices that if the central bank has no concern to stabilize the output gap ($\lambda_x = 0$) or the nominal interest rate ($\lambda_i = 0$), $\psi_\pi^{*,RE}$ is infinity. This is because if the central bank suffers no loss upon output

[12] The analogy of the Kalman gain from the Kalman filter does not prove helpful either because it requires a steady state forecast error variance matrix which is not available in a learning context.

variation, then the fact that the divine coincidence doesn't hold does not pose a problem. Similarly, if the monetary authority is willing to allow the nominal interest rate to fluctuate vastly in order to stabilize inflation, this also allows the central bank to be infinitely aggressive on inflation.

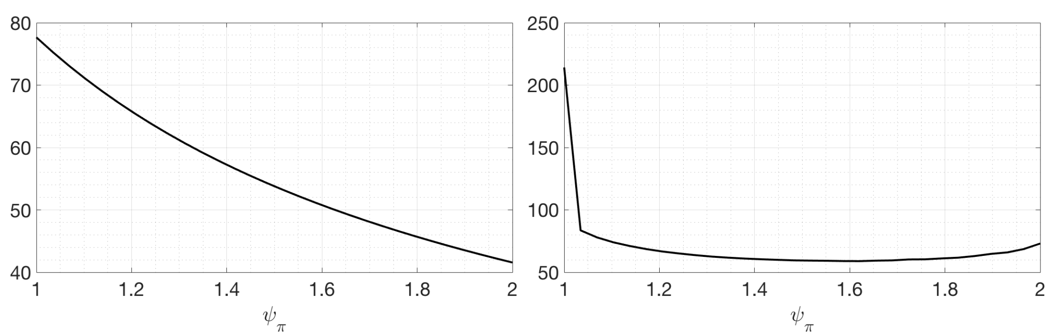
Table 2: Optimal coefficient on inflation, RE against learning for alternative parameters

	$\psi_{\pi}^{*,RE}$	$\psi_{\pi}^{*,learn}$
Baseline	∞	1.6243
$\lambda_x = 1$	2.1042	1.0571
$\lambda_i = 1$	1.1	1.0978

The main observation however is that ψ_π is always lower for the anchoring model than for the RE model. This is reinforced in Fig. 3 which plots the central bank's loss in the RE and learning models for various values of ψ_π but otherwise the baseline specification. The message is clear: while for rational expectations, the loss is strictly decreasing in ψ_π, this is not the case for the anchoring model. Why does the anchoring expectation formation induce the central bank to optimally be less aggressive on inflation?

The answer is that the anchoring mechanism introduces a novel tradeoff for the central bank. On the one hand, as we will see shortly in Fig. 4, having unanchored expectations increases the volatility of the observables. This results in the central bank wishing to anchor expectations. As Fig. 5 makes clear, this requires raising ψ_π. But in an environment where agents know the Taylor rule, a higher coefficient on inflation will lead to initially higher volatility due to the agents anticipating the endogenous responses of the nominal interest rate far in the future.

