De-Anchored Inflation Expectations and Monetary Policy*

Johannes J. Fischer†

European University Institute

Click here for the most recent version

This version: October 13, 2021

Abstract

This paper studies the conduct of monetary policy in a model with an endogenous degree of expectations anchoring. I use an estimated New Keynesian model with endogenous forecast switching to replicate the time-varying excess sensitivity of long-term inflation expectations to inflation surprises as well as the resulting movements of long-term inflation expectations. In this model de-anchoring leads to increased inflation volatility and can cause deflationary spirals when the zero lower bound (ZLB) is binding. This can be prevented by an asymmetric monetary policy rule which responds more aggressively to below-target inflation. Price Level Targeting on the other hand can increase the risk of deflationary spirals near the ZLB.

JEL Classification: E31, E52, E70

Keywords: Monetary policy, Inflation Expectations, Non-rational Expectations

*I am indebted to my supervisors Ramon Marimon and Edouard Challe for their continuous guidance and support. I, furthermore, thank Leonardo Melosi, Martin Hellwig, Mirko Wiederholt, Michael Bauer, Matthias Rottner, Oliver Pfäuti, Wolfram Horn, as well as conference and seminar participants of the the 52^{nd} Annual Conference of the Money, Macro and Finance Society, the 2021 Annual Meeting of the Central Bank Research Association, the 55^{th} Annual Conference of the CEA, EcoMod2021, the 2021 Annual Congress of the Swiss Society of Economics and Statistics, the DIW, and the EUI for helpful discussions and comments.

†Author's contact: European University Institute, Department of Economics, Via delle Fontanelle 18, 50014 San Domenico di Fiesole (FI), Italy; Email: johannes.fischer@eui.eu

1 Introduction

The anchoring of inflation expectations is a central tenet of both monetary theory and monetary policy making. Most central banks in advanced economies try to anchor long-term inflation expectations to their inflation target. Anchored long-term expectations reflect the trust that the central bank will offset any inflationary shocks and eventually brings inflation back to target, thus preventing these shocks from feeding into wage and price setting. When households lose trust in the central bank's ability or willingness to bring inflation back to target, they increasingly believe that temporary inflation movements will be permanent and their long-term expectations de-anchor from the central banks target (see Figure 1). This belief can be justified if the central bank indeed does not enact appropriate policy and in this case the policy prescription is straightforward. However, this belief can also arise even if the underlying shock is temporary and the central bank attempts to bring back inflation to target.[1] This model-inconsistent (and thus non-rational) belief will be the focus of this paper because the policy recommendations are less straightforward in this case. Furthermore, the COVID-19 pandemic has renewed the importance of this question as the monetary responses to the pandemic have nourished renewed fears of de-anchoring (Reis, 2021).

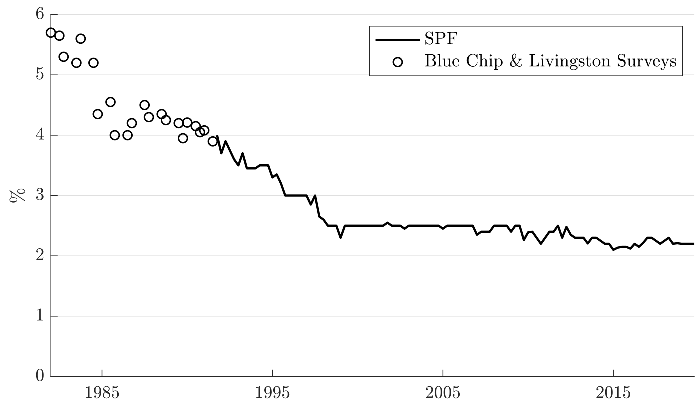
Therefore, this paper studies the conduct of monetary policy when the anchoring of expectations is evolving endogenously. First, I illustrate the time-varying excess sensitivity of (long-term) inflation expectations to inflation surprises using survey data from the U.S. This behaviour is consistent with forecasters perceiving temporary shocks as having permanent effects even though the Fed arguably was fully committed to stabilising inflation when expectations appeared to be de-anchored. To explain this pattern of de-anchoring, I introduce adaptive learning[2] (Evans and Honkapohja, 2001) into the heuristic switching approach of Brock and Hommes (1997) embedded in an otherwise standard New Keynesian (NK) model with sticky prices. By combining adaptive learning with naive forecasting in a model with optimal forecast selection, I provide a novel, quantitatively realistic model of expectation de-anchoring. Using this model, I discuss the stability properties of two monetary

[1]Unfortunately, there is no widely agreed upon definition of de-anchored expectation (Kumar et al., 2015) Instead, researchers work with a set of characteristics that are commonly thought to characterise anchored expectations, such as closeness to the central bank's target, no pass-through from short- to long-term expectations, no sensitivity to inflation news. A common theme of de-anchored expectations is that households perceive temporary shocks to have persistent effects, even if this is not justified by the fundamental structure of the economy. These subjective beliefs then can lead to self-enforcing inflation movements. Because these expectations are not model consistent, de-anchored expectations represent deviations from rational expectations.

[2]Adaptive learning replaces rational expectations by assuming that forecasters use an ad-hoc forecasting rule to form expectations using a given information set. The coefficients of this forecasting rule are updated in every period using recursive estimation techniques

policy frameworks, Inflation Targeting (IT) and Price Level Targeting (PLT). Finally, I take the model to the data and provide the first general equilibrium model with endogenous de-anchoring that is fully estimated using likelihood methods. I use this model to quantitatively evaluate the two different policy frameworks in their ability to stabilise inflation.

Figure 1: 10-Year Ahead U.S. Inflation Expectations



To motivate my modelling choices, I first demonstrate the pattern of de-anchoring in the expectations reported by the U.S. survey of professional forecasters (SPF). Importantly, the excess sensitivity of expectations to inflation surprises arises jointly in short-term and long-term expectations. My main contribution is, therefore, a novel model of forecast switching that directly replicates the time-varying excess sensitivity of short-term expectations to inflation and also matches the pattern in long-term expectations. In this model, households do not know the true structure of the economy so they form expectations using two competing forecasting heuristics: One group uses historical data to learn the relationship between exogenous news and endogenous realisations over time, updating their beliefs through constant-gain learning (the standard approach of Evans and Honkapohja (2001)). The second group uses a naive forecasting heuristic which uses the the previous period's realisation as forecast for the next period. The share of each group evolves according to its relative forecasting accuracy. Therefore, the sensitivity of expectations to past realisations depends on the historical forecasting performance. Since the naive forecasters are the ones who perceive changes of inflation to be permanent, the share of naive forecasters can be directly interpreted as the degree of de-anchoring.

The model features the same steady state as under rational expectations (RE) but when expectations de-anchor, the volatility of output growth and inflation increases. Monetary policy can prevent expectations de-anchoring from causing inflationary or deflationary spirals as long as its not constrained by the zero lower bound if it reacts more than one-to-one to movements in inflation. At the zero lower bound, however, de-anchoring can lead to self-reinforcing deflationary spirals (a feature absent from the standard rational expectations model and different from the self-fulfilling liquidity trap of, e.g., Benhabib et al. (2002)). Thus, the potential welfare loss of de-anchoring is asymmetric and bigger in a low interest rate environment. My results have implications regarding the monetary policy regime most likely to ensure price and macroeconomic stability. I find that Inflation Targeting ensures stationary model dynamics even when expectations are fully de-anchored. However, under Price Level Targeting there exists a threshold level of de-anchoring, above which the model becomes unstable. This is the case because the combination of naive forecasting and a history dependent policy rule leads to explosive dynamics if naive forecasters dominate the economy. Thus, my findings contradict the notion that PLT is an optimal policy to address liquidity traps (e.g. Evans, 2012).

As an empirical contribution I subsequently estimate the model using a particle filter and observed expectations from the U.S. Here, I first establish using the estimated marginal likelihood that indeed, the model with forecast switching overall fits the data much better than a simple rational expectations model. Furthermore, the model-implied long-term inflation expectations fit the data quite well even though long-term expectations were not used in the estimation. As a result, the model-implied expectations, both short-term and long-term, exhibit the same degree of time-varying sensitivity to news as the observed forecasts of the SPF. Finally, the model-implied share of naive forecasters suggests that expectations were highly de-anchored from the 1980s until the early 1990s. After a period of re-anchoring the Global Financial Crisis led to another uptick of de-anchoring. This uptick, however, was small and short-lived.

After having established that the estimated model provides a realistic description of expectation formation, I use simulation methods to investigate the performance of the two policy regimes. Price Level Targeting leads to higher welfare than Inflation Targeting in a high nominal interest rate environment. However, in a low nominal interest rate environment (i.e. close to the ZLB), Price Level Targeting entails a high risk of deflationary spirals. This is not the case under Inflation Targeting. Since the biggest risk of de-anchoring is the emergence of belief-driven deflationary spirals at the zero lower bound, I argue that an asymmetric Inflation Targeting framework (in form of the asymmetric Taylor type rule of Bianchi et al. (2019)) can address this directly: By reacting more forcefully to below-target than to above-target inflation, this policy stabilises inflation and prevents deflationary spirals. Furthermore, it leads to higher welfare than a symmetric Inflation Targeting framework, both in low and in high nominal interest rate environments. Therefore, the asymmetric policy rule is more robust than Price Level Targeting across different nominal interest rate environments and dominates a symmetric policy rule in terms of welfare.

The paper contributes to the literature in several ways: On the theoretical side, I incorporate adaptive learning into a general equilibrium model featuring heuristic switching. This extends the existing literature on central bank credibility (e.g. Hommes and Lustenhouwer, 2019; Honkapohja and Mitra, 2020), which commonly constrains heuristics to be either naive or anchored to the central banks target. This framework is not well suited to study de-anchoring: Any deviation from the central banks target leads to the adoption of the naive forecasting rule as households are not aware of the stochastic processes in the model economy. In contrast, in my setup agents can make use of either current news to form expectations (i.e. the solution of the standard NK model under rational expectations) or naively forecast last period's realisations. Time-varying shares generate dynamic interactions between the two heuristics that affect not only the mean of the process as in Goy et al. (2020) but also the endogenous amplification of shocks. These two heuristics are reasonably close to the specification used in Branch (2004), who provides empirical evidence for dynamic choice between heterogeneous predictors among inflation forecasters.

On the empirical side, my contribution is twofold. First, I show that the patterns of de-anchoring arise jointly in short- and long-term expectations. Second, this is the first paper to structurally estimate the parameters of the heuristic switching approach in a general equilibrium setting (and the first fully estimated paper with expectation de-anchoring). Previous studies have only done so in a partial equilibrium setting (e.g. Cornea-Madeira et al., 2019). I do so explicitly accounting for the model non-linearities by evaluating the model likelihood using the particle filter. Closely related is the recent paper by Ilabaca and Milani (2020) who estimate a model with time-varying, heterogeneous expectations. However, the variation in the shares of different heuristics is not endogenous but a results of the time-varying estimation approach.

Finally, I improve over the simulation exercises used in the existing literature: Simulating the linearised model does not adequately capture the performance of monetary policy in a model with many non-linearities such as the ZLB. Therefore, I simulate the heuristic switching model using the time iteration approach of Richter et al. (2014).

My paper is most closely related to the literature studying the de-anchoring of expectations. On the empirical side, a wide range of papers investigates whether expectations are anchored or not (e.g. Reis, 2021; Łyziak and Paloviita, 2017; Strohsal et al., 2016). On the theoretical side, Eusepi et al. (2020) and Eusepi et al. (2021) study the conduct of monetary policy for different degrees of the sensitivity of long-term inflation expectations to short-run forecast errors, which is their proxy for de-anchoring. In a related paper Jorgensen and Lansing (2019) study how the anchoring of the expectations contributed to the flattening of the Philips curve. Carvalho et al. (2020) and Gáti (2020) extend this framework and endogenise the sensitivity of long-term inflation expectations by using the endogenous gain learning approach introduced by Marcet and Nicolini (2003). I deviate from this literature and instead use the heuristic switching approach of Brock and Hommes (1997). This approach directly models the underlying property of de-anchoring - the belief that temporary shocks to inflation are permanent - instead of a symptom (a higher sensitivity to inflation surprises). Furthermore, my model fits the empirical properties of expectation formation (see Angeletos et al., 2020) better than the adaptive learning framework (see Appendix D.2). The heuristic switching approach of Brock and Hommes (1997) is commonly used in the closely related strand of literature that studies central bank credibility (e.g. Hommes and Lustenhouwer, 2019; Honkapohja and Mitra, 2020). In this literature agents typically chose between naively forecasting the past into the future or by forecasting the central banks target. As discussed before, this framework is not well suited to study de-anchoring: Any deviation from the central banks target leads to the adoption of the naive forecasting rule as households are not aware of the exogenous shocks. Therefore, I extend this framework and introduce adaptive learning to endow households with more information about the underlying states of the model. Most similar to my paper is Gibbs and Kulish (2017), who study disinflations under imperfectly anchored expectations. As in my paper, their model features heterogeneous forecasting rules. De-anchoring is driven by the the share of adaptive, backward looking forecasters relative to forecasters using rational expectations. Unlike in my paper, however, the degree of de-anchoring is not determined endogenously.

More generally, my paper is connected to the literature that attempts to explain features of expectations data that do not coincide with the predictions of the standard full-information rational expectations hypothesis (e.g. predictability of forecast revisions (Coibion and Gorodnichenko, 2015), over- and under-reaction to news (Bordalo et al., 2018), heterogeneity of forecasts (Mankiw et al., 2003), heterogeneity of forecasting tools (Pfajfar and Santoro, 2010)). There is a wide range of literature that attempts to explain these patterns (e.g. the incomplete/noisy information assumption of Mankiw and Reis (2002) or the adaptive learning framework of, among others, Sargent (1999)). Angeletos et al. (2020) provide stylised facts of expectations to distinguish among these competing theories. My analysis embeds both incomplete information and adaptive learning. My results suggest that both are important to understand the de-anchoring of expectations and to match the observations of Angeletos et al. (2020).

The paper proceeds as follows. Section 2 documents the empirical patterns associated with de-anchored inflation expectations. Section 3 introduces the model with heterogeneous expectations. Section 4 describes the model solution and stability properties. Section 5 describes the estimation procedure and illustrates the dynamics of the estimated model.

Section 6 provides simulation evidence on the performance of different policy rules and Section 7 concludes.

2 De-Anchoring in Practice

When doubts arise regarding the central bank's ability (or willingness) to offset inflation shocks, these movements are increasingly perceived to be permanent even if this is not justified. When this happens, inflation expectations reflect this perceived increased persistence by becoming significantly reactive to forecast errors (e.g. Strohsal et al., 2016) and to past inflation realisations (e.g. Ehrmann, 2015). In the following, I illustrate these two patterns and show that they arise jointly in short- and long-term expectations.

Denoting $\hat{\mathbb{E}}_{i,t}\pi_{t+k}$ as the k quarter ahead subjective inflation forecast of forecaster i at time t reported in the Philadelphia Fed's Survey of Professional Forecasters (SPF), I estimate

$$\hat{\mathbb{E}}_{i,t}\pi_{t+k} = \alpha^w_{1,i} + \beta^w_1(\pi_{t-1} - \hat{\mathbb{E}}_{i,t-5}\pi_{t-1}) + \varepsilon_{1,t} \tag{1}$$

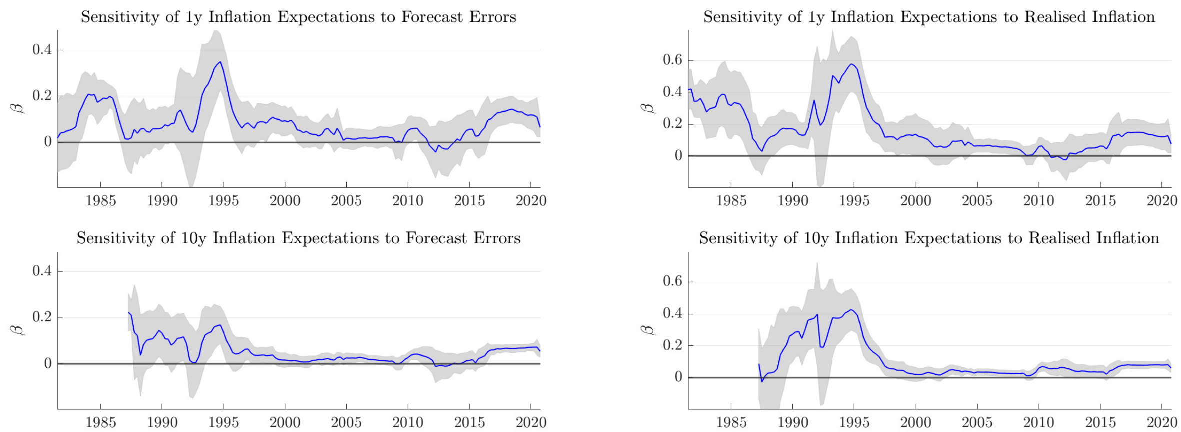
where w indexes rolling windows of 20 quarters, α^w_i are forecaster fixed effects, and $k \in (4, 40)$. That is, I regress the 1-year and 10-year ahead (CPI) inflation forecasts of the panel of the SPF forecasters on lagged[3] individual forecast errors. Figure 2 plots the time series of the estimated coefficient β^w along with 95% confidence intervals.[4] As the left column shows, both long *and* short term expectations display time-varying sensitivity to forecast errors: Starting from 1981, 1-year ahead inflation expectations react significantly positive to lagged forecast errors up until the early 2000s (with brief exceptions). Similarly, 10-year ahead expectations (for which data starts in 1991 only) react significantly to forecast errors up until the early 2000's with the exception of the first few quarters of the sample. With the beginning of the early 2000's, expectations over both horizons became much more anchored. In the years following the Great Recession, expectations at times became responsive to news again, but only temporarily and with rather small magnitude.

A direct consequence of de-anchoring is that the sensitivity of inflation expectations to

[3] I lag the respective regressors to account for the lagged publication of inflation realisations so to make sure that the respondents in the SPF have access to at least the first estimates of realised inflation.

