DNB Working Paper

No. 614 / December 2018

Forward Guidance and the Role of Central Bank Credibility under Heterogeneous Beliefs

Gavin Goy, Cars Hommes and Kostas Mavromatis

DeNederlandscheBank

EUROSYSTEEM

Forward Guidance and the Role of Central Bank Credibility under Heterogeneous Beliefs

Gavin Goy, Cars Hommes and Kostas Mavromatis *

* Views expressed are those of the authors and do not necessarily reflect official positions of De Nederlandsche Bank.

Working Paper No. 614

December 2018

De Nederlandsche Bank NV
P.O. Box 98
1000 AB AMSTERDAM
The Netherlands

Forward Guidance and the Role of Central Bank Credibility under Heterogeneous Beliefs*

Gavin Goy[†a,b,c], Cars Hommes[a,b], and Kostas Mavromatis[c]

[a] *CeNDEF, University of Amsterdam*
[b] *Tinbergen Institute*
[c] *De Nederlandsche Bank*

December 2018

Abstract

This paper studies the macroeconomic effects of central bank forward guidance when central bank credibility is endogenous. In particular, we take a stylized New Keynesian model with an occasionally binding zero lower bound constraint on nominal interest rates and heterogeneous and boundedly rational households. The central bank uses a bivariate VAR to forecast, not taking into account the time-variation in the distribution of aggregate expectations. In this framework, we extend the central bank's toolkit to allow for the publication of its own forecasts (Delphic guidance) and the commitment to a future path of the nominal interest rate (Odyssean guidance). We find that both Delphic and Odyssean forward guidance increase the likelihood of recovery from a liquidity trap. Even though Odyssean guidance alone appears more powerful, we find it to increase ex post macroeconomic volatility and thus reduce welfare.

JEL-Classification: E03, E12, E58

Keywords: Forward Guidance, Heterogeneous Beliefs, Bounded Rationality, Central Bank Learning

*We wish to thank Marco Del Negro, Paul Levine, Joep Lustenhouwer, Domenico Massaro, Bruce Preston and Micheal Woodford for valuable discussions. Also, we thank seminar participants at the European Central Bank, De Nederlandsche Bank and the University of Amsterdam, as well as conference participants at the 21st Workshop on the Economic Science with Heterogeneous Interacting Agents, the 22nd Conference Computing in Economics and Finance, the 12th Dynare Conference, the 4th meeting of the German Network for New Economic Dynamics, the 2016 MacFin Robots Workshop in Como, Italy, the 22nd Spring Meeting of Young Economists, the 72nd European Meeting of the Econometric Society and the Conference on Expectations in Dynamic Macroeconomic Models at the Federal Reserve Bank of St. Louis in 2017 for comments and helpful discussions. Lastly, we are also thankful for financial support from the EU FP7 project "Integrated Macro-Financial Modeling for Robust Policy Design" (MACFINROBODS, grant 612796). The views expressed do not reflect the official position of De Nederlandsche Bank or the Eurosystem.

†Corresponding author: Gavin Goy (E-mail: g.w.goy@uva.nl), Roetersstraat 11, 1018WD, Amsterdam, The Netherlands.

1 Introduction

With policy rates in most developed economies being constrained by the zero lower bound (ZLB), several central banks made use of unconventional policy tools, such as large scale asset purchases, long-term liquidity provisions and forward guidance. The latter can be seen as an additional tool for policy makers to influence the public's expectations−which play a crucial role in the transmission mechanism of monetary policy−especially once the policy rate hits the ZLB.

The effectiveness of forward guidance is, however, not a priori clear, because it depends on how agents interpret the announcement: "is it a signal of additional monetary stimulus, or rather a sign that the central bank's economic outlook became worse?" (see De Graeve et al., 2014, p.3). Empirical evidence finds that forward guidance has a considerably impact on private sector expectations, yet also reveals that the public moves its expectations only partially in the direction of the announcement (Ferrero and Secchi, 2009; Hubert, 2014, 2015*a*,*b*).[1] Clearly, a model with perfect information and a fully credible central bank cannot address this question.

The main contribution of this paper is to analyze forward guidance under bounded rationality, where we link the effectiveness of forward guidance to the central bank's *credibility* in a heterogeneous expectation framework. To this end, we depart from the standard New Keynesian model in four ways. First, credibility is endogenous. Second, expectation of household-firms are heterogeneous, with agents' beliefs switching endogenously between two rules (*adaptive expectations* and *credibility believers*). Third, we use N-step ahead Euler equation learning to study the effects of forward guidance announcements, and fourth, the central bank cannot observe contemporaneous inflation and output, but instead uses a vector autoregression (VAR) model, which is miss-specified because of time variation in the distribution of aggregate expectations.