Figure 3: Central bank loss function as a function of ψ_π

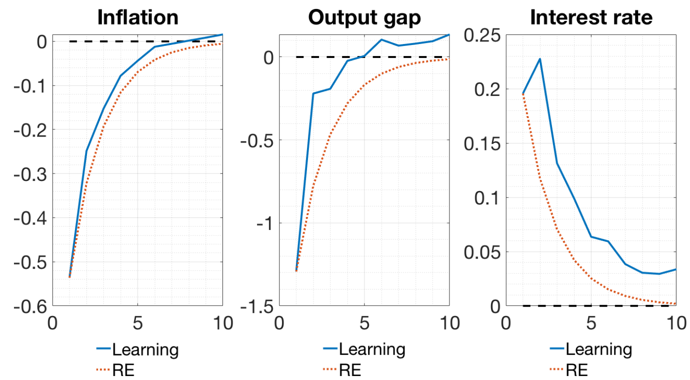


(a) RE

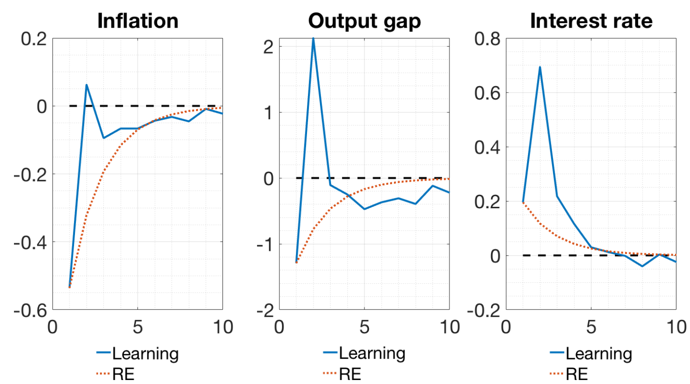
(b) Learning

To understand what's going on in the model in detail, consider Fig. 4, portraying the impulse responses of the model after a contractionary monetary policy shock. The red dashed lines show the responses of the observables in the rational expectations version of the model. The blue lines show the responses in the learning model, on panel (a) conditional on expectations being anchored when the shock hits, on panel (b) being unanchored upon the arrival of the shock.

Figure 4: Impulse responses after a contractionary monetary policy shock



(a) RE against learning, expectations anchored



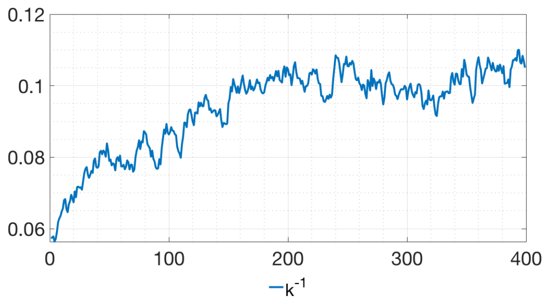
(b) RE against learning, expectations unanchored

Shock imposed at $t = 25$ of a sample length of $T = 400$ (with 100 initial burn-in periods), cross-sectional average with a cross-section size of $N = 100$. For the rest of the paper, I keep these simulation values unless otherwise stated. For the learning model, the remark refers to whether expectations are anchored at the time the shock hits.

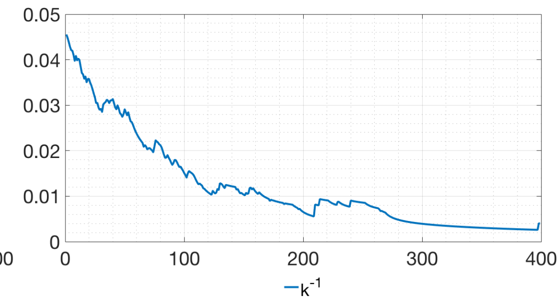
Not only do the impulse responses show the usual behavior of learning models - dampened responses and increased persistence. More importantly, responses differ strongly depending on whether expectations are anchored or not when the shock hits. In particular, if expectations are anchored, responses are closer to rational expectations than when expectations are unanchored. Moreover, when expectations are unanchored, the endogenous responses of the observables become much more volatile, indeed, oscillatory. This makes intuitive sense: expectations being unanchored reflects the fact that firms and households are confronted with an environment that does not line up with their currently held perceived law of motion. They thus believe that a structural change has occurred and are therefore revising their expectations. Expectations are therefore fluctuating strongly, and as they feed back to the observables, the latter inherit their volatility.

Therefore, to avoid the volatility that results from unanchored expectations, the central bank wishes to anchor expectations. What choice of ψ_π will do the job? Fig. 5 provides the answer. The figure shows the cross-sectional average of inverse gains that result when ψ_π takes on different values. Clearly,

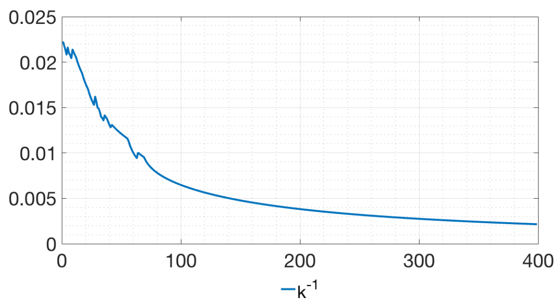
Figure 5: Cross-sectional average inverse gains for various values of ψ_π