[4] Standard errors are estimated by clustering on the forecaster and quarter level

Figure 2: Time-Varying Sensitivity of Expectations to Inflation Surprises and Inflation



Note: The panels show the estimated time-varying sensitivity of short-run (upper row) and long-run (lower row) inflation expectations from the Survey of Professional Forecasters to inflation surprises (left column) and realised inflation (right column). Solid blue lines depict the estimated time-varying coefficient with 95% confidence intervals shaded in grey.

realised inflation increases as well.[5] To demonstrate this, I estimate the following model

$$\hat{\mathbb{E}}_{i,t}\pi_{t+k} = \alpha^w_{2,i} + \beta^w_2 \pi_{t-1} + \varepsilon_{2,t} \tag{2}$$

in the same fashion as Equation 1. As the right column of Figure 2 shows, the sensitivity of inflation expectations to realised inflation behaves almost the same way as the sensitivity to forecast errors. This is not surprising for two reasons: First, when households perceive temporary inflation shocks to have permanent effects because they doubt the centrals bank's ability to offset these shocks, forecasts should reflect this perceived increased persistence. Second, the only difference between the two regressions is the inclusion of the lagged inflation forecast.

As a robustness check, I show in Appendix A.1 that the same pattern arises when using the inflation expectations of the Michigan Fed's Survey of Consumers instead of professional forecasts. Furthermore, in Appendix A.2 I show that the same pattern arises in the euro zone.

[5] Sensitivity to past inflation realisations can be explained in a rational expectations framework by a history dependent monetary policy rule or highly persistent inflation shocks, but sensitivity to forecast errors cannot be explained this way. Nonetheless, we would expect a higher sensitivity of inflation expectations to realised inflation as a consequence of de-anchoring. This is because of 1) mechanical reasons and 2) because households perceive temporary inflation shocks to have permanent effects because they doubt the centrals bank's ability to offset these shocks. Therefore, testing the sensitivity of inflation expectations to forecast errors enables us to test whether a high sensitivity to realised inflation is a consequence of de-anchoring or borne out by structural forces.

To summarise, short-run inflation expectations simultaneously exhibit the same signs of de-anchoring as long-run expectations: They become increasingly sensitive to inflation news such as realised inflation or forecast errors in the same periods as long-term expectations do. In the next Section I build a model of de-anchoring using the heuristic switching approach of Brock and Hommes (1997) to replicate this pattern. I focus on replicating the time-varying excess sensitivity of short term expectations and then show that this accurately captures the behaviour of long-term expectations as well.

3 A Model of De-anchoring

I use a standard New Keynesian (NK) model to study the de-anchoring of expectations. To isolate the effect of de-anchoring, I keep the model simply and only introduce one distortion, sticky prices à la Rotemberg (1982). In the following, I first lay out the non-linear version of the model before turning to the linearised version and discussing the expectation formation.

3.1 The New Keynesian Framework

The model is populated by a representative household, intermediate and final goods producers, a central bank, and the government. The representative **household**[6] chooses consumption C_t, labour H_t, and government bonds B_t to maximise the expected discounted stream of utility

$$\hat{\mathbb{E}}_0 \sum_{t=1}^{\infty} \beta^t \Xi_{t-1} \left[\frac{C_t^{1-\sigma}}{1-\sigma} - \chi \frac{H_t^{1+\eta}}{1+\eta} \right] \tag{3}$$

subject to the budget constraint

$$P_t C_t + B_t = P_t W_t H_t + R_{t-1} B_{t-1} + P_t Div_t + T_t \tag{4}$$

where $\hat{\mathbb{E}}$ is the non-rational expectations operator, Ξ_t is a shock to the discount factor, P_t is the price level, W_t is the real wage, R_t is the gross interest rate, T_t are lump-sum taxes and Div_t are real profits from the intermediate good firms. B_t denotes the one-period government bonds in zero net supply. Solving the representative household's problem yields the Euler

[6] To simplify notation I use the representative household framework to derive the model conditions and then impose heterogeneous expectations only ex-post. For a microfounded approach with heterogeneous expectations see Branch and McGough (2009). They propose a set of assumptions that restrict agents' expectations so that they satisfy the law of iterated expectations at both the individual and aggregate level, which enables aggregation.

equation

$$1 = \beta R_t \hat{\mathbb{E}}_t \frac{\Xi_t}{\Xi_{t-1}} \left(\frac{C_t}{C_{t+1}} \right)^{\sigma} \frac{1}{\Pi_{t+1}} \tag{5}$$

and the labor supply

$$W_t = \chi H_t^{\eta} C_t^{\sigma} \tag{6}$$

The **final goods producers** transform intermediate goods into the homogeneous good, which is obtained by aggregating intermediate goods using the following technology

$$Y_t = \left(\int_0^1 Y_t(j)^{\frac{\epsilon-1}{\epsilon}} dj \right)^{\frac{\epsilon}{\epsilon-1}} \tag{7}$$

where $Y_t(j)$ is the intermediate good of firm j. The price index for the aggregate homogeneous good is

$$P_t = \left[\int_0^1 P_t(j)^{1-\epsilon} dj \right]^{\frac{1}{1-\epsilon}} \tag{8}$$

and the demand for the differentiated good $j \in (0, 1)$ is

$$Y_t(j) = \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} Y_t \tag{9}$$

The **intermediate goods producer** j produces using labour as the only input

$$Y_t(j) = A_t H_t^{1-\alpha} \tag{10}$$

where A_t denotes the total factor productivity, which follows an AR(1) process in logs like the discount factor shock. The firm j sets the price $P_t(j)$ of its differentiated good so as to maximise its profits

$$\mathrm{Div}_t(j) = P_t(j) \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} \frac{Y_t}{P_t} - \alpha\, mc_t \left(\frac{P_t(j)}{P_t} \right)^{-\epsilon} Y_t - \frac{\varphi}{2} \left(\frac{P_t(j)}{\bar{\Pi} P_{t-1}(j)} - 1 \right)^2 Y_t \tag{11}$$

subject to the demand curve for intermediate goods and where $\Pi \geq 1$ is the steady state inflation rate. The parameter $\varphi \geq 0$ measures the cost of price adjustment in units of the final good with $\varphi = 0$ leading to the flexible price output level Y_t^*. Solving the intermediates firm problem can be shown to yield the New Keynesian Phillips curve

$$\varphi \left(\frac{\Pi_{t+1}}{\bar{\Pi}} - 1 \right) \frac{\Pi_t}{\bar{\Pi}} = (1 - \epsilon) + \epsilon \alpha M C_t + \varphi \hat{\mathbb{E}}_t \Lambda_{t,t+1} \left(\frac{\Pi_{t+1}}{\bar{\Pi}} - 1 \right) \frac{\Pi_{t+1} Y_{t+1}}{\bar{\Pi}} \tag{12}$$

where $\Lambda_{t,t+1}$ is the household's stochastic discount factor.

The fiscal authority sets taxes to balance the budget in every period, so that the **aggregate resource constraint** is

$$C_t = Y_t \left[1 - \frac{\varphi}{2} \left(\frac{\Pi_t}{\bar{\Pi}} - 1 \right)^2 \right] \tag{13}$$

Combining and linearising the Equations 3 - 13 around the zero inflation steady state yields the familiar 3 equation system (where small letters symbolise deviations of the respective variable from its steady state)

$$\begin{aligned} y_t &= \hat{\mathbb{E}}_t \left\{ y_{t+1} \right\} - \tfrac{1}{\sigma} \left(r_t - \hat{\mathbb{E}}_t \pi_{t+1} + \xi_t \right) \\ \pi_t &= \beta \hat{\mathbb{E}}_t \left\{ \pi_{t+1} \right\} + \underbrace{\frac{\varepsilon - 1}{\varphi} \left(\sigma + \frac{\chi\eta + \alpha}{1 - \alpha} \right)}_{\equiv \kappa} \left[y_t - \underbrace{\tfrac{1+\chi\eta}{1-\alpha} \big/ (\sigma + \tfrac{\chi\eta+\alpha}{1-\alpha})}_{\equiv \vartheta} a_t \right] \end{aligned} \tag{14}$$

with

$$\begin{aligned} a_t &= \rho_a a_{t-1} + \varepsilon_t^a \\ \xi_t &= \rho_\xi \xi_{t-1} + \varepsilon_t^\xi \end{aligned} \tag{15}$$

where $\xi_t = \log \left(\frac{\Xi_t}{\Xi_{t-1}} \right)$. The shocks ε_t^a and ε_t^ξ are distributed normally with standard deviations σ_a and σ_ξ around a mean of zero.

To close the model, I consider two different monetary policy frameworks for the **central bank**, Inflation Targeting (IT) and Price Level Targeting (PLT). Under Inflation Targeting, the central bank follows a standard Taylor rule subject to the zero lower bound (ZLB)

$$r_t = \max \left[0, \bar{\pi} + \phi_\pi (\pi_t - \bar{\pi}) + \phi_y \hat{y}_t + m_t \right] \tag{16}$$

where m_t evolves according to

$$m_t = \rho_m m_{t-1} + \varepsilon_t^r \quad \varepsilon_t^r \sim \mathcal{N}(0, \sigma_r)$$

and y^* is the output gap

$$\hat{y}_t = y_t - \vartheta a_t \tag{17}$$

Under Price Level Targeting the central bank reacts to the deviation of the price level p_t

from the target price level path p_t^*

$$r_t = \max\left[0, \bar{\pi} + \phi_\pi \hat{p}_t + \phi_y \hat{y}_t + m_t\right] \tag{18}$$

where $\hat{p}_t$ evolves according to $\hat{p}_t = \pi_t - \bar{\pi} + \hat{p}_{t-1}$ and $\hat{p}_{t-1}$ enters the model as an additional state variable.

3.2 Heterogeneous Expectations

To model the de-anchoring of expectations I deviate from the rational expectations benchmark and assume that agents do not know the true model of the economy. Instead, the environment is populated by two kinds of agents: One group of agents forms expectations by following a naive forecast heuristic; a second group acts as econometrician and tries to estimate the law of motion of the economy using constant gain learning. Both groups make use of the same information set $\boldsymbol{I}_t = \{\boldsymbol{z}_{\{t-1,t-\infty\}}, \boldsymbol{w}_{\{t,t-\infty\}}\}$ but form disparate forecasts. That is, both groups observe the contemporaneous values of the exogenous processes $\boldsymbol{w}$ but observe the endogenous realisations $\boldsymbol{z}$ only with a lag of one period.[7] The average forecast is a weighted average of the two heuristics, i.e.

$$\hat{\mathbb{E}}_t \mathbf{z}_{t+1} = (1 - n_t)\hat{\mathbb{E}}_t^a \mathbf{z}_{t+1} + n_t \hat{\mathbb{E}}_t^n \mathbf{z}_{t+1} \tag{19}$$

where $\hat{\mathbb{E}}^a$ denotes the expectations of the adaptive learners, $\hat{\mathbb{E}}^n$ denotes the expectations of the naive forecasters, n_t is the share of naive forecasters in the economy, and $\mathbf{z}_t$ is a vector containing the endogenous realisation y_t, π_t (and $\hat{p}_t$ in the case of Price Level Targeting).

Adaptive learners assume that the endogenous variables $\mathbf{z}_t$ evolve according to the following linear perceived law of motion (PLM):

$$\mathbf{z}_t = \Psi_t \mathbf{x}_{t-1} + \epsilon_t \tag{20}$$

where $\boldsymbol{\Psi}_t = [\mathbf{a}_t,\, \mathbf{b}_t]$ and $\mathbf{x}_t = (1,\, \xi_t,\, a_t,\, m_t)'$ so that the corresponding forecast becomes

$$\hat{\mathbb{E}}_t^a \mathbf{z}_{t+1} = \Psi \mathbf{x}_t \tag{21}$$

Under Price Level Targeting the adaptive forecasting heuristic additionally makes use of

[7] The learning rule introduced here is not derived at the micro level and then aggregated. Instead, they are imposed ex post, on the demand and supply equations. This has also been the approach in the learning literature pioneered by Evans and Honkapohja (2001)

the lagged price level gap to forecast $\mathbf{z}$ so that $\mathbf{\Psi}_t = [\mathbf{a}_t,\, \mathbf{b}_t,\, \tilde{\mathbf{c}}_t]$ and $\mathbf{x}_t = (1,\, \xi_t,\, a_t,\, m_t,\, \hat{p}_{t-1})'$.[8] In both cases the available information set is equivalent to the minimum state variable solution under rational expectations. However, the adaptive learners do not know the exact structural relationship between the exogenous shocks and the endogenous variables. Furthermore, they do not observe the share of naive forecasters n_t. Therefore, the adaptive learners re-estimate their perceived law of motion and update the parameters $\mathbf{\Psi}_t$ in each period when new information becomes available using:

$$\begin{gathered}\mathbf{\Psi}'_t = \mathbf{\Psi}'_{t-1} + \gamma \boldsymbol{R}^{-1}_{t-1} \boldsymbol{x}_{t-2} \left(\boldsymbol{z}_{t-1} - \mathbf{\Psi} \boldsymbol{x}_{t-2}\right)' \\ \mathbf{R}_t = \mathbf{R}_{t-1} + \gamma \left(\mathbf{x}_{t-2}\mathbf{x}'_{t-2} - \mathbf{R}_{t-1}\right)\end{gathered} \tag{22}$$

where γ is the constant gain parameter determining the weight given to new information. A constant γ implies that recent observations receive relatively higher weights than older observations. This loss of distant memory is supported by the experience based learning documented by Malmendier and Nagel (2016). If instead $\gamma = {}^1\!/_t$, each observation would be equally weighted and the updating of coefficients would be equivalent to running a recursive OLS regression in every period. In either case, if the economy was only populated by adaptive learners, the coefficient matrix $\mathbf{\Psi}_t$ would eventually converge to its rational expectations counterpart. However, the presence of the naive forecasters prevents this from happening. These agents simply forecast

$$\hat{\mathbb{E}}^n_t \mathbf{z}_{t+1} = \hat{\mathbb{E}}^n_t \mathbf{z}_t = \mathbf{z}_{t-1} \tag{23}$$

Since the share of naive forecasters is time-varying, the model has to be closed by specifying a law of motion for n_t: Agents evaluate each heuristic based on its respective forecasting performance in terms of the sum of its squared forecast errors. A standard way to formalise this is to assume that the probability of choosing the naive forecasting rule follows a logistic distribution (Brock and Hommes, 1997).[9] This smooth transition function maps the forecast errors into the share of the population n_t that follows the naive forecasting rule:

$$n_t = \frac{\exp\left(-\theta e^n_{t-1}\right)}{\exp\left(-\theta e^n_{t-1}\right) + \exp\left(-\theta e^a_{t-1}\right)} \tag{24}$$

with

$$e^i_t = (1-\gamma) e^i_{t-1} + \gamma \varepsilon^i_t \quad i \in \{a, n\} \tag{25}$$

[8] $\tilde{\mathbf{c}}_t$ is a 3×1 vector of coefficients assigned to the lagged price level gap and denoted by a tilde because later it will be convenient to define $\mathbf{c} = [\mathbf{0}_{3\times 2}\ \tilde{\mathbf{c}}]$

[9] This can be formalised by assuming that the forecasting performance of both heuristics is publicly available, but subject to noise. Assuming that the noise is logistically distributed, the probability that an agent chooses a given heuristic is then given by the logistic function (e.g. Anufriev et al., 2013; De Grauwe and Ji, 2019).

where e^i_{t-1} is the recursive weighted average of the sum of squared forecast errors ε^i_t. That is, agents avoid making systematic forecasting mistakes by switching to the better performing forecast rule. The two parameters γ and θ jointly determine how the shares of agents using each forecasting model evolve over time: γ controls the rate at which previous forecast errors decay. For simplicity, I assume that the recursive weihted moving average of forecast errors is governed by the same updating parameter as the updating equations (22). θ determines the sensitivity to differences in the forecasting performance: When $\theta \to \infty$, agents immediately switch to the forecasting rule with the lower forecast error. Conversely, with $\theta = 0$ agents never switch and the shares remain fixed at 0.5. Any intermediate value implies an imperfect adoption of the better performing forecast rule. In the steady state, that is when neither heuristic makes any forecast error because there are no exogenous shocks, each group makes up half of the economy and $n = 0.5$. Unlike, for example, De Grauwe and Ji (2019), I assume that the two forecasting rules are compared on their overall performance and not separately for each endogenous variable, i.e. n_t is a scalar and not a vector. This expectation formation process has empirically plausible properties in the spirit of Angeletos et al. (2020): In response to a shock, expectations initially under-react, followed by a delayed over-reaction (see Appendix D.2).

One way to interpret this information setup is that there exists a subgroup of informed agents in the population that reads current news and makes their forecasts accordingly (e.g. increasing inflation expectations when reading about a positive oil price change). The second group on the other hand is uninformed and simply forecasts based on recent experience. This difference could be due to some cost of information acquisition which the group of adaptive learners considers worthwhile whereas naive forecasters do not. Alternatively, this information setup can of course also be interpreted as one representative agent using a weighted average of two forecasting heuristics.

Both of these forecasting heuristics are of course biased in the sense that they do not incorporate all available information when forming expectations. In fact, each heuristic makes use of only one subset of relevant predictors, namely either exogenous processes or lagged endogenous realisations. However, as De Grauwe (2011) points out, the bias of the forecasting heuristics does not necessarily imply that agents are irrational: While the individual heuristics may be biased, agents try to reduce that bias and react to forecast errors by adapting better performing heuristics.

This heuristic switching mechanism provides an easy statistic measuring the degree of expectation anchoring: When expectations are well anchored, the adaptive rule will produce good forecasts and the share of naive forecasters will go down. When expectations are not well anchored, for example because the central bank does not react enough to offset shocks that move away output and inflation from their target, the adaptive rule will produce bad forecasts and the share of naive forecasters will go up. The role of monetary policy in this context therefore is to send correct signals for the evolutionary selection of strategies and induce stable dynamics converging to the rational expectations steady state.

3.3 Implied Long-Term Expectations

The heuristic switching model implicitly also pins down long-term expectations. The long-term forecast (i.e. the forecast for a horizon k sufficiently large) of the adaptive heuristic is simply the intercept of the learning rule:

$$\hat{\mathbb{E}}^a_t \mathbf{z}_{t+k} = \mathbf{a} \tag{26}$$

Naive forecasters extend their forecasts over the relevant horizon using the following formulation

$$\hat{\mathbb{E}}^n_t \mathbf{z}_{t+k} = (1 - {}^1\!/_k)\hat{\mathbb{E}}^n_{t-1} \mathbf{z}_{t+k-1} + {}^1\!/_k\, \mathbf{z}_{t-1} \tag{27}$$

That is, the naive heuristic specifies forecasts over horizon k as the recursive weighted average of realisations over the previous k periods. For $k = 1$ this heuristic nests the one quarter ahead forecast discussed previously. Furthermore, for $k \to \infty$ the naive forecast will converge to the mean of the process $\mathbf{z}$, which corresponds to the long-run forecast of the adaptive heuristic.[10]

The average long-run forecast is a weighted average of the two heuristics. As before, the share of agents using a naive forecast heuristic is pinned down by the accuracy of the one-quarter ahead forecast. The same share then pins down the weights for the long-term forecast:

$$\begin{aligned} \hat{\mathbb{E}}_t \mathbf{z}_{t+k} &= (1 - n_t)\hat{\mathbb{E}}^a_t \mathbf{z}_{t+k} + n_t \hat{\mathbb{E}}^n_t \mathbf{z}_{t+k} \\ &= (1 - n_t)\mathbf{a} + n_t((1 - {}^1\!/_k)\hat{\mathbb{E}}^n_{t-1} \mathbf{z}_{t+k-1} + {}^1\!/_k\, \mathbf{z}_{t-1}) \end{aligned} \tag{28}$$

This expectation formation process is empirically plausible: Following a monetary policy shock, these expectations react in the same way as observed long-run expectations (see Appendix D.1).