The first two novelties are closely related. We use a heterogeneous expectations heuristics switching model with two types, as in Hommes and Lustenhouwer (2015), with one type believing in the targets of the central bank (the credibility believers) and another type using a simple backward looking adaptive expectations rule. We extend the model to allow for two types of forward guidance: *Delphic* and *Odyssean* guidance (see Campbell et al., 2012). Under Delphic forward guidance the central bank communicates to the public its forecasts of the economic outlook and the expected monetary policy action consistent with this outlook. Contrary, Odyssean forward guidance can be interpreted as a "lower-for-long" policy (Eggertson and Woodford, 2003; Haberis et al., 2014), i.e. the central bank's commitment to temporarily deviate from its policy rule and not respond to rising inflation and growth, but instead keep rates close to zero for a longer period.[2] The dynamics under the heuristic switching model have been widely studied (see e.g. Brock and Hommes (1997); Branch (2004) and more recently Hommes and Lustenhouwer (2015)) and the two-type version that we use is empirically appealing and closely resembles the recently estimated switching model

[1] For instance Ferrero and Secchi (2009) show that market expectations of the short term interest rate respond in a significant and consistent way to the unexpected component of the published path of the Reserve Bank of New Zealand's (RBNZ) interest rate projections, even though the adjustment is not complete.

[2] Some argue that central banks do not commit to the future path of interest rates and stress the conditionality of these statements (see e.g. Moessner et al., 2017), however, we considered it a relevant theoretical distinction and do not consider the time-inconsistency in the central banks' action.

with fundamentalists versus naive expectations in Cornea-Madeira et al. (2017).[3] This two-type heterogeneous expectations switching model allows us to study the role of endogenous credibility and the implications for central bank communication policies.

Ultimately, the fraction of households adjusting their own forecasts in response to forward guidance will depend on how close past inflation was to the central bank's target. It is therefore, that this fraction of households can be seen as an endogenous (and thus time-varying) measure of central bank credibility. In fact, a credible policy regime will be characterized by a disconnect between inflation dynamics and inflation expectation similar to Demertzis et al. (2008), as high levels of credibility in our model imply that inflation expectations are centered around the inflation target. Similarly, Bernanke and Posen (2001) discuss how the behavior of survey forecasts relative to the central banks's inflation target provides information about credibility.

Hence, our approach offers a natural way to relax the ad hoc assumption of exogenous credibility and endogenizes the central bank's credibility in a heterogeneous expectations framework with time-varying fractions. Eventually, this assumption leads to heterogeneity in private sector expectations with a time-varying distribution.[4]

The third novelty−to use N-step ahead Euler equation learning−is needed to implement forward guidance under bounded rationality, because this approach ensures that expected future interest rate matter. The learning literature proposed two generalizations of the standard Euler equation learning approach by Evans and Honkapohja (2001), namely N-step Euler equation learning (Ferrero and Secchi, 2011; Branch et al., 2012) and infinite horizon learning (see e.g. Sargent, 1999; Eusepi, 2005; Eusepi and Preston, 2018). Despite being attractive, we show in Appendix J that the infinite horizon approach together with the heuristic switching model does not satisfy asymptotic stationarity properties. Therefore, we focus on the N-step ahead Euler equation concept and provide robustness checks showing that the policy implications do not hinge on the specific value of the forward-looking horizon N, as long as N is not too large.[5]

Arguably, measuring the forward-looking horizon is a challenging task. Thus, we rely on self-reported financial planing horizons from survey data. For instance, according to the Survey of Consumer Finances (SCF) or the Health and Retirement Study (HRS) 60-70% of US households indicate to have planning horizons ranging from a few month to less than five years, while 40% even indicated horizons of no more than one year, when asked directly about their financial planning horizon in (see Dow Jr and Jin, 2013; Rodriguez de Rubio, 2015). Obviously, there might be several reasons for these relatively short planning horizons, including cognitive limitations, the educational level in general, other household characteristics such as age, but also financial frictions, such as

[3] The empirical evidence in favor of the heuristics switching model is compelling. For instance, Branch (2004, 2007) finds that survey data on inflation expectations are consistent with a dynamic choice between statistical predictor functions. Further, Anufriev and Hommes (2012*a*,*b*) fit the heuristics switching model to the data of asset pricing learning-to-forecast experiments (see Hommes, 2011, for a survey of laboratory experiments) and find that already four simple heuristics explain most of the observations. Also, Hommes et al. (2005) argue that laboratory experiments with human subjects are well suited to discipline the class of individual heuristics that boundedly rational subjects may use in their decision making process. Lastly, Assenza et al. (2013) use the same heuristics switching model as in Anufriev and Hommes (2012*a*,*b*) and find that the simple heterogeneous expectations switching model also fits individual learning and aggregate outcomes in the standard New Keynesian macroeconomic setting.

[4] The empirical evidence of heterogeneity in expectations is numerous. To name a few, Carroll (2003), Mankiw et al. (2003), Pfajfar (2009) and Pfajfar and Santoro (2010) provide empirical support for heterogeneity in expectations using survey data on inflation expectations. Furthermore, Hommes et al. (2005), Adam (2007), Hommes (2011), Pfajfar and Zakelj (n.d.) and Assenza et al. (2013) find evidence for heterogeneity in learning-to-forecast laboratory experiments with human subjects. Importantly, also the distribution of heterogeneity evolves over time in response to economic volatility (see Mankiw et al., 2003) a feature well captured by our heuristic switching model.

