Stockholm School of Economics
Department of Economics
5350 Master's Thesis in Economics
Spring 2026

When Knowing Less Fits More: Comparing Rational Expectations and Adaptive Learning in DSGE Models

Hugo Hennel (42867) & Lukas Jung (42850)

Abstract

A central feature of modern macroeconomics is that agents form expectations about future economic variables under rational expectations (RE). However, evidence suggests agents form expectations based on more limited information than RE assumes. This thesis therefore compares the empirical performance of a DSGE model under RE against one with adaptive learning (AL). Under AL, agents form expectations using a univariate AR(2) forecasting model and update their beliefs recursively using a Kalman Filter. We estimate both the RE and AL versions of the model using Bayesian methods on quarterly euro area data from the Area-Wide Model Database covering 1970–2019. Our results indicate that the AL specification fits the euro area data substantially better than the RE benchmark. Adaptive learning generates endogenous persistence by making agents more backward-looking, reducing the model's reliance on mechanical price stickiness and yielding more realistic propagation of monetary policy shocks. Interestingly, the estimated degree of active belief updating is small, suggesting that the empirical gains stem primarily from the bounded rationality embedded in rule-of-thumb AR(2) forecasting structure rather than from the Kalman Filter learning itself.

Keywords: DSGE, adaptive learning, rational expectations, Kalman Filter learning, New Keynesian model, monetary policy, euro area

JEL: E30, E52, E71, D84

Supervisor:	Álvaro Jañez Garcia
Date submitted:	2026-05-18
Date examined:	2026-05-27
Discussants:	Pippa Johns (25399) & Jan David Wechselberger (42871)
Examiner:	Sampreet Singh Goraya

Acknowledgments

We thank our supervisor Álvaro Jañez Garcia at the Stockholm School of Economics for his guidance and support throughout the thesis writing.

We are also grateful to Anders Warne at the European Central Bank for generously taking the time to answer our many questions with great care and detail, which has been invaluable to our work.

Finally, we thank Arne Gieseck, formerly at the European Central Bank, for his thoughtful feedback and comments.

AI Declaration

For this thesis, generative AI has been used to enhance the grammar, phrasing, and readability of the written text, as well as to support the Python coding for the local projections benchmark through debugging and syntax corrections. For all tasks, the Claude Sonnet 4.6 model has been utilised.

Contents

List of Figures

List of Tables

1 Introduction

A central feature of modern macroeconomic models is that agents forecast future economic variables under rational expectations (Sargent, 1993; Massaro, 2013). Under this assumption, all agents possess full information about the current state of the economy and understand the structural laws governing it, allowing them to form the most accurate possible predictions about future variables. The rational expectations (RE) assumption has proven a successful and tractable modelling choice, and is widely adopted for its analytical convenience and empirical fit. However, a substantial body of research suggests that the assumption might be too strong, as evidence shows that households and firms systematically fail to predict key macroeconomic variables (Carroll, 2003; Coibion and Gorodnichenko, 2015). Relaxing this assumption may therefore be appealing on two grounds. First, modelling agents as forming expectations in a more realistic way might yield a more faithful description of the economy in which expectations are taken seriously instead of being imposed by convenience. Second, this added realism may improve the model's fit to data as well as its forecasting performance. An alternative way of expectation formation is to assume that agents are unaware of the true structure of the economy and instead adaptively learn about their economic environment and how best to forecast it. A growing literature (Slobodyan and Wouters, 2012; Rychalovska et al., 2025; Warne, 2026) shows that adaptive learning (AL) constitutes a more credible description of expectation formation and, crucially, that this modelling choice carries empirical value, as it improves several measures on in- and out-of-sample fit compared to the rational expectations benchmark. Motivated by these findings, this thesis estimates a small-scale New Keynesian Dynamic Stochastic General Equilibrium (NK DSGE) model on euro area data, compares the empirical performance of adaptive learning against rational expectations, and analyses how macroeconomic dynamics change under learning, with the aim of identifying which features of adaptive learning drive its potential empirical advantage.

To carry out this analysis, we build on the DSGE model developed by Del Negro and Schorfheide (2013), which is a small-scale version of the influential medium-scale framework of Smets and Wouters (2007). Our model retains the core features of a modern macroeconomic model, having optimising households, intermediate goods firms with Calvo-style price stickiness, and a central bank following a Taylor rule, while remaining sufficiently transparent so that differences in model outcomes can more easily be attributed to the expectation formation mechanism rather than to other modelling choices. After setting up the framework, we deviate from the standard implementation in Del Negro and Schorfheide (2013) by modifying the way agents form expectations. In the benchmark case, agents have rational expectations and know the true model of the economy. In the alternative case, agents are adaptive learners who form expectations using a parsimonious forecasting model which they recursively update via a Kalman Filter whenever their forecasts prove incorrect.

We estimate the model on quarterly euro area data covering 1970-2019 using Bayesian methods, following the influential approach of Smets and Wouters (2003, 2007). Rather than estimation by matching specific moments in the data, this approach fits the model directly to macroeconomic time series. It also allows existing information on the model's parameter values to be incorporated through the use of priors. The euro area is a suitable subject of study, since treating it as a closed economy is a reasonable and common approximation at the aggregate level (Smets and Wouters, 2003), as internal trade integration reduces the relative importance of external dynamics and allows us to abstract from open economy considerations. The sample begins as early as the data permit and ends in 2019, a choice made to exclude the extreme volatility of the COVID-19 pandemic, that our small-scale model is not equipped to handle and which would confound the estimation of structural parameters. Importantly, the sample spans episodes of considerable macroeconomic heterogeneity, including the OPEC oil shock, the great moderation, and the global financial crisis, allowing us to study the role of adaptive learning in an economic environment that exhibit structural variation over time and giving the agents the possibility to adaptively learn about changing economic landscapes.

Our results are broadly consistent with the adaptive learning literature. In-sample fit improves under AL on both metrics we consider: the log marginal likelihood, replicating the findings of prior work (Milani, 2011; Dizioli and Wang, 2024; Rychalovska et al., 2025), and the replication of empirically identified impulse responses to monetary policy shocks benchmarked against high-frequency local projections, which to our knowledge is a novel contribution. For out-of-sample forecasting performance, the AL model is competitive with the rational expectations benchmark. When delving into dynamics, we find that adaptive learning introduces endogenous persistence by making agents more backward-looking when forming expectations. This reduces the model's reliance on the excessive mechanical persistence devices that RE models typically require, improving both the marginal likelihood and the realism of monetary policy transmission dynamics. A second notable finding is that we estimate the optimal degree of learning to be very limited, suggesting that the empirical gains from the AL specification are not primarily driven by active belief updating, but by other ways that the learning set-up changes how agents form expectations. This finding challenges previous conclusions in the adaptive learning literature that emphasize the importance of active learning for increased model fit (Milani, 2011; Slobodyan and Wouters, 2012).

The remainder of the paper proceeds as follows. Section 2 reviews the literature on NK DSGE models and adaptive learning, and positions our contribution relative to existing work. Section 3 derives the small-scale NK DSGE model and describes how rational expectations and adaptive learning are implemented. Section 4 presents the Bayesian estimation approach, the data, and the calibration and priors. Section 5 presents and interprets the Bayesian estimation results for key application areas of DSGE models. Section 6 discusses those results and highlights our three main conceptual findings. Section 7 concludes.

2 Literature review

New Keynesian DSGE models emerged from the *new neoclassical synthesis* of the 1990s (Goodfriend and King, 1997), combining neoclassical real business cycle theory with nominal rigidities such as Calvo-style price stickiness and wage frictions. The resulting framework features optimising households and firms, a central bank following a policy rule, and stochastic shock processes that drive business cycle fluctuations, which are all derived from explicit microeconomic foundations. Among the models in this tradition, the medium-scale rational expectations framework of Smets and Wouters (2007) has become particularly influential. It paved the way for explicitly estimating DSGE models based on macroeconomic time series using Bayesian methods and proved it was capable of generating a time series fit comparable to those of less restrictive vector autoregressions (Herbst and Schorfheide, 2016). Following this, Del Negro and Schorfheide (2013) derived a small-scale version of the seminal Smets and Wouters (2007) model, with the aim of making the model more tractable while retaining most of its good time series fit and forecasting ability. The difference in scales of the two models is reflected by the number of time series they require. The original model of Smets and Wouters (2007) requires seven, while the small-scale version suffices with three. In our paper, we deliberately choose a smaller model, as our main interest lies not in achieving the best possible absolute model fit, which would warrant a more complex framework, but in isolating the relative empirical performance of two expectation types and studying how the model dynamics depend on them. In this regard, the framework of Del Negro and Schorfheide (2013) is particularly appealing given its relative simplicity, while still being closely related to the widely-used model of Smets and Wouters (2007). To our knowledge, this makes us the first to study the dynamics of adaptive learning in the small-scale framework of Del Negro and Schorfheide (2013).

The rational expectations hypothesis has been the dominant expectation formation mechanism since the early days of DSGE modelling. Its main advantages being that it yields a tractable linear solution, imposes cross-equation restrictions on model dynamics, which improves identification, and enforces internal consistency on agents' expectations which guarantees that the ex ante solution to their intertemporal problem also turns out to be optimal when implemented (Evans and Honkapohja, 2001; Slobodyan and Wouters, 2012).

Despite these advantages, RE has received substantial criticism. First, the assumption that agents possess full knowledge of the structural model and use it to form unbiased forecasts is difficult to defend empirically, as survey evidence consistently documents that households and firms systematically fail to predict future inflation (Carroll, 2003; Coibion and Gorodnichenko, 2015). Additionally, RE models have been found to require unrealistic amounts of frictions and implausibly high persistence in exogenous shock processes in order to match the patterns found in macroeconomic time series (Chari et al., 2009; Rychalovska et al., 2025). Cogley and Sbordone (2008) and others have argued that the excessive reliance on these model features under RE may in fact reflect misspecified expectation dynamics. This has motivated a literature exploring less over-idealized specifications of how agents form expectations, with adaptive learning as one of the frontrunners.[1] Pioneered by Marcet and Sargent (1989), adaptive learning has been shown to produce expectations more in line with the survey evidence, most notably by exhibiting serial correlation and sluggish responses to shocks (Carroll, 2003; Coibion and Gorodnichenko, 2015; Dizioli and Wang, 2024). More recent research have demonstrated that adaptive learning can also be a source of persistence in itself, and thereby render some of the unrealistic exogenous shock processes and frictions redundant. For example, in a small-scale model, Milani (2007) showed that adaptive learning was capable of replacing the unreasonable levels of habit formation and price indexation used in RE models, while Slobodyan and Wouters (2012) showed that adaptive learning decreased the need for high persistence in price and wage mark-up shocks. As such, AL is capable of addressing the two mentioned objections directed toward RE and may therefore have the potential of bringing DSGE models closer to reality.

In the literature, there are two dominating approaches to adaptive learning, both

[1] Alternative approaches to relaxing the rational expectations assumption include incorporating survey expectations directly into the estimation (Carvalho et al., 2023), allowing for heterogeneous expectation formation (Poledna et al., 2023), and modelling imperfect information with endogenous learning gains (Gáti, 2023). Rychalovska et al. (2025) provide a recent comprehensive comparison of such alternatives, finding that the choice of expectation formation mechanism significantly affects both in-sample fit and forecasting performance, consistent with the earlier findings of Milani and Rajbhandari (2012).

of which update agents' beliefs recursively in response to forecast errors. Orphanides and Williams (2005) and Milani (2007, 2011) used constant gain (CG) learning, which is the more parsimonious of the two as it assigns the same informational value to all new forecast errors. The other, more sophisticated approach, put forward by Slobodyan and Wouters (2012), is Kalman filter learning. It allows for the informational value to endogenously increase during periods when the underlying data generating process is changing. In their paper, Slobodyan and Wouters (2012) implement Kalman filter learning in the medium-scale model from Smets and Wouters (2007) and show that introducing learning helps explain the dynamics in US data better than with rational expectations, while also yielding a better fit than CG learning.

Motivated by this, we implement the richer adaptive learning model from Slobodyan and Wouters (2012) into the small-scale DSGE model from Del Negro and Schorfheide (2013). We follow Milani (2011) in keeping the underlying DSGE model simple so that the contribution of adaptive learning is not confounded by other model features, while extending their approach by adopting a more sophisticated learning mechanism. This positions our paper at the intersection of Milani's (2007; 2011) small-scale approach and Slobodyan and Wouters' (2012) learning approach, and allows us to examine whether the conclusions from Slobodyan and Wouters (2012), that Kalman filter learning plays an important role in business cycle dynamics, applies when learning is implemented in a simpler, small-scale model where there are fewer macroeconomic variables for agents to learn about.

3 Model description

In this section, we present and derive the small-scale New Keynesian DSGE model that forms the basis of our analysis. The model is based on the seminal DSGE model by Smets and Wouters (2007) which uses wage and price frictions and seven exogenous shocks. Del Negro and Schorfheide (2013) simplify the framework by ignoring capital accumulation, wage stickiness, and habit formation of the consumers, and reduce the number of exogenous shocks down to three. Thereby, the small-scale version retains the core macroeconomic dynamics and allows us to focus on the different expectation formations.

As is standard in the DSGE literature, the households maximize their intertemporal utility by deciding on their consumption and labour hours worked. They are able to smooth consumption over time by borrowing and saving through a one period bond. Furthermore, it is assumed that households are the owners of the firms and thus receive dividends.

On the production side, there are two types of firms. Intermediate good firms take labour as an input and have market power when selling their good. They face sticky prices as described by Calvo (1983) such that only a share of intermediate firms is able to readjust their prices each period. Firms that cannot re-optimize their prices instead have them partially indexed to a weighted average of past and steady-state inflation. Final good firms then use the intermediate good as their only input to produce the final good. In contrast to intermediate firms, they operate under perfect competition by taking prices as given.

Finally, the model is closed by introducing a central bank that sets the nominal interest rate following a monetary policy rule. To keep the economy stable, they follow a Taylor rule to react sufficiently to deviations from steady state in inflation and output gap. The model is then solved by deriving optimal conditions for all parties and log-linearising them around the steady state.

To understand the macroeconomic dynamics of different shocks to our model, the model incorporates three types of exogenous shocks. First, monetary policy shocks enter the monetary policy rule and thus generate a disturbance to households' intertemporal consumption-smoothing decisions. Second, government spending shocks enter the aggregate resource constraint of the economy and affect output by generating additional demand. Finally, productivity shocks enter the intermediate firm's production function, altering how labour translates into output and thereby affecting the level of output in the economy.

After having set up the DSGE model, we treat it as a laboratory for examining the model dynamics under different ways of letting agents form their expectations of future variables. We compare the standard case of rational expectations to the alternative of adaptive learning.

To shortly summarize the two cases, rational expectations assumes that agents have full information about the model which they inhabit. Essentially, this allows them to form optimal expectations about the future states of the economy, with the only unknowns being the realizations of future exogenous shocks. In our thesis, adaptive learning is modelled through Kalman filter learning. In its essence, this means that agents are unaware of the correct forecasting model for output, consumption, and inflation, and instead use a simple autoregressive forecasting rule, which they iteratively revise as new data arrive. The revision in each period is increasing in their forecast error, meaning that the larger the forecast error, the larger the update. The weight placed on each new forecast error is not fixed but adjusts endogenously: agents update their beliefs more aggressively when their forecasting model appears to no longer provide a good fit to data.

This makes the learning mechanism particularly responsive during periods of structural change, such as the shift in inflation dynamics following the 1970s oil shocks. To avoid the kind of unstable model dynamics that unhindered learning could give rise to, the learning is disciplined by another autoregressive process that anchors the coefficients in the autoregressive forecasting models around some pre-specified baseline values. Both the configuration for RE and AL are described in detail in section 3.2.

3.1 Deriving the general version of model

The following section gives a brief overview of the derivation of the small-scale DSGE model, developed by Del Negro and Schorfheide (2013). Readers interested in the full derivations of the household and firm optimisation problems, the the steady-state relationships, and log-linearisation procedure are referred to Appendix A, where all steps are carried out in detail.

3.1.1 The household sector

Each household $j \in [0, 1]$ chooses consumption $C_t(j)$, hours worked $L_t(j)$, and bond holdings $B_t(j)$ to maximise the objective function given by:

$$E_t \sum_{s=0}^{\infty} \beta^s \left\{ \frac{1}{1-\sigma_c} (C_{t+s}(j))^{1-\sigma_c} \exp\left(\frac{\sigma_c - 1}{1+\sigma_l} L_{t+s}(j)^{1+\sigma_l} \right) \right\} \tag{1}$$

where the household's utility is increasing in consumption $C_{t+s}(j)$, decreasing in hours worked $L_{t+s}(j)$, and notably, inseparable in the two arguments. The household maximizes the intertemporal utility function subject to the following budget constraint:

$$C_{t+s}(j) + \frac{B_{t+s}(j)}{R_{t+s}P_{t+s}} \leq \frac{B_{t+s-1}(j)}{P_{t+s}} + \frac{W^h_{t+s}(j)L_{t+s}(j)}{P_{t+s}} + \frac{Div_{t+s}}{P_{t+s}} \tag{2}$$

This indicates that the period's expenditures, consisting of consumption $C_t(j)$ and purchased/borrowed bonds $B_t(j)$ must be equal or smaller than the available resources. Here, $B_t(j)$ are deflated and discounted by the policy rate (nominal return) R_t. The available resources consist of the bond holdings from last period, the income from labour $(W^h_{t+s}(j)L_{t+s}(j))$ and dividends (Div_{t+s}) - all values indicated in real terms.

Solving this constrained optimization problem yields the households' labour supply (3), that represents the intratemporal optimality condition, and the Euler equation (4), that states the intertemporal optimality condition.

$$C_t L_t^{\sigma_l} = \frac{W_t}{P_t} \tag{3}$$

$$E_t[{C_{t+1}}^{-\sigma_c}] = \frac{1}{\beta R_t E_t[\pi_{t+1}^{-1}]} {C_t}^{-\sigma_c} E_t[\frac{\exp\left(\frac{\sigma_c-1}{1+\sigma_l} L_t^{1+\sigma_l}\right)}{\exp\left(\frac{\sigma_c-1}{1+\sigma_l} L_{t+1}^{1+\sigma_l}\right)}] \tag{4}$$

The intratemporal optimality condition (3) equates the marginal rate of substitution between consumption and leisure to the real wage. The Euler equation (4) states the intertemporal optimality condition, relating current to future consumption by equating the marginal utility of consuming today to the expected discounted marginal utility of consuming tomorrow, adjusted for the return on bonds R_t and expected inflation $E_t\left[\pi_{t+1}\right]$.

3.1.2 The firm sector

Final goods producers

The final goods producer uses a continuum of differentiated intermediate goods $Y_t(i)$ in their production and packages them into a single final good Y_t to sell it in a perfectly competitive market at price P_t. The final goods firm makes zero profit in equilibrium, its role consists only of aggregating inputs. Final goods firms maximize the following profit function:

$$\max_{Y_t,\, Y_t(i)} \; P_t Y_t - \int_0^1 P_t(i) Y_t(i)\, di \tag{5}$$

Under the assumption of identical firms and that intermediate goods are turned into a final good via Dixit-Stiglitz aggregation. We obtain the following demand curve for intermediate good i:

$$Y_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-(1+\varepsilon_p)} Y_t \tag{6}$$

The demand curve relates the demand for intermediate input $Y_t(i)$ to its relative price and to aggregate output Y_t. Aggregating across all intermediate firms, the price index in the final goods sector is given by:

$$P_t = \left(\int_0^1 P_t(i)^{-\varepsilon_p} \, di \right)^{-\frac{1}{\varepsilon_p}} \tag{7}$$

Intermediate goods producers

Unlike the final goods producers, intermediate goods producers have market power and thus act as price setters. Thereby, they end up with smaller profits instead of losses when unable to re-optimize prices due to price stickiness. Intermediate firm i produces a differentiated good $Y_t(i)$ using labour as the only input, with the production function:

$$Y_t(i) = \epsilon_t^a L_t(i) - \Phi \tag{8}$$

where ϵ_t^a is the total factor productivity (TFP), $L_t(i)$ is labour hired by firm i, and $\Phi > 0$ is a fixed cost of production. TFP follows an exogenous AR(1) process:

$$\ln \epsilon_t^a = \rho_a \ln \epsilon_{t-1}^a + \sigma_a \eta_t^a, \qquad \eta_t^a \sim \mathcal{N}(0, 1) \tag{9}$$

The firm's problem is solved by taking the nominal wage W_t as given and minimizing total labour cost $W_t L_t(i)$ subject to meeting demand $Y_t(i)$. In the flex price case, the firm simply sets price as a constant markup over marginal cost:

$$P_t(i) = \frac{(1 + \varepsilon_p)}{\varepsilon_p} MC_t \tag{10}$$

However, we follow Smets and Wouters (2007) and introduce nominal price rigidity via the Calvo (1983) mechanism: in each period, a fraction $\xi_p \in (0, 1)$ of firms cannot re-optimise their price and instead index it to a weighted average of past and steady-state inflation:

$$P_{t+s}(i) = \tilde{P}_t(i) \cdot X_{t,s}, \qquad X_{t,s} = \prod_{l=1}^{s} \pi_{t+l-1}^{\iota_p} \bar{\pi}^{1-\iota_p} \tag{11}$$

where $\tilde{P}_t(i)$ is the price chosen at time t, $\iota_p \in [0, 1]$ is the degree of indexation to past inflation π_{t-1}, and $\bar{\pi}$ is steady-state inflation. When $\iota_p = 1$, prices are fully indexed to past inflation. When $\iota_p = 0$, non-optimizing firms keep their price fixed at $\tilde{P}_t(i)$.

The maximization problem for the fraction $(1 - \xi_p)$ of firms that can re-optimize in period t is now different. They choose $\tilde{P}_t(i)$ to maximize the present discounted value of profits over all future periods in which they may be stuck with this price:[2]

$$\max_{\tilde{P}_t(i)} \; E_t \sum_{s=0}^{\infty} (\beta \xi_p)^s \frac{\lambda_{t+s}}{\lambda_t} \frac{P_t}{P_{t+s}} \left[\tilde{P}_t(i) X_{t,s} - MC_{t+s} \right] Y_{t+s}(i) \tag{12}$$

[2]The term $\frac{\lambda_{t+s}}{\lambda_t} \frac{P_t}{P_{t+s}}$ is the real stochastic discount factor, where λ_t is the marginal utility of wealth from the household's optimisation problem (equation 4).

subject to the demand constraint faced in each period $t+s$, which follows from the final goods producer's optimisation problem in equation (6):

$$Y_{t+s}(i) = \left(\frac{\tilde{P}_t(i) X_{t,s}}{P_{t+s}} \right)^{-(1+\varepsilon_p)} Y_{t+s} \tag{13}$$

Solving the intermediate firm's optimisation problem and aggregating across all firms yields the following law of motion for the aggregate price level:

$$1 = (1-\xi_p) \left(\frac{\tilde{P}_t}{P_t} \right)^{1-\varepsilon_p} + \xi_p \left(\frac{\pi_{t-1}^{\iota_p} \bar{\pi}^{1-\iota_p}}{\pi_t} \right)^{1-\varepsilon_p} \tag{14}$$

The first term aggregates over the fraction $(1-\xi_p)$ of firms that re-optimise, who all set the same reset price $\tilde{P}_t$. The second term aggregates over the fraction ξ_p of firms that cannot re-optimise and instead mechanically index their price by $\pi_{t-1}^{\iota_p} \bar{\pi}^{1-\iota_p}$, so their relative price this period is their relative price last period scaled by this indexation factor, which in relative terms equals $\pi_{t-1}^{\iota_p} \bar{\pi}^{1-\iota_p} / \pi_t$.

3.1.3 Monetary policy rule

To close the model, we add a central bank that follows the following Taylor rate rule by changing the interest rate as a response to deviations of inflation and output from their target:

$$\frac{R_t}{R^*} = \left(\frac{R_{t-1}}{R^*} \right)^{\rho_{\mathrm{mp}}} \left[\left(\frac{\pi_t}{\pi^*} \right)^{\psi_\pi} \left(\frac{Y_t}{Y_t^*} \right)^{\psi_y} \right]^{1-\rho_{\mathrm{mp}}} \varepsilon_t^r \tag{15}$$

Here, R^*, π^* and Y^* represent the steady state nominal policy rate, inflation rate and output level respectively. The parameter ρ_{mp} represents the degree to which the central bank engages in interest rate smoothing over time. A higher ρ_{mp} means the central bank puts more weight on keeping the interest rate close to its previous value and adjusts more gradually in response to inflation and output gap developments. Finally, ε_t^r represents a monetary policy shock that follows an AR(1) process:

$$\ln \varepsilon_t^r = \rho_r \ln \epsilon_{t-1}^r + \sigma_r \eta_t^r, \qquad \eta_t^r \sim \mathcal{N}(0,1) \tag{16}$$

3.1.4 Equilibrium & steady state

A competitive equilibrium in this economy consists of sequences of quantities $\{Y_t, C_t, L_t\}$ and prices $\{W_t, \pi_t, R_t\}$ such that (i) households maximize lifetime utility subject to their budget constraint, (ii) final goods producers maximize profits subject to the aggregation technology, (iii) intermediate goods producers minimize costs and set prices optimally subject to the Calvo mechanism, (iv) the central bank follows the Taylor rule (15), and (v) all markets clear. Thus we end up with the following conditions:

$$L_t^s = L_t^d = L_t \tag{17}$$

$$B_t = 0 \tag{18}$$

$$Y_t = C_t \tag{19}$$

Because we abstract from capital accumulation and investment, the aggregate resource constraint is naturally the same as the goods market clearing condition. The equilibrium is then fully determined by the optimality conditions of households and firms derived above, the Taylor rule (15) which determines the nominal interest rate, and the exogenous shock processes which drive fluctuations around the steady state.

The steady state is defined as the equilibrium in which all variables take constant values so that $X_t = X^*$, and trending variables grow at the constant trend rate $\bar{\gamma}$, while shocks take the value of one, as they enter the model multiplicatively. Evaluating the equilibrium conditions derived above at these values delivers the steady-state relationships of the model.

From the private bond holding condition we derive expression that pins down the steady-state nominal interest rate:

$$R^* = \frac{\bar{\pi}}{\beta} \tag{20}$$

In steady state all intermediate firms are identical and set the same price, so $P_t(i) = P_t = P^*$ for all i, meaning that we can derive the steady-state real marginal costs and the steady-state real wage:

$$mc^* = \frac{\varepsilon_p}{1+\varepsilon_p} = w^* \tag{21}$$

From the production function (8), evaluated at steady state:

$$Y^* = \varepsilon^a L^* - \Phi = L^* - \Phi \tag{22}$$

where the fixed cost Φ is calibrated such that steady-state profits are zero.