(a) $\psi_\pi = 1.01$



(b) $\psi_\pi = 1.5$



(c) $\psi_\pi = 2$

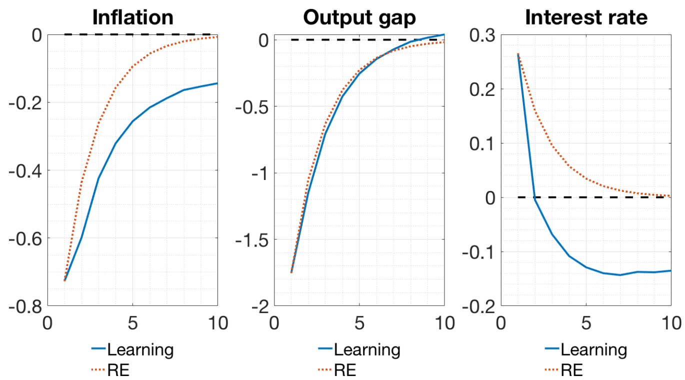
a higher ψ_π results in lower and decreasing gains.[13] Thus a central bank aiming to anchor expectations needs to employ a high ψ_π. This is also intuitive. A more aggressive central bank can signal to the public that it is determined to achieve the announced inflation target. Thus agents can rest assured that their believed inflation target is indeed the one the central bank can and will implement. Whenever ψ_π is low, however, the central bank is willing to tolerate bigger deviations from the target. This opens the door to speculation about whether the central bank is really committed to the target. In this case, deviations from the believed target of the same magnitude can cast doubt on the central bank's commitment, so that agents decide to monitor recent data closely to learn the seemingly shifting average value of inflation.

But if unanchored expectations cause heightened volatility, and being aggressive on inflation is able to anchor expectations, why does the monetary authority optimally choose a lower value for ψ_π than under rational expectations? The reason is that conditional on being unanchored, a higher ψ_π actually causes higher volatility than a lower one. This can be seen on Fig. 6 which depicts the same impulse responses to a contractionary monetary policy shock as Fig. 4, focusing however only on responses conditional on expectations being unanchored upon the shock. It shows these responses for three different values of ψ_π.

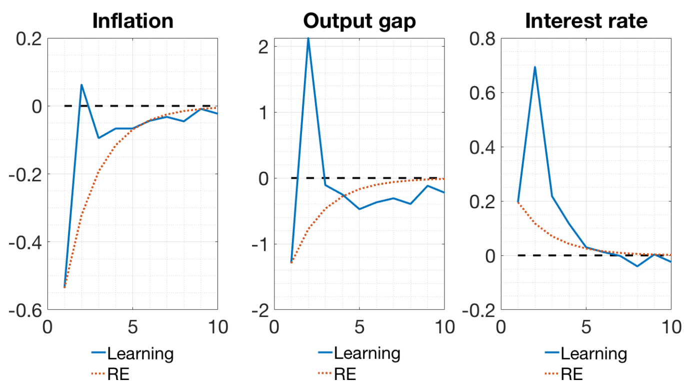
As the figure shows, a high ψ_π leads to more volatility than a low one does. The intuition is a little subtle. Since expectations are unanchored, they are also volatile. This implies that inflation far ahead

[13] As I remark in Section 3, if one uses the anchoring criterion of Carvalho et al. (2019), this conclusion is overturned.

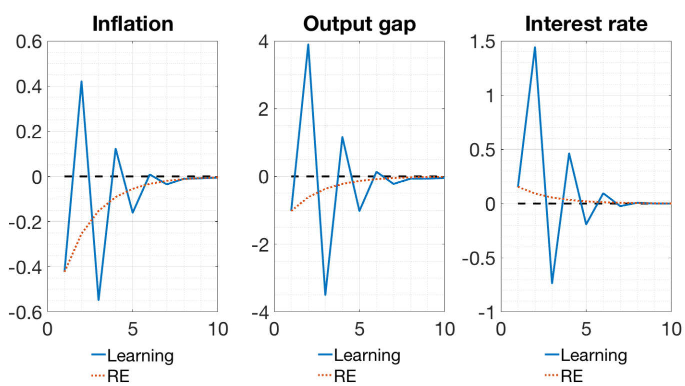
Figure 6: Impulse responses for unanchored expectations for various values of ψ_π



(a) $\psi_\pi = 1.01$



(b) $\psi_\pi = 1.5$



(c) $\psi_\pi = 2$

in the future is expected to fluctuate strongly. Since agents know the Taylor rule, this also means that they expect the nominal interest rate far in the future to respond. The more aggressive the central bank, the stronger an interest rate response will the agents expect. This however feeds back into current output gaps and thus inflation. Higher overall volatility is the result.

Unexpectedly, the model dynamics here echo the predictions of Ball (1994). But the underlying channels are quite different. Ball (1994) observes that, contrary to conventional wisdom, rational expectations New Keynesian models imply expansionary disinflations. To reconcile this model feature with data pointing to the costliness of disinflations, he concludes that central bank announcements must suffer from credibility issues.