[5] We discuss the modeling implications in detail in Appendix D, G.1, G.2 and I.2.

credit constraints, that prevent households from smoothing consumption over longer horizons. In this paper, we focus more on the bounded rationality reason to finite planing horizons, while e.g. McKay et al. (2016) focus on financial friction dimension.

The fourth novelty–learning of the Central Bank–is introduced for two reasons. First, it is more realistic to assume that the central bank cannot observe current state variables, as has been argued in the literature (see Orphanides, 2001; Aoki and Nikolov, 2006; Lubik and Matthes, 2016). The second reason is tractability. Solving the stochastic model with the central bank having rational expectations under both heuristic switching of boundedly rational households and an additional nonlinearity through the ZLB constraint is a highly complex problem and non-tractable. It is important to note that under forward guidance, the central bank's multiple period ahead expectations appear in the model in a nonlinear fashion through the heuristic switching and the ZLB constraint. With our assumption of central bank learning, the problem becomes tractable and moreover fits within our boundedly rational framework.

In this bounded rationality framework, we analyze the effectiveness of both Delphic and Odyssean forward guidance. Both policies are of a particular interest as the central bank's credibility evolves endogenously in our model. Firstly, we find that both policies jointly enlarge the basin of attraction of the targeted steady state and thus increase the likelihood of recovery from a liquidity trap. Different from rational expectations, however, recovery is not ensured and depends on the credibility of the central bank. Monte Carlo simulations support this theoretical result, and moreover, suggest that the lower-for-long policy alone, that is without the publication of the central bank's forecasts, might be even more effective in inducing recovery, yet at the cost of increased average macroeconomic volatility (and thus lower welfare). We attribute this result to what Melosi (2017) calls the *signaling channel* of monetary policy.[6] In our model, some households downward-adjust their inflation and output expectations, if the central bank's outlook is worse than their own. On the other hand, expectations for the nominal interest rate become more expansionary.[7] Ultimately, if the fraction of households doing so (which is our measure of credibility) is still large, the beneficial effects of forward guidance will be dampened.

The paper continues as follows. Section 2 briefly discusses the related literature. In section 3, we present the ingredients of our behavioral New Keynesian DSGE model followed by an analysis of the dynamics under learning in section 4. In section 5, we introduce the ZLB constraint and discuss the model with and without forward guidance announcements. Our numerical results are illustrated in section 6 and, finally, section 7 concludes.

[6] In Melosi (2017), the policy rate signals non-redundant information to firms and hence directly influences their beliefs about macroeconomic developments. Here, this information extraction is not needed, because the central bank provides its forecasts through the Delphic guidance announcements. However, the channel remains the same: If the central bank, for instance, lowers the policy rate it might be interpreted as a pessimistic signal, leading to a downward adjustment of households' expectations. Similarly, in our model, credibility believers are too optimistic in a recession, hence adopting the central bank's realistic but lower expectations leads to lower aggregate inflation expectations.

[7] This effect of forward guidance is broadly in line with empirical evidence, that shows forward guidance to have a considerably impact on private sector expectations (Ferrero and Secchi, 2009; Hubert, 2014, 2015*a*,*b*), Also Campbell et al. (2016) show that the private information content of FOMC's statements (i.e. what the authors interpret as Delphic guidance) influences markets' expectations of short-run interest rates and induces substantial revision of markets' forecasts. Lastly, Arifovic and Petersen (2015) and Ahrens et al. (2016) find strong evidence that a central bank publishing its own projections fosters expectational coordination in laboratory experiments with human subjects.

2 Related literature

This paper contributes to a growing literature evaluating the role of communications policies when agents have incomplete information about the economy and when central banks have imperfect credibility. Eusepi (2010); Eusepi and Preston (2010); Honkapohja and Mitra (2014, 2015) all consider communications policies when the zero lower bound on nominal interest rates is a relevant constraint. For example Eusepi and Preston (2010) show that a transparent central bank, which gives full information about its policy rule, enhances the predictability of the nominal interest rate and thus stabilizes expectations. Similar conclusions are found in Eusepi (2010), but in a nonlinear framework. Eusepi (2005); Orphanides and Williams (2005); Faust and Leeper (2005); Preston (2006); Eusepi and Preston (2010) all study announcements about monetary policy which are Delphic in flavor. Contrary to those papers, however, we allow the central bank to publish its forecasts and to commit to a future path for the nominal interest rate, while the functional form of the policy rule is known.