Since we abstract from capital, the goods market clearing is:

$$Y^* = C^* \tag{23}$$

At steady state, labour market clearing combined with the household intratemporal optimality condition requires that the real wage equals the marginal rate of substitution between leisure and consumption:

$$w^* = C^*(L^*)^{\sigma_l} \tag{24}$$

Together with the production function (22) and goods market clearing (23), this pins down the steady-state values of L^* and C^* given the known steady-state real wage w^* from (21). Equations (20)–(24) thus jointly determine the steady-state values $\{Y^*, C^*, L^*, w^*, R^*, mc^*\}$ given the structural parameters $\{\beta, \sigma_c, \sigma_l, \varepsilon_p, \bar{\pi}, \Phi\}$.

Finally, evaluating the Taylor rule (15) at steady state, where all gaps are zero and $\varepsilon_t^r = 1$ yields a condition that is satisfied for all parameter values:

$$\frac{R^*}{R^*} = \left(\frac{R^*}{R^*}\right)^{\rho_{\mathrm{mp}}} \left[\left(\frac{\bar{\pi}}{\bar{\pi}}\right)^{\psi_\pi} \left(\frac{Y^*}{Y^*}\right)^{\psi_y}\right]^{1-\rho_{\mathrm{mp}}} = 1 \tag{25}$$

3.1.5 Log-linearised equilibrium conditions

The model is solved by log-linearising all equilibrium conditions around the steady state derived above. For any variable X_t, the log-deviation from steady state is defined as $\hat{x}_t = \ln X_t - \ln X^*$, which for small deviations approximates the percentage deviation from steady state.[3]

Combining the log-linearised Euler equation (4) and the log-linearised production function (8) yields the dynamic IS curve.

$$\hat{c}_t = E_t[\hat{c}_{t+1}] - \frac{\sigma_c - 1}{\sigma_c} E_t[\hat{\tau}_{t+1}] - \frac{1}{\sigma_c}(\hat{r}_t - E_t[\hat{\pi}_{t+1}]) + \frac{\sigma_c - 1}{\sigma_c}(\hat{y}_t - E_t[\hat{y}_{t+1}]) \tag{26}$$

Equation (26) states that current consumption depends on four terms. First, expected future consumption $E_t[\hat{c}_{t+1}]$ reflects the forward-looking nature of the household's savings decision. Second, expected TFP growth $E_t[\hat{\tau}_{t+1}] \equiv \hat{\varepsilon}_{t+1}^a - \hat{\varepsilon}_t^a$ enters because households anticipate future productivity gains and adjust consumption today accordingly - a higher expected growth rate raises expected future income, encouraging more consumption today. Third, the ex-ante real interest rate $\hat{r}_t - E_t[\hat{\pi}_{t+1}]$ governs the intertemporal substitution of consumption: a higher real interest rate makes saving more attractive and reduces current consumption. The parameter $1/\sigma_c$ is the intertemporal elasticity of substitution, governing the sensitivity of consumption to changes in the real interest rate. Fourth, the term $\frac{\sigma_c-1}{\sigma_c}(\hat{y}_t - E_t[\hat{y}_{t+1}])$ reflects the non-separability between consumption and labour in the utility function. When output is expected to grow, future labour supply increases, which affects the marginal utility of consumption today through the non-separable utility. When $\sigma_c = 1$ this term vanishes entirely, recovering the standard separable case.

After some substitutions the log-linearized real marginal costs become:

$$\hat{mc}_t = \hat{c}_t + \sigma_l \hat{y}_t - (\sigma_l + 1)\hat{\varepsilon}_t^a \tag{27}$$

where real marginal cost $\hat{mc}_t$ increases with both consumption $\hat{c}_t$ and output y_t. The consumption term reflects that higher consumption raises the marginal utility of wealth and therefore the opportunity cost of working, pushing up wages and hence marginal cost. The output term $\sigma_l \hat{y}_t$ reflects that higher output requires more labour, which raises the marginal disutility of labour and therefore wages.

Deriving the NK Phillips Curve requires combining the log-linearised Calvo pricing FOC derived in the intermediate goods section with the log-linearised aggregate price index with Calvo indexation (14) and yields:

$$\hat{\pi}_t = \frac{\iota_p}{1 + \iota_p \beta\gamma}\hat{\pi}_{t-1} + \frac{\beta\gamma}{1 + \iota_p \beta\gamma} E_t[\hat{\pi}_{t+1}] + \frac{(1-\xi_p)(1-\beta\gamma\xi_p)}{\xi_p(1 + \iota_p\beta\gamma)}\hat{mc}_t \tag{28}$$

[3]Note that equations (26) and (27) differ slightly from Del Negro and Schorfheide (2013), as we follow the Smets and Wouters (2007) utility function which features non-separable consumption and labour.

The Phillips curve relates current inflation to three terms. The backward-looking term $\frac{\iota_p}{1+\iota_p\beta}\hat{\pi}_{t-1}$ reflects price indexation. Firms that cannot re-optimise mechanically adjust their prices to past inflation, generating intrinsic inflation persistence. The forward-looking term $\frac{\beta}{1+\iota_p\beta}E_t[\hat{\pi}_{t+1}]$ reflects that optimising firms are forward-looking and set prices based on expected future costs. The slope $\frac{(1-\xi_p)(1-\beta\xi_p)}{\xi_p(1+\iota_p\beta)}$ determines how strongly inflation responds to real marginal cost $\hat{mc}_t$: a higher degree of price stickiness ξ_p means fewer firms re-optimise each period, flattening the Phillips curve. The appearance of γ in the discount factor reflects that the model accounts for a deterministic balanced growth path at rate γ, so future profits must be discounted not only by β but also by the trend growth rate, yielding an effective discount factor of $\beta\gamma$.

Log-linearising the non-linear Taylor rule (15) around the steady state delivers:

$$\hat{r}_t = \rho_{\text{mp}}\hat{r}_{t-1} + (1-\rho_{\text{mp}})\psi_\pi\hat{\pi}_t + (1-\rho_{\text{mp}})\psi_y dy_t + \hat{\varepsilon}^r_t \tag{29}$$

Note that the central bank responds to output growth dy_t rather than the output gap $\tilde{y}_t$. This reflects the Del Negro and Schorfheide (2013) specification, in which the central bank targets changes in output rather than deviations from potential and allows us to abstract from modelling the flexible-price economy.

The central bank smooths the interest rate with weight ρ_{mp} on the lagged rate, and responds to current inflation with weight $(1-\rho_{\text{mp}})\psi_\pi$ and to output growth with weight $(1-\rho_{\text{mp}})\psi_y$. A higher smoothing parameter ρ_{mp} implies more inertia. The central bank adjusts gradually rather than responding aggressively to current conditions, as both response coefficients are scaled down by $(1-\rho_{\text{mp}})$. The Taylor principle requires $\psi_\pi > 1$ to ensure a unique rational expectations equilibrium. If the central bank raises nominal rates more than one-for-one with inflation, it raises real rates and stabilises inflation.

In the log-linearised model, we introduce an exogenous government spending shock ε^g_t that shifts aggregate demand. The aggregate resource constraint then becomes:

$$\hat{y}_t = \hat{c}_t + \varepsilon^g_t \tag{30}$$

where ε^g_t is an exogenous demand shock that follows an AR(1) process:

$$\varepsilon^g_t = \rho_g\varepsilon^g_{t-1} + \sigma_g\eta^g_t, \quad \eta^g_t \sim \mathcal{N}(0,1) \tag{31}$$

with persistence parameter $\rho_g \in (0,1)$, and shock standard deviation σ_g. Note that ε^g_t is introduced directly at the log-linearised stage as a reduced-form demand disturbance rather than being derived from a non-linear government spending process, following Del Negro and Schorfheide (2013). Since government spending is entirely exogenous, any increase in ε^g_t must be matched by an increase in output or a decrease in consumption. There is no investment or net exports to absorb the shock, so government spending crowds out consumption in the absence of a monetary policy response.

Since the model is solved in detrended form but matched to data on output growth in levels, Del Negro and Schorfheide (2013) define observed output growth as:

$$dy_t = \hat{y}_t - \hat{y}_{t-1} + \hat{\tau}_t \tag{32}$$

where TFP growth is defined as:

$$\hat{\tau}_t = \hat{\varepsilon}^a_t - \hat{\varepsilon}^a_{t-1} \tag{33}$$

To sum up, the model is driven by three structural disturbances defined as standard AR(1) processes:

$$\varepsilon_t^g = \rho_g \varepsilon_{t-1}^g + \sigma_g \eta_t^g \tag{34}$$
$$\varepsilon_t^r = \rho_r \varepsilon_{t-1}^r + \sigma_r \eta_t^r \tag{35}$$
$$\varepsilon_t^a = \rho_a \varepsilon_{t-1}^a + \sigma_a \eta_t^a \tag{36}$$
$$\tag{37}$$

where $\eta_t^g, \eta_t^a, \eta_t^r \overset{iid}{\sim} \mathcal{N}(0,1)$ are uncorrelated structural innovations and $\sigma_g, \sigma_a, \sigma_r$ are the standard deviations of the respective shocks. The persistence parameters ρ_g, ρ_r, and ρ_a govern how long the respective shocks last.

Equations (26), (27), (28), (15), (30), (32) and (33) together with the shock processes (34)–(36) form a closed system in the endogenous variables $\{\hat{c}_t, \hat{\tau}_t, \hat{r}_t, \hat{\pi}_t, \hat{m}c_t, \hat{y}_t, dy_t, \hat{\varepsilon}_t^a, \hat{\varepsilon}_t^r, \varepsilon_t^g\}$. This system constitutes the small-scale New Keynesian model of Del Negro and Schorfheide (2013) that is taken to the data in the subsequent sections.

3.2 Two paradigms of expectation formation

The standard procedure when solving DSGE models is to assume that agents have rational expectations. In this thesis we deviate from this assumption by assuming that agents are adaptive learners. In our model, this means that agents have to forecast future variables based on limited information from today. Additionally, the agents' forecasting model is not static. Instead, agents evaluate its performance and subsequently update the model's coefficients in order to increase its forecasting accuracy. This makes agents adaptive learners in the sense that they continuously adapt their beliefs about the true forecasting model by learning from their forecasting errors. Adaptive learning is a well-established framework in macroeconomics with a substantial body of literature. For an introduction and overview, see Evans and Honkapohja (2009). We base our model of adaptive learning on the work of Slobodyan and Wouters (2012), and will therefore mostly stick to their notation.

3.2.1 Model solution under rational expectations

In order to understand how adaptive learning differs from rational expectations, it is instructive to first review the general solution to a DSGE model under rational expectations. In the general case, after log-linearizing a DSGE model around its deterministic steady state, it can be represented by the following system of equations in matrix form:

$$\mathbf{A}_0 \begin{bmatrix} \mathbf{y}_{t-1} \\ \mathbf{w}_{t-1} \end{bmatrix} + \mathbf{A_1} \begin{bmatrix} \mathbf{y}_t \\ \mathbf{w}_t \end{bmatrix} + \mathbf{A_2} E_t \mathbf{y}_{t+1} + \mathbf{B_0} \boldsymbol{\epsilon_t} = \text{constant} \tag{38}$$

which is called the structural form. $\mathbf{y_t}$ and $\mathbf{w_t}$ are vectors containing the endogenous and exogenous variables respectively, and together make up the full set of model variables. $\boldsymbol{\epsilon_t}$ is a vector of i.i.d. normally distributed structural shocks with zero mean and identity covariance matrix. Solving this system under rational expectations means assuming that agents have model-consistent expectations. This means that the expectation operator E_t is formed using the history of all relevant variables available at time t together with full knowledge of all parameters of the DSGE model. That is, agents are assumed not only to know the history of all variables, but also the exact structural equations governing the economy they inhabit. Under this assumption, the solution to (38) becomes

$$\begin{bmatrix} \mathbf{y}_t \\ \mathbf{w}_t \end{bmatrix} = \mu + \mathbf{T} \begin{bmatrix} \mathbf{y}_{t-1} \\ \mathbf{w}_{t-1} \end{bmatrix} + \mathbf{R}\epsilon_\mathbf{t} \tag{39}$$

This is the general expression of the RE equilibrium and corresponds to the law of motion under model-consistent expectations. The solution expresses contemporaneous endogenous variables as a function of their lagged values and the structural innovations that shock the system out of equilibrium. Notably, the forward-looking expectations term $E_t\mathbf{y}_{t+1}$ that appears in the structural form (38) has disappeared: under rational expectations, agents know the true law of motion of the economy, so their expectations of future variables can be substituted out and expressed as functions of current and lagged variables. The vector $\boldsymbol{\mu}$ allows for deterministic growth in steady state, while the matrices $\mathbf{T}$ and $\mathbf{R}$ are functions of the structural parameters and govern the law of motion, which determines how structural shocks propagate throughout the economy.

3.2.2 Perceived law of motion

Following Slobodyan and Wouters (2012) and Warne (2026), we implement adaptive learning by assuming that agents (i.e. consumers and firms) behave as Bayesian econometricians in the sense that they employ a simple forecasting model and sequentially update its coefficients as new data arrive. We further assume that agents forecast only one period ahead, focusing exclusively on the variables that appear as expectations in the model's Euler equations. This approach, known as Euler equation learning and popularised by Evans and Honkapohja (2001), keeps the learning problem tractable by limiting agents to forecasting only the variables directly relevant to their current decisions, rather than requiring them to form expectations about the full future path of all model variables.

The agents' forecasting model is assumed to be linear, and can be expressed as $y_j^f = \mathbf{X}_j^T\boldsymbol{\beta}_j$ for $j = y, c, \pi$, where $\mathbf{y}^f$ denotes the sub-vector of $\mathbf{y}$ containing the forward-looking variables that agents must forecast. In our model that is output (y), consumption (c), and inflation (π). The matrix $\mathbf{X}_j$ represents the information set available to agents when forecasting variable j. Under RE, $\mathbf{X}_j$ contains everything relevant to the model, consistent with agents having full structural knowledge. Under AL, $\mathbf{X}_j$ is considerably more parsimonious, reflecting the imperfect information that agents base their expectations on in the real world. Specifically, we assume that each variable is forecast using a univariate AR(2) model, so that the forecasting model takes the form

$$E_t y^f_{t+1,j} = \beta_{0,t,j} + \beta_{1,t,j} y^f_{t,j} + \beta_{2,t,j} y^f_{t-1,j}, \quad \text{j = y, c, } \pi \tag{40}$$

which is the explicit functional form of the Perceived Law of Motion (PLM) assumed for agents in our model. Under this specification, agents forecast $y^f_{t+1,j}$ using only a constant and two lags of the variable itself, with no reference to any other model variables or structural relationships. The term perceived law of motion reflects the fact that this forecasting rule represents agents' own beliefs about how variables evolve, which need not coincide with the actual law of motion of the economy.

3.2.3 Kalman filter learning

The key feature of the setup is that agents are adaptive learners who update the $\boldsymbol{\beta}$-coefficients in their forecasting model (40) based on past forecast errors. We implement this updating through Kalman filter (KF) learning, which is an instance of a broader class of recursive statistical learning rules. The common principle underlying all such rules is that agents forecast variables using a rule-of-thumb and recursively revise it in response to forecast errors. A simpler case of recursive statistical learning is constant gain learning, in which every forecast error is assigned the same informational value. This is achieved by holding the, so called, gain parameter fixed throughout the sample, so that each new observation receives equal weight when the forecasting rule is revised.

Kalman filter learning is a more sophisticated approach, which shares the same fundamental principle as CG learning in that agents recursively revise their forecasting rule based on forecast errors. What distinguishes it is that the gain parameter is no longer fixed but varies endogenously over time. Specifically, the gain depends on agents' confidence in their current forecasting model. When confidence is low, forecast errors are treated as more informative and lead to larger revisions of the forecasting rule.

Consequently, the gain will tend to be larger during periods of disruption and structural change, allowing agents to converge more quickly to a more appropriate forecasting model. Slobodyan and Wouters (2012) highlight a clear example of when this feature is empirically important: the inflation dynamics of the 1970s in the US. Following the OPEC oil shock, the dynamics driving inflation changed abruptly and inflation became persistently elevated. During such a period, the large positive forecast errors temporarily reduce agents' confidence in their forecasting model and raise the Kalman gain, allowing KF learning to adapt to the new inflation regime faster than CG learning would. More generally, regardless of the specific type of recursive statistical learning employed, a key appeal of the framework is that it allows for time-varying forecasting models that can adapt to changing dynamics in macroeconomic time series, making it a suitable tool to study economies that have undergone macroeconomic "regime" shifts and structural changes.

Another dimension of the learning setup concerns the information available to agents when making their forecasts. By restricting the variables used in the forecasting rule, agents are not only adaptive learners but also informationally constrained, using a much smaller information set than a fully rational agent would. The result is that expectations will generally be less accurate than under RE. However, this is not necessarily a shortcoming. A useful macroeconomic model should not aim to equip agents with the most accurate possible expectations, but rather with expectations that resemble those of real-world agents as closely as possible.

Finally, the extent to which learning actively shapes agents' expectations is not fixed a priori but determined through estimation. In simplified terms, the model spans a spectrum of specifications ranging from completely fixed rule-of-thumb forecasting at one extreme, where no learning takes place, to unbounded belief updating at the other, where the forecasting model is free to follow wherever the data lead. The estimation procedure determines where along this spectrum the model ends up, allowing the data to speak on how much active learning is needed to best describe the dynamics of the studied economy.

We now turn to the formal definition of Kalman filter learning.[4] In each period,

[4] Warne (2025) (chpt. 17) provides a more exhaustive treatment of Kalman filter learning in the context of DSGE models than is given here.

agents update the $\boldsymbol{\beta}$-coefficients in their AR(2) forecasting model (40) based on the contemporaneous forecast error. Stacking all belief coefficients into the vector $\boldsymbol{\beta} = [\beta_{0,y}, \beta_{1,y}, \beta_{2,y}, \beta_{0,c}, \beta_{1,c}, \beta_{2,c}, \beta_{0,\pi}, \beta_{1,\pi}, \beta_{2,\pi}]$ provides a compact representation of the agents' three forecasting models. Agents are further assumed to believe that $\boldsymbol{\beta}$ follows a vector autoregressive process around a baseline belief vector $\bar{\boldsymbol{\beta}}$, which must be exogenously specified and is discussed in section 4.3. This VAR(1) process does not tell us anything about the learning mechanism, instead its role is simply to impose that beliefs are mean-reverting toward a pre-specified baseline. The VAR(1) is defined as follows:

$$\boldsymbol{\beta}_t - \bar{\boldsymbol{\beta}} = \rho_{AL} \boldsymbol{I}_9 [\boldsymbol{\beta}_{t-1} - \bar{\boldsymbol{\beta}}] + \boldsymbol{v}_t \tag{41}$$

where $\boldsymbol{I}_9$ is the 9-dimensional identity matrix, $\boldsymbol{v}_t$ is an i.i.d. normally distributed error vector with covariance matrix $\boldsymbol{V}$, and $\rho_{AL} < 1$ is the autoregressive parameter governing the speed at which $\boldsymbol{\beta}_t$ reverts toward the baseline beliefs $\bar{\boldsymbol{\beta}}$. This parameter plays a central role in the learning setup: when ρ_{AL} is small, any deviation of $\boldsymbol{\beta}_t$ from the baseline as a result of the adaptive learning mechanism disappears quickly, meaning that the pre-specified baseline $\bar{\boldsymbol{\beta}}$ ends up being the dominant determinant of agents' beliefs rather than the learning dynamics themselves. In contrast, a value of ρ_{AL} close to unity allows beliefs to drift persistently away from the baseline, giving the learning mechanism more leeway to shape agents' expectations over time.

Before specifying how $\boldsymbol{\beta}_t$ is updated through KF learning, it is necessary to state the combined forecasting model by stacking the individual forecasting models for j = y, c, π (from section 3.2.2). This yields:

$$\begin{bmatrix} y^f_{y,t} \\ y^f_{c,t} \\ y^f_{\pi,t} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^T_{y,t-1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{X}^T_{c,t-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{X}^T_{\pi,t-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta}_{y,t-1} \\ \boldsymbol{\beta}_{c,t-1} \\ \boldsymbol{\beta}_{\pi,t-1} \end{bmatrix} + \begin{bmatrix} u_{y,t} \\ u_{c,t} \\ u_{\pi,t} \end{bmatrix} = \boldsymbol{X}^T_{t-1} \boldsymbol{\beta}_{t-1} + \boldsymbol{u}_t \tag{42}$$

where the $\boldsymbol{\beta}_{j,t-1}$'s are 3-dimensional vectors stacked into a single 9-dimensional vector identical to the previously defined $\boldsymbol{\beta}$, and the $\mathbf{X}^T_{j,t-1}$'s are unpacked into a 3×9 matrix denoted $\mathbf{X}^T_{t-1}$. The three errors $u_{j,t}$ are linear combinations of the true model innovations $\boldsymbol{\epsilon}_t$, meaning that they are in general correlated across equations, so that the error covariance matrix $\boldsymbol{\Sigma} = E\left[\mathbf{u}_t \mathbf{u}_t^T\right]$ is likely non-diagonal. This necessitates stacking the three individual forecasting models into a single system in a Seemingly Unrelated Regression Equations (SURE) format when estimating, as that allows for accounting for the bias that it introduces.

With the notation established above, we can now state the Kalman filter learning mechanism. It consists of two sets of equations: an updating and a transition equation for the belief parameters $\boldsymbol{\beta}_t$, given in (43) and (44) respectively, and an analogous pair for the covariance matrix $\boldsymbol{P}_t$, given in (45) and (46). The matrix $\boldsymbol{P}_t$ is defined as the mean squared error of the belief coefficients and reflects agents' uncertainty about their own forecasting rule: a larger $\boldsymbol{P}_t$ indicates that agents place less confidence in their current belief coefficients.

$$\boldsymbol{\beta}_{t|t} = \boldsymbol{\beta}_{t|t-1} + \boldsymbol{K}_t [\mathbf{y}^f_t - \mathbf{X}^T_{t-1} \boldsymbol{\beta}_{t-1|t-1}] \tag{43}$$

$$\boldsymbol{\beta}_{t+1|t} = \bar{\boldsymbol{\beta}} + \rho_{AL} \boldsymbol{I}_9 [\boldsymbol{\beta}_{t|t} - \bar{\boldsymbol{\beta}}] \tag{44}$$

$$\mathbf{P}_{t|t} = \mathbf{P}_{t|t-1} - \boldsymbol{K}_t \mathbf{X}_{t-1}^T \mathbf{P}_{t|t-1} \tag{45}$$

$$\mathbf{P}_{t+1|t} = \rho_{AL} \mathbf{I}_9 \mathbf{P}_{t|t} [\rho_{AL} \mathbf{I}_9]^T + \mathbf{V} \tag{46}$$

$$\boldsymbol{K}_t = \mathbf{P}_{t|t-1} \mathbf{X}_{t-1} [\boldsymbol{\Sigma} + \mathbf{X}_{t-1}^T \mathbf{P}_{t|t-1} \mathbf{X}_{t-1}]^{-1} \tag{47}$$

where $\boldsymbol{K}_t$ is the Kalman gain. The subscript following the conditioning sign "|" indicates the information set used in estimation: for example, $\boldsymbol{\beta}_{t|t-1}$ denotes the belief parameters for period t estimated using information available through period $t-1$.

The updating equation (43) and the transition equation (44) together govern the evolution of the belief coefficients. The updating equation revises the belief coefficients from $\boldsymbol{\beta}_{t|t-1}$ to $\boldsymbol{\beta}_{t|t}$ by adding the product of the current forecast error $\mathbf{y}_t^f - \mathbf{X}_{t-1}^T \boldsymbol{\beta}_{t-1|t-1}$ and the Kalman gain $\boldsymbol{K}_t$, so that the magnitude of the revision is increasing in both the size of the forecast error and the gain. The transition equation then maps $\boldsymbol{\beta}_{t-1|t-1}$ forward to $\boldsymbol{\beta}_{t|t-1}$ according to the VAR(1) process in (41), providing the prior belief entering the next period's updating step. The same two-step logic applies to the covariance matrix $\boldsymbol{P}_t$: it is updated via (45) and mapped forward via (46). Finally, the endogeneity of the gain is reflected in (47), where $\boldsymbol{K}_t$ is recalculated each period as an increasing function of $\boldsymbol{P}_{t|t-1}$, the endogenous measure of belief uncertainty, so that belief revisions are automatically larger precisely when agents are most uncertain about their forecasting rule.

3.2.4 Model solution under adaptive learning

With the recursive learning setup for the belief parameters established, we can replace the $\boldsymbol{\beta}_{t-1}$ in the combined forecasting model (42), with the Kalman filter updated beliefs for $\boldsymbol{\beta}_{t-1|t-1}$. This substitution yields an explicit expression for the expectations of the forward-looking variables, $E_t \mathbf{y}_{t+1}^f$, under Kalman filter learning. Plugging this expression back into the structural form (38) then eliminates all forward-looking terms, yielding a purely backward-looking representation of the DSGE model that can be solved directly. The solution takes the form:

$$\begin{bmatrix} \mathbf{y}_t \\ \mathbf{w}_t \end{bmatrix} = \boldsymbol{\mu}_t + \mathbf{T}_t \begin{bmatrix} \mathbf{y}_{t-1} \\ \mathbf{w}_{t-1} \end{bmatrix} + \mathbf{R}_t \boldsymbol{\epsilon}_t \tag{48}$$

where the matrices $\boldsymbol{\mu}_t$, $\mathbf{T}_t$, and $\mathbf{R}_t$ are now time-varying. Apart from this, the AL solution takes the same general form as the RE solution (39). The reason for time-variation in the AL solution is that the solution is no longer fully pinned down by the fixed structural parameters, as in RE, but now also by the time-varying belief parameters, which are continuously updated by the Kalman filter learning mechanism, causing the solution matrices to change over time. The solution (48) represents the Actual Law of Motion (ALM) of the system and describes how the endogenous variables actually evolve and respond to structural shocks under adaptive learning. However, the solution form points to a fundamental difference between AL and RE. Under RE, the PLM and ALM coincide by construction, as agents' beliefs are consistent with the true law of motion of the economy. Under AL, the two generally differ, and it is precisely this gap between what agents believe and what actually happens that gives rise to the endogenous learning dynamics that this thesis sets out to study. Importantly, this discrepancy is not merely a technical artefact but the central mechanism through which adaptive learning influences the real economy: since expectations enter directly into the decision problems of both households and firms, non-rational beliefs feed back into actual outcomes, shaping consumption, pricing, and ultimately aggregate dynamics. This is perhaps most clearly illustrated by the time-dependence of the solution matrices in (48), which shows how the ALM itself becomes time-varying as a direct consequence of its contingency on agents' time-varying PLMs.

As the learning setup is recursive, it must be initialized before estimation can proceed. This requires specifying the parameters $\boldsymbol{\Sigma}$, $\boldsymbol{V}$, $\boldsymbol{P}_{1|0}$, $\boldsymbol{\beta}_{1|0} = \bar{\boldsymbol{\beta}}$, and ρ_{AL}, the details of which are discussed in section 4.3.