[10] This setup assumes that 1) the adaptive learners do not internalise future revisions of the learning rule leading to changes in the intercept (this assumption is known in the learning literature as anticipated utility (Kreps, 1997)) and that 2) neither group anticipates switching heuristics in the future. This of course imposes a higher degree of non-rationality but keeps the model tractable.

4 Model Equilibria and Stability

In this Section I solve the model under the two different monetary policy frameworks, Inflation Targeting and Price Level Targeting, and discuss the stability properties both away from and at the ZLB. In each case, I first derive the fixed points of the adaptive forecasting heuristic under some fixed share n of naive forecasters. In a second step, I use these fixed points to characterise the steady state under time varying n_t. In a third step I discuss the E-Stability of this steady state, i.e. whether the steady state can be learned.[11]

4.1 Dynamics under Inflation Targeting

To simplify the analysis I assume that the central bank only reacts to deviations of inflation from target (i.e. $\phi_y = 0$). Furthermore, I initially consider an environment without the zero lower bound in 4.1.1 before reintroducing it in 4.1.2.

4.1.1 Away from the ZLB

Substituting the monetary policy rule as well as the perceived laws of motion (20) & (23) into the system of equations (14) and rearranging yields the Actual Law of Motion (ALM)

$$\begin{aligned} y_t &= (1-n_t)[a_y + b_{y,a}a_t + b_{y,\xi}\xi_t + b_{y,m}m_t] + n_t y_{t-1} - \frac{1}{\sigma}(\bar{\pi} + \phi_\pi(\pi_t - \bar{\pi}) \\ &\quad + m_t - (1-n_t)[a_\pi + b_{\pi,a}a_t + b_{\pi,\xi}\xi_t + b_{\pi,m}m_t] - n_t y_{t-1} + \xi_t) \\ \pi_t &= \beta[(1-n_t)[a_\pi + b_{\pi,a}a_t + b_{\pi,\xi}\xi_t + b_{\pi,m}m_t] + n_t\pi_{t-1}] + \kappa(y_t - \vartheta a_t) \end{aligned} \tag{29}$$

or in matrix notation

$$\Leftrightarrow \mathbf{z}_t = \mathbf{A} \underbrace{[(1-n_t)(\mathbf{a} + \mathbf{b}\mathbf{w}_t) + n_t\mathbf{z}_{t-1}]}_{\hat{\mathbb{E}}_t \boldsymbol{z}_{t+1}} + \mathbf{B}\mathbf{w}_t + \mathbf{C}\bar{\mathbf{z}} \tag{30}$$

Substituting (30) into the ODE associated with the updating equations (22) yields

$$\frac{\partial \boldsymbol{\Psi}'}{\partial \tau} = \mathbf{R}^{-1} E\mathbf{x}_{t-2} \left(\mathbf{A} \left[(1-n)\boldsymbol{\Psi}\mathbf{x}_{t-1} \right) \right] + \mathbf{A}n\mathbf{z}_{t-2} + \mathbf{B}\mathbf{w}_t + \mathbf{C}\bar{\mathbf{z}}) - \boldsymbol{\Psi}\mathbf{x}_{t-2})' \tag{31}$$

where τ denotes "notional" time (Evans and Honkapohja, 2001). As long as the ALM in Equation (30) is asymptotically stationary, i.e. has roots within the unit circle so that it is mean-reverting

$$\lambda < 1 \quad \forall \lambda \in \Lambda = \{\Lambda : |\mathbf{I} - n\mathbf{A} - \boldsymbol{\Lambda}\mathbf{I}| = 0\} \tag{32}$$

[11] E-stability ensures that estimated coefficients of the learners converge to their respective fixed points

and when holding n fixed, the coefficients of the adaptive PLM converge to their respective fixed points (for a detailed derivation see Appendix B)

$$vec(\bar{\boldsymbol{a}}') = (\boldsymbol{I} - \boldsymbol{A})^{-1} vec\left((\boldsymbol{C}\bar{\boldsymbol{z}})'\right) \tag{33}$$

$$vec(\bar{\boldsymbol{b}}') = (\boldsymbol{I} - \mathbf{G}_1)^{-1} \mathbf{G}_2 vec(\mathbf{B}') \tag{34}$$

where

$$\begin{aligned}\mathbf{G}_1 = &\ \mathbf{A}(1-n) \otimes \mathbf{F} + \mathbf{A}n \otimes \boldsymbol{\Sigma}_{\mathbf{w}}{}^{-1}\mathbf{A}(1-n) \otimes \boldsymbol{\Sigma}_{\mathbf{w}} \\ &+ \mathbf{A}n \otimes \boldsymbol{\Sigma}_{\mathbf{w}}{}^{-1}\mathbf{A}n \otimes \mathbf{I}\left[\mathbf{I} - n\mathbf{A} \otimes \mathbf{F}\right]^{-1}\left[\mathbf{A}(1-n) \otimes \mathbf{F}\boldsymbol{\Sigma}_{\mathbf{w}}\right]\end{aligned}$$

$$\begin{aligned}\mathbf{G}_2 = &\ \mathbf{I} \otimes \mathbf{F} + \mathbf{A}n \otimes \boldsymbol{\Sigma}_{\mathbf{w}}{}^{-1}\mathbf{I} \otimes \boldsymbol{\Sigma}_{\mathbf{w}} \\ &+ \mathbf{A}n \otimes \boldsymbol{\Sigma}_{\mathbf{w}}{}^{-1}\mathbf{A}n \otimes \mathbf{I}\left[I - \mathbf{A}n \otimes \mathbf{F}\right]^{-1}\mathbf{F}\boldsymbol{\Sigma}_{\mathbf{w}} \otimes \mathbf{I}\end{aligned}$$

Holding n fixed, the mapping T from the perceived to the actual law of motion can therefore be characterised as

$$T(\boldsymbol{\Psi}', n) = \begin{pmatrix} T(\mathbf{a}', n) \\ T(\mathbf{b}', n) \end{pmatrix} = \begin{pmatrix} ((\boldsymbol{I} - \boldsymbol{A}n)^{-1}\boldsymbol{A}(1-n))' \\ \mathbf{G}_1 \end{pmatrix} \tag{35}$$

This result is summarised in the following proposition:

Proposition 1. *The mapping (35) from the PLM to the ALM has a unique fixed point* $\forall n \in [0,1]$ *if the Taylor principle is satisfied by the central bank, i.e.*

$$\phi_\pi > 1$$

Proof. See Appendix C.1. □

The fixed point $\bar{\mathbf{a}}$ is independent of n and equivalent to the steady state under rational expectations:

$$\bar{\mathbf{a}} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{C}\bar{\mathbf{z}} = \begin{pmatrix} 0 & \frac{1-\beta}{\kappa} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \bar{y} \\ \bar{\pi} \end{pmatrix} \tag{36}$$

On the other hand, $\bar{\mathbf{b}}$ is a non-linear function in the share of naive forecasters. Intuitively, there are two offsetting effects that determine that magnitude of $\bar{\mathbf{b}}$: As n decreases $\mathbf{w}_t$ has a stronger effect on $\mathbf{z}_t$ because the average forecast of $\mathbf{z}_{t+1}$ loads stronger on $\mathbf{w}_t$. However, as n decreases, the correlation between $\mathbf{z}_{t+1}$ and $\mathbf{z}_t$ decreases so that $\mathbf{w}_t$ has a smaller effect on $\mathbf{z}_{t+1}$. Now, if n_t is allowed to vary over time, only $n_t = 0.5$ is consistent with a steady state in which both heuristics do not make any forecast error. Therefore, this fixed point implies a unique steady state under heterogeneous expectations:

Proposition 2. *The dynamic system (29) has a unique steady state with* $\pi^* = \bar{\pi}$, $y^* = \frac{1-\beta}{\kappa}\bar{\pi}$, $r^* = \bar{\pi}$ *and* $n^* = 0.5$

Proof. See Appendix C.2 □

The question remains whether this steady state is E-stable under learning, even if all eigenvalues of $(n\mathbf{A})$ lie within the unit circle. According to the E-stability principle (Evans and Honkapohja, 2001, Chapter 13), convergence under learning requires that all eigenvalues of

$$DT_{\mathbf{\Psi}}(\mathbf{\Psi}', n)) - \mathbf{I} \tag{37}$$

need to have negative real parts to ensure that estimated coefficients of the learners converge to their respective fixed points. Holding n^* fixed at the steady state, the stationarity and E-stability requirements are always satisfied for the steady state away from the ZLB:

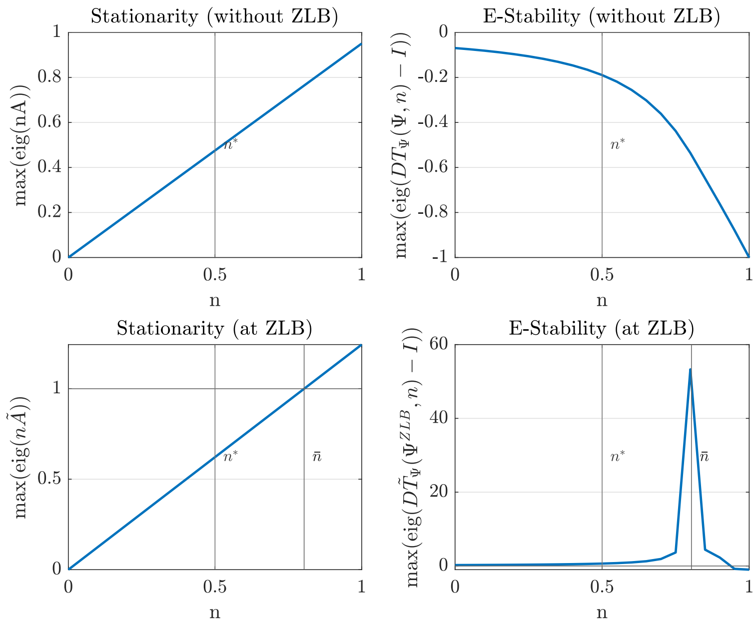
Proposition 3. *The steady state away from the zero lower bound is asymptotically stationary and E-stable*

Proof. See Appendix C.3. □

However, convergence to the steady state may not occur because the true model features a time-varying share of naive forecasters. Only asymptotically a so-called restricted perception equilibrium (RPE, see Evans and Honkapohja, 2001; Berardi, 2009) can be attained. In this steady state expectations might not be rational, but they are optimal given the underparameterised information set that the adaptive learners are endowed with.[12] Away from the steady state the forecast switching can create, or at least amplify, booms and busts. In particular, if the adaptive learners underestimate inflation over several quarters, the resulting deterioration in forecasting performance will increase the share of the naive forecasters. The resulting drift of inflation and output leads to further coordination away from the minimum state variable forecast. This can lead to self-reinforcing cycles that amplify the drift and poses additional challenges to monetary policy not present in the standard rational expectations framework.

[12] That is, the forecast errors are orthogonal to the variables in the information set. However, the forecasting function is still biased because it is based on an underparameterised information set. If agents were able to step out of the model they would be able to detect this but, given the model they are endowed with, this is the best they can do.

Figure 3: Stability Conditions



Note: This Figure shows the stability properties of the model under Inflation Targeting as a function of the share of naive forecasters without and with the ZLB. The maximum eigenvalue needs be a) smaller than 1 in absolute value for stationarity (left panel); b) negative for E-stability (right panel).

4.1.2 At the ZLB

Re-introducing the ZLB into the system, the ALM at the ZLB becomes

$$\begin{aligned} y_t &= (1-n_t)[a_y + b_{y,a}a_t + b_{y,\xi}\xi_t + b_{y,m}m_t] + n_t y_{t-1} \\ &\quad - \frac{1}{\sigma}(0 - (1-n_t)[a_\pi + b_{\pi,a}a_t + b_{\pi,\xi}\xi_t + b_{\pi,m}m_t] - n_t y_{t-1} + \xi_t) \\ \pi_t &= \beta[(1-n_t)[a_\pi + b_{\pi,a}a_t + b_{\pi,\xi}\xi_t] + n_t\pi_{t-1}] + \kappa(y_t - \vartheta a_t) \end{aligned} \tag{38}$$

or in matrix form

$$\Leftrightarrow \mathbf{z}_t = \tilde{\mathbf{A}}\left[(1-n_t)(\mathbf{a} + \mathbf{b}\mathbf{w}_t) + n_t\mathbf{z}_{t-1}\right] + \tilde{\mathbf{B}}\mathbf{w}_t + \tilde{\mathbf{C}}\bar{\mathbf{z}} \tag{39}$$

Following the same approach as before, it can be shown that the ALM has a fixed point at

$$vec(\tilde{\bar{\mathbf{a}}}) = (\boldsymbol{I} - \tilde{\boldsymbol{A}})^{-1} vec\left((\tilde{\boldsymbol{C}}\bar{\boldsymbol{z}})'\right) \tag{40}$$

$$vec(\tilde{\bar{\mathbf{b}}}') = \left(\boldsymbol{I} - \tilde{\mathbf{G}}_1\right)^{-1} \tilde{\mathbf{G}}_2 vec(\tilde{\mathbf{B}}') \tag{41}$$

where

$$\tilde{\mathbf{G}}_1 = \tilde{\mathbf{A}}(1-n) \otimes \mathbf{F} + \tilde{\mathbf{A}}n \otimes \mathbf{\Sigma_w}^{-1} \tilde{\mathbf{A}}(1-n) \otimes \mathbf{\Sigma_w}$$
$$+ \tilde{\mathbf{A}}n \otimes \mathbf{\Sigma_w}^{-1} \tilde{\mathbf{A}}n \otimes \mathbf{I} \left[\mathbf{I} - n\tilde{\mathbf{A}} \otimes \mathbf{F}\right]^{-1} \left[\tilde{\mathbf{A}}(1-n) \otimes \mathbf{F}\mathbf{\Sigma_w}\right]$$

$$\tilde{\mathbf{G}}_2 = \mathbf{I} \otimes \mathbf{F} + \tilde{\mathbf{A}}n \otimes \mathbf{\Sigma_w}^{-1} \mathbf{I} \otimes \mathbf{\Sigma_w}$$
$$+ \tilde{\mathbf{A}}n \otimes \mathbf{\Sigma_w}^{-1} \tilde{\mathbf{A}}n \otimes \mathbf{I} \left[I - \tilde{\mathbf{A}}n \otimes \mathbf{F}\right]^{-1} \mathbf{F}\mathbf{\Sigma_w} \otimes \mathbf{I}$$

Holding n fixed, the mapping $\tilde{T}$ from the perceived to the actual law of motion for the PLM coefficients at the ZLB is now given by

$$\tilde{T}(\mathbf{\Psi}', n) = \begin{pmatrix} \tilde{T}(\mathbf{a}', n) \\ \tilde{T}(\mathbf{b}', n) \end{pmatrix} = \begin{pmatrix} \left((\boldsymbol{I} - \tilde{\boldsymbol{A}}n)^{-1} \tilde{\boldsymbol{A}}(1-n)\right)' \\ \tilde{\mathbf{G}}_1 \end{pmatrix} \tag{42}$$

However, the this mapping has a fixed point only for certain values of n:

Proposition 4. *The mapping (42) from the PLM to the ALM has a unique fixed point for all* $n \in [0, \bar{n})$

Proof. See Appendix C.4 □

If $\bar{n} > 0.5$, the fixed point $\bar{\tilde{\mathbf{a}}}$ exists. The fixed point is independent of n and equivalent to the steady state under rational expectations:

$$\bar{\tilde{\mathbf{a}}} = (\mathbf{I} - \tilde{\mathbf{A}})^{-1} \tilde{\mathbf{C}} \bar{\mathbf{z}} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{y} \\ \bar{\pi} \end{pmatrix} \tag{43}$$

Again, if n_t is allowed to vary over time, only $n_t = 0.5$ is consistent with a steady state in which neither heuristics makes any forecast error, so this fixed point implies a second steady state:

Proposition 5. *If* $\bar{n} > 0.5$, *there exists a second steady state with* $\tilde{\pi}^* = 0$, $\tilde{y}^* = 0$, $\tilde{r}^* = 0$ *and* $n^* = 0.5$

Proof. See Appendix C.5 □

However, even if the ZLB steady state exists, it is not stable when holding n^* fixed at the steady state:

Proposition 6. *The zero lower bound steady state is not E-stable.*

Proof. See Appendix C.6 □

In general, at the ZLB both conditions (asymptotic stationarity and E-stability) never hold at the same time. This is the case because the eigenvalues of $D\tilde{T}_{\mathbf{\Psi}}(\mathbf{\Psi}', n) - \mathbf{I}$ have negative real parts only if the eigenvalues of $n\tilde{\mathbf{A}}$ are outside the unit circle, as Figure 3 shows.

To summarise, under Inflation Targeting the model has two steady states: One steady state away from the ZLB and one at the ZLB, but only the former is stable. These steady states are identical to the ones under rational expectations. However, the model features complex dynamics away from the steady state: Whenever the share of naive forecasters changes, the relationship between exogenous and endogenous realisations changes so that the learners have to update their beliefs. This in turn causes the share of naive forecasters to change again, creating fluctuations around the steady state.

4.2 Dynamics under Price Level Targeting

Under Price Level Targeting the fixed points $\bar{\mathbf{a}}, \bar{\mathbf{b}}, \bar{\mathbf{c}}$ of the adaptive forecasting heuristic are all (non-linearly) dependent of each other so that no closed form solution can be derived. To characterise the solution at least computationally, I keep the assumption of $\phi_y = 0$ and furthermore assume that the shocks $\mathbf{w}_t$ are iid. Adaptive learners therefore only forecast the mean values of the state variables, i.e. they engage in steady state learning. This does not change the stability properties of the system but simplifies the analysis. As before, I first consider the steady state in an environment without the zero lower bound before reintroducing it.