Although the channel through which forward guidance operates is well understood, this does not hold for its effectiveness. In fact, the canonical NKM highly overestimates the effects of forward guidance (Carlstrom et al., 2015), which became known as the *forward guidance puzzle* (Del Negro et al., 2012). A vast empirical literature, on the other hand, concludes forward guidance to have stimulative and non-trivial, yet not huge effects (see e.g. Raskin, 2013; Kiley, 2016; Del Negro et al., 2012). Eventually, several attempts have been made to solve the puzzle. For instance, McKay et al. (2016) and Caballero and Farhi (2013) use an incomplete market approach. Andrade et al. (2015) argue that it is uncertainty in the nature of the policy, which causes the muted response. Contrary, Cochrane (2013) argues that it is the failure of the New Keynesian theory that causes this result, while García-Schmidt and Woodford (2015) and Gabaix (2016) attribute it to the assumption of perfect foresight or rational expectations (RE). García-Schmidt and Woodford (2015), for instance, allow for different levels of reflection and find that maintaining a low nominal interest rate for longer always has an expansionary effect, but the effect under low levels of reflection is considerably smaller than under prefect foresight. Similarly, Gabaix (2016) proposes a behavioral New Keynesian model in which myopic agents not fully understand future policies and their impact. Lastly, Ferrero and Secchi (2011); Cole (2015) and Honkapohja and Mitra (2015) study the effects of forward guidance under adaptive learning in a representative household framework. We add to this growing literature of bounded rationality by studying the effectiveness of not only Odyssean, but also Delphic guidance in a heterogeneous expectations framework.

Most closely, our paper relates to Ferrero and Secchi (2011), Haberis et al. (2014); Busetti et al. (2017) and Honkapohja and Mitra (2015). Busetti et al. (2017) consider imperfect credibility for stabilization policy after the crisis using the heuristic switching model, in which credibility is measured by how well long-term inflation expectations are anchored (instead of the distribution of current inflation expectations). While this definition is arguably more realistic, it comes at the cost of being less tractable analytically, yet conceptually not too different as credibility in our model is measured by the fraction of households whose inflation expectations are anchored at the target. Also, Busetti et al. (2017) do not consider how forward guidance affects expectations and thus recovery.

Although Haberis et al. (2014) and arguably Ferrero and Secchi (2011) also consider the case with imperfect credibility, these paper together with Honkapohja and Mitra (2015) treat central bank credibility as exogenous. More specifically, Ferrero and Secchi (2011) study the effects central bank publications of macroeconomic projections have on the dynamic properties of an economy in which agents are learning. Similar to our approach, Ferrero and Secchi (2011) also iterate the New Keynesian IS and Philips curve to implement forward guidance, however, their approach differs in multiple dimensions. In particular, a representative household uses recursive least squares learning, but no such informational constraint is imposed on the side of the central bank. Under forward guidance, the agent uses a weighted average of own and central bank predictions with fixed weights. Contrary, our paper uses the heuristic switching model to endogenize the fraction of agents who believe the policy announcements (which can alternatively be interpreted as a weight attached to competing models). Lastly, the authors ignore the ZLB constraint on nominal interest rates, which we believe to be an important nonlinearity in the context of forward guidance.

In a similar vein, Honkapohja and Mitra (2015) examine the global dynamics of New Keynesian model under learning, when the central bank targets either price level or nominal GDP. The fact that both policy regimes make monetary policy history-dependent allows the authors to implement forward guidance announcements into their model. Honkapohja and Mitra (2015) show that when the policy announcements are incorporated into the private agents' learning, the phase diagram of the targeted steady state under price level targeting increases substantially, as compared to inflation targeting. However, the extreme effects found by Honkapohja and Mitra (2015) seem to be (at least partly) driven by the ad hoc assumption that the central bank is perfectly credible in achieving either the nominal GDP or price level target.

A different approach is taken by Haberis et al. (2014), who show that the macroeconomic effects of a transient interest rate peg can be significantly dampened once the peg is believed to be only imperfectly credible. In fact, the authors assume that the central bank reneges on its announcement of keeping nominal interest rates at zero for a finite number of periods. This decision of when to renege on the promise or not is not modeled explicitly, but stochastic with exogenous probability. Ultimately, the private sector's expectations of endogenous variables are a convex combination of the endogenous variables resulting from no peg and those which are consistent with the interest rate peg, both under RE. Quite intuitively, as the likelihood of reneging decreases, the effects of forward guidance tends imitate those backward-explosive dynamics outlined by Carlstrom et al. (2015), while as the likelihood of reneging increases the effects of the peg tend to cease away. Although, our model also allows for imperfect credibility, we endogenize the households' perceived credibility of the central bank, while assuming that the central bank honors its promise as long as expected inflation does not exceed the target by too much. Moreover, we additionally consider the effects of Delphic guidance, besides Odyssean guidance. Lastly, we deviate from the RE hypothesis and also allow for heterogeneity in expectations.

Therefore, we see the main contribution of our paper in shedding more light on the effectiveness of forward guidance when the central bank's credibility itself is endogenous and agents are heterogeneous and boundedly rational. From a more methodological view point, we are−to the best of our knowledge−the first to combine N-step Euler equation learning and the heuristic switching model and use the model to study the role of forward guidance.