4 Estimation

4.1 Estimation approach

The following section describes the estimation procedure for our model, which requires some additional steps relative to the standard rational expectations case. The conventional approach to estimating RE DSGE models from macroeconomic time series, developed by Ireland (2004) and further popularised by Smets and Wouters (2003, 2007), is to apply Kalman filtering to the RE solution (39) in order to evaluate the likelihood of the model given the observed data. It should be noted that this application of the Kalman filter is different from the Kalman filter learning mechanism described in section 3.2. The two are separate applications of the same mathematical tool.

The Kalman filter is a general tool for estimating dynamic unobserved state variables from observables that are linearly related to them, and conveniently delivers the log-likelihood function as a byproduct. This likelihood can then be used either directly in a maximum likelihood framework or combined with prior distributions over the structural parameters to produce Bayesian posterior estimates. Following the influential work of Smets and Wouters (2003, 2007), the Bayesian approach has become one of the dominant estimation paradigm in the DSGE literature, for two main reasons. First, priors allow the researcher to incorporate external micro-evidence on specific parameters, for instance using price-setting frequencies from scanner data to discipline the Calvo parameter. Second, and more importantly, the likelihood surface of a DSGE model is typically ill-behaved, exhibiting multiple local maxima and near-singularities that can cause maximum likelihood estimators to converge to economically implausible parameter values. Prior distributions help anchor the optimisation algorithm to regions of the parameter space that are both data-consistent and economically meaningful.

We follow this tradition and estimate our model within a Bayesian framework, using the log-likelihood function obtained from the Kalman filter representation of the model. However, the presence of adaptive learning agents who themselves update their beliefs via a Kalman filter necessitates some additional steps relative to the standard estimation procedure, which we now describe.

A Kalman filter is always expressed in a state-space representation consisting of two equations. The first, the state equation, describes the dynamics of the unobserved states, which in the DSGE context correspond to the model variables' log-deviations from steady state. The second, the measurement equation, relates these unobserved states to the observed time series. In the standard RE case, the state equation is given by the time-invariant RE solution (39), and a single Kalman filter applied to this state-space representation delivers the likelihood function directly.

In our AL case, however, the state equation is given by the ALM (48), whose solution matrices are time-varying because they depend on the belief parameters $\boldsymbol{\beta}_t$ that are themselves being updated each period by the Kalman filter learning mechanism described in section 3.2. This means that a single standard Kalman filter cannot be applied directly, as the state equation no longer has the linear time-invariant form that the standard filter requires. Instead, the estimation requires two nested Kalman filters running in parallel: an inner filter that updates the belief parameters $\boldsymbol{\beta}_t$ according to the learning mechanism, and an outer filter that tracks the model states given the time-varying ALM that results from the evolving beliefs. These two filters are combined into a joint Kalman filter,[5] from which the log-likelihood function is obtained. For our model with three time series, the measurement equation linking the unobserved model states to the observables takes the form:

$$\begin{bmatrix} dlGDP_t \\ dlP_t \\ RATE_t \end{bmatrix} = \begin{bmatrix} \bar{\gamma} \\ \bar{\pi} \\ \bar{r} \end{bmatrix} + \begin{bmatrix} \hat{y}_t - \hat{y}_{t-1} + \hat{\tau}_t \\ \hat{\pi}_t \\ \hat{R}_t \end{bmatrix} = \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{H} \begin{bmatrix} \mathbf{y}_t \\ \mathbf{w}_t \end{bmatrix} + \boldsymbol{w}_t \tag{49}$$

where the first equality shows that the observed variables equal their deterministic steady-state values, with $\bar{\gamma}$ capturing the trend growth rate along the balanced growth path, plus the cyclical deviations of the model states from steady state.[6] l and d denote the natural logarithm and first difference operators respectively. The observed time series are described in section 4.2. $\boldsymbol{x}_t$ is a vector of deterministic variables, $\boldsymbol{H}$ is the observation matrix linking the unobserved model states to the observable time series, and $\boldsymbol{w}_t$ is a white noise measurement error vector.

From the previously described joint Kalman filter we obtain the log-likelihood function, which is then combined with the prior distributions to form the posterior. Since no closed-form expression exists for the posterior, numerical methods are required for optimisation, and the parameter values that maximise the posterior density are taken as the Bayesian point estimates. The Metropolis-Hastings algorithm is then used to explore the full posterior distribution, construct credible intervals, and compute the marginal likelihood of the model for model comparison purposes.[7] Before turning to the prior specifications, the next section describes the data used in estimation.

4.2 Data

We estimate the model using euro area data from the Area-Wide Model Database (AWMD), maintained and made available by İpek and Kısacıkoğlu (2025). The database draws on official data collected and revised by the ECB and Eurostat for the 20 euro area member states (as of 2025) and covers quarterly observations from 1970Q1 to 2019Q4. We use three time series from the database: real GDP, the GDP deflator, and the nominal short-term

[5]The exact procedure is described in Warne (2025) (Chpt. 17.5).

[6]The difference in the measurement equation for *dlGDPt* relative to the other two observables follows directly from equation (32), which defines output growth as the first difference in output plus TFP growth.

[7]The estimation and all subsequent analysis are carried out in the MATLAB program YADA. For more information see `https://www.texlips.net/yada/index.html` or Warne (2025).

interest rate. Both GDP and the deflator are seasonally adjusted and corrected for calendar effects.

Real GDP is converted into an approximation of real GDP per capita by dividing by the labour force, which is also provided in the AWMD. As specified in the measurement equation (49), both GDP per worker and the GDP deflator enter the model in log-differences, while the nominal short-term interest rate is measured as the quarterly average of the 3-month Euribor rate. The Appendix B.1 contains line plots of the three time series.

As emphasised by Milani (2011) and Slobodyan and Wouters (2012), the use of revised rather than real-time data is potentially problematic when estimating DSGE models with adaptive learning agents, since revised data can differ substantially from the data agents actually had available when forming their forecasts and updating their beliefs. This discrepancy is a limitation of our study, that at best introduces additional noise into the estimates, and at worst it introduces bias. While it is possible to conduct estimation using real-time data, as demonstrated by Warne (2026), this lies outside the scope of this thesis.

4.3 Priors & calibration

In Bayesian estimation of DSGE models, it is standard practice to calibrate a subset of parameters rather than estimate all of them. For the structural parameters of our model, we follow the specification of Del Negro and Schorfheide (2013), with two exceptions. First, we calibrate the monetary policy response to inflation to $\psi_\pi = 1.5$ rather than estimating it, consistent with commonly reported values in the literature (Smets and Wouters, 2007; Galí, 2015). Second, we calibrate the autoregressive parameter of the monetary policy shock process to $\rho_r = 0.1$, again broadly in line with previous estimates (Smets and Wouters, 2007; Warne, 2026).

For the estimated structural parameters, we use the same priors as Del Negro and Schorfheide (2013), which are reported in Table 1. For the adaptive learning parameters, we follow the calibration approach of Slobodyan and Wouters (2012) and Warne (2026), who initialise the learning setup using the rational expectations equilibrium, so that all learning parameters are RE-consistent at the start of the sample. Specifically, we set the initial beliefs equal to the baseline beliefs, $\boldsymbol{\beta}_{1|0} = \bar{\boldsymbol{\beta}} = (E[\mathbf{X}^T\mathbf{X}])^{-1}E[\mathbf{X}^T\mathbf{y}]$, and compute the forecast error covariance matrix as $\boldsymbol{\Sigma} = E[(\mathbf{y}_t^f - \mathbf{X}_{t-1}^T\boldsymbol{\beta}_{1|0})(\mathbf{y}_t^f - \mathbf{X}_{t-1}^T\boldsymbol{\beta}_{1|0})^T]$. The initial covariance matrix of the belief coefficients $\mathbf{P}_{1|0}$ and the covariance matrix of the belief innovation $\boldsymbol{V}$ are both assumed proportional to $(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$, giving $\mathbf{P}_{1|0} = \sigma_0(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$ and $\boldsymbol{V} = \sigma_v(\mathbf{X}^T\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}$. This leaves three parameters to be determined: σ_0, σ_v, and ρ_{AL}. Since σ_0 and σ_v cannot be jointly identified from the data given the structure of the learning setup, we follow Slobodyan and Wouters (2012) and calibrate them to $\sigma_0 = 0.03$ and $\sigma_v = 0.003$, leaving only the learning persistence parameter ρ_{AL} to be estimated, for which we assume a uniform prior over the unit interval.

Table 1: Prior distributions for the structural parameters of the RE and AL versions of our small-scale model.

parameter	density	RE P_1	RE P_2	AL P_1	AL P_2
σ_c	N	1.500	0.375	1.500	0.375
σ_l	N	2.000	0.500	2.000	0.500
ξ_p	β	0.500	0.100	0.500	0.100
$\imath_p$	β	0.500	0.150	0.500	0.150
ρ_{mp}	β	0.750	0.100	0.750	0.100
ψ_y	N	0.125	0.050	0.125	0.050
$\bar{\pi}$	Γ	0.625	0.100	0.625	0.100
β^{-1}	Γ	0.250	0.100	0.250	0.100
$\bar{\gamma}$	N	0.400	0.100	0.400	0.100
ρ_g	β	0.500	0.200	0.500	0.200
ρ_a	β	0.500	0.200	0.500	0.200
σ_g	Γ^{-1}	0.100	2.000	0.100	2.000
σ_a	Γ^{-1}	0.100	2.000	0.100	2.000
σ_r	Γ^{-1}	0.100	2.000	0.100	2.000
ρ_{AL}	$\mathcal{U}$	–		0.000	1.000
ρ_r	–	0.100		0.100	
ψ_π	–	1.500		1.500	
σ_0	–	–		0.030	
σ_v	–	–		0.003	

Notes: The columns P_1 and P_2 refer to the mean and the standard deviation of the normal (N), standardized beta (β), gamma (Γ), and inverse gamma (Γ^{-1}) distributions. For the uniform ($\mathcal{U}$) distribution, P_1 and P_2 denote the lower and upper bounds instead.
Source: Authors' specification following priors of Del Negro and Schorfheide (2013) and Slobodyan and Wouters (2012).

5 Results

The results section proceeds in six steps. We first examine the posterior estimates and their implied model equations, before turning to formal Bayesian model comparison. We then investigate the belief dynamics of the adaptive learning model, followed by an analysis of monetary policy transmission through impulse response functions benchmarked against local projections, and an out-of-sample forecasting comparison. The section closes with a number of robustness checks.

5.1 Posterior estimates

After combining prior information with the euro area data following the Bayesian estimation approach outlined in section 4.1, we now examine the resulting posterior estimates. These posteriors are central to what follows since they pin down the structural equations of the DSGE model and thus govern how agents behave and how the economy responds to shocks. Comparing the RE and AL posteriors is also interesting in its own right, as the differences reveal how each expectation formation assumption affects model parameters that themselves have direct economic interpretations. Table 2 reports the posterior mean, mode, and 5% and 95% quantiles for all estimated parameters under both specifications, based on 100,000 draws from the Random Walk Metropolis sampler.

Table 2: Posterior estimates of the structural parameters for the euro area RE and AL versions of our DSGE small-scale model for the sample 1970Q2-2019Q4.

		RE				AL			
	Description	mean	mode	5%	95%	mean	mode	5%	95%
Structural parameters									
σ_c	Substitution elasticity	2.196	2.186	1.956	2.444	4.838	4.824	4.514	5.163
σ_l	Inverse Frisch elasticity	3.140	2.985	2.680	3.620	0.152	0.152	0.113	0.192
ξ_p	Price stickiness	0.711	0.709	0.676	0.745	0.350	0.350	0.323	0.379
$\imath_p$	Price indexation	0.118	0.109	0.049	0.198	0.845	0.841	0.770	0.912
ρ_{mp}	Policy smoothing	0.608	0.611	0.556	0.657	0.886	0.883	0.864	0.908
ψ_y	Output response	0.213	0.215	0.148	0.278	0.142	0.142	0.079	0.204
$\bar{\pi}$	Steady-state inflation	0.518	0.517	0.425	0.616	0.512	0.509	0.410	0.620
β^{-1}	Inverse Discount factor	0.139	0.131	0.084	0.199	0.130	0.131	0.068	0.200
$\bar{\gamma}$	Trend growth	0.310	0.316	0.273	0.345	0.171	0.165	0.131	0.211
ρ_g	Demand persistence	0.955	0.955	0.946	0.963	0.986	0.985	0.980	0.991
ρ_a	Technology persistence	0.968	0.967	0.948	0.986	0.995	0.996	0.991	0.998
σ_g	Government spending volatility	2.555	2.495	2.177	2.955	0.390	0.386	0.341	0.442
σ_a	Technology volatility	0.755	0.737	0.681	0.836	0.823	0.820	0.758	0.891
σ_r	Monetary volatility	1.275	1.260	1.121	1.437	0.654	0.652	0.605	0.706
ρ_{AL}	Belief persistence	–				0.278	0.316	0.056	0.541
Calibrated parameters									
ρ_r	Rate persistence	0.100				0.100			
ψ_π	Inflation response	1.500				1.500			
σ_0	Initial belief uncertainty	–				0.030			
σ_v	Long-run belief uncertainty	–				0.003			

Notes: The columns display the posterior mean, mode, and 5% and 95% quantiles respectively, based on 100,000 MCMC draws from the Random Walk Metropolis sampler. The learning parameters ρ_{AL}, σ_0 and σ_v are specific to the AL version only. The Table also shows the calibrated parameters. ($T = 199$)
Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025).

A feature that stands out immediately upon inspecting Table 2 is the estimates for the inverse elasticity of intertemporal substitution (σ_c) and the inverse Frisch elasticity (σ_l). Both parameters differ substantially across specifications, with the 90% credible intervals not overlapping between RE and AL. This is difficult to reconcile with economic theory. σ_c and σ_l are parameters governing household preferences and should therefore in principle be unaffected by the assumption made about expectation formation. A higher σ_c, for instance, implies a stronger preference for consumption smoothing, which should not depend on whether agents form expectations rationally or adaptively. The same applies to the trend growth parameter $\bar{\gamma}$, describing the steady-state output growth that also displays significant differences between RE and AL estimates.

That these parameters shift so markedly when only the expectation formation mechanism is changed casts doubt on our estimated specification, and caution is warranted. However, the finding is not without precedent: Milani and Rajbhandari (2012) explicitly document that posterior distributions of structural parameters shift significantly when the RE assumption is modified, and similarly find an increase in the intertemporal elasticity of substitution under learning. A natural mechanism behind this lies in the identification structure of the model. Under RE, model-consistent expectations impose tight cross-equation restrictions that anchor the mapping from reduced-form dynamics to structural parameters, while under AL these restrictions are relaxed, giving the likelihood greater freedom to redistribute persistence between utility curvature and expectational dynamics.[8]

The two price-setting parameters, ξ_p and $\imath_p$, are central to New Keynesian modelling

[8]This is further compounded by the small estimated ρ_{AL}. This makes the beliefs tightly anchored around the RE-consistent baseline and expectations effectively collapse to AR(2) projections of the lagged observables. Thus, the dynamic structure of the model through which σ_c and σ_l are identified changes which makes direct comparability with their RE counterparts difficult.

as they generate price stickiness and thereby inflation persistence and non-neutrality of monetary policy (Galí, 2015). The Calvo parameter ξ_p is halved from 0.71 under RE to 0.35 under AL, moving from a value broadly in line with estimates from canonical medium-scale rational expectations models such as Smets and Wouters (2003, 2007) toward a lower degree of price-stickiness that implies that firms are substantially more flexible in their price setting. The price indexation parameter $\imath_p$, which governs how much firms unable to readjust their prices adjust them in line with past inflation, moves in the opposite direction, from near zero under RE to 0.845 under AL. Together these two changes imply a considerable shift in price setting dynamics in the AL model, which we examine through the implied Phillips curves below.

The parametrization of the monetary policy rule shapes how shocks are dampened by monetary policy as they propagate through the economy. A notable difference across specifications is the interest rate smoothing parameter ρ_{mp}, which is higher under AL (0.886) than under RE (0.608). A higher ρ_{mp} implies that the central bank adjusts rates more gradually in response to inflation and output developments, as both response coefficients are scaled down by $(1-\rho_{\mathrm{mp}})$. The AL monetary policy rule therefore reacts less aggressively to contemporaneous shocks and instead distributes the adjustment over time, which will have implications for the persistence of impulse responses discussed below.

Regarding the shock processes, more informative than the near-unit-root persistence parameters (ρ_g and ρ_a) are the shock standard deviations, which govern how strongly each structural innovation enters the AR(1) processes. Here the two specifications differ considerably. Monetary policy shock volatility σ_r is roughly twice as large under RE (1.275) as under AL (0.654), while government spending volatility σ_g is more than six times larger under RE (2.555) than under AL (0.390). These differences already hint at large disparities in how the two specifications model shocks, which will become apparent in the impulse response analysis below.

Finally, a parameter of particular importance for the adaptive learning specification is ρ_{AL}, which governs the speed at which belief coefficients revert back toward the RE-consistent baseline. Our baseline estimate of $\rho_{AL} = 0.278$ is small. This implies that deviations of beliefs from the baseline are short-lived, after only two quarters $1 - 0.278^2 \approx 92\%$ of a belief shift induced by learning has already disappeared. In practice, this makes agents closer to fixed rule-of-thumb forecasters than active learners. As ρ_{AL} is subject to estimation, its low value is a consequence of optimization, meaning that excessive learning from past mistakes reduces model fit as opposed to improving it. The drivers and implications of this finding is analysed in depth in the discussion section 6.3.

The differences in posterior estimates across specifications feed directly into the structural model equations. While the implications for the Taylor rule are already apparent from the parameter differences discussed above, the IS curve and Phillips curve has more to tell us about how shocks propagate differently under RE and AL. Notably, these differences do not only stem from shifting parameter estimates, but also from the fact that the curves' expectations terms will take on different values due to the differing expectation formations.

The implied IS curve is primarily shaped by σ_c. As shown in equations (50) and (51), the substantially higher σ_c under AL lowers the intertemporal elasticity of substitution, reducing households' willingness to shift consumption across periods in response to interest rate changes. Consequently, the coefficient on the real interest rate falls from 0.46 to 0.21, while the coefficients on expected productivity growth and the output growth spread rise from 0.54 to 0.79. Under AL, consumption decisions are thus driven more by expected income dynamics and less by the real interest rate, dampening the direct channel through which monetary policy affects demand.

Rational expectations IS curve:

$$\hat{c}_t = E_t[\hat{c}_{t+1}] - 0.54E_t[\hat{\tau}_{t+1}] - 0.46(r_t - E_t[\pi_{t+1}]) + 0.54(y_t - E_t[y_{t+1}]) \tag{50}$$

Adaptive learning IS curve:

$$\hat{c}_t = E_t[\hat{c}_{t+1}] - 0.79E_t[\hat{\tau}_{t+1}] - 0.21(r_t - E_t[\pi_{t+1}]) + 0.79(y_t - E_t[y_{t+1}]) \tag{51}$$

The AL Phillips curve differs from its RE counterpart along two dimensions. The higher $\imath_p$ shifts the inflation weights toward the past, with the backward-looking term rising from 0.11 to 0.46 and the forward-looking term falling correspondingly, as non-optimising firms rely more heavily on last period's price level. The lower ξ_p, meanwhile, steepens the slope on real marginal cost from 0.10 to 0.65. With more firms able re-optimise each period, inflation reacts more strongly to output and cost conditions. This has a favourable implication for monetary policy since a steeper Phillips curve means a smaller output contraction is needed to achieve a given disinflation.

Rational expectations Phillips curve:

$$\pi_t = 0.11\pi_{t-1} + 0.90E_t[\pi_{t+1}] + 0.10mc_t \tag{52}$$

Adaptive learning Phillips curve:

$$\pi_t = 0.46\pi_{t-1} + 0.54E_t[\pi_{t+1}] + 0.65mc_t \tag{53}$$

5.2 Bayesian model comparison

Having examined the posterior estimates and their implied model equations, we now turn to a formal comparison of empirical fit, asking which of the two specifications is better supported by the euro area data and consequently whether the parameter differences documented above have meaningful economic interpretations or not. A natural way to compare empirical fit is through the log marginal likelihoods reported in Table 3, which displays the log likelihood at the mode, the log prior and posterior density kernels, and the Laplace approximation of the log marginal likelihood for both the RE and AL specifications.

Table 3: Prior and posterior densities, and likelihoods, for RE vs AL

	RE	**AL**
Log likelihood at θ mode	−759.847	−666.273
Log prior density of θ	−23.130	−61.676
Log posterior density kernel of θ	−782.977	−727.949
Log Jacobian ($d\theta/d\phi$)	−13.782	−21.262
Laplace approx. of log marginal likelihood	−818.935	−767.168

Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025).

The log likelihood at the θ mode measures how well each model fits the data at its optimal parameter values, disregarding prior beliefs. The AL model fits the data substantially better than RE (-666 vs. -760), suggesting that the learning mechanism generates dynamics that match the euro area data more closely at their respective best fitting parameters. Turning to the log prior density, the more negative value for AL (-62 vs. -23) indicates a larger prior penalty, meaning the optimizer was pushed towards regions of the parameter space that are a priori less plausible. This is already visible in Table 2, where the posterior modes of σ_c and σ_l under AL lie far outside the 90% credible intervals of their priors, while the RE modes lie inside or close to the boundary of the 90% credible intervals of their priors. Both pieces of information are combined in the log posterior density kernel, which reflects the trade-off between data fit and prior plausibility. Notably, AL still achieves a substantially higher posterior kernel (-728 vs. -783), confirming that the AL model dominates RE not only on data fit alone but also when the larger prior penalty is accounted for.

While the log posterior kernel evaluates model fit at a single point estimate, this risks rewarding complexity, as a more flexible model can always find a better mode simply by having more parameters to exploit. The log marginal likelihood addresses this by integrating the posterior over the entire parameter space, automatically penalizing model complexity and providing a more robust criterion for model comparison. The AL model wins decisively on this metric as well (-767 vs. -819), confirming that even when fully accounting for AL's larger complexity, the learning model is still preferred.

5.3 Belief dynamics and the role of learning

Having established a significant improvement in marginal likelihood under the AL specification, we now examine the model's internal expectation dynamics to better understand what drives this result. We begin by investigating the belief persistence parameter ρ_{AL} and its implications for how agents form expectations, before turning to the laws of motion that agents perceive versus what actually governs the economy.

The estimated belief persistence parameter $\rho_{AL} = 0.278$ is central to understanding how adaptive learning operates in our model. As discussed in section 5.1, this small value tightly anchors agents' forecasting rules around the RE-consistent AR(2) baseline, rendering them closer to static rule-of-thumb forecasters than active learners. This creates an important interplay. While the specification of the baseline beliefs $\bar{\beta}$ shapes the magnitude of ρ_{AL} (a less accurate anchor necessitating a larger ρ_{AL} to allow beliefs sufficient flexibility to track the data), it is in turn ρ_{AL} that determines how closely the actual Perceived Law of Motion tracks this baseline.

The RE-consistent baseline beliefs $\bar{\beta}$ are thus central to the model's dynamics, and it is instructive to examine what they actually imply for agents' forecasting rules in our estimated specification. Following Slobodyan and Wouters (2012) and Warne (2026), and as described in section 3.2, we set $\bar{\beta}$ to its RE-consistent value, calibrating the implied AR(2) forecasting rules to be as close as possible to the belief system of fully rational agents. This choice is motivated by the assumption that agents possess a basic understanding of the economic environment they inhabit, sufficient to form reasonable priors over how macroeconomic variables evolve without requiring full rationality.

The three AR(2) processes implied by the full-sample RE solution are shown in equations (54) to (56). All three processes have a zero intercept, consistent with the variables being modelled as log-deviations from steady state. The persistence coefficients on the first lag differ considerably across variables: inflation exhibits the strongest autoregressive component (0.656), followed by output (0.463), while consumption displays a notably smaller first-lag coefficient (0.158). The second-lag coefficients are negative for both consumption and inflation (−0.137 and −0.206 respectively), implying dampened oscillatory dynamics for these two variables as their signs on average alternate each period. Output, by contrast, carries a positive second-lag coefficient (0.160), indicating smoother and more monotonic propagation.

$$E_t y^f_{t+1,c} = 0.1575 y^f_{t,c} - 0.1366 y^f_{t-1,c} \tag{54}$$

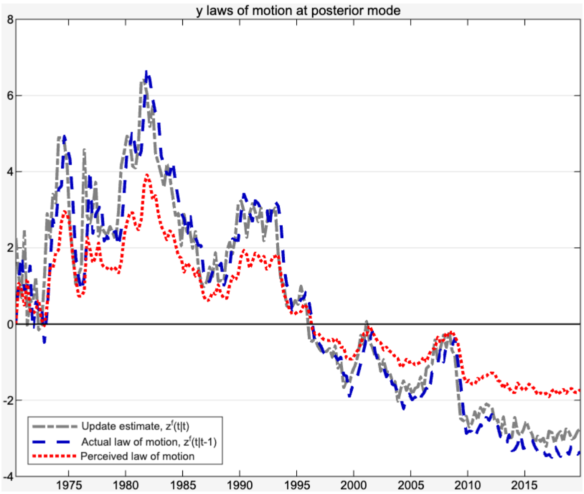
$$E_t y^f_{t+1,\pi} = 0.6558 y^f_{t,\pi} - 0.2061 y^f_{t-1,\pi} \tag{55}$$

$$E_t y^f_{t+1,y} = 0.4628 y^f_{t,y} + 0.1598 y^f_{t-1,y} \tag{56}$$

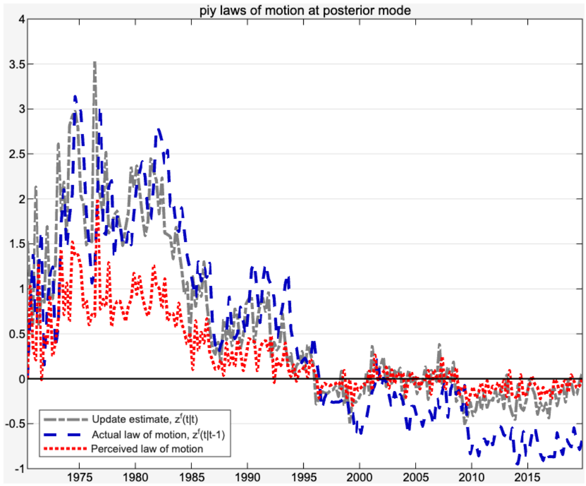
The low estimated belief persistence ensures that beliefs are rapidly pulled back toward this AR(2) process, keeping agents close to the rule-of-thumb forecasters described above. As shown and discussed in Appendix B.3, the belief coefficients remain tightly anchored around the RE-consistent parameter values throughout the sample, with only small fluctuations emerging in response to the macroeconomic volatility of the early sample period.

Even though the belief coefficients display limited variation around the baseline, this does not mean that adaptive agents' expectations are unresponsive to changing economic conditions. Since the AR(2) forecasting rule takes current economic conditions as its input each period, it generates time-varying expectations even when its coefficients remain anchored, and thus appears to provide a sufficiently flexible approximation of the true dynamics. It is therefore instructive to examine how closely the perceived law of motion tracks the actual law of motion, the model's one-period-ahead prediction conditional on last period's state variables rather than realised values. We examine this in Figure 1.