4.2.1 Away from the ZLB

Substituting the monetary policy rule (18) as well as the perceived laws of motion (23) & (20) into the system of equations (14) and rearranging yields the ALM

$$\begin{aligned}
y_t &= (1-n_t)[a_y + b_{y,a}a_t + b_{y,\xi}\xi_t + b_{y,m}m_t + c_{y,\hat{p}}\hat{p}_{t-1}] + n_t y_{t-1} - \frac{1}{\sigma}(\bar{\pi} + \phi_\pi(\pi_t - \bar{\pi} + \hat{p}_{t-1}) \\
&\quad + m_t - (1-n_t)[a_\pi + b_{\pi,a}a_t + b_{\pi,\xi}\xi_t + b_{\pi,m}m_t + c_{\pi,\hat{p}}\hat{p}_{t-1}] - n_t y_{t-1} + \xi_t) \\
\pi_t &= \beta[(1-n_t)[a_\pi + b_{\pi,a}a_t + b_{\pi,\xi}\xi_t + b_{\pi,m}m_t + c_{\pi,\hat{p}}\hat{p}_{t-1}] + n_t\pi_{t-1}] + \kappa(y_t - \vartheta a_t) \\
\hat{p}_t &= \pi_t - \bar{\pi} + \hat{p}_{t-1}
\end{aligned} \tag{44}$$

Or in matrix notation:

$$\boldsymbol{z}_t = \boldsymbol{A}[(1-n_t)(\boldsymbol{a}+\boldsymbol{b}\boldsymbol{w}_t+\mathbf{c}\mathbf{z}_{t-1})+n_t\boldsymbol{z}_{t-1}]+\boldsymbol{D}\boldsymbol{z}_{t-1}+\boldsymbol{B}\boldsymbol{w}_t+\boldsymbol{C}\bar{\boldsymbol{z}} \tag{45}$$

where $\mathbf{c} = [\mathbf{0}_{3\times 2}\ \tilde{\mathbf{c}}]$. Substituting this into the ODE associated with the updating equations (22) yields[13]

$$\frac{\partial \boldsymbol{\Psi}'}{\partial \tau} = \mathbf{R}^{-1}E\mathbf{x}_{t-2}\left(\mathbf{A}\left[(1-n)(\mathbf{a}+\mathbf{b}\mathbf{w}_t+\mathbf{c}\mathbf{z}_{t-2})\right]+\mathbf{A}n\mathbf{z}_{t-2}+\mathbf{B}\mathbf{w}_t+\mathbf{C}\bar{\mathbf{z}})-\boldsymbol{\Psi}\mathbf{x}_{t-2}\right)' \tag{46}$$

To characterise the fixed points of the adaptive forecasting heuristic, I start with the fixed point of $\mathbf{c}$ since this determines the other two fixed points. The associated mapping

$$\frac{\partial \mathbf{c}'}{\partial \tau} = \boldsymbol{\Sigma}_p^{-1}\boldsymbol{\Sigma}_{p_{t-2},p_{t-1}}^{-1}(\mathbf{A}(1-n)\mathbf{c}_{:,3}+\mathbf{D}_{:,3}+n\mathbf{A}_{:,3})+\boldsymbol{\Sigma}_p^{-1}\boldsymbol{\Sigma}_{p_{t-2},(y_{t-1},\pi_{t-1})}(\mathbf{D}_{1:2}+n\mathbf{A}_{1:2}) \tag{47}$$

can be solved numerically for $\bar{\mathbf{c}}$. Having pinned down $\bar{\mathbf{c}}$, we can write (holding $(\boldsymbol{\Psi}, n)$ fixed) $\lim_{t\to\infty} E\mathbf{z_t}$ as

$$\mathbf{z}_t(\boldsymbol{\Psi}, n) = \left((\mathbf{I}-\mathbf{A}(1-n)\mathbf{c}-n\mathbf{A}-\mathbf{D})^{-1}\left(\mathbf{A}\left[(1-n)\mathbf{a}\right]+\mathbf{C}\bar{\mathbf{z}}\right)\right)' \tag{48}$$

as long as the ALM in Equation (45) is asymptotically stationary, i.e.

$$\lambda < 1 \quad \forall \lambda \in \Lambda = \{\Lambda : |\mathbf{I}-\mathbf{A}(1-n)\mathbf{c}-n\mathbf{A}-\mathbf{D}-\boldsymbol{\Lambda}\mathbf{I}| = 0\} \tag{49}$$

The fixed point $\bar{\mathbf{a}}$ then becomes

$$vec(\bar{\boldsymbol{a}}') = \left[(\boldsymbol{I}+(\boldsymbol{A}(1-n)\bar{\mathbf{c}}+n\mathbf{A}+\mathbf{D})(\boldsymbol{I}-(\boldsymbol{A}(1-n)\bar{\mathbf{c}}-n\mathbf{A}-\mathbf{D})^{-1}-\mathbf{A}(1-n)\right]^{-1} vec\left((\boldsymbol{C}\bar{\boldsymbol{z}})'\right) \tag{50}$$

Finally, the fixed point $\bar{\mathbf{b}}$ is then given by

$$vec(\bar{\boldsymbol{b}}') = [\boldsymbol{I}-(\mathbf{A}(1-n)\bar{\mathbf{c}}+\mathbf{D}+n\mathbf{A})\otimes\mathbf{I}\,\mathbf{A}(1-n)\otimes\mathbf{I}]^{-1}(\mathbf{A}(1-n)\bar{\mathbf{c}}+\mathbf{D}+n\mathbf{A})\otimes\mathbf{I}\,vec(\mathbf{B})' \tag{51}$$

Holding n fixed, the mapping $\tilde{T}$ from the perceived to the actual law of motion for the PLM coefficients at the ZLB is now given by

$$\mathrm{T}(\Psi', n) = \begin{pmatrix} ((\mathbf{I}+(\mathbf{A}(1-n)\bar{\mathbf{c}}+n*\mathbf{A}+\mathbf{D})(\mathbf{I}-\mathbf{A}(1-n)\bar{\mathbf{c}}+n*\mathbf{A}+\mathbf{D})^{-1})\mathbf{A}(1-n))' \\ ((\mathbf{A}(1-n)\bar{\mathbf{c}}+\mathbf{D}+n\mathbf{A})\otimes\mathbf{I}\,\mathbf{A}(1-n)\otimes\mathbf{I})' \\ \boldsymbol{\Sigma}_p^{-1}\boldsymbol{\Sigma}_{p_{t-2},p_{t-1}}^{-1}(\mathbf{A}(1-n)\mathbf{c}_{:,3}+\mathbf{D}_{:,3}+n\mathbf{A}_{:,3})'+\boldsymbol{\Sigma}_p^{-1}\boldsymbol{\Sigma}_{p_{t-2},(y_{t-1},\pi_{t-1})}(\mathbf{D}_{:,1:2}+n\mathbf{A}_{:,1:2})' \end{pmatrix} \tag{52}$$

[13]where $\mathbf{c}_{:,3}$ denotes the third column of $\mathbf{c}$, i.e. $\mathbf{c}_{:,3} = \tilde{\mathbf{c}}$.

Proposition 7. *The mapping (52) from the PLM to the ALM has a unique fixed point for all* $n \in [0, \bar{n})$

Proof. See Appendix C.7 □

Figure 4 shows, that under the estimated parameters (see next Section) $\bar{n} \approx 0.95$, so that for $n = 0.5$ a fixed point indeed exists. Furthermore, the fixed point $\bar{\mathbf{a}}$ itself is independent of n and equivalent to the steady state under rational expectations:

$$\bar{\mathbf{a}} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{C}\bar{\mathbf{z}} = \begin{pmatrix} 0 & \frac{1-\beta}{\kappa} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{y} \\ \bar{\pi} \\ \bar{p} \end{pmatrix} \tag{53}$$

On the other hand, $\bar{\mathbf{b}}$ and $\bar{\mathbf{c}}$ are non-linear functions of the share of naive forecasters. Again, if n_t is allowed to vary over time, only $n_t = 0.5$ is consistent with a steady state in which neither heuristics makes any forecast error, so this fixed point implies a second steady state

Proposition 8. *If* $\bar{n} > 0.5$*, the dynamic system (55) has a unique steady state with* $\pi^* = \bar{\pi}$*,* $y^* = \frac{1-\beta}{\kappa}\bar{\pi}$*,* $r^* = \bar{\pi}$*,* $\hat{p}^* = 0$ *and* $n^* = 0.5$

Proof. See Appendix C.8. □

E-stability then requires that the eigenvalues of

$$DT_{\mathbf{\Psi}}(\mathbf{\Psi}', n)) - \mathbf{I} \tag{54}$$

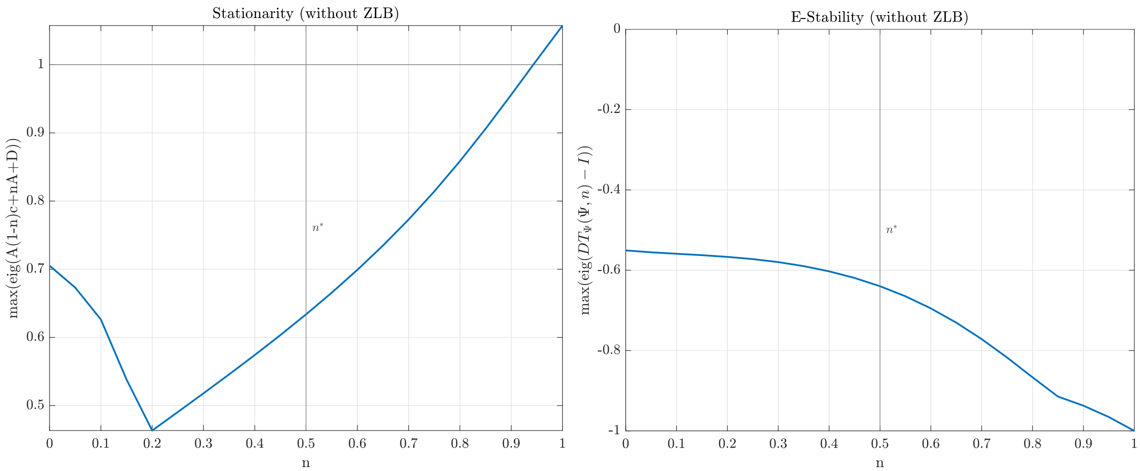
need to have negative real parts to ensure that the perceived law of motion converges to its fixed point and the steady state is learnable. Figure 4 shows that this is indeed the case under the estimated parameters (see next Section) for $n = 0.5$.

4.2.2 At the ZLB

At the ZLB the ALM becomes

$$\begin{aligned} y_t &= (1 - n_t)[a_y + b_{y,a}a_t + b_{y,\xi}\xi_t + b_{y,m}m_t + c_{y,\hat{p}}\hat{p}_{t-1}] + n_t y_{t-1} \\ &\quad - \frac{1}{\sigma}(0 - (1 - n_t)[a_\pi + b_{\pi,a}a_t + b_{\pi,\xi}\xi_t + b_{\pi,m}m_t + c_{\pi,\hat{p}}\hat{p}_{t-1}] - n_t y_{t-1} + \xi_t) \\ \pi_t &= \beta[(1 - n_t)[a_\pi + b_{\pi,a}a_t + b_{\pi,\xi}\xi_t + b_{\pi,m}m_t + c_{\pi,\hat{p}}\hat{p}_{t-1}] + n_t\pi_{t-1}] + \kappa(y_t - \vartheta a_t) \\ \hat{p}_t &= \pi_t - \bar{\pi} + \hat{p}_{t-1} \end{aligned} \tag{55}$$

Figure 4: Stability Conditions under Price Level Targeting



Note: This Figure shows the stability properties of the model under Price Level Targeting as a function of the share of naive forecasters. The maximum eigenvalue needs be a) smaller than 1 in absolute value for stationarity (left panel); b) negative for E-stability (right panel).

or in matrix form

$$\boldsymbol{z}_t = \tilde{\boldsymbol{A}} \left[(1 - n_t)(\boldsymbol{a} + \boldsymbol{b}\boldsymbol{w}_t + \mathbf{c}\mathbf{z}_{t-1}) + n_t \boldsymbol{z}_{t-1} \right] + \tilde{\boldsymbol{D}} \boldsymbol{z}_{t-1} + \tilde{\boldsymbol{B}} \boldsymbol{w}_t + \tilde{\boldsymbol{C}} \bar{\boldsymbol{z}} \tag{56}$$

In this case no equilibrium can exist. Note that the price level gap at the zero lower bound follows a unit root because no deviation of inflation from target can be offset by the central bank.

Proposition 9. *Under Price Level Targeting no steady state exists at the zero lower bound*

Proof. See Appendix C.9. □

To summarise, under Price Level Targeting the model features only the steady state away from the zero lower bound. However, this steady state is not stable for all $n \in [0, 1]$ but only for $n \in [0, \bar{n})$. The question remains whether this threshold will ever be crossed and Price Level Targeting thereby can lead to unstable dynamics. I will address this using simulation methods in Section 6.

5 Estimation and Model Dynamics

5.1 Estimation

To assign numerical values to the model parameters, in particular to the non standard parameters φ, γ, θ, I estimate the model using Bayesian methods. Besides estimating the relevant parameters, this also allows me to extract an estimate of the unobserved degree of de-anchoring over time. Furthermore, Bayesian estimation provides a coherent framework to compare the data fit of the heuristic switching model with the standard rational expectations model. Due to the non-linear elements of the heuristic switching model, I compute the model likelihood $\ln p(Y \mid \Omega_{HSM})$ of observing the data Y given the parameter vector Ω_{HSM} using the bootstrap particle filter with 20,000 particles. The likelihood of the rational expectations model on the other hand can be computed using the Kalman filter because the model is fully linear.

The sample period runs from 1982Q1 to 2007Q4. I exclude the zero lower bound period because this would add yet another source of non-linearity to an already computationally complex estimation process. I estimate the model using observable data on quarter-on-quarter real GDP growth, the quarter-on-quarter growth of the implied GDP price deflator, the Federal Funds Rate (adjusted to quarterly frequency), median 1Q ahead real GDP growth expectations (SPF), and the median 1Q ahead expectations of the change of the implied GDP price deflator (SPF). As in Carvalho et al. (2020), I do not use observed long-term expectations in the estimation due to limited data before 1990. Instead, I will back out the model-implied long-term expectations in the next Section and compare it to the available data. The measurement equation, therefore, is

$$\mathbf{Y}_t = \mathbf{\Gamma}_0 + \mathbf{\Gamma}_1 \mathbf{Z}_t + \mathbf{H}\epsilon_t \tag{57}$$

where the vector $\mathbf{Y}_t$ includes the five observable data series; $\mathbf{\Gamma}_0$ contains the means of real GDP growth, the quarterly Federal Funds Rate, the expected real GDP growth, and zeros for (expected) inflation; the observation matrix $\mathbf{\Gamma}_1$ selects the corresponding variables from the state vector $\mathbf{Z}_t$. Since the particle filter requires some measurement error ϵ_t to avoid degeneracy, the variance $\mathbf{H}'\mathbf{H}$ of the measurement error is set to 50% of variance of $\mathbf{Y}_t$.[14]

To estimate the posterior distribution of the rational expectations model, I first maximise the posterior likelihood and take the resulting posterior mode Ω_{RE} as starting point for the Metropolis-Hastings algorithm. To estimate the posterior distribution of the heuristic switching model I follow a two step procedure: I first use the linear Kalman filter to maximise the posterior of the model and take the resulting posterior mode as starting point for the Metropolis-Hastings algorithm (using the identity matrix as proposal distribution). In a second step I use the results as starting points for the estimation that uses the particle filter to compute the model likelihood. In each case, I use the Metropolis-Hastings algorithm with 10,000 draws and 4 separate chains to compute the full distribution.[15] I scale the proposal

[14] For comparability, I use the same assumption when estimating the RE model using the Kalman Filter.

[15] The particle filter generates random numbers at several steps of the computation so I reset the seed of

distribution to target an overall acceptance rate of roughly 30%. Finally, I discard the first 20% of draws as burn-in, keep every 5th remaining draw, and average over the four chains to compute the relevant moments.

Table 1: Prior Distributions and Posterior Estimates

Parameter		Prior			Posterior (RE)		Posterior (HSM)	
		Prior Dist.	Mean	Std.	Mean	Std.	Mean	Std.
Discount factor	β	B	0.99	0.005	0.992	0.004	0.991	0.004
Inv. IES	σ	$\mathcal{N}$	2	0.75	2.058	0.099	2.531	0.212
Inv. Frisch Elast.	$\frac{1}{\eta}$	$\mathcal{N}$	3	0.3	3.000	0.032	2.928	0.034
Elast. Demand	$\frac{\epsilon}{\epsilon-1}$	Γ	1.35	0.2	1.431	0.141	1.521	0.250
Rotemberg Adj. Cost	ϕ	$\mathcal{U}(0,1000)$	-	-	51.865	0.234	121.301	0.548
AR(1) Demand Shock	ρ_ζ	B	0.5	0.2	0.882	0.032	0.854	0.062
AR(1) TFP shock	ρ_a	B	0.5	0.2	0.631	0.046	0.511	0.124
AR(1) MP shock	ρ_m	B	0.5	0.2	0.642	0.097	0.910	0.039
Std. Demand Shock	σ_ζ	Γ^{-1}	1	0.5	0.283	0.027	0.446	0.058
Std. TFP Shock	σ_a	Γ^{-1}	1	0.5	0.408	0.054	0.559	0.094
Std. MP Shock	σ_m	Γ^{-1}	1	0.5	0.389	0.053	0.374	0.047
Feedback Output	ϕ_y	$\mathcal{N}$	0.25	0.125	0.422	0.097	0.329	0.108
Feedback Inflation	ϕ_π	$\mathcal{N}$	1.5	0.5	2.807	0.146	1.937	0.141
Inflation Target	$\bar{\pi}$	B	0.5	0.1	0.854	0.036	0.571	0.093
Constant Gain	γ	B	0.03	0.01	-	-	0.034	0.005
Switching Intensity	θ	$\mathcal{U}(0,1000)$	-	-	-	-	78.586	0.194
Initial De-Anchoring	n_0	$\mathcal{N}$	0.75	0.125	-	-	0.694	0.099
NKPC slope (implied)	κ	-	0.059	-	0.241	-	0.125	-
Posterior Likelihood	$\ln p(Y \mid \Omega^*)$	-	-	-	-122.631		-92.053	
Marginal Likelihood ($\tau = 0.5$)	$\ln p(Y)$	-	-	-	-59.910		-33.388	

Note: The marginal likelihood is estimated using Geweke's Harmonic Mean Estimator with a tuning parameter of $\tau = 0.5$. The estimates do not change significantly for different values of τ.