3 The Model

3.1 New Keynesian framework

The model features a continuum of heterogeneous household-firms that differ in the way expectations are formed. Since households are boundedly rational, they are not able to fully optimize over an infinite horizon. Instead, we follow Branch et al. (2012) and assume that our boundedly rational households use N-step Euler equation learning, implying that households use the marginal costs versus marginal benefits trade-off of the Euler equation to make decisions given their budget constraint, and given their subjective forecasts of aggregate variables. Further, the model features monopolistic competition in the goods market as well as sticky prices due to price adjustment costs á la Rotemberg (1982). Log-linearizing around a deterministic steady state, the model can be summarized by the following two equations[8]

$$x_t = \tilde{E}_t x_{t+N} - \frac{1}{\sigma}\tilde{E}_t \sum_{j=0}^{N-1} \left(i_{t+j} - \pi_{t+j+1} - \bar{r}\right) + e_t, \tag{1}$$

$$\pi_t = \beta^N \tilde{E}_t \pi_{t+N} + \tilde{E}_t \sum_{j=0}^{N-1} \beta^j \kappa x_{t+j} + u_t, \tag{2}$$

where $\tilde{E}_t$ denotes the heterogeneous expectations operator to be specified below. The IS curve (1) and Philips curve (2) pin down output x_t and inflation π_t, given a nominal interest rate i_t. The term $\bar{r}$ is the steady state real interest rate, given by $\bar{r} = \frac{1}{\beta} - 1$. The variables e_t and u_t represent *iid.* demand (or real interest rate) and cost-push shocks with variances σ_x^2 and σ_π^2, respectively. The parameters β and σ are the households' discount factor and the inverse intertemporal elasticity of substitution, while κ is a composite parameter of price rigidity indicating the slope of the Philips curve.

To close the model, we follow Orphanides (2001) and assume the central bank sets the nominal interest rate according to a simple expected contemporaneous Taylor-type rule to capture the informational problem faced by the central bank. In particular, Orphanides (2001) shows that there is substantial uncertainty around real-time measures of current endogenous variables, such that the central bank in fact reacts to its inflation nowcasts. Denote the nominal interest rate implied by this reaction function as i_t^{mp}, which might differ from the nominal interest rate ultimately set by the central bank, i_t, due to the ZLB constraint and forward guidance. Formally the reaction function is given by

$$i_t^{mp} = \max\{0, \bar{r} + \bar{\pi} + \phi(\pi_{t|t}^{e,cb} - \bar{\pi})\}, \tag{3}$$

where the max-term represents the ZLB constraint and $\pi_{t|t}^{e,cb}$ denotes the central bank's real-time inflation projection, made at the beginning of period t, that is, before endogenous aggregate outcomes are realized. Lastly, $\bar{\pi}$ denotes the central bank's inflation target.

For the *unconventional* policy, we equip the central bank with two additional policy tools. Firstly, the central bank can publish its own j-periods ahead projections for inflation, the nominal interest rate and output (denoted as $\pi_{t+j|t}^{e,cb}, i_{t+j|t}^{e,cb}$ and $x_{t+j|t}^{e,cb}$, respectively), which we interpret as Delphic forward guidance. Secondly, the central bank can announce a future path for the nominal interest rate i_{t+j}, which we regard as Odyssean forward guidance. In other words, Odyssean

[8] See Appendix A for derivation.

guidance is the conditional promise (or commitment) of the central bank to keep nominal interest rates at zero for a prolonged period of time(here q^O periods), as long as the central bank's own projections of next period's inflation do not exceed its target $\bar{\pi}$ too much.[9] Therefore, let $\tilde{\pi}$ be a threshold level of inflation with $\tilde{\pi} > \bar{\pi}$, for which it holds that if $\pi^{e,cb}_{t|t} > \tilde{\pi}$, then the central bank will revert back to its usual reaction function given by Equation (3).[10] Formally, the interest rate policy under Odyssean forward guidance is

$$i_{t+j} = \begin{cases} i^{mp}_{t+j} & \text{if} \quad i^{mp}_{t-j} > 0, \quad \forall j = 1, \dots, q^O \quad \text{or} \quad \pi^{e,cb}_{t+j|t+j} > \tilde{\pi} \\ 0 & \text{else} \end{cases}, \tag{4}$$

with q^O denoting the horizon of Odyssean forward guidance. Note that the interest rate policy (4) is written in a backward-looking fashion, stating that if the ZLB was not binding in the last q^O periods, or it was, but period-$t + j$ inflation nowcast $\pi^{e,cb}_{t+j|t+j}$ exceeds the tolerance level $\tilde{\pi}$, then set $i_{t+j} = i^{mp}_{t+j}$ in any period $t + j$, while if the ZLB is or was binding less than q^O periods ago, then continue with the lower-for-long policy. Regarding the frequency of announcements, we assume that the central bank only once announces its Delphic and/or Odyssean guidance for q^D and q^O periods, respectively, in the period it is first constrained by the ZLB. It then chooses the guidance horizons according to

$$q^f = \begin{cases} q & \text{if} \quad \pi^{e,cb}_{t|t} \leq \bar{\pi} - \phi^{-1}(\bar{r} + \bar{\pi}) \quad \text{but} \quad \pi^{e,cb}_{t-j|t-j} > \bar{\pi} - \phi^{-1}(\bar{r} + \bar{\pi}), \quad \forall j = 1, \cdots, q \\ 0 & \text{else} \end{cases},$$

for $f = O, D$. Announcements can be renewed if the economy still has not recovered after q periods.