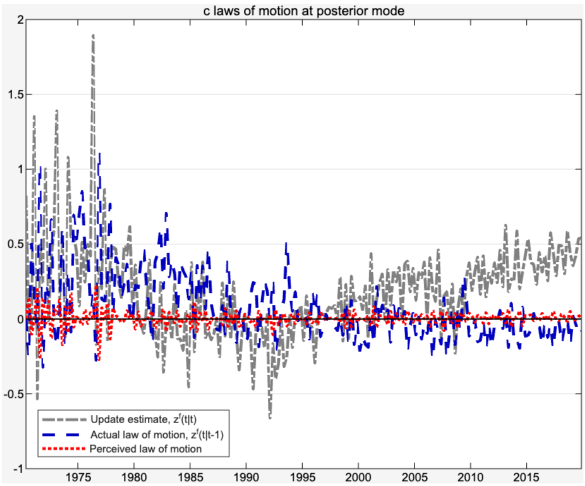
The laws of motion for inflation π and output y allow us to link the model's dynamics to well-known historical episodes. Both series display elevated levels and high volatility from the early 1970s through to the early 1990s, reflecting the OPEC oil shocks, the associated stagflation, and the subsequent European expansion and integration boom, before stabilising around their steady-state values during the great moderation. The global financial crisis and European sovereign debt crisis then push the ALM into increasingly negative territory. For output, the PLM tracks the ALM closely throughout, suggesting that a static AR(2) provides an adequate approximation of the model's dynamics. Inflation tells a different story. A persistent gap between PLM and ALM reveals the approximation error inherent in rule-of-thumb forecasting. Agents systematically underestimate the magnitude of inflation during the high-inflation episode of the 1970s and 1980s, while the pattern reverses in the post-crisis period, where deflationary pressures pull the ALM below what agents' backward-looking rule anticipates. For consumption, the law of motion fluctuates around zero throughout the sample with no clear link to historical episodes, though the elevated volatility in the 1970s and 1980s still reflects the macroeconomic



(a) Output growth (y)



(b) Inflation (π)



(c) Consumption (c)

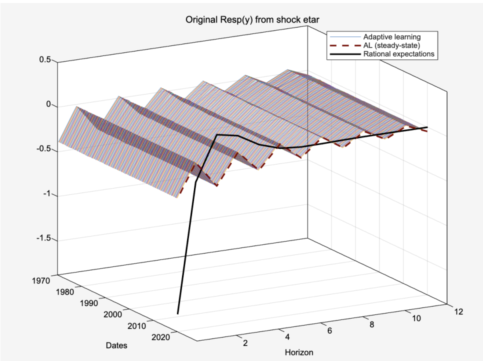
Figure 1: Laws of motion for consumption, inflation and output at the DSGE posterior mode over time.

Notes: The figure plots the perceived law of motion (red dotted), the actual law of motion $z^f(t|t-1)$ (blue dashed), and the updated estimate $z^f(t|t)$ (grey dashed) across the sample period 1970–2019. Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025).

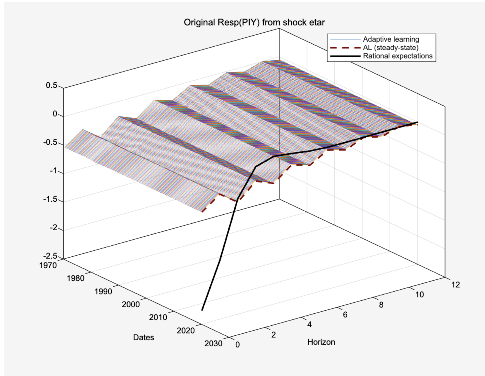
turmoil of the period. The PLM remains tightly anchored around zero but converges closer to the ALM after the 2000s as overall volatility declines.[9]

5.4 Monetary policy transmission

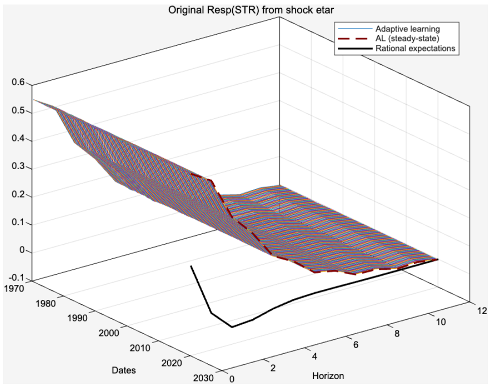
Impulse response functions (IRFs) are a central tool for understanding how exogenous shocks propagate through the economy in NK DSGE models. Since expectation formation plays a key role in shock transmission, and since AL fundamentally alters how agents perceive and respond to shocks, comparing the IRFs of the RE and AL specifications offers direct insight into how the choice of expectation formation assumption shapes macroeconomic dynamics. Of particular interest to central banks is the response to a monetary policy tightening: how strongly and how persistently does a one-standard-deviation increase in the policy rate affect output and inflation, and over what horizon does the effect dissipate.



(a) Output Growth



(b) Inflation



(c) Short-Term Rate

Figure 2: NK DSGE model impulse response functions to a monetary policy contraction.
Notes: The black line shows the RE impulse response, which is time-invariant by construction. The red surface displays the AL impulse response across calendar time.
Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025).

[9]The gap between the laws of motion for consumption and output is somewhat puzzling given that $y_t = c_t + \epsilon_t^g$ with ϵ_t^g fluctuating around zero. This likely reflects a disconnect inherent in the model's design: while the observed time series used for estimation are inflation, output growth, and the interest rate, the forward-looking variables agents actually forecast are inflation, output, and consumption (all in deviation from steady state). This might introduce an inconsistency between the variables driving estimation and those entering the learning problem.

Figure 2 presents the impulse response functions of output growth, inflation, and the euro area short-term rate to a one-standard-deviation monetary policy tightening shock, estimated over the sample period 1970–2019. The three-dimensional surface follows naturally from the time-varying solution matrices in equation 48, which depend on the evolving belief parameters β_t and thus vary with calendar time - with the impulse response horizon on one axis and calendar time on the other.[10] To isolate the role of expectation formation, structural parameters are fixed to the AL posterior estimates for both the RE and AL IRFs.

For output growth (Figure 2a), the RE model produces a sharp contraction on impact, followed by a quick recovery that briefly overshoots the steady state before returning to zero within approximately four quarters. The AL model, by contrast, generates a considerably more muted initial response but exhibits substantial persistence, with oscillating deviations around the steady state remaining visible even after three years. Inflation (Figure 2b) displays a qualitatively similar pattern. RE produces a stronger initial decline that reverts cleanly to steady state, while AL yields a more gradual and persistent disinflationary path that approaches zero from below. For the short-term rate (Figure 2c), the pattern reverses. The RE model implies a relatively modest and short-lived increase in the policy rate that dissipates within about a year, whereas the AL model generates a stronger and considerably more persistent positive deviation from steady state, remaining elevated well beyond the three-year horizon.

The differing responses between expectation formation schemes reflects fundamental differences in how agents process and propagate the monetary shock. Under rational expectations, agents immediately and correctly perceive the full future path of the shock's dissipation. Since the monetary policy shock has an autoregressive coefficient of only $\rho_r = 0.1$, rational agents know the shock will dissipate within a quarter, and adjust their behaviour accordingly, producing a large front-loaded response as they anticipate a quick return to steady state. This contemporaneous reaction is further amplified by the estimated monetary volatility parameter σ_r, which is roughly twice as large under RE as under AL, meaning that a one-standard-deviation shock enters the system with considerably greater force under rational expectations. Under adaptive learning, update their expectations incrementally using a perceived law of motion estimated from past data, implying that the shock becomes incorporated into their forecasting rule only with a delay. This expectational stickiness dampens the initial response (agents do not immediately internalize the full magnitude of the shock) but simultaneously prolongs the propagation, as the shock feeds into agents' expectations of future inflation and output, sustaining the impulse over subsequent periods.

To assess which of the two specifications generates more empirically plausible dynamics, we complement the comparison with an external benchmark: local projections (LP), which identify impulse responses directly from the data without imposing the equation structure of a DSGE model.[11] Local projections (Jordà, 2005) identify impulse responses by directly

[10] The limited variation of the AL IRFs across calendar time is a direct consequence of the small estimated ρ_{AL}: since belief coefficients remain tightly anchored around the RE-consistent baseline throughout the sample, as documented in section 5.3, the time-varying solution matrices are nearly constant, producing IRF surfaces that are stable across time. For the RE model, which has no time-varying belief parameters, the IRF surface is by construction flat across calendar time.

[11] This comparison also serves as a robustness check on the likelihood-based results, which rely on a Laplace approximation of the log marginal likelihood. Given the non-linear, time-varying structure of the AL model, this approximation may be unreliable. Local projections provide a complement as they are

regressing the outcome variable at each future horizon h on the shock of interest today, yielding one coefficient β_h per horizon that traces out the impulse response function. The specification controls for lags of both the dependent variable and the shock, thus accounting for existing trends. Since the error term ξ_{t+h} appears across multiple horizons by construction, it is serially correlated, and we therefore use Newey-West standard errors to avoid biased standard errors.

To avoid endogeneity, we do not extract the monetary policy shock directly from the estimated Taylor rule residual but instead use the high-frequency identification approach of Gertler and Karadi (2015). The idea is to observe residual movements in asset prices in a small window around a policy announcement, which, given the short time frame, can be attributed solely to the unexpected component of the announcement. This residual movement thus yields an exogenous monetary policy shock. Since policy announcements may also convey information about the central bank's economic outlook, we disentangle the two components using the euro area shocks of Jarociński and Karadi (2020), who explicitly account for this informational confounder.[12] A more detailed description of the local projection specification and the instrument is provided in Appendix B.4.

We estimate local projections for the two forward-looking variables central to our model, log-linearised output growth and quarterly inflation, on the same euro area dataset used for the DSGE estimation, identified using the high-frequency monetary policy surprises of Jarociński and Karadi (2020). Figure 3 plots the resulting impulse response functions alongside the DSGE model IRFs under RE and AL, and shows the dynamic response of each variable to a one-standard-deviation contractionary monetary policy shock.

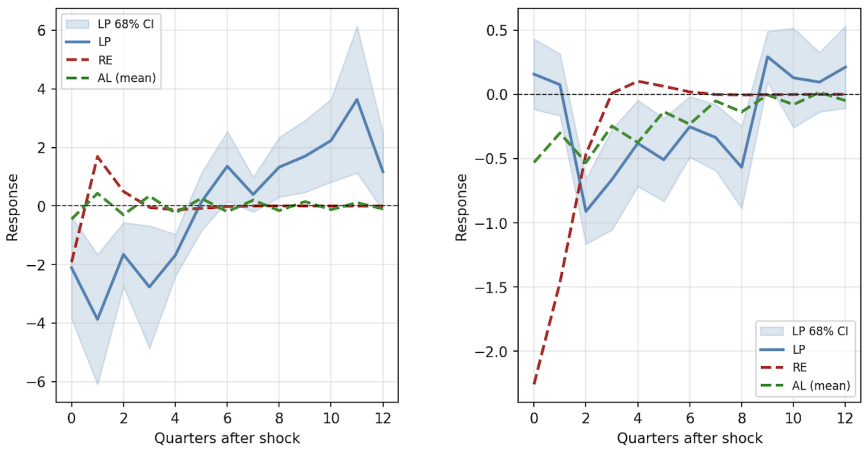


Figure 3: Impulse response functions to a contractionary monetary policy shock for output growth (left) and inflation (right).

Notes: LP benchmark vs. RE and AL DSGE models. ECB monetary policy shock, 1999 Q1–2019 Q4. Shaded bands denote Jordà (2005) 68% confidence intervals.

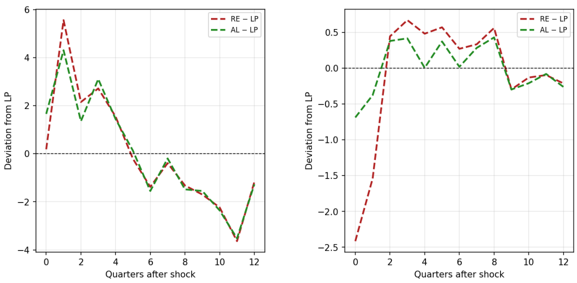
Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025) and Jarociński and Karadi (2020) monetary policy shocks.

untargeted by optimization and identified through plausibly exogenous high-frequency shocks.

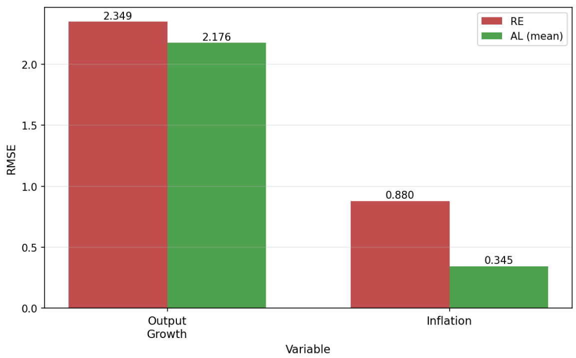
[12]These high-frequency identified shocks are conceptually different from the structural monetary policy shock ϵ_t^r in our DSGE model. The LP comparison should therefore be interpreted as a plausibility check rather than a formal test. However, Castellanos (2025) argues that LP benchmarking can still be informative under imperfect shock correspondence.

The estimated local projections following a contractionary monetary policy shock is shown in Figure 3. For the LP responses, output growth falls sharply on impact and remains negative throughout the first year before reverting toward zero and then briefly turning significantly positive. Turning to inflation, the LP response is initially insignificant before turning negative and reaching a bottom of of approximately -0.9% at $h = 2$, after which it gradually disappears, becoming statistically indistinguishable from zero by around eight quarters. Both responses are broadly consistent with the standard findings on monetary transmission that a surprise rate increase dampens demand and compresses output and inflation.

To systematically compare the RE and AL model IRFs against the LP benchmark, Figure 4a plots the deviation of each DSGE specification from the LP estimate for each quarter, while Figure 4b summarises these deviations based on the root mean squared error.[13]



(a) DSGE deviations from LP benchmark for output growth (left) and inflation (right) as response to ECB monetary policy shock, 1999 Q1-2019 Q4, LP: Jordà (2005)



(b) RMSE of DSGE IRFs relative to LP benchmark

Figure 4: Comparison DSGE vs. LP IRF differences in response to a monetary policy contraction

Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025) and Jarociński and Karadi (2020) monetary policy shocks.

For output growth, the performance of the two models are equally poor. They both fail to reproduce the initial persistent fall in output that the LP impulse response exhibit. The RE model closely matches the LP on impact, producing a contraction of similar magnitude but reverts back to zero too quickly and then briefly swings over in the positive direction. The AL model, by contrast, underestimates the response and lies close to zero for the whole period. The advantageous fit for RE during $h = 0$ is cancelled out by its worse fit later on, resulting in very similar, and large, RMSEs for AL and RE.

For inflation, the picture is clearer. Neither DSGE specification manages to reproduce the hump-shaped response seen in the local projection, but the two models fail in distinctly different ways. The RE model generates a strong contemporaneous decline that reverts to zero within around four quarters, exhibiting the same pattern of a strong front-loaded response with insufficient persistence seen for output growth. The AL model also misses the LP's slow initial response, but generates a more gradual and persistent convergence back to steady state that better follows the shape of the empirical benchmark, and it

[13] $\text{RMSE} = \sqrt{\sum_t \left(IRF_{t,\,\text{data}} - IRF_{t,\,\text{model}}\right)^2 / T}$

manages to track the LP estimate closely from around $h = 3$ onward. That AL improves the model's ability to replicate the empirically identified monetary transmission dynamics for inflation is further confirmed by the RMSE, which is less than half as large for AL as for RE.

5.5 Out-of-sample forecast performance

A core application of DSGE models is forecasting macroeconomic developments, making out-of-sample forecast performance an important dimension to evaluate our two model specifications along. To do this, we split the sample into an estimation period and a forecasting period, comparing the models' forecasts against realised data.

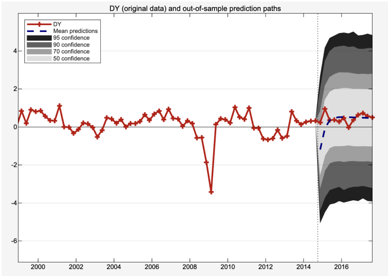
The out-of-sample forecasting performance is evaluated by estimating both models on the 1970Q2-2014Q3 sample and generating 12-quarter-ahead forecasts from 1,000 simulated paths per parameter draw for output growth, inflation, and the short-term interest rate. A 12-quarter horizon is a standard choice in the DSGE forecasting literature, as it is a window over which structural models can be expected to outperform naive benchmarks (Slobodyan and Wouters, 2012). The choice of estimation sample end date is largely arbitrary. However, as we confirm below, shifting this cut-off do not change the qualitative comparison between RE and AL substantially.

A visual inspection of Figure 5 reveals that neither model forecasts inflation or the short-term interest rate well at all.[14] Both specifications predict inflation that is too negative and a short-term rate that is too positive relative to the realized values. The reason for that might be the simplicity of our small-scale framework likely limits absolute forecast accuracy relative to larger specifications such as Warne (2026), which benefits from a richer shock structure and more observables. The picture is better for output growth. The RE model produces a sharp initial drop followed by a rapid reversion toward the actualized values, capturing the broad level around which output growth fluctuates but missing the short-run dynamics. The AL model performs better on the short-run dynamics. It remains in the right region in the first year of the forecast horizon which is consistent with the more slow-moving dynamics documented for impulse responses in section 5.4.

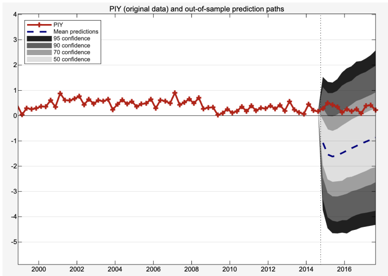
To complement the visual inspection with a quantitative comparison, we compute marginal predictive likelihoods for both the RE and AL models across multiple forecast horizons (h = 1, 2, 4, 8, and 12) and estimation sample specifications (1995,2006,2014), with results displayed in Table 4 for each of the three observed variables. The average across sample specifications broadly mirrors the findings of Slobodyan and Wouters (2012). Two patterns emerge. First, the out-of-sample performance of the AL model is generally competitive with RE as marginal predictive likelihoods are not systematically higher for the RE specification, confirming that the forecasting advantage of adaptive learning documented in larger models carries over to our parsimonious framework.

Second, a horizon-dependent pattern is visible within the results. At short horizons (h = 1, 2, and 4), the AL model outperforms RE for output growth and inflation, reflecting the value of AL when the forecast window is short enough for learning to remain informative.

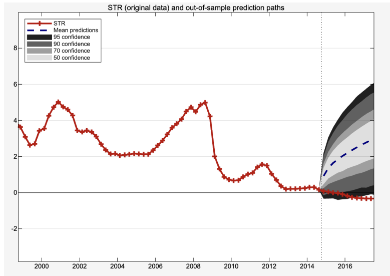
[14]The poor forecast performance for both inflation and the short-term interest rate likely reflects the ECB's unconventional monetary policy measures introduced after 2012, including large-scale asset purchase programmes that pushed the effective monetary policy stance well below what the nominal short-term rate suggests. Shadow interest rate estimates for the euro area during this period are significantly negative (Wu and Xia, 2016), a dimension our three-shock model has no mechanism to capture.



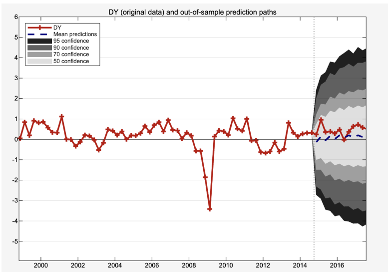
(a) dy – Rational Expectations



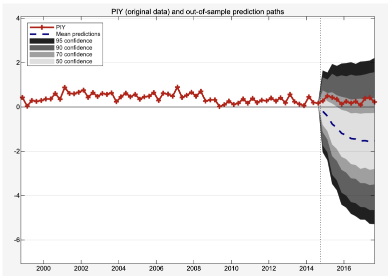
(b) π – Rational Expectations



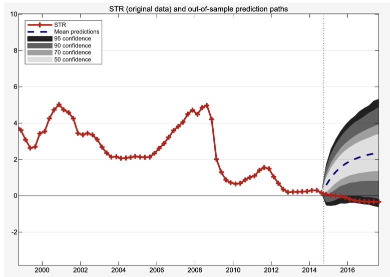
(c) r – Rational Expectations



(d) dy – Adaptive Learning



(e) π – Adaptive Learning



(f) r – Adaptive Learning

Figure 5: 12-step ahead forecasts for output growth (dy), inflation (π), and short-term interest rate (r) under Rational Expectations (top) and Adaptive Learning (bottom).

Notes: The red solid line shows the actual realised values, the blue dashed line denotes mean predictions, and shaded bands represent posterior predictive uncertainty at the 50%, 70%, 90%, and 95% confidence levels.

Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025).

Table 4: Marginal Predictive Log-Likelihoods: Rational Expectations vs. Adaptive Learning

	Rational Expectations				**Adaptive Learning**			
	1996	2005	2014	Avg.	1996	2005	2014	Avg.
Output Growth (Δy)								
$h = 1$	−1.58	−1.95	−1.85	−1.79	**−1.20**	**−1.45**	**−1.24**	**−1.30**
$h = 2$	−1.76	−1.81	−1.81	−1.79	**−1.52**	**−1.43**	**−1.51**	**−1.49**
$h = 4$	−1.76	−1.77	−1.72	−1.75	**−1.61**	**−1.60**	**−1.54**	**−1.58**
$h = 8$	**−1.72**	**−1.72**	**−1.72**	**−1.72**	−1.65	−1.65	−1.66	−1.65
$h = 12$	**−1.76**	−2.18	**−1.72**	−1.89	−1.75	**−2.16**	**−1.72**	**−1.88**
Inflation (π)								
$h = 1$	−1.20	−1.27	−1.72	−1.40	**−0.83**	**−0.83**	**−0.98**	**−0.88**
$h = 2$	−1.32	−1.63	−2.25	−1.73	**−0.92**	**−0.99**	**−1.35**	**−1.09**
$h = 4$	−1.38	−1.62	−2.10	−1.70	**−1.28**	**−1.36**	**−1.68**	**−1.44**
$h = 8$	**−1.43**	**−1.65**	**−1.80**	**−1.63**	−1.49	−1.71	−1.95	−1.72
$h = 12$	**−1.46**	**−1.53**	**−1.66**	**−1.55**	−1.55	−1.67	−1.98	−1.73
Short-term Interest Rate (r)								
$h = 1$	−0.52	−0.55	−1.32	−0.80	**−0.44**	**−0.37**	**−0.76**	**−0.52**
$h = 2$	**−0.83**	−0.79	−1.96	−1.19	−0.84	**−0.69**	**−1.42**	**−0.98**
$h = 4$	**−1.05**	−1.05	−2.44	−1.51	**−1.04**	**−0.96**	**−2.18**	**−1.39**
$h = 8$	−1.60	−1.54	−3.19	−2.11	**−1.80**	**−1.47**	**−2.95**	**−2.07**
$h = 12$	−1.77	**−1.40**	−3.56	−2.24	**−1.96**	−1.36	**−2.93**	**−2.08**

Notes: Marginal predictive log-likelihoods evaluated at forecast origins 1996Q4, 2005Q4, and 2014Q3 for horizons $h = 1, 2, 4, 8, 12$ quarters ahead. "Avg." denotes the simple average across the three forecast origins. Bold entries indicate the better-performing model at each horizon and origin. Higher values (less negative) indicate better forecast performance.

Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025).

At longer horizons, however, RE recovers and tends to dominate for inflation in particular. This is consistent with the warning in Slobodyan and Wouters (2012) that Euler-equation learning, which only forms expectations one period ahead, may struggle to be informative over longer forecast horizons. The short-term interest rate tells a different story. Here, AL maintains a forecasting advantage even at longer horizons. A plausible explanation is that during periods of prolonged interest rate stability, such as the great moderation and the near-zero rate environment following the global financial crisis, the AR(2) rule-of-thumb forecaster simply extrapolates recent stability forward. RE agents, by contrast, anticipate mean reversion implied by the model's structural dynamics, which proves counterproductive when rates remain persistently flat for reasons the model did not anticipate.

5.6 Robustness checks

The results presented above depend on a range of assumptions and model choices. To ensure that our findings are not artifacts of any single specification decision, we conduct thorough robustness checks by systematically varying the model features that most plausibly affect the results. The following subsection summarizes the effects of alternative estimation samples, changes in model calibration and estimation, and modifications to the adaptive learning specification. We focus on the marginal likelihoods as the primary criterion for robustness, together with changes in the nominal rigidity and learning parameters, which are central drivers of model dynamics under adaptive learning and on which the conclusions in the discussion section 6 are based on.

Our baseline estimation sample runs from 1970Q2 to 2019Q4. Given existing evidence on the sensitivity of adaptive learning models to changes in sample periods (Milani, 2007), we assess robustness to sample changes. The broad finding is that results are stable for minor adjustments but sensitive to changes that alter the composition of macroeconomic regimes in the sample. Removing one year from either end leaves our key findings unchanged, that are that AL outperforms RE on marginal likelihood, ξ_p remains near 0.35, and ρ_{AL} stays around 0.3. The same applies to cutting the last decade (2010-2019). Dropping the first decade (1970-1979), however, substantially changes the estimates. While AL still dominates RE on marginal likelihood, σ_l increases to around 4.3, ξ_p reverts to 0.7, and ρ_{AL} falls to 0.17. This highlights the OPEC episode as an important driver of the baseline result that learning decreases Calvo price stickiness. This becomes even clearer when isolating the volatile OPEC subsample (1970-1979), which raises ρ_{AL} to 0.41 and proves stable for windows extending to 1985, while the great moderation period (1980–2005) halves ρ_{AL} to 0.20 and yields the strongest AL advantage on marginal likelihood (-316 vs. -655). Together, this shows how the estimated degree of learning tracks the volatility of the estimation window, with our baseline ρ_{AL} of 0.28 being roughly in the middle of these two extremes. One might further argue that the marginal likelihood gains are driven by the volatility of our estimation window rendering the results specific to those episodes. Restricting the sample to a single monetary policy and exchange rate regime by taking the ECB subsample from 1999Q1 to 2019Q4 addresses this concern directly: the AL model continues to yield substantial improvements in marginal likelihood while preserving a low estimated ρ_{AL}, lending further support to the robustness of our findings. Extending the sample through the COVID period causes gradient explosion across virtually all specifications, as the model's simple three-shock structure, together with the linear approximation around its steady state, most likely means that it is incapable of handling complex shocks of that magnitude.

A more fundamental robustness check tests external validity across regions by re-estimating the model on US data. Using the Slobodyan and Wouters (2012) time series from 1970-2009, the marginal likelihoods for RE and AL are nearly identical (-732 vs. -735). This pattern also holds when restricting our euro area sample to the same period, suggesting this convergence is attributable to the exclusion of the post-2009 years rather than to the data region. Table 5 reports posterior estimates under RE and AL for both areas. The structural parameters shift from RE to AL in broadly the same way for the US data as for our original euro area data, which supports the generalisability of our parameter specific conclusions. Most notably, price stickiness ξ_p remains substantially lower under AL than RE in both cases and the belief persistence ρ_{AL} remains small on US data.