The independent prior distributions are described in Table 1. The discount factor β follows a relatively tight Beta distribution. The intertemporal elasticity of substitution and the Frisch elasticity are centred around 2 & 3, respectively. The elasticity of demand is centred around an equilibrium markup of 35%, approximately the midpoint of the aggregate markups estimated by De Loecker et al. (2020). All the autoregressive parameters follow a Beta distributions centred around 0.5, with the variances following an inverse Gamma distribution

the random number generator every time the likelihood is evaluated to avoid injecting randomness in the calculation of the likelihood (see e.g. Fernández-Villaverde and Rubio-Ramírez, 2007; Carvalho et al., 2020). I thereby fix the random initial conditions for the nonlinear state variables; the random draws to compute shocks in the nonlinear prediction step; and random draws in the resampling step so that the particle filter uses the same particles for every new parameter evaluation. The MCMC sampler draws from a different random number generator so that resetting the seed does not affect the parameter draws of the sampler.

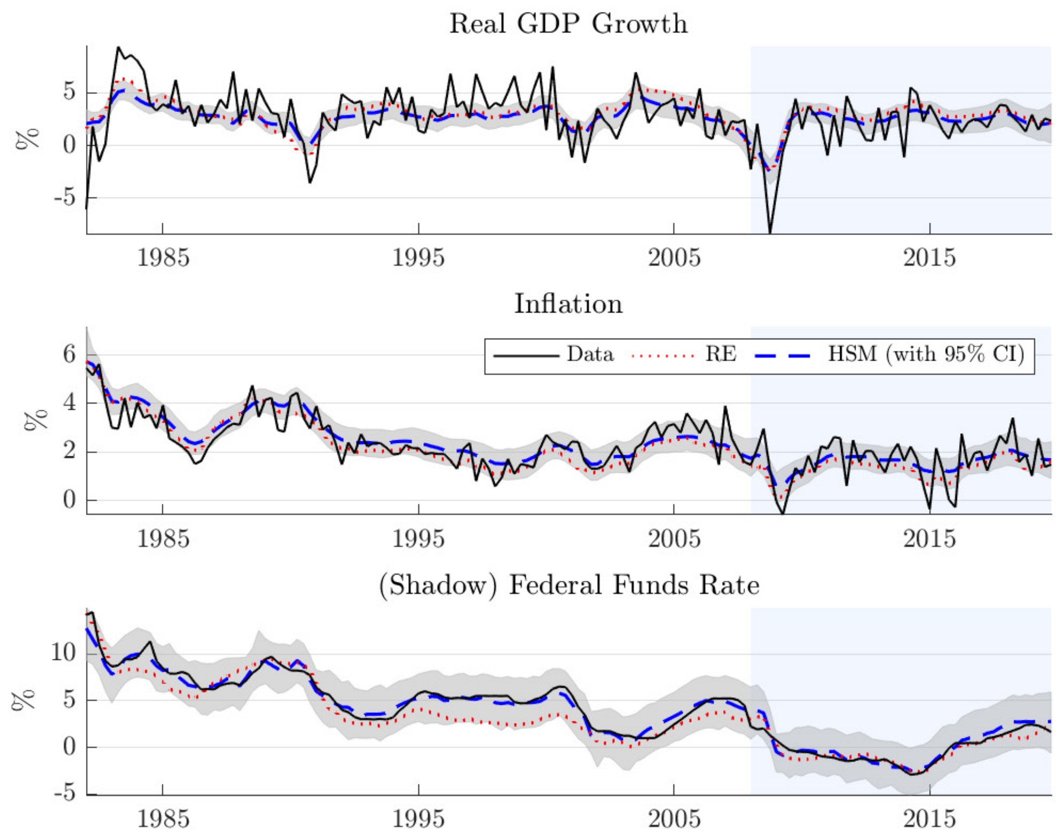
centred around 1. The coefficients of the Taylor rule are centred around the standard values used in the literature and the prior inflation target is set to an annualised value of 2%. For the constant gain parameter the literature on adaptive learning suggests a range of 0.01 (see Milani, 2007) to 0.06 (see Branch and Evans, 2006). I pick an intermediate value of 0.03 as prior for γ, which follows a Beta distribution with standard deviation 0.01. For the non-standard parameters φ, θ I impose a very loose prior in the form of a uniform distribution reflecting the lack of prior knowledge about these parameters. Both are initialised at a value of 100. I calibrate the steady-state labour to 1/3 of the available time so that the composite parameters such as κ or χ (omitted) are pinned down by the respective underlying prior (posterior) values. I also estimate the initial share of naive forecasters n_0 and build on the estimation results of Section 2: Given that inflation forecasts were not perfectly anchored during the 1980s, I assume that n_0 lies above its steady state value and assume a normal distributed prior centred at 0.75. This initial share in turn pins down the initial values for $\mathbf{\Psi}$ and yields values for the initial forecast errors e_0^i.

Table 1 also presents the posterior mean and standard deviations for both models. Relative to the RE model, the HSM suggests a much higher real interest rate over the sample period (3.7% vs 3.2%). This is offset by a much lower implied inflation target, so that the HSM suggests a lower equilibrium nominal interest rate (6% vs 6.7%). The HSM model suggests a lower intertemporal elasticity of substitution (0.40 vs 0.49) but yields very similar estimates of the inverse Frisch elasticity. The estimates of the autoregressive components differ a lot, with the HSM suggesting a (slightly) lower degree of autocorrelation for the discount factor and productivity shock processes, but a much higher degree of autocorrelation for the monetary policy shock process. This is somewhat surprising since previous literature (e.g. Milani, 2007) has observed that adding adaptive learning to a standard NK model diminishes the need for mechanical sources of persistence. It appears that forecast switching still requires mechanical sources of persistence to match the data, although along different dimensions than under RE. The estimated standard deviations of the exogenous processes for demand & TFP shocks are larger under HSM than under RE, but the opposite is the case of the standard deviations of the monetary policy shocks. The rational expectations model implies a much more aggressive monetary policy reaction. The posterior estimate of $\gamma = 0.034$ implies that firms and households rely on the last $\sim 7^1/_2$ years of data. Finally, the implied NKPC slope estimate under rational expectations is much larger than under forecast switching.

At the posterior mean, the HSM model fits the observed data better than the RE model with a log likelihood of $\ln p(\Omega^*_{HSM} \mid Y) = -92.053$ vs. $\ln p(\Omega^*_{RE} \mid Y) = -122.631$ (where Ω^*_i denotes the posterior mean of the respective model). For an overall measure of the model fit (by marginalising out the influence of the parameters Ω) I compute the marginal Likelihood $\ln p(Y \mid \mathcal{M})$ using the Modified Harmonic Mean Estimator of Geweke (1999). The resulting estimates of the marginal density in the lower panel of Table 1 suggest that the heuristic switching model fits the data much better than the standard rational expectations model.

5.2 Model Predictions

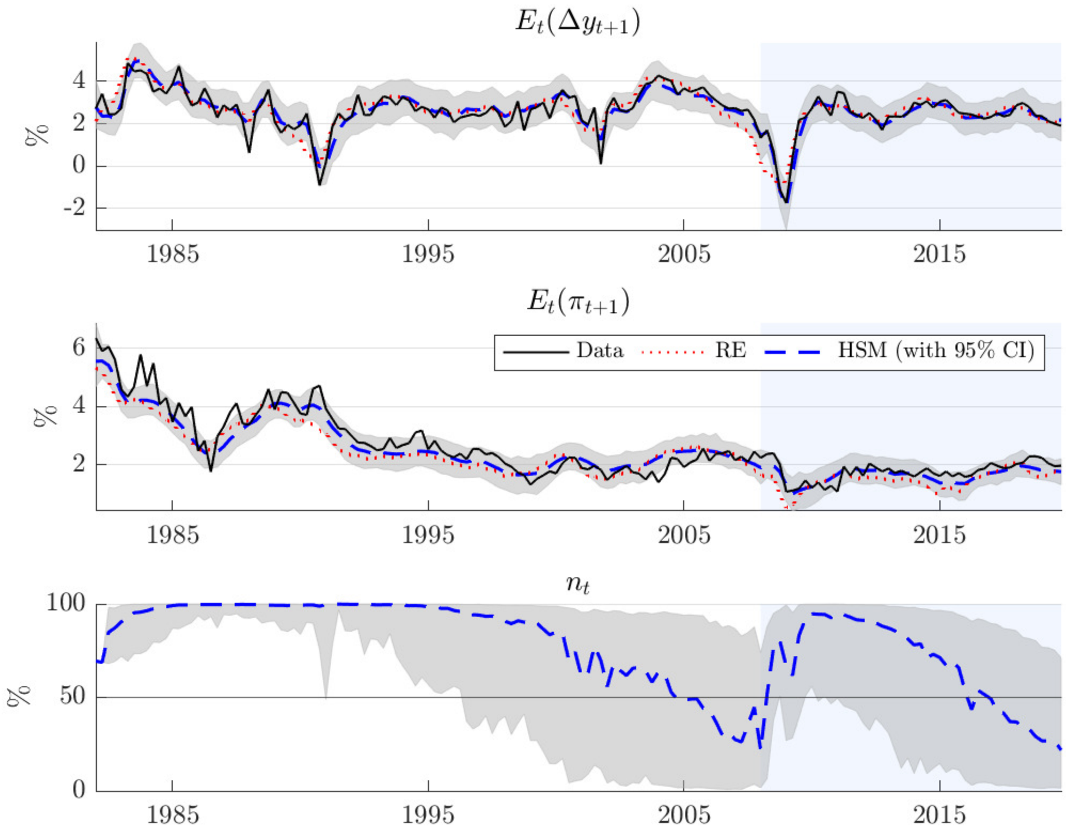
Figure 5: Model Predictions - Real Variables



Note: The panels show model predictions for real GDP growth (top), inflation (middle) and the (shadow) federal fund rates (bottom). The black solid line denotes the data; the blue dashed line depicts the median model prediction of the heuristic switching model along with 95% confidence intervals in grey; the red dotted line depicts the median model prediction under rational expectations (RE). The blue shaded area denotes the out-of-sample forecast period.

I back out the model-implied states using the non-linear smoother of Godsill et al. (2004). As an extension, I compute the out-of-sample model forecasts up until 2019Q4. That is, I estimate the model on the sample from 1982Q1:2007Q4 (where the ZLB is not binding) and use the resulting parameters with an extended set of observables to compute the out-of-sample model predictions until 2019Q4. To circumvent the problem of the ZLB and to

Figure 6: Model Predictions - Expectations



Note: The panels show model predictions for 1q ahead expected real GDP growth (top), 1q ahead expected inflation (middle) and the share of naiv forecasters (bottom). The black solid line denotes the data; the blue dashed line depicts the median model prediction of the heuristic switching model along with 95% confidence intervals in grey; the red dotted line depicts the median model prediction under rational expectations (RE). The blue shaded area denotes the out-of-sample forecast period.

account for the fact that central banks have more tools than just the interest rate, I use the shadow interest rate of Wu and Xia (2016) starting from 2008Q1 instead of the realised Federal Funds Rate. The resulting smoothed state estimates are plotted (in blue) alongside the data (in black) and the RE states (in red) in Figures 5 & 6. Grey shaded areas indicate the 95% confidence intervals of the heuristic switching model and the blue shaded area indicates the out-of-sample forecast period. In general, neither model captures the high frequency fluctuations of the observable data well as they are attributed to measurement error. Instead, both models track the apparent underlying low frequency movements quite well.

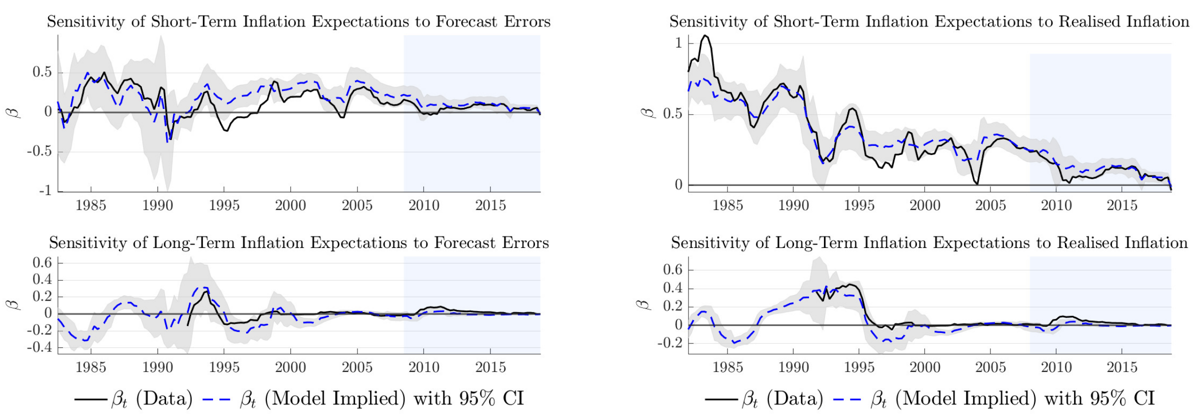
In terms of root mean squared error (RMSE) the heuristic switching model significantly outperforms the rational expectations model when it comes to matching the interest rate and expectations for the whole period between 1982 and 2019 (see Table 2). It performs slightly worse when it comes to matching inflation and output growth. This result does not change when considering only the in-sample years 1982 - 2007. However, in the out-of-sample period 2008-2019, the HSM model actually outperforms the RE model in all series except the shadow interest rate. That is, the heuristic model achieves its purpose of modelling the joint evolution of real variables and expectations, even though the rational expectations might have a small edge in terms of forecasting output and inflation depending on the period. Finally, there does not seem to be any significant difference in model fit between the in- and out-of-sample forecasts.

Table 2: Individual Data Series Fit (Root Mean Squared Error)

	Δy_t	π_t	r_t	$E_t(\Delta y_{t+1})$	$E_t(\pi_{t+1})$
RE (1982-2019)	1.940	0.577	1.374	0.527	0.539
1982-2007	1.995	0.514	1.608	0.543	0.581
2008-2019	1.815	0.694	0.613	0.490	0.433
HSM (1982-2019)	1.947	0.580	0.641	0.373	0.392
1982-2007	2.059	0.527	0.638	0.418	0.443
2008-2019	1.679	0.681	0.649	0.251	0.251

The predicted share of naive forecasters - the gauge of expectations anchoring - lies at (or close to) at 100% and suggests that even professional forecasters adopted a random walk like forecasting function in the 1980s and early 1990s, i.e. a period characterised by a high degree of inflation persistence. In the mid 1990s expectations begin to re-anchor, until the Global Financial Crisis leads to another increase in the share naive forecasters. However, this uptick of de-anchoring is small and short-lived. Again, it is important to note that, due to the presence of the measurement error in the estimation, the model households do not forecast headline inflation directly, but rather the underlying (much smoother) inflation trend. As a result, the correlation between model-implied forecasts and realised lagged headline inflation is less than one and tracks the correlation between observed inflation expectations and the realised lagged headline inflation quite well (see Figure 7):[16] Model-implied short-term forecasts respond significantly to lagged inflation realisations but the estimated coefficient experiences a steady downward trend with a short-lived uptick in 2009. Long-term expectations respond significantly at first but become anchored over time as well, without experiencing a similar uptick in 2009. A similar pattern emerges for the sensitivity to lagged forecast errors. Overall, this suggests that the Great Recession led to some degree of de-anchoring but this process did not feed through to long-term expectations.

Figure 7: Sensitivity of Implied Expectations to Inflation Surprises and Inflation



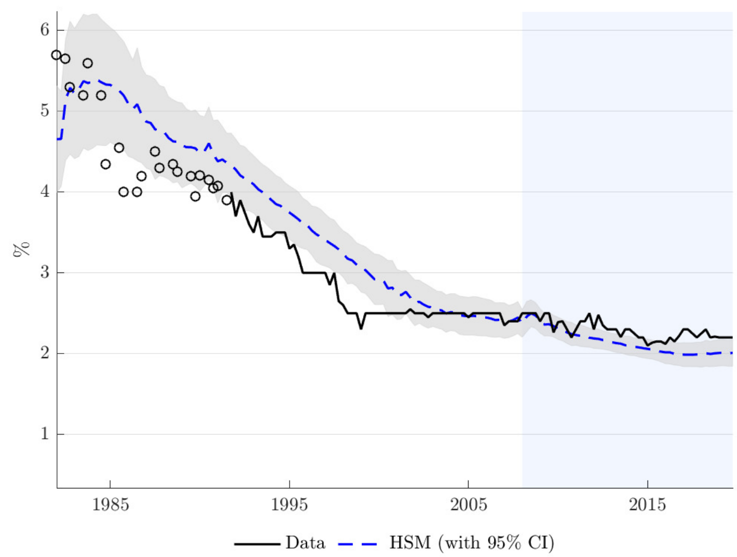
Note: The panels show the estimated time-varying sensitivity of short-run (upper row) and long-run (lower row) inflation expectations to inflation surprises (left column) and realised inflation (right column). Solid black lines depict the time-varying sensitivity of model-implied inflation expectations with 95% credible intervals shaded in grey. Finally, blue dashed lines depict the time-varying sensitivity of median observed inflation expectations from the SPF.

Despite using long-run expectations as observable data in the estimation, I can back out the long-run inflation expectations (see Section 3.3). Figure 8 plots the evolution of these implied long-run inflation expectations together with the observed median 10y ahead CPI inflation expectations obtained from the SPF (the longest continuously available long-run forecast available in the SPF). The model-implied long-run inflation expectations appear to match an underlying trend present in the SPF forecasts. They follow the steady downward

[16] The sensitivity of the smoothed 1Q ahead inflation expectations to the model-implied forecast error as well as lagged, observed, inflation realisations is estimated using a rolling regression model as in Section 2. Differences in the estimated sensitivity of the observed inflation expectations arise from the fact that here I use GDP price inflation and its corresponding forecasts as measure of inflation. Furthermore, I only use median inflation expectations instead of exploiting the whole panel structure to keep the estimates comparable.

trend in the first half of the sample but miss the (almost) discrete drop in the late 1990's. Furthermore, when observed inflation expectations become increasingly volatile again following the Great Recession, the model implies a decrease in long-run expectations which slightly underpredicts observed expectations. Despite these differences in terms of the level forecast, the sensitivity of model-implied inflation forecasts is quite close to the sensitivity present in the SPF forecasts, as discussed earlier. Finally, the model-implied IRF of long-term inflation expectations to a 1pp monetary policy shock matches the IRF estimated on observed expectations rather well (see Figure D.1).