3.2 Private sector expectations

At this point, we have to specify how private households' expectations are formed. As mentioned above, our assumption on the timing of expectation formation is that private agents as well as the central bank do not observe current aggregate outcomes (i.e. endogenous variables), but only lagged realizations and form their expectations at the beginning of each period. Also, shocks are not observed. Thus, private agents and the central bank share the same information set.

Private households are assumed to use simple forecasting heuristics to form their expectations about key macroeconomic variables. Further, we let the private agents choose these heuristics endogenously out of a set of forecasting heuristics according to their relative performance in recent past. This idea goes back to Herbert Simon (1984) who proposed to model human decision making as a rational choice between simple forecasting heuristics and has been formalized in the heuristic switching model (HSM) of Brock and Hommes (1997). Let us denote the fraction of agents using a specific forecasting heuristic h out of the set of forecasting heuristic $\mathcal{H}$ at time t by $n_{h,t}$, which

[9] This assumption is in line with the Fed's announcement from December 12, 2012. The relevant part reads: "*[T]he Committee decided to keep the target range for the federal funds rate at 0 to 1/4 percent and currently anticipates that this exceptionally low range for the federal funds rate will be appropriate at least as long as ... inflation between one and two years ahead is projected to be no more than a half percentage point above the Committee's 2 percent longer-run goal, and longer-term inflation expectations continue to be well anchored*". While short term inflation expectations of the credibility believers can be influenced through Delphic guidance statements, their longer term inflation expectations are anchored at the target in our model as long as N exceeds the Delphic guidance horizon q^D.

[10] Generally, it is possible to micro-found this threshold inflation level $\tilde{\pi}$ by formulating the central bank's trade-off between the loss of credibility when continuing the lower-for-long policy due to higher inflation versus the loss of credibility from reneging from the previous promise. King (2010), for instance, shows that both inflation and output gap are optimally higher under commitment, when the economy exits the ZLB.

follows a logistic distribution of the form

$$n_{h,t} = \frac{\exp(bU_{h,t-1})}{\sum_h \exp(bU_{h,t-1})} \quad \text{for } h \in \mathcal{H}, \tag{5}$$

$$U_{h,t-1} = -\sum_{z\in\mathcal{Z}} \left(z_{t-1} - \tilde{E}_{h,t-2}z_{t-1}\right)^2 \quad \text{with } \mathcal{Z} = \{x, \pi, i\}, \tag{6}$$

where $U_{h,t}$ is the performance measure here defined as the negative sum of squared forecast errors. The parameter $b \in [0, \infty)$ is called the intensity of choice and it governs the sensitivity towards forecast errors, i.e. how fast agents switch to the optimal forecasting heuristic. In the special case $b = 0$ agents never switch their strategy such that all fractions will be constant and equal to $\frac{1}{|\mathcal{H}|}$. Contrary, in the other extreme case of $b \to \infty$, all agents will use the same optimal strategy in each period. The latter case is sometimes referred to as *neoclassical limit*, because it represents the highest degree of rationality. Aggregate expectations are then given by the weighted average of individual expectations

$$\tilde{E}_t z_{t+j} = \sum_{h\in\mathcal{H}} n_{h,t}\tilde{E}_{h,t}z_{t+j} \qquad \text{for } z_{t+j} \in \{x_{t+j}, \pi_{t+j}, i_{t+j}\},$$

In the rest of the paper we will confine ourselves to the case with $|\mathcal{H}| = 2$. Precisely, we assume two types of agents, namely *adaptive learners* and *credibility believers*, which can well explain survey data on inflation expectations of professional forecasters in the United States (Cornea-Madeira et al., 2017). Adaptive learners form their expectations for all variables (output, inflation and the nominal interest rate) using an adaptive expectations rule. In fact, this rule can be derived from agents using steady state learning with a constant gain parameter when the exogenous shocks are *iid*.[11] In other words, adaptive learners treat all variables as *iid* processes with an unknown mean, which they try to estimate by least squares. On the other hand, credibility believers fully believe in the central bank's ability to achieve its target of price stability, that is inflation to be at target $\bar{\pi}$, which coincides with the rational expectation equilibrium (REE)[12]. Likewise, credibility believers expect output and nominal interest rate to equal their REE steady state levels $\bar{x}$ and $\bar{i}$, respectively. However−and this is a key novelty of our paper−credibility believers adopt the central bank's projections of these variables in case of forward guidance.

In the following, we formally describe private households' forecasting heuristics. Let the expectations of all households using forecasting heuristic h for variable $z_{t+j} = x_{t+j}, \pi_{t+j}$ with $j \geq 1$ at time t be denoted as $\tilde{E}_{h,t}z_{t+j}$. We then have a credibility believer's ($h = 1$) expectations given by

$$\tilde{E}_{1,t}z_{t+j} = \begin{cases} z^{e,cb}_{t+j|t}, & \forall j = 0, \ldots, q^D \\ \bar{z}, & \forall j = q^D + 1, \ldots, N \end{cases}, \tag{7}$$

and those of the adaptive learner ($h = 2$)

$$\tilde{E}_{2,t}z_{t+j} = \tilde{E}_{2,t-1}z_t + \omega\big(z_{t-1} - \tilde{E}_{2,t-1}z_t\big), \quad \forall j = 1, \ldots, N, \tag{8}$$

respectively, where the first row in Equation (7) corresponds to the central bank's Delphic guidance, i.e. the announcement of its own projections $z^{e,cb}_{t+j|t}$ for the next q^D periods. How the central bank

[11] See Evans et al. (2008) for a more detailed discussion.