Note that in the present context, when expectations are anchored (panel (a) of Fig. 4), impulse responses do not exhibit this feature. However, when expectations are unanchored (panel (b) of Fig. 4), impulse responses look exactly as Ball (1994) predicts: we obtain an expansionary disinflation.

The reason this is happening is that when agents know the Taylor rule, long-horizon expectations of the interest rate move in tandem with the same expectations of inflation in the far future. A current disinflation lowers long-horizon inflation expectations, leading the public to expect low interest rates far out in the future. Through the NKIS-curve (Equation 9), this stimulates current output.[14] But the absence of the "Ball-effect" from the anchored expectations impulse responses indicates that the channel is only operational when expectations are moving sufficiently. Thus I arrive at a different conclusion than Ball (1994); instead of credibility issues, it is anchored expectations that are responsible for the absence of expansionary disinflations of the type seen on Fig 4, panel (b).

Thus the presence of an expectation formation that allows for the anchoring and unanchoring of expectations introduces several unexpected new features to the New Keynesian model. Unsurprisingly, it is desirable for the central bank to anchor expectations. It is also intuitive that being aggressive on inflation helps to anchor expectations. However, less intuitive is the fact that the optimal degree of aggressiveness on inflation is lower than under rational expectations. This has to do with the heightened volatility of the expectations process when ψ_π is high. A higher ψ_π increases the response of future nominal interest rate expectations, thus raising the feedback from expectations to current observables. Thus the central bank faces an intertemporal tradeoff: to reduce volatility in the long-run, it seeks to anchor expectations. However, the price the bank has to pay in order to get expectations anchored is higher short-run volatility. Thus, the monetary authority trades off the short-run cost with the long-run benefit of anchoring expectations.

It is immediate that if the monetary authority is allowed to reoptimize at a particular date τ, it will choose Taylor-rule coefficients depending on whether expectations are anchored or not. More precisely, the current value of the gain determines the short-run costs of anchoring expectations. As the gain varies over time, it follows that time-zero optimal policy will be time-inconsistent.

As emphasized at the end of Section 2, this result bears resemblance to that of Lubik and Matthes (2016) who show that central bank learning and data misperceptions lead to time-varying Taylor-rule coefficients. But it is important to point out that this result would not arise in a setting where the public sector is learning with a constant or a decreasing gain specification. The reason is that in such a model, the monetary authority would not find it optimal to reset its coefficient at different points in time because this choice would have no impact on the learning process. With constant gain learning, the central bank would just have to accept heightened volatility and accordingly set a lower ψ_π than under rational expectations. With decreasing gain learning, the bank would have to accept higher

[14] The extension in which the public has to learn the Taylor rule is interesting in this regard. As expected, the Ball-type disinflationary boom does not initially show up in impulses responses obtained in that extension. However, as the agents are learning the Taylor rule, the expansionary disinflation slowly reemerges in the impulse responses.

volatility initially as the learning process converges to RE. Either case can be thought of as simply scaling the path of volatility, but with the central bank having no influence on the slope of the path. Under anchoring, the authority can exert direct influence on the volatility path, and its options of doing so are more or less costly depending on whether expectations are currently anchored or not. In this manner, the anchoring mechanism not only introduces a novel tradeoff for policy, but also a new mechanism underlying the time-inconsistency of time-zero optimal commitment.

Already without the analytical solutions for policy at hand, one can thus use the model to provide fresh interpretations of the current, 2019-2020 monetary episode in the US. The model suggests that the downward drift of the public's long-horizon inflation expectations is a sign of expectations threatening to become unanchored. The Fed seems to have internalized that anchoring expectations will be much more costly once they are truly unanchored, and thus moved swiftly to lower interest rates in the fall of 2019. In so doing, it communicated to the public its commitment to the 2% target. Through the lens of my model, then, we can interpret this counterintuitive easing at the height of an expansion as an effort on the Fed's part to keep expectations anchored.

5 Conclusion

Central bankers frequently voice a concern to anchor expectations. The fact that rational expectations New Keynesian models have nothing to say about this aspect of monetary is a gap in the macroeconomic literature. Absent a theory of anchored expectations, it is difficult for macroeconomists to understand periods where central banks are clearly off the Taylor rule. The current stance of US monetary policy is an example of such an episode: the business cycle calls for monetary tightening, yet inflation lags below the target and if anything, the Fed is expansionary. My work suggests that the Fed reads the downward drift of long-run inflation expectations as a threat that expectations may become unanchored. In order to prevent that from happening, the Fed therefore is anxious to signal that it is determined to achieve its 2% inflation target. Its expansionary actions are thus not intended to stimulate an already tight labor market; instead the Fed's present objective is to keep expectations anchored.