Figure 8: Model-Implied and Observed 10y Inflation Expectations



Note: This Figure shows the model-implied 10y inflation expectations as the dashed blue line with 95% confidence bands shaded in grey. The continuous black line depicts the median 10y ahead CPI forecast from the Survey of Professional Forecasters. The circled data points are obtained by combining the 10y ahead forecasts from the *Livingston Survey* and the *Blue Chip Economic Indicators*, which can be found here. The blue shaded area denotes the out-of-sample forecast period.

The implied moments of the estimated model as a function of the share of naive agents are displayed in Figures D.3 & D.4. The variance of output and inflation is increasing in the degree of de-anchoring, just like the autocorrelation of the two variables. The contemporaneous impact of the exogenous processes on output and inflation is highly non-linear: Output reacts least (most) in an absolute sense to discount factor and monetary policy shocks (productivity shocks) for intermediate values of n. Inflation, on the other hand, reacts most in an absolute sense to each shock for intermediate values of n.

5.3 De-Anchoring in the Model

To illustrate the process of de-anchoring I simulate the response of the estimated model to four consecutive σ_ζ discount factor shocks when the zero lower bound is present (see Figure 9). This series of shocks decreases demand enough to force the central bank to the zero lower bound for a prolonged period and leads to significant drops in output and inflation.[17] As the lower right panel illustrates, the share of naive forecasters initially drops but, instead of reverting back to its steady state value, it quickly increases up to 100%. Upon impact of the initial contractionary shocks, the adaptive learning rule initially predicts the resulting drop in activity better than the naive forecasters as the latter adjust their expectations only with a lag. As the shocks start to level off, however, the adaptive heuristic yields worse forecasts. This is the case because, coming out of steady state, it misjudges the relation between the exogenous processes and the endogenous realisations. That is, because the adaptive forecasting heuristic starts with $\bar{\boldsymbol{\Psi}}_0$ at its steady state value where $n = 0.5$ and where monetary policy can mediate the effect of exogenous shocks, it fails to accurately reflect the new environment where the share of adaptive learners is higher and the ZLB is binding. Therefore, the share of naive forecasters gradually increases and ultimately pushes the economy into a deflationary spiral.

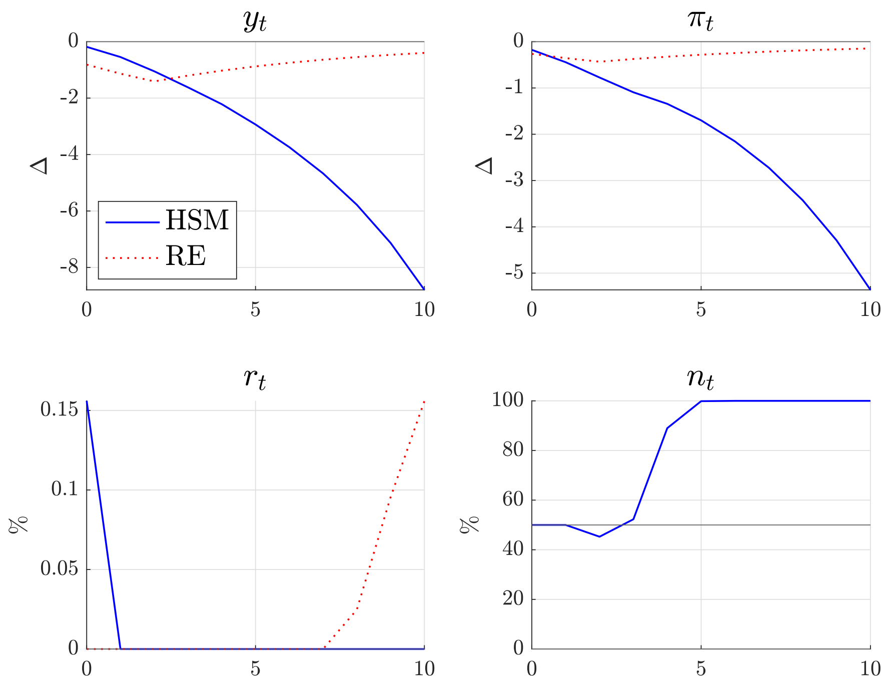
Note that, upon impact, the HSM model yields the same response as the rational expectations model (plotted in red).[18] The key difference is that forecast switching can generate endogenous, belief-driven deflationary spirals at the zero lower bound. The rational expectations model on the other hand returns back to the steady state once the initial shocks disappear and the central bank leaves the zero lower bound after 8 quarters.

The deflationary spiral is driven by the forecast switching behaviour. In fact, without forecast switching (i.e. keeping $n = 0.5$ constant) the economy would very slowly return to its pre-crisis level as the counterfactual analysis in Figure D.5 shows. Note that this scenario occurs only when the central bank is constrained by the zero lower bound. If monetary policy is unconstrained, the system remains stable even if the share of naive forecasters approaches 100%. This is why the estimated process remains stable throughout the 1980s & 1990s. However, even when monetary policy remains unconstrained and is able to prevent deflationary spirals, the economy still experiences elevated volatility of output and inflation. Therefore, the potential downside posed by de-anchoring is highly asymmetric with regard to the proximity to the ZLB, i.e. depends on the level of the equilibrium interest rate.

[17] I plot the IRFs for 10 quarters only. For a longer time span the rational expectations response becomes hard to discern due to the deflationary spiral.

[18] The rational expectations IRF is computed using the parameters of the HSM posterior for better comparability

Figure 9: De-Anchoring at the ZLB



Note: IRF to four consecutive σ_ζ shocks to the discount factor ζ_t. The solid blue line shows the IRFs of the heuristic switching model (HSM), the dotted red line shows the IRFs of the rational expectations model (RE)

6 Policy Comparison

As discussed in the previous Section, sequences of adverse shocks can push the economy to the zero lower bound and lead to deflationary spirals by triggering expectations de-anchoring. How can monetary policy prevent this scenario? And how do the two previously discussed monetary rules compare in terms of their welfare implications? In this section, I address these question with stochastic simulations. Due to the asymmetric risk posed by de-anchoring (i.e. the risk of deflationary spirals at the ZLB), I also consider a version of Inflation Targeting that features an asymmetric response to below-target inflation (Bianchi et al., 2019)

$$r_t = \max\left[0, \bar{\pi} + \mathbf{1}_{\pi_t<\bar{\pi}}\underline{\phi}_\pi(\pi_t - \bar{\pi}) + (1 - \mathbf{1}_{\pi_t<\bar{\pi}})\bar{\phi_\pi}(\pi_t - \bar{\pi}) + \phi_y \hat{y}_t\right]$$

with $\underline{\phi}_\pi = \pi + 0.5$ & $\bar{\phi_\pi} = \pi - 0.5$. I evaluate the performance of the three policy frameworks using the welfare function $\mathcal{W}$

$$\mathcal{W} = -\sum_{t=0}^{T} \beta^t \left[(\pi_t - \bar{\pi})^2 + \frac{\kappa}{\epsilon}\hat{y}_t^2\right]$$

I simulate the model economy using the estimated parameters over 2,000 periods, discarding the first 20% of them. However, instead of simulating the linearised model of Section 4.1.1, I simulate the non-linear model described in Section 3.1. The model is solved with the time iteration approach of Richter et al. (2014), which is a version of policy function iteration. This approach solves for the labour and inflation policy functions given set of subjective expectations, that set the error of the two inter-temporal conditions (5) & (12) to zero. I maintain the assumption that households use linear predictors to forecast the *deviations* of output and inflation from steady state, but translate those expectations back into *levels* of output and inflation to solve the non-linear model. For simplicity, I impose market clearing to derive the consumption forecast for the Euler equation (5). I simulate the economy under two different environments: The first environment features a discount factor β equal to the estimated parameters presented in Table 1 so that, along with an inflation target of 2%, the equilibrium nominal interest rate is $\sim 6\%$. In the second environment I increase the discount rate to $\beta = 0.9975$ so that the equilibrium nominal interest rate decreases to $\sim 3\%$ (annualised). Importantly, I simulate the model assuming that the zero lower bound on interest rates can be binding and do not account for the many other tools central banks have at their disposal at the ZLB such as QE, forward guidance etc. These instruments, however, potentially bear risks for financially stability as well as central bank independence and, finally, might face decreasing returns. Thus, even if central banks are not completely constrained by the ZLB (e.g. Wu and Xia, 2016), it is in the interest of the central banks to optimise its operating procedure in normal times to minimise the frequency and severity of those ZLB episodes.

Table 3 presents the simulation results. In a high nominal interest rate world the risk of a binding zero lower bound is relatively small and the risk of deflationary spirals is (unsurprisingly) basically non existent. Among the three policy rules, Price Level Targeting minimises the volatility of output and inflation. The asymmetric Inflation Targeting framework is a second best in terms of welfare, improving over the basic symmetric Inflation Targeting. However, in a low nominal interest rate environment, Price Level Targeting leads to a large degree of macroeconomic instability and is associated with an almost 21% probability of deflationary spirals.[19] The threshold level of de-anchoring (see Section 4.2) appears to be crossed repeatedly in a low nominal interest rate environment. In this environment, the asymmetric Inflation Targeting framework still outperforms the symmetric one and leads to the highest welfare of the three policy rules. Therefore, the asymmetric Inflation Targeting framework dominates the symmetric Inflation Targeting framework in terms of welfare and

[19] A deflationary spiral is defined as a quarter with inflation of less than -20%. If a deflationary spiral occurs, I reset all the relevant learning parameters and continue the simulation.

is more robust than Price Level Targeting across different levels of the equilibrium nominal interest rate.

Table 3: Simulation Results

	R*=6%			R*=3%		
	IT	PLT	asym. IT	IT	PLT	asym. IT
$\sigma(\pi)$	0.011	0.001	0.007	0.012	0.019	0.007
$\sigma(y)$	0.039	0.007	0.014	0.041	0.128	0.024
ZLB frequency (%)	8.188	2.250	7.688	35.750	92.750	28.500
ZLB length	6.700	4.966	5.528	12.344	82.364	9.486
Welfare $\mathcal{W}$	-0.138	-0.001	-0.025	-0.172	-0.546	-0.048
Risk of defl. Spiral (%)	0.063	0.000	0.000	0.063	21.375	0.063

Note: This table shows the simulation results under inflation targeting (IT), price level targeting (PLT), and asymmetric inflation targeting (asym. IT) á la Bianchi et al. (2019) under a high nominal interest rate environment (i.e. $R^* \sim 6\%$) and under a low nominal interest rate environment (i.e. $R^* \sim 3\%$).

7 Conclusion

Central bankers frequently voice concerns about the possibility of de-anchored inflation expectations, that is, the risk that households perceive temporary inflation shocks as being permanent so that short-term developments in inflation feeding into long-term inflation expectations. The contribution of this paper is to study the anchoring of inflation expectations. For this purpose, I build a model in which agents hold competing forecasting heuristics - a naive forecasting rule and an adaptive learning rule. The time-varying share of agents holding the naive forecasting heuristic determines the sensitivity of short- and long-run expectations to short-run conditions. The model has the same steady states as under rational expectations but features complex dynamics away from the steady state that are non-linear in the degree of anchoring: When expectations de-anchor, the volatility of output growth and inflation increases. Monetary policy can prevent expectations de-anchoring from causing inflationary or deflationary spirals as long as the Taylor principle is satisfied and the zero lower bound is not binding. At the zero lower bound, however, de-anchoring can lead to a self-fulfilling deflationary spirals. Thus, the potential welfare loss of de-anchoring is asymmetric and bigger in a low interest rate environment. I estimate the model using the non-linear particle filter on U.S. data and use the estimated model to explore the implications for monetary policy, both in high and low nominal interest rate environments. A striking result is that price level targeting can have serious de-stabilising effects when employed near the zero lower bound, due to its more restrictive stability requirements. However, in high nominal interest rate environments it successfully stabilises the economy. An asymmetric inflation targeting regime that responds more aggressively to below-target inflation on the other hand is particularly well suited to prevent the risk of deflationary spirals and improves over a symmetric inflation targeting framework in terms of welfare.

References

Angeletos, G.-M., Z. Huo, and K. A. Sastry (2020). Imperfect Macroeconomic Expectations: Evidence and Theory. In *NBER Macroeconomics Annual 2020, volume 35*, pp. 1–86. University of Chicago Press.

Anufriev, M., T. Assenza, C. Hommes, and D. Massaro (2013). Interest rate rules and macroeconomic stability under heterogeneous expectations. *Macroeconomic Dynamics 17*(8), 1574–1604.

Benhabib, J., S. Schmitt-Grohe, and M. Uribe (2002). Avoiding Liquidity Traps Avoiding Liquidity Traps Stephanie Schmitt-Grohe Martı ´ n Uribe. *Journal of Political Economy 110*(3), 535–563.

Berardi, M. (2009). Monetary policy with heterogeneous and misspecified expectations. *Journal of Money, Credit and Banking 41*(1), 79–100.

Bianchi, F., L. Melosi, and M. Rottner (2019). Hitting the Elusive Inflation Target.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018). Diagnostic Expectations and Credit Cycles. *Journal of Finance 73*(1), 199–227.

Branch, W. A. (2004). The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations. *Economic Journal 114*(497), 592–621.

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

Branch, W. A. and B. McGough (2009). A New Keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control 33*(5), 1036–1051.

Brock, W. A. and C. H. Hommes (1997). A Rational Route to Randomness. *Econometrica 65*(5), 1059–1096.

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

Carvalho, C., S. Eusepi, E. Moench, and B. Preston (2020). Anchored Inflation Expectations.

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

Cornea-Madeira, A., C. Hommes, and D. Massaro (2019). Behavioral Heterogeneity in U.S. Inflation Dynamics. *Journal of Business and Economic Statistics 37*(2), 288–300.

De Grauwe, P. (2011). Animal spirits and monetary policy. *Economic Theory 47*(2-3), 423–457.

De Grauwe, P. and Y. Ji (2019). Inflation Targets and the Zero Lower Bound in a Behavioural Macroeconomic Model. *Economica 86*(342), 262–299.

De Loecker, J., J. Eeckhout, and G. Unger (2020). The rise of market power and the macroeconomic implications. *Quarterly Journal of Economics 135*(2), 561–644.

Ehrmann, M. (2015). Targeting inflation from below: How do inflation expectations behave? *International Journal of Central Banking 11*(4), 213–249.

Eusepi, S., M. Giannoni, and B. Preston (2020). On the Limits of Monetary Policy.

Eusepi, S., C. Gibbs, and B. Preston (2021). Forward guidance with unanchored expectations.

Evans, C. L. (2012). Monetary Policy in a Low-Inflation Environment: Developing a State-Contingent Price-Level Target. *Journal of Money, Credit and Banking 44*(SUPPL. 1), 147–155.

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

Fernández-Villaverde, J. and J. F. Rubio-Ramírez (2007). Estimating macroeconomic models: A likelihood approach. *Review of Economic Studies 74*(4), 1059–1087.

Gáti, L. (2020). *Monetary Policy & Anchored Expectations An Endogenous Gain Learning Model.* Ph. D. thesis.

Geweke, J. (1999). Using simulation methods for bayesian econometric models: Inference, development, and communication. *Econometric Reviews 18*(1), 1–73.

Gibbs, C. G. and M. Kulish (2017). Disinflations in a model of imperfectly anchored expectations. *European Economic Review 100*, 157–174.

Godsill, S. J., A. Doucet, and M. West (2004). Monte carlo smoothing for nonlinear time series. *Journal of the American Statistical Association 99*(465), 156–168.

Goy, G., C. Hommes, and K. Mavromatis (2020). Forward guidance and the role of central bank credibility under heterogeneous beliefs. *Journal of Economic Behavior and Organization*.

Gürkaynak, R. S., B. Sack, and E. Swanson (2005). The sensitivity of long-term interest rates to economic news: Evidence and implications for macroeconomic models. *American Economic Review 95*(1), 425–436.

Hommes, C. and J. Lustenhouwer (2019). Inflation targeting and liquidity traps under endogenous credibility. *Journal of Monetary Economics 107*(C), 48–62.

Honkapohja, S. and K. Mitra (2020). Price level targeting with evolving credibility. *Journal of Monetary Economics 116*(C), 88–103.

Ilabaca, F. and F. Milani (2020). Heterogeneous Expectations, Indeterminacy, and Postwar US Business Cycles.

Jorgensen, P. L. and K. J. Lansing (2019). Anchored Inflation Expectations and the Slope of the Phillips Curve.

Kreps, D. M. (1997). Anticipated Utility and Dynamic Choice. In *Frontiers of Research in Economic Theory*, Chapter Anticipate, pp. 242 – 274.

Kumar, S., H. Afrouzi, O. Coibion, and Y. Gorodnichenko (2015). Inflation Targeting Does Not Anchor Inflation Expectations: Evidence from Firms in New Zealand. *National Bureau of Economic Research* (Solow 1979).

Łyziak, T. and M. Paloviita (2017). Anchoring of inflation expectations in the euro area: Recent evidence based on survey data. *European Journal of Political Economy 46*(C), 52–73.

Malmendier, U. and S. Nagel (2016). Learning from inflation experiences. *Quarterly Journal of Economics 131*(1), 53–87.

Mankiw, N. G. and R. Reis (2002). Sticky information versus sticky prices: A proposal to replace the new Keynesian Phillips curve. *Quarterly Journal of Economics*.

Mankiw, N. G., R. Reis, and J. Wolfers (2003). *Disagreement about inflation expectations*, Volume 18. National Bureau of Economic Research, Inc.

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

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

Pfajfar, D. and E. Santoro (2010). Heterogeneity, learning and information stickiness in inflation expectations. *Journal of Economic Behavior and Organization 75*(3), 426–444.

Reis, R. (2021). Losing the Inflation Anchor. In *BPEA Conference Drafts*. Brookings Institution.

Richter, A. W., N. A. Throckmorton, and T. B. Walker (2014). Accuracy, speed and robustness of policy function iteration. *Computational Economics 44*(4), 445–476.

Romer, C. D. and D. H. Romer (2004). A New Measure of Monetary Shocks: Derivation and Implications. *American Economic Review 94*(4), 1055–1084.

Rotemberg, J. J. (1982). Sticky Prices in the United States. *Journal of Political Economy 90*(6), 1187–1211.

Sargent, T. J. (1999, 3). *The Conquest of American Inflation*. Princeton University Press.