[12] See Section B in the Appendix for detailed calculations.

obtains its projections will be specified below. Moreover, note that in the absence of Delphic forward guidance ($q^D = 0$) or at longer horizons ($j > q^D$) credibility believers expectations are anchored at the target, i.e. the REE value. The parameter $\omega \in (0, 1]$ in the adaptive expectations rule (8) is the gain parameter, which governs the speed with which forecasts are updated. For $\omega = 1$ these expectations reduce to what is generally called *naive* expectations.

Next, recognize that, due to the assumption of N-step Euler equation learning, expectations about the nominal interest rate show up explicitly in the IS curve (1). For tractability, we assume that households are aware of the ZLB constraint and, moreover, know the functional form of central bank's reaction function (3). Using their own inflation expectations, private households then determine the nominal interest rate that is consistent with their beliefs.[13] Formally, the nominal interest rate expectations of credibility believers $\tilde{E}_{1,t} i_{t+j}$ for $j \geq 1$ are given by

$$\tilde{E}_{1,t} i_{t+j} = \begin{cases} i^{e,cb}_{t+j|t} = \max\{0, \bar{r} + \bar{\pi} + \phi(\pi^{e,cb}_{t+j|t} - \bar{\pi})\}, & \forall j = 0 \dots, q^D \\ \bar{\pi} + \bar{r}, & \forall j = q^D + 1, \dots, N \end{cases}, \tag{9}$$

under Delphic guidance only, while they become

$$\tilde{E}_{1,t} i_{t+j} = \begin{cases} 0 & \forall j = 0 \dots, q^O \\ \bar{\pi} + \bar{r}, & \forall j = q^O + 1, \dots, N \end{cases}, \tag{10}$$

if the central bank announces a lower-for-long policy (i.e. either Odyssean guidance only or both). That is, if the central bank commits to temporarily deviate from its policy rule by keeping interest rates at zero, credibility believers will fully incorporate this announcement about future interest rates in their expectation formation. Hence, credibility believers expectations need not to be consistent with their own inflation and/or output expectations due to the central bank's explicit commitment under Odyssean guidance. In the absence of any such policies (i.e. $q^O = q^D = 0$) or at longer horizons, credibility believers expect the nominal interest rate to be at its target $\bar{r} + \bar{\pi}$.

Contrary, the expectations of adaptive learners are given by

$$\tilde{E}_{2,t} i_{t+j} = \max\{0, \bar{r} + \bar{\pi} + \phi(\tilde{E}_{2,t}\pi_{t+j} - \bar{\pi})\}, \quad \forall j = 1, \dots, N. \tag{11}$$

3.3 Central bank learning

Lastly, we specify how the central bank forms its forecasts. It seems to be a reasonable assumption that the central bank has the most sophisticated forecasting model, as central banks generally devote large amounts of resources on forecasting. For this reason, we assume that the central bank uses a parsimonious first-order bivariate VAR model, whose coefficients are updated each period, to estimate future inflation and output. Formally, the central bank estimates the following VAR model

$$\begin{aligned} x^{e,cb}_{t|t} &= a_{10} + a_{11} x_{t-1} + a_{12} \pi_{t-1}, \\ \pi^{e,cb}_{t|t} &= a_{20} + a_{21} x_{t-1} + a_{22} \pi_{t-1}, \end{aligned}$$

[13] This assumption, which only affects the expectations of the adaptive learners, is made to limit the size of the state space and thus keep the model analytically tractable. We provide some robustness checks in the Appendix I.3 to convince the reader, that our results are qualitatively robust to this assumption, and, moreover, confirm earlier work by e.g. Eusepi (2005); Eusepi and Preston (2010) that knowing the policy rule of the central bank has a stabilizing effect.

which can be simplified to

$$y^{e,cb}_{t|t} = A'w_{t-1}, \tag{12}$$

where $y^{e,cb}_{t,t} = [x^{e,cb}_{t|t}, \pi^{e,cb}_{t|t}]'$ and $w_{t-1} = [1, x_{t-1}, \pi_{t-1}]'$. In the learning literature Equation (12) is also called the Perceived Law of Motion (PLM). Note that this model is correctly specified in the absence of heuristic switching, while it is miss-specified under switching. The matrix A is the corresponding 3×2 coefficient matrix, whose elements are updated each period using the following recursive least squares (RLS) algorithm

$$A_t = A_{t-1} + \gamma_t R_t^{-1} w_{t-1}\big(y_t - A'_{t-1}w_{t-1}\big)', \tag{13}$$

$$R_t = R_{t-1} + \gamma_t\big(w_{t-1}w'_{t-1} - R_{t-1}\big), \tag{14}$$

where R_t is a moment matrix and γ_t is the gain parameter, which is set $\gamma_t = \frac{1}{t}$ under recursive least squares. We follow the literature and assume anticipated utility behavior of the central bank. That is, the central bank believes that the parameter estimates will remain unchanged in the future and does not take into account the fact that it is likely to revise them subsequently. Hence, the central bank's behavior deviates from full rationality as the bank does not take into account the effects its decisions have on future learning and it ignores the period-by-period model misspecification (Sargent, 1999). Moreover, the heterogeneity in private sector's expectation creates time-variation in the coefficients of the economic model which the central bank ignores by using a time-invariant PLM.[14] In doing so, the central bank effectively assumes homogeneity in agent expectations. Thus, the misspecified PLM can only converge to its REE, if the fractions of households converge.