Table 5: Posterior Mode Estimates – US vs. Euro Area Data

	US Data		**Euro Area Data**	
Parameter	**RE**	**AL**	**RE**	**AL**
σ_c	1.576	4.378	2.074	4.124
σ_l	1.712	0.178	2.486	0.197
ξ_p	0.742	0.437	0.653	0.351
$\imath_{hp}$	0.182	0.799	0.172	0.796
ρ_{mp}	0.430	0.673	0.533	0.787
ψ_y	0.201	0.142	0.190	0.150
$\bar{\pi}$	0.616	0.551	0.580	0.537
β^{-1}	0.218	0.163	0.183	0.145
$\bar{\gamma}$	0.405	0.178	0.405	0.234
ρ_g	0.912	0.957	0.951	0.957
ρ_a	0.919	0.994	0.903	0.992
σ_g	2.850	0.614	2.528	0.417
σ_a	1.003	1.146	0.661	0.847
σ_r	1.669	1.241	1.230	0.792
ρ_{AL}	—	0.335	—	0.324

Notes: All values are posterior modes.
Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025) and Slobodyan and Wouters (2012) US time series.

We also assess robustness to alternative calibration and estimation choices. Freeing both r_π and ρ_{MP} for estimation preserves the AL advantage on marginal likelihood, but yields structurally implausible estimates (e.g. $\sigma_l = 5.7$ and $r_\pi = 2.38$) alongside unreasonable posterior mode curvature[15] that are difficult to reconcile with the literature (Slobodyan and Wouters, 2012). This supports fixing these parameters to theoretically grounded values, which yields a more stable and interpretable model. Fixing ρ_{AL} at the extremes of its range confirms the sensitivity of the AL specification. Values of 0.8 or above cause gradient explosion, while $\rho_{AL} = 0$ (pure rule-of-thumb forecasting) can be reconciled with the data and still yields an AL advantage on marginal likelihood, albeit with more extreme estimates elsewhere (e.g. $\sigma_l = 6.06$). More generally, the AL model

[15]The posterior mode curvature is unreasonable for most parameters which undermines the reliability of the Laplace approximation underlying our marginal likelihood calculation and casting doubt on the reported Bayes factor.

is considerably less numerically stable than its RE counterpart, frequently encountering gradient explosion and poor convergence. This instability can sometimes be mitigated by reinitialising from the RE posterior,[16] though this should be treated as a practical fix rather than a principled solution.

We also assess robustness to the specification of the PLM as the structural assumption governing how agents form expectations. Replacing the baseline univariate AR(2) with alternative forecasting rules yields informative variation in model fit and parameter estimates. Including all three observables (dy, π, r) as regressors in an AR(2) yields a marginal likelihood of -807, improving on RE (-818) but falling short of the baseline AR(2). Price stickiness ξ_p reverts toward RE levels and σ_l rises substantially, suggesting that richer information sets alter how persistence is distributed across the model's structural parameters. Switching to a univariate AR(1) improves fit markedly (-632) but similarly produces high ξ_p and σ_l (4.33), while including all observables as AR(1) regressors yields a marginal likelihood of -821, just above the RE benchmark which indicates that providing agents with more information does not necessarily improve aggregate model fit. This last finding is consistent with Vázquez and Aguilar (2021), who show that enriching the PLM with term structure information does not uniformly improve model fit, which suggests that the relationship between information set size and aggregate performance is not monotonic.

Across all PLM specifications, the estimated degree of learning remains modest, lending support to the interpretation that bounded rationality rather outperforms active belief updating under different expectation formations and adding robustness to our results. That said, the sensitivity of structural parameter estimates to PLM choice warrants caution. Different specifications tell meaningfully different stories about model dynamics raising the question which of the arbitrary structural PLM choices is the most suitable. The justification for our baseline univariate AR(2) rests on its parsimony and ability to capture a constant, trend, and persistence in a single compact rule, consistent with its predominance in the literature (Slobodyan and Wouters, 2012; Warne, 2026). The key reassurance for our results is that AL remains competitive against its RE counterpart across all specifications considered, consistent with Vázquez and Aguilar (2021), and that limited learning is a robust feature rather than an artifact of any single PLM choice.

Following Slobodyan and Wouters (2012), the baseline belief specification and the initialisation of the Kalman filter are natural robustness targets given their centrality to the model's empirical performance. When calibrating both to the RE-consistent AR(2) process estimated over shorter pre-sample windows (1970–1976, 1970–1979, and 1970–1989), the AL model retains a superior marginal likelihood across all three specifications and ρ_{AL} remains relatively stable and low (below 0.45), while the structural parameters vary unpredictably and price stickiness ξ_p is generally pushed above 0.7. To isolate the initialisation channel, we hold the anchor fixed at its full-sample RE-consistent value while varying only the starting condition. Using the same training samples reveals greater sensitivity: the marginal likelihood falls below the RE benchmark for the 1970–1976 initialisation but recovers and outperforms RE for 1970–1979. Belief persistence ρ_{AL} is robustly low across both (around 0.11) which likely reflects the optimiser pushing beliefs rapidly back toward the well-specified anchor to compensate for the poor starting condition. Structural parameters

[16]Importantly, these findings highlight the central role of initialisation. Even without changing distributional assumptions, different starting values can materially alter posterior estimates and model implications, likely reflecting the difficulty of optimising over a high-dimensional likelihood surface prone to local optima.

again prove unstable. For the 1970–1989 initialisation, the estimation encounters an exploding gradient.

Taken together, the robustness checks confirm two things: the AL advantage on marginal likelihood is robust across most feasible specifications, and variation in ρ_{AL} follows a predictable pattern tied to sample volatility and quality of $\bar{\beta}$. Structural parameter estimates are more sensitive to specification choices. This is particularly true around the OPEC period, which emerges as a central driver of baseline results. They should thus be interpreted with caution. That said, our baseline assumptions are well-rationalised, consistent with the literature, and replicate in broadly the same pattern on US data, partially mitigating these concerns.

6 Discussion

6.1 Adaptive learning improves model fit

Having discussed the estimation and outcomes of the small-scale model of Del Negro and Schorfheide (2013) under two expectation assumptions, we now draw the evidence together to answer our central research question: does introducing adaptive learning improve the empirical fit of a small-scale New Keynesian DSGE model for euro area data? We address this along two dimensions, in-sample fit and out-of-sample forecast performance, allowing us to draw conclusions that are robust to overfitting concerns while also speaking to the practical value of learning for DSGE-based forecasting.

On both dimensions, we find that adaptive learning improves upon the rational expectations benchmark. The in-sample comparison is unambiguous, as documented in section 5.2, the log marginal likelihood is substantially higher for the AL model (-767) than for RE (-819), yielding a log Bayes factor of 52. On the scale of Kass and Raftery (1995), this constitutes decisive evidence in favour of the adaptive learning specification. Crucially, the marginal likelihood already penalizes model complexity through its integration over the full parameter space, hence, this advantage cannot be attributed to the additional flexibility that learning introduces. As shown in section 5.6, this result holds across a range of alternative specifications, such as different estimation samples, belief models, and initialisation choices, suggesting that the finding is not specific to any single modelling decision.

A second, complementary assessment of in-sample fit comes from comparing model-implied impulse response functions to the high-frequency identified local projections described in section 5.4. Since DSGE models are among the primary tools central banks use to assess the transmission of monetary policy, the ability to replicate empirically plausible impulse responses is arguably as important as the likelihood-based fit. Here the picture is more nuanced. For output growth, neither specification clearly dominates, and both models struggle to reproduce the empirical benchmark. For inflation, however, the AL model generates substantially more realistic propagation dynamics, with an RMSE less than half that of the RE specification.

These findings are consistent with the existing literature. For adaptive learning in general, Milani (2011) and Dizioli and Wang (2024) find that it improves fit. For Kalman filter learning, Slobodyan and Wouters (2012) and Rychalovska et al. (2025) demonstrate that its implementation raises the marginal likelihood in a medium-scale model estimated on US data. Our results show that these improvements carry over to a small-scale framework estimated on euro area data up to 2019, showing that the benefits of adaptive learning are not specific to the US or to earlier sample periods. Beyond replicating the marginal likelihood result, we are, to our knowledge, the first to benchmark adaptive learning IRFs against high-frequency identified local projections. The finding that learning improves inflation dynamics under monetary policy shocks is therefore a unique contribution, suggests that adaptive learning enhances not only the likelihood-based fit, which is the direct target of estimation, but also model features such as impulse responses, that are not directly optimized during estimation.

The out-of-sample evidence, reported in section 5.5, is more mixed but broadly in favour of adaptive learning. Evaluated through marginal predictive likelihoods across multiple forecast origins and horizons, the AL model is overall competitive with RE and outperforms it at short horizons for both output growth and inflation. At longer horizons, RE recovers and outperforms AL, consistent with the limitations of Euler-equation learning, which forms expectations only one period ahead and therefore struggles to reliably form expectations over longer horizons. Notably, this horizon-dependent pattern is broadly consistent with findings in the adaptive learning literature (Rychalovska et al., 2025; Warne, 2026).

Taken together, the evidence establishes adaptive learning as a viable alternative to rational expectations in the small-scale DSGE setting, and suggests, together with previous research, that AL is generally a competitive specification across many different kinds of DSGE models. Along the dimensions where the two specifications differ most sharply, namely in-sample fit and the replication of monetary policy transmission dynamics, AL is the preferred specification. Where the advantage is less clear, AL remains competitive rather than substantially inferior, meaning that the case for adaptive learning does not rest on selective evidence but holds for all evaluation criteria we consider.

That AL improves model performance is more than a statistical result. The improved fit suggests that adaptive learning offers a more accurate description of how households and firms actually process information and form expectations in the real world, which motivates a closer examination of how model dynamics change when rational expectations are replaced by adaptive learning. Section 6.2 thus investigates how adaptive learning generates endogenous persistence, and to what extent this reflects the real-world sources of sluggishness in macroeconomic dynamics. Section 6.3 then turns to the nature of the learning process itself, examining why adaptive learning improves model fit despite the estimated degree of learning in our model being very limited.

6.2 Endogenous persistence under adaptive learning

Macroeconomic time series move with sluggishness and exhibit high persistence. Thus, any credible macroeconomic model must be capable of generating a realistic amount of it. This has proven a persistent challenge for rational expectations models, which tend to produce impulse responses that revert to steady state faster than the data suggests (Dizioli and Wang, 2024). The standard remedy, exemplified by Christiano et al. (2005) and Smets and Wouters (2007), has been to incorporate a wide range of frictions and highly persistent shock processes, including habit formation, investment adjustment costs, price stickiness, price indexation, and near-unit-root shock autocorrelations. While effective at matching the data, these devices are frequently criticised as mechanical and difficult to reconcile with micro-evidence (Carroll, 2003; Coibion and Gorodnichenko, 2015).

Our findings suggest that adaptive learning constitutes a superior source of persistence, and two pieces of evidence support this claim. First, the endogenous persistence introduced by learning seems to be capable of substituting for the mechanical frictions that rational expectations models typically require. In our parsimonious framework, the Calvo pricing mechanism is the key source of persistence, generating sluggishness through the price stickiness coming from the fraction of firms prohibited from re-optimising their prices each period. As shown by the posterior estimates in section 5.1, introducing adaptive learning halves the Calvo stickiness parameter from $\xi_p = 0.711$ under RE to $\xi_p = 0.350$, indicating that the model relies considerably less on price stickiness for its fit when learning is present. The AL estimate moves the Calvo parameter toward micro-evidence on price stickiness, somewhat undershooting the 0.41–0.50 range implied by Bils and Klenow (2004) for US CPI scanner data, while the RE estimate of 0.71 aligns with posterior values typically found in canonical medium-scale models (Christiano et al., 2005; Smets and Wouters, 2007). Second, the persistence introduced by adaptive learning yields impulse responses to monetary policy shocks that more closely match the empirical benchmark. As shown in section 5.4, the expectational stickiness both dampens the initial impact response and produces a slower, more gradual reversion to steady state. Both of these features bring the impulse responses closer to the empirical benchmark from the local projections, showing that the persistence matters for improving the realism of the impulse responses. To understand whether these improvements reflect something genuine about the economy, we now turn to the mechanisms through which adaptive learning generates endogenous persistence.

We find that the source of endogenous persistence is expectational stickiness, which in turn arises from the backward-looking nature of expectation formation in our model. In the AL setup, agents forecast future variables using only their lagged values, making expectations completely backward-looking. Thus, rather than relying on any understanding of the underlying structure of the economy, agents simply extrapolate historical AR(2) patterns forward.[17] Thus, when hit by a shock, agents cannot anticipate the speed at which it will fade away, and can therefore not directly incorporate the future path of the shock into their expectations. Instead, the shock enters agents' expectations by affecting output and inflation in the current period, and then feeds into their forecasting rules in subsequent periods, since it is only through its effect on past values of these variables that the shock is incorporated into adaptive learners' backward-looking expectations. In other words, rather than a shock being immediately and fully processed as under RE, it is absorbed gradually into beliefs, which both dampens the initial response and prolongs its subsequent propagation, giving rise to a type of expectational stickiness.

Our results are consistent with the main findings of the adaptive learning literature. Orphanides and Williams (2005), Milani (2007), and Slobodyan and Wouters (2012) find that excess price stickiness decreases with the introduction of AL. They also explain these findings with the expectational stickiness introduced by learning generating endogenous persistence that can render many of these mechanical sources redundant. However, their richer model specifications allow them to document a broader substitution of sources of persistence. Slobodyan and Wouters (2012) find that the wage indexation as well as markup shock processes for prices and wages become redundant, while Milani (2007) shows that habit formation and price indexation is no longer needed to match the data.

[17] This stands in sharp contrast to the RE setup, where agents are assumed to know the full structure of the economy, allowing them to form forward-looking expectations that incorporate information on how relevant variables, including shocks, will evolve.

Notably, this pattern holds even in our simple framework with only two nominal friction, highlighting that the role of adaptive learning as a source of persistence is not limited to richer model specifications. Our other finding, that impulse responses are more persistent under adaptive learning, is consistent with Dizioli and Wang (2024), who document the same result in another small-scale setting.

While we document a reduction in price stickiness under AL, the price indexation parameter $\imath_p$ moves sharply in the opposite direction. Since both parameters are standard sources of inflation persistence in New Keynesian models, their opposing movements under AL might appear contradictory. However, the two parameters generate persistence through fundamentally different mechanisms. Price stickiness reduces firms' reactivity to current cost conditions, whereas price indexation makes non-reoptimising firms mechanically carry past inflation forward, producing more backward-looking price dynamics. The AL configuration therefore shifts the source of nominal rigidity rather than eliminating it. As seen in the Phillips curve (53), the higher indexation gives past inflation a greater influence on contemporaneous inflation, while the lower Calvo parameter means that current cost conditions matter more. These two forces complement rather than cancel each other, and together reduce the weight on expected future inflation in the AL Phillips curve relative to RE, making inflation dynamics less forward-looking. The higher price indexation under AL is not a coincidence but a natural complement to the expectational stickiness introduced by the AR(2) forecasting rule, as both generate persistence through the same mechanism. Essentially, both features increase the weight of π_{t-1} in the Phillips curve, making inflation dynamics more backward-looking.[18]

While these findings are in line with the literature, it is important to assess the robustness of our own persistence result before drawing definitive conclusions. The robustness checks in section 5.6 reveal that the structural parameters are sensitive to changes in the estimation sample, calibration choices, and the specification of the learning set-up. In particular, the reduction in price stickiness seems to be driven largely by the inclusion of the OPEC period (roughly 1970–1985), suggesting that this episode plays an important role in identifying this substitution of persistence sources. Nonetheless, there are reasons to believe these findings. First, as demonstrated in the robustness checks (5.6), the same pattern emerges for US data over a different sample period (1966–2009), suggesting that the result is not limited to only the euro area. Second, the mechanism by which learning substitutes for structural persistence is well-documented, with studies such as Slobodyan and Wouters (2012) and Dizioli and Wang (2024) consistently finding that introducing adaptive learning reduces the need for mechanical frictions like price stickiness to match the data.

The shift from rational to adaptive expectations represents a fundamental change in the story the model tells about how agents understand the economy. Under RE, agents are perhaps unrealistically omniscient. In our AL specification, by contrast, agents rely entirely on backward-looking extrapolation, possessing no genuine forward-looking

[18]This complementarity is most clearly visible by substituting the explicit AR(2) forecasting rule into the AL Phillips curve (53). The lagged inflation term π_{t-1} then enters both through the indexation parameter and through the second lag coefficient of the AR(2), making the two more or less interchangeable as sources of backward-lookingness. This pattern of substitutability is consistent when comparing our estimates to those of Slobodyan and Wouters (2012). In our AL specification the second lag coefficient is approximately -0.2 and indexation is 0.845, whereas in Slobodyan and Wouters (2012) the coefficient hovers around 0.2 and indexation is considerably lower at 0.29, suggesting that the two parameters trade off against each other in delivering a similar overall degree of backward-looking inflation.

understanding of the economy's structure whatsoever. One may therefore ask whether we have simply replaced one unrealistic assumption with another, and whether the improved fit reflects a genuinely more accurate description of expectation formation or merely a more flexible way of introducing persistence. There are reasonable arguments on both sides. Against the backward-looking specification, real-world agents do monitor central bank communications, macroeconomic indicators, and professional forecasts, all of which introduce a forward-looking dimension that our AR(2) rule entirely disregards. In its favour, survey evidence consistently documents that households and firms systematically fail to form model-consistent expectations (Coibion and Gorodnichenko, 2015) and exhibit precisely the backward-looking dependence that AL captures (Roberts, 1998; Carroll, 2003). Furthermore, the fact that we find that AL improves both in-sample fit and impulse response accuracy suggests that it captures something genuine in the way agents form expectations. While this thesis' results suggest that agents' expectations are not as forward-looking as the RE assumption postulates, to what degree agents' expectations should be modelled as backward-looking remains an open question in the DSGE literature.

If part of the mechanical persistence introduced through structural parameters is substituted by expectational stickiness, this has direct implications for how monetary policy ought to be conducted. Under adaptive learning, inflation expectations exhibit inertia and become harder to control once they begin drifting from target, since agents lack the model-consistent anchoring mechanism that rational expectations provides. The muted but persistent response to monetary policy shocks we document, consistent with Dizioli and Wang (2024), suggests it is optimal for central banks to front-load their policy response to inflationary surprises, acting more aggressively early on to arrest the drift in agents' expectations before it becomes entrenched.

6.3 Bounded rationality without active learning

Our adaptive learning model is estimated with a belief persistence parameter of $\rho_{AL} = 0.278$, pulling beliefs quickly back to the RE-consistent baseline $\bar{\beta}$ and rendering agents closer to static rule-of-thumb forecasters than active learners.[19] This result is striking as a central feature that distinguishes our AL specification from RE, the ability of agents to continuously update their forecasting rules in response to forecast errors, is estimated to play only a minimal role. Two questions naturally follow. First, why does the optimiser actively prefer this parsimonious fixed AR(2) rule over a more flexible and continuously updating one? Second, if active learning plays so limited a role, why does the AL specification still outperform the more sophisticated RE counterpart?

To address the first question, it is instructive to compare our estimate to the near-unit-root $\rho_{AL} = 0.97$ reported by Slobodyan and Wouters (2012) on US data, implying that beliefs drift persistently away from the baseline over extended periods. Since Slobodyan and Wouters (2012) broadly finds similar evidence as us for improved fit and endogenous persistence, investigating the source of this discrepancy could offer insight into the role of adaptive learning in DSGE models. There are three potential explanations. First, the time series we employ may exhibit fundamentally different dynamics, in particular, dynamics that can be well approximated by a fixed AR(2) process, leaving little room for belief updating to improve fit. Second, the additional time series included in Slobodyan and

[19]While the baseline is initialised to be as close as possible to the belief system of fully rational agents, it remains a parsimonious univariate approximation of the true (model-consistent) multivariate ALM and the gap between PLM and ALM visible in Figure 1 reflects precisely this simplification.

Wouters (2012) may exhibit regime-shifts or other patterns that a fixed AR(2) cannot adequately capture, generating persistent forecast errors that push ρ_{AL} upward. Third, differences in model specification may matter independently of the data, as a a larger model with more forward-looking variables and with potentially less accurate baseline beliefs may inherently generate greater demand for active learning. While the first two explanations can be partially assessed by examining how well a fixed AR(2) describes the relevant time series, the model-based explanations are harder to disentangle, and this comparison therefore remains necessarily preliminary.

Slobodyan and Wouters (2012) estimate their model on US data spanning 1966–2008, while we use euro area data over 1970–2019, leaving open the possibility that differences in either the country or the sample period drive the discrepancy in ρ_{AL}. To assess this, we examine whether the two datasets differ in how well a fixed-coefficient AR(2) process can describe the main forward-looking variables (output growth and inflation) with full details provided in Appendix B.5. The results suggest no meaningful difference: both datasets yield a similar goodness of fit, with inflation well-described by a fixed AR(2) and output growth poorly described by one in both cases. This finding is further verified by re-estimating our DSGE model with the time series used by Slobodyan and Wouters (2012), yielding $\rho_{AL} = 0.2$ which is close to our baseline estimate and far from their original estimate of 0.97.

Since differences in data dynamics are unlikely to explain the discrepancy, the four additional time series in Slobodyan and Wouters (2012) emerge as the more plausible driver. This matters because a single belief persistence parameter governs the updating of all forecasting rules simultaneously. So if even a subset of the additional variables benefits from active learning, this pushes ρ_{AL} upward for all forward-looking variables, including those already well-described by a fixed AR(2). As documented in Appendix B.5, fitting AR(2) models to the four additional variables (labour input, consumption, investment, and wages) reveals an extremely poor fit for three of the four, suggesting that a fixed AR(2) is inadequate for these series. This contrast offers a plausible, if not conclusive, explanation for our smaller ρ_{AL}. Whereas the additional volatile series in Slobodyan and Wouters (2012) create pressure for persistent belief updating, our model's agents forecast only output growth and inflation (of which inflation is already well-described by a fixed AR(2)) leaving insufficient pressure to push beliefs away from the RE-consistent anchor. Whether this difference is attributable to model size alone, or also reflects the country and sample period, remains an open question.

Beyond time series dynamics, the estimation of belief persistence is fundamentally shaped by the structural assumptions underlying our expectation formation mechanism, specifically the specification of the baseline beliefs ($\bar{\beta}$) and the belief initialisation ($\beta_{1|0}$). As they are calibrated to be RE-consistent, a small ρ_{AL} implies that this specification endows agents with close to optimal beliefs from the outset, leaving little room for the Kalman filter to improve upon them. As documented in the robustness checks (5.6), the estimated degree of learning is consistently low (0.1-0.5) across specifications. Limited active learning implies that the baseline beliefs are given greater influence over the actual forecasting models that agents end up using, than what is the case with a unit-root ρ_{AL}. This motivates a more critical examination of the RE-consistent baseline assumption.

While this choice is standard in the adaptive learning literature (Slobodyan and Wouters, 2012; Warne, 2026), there is an inherent tension in assuming that boundedly rational agents initialise and anchor their forecasting rules around coefficients that require knowledge of the model's RE solution over the entire sample period. This implicitly endows agents at the start of the sample with a forecasting rule calibrated to data they have not yet observed, introducing a form of look-ahead bias that is difficult to reconcile with the spirit of adaptive learning.[20] That said, the RE-consistent baseline is necessarily a simplification, much like rational expectations itself, and should be judged not by literal accuracy but by whether it constitutes a useful and disciplined approximation. Indeed, as shown in Figure 1, the RE-consistent expectations that emerge from the Kalman filter optimisation leads AL agents to systematically underestimate inflation during high-inflation episodes and overestimate it during low-inflation periods, consistent with survey evidence documenting similar expectational inertia during that period (Carroll, 2003; Coibion and Gorodnichenko, 2015). Viewed this way, it represents a disciplined step toward reconciling model expectations with reality while also delivering meaningful improvements in model performance.

To address the second question of how a static AR(2) forecasting heuristic with demonstrably less accurate beliefs can outperform a fully rational RE specification, it is important to distinguish between the accuracy of agents' beliefs and the empirical fit of the model as a whole. Consistent with the broader learning literature (Vázquez and Aguilar, 2021; Rychalovska et al., 2025; Warne, 2026), our results suggest that more realistic, and thus less accurate, expectations can improve aggregate model fit. Three complementary mechanisms help explain how replacing an internally consistent but empirically questionable assumption with a more flexible and realistic alternative can consistently improve model performance.

First, replacing model-consistent expectations with a parsimonious AR(2) rule relaxes the tight cross-equation restrictions imposed under RE. In a rational expectations model, the assumption that agents know the true structural model links coefficients across equations. The IS curve, the Phillips curve, and the Taylor rule must all be mutually consistent with agents having full model knowledge simultaneously. Under AL, this rigidity is relaxed, giving the model additional flexibility in fitting the data that is independent of the persistence channel.

Second, as explored in section 6.2, limited adaptive learning introduces sufficient endogenous persistence to better match the sluggish dynamics observed in the data, without requiring unreasonable assumptions on structural parameters. More broadly, a model whose assumptions more closely reflect the actual data-generating process appears better able to improve fit by generating more realistic propagation dynamics and by aligning more naturally with microeconomic evidence on how shocks transmit through the economy.

Third, as outlined above, the flexibility of the adaptive learning framework itself constitutes an independent source of empirical improvement. By scaling belief persistence up or down through ρ_{AL}, the model can accommodate both actively updating agents and near-static rule-of-thumb forecasters, adapting to whatever the data demand without changing the underlying specification. If the time series are marked by volatility and

[20] The standard defence appeals to the stability of deep structural relationships. If the economy evolves slowly enough, agents may plausibly have internalised its regularities through long experience, making the RE-consistent initialisation not a claim about observing future data but about having learned the deep structure of a slowly evolving environment. This defence is, however, weakest precisely where it is most needed: in the kind of shifting macroeconomic environment that motivates adaptive learning in the first place.

shifting macroeconomic regimes, a high ρ_{AL} allows beliefs to drift persistently away from the baseline, enabling agents to actively track and adapt to changing dynamics. In settings like ours, where macroeconomic volatility is more limited as a consequence of the small-scale model with two relatively stable forward-looking series, a low ρ_{AL} reduces agents to near-static rule-of-thumb forecasters, yet the resulting forecasting structure alone introduces sufficient persistence to improve model fit. This versatility, we argue, is a key reason why the marginal likelihood advantage of AL proves robust across specifications in our model and across models and samples in the broader learning literature (Slobodyan and Wouters, 2012; Dizioli and Wang, 2024). While a small ρ_{AL} is not without precedent (Warne, 2026), explicitly identifying this flexibility as a distinct source of empirical improvement and tracing its implications for why AL outperforms RE in small-scale models constitutes a contribution of this paper.