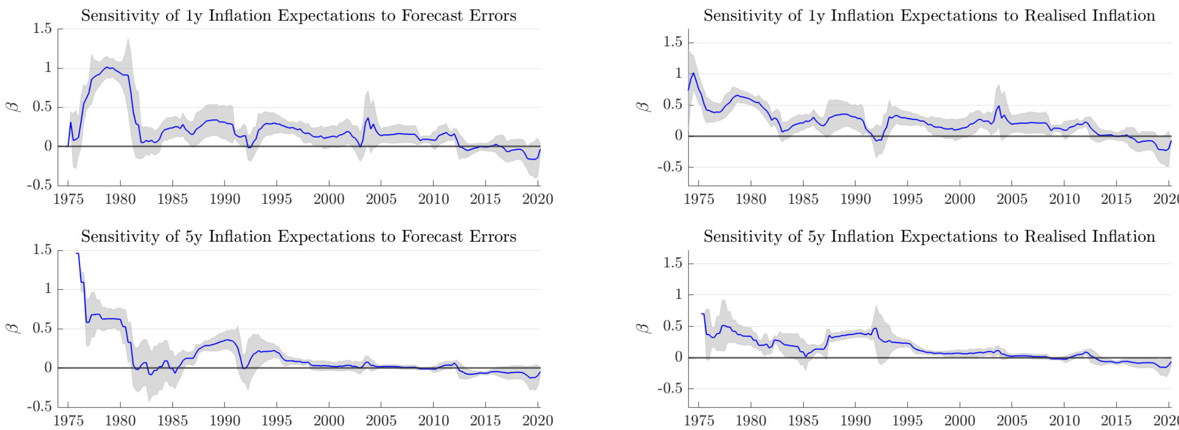
Strohsal, T., R. Melnick, and D. Nautz (2016). The time-varying degree of inflation expectations anchoring. *Journal of Macroeconomics 48*(C), 62–71.

Wu, J. C. and F. D. Xia (2016). Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. *Journal of Money, Credit and Banking 48*(2-3), 253–291.

A De-Anchoring in Practice

A.1 Alternative measures of inflation expectations

Figure A.1: Sensitivity of Household Expectations to Realised Inflation & Forecast Errors

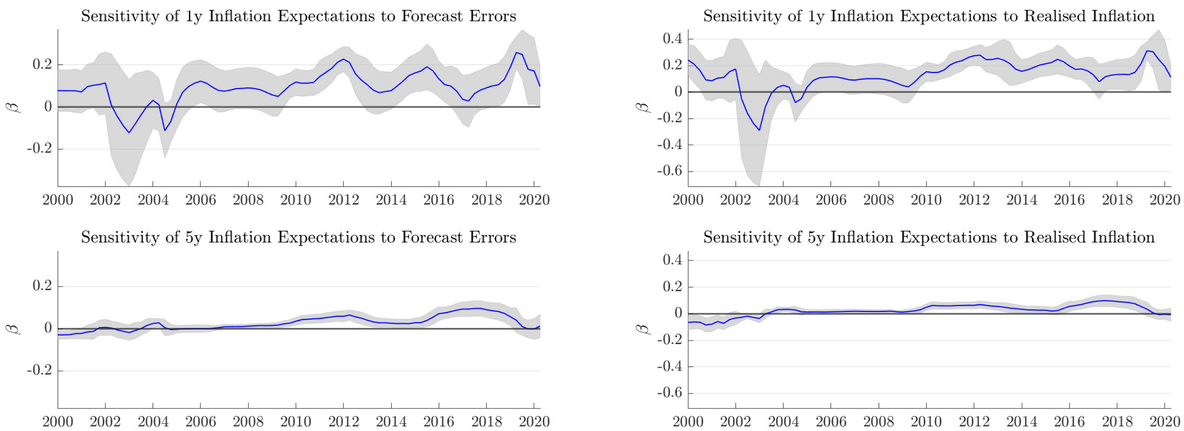


Note: The panels show the estimated time-varying sensitivity of the median short-run (upper row) and median long-run (lower row) inflation expectations from the Michigan Fed's Survey of Consumers to inflation surprises (left column) and realised inflation (right column). Forecast errors are computed as the difference between realised CPI and the corresponding median forecast for the same period. Solid blue lines depict the estimated time-varying coefficient with 95% confidence intervals shaded in grey.

A.2 De-Anchoring in the euro area

Inflation expectations of professional forecasters in the ECB's SPF display a similar pattern of time-varying sensitivity to short term news (see Figure A.2). 1-year and 5-year expectations do not react significantly to short term news in the pre-crisis sample, but they exhibit signs of de-anchoring first during the euro crisis and later during 2015-17. This latter period coincided with the ECB's large scale asset purchase programmes that were explicitly started due to worries about de-anchoring & deflation.

Figure A.2: Sensitivity of Expectations to Forecast Errors and Inflation Realisations (euro area)



Note: The panels show the estimated time-varying sensitivity of short-run (upper row) and long-run (lower row) inflation expectations from the ECB's Survey of Professional Forecasters to inflation surprises (left column) and realised inflation (right column). Solid blue lines depict the estimated time-varying coefficient with 95% confidence intervals shaded in grey.

B Derivation of the Fixed Point

Take the ODE

$$\frac{\partial \mathbf{\Psi}'}{\partial \tau} = \mathbf{R}^{-1} E\mathbf{x}_{t-2} \left(\mathbf{A} \left[(1-n)\mathbf{\Psi}\mathbf{x}_{t-1}) \right] + \mathbf{A} n \mathbf{z}_{t-2} + \mathbf{B}\mathbf{w}_t + \mathbf{C}\bar{\mathbf{z}} \right) - \mathbf{\Psi}\mathbf{x}_{t-2})' \tag{58}$$

and work separately with $\mathbf{a}, \mathbf{b}$. Given that $\lim_{t\to\infty} E\mathbf{x}_t\mathbf{x}_t = \Sigma_x = \mathbf{R} = diag(1, \mathbf{\Sigma_w})$, it is possible to write the ODE as

$$\frac{\partial \mathbf{a}'}{\partial \tau} = \left(\mathbf{A} \left[(1-n)\mathbf{a} \right] + n\mathbf{A}\mathbf{z}_{t-2} + \mathbf{C}\bar{\mathbf{z}} \right)' - \mathbf{a}' \tag{59}$$

$$\frac{\partial \mathbf{b}'}{\partial \tau} = \mathbf{\Sigma_w}^{-1} \mathbf{\Sigma}_{\mathbf{w}_{t-2},\mathbf{w}_{t-1}} (\mathbf{A} \left[(1-n)\mathbf{b}) \right] + \mathbf{B})' + \mathbf{\Sigma}_w^{-1} \mathbf{\Sigma}_{\mathbf{w}_{t-2},\mathbf{z}_{t-2}} (\mathbf{A}n)' - \mathbf{b}' \tag{60}$$

As long as the ALM in Equation (30) is asymptotically stationary, i.e.

$$\lambda < 1 \quad \forall \lambda \in \Lambda = \{ \Lambda : |\mathbf{I} - n\mathbf{A} - \mathbf{\Lambda}\mathbf{I}| = 0 \} \tag{61}$$

we can write, holding $(\mathbf{\Psi}', n)$ fixed, $\lim_{t\to\infty} E\mathbf{z_t}$ as

$$\mathbf{z}_t(\mathbf{\Psi}, n) = \left((\mathbf{I} - n\mathbf{A})^{-1} \left(\mathbf{A} \left[(1-n)\mathbf{a} \right] + \mathbf{C}\bar{\mathbf{z}} \right) \right)' \tag{62}$$

so that the ODE for $\mathbf{a}$ becomes

$$\frac{\partial \mathbf{a}'}{\partial \tau} = \left(\mathbf{A}\left[(1-n)\mathbf{a} + n(\mathbf{I} - n\mathbf{A})^{-1}\left(\mathbf{A}(1-n)\mathbf{a} + \mathbf{C}\bar{\mathbf{z}}\right)\right] + \mathbf{C}\bar{z}\right)' - \mathbf{a}' \tag{63}$$

yielding the fixed point

$$vec(\bar{\boldsymbol{a}}') = (\boldsymbol{I} - \boldsymbol{A})^{-1} vec\left((\boldsymbol{C}\bar{\boldsymbol{z}})'\right) \tag{64}$$

Turning our attention to the ODE for $\mathbf{b}$ we note that $\mathbf{\Sigma}_{\mathbf{w}_{t-2},\mathbf{w}_{t-1}} = \mathbf{F}\mathbf{\Sigma}_{\mathbf{w}} = \mathbf{\Sigma}_{\mathbf{w}}\mathbf{F}$ (because both are diagonal) and that $\Sigma_{\mathbf{w}_{t-2},\mathbf{z}_{t-2}} = \Sigma_{\mathbf{w}_t,\mathbf{z}_t}$ is endogenously determined:

$$\mathbf{\Sigma}_{\mathbf{w}_t,\mathbf{z}_t} = E(\mathbf{w}_t)(\mathbf{A}\left[(1-n)\mathbf{\Psi}\mathbf{x}_{t-1})\right] + \mathbf{A}n\mathbf{z}_{t-2} + \mathbf{B}\mathbf{w}_{t-1} + \mathbf{C}\bar{z})' \tag{65}$$

$$\Leftrightarrow \mathbf{\Sigma}_{\mathbf{w}_t,\mathbf{z}_t} = \Sigma_{\mathbf{w}}\mathbf{b}'(\mathbf{A}(1-n))' + \mathbf{\Sigma}_{\mathbf{w}}\mathbf{B}' + \Sigma_{\mathbf{w}_t,\mathbf{z}_{t-1}}(\mathbf{A}n)' \tag{66}$$

where $\Sigma_{\mathbf{w}_t,\mathbf{z}_{t-1}}$ again is endogenous

$$\mathbf{\Sigma}_{\mathbf{w}_t,\mathbf{z}_{t-1}} = E(\mathbf{F}\mathbf{w}_{t-1})(\mathbf{A}\left[(1-n)\mathbf{\Psi}\mathbf{x}_{t-1})\right] + \mathbf{A}n\mathbf{z}_{t-2} + \mathbf{B}\mathbf{w}_{t-1} + \mathbf{C}\bar{\mathbf{z}})' \tag{67}$$

$$\Leftrightarrow \mathbf{\Sigma}_{\mathbf{w}_t,\mathbf{z}_{t-1}} = \mathbf{F}\mathbf{\Sigma}_{\mathbf{w}}(\mathbf{A}(1-n)\mathbf{b} + \mathbf{B})' + \mathbf{F}\mathbf{\Sigma}_{\mathbf{w}_{t-1},\mathbf{z}_{t-2}}(\mathbf{A}n)' \tag{68}$$

$$\Rightarrow vec(\mathbf{\Sigma}_{\mathbf{w}_t,\mathbf{z}_{t-1}}) = (\mathbf{I} - \mathbf{A}n \otimes \mathbf{F})^{-1}\left(\mathbf{A}(1-n) \otimes \mathbf{F}\mathbf{\Sigma}_{\mathbf{w}} vec(\mathbf{b}') + vec(\mathbf{F}\mathbf{\Sigma}_{\mathbf{w}}\mathbf{B}')\right) \tag{69}$$

Substituting this into $vec\left(\Sigma_{\mathbf{w}_t,\mathbf{z}_t}\right)$ and then into $vec({}^{\partial \mathbf{b}'}/_{\partial\tau})$ yields the fixed point

$$vec(\bar{\boldsymbol{b}}') = (\boldsymbol{I} - \mathbf{G}_1)^{-1}\mathbf{G}_2 vec(\mathbf{B}') \tag{70}$$

where

$$\begin{aligned}\mathbf{G}_1 = {} & \mathbf{A}(1-n) \otimes \mathbf{F} + \mathbf{A}n \otimes {\mathbf{\Sigma}_{\mathbf{w}}}^{-1}\mathbf{A}(1-n) \otimes \mathbf{\Sigma}_{\mathbf{w}} \\ & + \mathbf{A}n \otimes {\mathbf{\Sigma}_{\mathbf{w}}}^{-1}\mathbf{A}n \otimes \mathbf{I}\left[\mathbf{I} - n\mathbf{A} \otimes \mathbf{F}\right]^{-1}\left[\mathbf{A}(1-n) \otimes \mathbf{F}\mathbf{\Sigma}_{\mathbf{w}}\right]\end{aligned}$$

$$\begin{aligned}\mathbf{G}_2 = {} & \mathbf{I} \otimes \mathbf{F} + \mathbf{A}n \otimes {\mathbf{\Sigma}_{\mathbf{w}}}^{-1}\mathbf{I} \otimes \mathbf{\Sigma}_{\mathbf{w}} \\ & + \mathbf{A}n \otimes {\mathbf{\Sigma}_{\mathbf{w}}}^{-1}\mathbf{A}n \otimes \mathbf{I}\left[I - \mathbf{A}n \otimes \mathbf{F}\right]^{-1}\mathbf{F}\mathbf{\Sigma}_{\mathbf{w}} \otimes \mathbf{I}\end{aligned}$$

The mapping T from the perceived to the actual law of motion for the PLM coefficients for a given n can therefore be characterised as

$$T(\mathbf{\Psi}', n) = \begin{pmatrix} T(\mathbf{a}', n) \\ T(\mathbf{b}', n) \end{pmatrix} = \begin{pmatrix} ((\boldsymbol{I} - \boldsymbol{A}n)^{-1}\boldsymbol{A}(1-n))' \\ \mathbf{G}_1 \end{pmatrix} \tag{71}$$

C Proofs

C.1 Proof of Proposition I

Proof. A fixed point exists if the process z_t is asymptotically stationary. Since the eigenvalues of $(n\mathbf{A})$ are increasing in n, a fixed point exists if the eigenvalues of A lie within the unit circle. Bullard and Mitra (2002) show that a necessary and sufficient condition for the eigenvalues of A to lie within the unit circle for a bivariate Taylor rule is

$$\kappa\,(\phi_\pi - 1) + (1-\beta)\phi_y > 0$$

Strict inflation targeting is the limit case with $\phi_y = 0$, thus reducing the requirement to

$$\phi_\pi > 1 \tag{72}$$

i.e. the standard Taylor principle. □

C.2 Proof of Proposition II

Proof. I guess that $\mathbf{z}^* = \bar{\mathbf{a}} = (\frac{1-\beta}{\kappa}\bar{\pi}, \bar{\pi})'$ (i.e. the rational expectations steady state) is a steady state. In the absence of structural shocks and given $\mathbf{z}_t = \mathbf{z}_{t-1}$ neither forecast heuristic makes prediction errors so that $n_t = n = 0.5$ and the system becomes

$$y^* = (1-n)y^* + ny^* - \frac{1}{\sigma}\left(\bar{\pi} + \phi_\pi(\pi^* - \bar{\pi}) - (1-n)\bar{\pi} - n\pi^*\right)$$

$$\pi^* = \beta[(1-n)\bar{\pi} + n\pi^*] + \kappa y^*$$

It is clear that $\pi^* = \bar{\pi}$ and $y^* = \frac{1-\beta}{\kappa}\bar{\pi}$ is a solution, thus confirming the initial guess. This is the case even if we were to fix the share of naive forecasters at any value $n_t = n \in (0, 1)$. However, due to the decay of past prediction errors and the forecast selection mechanism, only $n_t = n^* = 0.5$ can be a steady state. □

C.3 Proof of Proposition III

Proof. When all households are adaptive learners (i.e. $n_t = 0$), $DT_a(\mathbf{a}', n)$ reduces to $[\mathbf{A}]$. Given that the eigenvalues of $\mathbf{A}$ are smaller than one in absolute value, the eigenvalues of $[\mathbf{A} - \mathbf{I}]$ will be negative since we subtract 1 from each of the eigenvalues of $\mathbf{A}$. Similarly, $DT_b(\mathbf{b}', n)$ reduces to $[\mathbf{A} \otimes \mathbf{F}]$, so that the eigenvalues of $[\mathbf{A} \otimes \mathbf{F} - \mathbf{I}]$ will be negative as well. When all households are naive forecasters (i.e. $n_t = 1$), both $DT_a(\mathbf{a}', n)$ and $DT_b(\mathbf{b}', n)$ reduce to zero so that the associated eigenvalues are of course -1. Any combination in between for $n \in (0, 1)$ is a non-linear combination of the two preceding cases whose eigenvalues are decreasing in n and therefore still features negative eigenvalues as Figure 3 illustrates. □

C.4 Proof of Proposition IV

Proof. The eigenvalues of $\tilde{\mathbf{A}}$ are given by

$$\lambda_{1,2} = \frac{1}{2}\left(1 + \beta + \frac{\kappa}{\sigma} \pm \sqrt{\left(1 + \beta + \frac{\kappa}{\sigma}\right)^2 - 4\beta}\right)$$

which implies for all positive values of β, κ, and σ that $|\lambda_1| < 1$ and $\lambda_2 > 1$. Since the eigenvalues of $n\tilde{\mathbf{A}}$ are increasing in n, there exists $\bar{n} = \frac{2}{\left(1+\beta+\frac{\kappa}{\sigma}+\sqrt{(1+\beta+\frac{\kappa}{\sigma})^2-4\beta}\right)} \in (0, 1)$ such that only for $n < \bar{n}$ all eigenvalues of $n\tilde{\mathbf{A}}$ fall within the unit circle, the process $\mathbf{z}_t$ is asymptotically stationary, and the mapping $\tilde{T}$ has a fixed point. For $n \geq \bar{n}$ the process is not stationary and no fixed point exists. □

C.5 Proof of Proposition V

Proof. I again guess that $\tilde{\mathbf{z}}^* = \tilde{\bar{\mathbf{a}}} = (0, 0)'$ and assume that $\bar{n} > 0.5$. In the absence of structural shocks and given $\mathbf{z}_t = \mathbf{z}_{t-1}$ neither forecast heuristic makes prediction errors so that $n_t = n = 0.5$ and the system becomes

$$\tilde{y} = (1 - n)\tilde{y} + n_t\tilde{y} - \frac{1}{\sigma}\left(-((1 - n)\tilde{\pi} + n\tilde{\pi})\right)$$

$$\tilde{\pi} = \beta((1 - n)\tilde{\pi} + n\tilde{\pi}) + \kappa\tilde{y}$$

which is solved by $\tilde{\pi}^* = \tilde{y}^* = 0$, confirming the initial guess. However, if $\bar{n} \leq 0.5$, there exists no fixed point $\tilde{\bar{\mathbf{a}}}$ so that the ZLB steady state does not exist. □