References

Ball, L. (1994). Credible disinflation with staggered price-setting. *The American Economic Review 84*(1), 282–289.

Branch, W. A. and G. W. Evans (2006). A simple recursive forecasting model. *Economics Letters 91*(2), 158–166.

Brown, R. L., J. Durbin, and J. M. Evans (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (Methodological) 37*(2), 149–163.

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.

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

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 (2006). Monetary policy, expectations and commitment. *Scandinavian Journal of Economics 108*(1), 15–38.

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

Gaspar, V., F. Smets, and D. Vestin (2006). Adaptive learning, persistence, and optimal monetary policy. *Journal of the European Economic Association 4*(2-3), 376–385.

Lubik, T. A. and C. Matthes (2016). Indeterminacy and learning: An analysis of monetary policy in the great inflation. *Journal of Monetary Economics 82*, 85 – 106.

Lütkepohl, H. (2013). *Introduction to multiple time series analysis.* Springer Science & Business Media.

Marcet, A. and J. P. Nicolini (2003). Recurrent hyperinflations and learning. *American Economic Review 93*(5), 1476–1498.

Milani, F. (2007). Expectations, learning and macroeconomic persistence. *Journal of monetary Economics 54*(7), 2065–2082.

Milani, F. (2014). Learning and time-varying macroeconomic volatility. *Journal of Economic Dynamics and Control 47*, 94–114.

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

Orphanides, A. and J. C. Williams (2005). The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations. *Journal of Economic dynamics and control 29*(11), 1927–1950.

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

Preston, B. (2008). Adaptive learning and the use of forecasts in monetary policy. *Journal of Economic Dynamics and Control 32*(11), 3661 – 3681.

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

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

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

A Coefficient matrices in NK model

$$A_a = \begin{pmatrix} g_{\pi a} \\ g_{xa} \\ \psi_\pi g_{\pi a} + \psi_x g_{xa} \end{pmatrix} \quad A_b = \begin{pmatrix} g_{\pi b} \\ g_{xb} \\ \psi_\pi g_{\pi b} + \psi_x g_{xb} \end{pmatrix} \quad A_s = \begin{pmatrix} g_{\pi s} \\ g_{xs} \\ \psi_\pi g_{\pi s} + \psi_x g_{xs} + \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \end{pmatrix} \tag{31}$$

$$g_{\pi a} = (1 - \frac{\kappa\sigma\psi_\pi}{w}) \begin{bmatrix} (1-\alpha)\beta, \kappa\alpha\beta, 0 \end{bmatrix} \tag{32}$$

$$g_{xa} = \frac{-\sigma\psi_\pi}{w} \begin{bmatrix} (1-\alpha)\beta, \kappa\alpha\beta, 0 \end{bmatrix} \tag{33}$$

$$g_{\pi b} = \frac{\kappa}{w} \begin{bmatrix} \sigma(1-\beta\psi_\pi), (1-\beta-\beta\sigma\psi_x, 0 \end{bmatrix} \tag{34}$$

$$g_{xb} = \frac{1}{w} \begin{bmatrix} \sigma(1-\beta\psi_\pi), (1-\beta-\beta\sigma\psi_x, 0 \end{bmatrix} \tag{35}$$

$$g_{\pi s} = (1 - \frac{\kappa\sigma\psi_\pi}{w}) \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} (I_3 - \alpha\beta P)^{-1} - \frac{\kappa\sigma}{w} \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} (I_3 - \beta P)^{-1} \tag{36}$$

$$g_{xs} = \frac{-\sigma\psi_\pi}{w} \begin{bmatrix} 0 & 0 & 1 \end{bmatrix} (I_3 - \alpha\beta P)^{-1} - \frac{\sigma}{w} \begin{bmatrix} -1 & 1 & 0 \end{bmatrix} (I_3 - \beta P)^{-1} \tag{37}$$

$$w = 1 + \sigma\psi_x + \kappa\sigma\psi_\pi \tag{38}$$

B The observation matrix for learning

$$g^l = \begin{bmatrix} F & G \end{bmatrix} \tag{39}$$

with

$$F = \left(A_a \frac{1}{1-\alpha\beta} + A_b \frac{1}{1-\beta} \right) a_{t-1} \tag{40}$$

$$G = A_a b_{t-1} \left(I_3 - \alpha\beta h \right)^{-1} + A_b b_{t-1} \left(I_3 - \beta h \right)^{-1} + A_s \tag{41}$$