7 Concluding remarks

The rational expectations hypothesis is a cornerstone of macroeconomic modelling, but has long been criticised for imposing unrealistically strong informational requirements on agents. This thesis has explored a more agnostic alternative: agents are assumed to lack knowledge of the economy's structure and instead adaptively learn to forecast key macroeconomic variables by recursively updating a simple forecasting heuristic. By combining the tractable small-scale approach of Milani (2007) with the richer Kalman filter learning of Slobodyan and Wouters (2012), we investigated how adaptive learning affects model performance and macroeconomic dynamics, with the aim of identifying model features that drive its empirical advantage over rational expectations.

We find that AL outperforms RE across multiple dimensions (both in-sample and at shorter forecast horizons), which extends the findings of Milani (2007), Slobodyan and Wouters (2012) and Warne (2026) to the small-scale Del Negro and Schorfheide (2013) framework and more recent sample periods. We further benchmark model-implied impulse responses against high-frequency identified local projections, finding that AL produces more muted but also more persistent responses that better match the empirical benchmark than RE. To our knowledge this comparison is novel in the adaptive learning literature, though results should be interpreted with caution given the shorter LP sample and imperfect comparability of shock identifications across the two approaches.

The improved fit seems to be driven by the endogenous persistence that adaptive learning introduces through backward-looking expectation formation, which reduces the model's reliance on mechanical frictions and alters the optimal conduct of monetary policy toward more front-loaded responses. Interestingly, our estimated degree of learning is limited, suggesting that the gains are not driven by active learning in itself but by the bounded rationality that the rule-of-thumb forecasting structure imposes. Our conclusion therefore differs from Slobodyan and Wouters (2012) and suggests that in small-scale DSGE models with relatively stable time series, fixed rule-of-thumb forecasting may be sufficient to capture the empirically relevant departure from rational expectations. More broadly, this highlights the flexibility of the adaptive learning framework, which can improve fit through different mechanisms depending on the model and data, consistently outperforming the rigid RE benchmark.

Our study has several limitations. A first concern is that introducing AL changes some structural parameters that should in principle be invariant to the expectation formation assumption. On the data side, we use revised rather than real-time data, which sits uneasily with the assumption that agents observe contemporaneous macroeconomic variables immediately as part of their learning process, given the publication lags typical of GDP and inflation releases. More fundamentally, the purely backward-looking nature of expectations under AL raises the question of whether we have simply replaced one unrealistic assumption with another. While survey evidence supports the backward-looking inertia that AL captures (Carroll, 2003; Coibion and Gorodnichenko, 2015), real-world agents also draw on forward-looking information such as central bank communications and professional forecasts. Extending this framework by allowing the degree of backward-looking to be endogenously determined (Deak et al., 2023), or combining it with the survey-augmented learning structure of (Rychalovska et al., 2025) in a small-scale setting, could yield a more realistic and nuanced account of expectation formation.

References

Bils, M. and Klenow, P. J. (2004). Some evidence on the importance of sticky prices. *Journal of Political Economy*, 112(5):947–985.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of monetary Economics*, 12(3):383–398.

Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *The Quarterly Journal of Economics*, 118(1):269–298.

Carvalho, C., Eusepi, S., Moench, E., and Preston, B. (2023). Anchored inflation expectations. *American Economic Journal: Macroeconomics*, 15(1):1–47.

Castellanos, J. (2025). Local projections vs. VARs for structural parameter estimation. *Bank of England Staff Working Papers*, (1116).

Chari, V. V., Kehoe, P. J., and McGrattan, E. R. (2009). New keynesian models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics*, 1(1):242–66.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy*, 113(1):1–45.

Cogley, T. and Sbordone, A. M. (2008). Trend inflation, indexation, and inflation persistence in the new keynesian phillips curve. *American Economic Review*, 98(5):2101–26.

Coibion, O. and Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, 105(8):2644–2678.

Deak, S., Levine, P., Pearlman, J., and Yang, B. (2023). Reinforcement learning in a new keynesian model. *Algorithms*, 16(6).

Del Negro, M. and Schorfheide, F. (2013). DSGE model-based forecasting. In Elliott, G. and Timmermann, A., editors, *Handbook of Economic Forecasting*, volume 2, pages 57–140. Elsevier, New York.

Dizioli, A. and Wang, H. (2024). How do adaptive learning expectations rationalize stronger monetary policy response in brazil? *Latin American Journal of Central Banking*, 5(1):100119.

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

Evans, G. W. and Honkapohja, S. (2009). Learning and macroeconomics. *Annual Review of Economics*, 1(Volume 1, 2009):421–449.

Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.

Gáti, L. (2023). Monetary policy & anchored expectations—an endogenous gain learning model. *Journal of Monetary Economics*, 140:S37–S47.

Gertler, M. and Karadi, P. (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics*, 7(1):44–76.

Goodfriend, M. and King, R. G. (1997). The new neoclassical synthesis and the role of monetary policy. *NBER macroeconomics annual*, 12:231–283.

Herbst, E. P. and Schorfheide, F. (2016). *Bayesian estimation of DSGE models*. Princeton University Press.

Ireland, P. N. (2004). A method for taking models to the data. *Journal of Economic Dynamics and Control*, 28(6):1205–1226.

Jarociński, M. and Karadi, P. (2020). Deconstructing monetary policy surprises—the role of information shocks. *American Economic Journal: Macroeconomics*, 12(2):1–43.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, 95(1):161–182.

Kass, R. E. and Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795.

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

Massaro, D. (2013). Heterogeneous expectations in monetary dsge models. *Journal of Economic Dynamics and Control*, 37(3):680–692.

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

Milani, F. (2011). Expectation shocks and learning as drivers of the business cycle. *Economic Journal*, 121(552):379–401.

Milani, F. and Rajbhandari, A. (2012). Expectation formation and monetary dsge models: Beyond the rational expectations paradigm.

Orphanides, A. and Williams, J. C. (2005). Inflation scares and forecast-based monetary policy. *Review of Economic Dynamics*, 8(2):498–527.

Poledna, S., Miess, M. G., Hommes, C., and Rabitsch, K. (2023). Economic forecasting with an agent-based model. *European Economic Review*, 151:104306.

Roberts, J. M. (1998). Inflation expectations and the transmission of monetary policy. *Finance and Economics Discussion Series*, (1998-43).

Rychalovska, Y., Slobodyan, S., and Wouters, R. (2025). Survey expectations, learning and inflation dynamics. *European Economic Review*, page 105118.

Sargent, T. (1993). *Bounded Rationality in Macroeconomics: The Arne Ryde Memorial Lectures*. Arne Ryde memorial lectures. Clarendon Press.

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.

Smets, F. and Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association*, 1(5):1123–1175.

Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97(3):586–606.

Vázquez, J. and Aguilar, P. (2021). Adaptive learning with term structure information. *European Economic Review*, 134:103689.

Warne, A. (2025). *YADA Manual — Computational Details*. Manuscript, European Central Bank.

Warne, A. (2026). Dsge model forecasting: Rational expectations versus adaptive learning. *Journal of Forecasting*.

Wu, J. C. and Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, 48(2-3):253–291.

İpek, M. and Kısacıkoğlu, B. (2025). Estimating euro area output gap dynamics: Evidence from the updated area-wide model database. *CEPR Discussion Paper No. 19913. CEPR Press, Paris London.*

A Model derivation

A.1 The household sector

Every household j chooses in time t consumption $C_t(j)$, hours worked $L_t(j)$ and bond holdings $B_t(j)$ in order to maximize the objective function given by:

$$E_t \sum_{s=0}^{\infty} \beta^s \left\{ \frac{1}{1-\sigma_c} (C_{t+s}(j))^{1-\sigma_c} \exp\left(\frac{\sigma_c - 1}{1+\sigma_l} L_{t+s}(j)^{1+\sigma_l} \right) \right\} \tag{57}$$

Here, the intertemporal utility is determined over the sum of all future expected utilities. The discount factor $\beta \in (0,1)$, common across the continuum of households, governs the degree to which they value future utility relative to the present. The household's utility depends positively on consumption $C_{t+s}(j)$ and negatively on the amount of hours worked $L_{t+s}(j)$. Notably, consumption and labour are inherently inseparable within this specification of the utility function. The way households smooth their consumption over different future periods is determined by σ_c whose inverse represents the elasticity of intertemporal substitution. A higher σ_c thus means that households increase their preference for a smooth consumption path. Labour is linked to σ_l, the inverse Frisch elasticity which measures how much households adjust their labour supply in response to a change in real wage, holding the marginal utility of wealth constant. A higher σ_l thus indicates that the labour supply is less elastic, household are less willing to substitute work across time.

The household maximizes the intertemporal utility function subject to the following budget constraint:

$$C_{t+s}(j) + \frac{B_{t+s}(j)}{R_{t+s}P_{t+s}} \leq \frac{B_{t+s-1}(j)}{P_{t+s}} + \frac{W^h_{t+s}(j)L_{t+s}(j)}{P_{t+s}} + \frac{Div_{t+s}}{P_{t+s}} \tag{58}$$

This indicates that the period's expenditures, consisting of consumption $C_t(j)$ and purchased/borrowed bonds $B_t(j)$ must be equal or smaller than the available resources. Here, $B_t(j)$ are deflated and discounted by the policy rate (nominal return) R_t. The available resources consist of the bond holdings from last period, the income from labour and dividends - all values indicated in real terms.

Thus, the household's maximization problem yields the following Lagrange equation with λ_t as Lagrange multiplier on the budget constraint:

$$\begin{aligned} \mathcal{L} = E_t \sum_{s=0}^{\infty} \beta^s \Bigg\{ & \frac{1}{1-\sigma_c} (C_{t+s}(j))^{1-\sigma_c} \exp\left(\frac{\sigma_c - 1}{1+\sigma_l} L_{t+s}(j)^{1+\sigma_l} \right) \\ & + \lambda_t \left[\frac{B_{t+s-1}(j)}{P_{t+s}} + \frac{W^h_{t+s}(j)L_{t+s}(j)}{P_{t+s}} + \frac{Div_{t+s}}{P_{t+s}} - C_{t+s}(j) - \frac{B_{t+s}(j)}{R_{t+s}P_{t+s}} \right] \Bigg\} \end{aligned} \tag{59}$$

Taking the first order conditions yields:

$$\frac{\partial \mathcal{L}}{\partial C_t} = \exp\left(\frac{\sigma_c - 1}{1+\sigma_l} L_t^{1+\sigma_l} \right) C_t^{-\sigma_c} = \lambda_t \tag{60}$$

$$\frac{\partial \mathcal{L}}{\partial L_t} = C_t^{1-\sigma_c} \exp\left(\frac{\sigma_c - 1}{1+\sigma_l} L_t^{1+\sigma_l}\right) L_t^{\sigma_l} = \lambda_t \frac{W_t^h}{P_t} \tag{61}$$

$$\frac{\partial \mathcal{L}}{\partial B_t} = \beta R_t E_t[\frac{1}{\pi_{t+1}}] = E_t[\frac{\lambda_t}{\lambda_{t+1}}] \tag{62}$$

Combining (60) and (61) yields the labour supply of households:

$$C_t L_t^{\sigma_l} = \frac{W_t}{P_t} \tag{63}$$

This represents the intratemporal optimality condition, equating the marginal rate of substitution between consumption and leisure to the real wage. Combining (60) and (62) yields in turn the Euler equation:

$$E_t[C_{t+1}{}^{-\sigma_c}] = \frac{1}{\beta R_t E_t[\pi_{t+1}^{-1}]} C_t{}^{-\sigma_c} E_t[\frac{\exp\left(\frac{\sigma_c-1}{1+\sigma_l} L_t^{1+\sigma_l}\right)}{\exp\left(\frac{\sigma_c-1}{1+\sigma_l} L_{t+1}^{1+\sigma_l}\right)}] \tag{64}$$

The Euler equation states the intertemporal optimality condition, relating current to future consumption by equating the marginal utility of consuming today to the expected discounted marginal utility of consuming tomorrow, adjusted for the return on bonds R_t and expected inflation $E_t\left[\pi_{t+1}^{-1}\right]$. The term relating labour today and labour tomorrow stems from the non-separability between consumption and labour in the utility function but cancels out in steady state with $L_t = L_{t+1} = L^*$.

A.2 The firm sector

Final goods producers

The final goods producer uses a continuum of differentiated intermediate goods $Y_t(i)$ in their production and packages them into a single final good Y_t to sell it in a perfectly competitive market at price P_t. The final goods firm makes zero profit in equilibrium — its role is purely to aggregate inputs. Intermediate goods are aggregated into a final good via Dixit-Stiglitz aggregation:

$$\int_0^1 \left(\frac{Y_t(i)}{Y_t}\right)^{\frac{\varepsilon_p}{1+\varepsilon_p}} di = 1 \tag{65}$$

Here, the ratio $\left(\frac{Y_t(i)}{Y_t}\right)$ captures the relative share of variety of firm i in total output. This ensures that if all intermediate inputs double, Y_t doubles too. The parameter ε_p governs the demand elasticity. Each intermediate firm faces a downward-sloping demand curve with constant elasticity $1+\varepsilon_p$: a one percent increase in the relative price $\left(\frac{P_t(i)}{P_t}\right)$ reduces demand for variety i by $1+\varepsilon_p$ percent.

The profit function is given by:

$$\max_{Y_t,\, Y_t(i)} \; P_t Y_t - \int_0^1 P_t(i) Y_t(i)\, di \tag{66}$$

Thus, the profit maximisation problem with Lagrangian multiplier $\lambda_{f,t}$ becomes:

$$\mathcal{L} = P_t Y_t - \int_0^1 P_t(i) Y_t(i)\, di + \lambda_t \left[\int_0^1 \left(\frac{Y_t(i)}{Y_t} \right)^{\frac{\varepsilon_p}{1+\varepsilon_p}} di - 1 \right] \tag{67}$$

Taking the FOC with respect to $Y_t(i)$, holding Y_t fixed:

$$\frac{\partial \mathcal{L}}{\partial Y_t(i)} = \lambda_{f,t} \cdot \frac{\varepsilon_p}{1+\varepsilon_p} \cdot \frac{1}{Y_t} \left(\frac{Y_t(i)}{Y_t} \right)^{-\frac{1}{1+\varepsilon_p}} = P_t(i) \tag{68}$$

Taking the FOC with respect to Y_t, noting that Y_t appears inside the aggregator through the ratio $Y_t(i)/Y_t$:

$$\frac{\partial \mathcal{L}}{\partial Y_t} = \lambda_{f,t} \frac{\varepsilon_p}{1+\varepsilon_p} \int_0^1 \left(\frac{Y_t(i)}{Y_t} \right)^{-\frac{1}{1+\varepsilon_p}} \frac{Y_t(i)}{Y_t^2}\, di = P_t \tag{69}$$

Dividing (68) by (69) eliminates $\lambda_{f,t}$:

$$\frac{P_t(i)}{P_t} = \frac{\left(\frac{Y_t(i)}{Y_t} \right)^{-\frac{1}{1+\varepsilon_p}}}{\int_0^1 \left(\frac{Y_t(i)}{Y_t} \right)^{-\frac{1}{1+\varepsilon_p}} \frac{Y_t(i)}{Y_t}\, di} \tag{70}$$

In a symmetric equilibrium (all intermediate firms identical) the denominator equals 1, so inverting and raising both sides to $-(1+\varepsilon_p)$ gives the CES demand curve:

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-(1+\varepsilon_p)} Y_t \tag{71}$$

The demand curve relates the demand for intermediate input $Y_t(i)$ to its relative price and to aggregate output Y_t. Specifically, the quantity of variety i demanded by the final goods producer is decreasing in its relative price $P_t(i)/P_t$: a higher price for variety i relative to the aggregate price level leads the final goods producer to substitute away from input i toward cheaper varieties. The strength of this substitution is governed by ε_p - a higher demand elasticity means varieties are more substitutable, so the final goods producer responds more strongly to any relative price increase. Finally, demand for each variety $Y_t(i)$ scales proportionally with aggregate output Y_t: whenever the economy produces more in aggregate, the final goods producer demands more of every intermediate input.

In order to derive the aggregate price index in the final good sector, one needs to substitute the demand curve (71) back into the aggregation constraint (65):

$$\int_0^1 \left(\frac{P_t(i)}{P_t} \right)^{-\varepsilon_p} di = 1 \tag{72}$$

Solving for P_t:

$$P_t = \left(\int_0^1 P_t(i)^{-\varepsilon_p}\, di \right)^{-\frac{1}{\varepsilon_p}} \tag{73}$$

This is the standard CES price index. In this specification is the aggregate price level a constant elasticity average of individual prices.

Intermediate goods producers

Unlike the final goods producers, intermediate goods producers have market power and thus act as price setters. Thereby, they do not end up having losses when unable to reset prices because of potential price stickiness or price indexation but they just lose parts of their mark-up. Intermediate firm i produces a differentiated good $Y_t(i)$ using labour as the only input. The production function is:

$$Y_t(i) = \epsilon_t^a L_t(i) - \Phi \tag{74}$$

where ϵ_t^a is the total factor productivity (TFP), $L_t(i)$ is labour hired by firm i, and $\Phi > 0$ is a fixed cost of production. The fixed cost ensures that profits are zero in steady state, pinning down the steady-state markup. Without it, profit would remain positive due to the monopoly markup and thereby complicate the model's steady-state calibration. TFP follows an exogenous AR(1) process:

$$\ln \epsilon_t^a = \rho_a \ln \epsilon_{t-1}^a + \sigma_a \eta_t^a, \qquad \eta_t^a \sim \mathcal{N}(0,1) \tag{75}$$

Each firm takes the nominal wage W_t as given and minimizes total labour cost $W_t L_t(i)$ subject to meeting demand $Y_t(i)$. From the production function (74), the labour requirement is:

$$L_t(i) = \frac{Y_t(i) + \Phi}{\epsilon_t^a} \tag{76}$$

Total nominal cost is therefore:

$$TC_t(i) = W_t \cdot \frac{Y_t(i) + \Phi}{\epsilon_t^a} \tag{77}$$

Nominal marginal cost is the derivative of total cost with respect to output, which is the same for all firms:

$$MC_t = \frac{\partial TC_t(i)}{\partial Y_t(i)} = \frac{W_t}{\epsilon_t^a} \tag{78}$$

Real marginal cost, obtained by deflating by the aggregate price level P_t, is:

$$mc_t = \frac{MC_t}{P_t} = \frac{W_t}{P_t \epsilon_t^a} = \frac{w_t}{\epsilon_t^a} \tag{79}$$

where $w_t = W_t / P_t$ is the real wage. Note that marginal cost is identical across all firms since they all face the same wage and TFP.

Each intermediate firm has monopoly power over its variety and sets its price subject to the demand curve (71) derived from the final goods sector. In the flexible price case, the firm simply sets price as a constant markup over marginal cost (this is derived in detailed in Appendix A.3):

$$P_t(i) = \frac{(1 + \varepsilon_p)}{\varepsilon_p} \cdot MC_t \tag{80}$$

However, we follow Smets and Wouters (2007) and introduce nominal price rigidity via the (Calvo, 1983) mechanism: in each period, a fraction $\xi_p \in (0,1)$ of firms cannot re-optimise their price and instead index it to a weighted average of past and steady-state inflation:

$$P_{t+s}(i) = \tilde{P}_t(i) \cdot X_{t,s}, \qquad X_{t,s} = \prod_{l=1}^{s} \pi_{t+l-1}^{\iota_p} \bar{\pi}^{1-\iota_p} \tag{81}$$

where $\tilde{P}_t(i)$ is the price chosen at time t, $\iota_p \in [0,1]$ is the degree of indexation to past inflation π_{t-1}, and $\bar{\pi}$ is steady-state inflation. When $\iota_p = 1$ prices are fully indexed to past inflation; when $\iota_p = 0$ non-optimizing firms keep their price fixed at $\tilde{P}_t(i)\bar{\pi}$.

The fraction $(1-\xi_p)$ of firms that can re-optimize in period t choose $\tilde{P}_t(i)$ to maximize the present discounted value of profits over all future periods in which they may be stuck with this price:

$$\max_{\tilde{P}_t(i)} E_t \sum_{s=0}^{\infty} (\beta\xi_p)^s \frac{\lambda_{t+s}}{\lambda_t} \frac{P_t}{P_{t+s}} \left[\tilde{P}_t(i) X_{t,s} - MC_{t+s}\right] Y_{t+s}(i) \tag{82}$$

subject to the demand curve in each period $t+s$:

$$Y_{t+s}(i) = \left(\frac{\tilde{P}_t(i) X_{t,s}}{P_{t+s}}\right)^{-(1+\varepsilon_p)} Y_{t+s} \tag{83}$$

The term $(\beta\xi_p)^s$ discounts future profits by both the time preference β and the probability ξ_p^s of still being stuck with price $\tilde{P}_t(i)$ at horizon s. The term $\frac{\lambda_{t+s}}{\lambda_t}\frac{P_t}{P_{t+s}}$ is the real stochastic discount factor, converting future nominal profits into current real values. Finally, $\left[\tilde{P}_t(i) X_{t,s} - MC_{t+s}\right]$ represents the per-unit profit margin of intermediate firms, so multiplying it by $Y_{t+s}(i)$ yields total profits at horizon s.

The demand constraint (83) is substituted directly into the objective (82) before differentiating, since for any price the firm sets it knows exactly how much it will have to produce. Define $\theta \equiv 1 + \varepsilon_p$ for convenience. Substituting and differentiating with respect to $\tilde{P}_t(i)$ gives the FOC (a detailed derivation is given in Appendix A.4):

$$E_t \sum_{s=0}^{\infty} (\beta\xi_p)^s \frac{\lambda_{t+s}}{\lambda_t} \frac{P_t}{P_{t+s}} Y_{t+s}(i) \left[\tilde{P}_t(i) X_{t,s} - \frac{(1+\varepsilon_p)}{\varepsilon_p} MC_{t+s}\right] = 0 \tag{84}$$

This condition says that the firm sets its reset price such that the present discounted value of the gap between its effective price $\tilde{P}_t(i) X_{t,s}$ and the desired markup $(1+\varepsilon_p)$ times marginal cost is zero across all future periods. In the static case ($s = 0$ only, no Calvo), this immediately delivers the flexible price result (80).

Finally, to get to the aggregated price of the intermediate firms, one has to differentiate their different price setting behaviors. In each period, a fraction $(1-\xi_p)$ of firms reset their price to $\tilde{P}_t$, while the remaining fraction ξ_p cannot re-optimise and instead index their price to past and steady-state inflation. Following (73) and splitting the price indexation into those two groups by inserting their relative prices under Dixit-Stiglitz aggregation into the equation, we end up with:

$$1 = (1-\xi_p)\left(\frac{\tilde{P}_t}{P_t}\right)^{1-\varepsilon_p} + \xi_p \left(\frac{\pi_{t-1}^{\iota_p}\bar{\pi}^{1-\iota_p} P_{t-1}}{P_t}\right)^{1-\varepsilon_p} = (1-\xi_p)\left(\frac{\tilde{P}_t}{P_t}\right)^{1-\varepsilon_p} + \xi_p \left(\frac{\pi_{t-1}^{\iota_p}\bar{\pi}^{1-\iota_p}}{\pi_t}\right)^{1-\varepsilon_p} \tag{85}$$

The first term aggregates over the fraction $(1-\xi_p)$ of firms that re-optimise, who all set the same reset price $\tilde{P}_t$. The second term aggregates over the fraction ξ_p of firms that cannot re-optimise and instead mechanically index their price by $\pi_{t-1}^{\iota_p}\bar{\pi}^{1-\iota_p}$, so their relative price this period is their relative price last period scaled by this indexation factor, which in relative terms equals $\pi_{t-1}^{\iota_p}\bar{\pi}^{1-\iota_p}/\pi_t$.

A.3 Derivation flexible price case for intermediate firm

In the flexible price case there is no Calvo friction, so $\xi_p = 0$ and every firm re-optimises its price every period. The firm therefore simply maximises current period profit:

$$\max_{\tilde{P}_t(i)} \left[\tilde{P}_t(i) - MC_t\right] Y_t(i) \tag{86}$$

subject to the demand curve:

$$Y_t(i) = \left(\frac{\tilde{P}_t(i)}{P_t}\right)^{-(1+\varepsilon_p)} Y_t \tag{87}$$

Substituting the demand curve into the objective and defining $\theta \equiv 1 + \varepsilon_p$, the profit expression becomes:

$$\Pi_t(i) = \left[\tilde{P}_t(i) - MC_t\right] \left(\frac{\tilde{P}_t(i)}{P_t}\right)^{-\theta} Y_t = P_t^{\theta} Y_t \cdot \tilde{P}_t(i)^{1-\theta} - MC_t P_t^{\theta} Y_t \cdot \tilde{P}_t(i)^{-\theta} \tag{88}$$

Differentiating with respect to $\tilde{P}_t(i)$ and setting equal to zero:

$$\frac{\partial \Pi_t(i)}{\partial \tilde{P}_t(i)} = (1-\theta) P_t^{\theta} Y_t \cdot \tilde{P}_t(i)^{-\theta} + \theta \cdot MC_t P_t^{\theta} Y_t \cdot \tilde{P}_t(i)^{-\theta-1} = 0 \tag{89}$$

Dividing out the strictly positive term $P_t^{\theta} Y_t \cdot \tilde{P}_t(i)^{-\theta-1} > 0$:

$$(1-\theta)\tilde{P}_t(i) + \theta \cdot MC_t = 0 \tag{90}$$

Solving for $\tilde{P}_t(i)$:

$$\tilde{P}_t(i) = \frac{\theta}{\theta - 1} MC_t \tag{91}$$

Substituting back $\theta = 1 + \varepsilon_p$, noting that $\theta - 1 = \varepsilon_p$:

$$\frac{\theta}{\theta - 1} = \frac{1+\varepsilon_p}{\varepsilon_p} \tag{92}$$

The flexible price result is therefore (80):

$$P_t(i) = \frac{1+\varepsilon_p}{\varepsilon_p} \cdot MC_t \tag{93}$$

This is the standard monopolistic competition result: the firm sets its price as a constant gross markup $\frac{1+\varepsilon_p}{\varepsilon_p}$ over marginal cost. A higher ε_p implies more substitutability between varieties, a lower markup, and a price closer to marginal cost. In the limit $\varepsilon_p \to \infty$, the markup converges to one and the perfectly competitive outcome $P_t(i) = MC_t$ is recovered.