C.6 Proof of Proposition VI

Proof. I again first consider the extreme cases: When all households are adaptive learners (i.e. $n_t = 0$), $D\tilde{T_a}(\mathbf{a}', n)$ reduces to $[\tilde{\mathbf{A}}]$. Given that one eigenvalue of $\tilde{\mathbf{A}}$ is bigger than one, not all eigenvalues of $[\tilde{\mathbf{A}} - \mathbf{I}]$ will be negative. The magnitude of eigenvalues of $D\tilde{T_b}(\mathbf{b}', n)$ on the other hand depends on the correlation matrix $\mathbf{F}$ and might be bigger or smaller than one in absolute value. When all households are naive forecasters (i.e. $n_t = 1$), both $D\tilde{T_a}(\mathbf{a}', n)$ and $D\tilde{T_b}(\mathbf{b}', n)$ reduce to zero so that the associated eigenvalues are of course -1. Since the relevant eigenvalues switch signs I need to consider the steady state case $n = 0.5$ explicitly. In that case the mapping becomes:

$$D\tilde{T_a}(\mathbf{a}', 0.5) = \left[(\mathbf{I} - {}^1\!/_2\tilde{\mathbf{A}})^{-1}\,{}^1\!/_2\tilde{\mathbf{A}}\right] = (2\mathbf{A}^{-1} - \mathbf{I})^{-1} \tag{73}$$

E-stability requires that

$$eig((2\tilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1} - \mathbf{I}) < 0$$

Where $eig(\mathbf{A})$ denotes the vector of eigenvalues λ associated with any matrix $\mathbf{A}$. The above is equivalent to

$$\Leftrightarrow eig((2\tilde{\mathbf{A}}^{-1} - \mathbf{I})^{-1}) - 1 < 0$$

$$\Leftrightarrow 1 \oslash eig((2\tilde{\mathbf{A}}^{-1} - \mathbf{I})) - 1 < 0$$

where $\oslash$ denotes the Hadamard division, i.e. elementwise division. Note further

$$\Leftrightarrow 1 \oslash (eig(2\tilde{\mathbf{A}}^{-1}) - 1) - 1 < 0$$

$$\Leftrightarrow 1 \oslash (2 \oslash eig(\tilde{\mathbf{A}}) - 1) - 1 < 0$$

$$\Leftrightarrow eig(\tilde{\mathbf{A}}) \oslash (2 - eig(\tilde{\mathbf{A}})) - 1 < 0$$

however, since one of the eigenvalues of $\tilde{\mathbf{A}}$ is larger than one, this condition cannot be satisfied. □

C.7 Proof of Proposition VII

Proof. Consider the case of $n = 1$. In that case the ALM collapses to:

$$\boldsymbol{z}_t = \boldsymbol{A} n_t \boldsymbol{z}_{t-1} + \boldsymbol{D}\boldsymbol{z}_{t-1} + \boldsymbol{B}\boldsymbol{w}_t + \boldsymbol{C}\bar{\boldsymbol{z}} > 0$$

Stationarity requires that all eigenvalues of $\mathbf{A} + \mathbf{D}$ lie within the unit circle or, equivalently, all eigenvalues $\mathbf{I} - \mathbf{A} - \mathbf{D}$ to be negative. However, the trace of $\mathbf{I} - \mathbf{A} - \mathbf{D}$ is positive as long as $\beta \leq 1$ and $\phi_\pi \geq 1$. Since the trace of a matrix equals the sum of its eigenvalues, this implies that at least one eigenvalue must be positive.

$$tr(\mathbf{I}-\mathbf{A}-\mathbf{D}) = (\mathbf{I}-\frac{\sigma}{\kappa\phi_\pi+\sigma}) + (\mathbf{I}-\frac{\kappa+\beta\sigma}{\kappa\phi_\pi+\sigma}) + (\mathbf{I}-\frac{\sigma}{\kappa\phi_\pi+\sigma})$$

Therefore, there must exist an $\bar{n} \in [0,1)$, above which the system becomes explosive and the mapping T has no fixed point. $\square$

C.8 Proof of Proposition VIII

Proof. I guess that $\mathbf{z}^* = \bar{\mathbf{a}} = (\frac{1-\beta}{\kappa}\bar{\pi}, \bar{\pi}, 0)'$ (i.e. the rational expectations steady state) is a steady state. In the absence of structural shocks and given $\mathbf{z}_t = \mathbf{z}_{t-1}$ neither forecast heuristic makes prediction errors so that $n_t = n = 0.5$ and the system becomes

$$y^* = (1-n)y^* + (1-n)c_{y,p}\hat{p}^* + ny^* - \frac{1}{\sigma}\left(\bar{\pi} + \phi_\pi(\pi_t - \bar{\pi} + \hat{p}^*) - (1-n)\pi^* - (1-n)c_{\pi,p}\hat{p}^* - n\pi^*\right)$$

$$\pi^* = \beta[(1-n)\pi^* + (1-n)c_{\pi,p}\hat{p}^* + n\pi^*] + \kappa y^*$$

$$\hat{p}^* = \pi^* - \bar{\pi} + \hat{p}^*$$

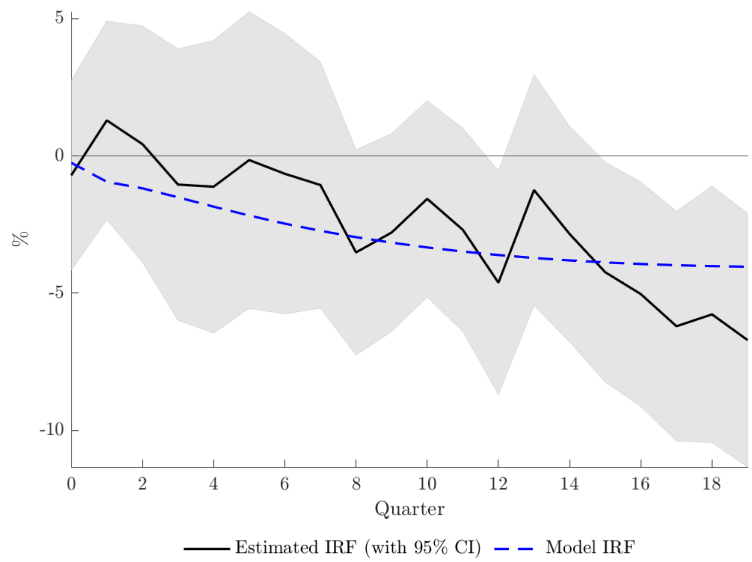
It is clear that $\pi^* = \bar{\pi}$, $y^* = \frac{1-\beta}{\kappa}\bar{\pi}$, and $\hat{p}^* = 0$ is a solution, thus confirming the initial guess. This is the case even if we were to fix the share of naive forecasters at any value $n_t = n < \bar{n}$. However, due to the decay of past prediction errors and the forecast selection mechanism, only $n_t = n^* = 0.5$ can be a steady state. However, if $\bar{n} \leq 0.5$, there exists no fixed point $\tilde{\bar{\mathbf{a}}}$ so that the steady state does not exist. $\square$

C.9 Proof of Proposition IX

Proof. Suppose the coefficient $\bar{\mathbf{c}}_{[3,3]}$ on the lagged price level gap converged to its true value one. Then the system is not stationary because at least eigenvalue of $\mathbf{A}(1-n)\bar{\mathbf{c}} + n\mathbf{A} + \mathbf{D}$ lies outside the unit circle. For values of $\bar{\mathbf{c}}_{[3,3]} < 1$ the ALM might be stationary, but this cannot be a fixed point of $\bar{\mathbf{c}}$. $\square$

D Model Dynamics

D.1 Long-run Inflation Expectations & Monetary Policy



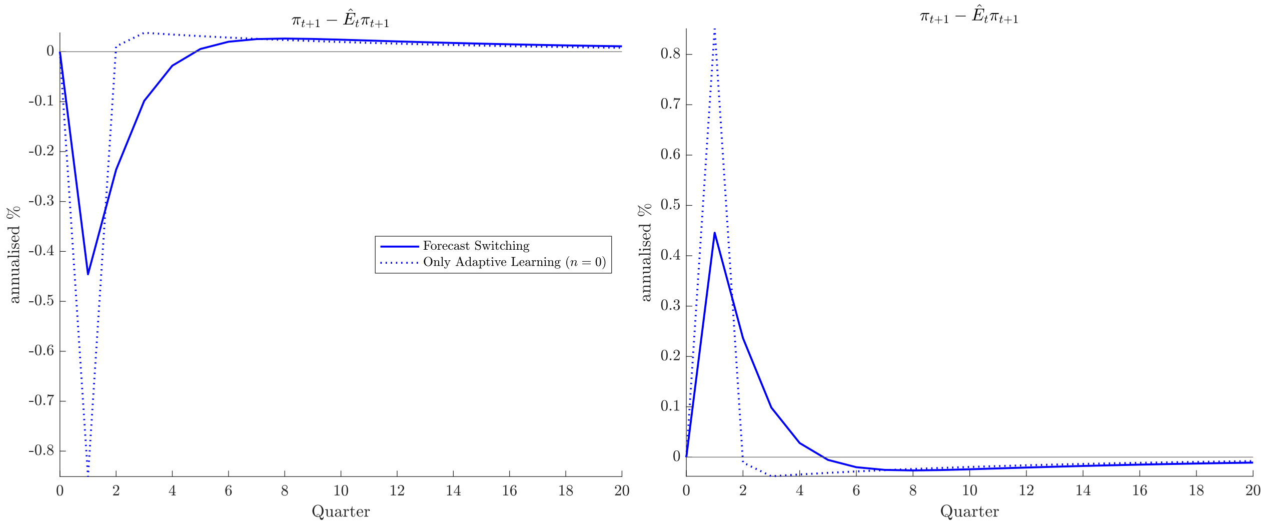
Note: This Figure plots the estimated IRF of (log) 10y inflation expectations to a 1pp monetary policy shock along with the equivalent model IRF. The data IRF is estimated using a local projection on a sample from 1990Q1:2007Q4 and uses Romer and Romer (2004) shocks to identify exogenous movements in monetary policy, which are used to instrument the 1y Treasury Rate. The model further includes lagged values of real GDP, CPI, the 1y Treasury Rate, and the dependent variable. Standard errors are heteroskedastic and autocorrelation consistent with 12 lags. IRF is robust to 1) including the monetary policy shock directly instead of using it as instrument; 2) using the combined data on long-run inflation expectations as in Figure 8; 3) using the high frequency identified shocks of Gürkaynak et al. (2005) and extending the sample to 1982Q1:2007Q4.

Figure D.1: IRF of 10y Inflation Expectations to a 1pp Monetary Policy Shock

D.2 Forecast Errors

Angeletos et al. (2020) document two stylised facts of aggregate expectations: They initially under-react but later over-shoot the actual outcomes. I test whether my set up of expectation formation fits these facts. I estimate the impulse response functions of average 1q ahead inflation forecast errors (i.e. $\pi_{t+1}\hat{\mathbb{E}}_t\pi_{t+1}$) to a monetary policy shock. The left (right) panel of Figure D.5 shows the IRF of the forecast error after a positive (negative) monetary policy shock that decreases (increases) inflation. In both cases we see an initial under-reaction of expectations. That is, after a positive (negative) monetary policy shock, households initially expect higher (lower) inflation than eventually realises. However, after 5 periods, the forecast errors flip signs, i.e. they over-react. Thus, my model with heuristic switching forecasts fits the stylised facts of Angeletos et al. (2020). As comparison, I also plot the forecast errors of the model if it was populated by adaptive learners only (i.e. $n_t = 0 \forall t$). In this case forecast errors are negative only upon impact of the shock, but then immediately over-shoot, thus not quite fitting the findings of Angeletos et al. (2020).

Figure D.2: IRF of Forecast Errors



Note: This Figure shows the IRF of the average inflation forecast errors to positive (right) and negative (left) 1pp monetary policy shock which occurs in period 1. The solid blue line depicts the forecast errors of the heuristic switching model, whereas the dotted blue line depicts the counterfactual forecast errors if the model contained adaptive learners only.

D.3 Model Moments

Figure D.3: Model Variances & Autocorrelation

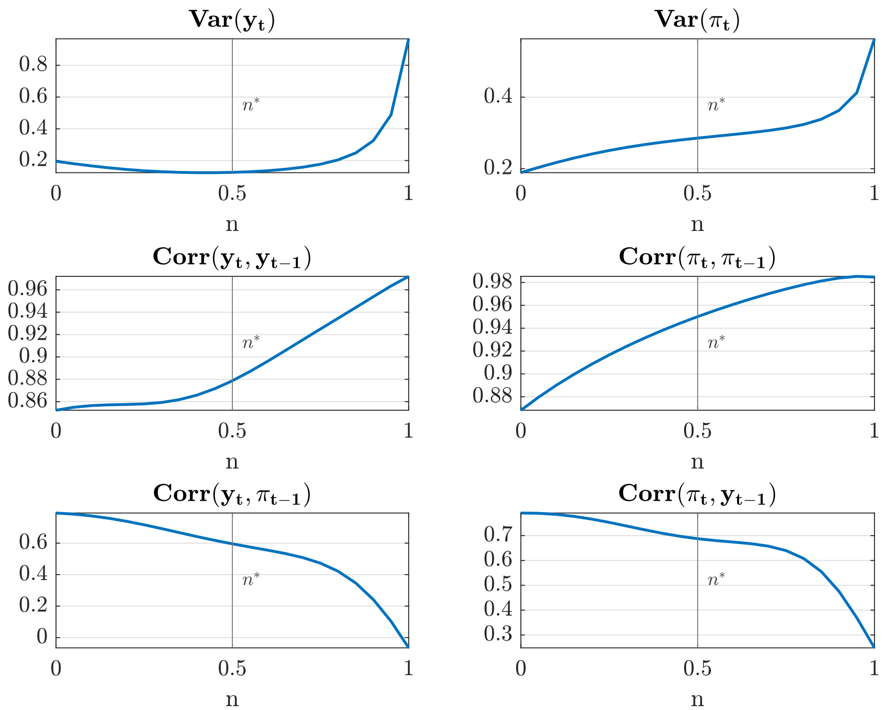
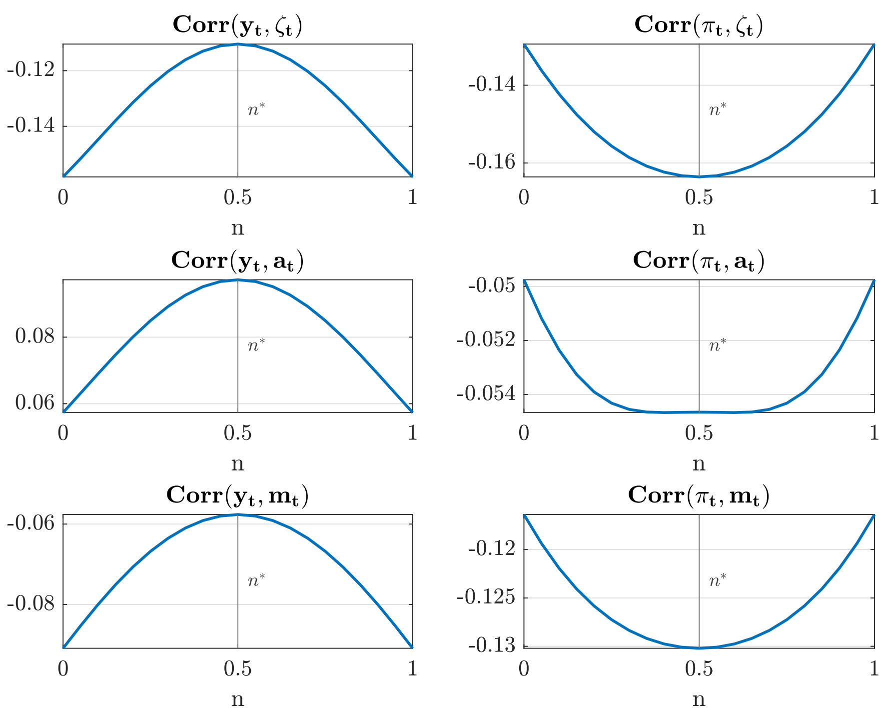
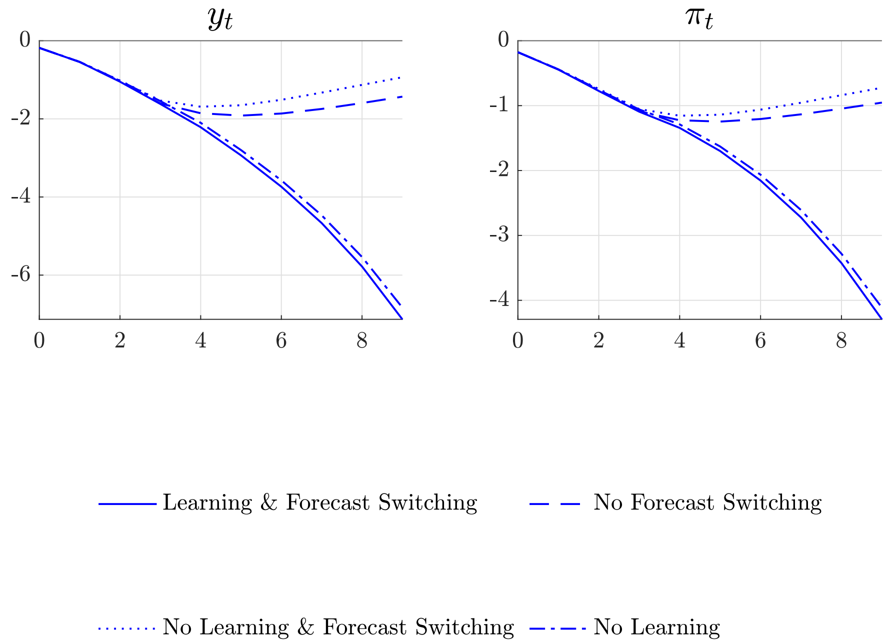


Figure D.4: Model Correlations with Exogenous Processes



D.4 Counterfactual IRFs

Figure D.5: Counterfactual IRF Shutting Down one Updating Mechanism at a Time



Note: This Figure shows the IRF of output (left) and inflation (right) to four consecutive σ_ζ shocks to the discount factor under different assumptions about the expectation formation. The solid blue line shows the response of the full model (i.e. the same response as in Figure 9); the dashed lined shows the IRF of the model with a constant share of naive forecasters (i.e. $n = 0$); the dotted lined shows the IRF of the model with a constant share of naive forecasters (i.e. $n = 0.5$) and no learning on the side of the adaptive forecasters (i.e. $\Psi_t = \Psi$). The dash-dotted lined shows the IRF of the model with a time varying share of naive forecasters but no learning on the side of the adaptive forecasters (i.e. $\Psi_t = \Psi$).