A.4 Derivation FOC for intermediate maximization problem

The fraction $(1 - \xi_p)$ of firms that can re-optimise in period t choose $\tilde{P}_t(i)$ to maximise the present discounted value of profits over all future periods in which they may be stuck with this price:

$$\max_{\tilde{P}_t(i)} E_t \sum_{s=0}^{\infty} (\beta \xi_p)^s \frac{\lambda_{t+s}}{\lambda_t} \frac{P_t}{P_{t+s}} \left[\tilde{P}_t(i) X_{t,s} - MC_{t+s}\right] Y_{t+s}(i) \tag{94}$$

subject to the demand curve in each period $t+s$:

$$Y_{t+s}(i) = \left(\frac{\tilde{P}_t(i) X_{t,s}}{P_{t+s}} \right)^{-(1+\varepsilon_p)} Y_{t+s} \tag{95}$$

where $X_{t,s} = \prod_{l=1}^{s} \pi_{t+l-1}^{\iota_p} \bar{\pi}^{1-\iota_p}$ captures the indexation of non-optimising firms to past and steady-state inflation, $\frac{\lambda_{t+s}}{\lambda_t} \frac{P_t}{P_{t+s}}$ is the real stochastic discount factor, and $(\beta\xi_p)^s$ discounts future profits by both the time preference β and the probability ξ_p^s of remaining unable to re-optimise at horizon s.

Since the demand constraint (95) holds for every s, it is substituted directly into the objective (94) before differentiating. Defining $\theta \equiv 1 + \varepsilon_p$, the profit term at each horizon s becomes:

$$\Pi_{t+s}(i) = \tilde{P}_t(i) X_{t,s} \left(\frac{\tilde{P}_t(i) X_{t,s}}{P_{t+s}} \right)^{-\theta} Y_{t+s} - MC_{t+s} \left(\frac{\tilde{P}_t(i) X_{t,s}}{P_{t+s}} \right)^{-\theta} Y_{t+s} \tag{96}$$

Expanding and collecting terms in $\tilde{P}_t(i)$:

$$\Pi_{t+s}(i) = P_{t+s}^{\theta} X_{t,s}^{1-\theta} Y_{t+s} \cdot \tilde{P}_t(i)^{1-\theta} - MC_{t+s} P_{t+s}^{\theta} X_{t,s}^{-\theta} Y_{t+s} \cdot \tilde{P}_t(i)^{-\theta} \tag{97}$$

Differentiating with respect to $\tilde{P}_t(i)$:

$$\frac{\partial \Pi_{t+s}(i)}{\partial \tilde{P}_t(i)} = (1-\theta) P_{t+s}^{\theta} X_{t,s}^{1-\theta} Y_{t+s} \cdot \tilde{P}_t(i)^{-\theta} + \theta \cdot MC_{t+s} P_{t+s}^{\theta} X_{t,s}^{-\theta} Y_{t+s} \cdot \tilde{P}_t(i)^{-\theta-1} \tag{98}$$

Factoring out $P_{t+s}^{\theta} X_{t,s}^{-\theta} Y_{t+s} \cdot \tilde{P}_t(i)^{-\theta-1}$:

$$\frac{\partial \Pi_{t+s}(i)}{\partial \tilde{P}_t(i)} = P_{t+s}^{\theta} X_{t,s}^{-\theta} Y_{t+s} \cdot \tilde{P}_t(i)^{-\theta-1} \left[(1-\theta) \tilde{P}_t(i) X_{t,s} + \theta \cdot MC_{t+s} \right] \tag{99}$$

Noting that the demand curve (95) implies $Y_{t+s}(i) = \tilde{P}_t(i)^{-\theta} X_{t,s}^{-\theta} P_{t+s}^{\theta} Y_{t+s}$, the factored term can be written as $Y_{t+s}(i) \cdot \tilde{P}_t(i)^{-1}$. Setting the present discounted sum of derivatives equal to zero and dividing out the strictly positive terms $Y_{t+s}(i) \cdot \tilde{P}_t(i)^{-1} > 0$:

$$E_t \sum_{s=0}^{\infty} (\beta\xi_p)^s \frac{\lambda_{t+s}}{\lambda_t} \frac{P_t}{P_{t+s}} Y_{t+s}(i) \left[(1-\theta) \tilde{P}_t(i) X_{t,s} + \theta \cdot MC_{t+s} \right] = 0 \tag{100}$$

Substituting back $\theta = 1 + \varepsilon_p$, so that $1 - \theta = -\varepsilon_p$, and dividing through by $-\varepsilon_p > 0$ yields (84):

$$E_t \sum_{s=0}^{\infty} (\beta\xi_p)^s \frac{\lambda_{t+s}}{\lambda_t} \frac{P_t}{P_{t+s}} Y_{t+s}(i) \left[\tilde{P}_t(i) X_{t,s} - \frac{1+\varepsilon_p}{\varepsilon_p} MC_{t+s} \right] = 0 \tag{101}$$

This condition states that the optimising firm sets its reset price such that the present discounted value of the gap between its effective price $\tilde{P}_t(i) X_{t,s}$ and the desired price $\frac{1+\varepsilon_p}{\varepsilon_p} MC_{t+s}$ is zero across all future periods. The term $\frac{1+\varepsilon_p}{\varepsilon_p}$ is the desired gross markup of price over marginal cost: in the static flexible price case ($\xi_p = 0$, $s = 0$ only), (101) immediately delivers $\tilde{P}_t = \frac{1+\varepsilon_p}{\varepsilon_p} MC_t$, the standard monopolistic competition result.

A.5 Monetary policy rule

To close the model, we add a central bank that follows the following Taylor rate rule by changing the interest rate as a response to deviations of inflation and output from their target:

$$\frac{R_t}{R^*} = \left(\frac{R_{t-1}}{R^*}\right)^{\rho_{\mathrm{mp}}} \left[\left(\frac{\pi_t}{\pi^*}\right)^{\psi_\pi} \left(\frac{Y_t}{Y_t^*}\right)^{\psi_y}\right]^{1-\rho_{\mathrm{mp}}} \varepsilon_t^r \tag{102}$$

Here, R^*, π^* and Y^* represent the steady state nominal policy rate, inflation rate and output level respectively. The parameter ρ_{mp} represents the degree to which the central bank engages in interest rate smoothing over time. A higher ρ_{mp} means the central bank puts more weight on keeping the interest rate close to its previous value and adjusts more gradually in response to inflation and output gap developments. Finally, ε_t^r represents a monetary policy shock that follows an AR(1) process:

$$\ln \varepsilon_t^r = \rho_r \ln \epsilon_{t-1}^r + \sigma_r \eta_t^r, \qquad \eta_t^r \sim \mathcal{N}(0,1) \tag{103}$$

A.6 Equilibrium

A competitive equilibrium in this economy consists of sequences of quantities $\{Y_t, C_t, L_t\}$ and prices $\{W_t, \pi_t, R_t\}$ such that (i) households maximize lifetime utility subject to their budget constraint, (ii) final goods producers maximize profits subject to the aggregation technology, (iii) intermediate goods producers minimize costs and set prices optimally subject to the Calvo mechanism, (iv) the central bank follows the Taylor rule (102), and (v) all markets clear.

Since there is no capital accumulation or investment in the small-scale model, the goods market clearing condition requires that all output is either consumed or absorbed by government spending:

$$Y_t = C_t \tag{104}$$

The labour market clears when aggregate labour supply of households equals aggregate labour demand of intermediate firms. Since wages are flexible, the real wage W_t/P_t adjusts in every period to equate supply and demand:

$$L_t^s = L_t^d = L_t \tag{105}$$

where $L_t = \int_0^1 L_t(i)\, di$ is aggregate labour demand across all intermediate firms.

Since the model is closed and bonds are claims between domestic households only, bonds are in zero net supply in equilibrium:

$$B_t = 0 \tag{106}$$

Combining the goods market clearing condition (104) with the household budget constraint and imposing bond market clearing (106), the aggregate resource constraint of the economy is:

$$Y_t = C_t \tag{107}$$

Because we abstract from capital accumulation and investment, the aggregate resource constraint is naturally the same as the goods market clearing condition . The equilibrium is then fully determined by the optimality conditions of households and firms derived above, the Taylor rule (102) which determines the nominal interest rate, and the exogenous shock processes which drive fluctuations around the steady state.

A.7 Steady state

The steady state is defined as the equilibrium in which all variables take constant values so that $X_t = X^*$, and trending variables grow at the constant trend rate $\bar{\gamma}$, while shocks take the value of one, as they enter the model multiplicatively. Evaluating the equilibrium conditions derived above at these values delivers the steady-state relationships of the model.

The steady state is defined as the equilibrium in which all variables take constant values so that $X_t = X^*$ for any variable X_t, and growth rates for inflation and the short-term rate are zero so that $X_t/X_{t-1} = 1$, while output, consumption, and the real wage grow at the constant trend rate $\bar{\gamma}$. Since all shocks enter the non-linear model multiplicatively, their steady-state values are unity, i.e. $\varepsilon_t^a = \varepsilon_t^r = 1$ for all t. Note that this corresponds to setting the log-deviations $\hat{\varepsilon}_t^a = \hat{\varepsilon}_t^r = 0$ in the log-linearised model, since $\ln(1) = 0$. Evaluating the equilibrium conditions derived above at these values delivers the steady-state relationships of the model.

Evaluating the bond Euler equation (62) at steady state, where $C_t = C_{t+1} = C^*$, $\pi_t = \bar{\pi}$ and $\varepsilon_t^b = 1$:

$$1 = \beta \frac{R^*}{\bar{\pi}} \tag{108}$$

Solving for the steady-state nominal interest rate delivers the Fisher equation:

$$R^* = \frac{\bar{\pi}}{\beta} \tag{109}$$

This states that the steady-state nominal interest rate equals the ratio of steady-state inflation to the discount factor. The implied steady-state real interest rate is $r^* = R^*/\bar{\pi} = 1/\beta$.

In steady state all intermediate firms are identical and set the same price, so $P_t(i) = P_t = P^*$ for all i. The steady-state markup condition follows from the pricing FOC (84), evaluated at steady state where all future periods are identical and the indexation term $X_{t,s} = 1$:

$$P^* = \frac{1+\varepsilon_p}{\varepsilon_p} MC^* \tag{110}$$

In terms of real variables, dividing through by P^*:

$$1 = \frac{1+\varepsilon_p}{\varepsilon_p} mc^* \tag{111}$$

so that steady-state real marginal cost is:

$$mc^* = \frac{\varepsilon_p}{1+\varepsilon_p} \tag{112}$$

From the real marginal cost expression (79), evaluated at steady state where $\varepsilon_t^a = 1$:

$$mc^* = \frac{w^*}{\varepsilon^a} = w^* \tag{113}$$

Combining (112) and (113), the steady-state real wage is:

$$w^* = \frac{\varepsilon_p}{1 + \varepsilon_p} \tag{114}$$

From the production function (74), evaluated at steady state:

$$Y^* = \varepsilon^a L^* - \Phi = L^* - \Phi \tag{115}$$

where the fixed cost Φ is calibrated such that steady-state profits are zero, ensuring that $Y^* > 0$.

From goods market clearing (104):

$$Y^* = C^* \tag{116}$$

At steady state, labour market clearing combined with the household intratemporal optimality condition requires that the real wage equals the marginal rate of substitution between leisure and consumption:

$$w^* = C^*(L^*)^{\sigma_l} \tag{117}$$

Together with the production function (115) and goods market clearing (116), this pins down the steady-state values of L^* and C^* given the known steady-state real wage w^* from (114).

Equations (109)–(117) jointly determine the steady-state values $\{Y^*, C^*, L^*, w^*, R^*, mc^*\}$ given the structural parameters $\{\beta, \sigma_c, \sigma_l, \varepsilon_p, \bar{\pi}, \Phi\}$.

Finally, evaluating the Taylor rule (102) at steady state, where all gaps are zero and $\varepsilon_t^r = 1$:

$$\frac{R^*}{R^*} = \left(\frac{R^*}{R^*}\right)^{\rho_{\mathrm{mp}}} \left[\left(\frac{\bar{\pi}}{\bar{\pi}}\right)^{\psi_\pi} \left(\frac{Y^*}{Y^*}\right)^{\psi_y}\right]^{1-\rho_{\mathrm{mp}}} = 1 \tag{118}$$

which is trivially satisfied, confirming that the steady state is consistent with the Taylor rule for any values of ρ_{mp}, ψ_π and ψ_y.

A.8 Log-linearised equilibrium conditions

The model is solved by log-linearising all equilibrium conditions around the steady state derived above. For any variable X_t, the log-deviation from steady state is defined as $\hat{x}_t = \ln X_t - \ln X^*$, which for small deviations approximates the percentage deviation from steady state.

Derivation dynamic IS curve

Combining the log-linearised Euler equation (64) and the log-linearised production function (74) yields the dynamic IS curve.

Starting from the non-linear Euler equation (64) derived in the household section:

$$C_t^{-\sigma_c} \exp\left(\frac{\sigma_c - 1}{1+\sigma_l} L_t^{1+\sigma_l}\right) = \beta R_t E_t \left[\frac{C_{t+1}^{-\sigma_c} \exp\left(\frac{\sigma_c - 1}{1+\sigma_l} L_{t+1}^{1+\sigma_l}\right)}{\pi_{t+1}} \right] \tag{119}$$

Taking logs of both sides and using the approximation $\ln E_t[X_{t+1}] \approx E_t[\ln X_{t+1}]$ for small fluctuations around 1:

$$\begin{aligned} -\sigma_c \ln C_t + \frac{\sigma_c - 1}{1+\sigma_l} L_t^{1+\sigma_l} &= \ln \beta + \ln R_t - E_t[\ln \pi_{t+1}] \\ &\quad - \sigma_c E_t[\ln C_{t+1}] + \frac{\sigma_c - 1}{1+\sigma_l} E_t[L_{t+1}^{1+\sigma_l}] \end{aligned} \tag{120}$$

Subtracting the steady-state version and writing in hat notation $\hat{x}_t = \ln X_t - \ln X^*$:

$$\begin{aligned} -\sigma_c \hat{c}_t + (\sigma_c - 1)(L^*)^{\sigma_l+1} \hat{L}_t &= \hat{R}_t - E_t[\hat{\pi}_{t+1}] \\ &\quad - \sigma_c E_t[\hat{c}_{t+1}] + (\sigma_c - 1)(L^*)^{\sigma_l+1} E_t[\hat{L}_{t+1}] \end{aligned} \tag{121}$$

Here, the labour term $\frac{\sigma_c-1}{1+\sigma_l} L_t^{1+\sigma_l}$ requires special attention since it is a power function of L_t rather than a logarithm, and therefore cannot be directly expressed in hat notation. Instead, it is linearised via a first-order Taylor expansion around the steady state L^*:

$$\begin{aligned} \frac{\sigma_c - 1}{1+\sigma_l} L_t^{1+\sigma_l} &\approx \frac{\sigma_c - 1}{1+\sigma_l} (L^*)^{1+\sigma_l} + \frac{\sigma_c - 1}{1+\sigma_l} (1+\sigma_l)(L^*)^{\sigma_l}(L_t - L^*) \\ &= \frac{\sigma_c - 1}{1+\sigma_l} (L^*)^{1+\sigma_l} + (\sigma_c - 1)(L^*)^{\sigma_l}(L_t - L^*) \end{aligned} \tag{122}$$

The first term is the steady-state value, which cancels when subtracting the steady-state version of the equation. The second term can be expressed in hat notation by dividing and multiplying by L^*:

$$(\sigma_c - 1)(L^*)^{\sigma_l}(L_t - L^*) = (\sigma_c - 1)(L^*)^{\sigma_l} \cdot L^* \cdot \frac{L_t - L^*}{L^*} \approx (\sigma_c - 1)(L^*)^{\sigma_l} \cdot L^* \cdot \hat{L}_t \tag{123}$$

where the approximation $\frac{L_t - L^*}{L^*} \approx \ln L_t - \ln L^* = \hat{L}_t$ holds for small deviations from steady state. The coefficient on $\hat{L}_t$ is therefore $(\sigma_c - 1)(L^*)^{1+\sigma_l}$.

Following Del Negro and Schorfheide (2013), steady-state labour supply is normalised to $L^* = 1$, consistent with the steady-state calibration, so that $(L^*)^{1+\sigma_l} = 1$. Including this in (121) and rearranging yields:

$$\sigma_c(\hat{c}_t - E_t[\hat{c}_{t+1}]) = (\sigma_c - 1)(\hat{L}_t - E_t[\hat{L}_{t+1}]) + E_t[\hat{\pi}_{t+1}] - \hat{R}_t \tag{124}$$

The production function log-linearises to $\hat{y}_t = \hat{\varepsilon}_t^a + \hat{L}_t$, so that $\hat{L}_t = \hat{y}_t - \hat{\varepsilon}_t^a$. The labour growth term therefore becomes:

$$E_t[\hat{L}_{t+1}] - \hat{L}_t = E_t[\hat{y}_{t+1}] - \hat{y}_t - (E_t[\hat{\varepsilon}_{t+1}^a] - \hat{\varepsilon}_t^a) = E_t[\hat{y}_{t+1}] - \hat{y}_t - E_t[\tau_{t+1}] \tag{125}$$

where $\tau_t \equiv \hat{\varepsilon}_t^a - \hat{\varepsilon}_{t-1}^a$ is the growth rate of TFP. Substituting into (124) yields:

$$\sigma_c(\hat{c}_t - E_t[\hat{c}_{t+1}]) = (\sigma_c - 1)(E_t[\hat{y}_t - \hat{y}_{t+1}] - E_t[\hat{\tau}_{t+1}]) + E_t[\hat{\pi}_{t+1}] - \hat{R}_t \tag{126}$$

Dividing through by σ_c, the dynamic IS curve is then (127):

$$\hat{c}_t = E_t[\hat{c}_{t+1}] - \frac{(\sigma_c - 1)}{\sigma_c} E_t[\hat{\tau}_{t+1}] - \frac{1}{\sigma_c}(\hat{r}_t - E_t[\hat{\pi}_{t+1}]) + \frac{\sigma_c - 1}{\sigma_c}(\hat{y}_t - E_t[\hat{y}_{t+1}]) \tag{127}$$

Equation (127) states that current consumption depends on four terms. First, expected future consumption $E_t[c_{t+1}]$ reflects the forward-looking nature of the household's savings decision. Second, expected TFP growth $E_t[\tau_{t+1}] \equiv \hat{\varepsilon}^a_{t+1} - \hat{\varepsilon}^a_t$ enters because households anticipate future productivity gains and adjust consumption today accordingly - a higher expected growth rate raises expected future income, encouraging more consumption today. Third, the ex-ante real interest rate $r_t - E_t[\pi_{t+1}]$ governs the intertemporal substitution of consumption: a higher real interest rate makes saving more attractive, reducing current consumption. The parameter $1/\sigma_c$ is the intertemporal elasticity of substitution, governing the sensitivity of consumption to changes in the real interest rate. Fourth, the term $\frac{\sigma_c - 1}{\sigma_c}(y_t - E_t[y_{t+1}])$ reflects the non-separability between consumption and labour in the utility function: when output is expected to grow, future labour supply increases, which through the non-separable utility affects the marginal utility of consumption today. When $\sigma_c = 1$ this term vanishes entirely, recovering the standard separable case.

Derivation real marginal cost

To derive the log-linearized real marginal costs, we take the household labour supply (63):

$$\hat{w}_t = \hat{c}_t + \sigma_l \hat{L}_t \tag{128}$$

We use the fact that $\hat{L}_t = \hat{y}_t - \hat{\varepsilon}^a_t$ from the log-linearized production function (74) and $\widehat{mc}_t = \hat{w}_t - \hat{\varepsilon}^a_t$ from log-linearized cost minimisation (79) to substitute for $\hat{w}_t$ and $\hat{L}_t$ to get:

$$m\hat{c}_t = \hat{c}_t + \sigma_l \hat{y}_t - (\sigma_l + 1)\hat{\varepsilon}^a_t \tag{129}$$

Real marginal cost mc_t increases with both consumption c_t and output y_t. The consumption term reflects that higher consumption raises the marginal utility of wealth and therefore the opportunity cost of working, pushing up wages and hence marginal cost. The output term $\sigma_l y_t$ reflects that higher output requires more labour, which raises the marginal disutility of labour and therefore wages. The TFP shock ε^a_t is separated out because it affects marginal cost directly through productivity - higher TFP means each worker produces more, reducing the labour required and hence marginal cost.

Derivation of the log-linearised NK Phillips Curve Deriving the NK Phillips Curve requires combining the log-linearised Calvo pricing FOC derived in the intermediate goods section (84) with the log-linearised aggregate price index with Calvo indexation (85) and can be separated into four steps.

Step 1: From the pricing FOC to the optimal reset price

Starting from the Calvo pricing FOC (84):

$$E_t \sum_{s=0}^{\infty} (\beta \xi_p)^s \frac{\lambda_{t+s}}{\lambda_t} \frac{P_t}{P_{t+s}} Y_{t+s}(i) \left[\tilde{P}_t(i) X_{t,s} - \frac{1 + \varepsilon_p}{\varepsilon_p} MC_{t+s} \right] = 0 \tag{130}$$

The real stochastic discount factor $\frac{\lambda_{t+s}}{\lambda_t} \frac{P_t}{P_{t+s}}$ equals 1 in steady state around the zero-inflation steady state, and $Y_{t+s}(i)/Y_t \approx 1$ in the symmetric equilibrium. Dividing through by Y_t, taking logs of the bracket, and denoting the desired gross markup as $\mu \equiv \frac{1+\varepsilon_p}{\varepsilon_p}$, the FOC in log form becomes:

$$E_t \sum_{s=0}^{\infty} (\beta\xi_p)^s \left[\ln \tilde{P}_t(i) + \ln X_{t,s} - \ln \mu - \ln MC_{t+s}\right] = 0 \tag{131}$$

Recalling that $X_{t,s} = \prod_{l=1}^{s} \pi_{t+l-1}^{\iota_p} \bar{\pi}^{1-\iota_p}$, around the zero-inflation steady state where $\bar{\pi} = 1$:

$$\ln X_{t,s} \approx \iota_p \sum_{l=1}^{s} \hat{\pi}_{t+l-1} \tag{132}$$

Expressing all terms in hat notation using $\ln \tilde{P}_t(i) = \hat{\tilde{p}}_t + \ln P_t$, $\ln MC_{t+s} = \widehat{mc}_{t+s} + \ln mc^* + \ln P_{t+s}$, and $\ln P_{t+s} - \ln P_t = \hat{p}_{t+s} - \hat{p}_t$:

$$\ln \tilde{P}_t(i) + \ln X_{t,s} - \ln MC_{t+s} \approx \hat{\tilde{p}}_t - \widehat{mc}_{t+s} - (\hat{p}_{t+s} - \hat{p}_t) \tag{133}$$

where the steady-state terms $\ln mc^*$ and $\ln \mu$ cancel when subtracting the steady state, and the indexation term $\iota_p \sum_{l=1}^{s} \hat{\pi}_{t+l-1} \approx 0$ in a first-order approximation around the zero-inflation steady state. Substituting back into the FOC:

$$E_t \sum_{s=0}^{\infty} (\beta\xi_p)^s \left[\hat{\tilde{p}}_t - \widehat{mc}_{t+s} - (\hat{p}_{t+s} - \hat{p}_t)\right] = 0 \tag{134}$$

Since $\hat{\tilde{p}}_t$ does not depend on s it factors out of the sum:

$$\hat{\tilde{p}}_t \sum_{s=0}^{\infty} (\beta\xi_p)^s = E_t \sum_{s=0}^{\infty} (\beta\xi_p)^s \left[\widehat{mc}_{t+s} + \hat{p}_{t+s} - \hat{p}_t\right] \tag{135}$$

The left-hand side is a geometric series with sum $\frac{1}{1-\beta\xi_p}$. The right-hand side cannot be simplified further since $\widehat{mc}_{t+s}$ and $\hat{p}_{t+s}$ differ at every horizon s. Multiplying both sides by $(1 - \beta\xi_p)$ delivers the optimal reset price equation:

$$\hat{\tilde{p}}_t = (1 - \beta\xi_p) \sum_{s=0}^{\infty} (\beta\xi_p)^s E_t \left[\widehat{mc}_{t+s} + \hat{p}_{t+s} - \hat{p}_t\right] \tag{136}$$

The reset price is a geometrically weighted average of all current and expected future real marginal costs, with weights $(1 - \beta\xi_p)(\beta\xi_p)^s$ that sum to one. A higher ξ_p shifts weight toward more distant horizons since the firm is more likely to be stuck with its reset price for longer.

Step 2: Log-linearise the Calvo price index

To derive the final NK Phillips Curve, one must furthermore log-linearise the Calvo price index of the intermediate firms. Log-linearising (85) around the zero-inflation steady state where $\bar{\pi} = 1$ and $\tilde{P}^*/P^* = 1$, using the first-order approximation $(1+x)^{1-\varepsilon_p} \approx 1 + (1-\varepsilon_p)x$ for small x:

$$0 = (1 - \xi_p)(1 - \varepsilon_p)\hat{\tilde{p}}_t + \xi_p(1 - \varepsilon_p)(\iota_p\hat{\pi}_{t-1} - \hat{\pi}_t) \tag{137}$$

where $\hat{\tilde{p}}_t = \ln \tilde{P}_t - \ln P_t$ is the log reset price relative to the current aggregate price level. Dividing through by $(1 - \varepsilon_p)$:

$$0 = (1 - \xi_p)\hat{\tilde{p}}_t + \xi_p\iota_p\hat{\pi}_{t-1} - \xi_p\hat{\pi}_t \tag{138}$$

Solving for $\hat{\tilde{p}}_t$:

$$\hat{\tilde{p}}_t = \frac{\xi_p \hat{\pi}_t - \xi_p \iota_p \hat{\pi}_{t-1}}{1-\xi_p} \tag{139}$$

$$\hat{\tilde{p}}_t = \frac{\xi_p}{1-\xi_p}(\hat{\pi}_t - \iota_p \hat{\pi}_{t-1}) \tag{140}$$

This states that the reset price relative to the current aggregate price level is a scaled version of current inflation adjusted for indexation. If current inflation is high, optimising firms must set a relatively high reset price to compensate for the fact that non-optimising firms are already indexing upward. Crucially, because the price index (85) is expressed in relative prices $\tilde{P}_t/P_t$ from the outset, no lagged price level $\hat{p}_{t-1}$ appears and the expression for $\hat{\tilde{p}}_t$ follows cleanly without any additional normalisation.

Step 3: Collapse the infinite sum recursively

Equation (136) contains an infinite sum over all future marginal costs, which is not directly tractable. To derive a recursive form, separate the $s=0$ term from the rest of the sum:

$$\hat{\tilde{p}}_t = (1-\beta\xi_p)\widehat{mc}_t + (1-\beta\xi_p)\sum_{s=1}^{\infty}(\beta\xi_p)^s E_t\left[\widehat{mc}_{t+s} + \hat{p}_{t+s} - \hat{p}_t\right] \tag{141}$$

Substituting $k = s-1$ so that $s = k+1$, re-indexing the sum:

$$\hat{\tilde{p}}_t = (1-\beta\xi_p)\widehat{mc}_t + \beta\xi_p(1-\beta\xi_p)\sum_{k=0}^{\infty}(\beta\xi_p)^k E_t\left[\widehat{mc}_{t+1+k} + \hat{p}_{t+1+k} - \hat{p}_t\right] \tag{142}$$

Decomposing the price level difference as $\hat{p}_{t+1+k} - \hat{p}_t = (\hat{p}_{t+1+k} - \hat{p}_{t+1}) + (\hat{p}_{t+1} - \hat{p}_t)$, equation (142) becomes:

$$\begin{aligned}\hat{\tilde{p}}_t = (1-\beta\xi_p)\widehat{mc}_t &+ \beta\xi_p(1-\beta\xi_p)\sum_{k=0}^{\infty}(\beta\xi_p)^k E_t\left[\widehat{mc}_{t+1+k} + \hat{p}_{t+1+k} - \hat{p}_{t+1}\right] \\ &+ \beta\xi_p(1-\beta\xi_p)\sum_{k=0}^{\infty}(\beta\xi_p)^k E_t\left[\hat{p}_{t+1} - \hat{p}_t\right]\end{aligned} \tag{143}$$

The last term involves $E_t[\hat{p}_{t+1} - \hat{p}_t] = E_t[\hat{\pi}_{t+1}]$, which is zero in a first-order approximation around the zero-inflation steady state and is therefore dropped. The remaining infinite sum is precisely (136) written at time $t+1$:

$$\hat{\tilde{p}}_{t+1} = (1-\beta\xi_p)\sum_{k=0}^{\infty}(\beta\xi_p)^k E_{t+1}\left[\widehat{mc}_{t+1+k} + \hat{p}_{t+1+k} - \hat{p}_{t+1}\right] \tag{144}$$

Taking expectations at time t and substituting into (142) delivers the recursive expression:

$$\hat{\tilde{p}}_t = (1-\beta\xi_p)\widehat{mc}_t + \beta\xi_p E_t[\hat{\tilde{p}}_{t+1}] \tag{145}$$

The infinite forward-looking sum has collapsed into a simple recursive expression: the reset price today equals a weighted average of current marginal cost and the discounted expected reset price tomorrow. The key insight is that the infinite tail of the sum at t exactly equals $\beta\xi_p$ times the entire sum at $t+1$ — shifting the summation index by one and recognising the resulting expression as (136) written at $t+1$ is what makes the recursion work, rather than a term-by-term cancellation.

Step 4: Substitute the price index and derive the NKPC

Substituting (140) at t and $t+1$ into (145):

$$\frac{\xi_p}{1-\xi_p}(\hat{\pi}_t - \iota_p\hat{\pi}_{t-1}) = (1-\beta\xi_p)\widehat{mc}_t + \beta\xi_p\frac{\xi_p}{1-\xi_p}(E_t[\hat{\pi}_{t+1}] - \iota_p\hat{\pi}_t) \tag{146}$$

Multiplying through by $\frac{1-\xi_p}{\xi_p}$:

$$\hat{\pi}_t - \iota_p\hat{\pi}_{t-1} = \frac{(1-\xi_p)(1-\beta\xi_p)}{\xi_p}\widehat{mc}_t + \beta(E_t[\hat{\pi}_{t+1}] - \iota_p\hat{\pi}_t) \tag{147}$$

Expanding the right-hand side and collecting all inflation terms on the left:

$$\hat{\pi}_t - \iota_p\hat{\pi}_{t-1} + \beta\iota_p\hat{\pi}_t = \frac{(1-\xi_p)(1-\beta\xi_p)}{\xi_p}\widehat{mc}_t + \beta E_t[\hat{\pi}_{t+1}] \tag{148}$$

$$\hat{\pi}_t(1+\beta\iota_p) = \iota_p\hat{\pi}_{t-1} + \beta E_t[\hat{\pi}_{t+1}] + \frac{(1-\xi_p)(1-\beta\xi_p)}{\xi_p}\widehat{mc}_t \tag{149}$$

Dividing through by $(1+\iota_p\beta)$, delivers the New Keynesian Phillips Curve (150):

$$\hat{\pi}_t = \frac{\iota_p}{1+\iota_p\beta\gamma}\hat{\pi}_{t-1} + \frac{\beta\gamma}{1+\iota_p\beta\gamma}E_t[\hat{\pi}_{t+1}] + \frac{(1-\xi_p)(1-\beta\gamma\xi_p)}{\xi_p(1+\iota_p\beta\gamma)}\widehat{mc}_t \tag{150}$$

The Phillips curve relates current inflation to three terms. The backward-looking term $\frac{\iota_p}{1+\iota_p\beta}\pi_{t-1}$ reflects price indexation: firms that cannot re-optimise mechanically adjust their prices to past inflation, generating intrinsic inflation persistence. The forward-looking term $\frac{\beta}{1+\iota_p\beta}E_t[\pi_{t+1}]$ reflects that optimising firms are forward-looking and set prices based on expected future costs. The slope $\frac{(1-\xi_p)(1-\beta\xi_p)}{\xi_p(1+\iota_p\beta)}$ determines how strongly inflation responds to real marginal cost mc_t: a higher degree of price stickiness ξ_p means fewer firms re-optimise each period, flattening the Phillips curve. The appearance of γ in the discount factor reflects that the model accounts for a deterministic balanced growth path at rate γ, so future profits must be discounted not only by β but also by the trend growth rate, yielding an effective discount factor of $\beta\gamma$.

Derivation log-linearised taylor rule

To log-linearise the non-linear Taylor rule (102) around the steady state, start with the original Taylor rule (102):

$$\frac{R_t}{R^*} = \left(\frac{R_{t-1}}{R^*}\right)^{\rho_{\text{mp}}}\left[\left(\frac{\pi_t}{\pi^*}\right)^{\psi_\pi}\left(\frac{Y_t}{Y_t^*}\right)^{\psi_y}\right]^{1-\rho_{\text{mp}}}\varepsilon_t^r \tag{151}$$

Taking logs of both sides:

$$\ln\frac{R_t}{R^*} = \rho_{\text{mp}}\ln\frac{R_{t-1}}{R^*} + (1-\rho_{\text{mp}})\left[\psi_\pi\ln\frac{\pi_t}{\pi^*} + \psi_y\ln\frac{Y_t}{Y_t^*}\right] + \ln\varepsilon_t^r \tag{152}$$

Defining log-deviations from steady state as:

$$\hat{r}_t = \ln\left(\frac{R_t}{R^*}\right), \qquad \hat{\pi}_t = \ln\left(\frac{\pi_t}{\pi^*}\right), \qquad dy_t = \ln\left(\frac{Y_t}{Y_t^*}\right) \tag{153}$$

Substituting these definitions delivers the log-linearised Taylor rule (154):

$$\hat{r}_t = \rho_{\text{mp}}\hat{r}_{t-1} + (1-\rho_{\text{mp}})\psi_\pi\hat{\pi}_t + (1-\rho_{\text{mp}})\psi_y\, dy_t + \hat{\varepsilon}^r_t \tag{154}$$

Because the original equation is multiplicative/exponential in structure, taking logs is exact - no first-order Taylor approximation is required. The shock $\hat{\varepsilon}^r_t$ enters multiplicatively in levels and thus additively in logs, where we apply the approximation $\ln \varepsilon^r_t \approx \hat{\varepsilon}^r_t$ around the steady state where $\varepsilon^r_t = 1$.

Note that the central bank responds to output growth dy_t rather than the output gap $\tilde{y}_t$. This reflects the Del Negro and Schorfheide (2013) specification, in which the central bank targets changes in output rather than deviations from potential.

The central bank smooths the interest rate with weight ρ_{mp} on the lagged rate, and responds to current inflation with weight $(1-\rho_{\text{mp}})\psi_\pi$ and to output growth with weight $(1-\rho_{\text{mp}})\psi_y$. A higher smoothing parameter ρ_{mp} implies more inertia: the central bank adjusts gradually rather than responding aggressively to current conditions, as both response coefficients are scaled down by $(1-\rho_{\text{mp}})$. The Taylor principle requires $\psi_\pi > 1$ to ensure a unique rational expectations equilibrium: if the central bank raises nominal rates more than one-for-one with inflation, it raises real rates and stabilises inflation.

Derivation aggregate resource constraint

In the log-linearised model, we introduce an exogenous government spending shock ε^g_t that shifts aggregate demand. The aggregate resource constraint is:

$$\hat{y}_t = \hat{c}_t + \varepsilon^g_t \tag{155}$$

where $\hat{\varepsilon}^g_t$ is an exogenous demand shifter that follows an AR(1) process:

$$\varepsilon^g_t = \rho_g\varepsilon^g_{t-1} + \sigma_g\eta^g_t \tag{156}$$

with $\eta^g_t \overset{iid}{\sim} \mathcal{N}(0,1)$, persistence parameter $\rho_g \in (0,1)$, and shock standard deviation $\sigma_g > 0$. Note that ε^g_t is introduced directly at the log-linearised stage as a reduced-form demand disturbance rather than being derived from a non-linear government spending process, following Del Negro and Schorfheide (2013). Since government spending is entirely exogenous, any increase in ε^g_t must be matched by an increase in output or a decrease in consumption. There is no investment or net exports to absorb the shock, so government spending crowds out consumption one-for-one in the absence of a monetary policy response.

Derivation output growth

In line with Del Negro and Schorfheide (2013), we define the observed output growth as:

$$dy_t = \hat{y}_t - \hat{y}_{t-1} + \hat{\tau}_t \tag{157}$$

where TFP growth is defined as:

$$\hat{\tau}_t = \hat{\varepsilon}^a_t - \hat{\varepsilon}^a_{t-1} \tag{158}$$

The complete log-linearised system

To sum up, the model is driven by three structural disturbances defined as standard AR(1) processes:

$$\begin{aligned}
\varepsilon_t^g &= \rho_g \varepsilon_{t-1}^g + \sigma_g \eta_t^g && (159)\\
\varepsilon_t^r &= \rho_r \varepsilon_{t-1}^r + \sigma_r \eta_t^r && (160)\\
\varepsilon_t^a &= \rho_a \varepsilon_{t-1}^a + \sigma_a \eta_t^a && (161)\\
& && (162)
\end{aligned}$$

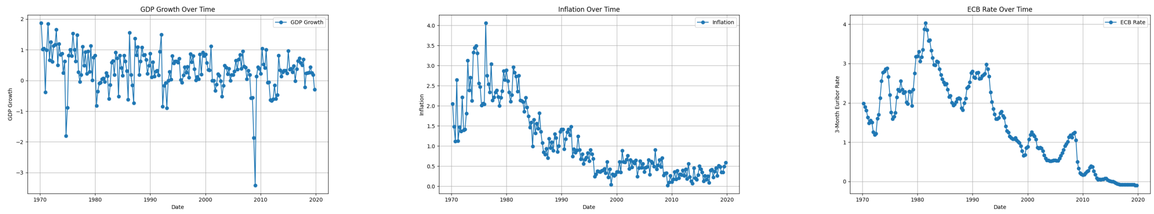
where $\eta_t^g, \eta_t^a, \eta_t^r \overset{iid}{\sim} \mathcal{N}(0,1)$ are mutually uncorrelated structural innovations and $\sigma_g, \sigma_a, \sigma_r > 0$ are the standard deviations of the respective shocks.

The persistence parameters ρ_g, ρ_r amd ρ_a govern how long the respective shocks last. All innovations are assumed to be mutually uncorrelated at all leads and lags, ensuring that the three shocks are structurally identified.

Equations (127), (129), (150), (154), (155), (157) and (158) together with the shock processes (159)–(161) form a closed system in the endogenous variables $\{\hat{c}_t, \hat{\tau}_t, \hat{r}_t, \hat{\pi}_t, \hat{m}c_t, \hat{y}_t, dy_t, \hat{\varepsilon}_t^a, \hat{\varepsilon}_t^r, \varepsilon_t^g\}$. This system constitutes the small-scale New Keynesian model of Del Negro and Schorfheide (2013).

B Additional material

B.1 Visualizing the three time series



(a) Output growth (dy) (b) Inflation (π) (c) 3-month Euribor rate (r)

Figure 6: Time series graphs for the three variables used in the model.

Source: Authors' rendering of AWMD data (İpek and Kısacıkoğlu, 2025).

B.2 Posterior distributions

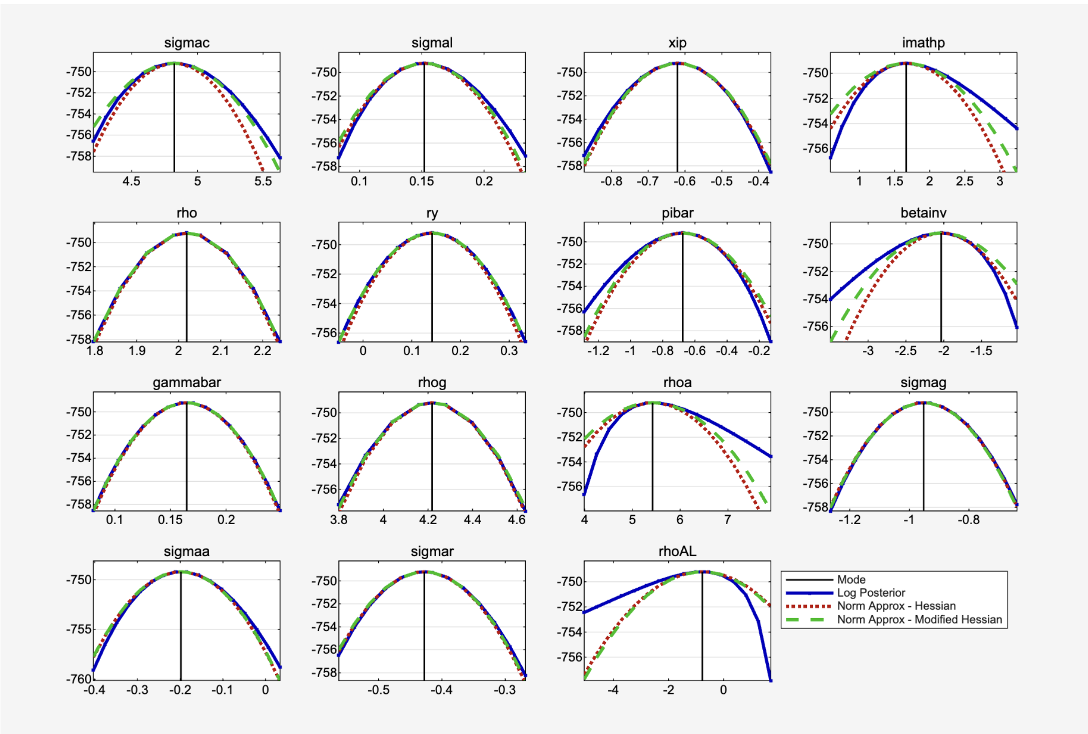


Figure 7: Posterior Distribution log-likelihood

Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025).

B.3 Belief coefficients

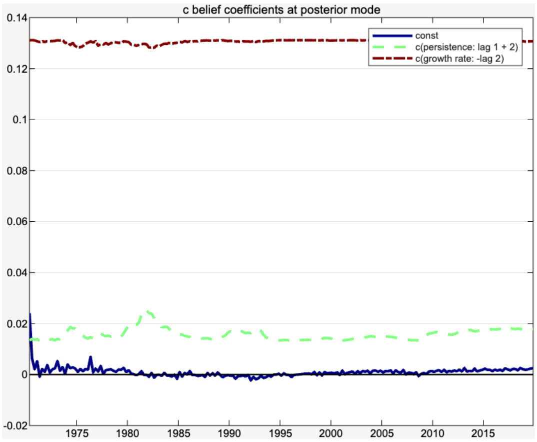
Our Figure 8 shows the beliefs about the constants, persistence (lag1+lag2), and growth rates (-lag2) evolve for the AR(2) models for the three forward-looking variables. To see where these interpretations come from, it is useful to rewrite AR(2) forecasting model as shown in 163.

The RE-consistent baseline beliefs described above are precisely the values around which the belief coefficients plotted in Figure 8 are anchored. The figure shows how the constant, persistence component (lag 1 + lag 2), and growth rate component (−lag 2) of the AR(2) forecasting models evolve over time for each of the three forward-looking variables. To interpret these components, it is useful to rewrite the AR(2) forecasting model in the alternative form shown in equation (163).[21]

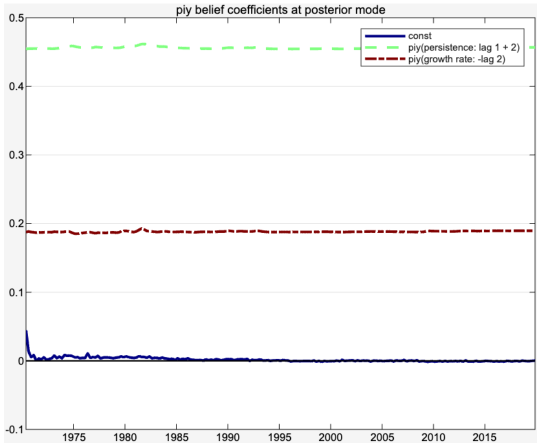
$$\mathbf{y}^f_{j,t} = \beta^1_{j,t} + (\beta^2_{j,t} + \beta^3_{j,t}) \times \mathbf{y}^f_{j,t-1} - \beta^3_{j,t} \times \Delta\mathbf{y}^f_{j,t-1} \tag{163}$$

A look at the belief coefficient evolution in Figure 8 yields two main takeaways. First, the coefficients remain tightly anchored around the RE-consistent baseline throughout the sample, consistent with the small estimated ρ_{AL}. Second, there are nonetheless small fluctuations, most notably in the consumption belief coefficients during the 1970s and 1980s. The constant term for consumption deviates from zero during this period before converging back, suggesting that agents' perceived mean of consumption was temporarily displaced from its steady-state value during the turbulent macroeconomic environment of the early sample, before stabilising as conditions normalised.

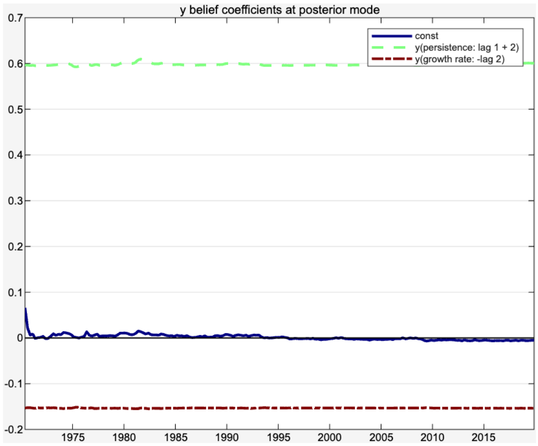
[21] This alternative form of the AR(2) forecasting model is obtained by adding and subtracting $\beta^3_{j,t} \times \mathbf{y}^f_{j,t-1}$ and then using $\Delta\mathbf{y}^f_{j,t-1} = \mathbf{y}^f_{j,t-1} - \mathbf{y}^f_{j,t-2}$ and then collecting terms.



(a) Consumption (c)



(b) Inflation (π)



(c) Output (y)

Figure 8: Belief coefficients at DSGE posterior mode over time. Each plot shows the constant (solid blue), persistence component (lag 1+2, dashed green), and growth rate component (lag −2, dash-dotted dark red).

Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025).

B.4 Local projections

To evaluate the empirical plausibility of our RE and AL model dynamics, we require an empirical benchmark against which the DSGE impulse response functions can be assessed. We obtain this benchmark by estimating impulse response functions directly from the data using local projections (LP), following Jordà (2005). An alternative would be to estimate IRFs via a vector autoregression (VAR). However, a VAR imposes a specific parametric structure on the entire data-generating process, and if that structure is even slightly misspecified, the resulting IRFs inherit those errors at every horizon. LPs avoid this pitfall by estimating each horizon h via a separate direct regression, so that any misspecification at one horizon does not contaminate the estimates at others. This robustness to misspecification makes LPs a natural choice for constructing an empirical benchmark in our setting.

The idea of local projections dates back to (Jordà, 2005) and is very simple: Instead of iterating our model forward to estimate IRFs (as in VAR), why not directly regress Y at each future horizon on the shock today. Thus one runs a simple OLS regression with H separate regression specification, one per horizon. The regression equation in horizon h can be written as:

$$y_{t+h} = \alpha_h + \beta_h \epsilon_t^r + X_t + \xi_{t+h} \tag{164}$$

Estimating this equation separately for each h and collecting the $\hat{\beta}_h$ coefficients across horizons delivers the impulse response function. In equation (164), y_{t+h} is the outcome variable, output growth, inflation, or the short-term rate, observed h periods after the shock, where $h = 0, 1, 2, \ldots, H$ indexes the horizon. α_h is a horizon-specific intercept. ϵ_t^r is the identified monetary policy shock at time t and β_h is the coefficient of interest, capturing the response of y at each horizon h. X_t is a vector of controls measured at or before time t, consisting of lags of the dependent variable and the shock, which absorb pre-existing dynamics without contaminating the estimated response. In our case, X_t consists of two lags of the dependent variable and two lags of the shock ϵ_t, which absorb pre-existing dynamics in the data without contaminating the estimated response to the shock. Finally, ξ_{t+h} is the error term. Since the same shock ϵ_t appears in regressions for multiple horizons h, the error terms ξ_{t+h} are serially correlated by construction. A shock at time t affects y_{t+1}, y_{t+2}, and so on, meaning the residuals across horizons are not independent. We therefore compute Newey-West standard errors, which remain consistent in the presence of this serial correlation, to construct valid confidence bands around the impulse response coefficients $\hat{\beta}_h$.

Using the monetary policy shock ε_t^r extracted directly from the estimated Taylor rule would likely confound the local projection estimates. The reason is twofold. First, the central bank may be endogenously reacting to economic developments not captured by the model's information set, so the residual conflates a genuine policy surprise with the central bank's systematic response to omitted variables. Second, even setting aside omitted variables, monetary policy announcements affect the economy through multiple channels simultaneously - not only through the interest rate itself but also through the information they convey about the central bank's economic outlook, introducing further bias.

The high-frequency identification approach offers a compelling solution to both problems. The key insight, introduced by Gertler and Karadi (2015), is that asset prices observed in a narrow window around a policy announcement have already incorporated all publicly available information about the state of the economy and the expected policy response. Any residual movement in that window therefore reflects only the unexpected component of the announcement which should compose a genuine exogenous shock. Specifically, Gertler and Karadi (2015) use changes in federal funds futures contracts measured within a 30-minute window around Federal Open Market Committee (FOMC) announcements as an external instrument to identify monetary policy shocks in a structural VAR, sidestepping the endogeneity and omitted variable problems inherent in model-based residuals.

However, Jarociński and Karadi (2020) show that even within this narrow window, a single high-frequency surprise conflates two distinct structural shocks. When the central bank raises rates and stocks simultaneously fall, markets interpret the move as a contractionary policy tightening. But when rates rise and stocks simultaneously rise, markets instead infer that the central bank has revealed positive private information about the economic outlook: a central bank information shock. Ignoring this distinction biases inference on monetary non-neutrality. Jarociński and Karadi (2020) disentangle the two components by imposing (poor man's) sign restrictions on the joint high-frequency co-movement of 3-month fed funds futures and the $S\&P$ 500 index, both measured within a 30-minute window around FOMC announcements, and identify the pure monetary policy shock as the component associated with a negative interest rate. We use the resulting series of identified monetary policy shocks provided by Jarociński and Karadi (2020) as the exogenous instrument ε_t in our local projection regressions.[22]

B.5 Analysis of AR(2) patterns

This appendix section shows our exploratory analysis of the autoregressive patterns in our time series, as well as the time series used by Slobodyan and Wouters (2012). This is done in order to investigate whether the apparent need for adaptive learning in the case of Slobodyan and Wouters (2012) (ρ_{AL} close to one) compared to the a smaller need in our model (smaller ρ_{AL}), could be driven by the patterns in the four additional time series that the agents in their model has to forecast.

Our analysis is done outside the context of any DSGE model and simply consists of fitting AR(2) models to all the time series. Then R^2 was calculated for the AR(2) for each variable to gauge how suitable a constant AR(2) is. R^2 has its drawback when autocorrelation is high and is therefore not commonly used when evaluating autoregressive models, instead measures as AIC is preferred. However, we wanted to facilitate a comparison of the fit between variables and not between different models for the same variable, which means that scale independence is crucial. Therefore we stick with R^2.

Table 6: R^2 for AR(2) models of the Euro area 1970-2019 time series used in our model and US 1966-2009 time series used in Slobodyan and Wouters (2012).

	Euro area 1970-2019	US 1966-2009
Output growth	0.14	0.08
Inflation	0.86	0.76
Labour input	-	0.96
Consumption	-	0.06
Investment	-	0.31
Wages	-	0.03

Source: Authors' calculations based on AWMD data (İpek and Kısacıkoğlu, 2025) and Slobodyan and Wouters (2012) US time series.

Table 6 shows R^2 for the two forward-looking variables in our model that has a corresponding time series (our third and final forward-looking variable is consumption which is determined within the model and therefore does not have a corresponding observed time series to analyze), and the same for Slobodyan and Wouters (2012). The fit is similar for output growth and inflation, meaning that the need for learning in Slobodyan and Wouters (2012) can probably not be attributed to any dynamics in any of those two time series. Instead, the difference could come from the four additional time series that

[22]While Jarociński and Karadi (2020) originally apply their methodology to the United States, they also provide an updated euro area shock series, constructed using surprises in overnight index swap rates within a 30-minute window around ECB Governing Council press statements and a 90-minute window around press conferences. It is this euro area shock series that we use in our local projection regressions.

Slobodyan and Wouters (2012) employ: labour input, consumption, investment and wages. In Table 6 we see that three of these have very small R^2 which means that the AR(2) models are incapable of accurately describing them. Possibly this mean that adaptive learning could improve the fit for these three time series. However, there is no guarantee for this at all, as it could also be that there are no autoregressive patterns at all. Assuming that the fit of these three time series could be improved upon by allowing for adaptive learning of the AR(2) coefficients, this would push the estimate of ρ_{AL} upwards. It does not matter that inflation is fairly well modeled by a fixed AR(2), as ρ_{AL} is estimated as the same for all forward-looking variables, and crucially, allowing for learning in an AR(2) model with very good fit will not notable depreciate its fit. The reason for this is the learning stems from forecast errors and a model with a good fit will not have any noticeable forecast error to that leads to changes in the model coefficients. The implication is a kind of asymmetry when adding more time series: adding a time series that does not require any adaptive learning to improve its fit will not put downward pressure on ρ_{AL}, but adding a time series that would benefit from learning will affect ρ_{AL} upwards. From this follows that a model with more time series are more likely to benefit from augmenting it with adaptive learning. We cannot confidently claim that this is the case for Slobodyan and Wouters (2012) but the fact that their additional four time series seems to not be well described by fixed AR(2) processes points towards the fact that the inclusion of them could be the driving factor behind the ρ_{AL} close to unity. If these variables are driving the need for learning, one should ask the question whether it is realistic to assume that agents adaptively learn about them. The rationale behind using adaptive learning is that private agents are aware of important macro variables (that receive attention in the news) such as growth and inflation, but based on that one could hardly argue that consumers and firms adaptively learn about about variables such as investment and wages that do not receive the same amount of attention. To summarize, we cannot conclude that the larger emphasis on learning in Slobodyan and Wouters (2012) compared to in our model (which ρ_{AL} implies), is due to the dynamics of the extra time series. Instead, we can only point towards it as a potential explanation for our starkly differing results.