To specify the central bank's own projections, we split the coefficient matrix A into a vector of constants A_0 and a 2×2 matrix A_1. Then, under the assumptions made above, the central bank's j-periods ahead forecasts made in the beginning of period t can be summarized by

$$y^{e,cb}_{t+j|t} = \big(I + A_{1,t-1} + A^2_{1,t-1} + \ldots + A^{j-1}_{1,t-1}\big)A_{0,t-1} + A^j_{1,t-1}y_{t-1}, \tag{15}$$

where I denotes a 2×2 identity matrix. These forecasts can be published by the central bank to influence private households' expectations. However, the fact that the central bank has imperfect knowledge about the structure of the economy−i.e. uses a misspecified VAR model plus the imprecision in its parameter estimates−may lead to policy mistakes that affect the performance of the central bank as well as its credibility (Aoki and Nikolov, 2006). In our framework, the central bank may publish its forecasts only once the ZLB is reached, but not in normal times.

4 Dynamics under learning

4.1 Existence of a steady state

Before turning to the E-stability analysis of the central bank's VAR learning, we will briefly state the stochastic dynamic system and discuss its deterministic steady state(s). For this purpose, it is helpful to first define the difference in fractions as $m_{t+1} \equiv n_{1,t+1} - n_{2,t+1}$. Evidently, the

[14] We restrict the PLM to be time-invariant, assuming the central bank is not aware of the time-variation in the distribution of private sector expectations. We relax this assumption in Section 6.2.2, however, such that the matrix A_t in Equation (12) is allowed to vary over time. This is, for instance, captured by constant gain learning under which agents assign a higher weight to the most recent information by discounting past data, or by Kalman filtering.

term m_{t+1} equals 1 when all households are credibility believers, and -1 when all households are adaptive learners. Similar to the fraction of credibility believers, we can interpret this expression as an endogenous measure of the central bank's credibility, because a higher m_{t+1} indicates higher confidence of the population in the central bank achieving its targets. Henceforth, we will use the term fractions and difference in fractions interchangeably for m_{t+1}.

To simplify the E-stability analysis of the central bank's VAR model we set the gain parameter ω in the adaptive learner's forecasting heuristic to $\omega = 1$. Hence, we consider the case with naive households. Note, that this assumption does not alter the steady state of the system, but significantly facilitates the analysis. Under this assumption, the stochastic dynamic system described by Equations (1) - (12) in normal times (i.e. without binding ZLB constraint and without forward guidance announcements) can be written as

$$
\begin{aligned}
x_t =& \Big[\frac{1-\beta}{\kappa}\big(\frac{1+m_t}{2}\big) + \frac{(N-1)\phi - N}{\sigma}\big(\frac{1-m_t}{2}\big) + \frac{\phi}{\sigma}\Big]\bar{\pi} \\
&+ \big(\frac{1-m_t}{2}\big)x_{t-1} - \Big[\frac{(N-1)\phi - N}{\sigma}\big(\frac{1-m_t}{2}\big)\Big]\pi_{t-1} - \frac{\phi}{\sigma}\pi^{e,cb}_{t|t}, \qquad (16)\\
\pi_t =& \Big[\frac{1+m_t}{2} + \frac{\kappa\big[(N-1)\phi - N\big]}{\sigma}\big(\frac{1-m_t}{2}\big) + \frac{\kappa\phi}{\sigma}\Big]\bar{\pi} \\
&+ \kappa\Big(\frac{1-\beta^N}{1-\beta}\Big)\frac{1-m_t}{2}x_{t-1} + \Big[\Big(\beta^N - \frac{\kappa\big[(N-1)\phi - N\big]}{\sigma}\Big)\frac{1-m_t}{2}\Big]\pi_{t-1} - \frac{\kappa\phi}{\sigma}\pi^{e,cb}_{t|t}, \qquad (17)
\end{aligned}
$$

after substituting for the policy rule i_t and private sector expectations $\tilde{E}_{t+j}$, where we omitted the random disturbance terms for simplicity. This can be done without loss of generality, because shocks are assumed to be stationary (or even *iid*) and not observed by agents and therefore do not affect the mapping. Note that the central bank's inflation expectation $\pi^{e,cb}_{t|t}$ enter the dynamic model only indirectly through the nominal interest rate rule. Writing the above system compactly in matrix notation, we get

$$
y_t \equiv \begin{pmatrix} x_t \\ \pi_t \end{pmatrix} = \Lambda_0(m_t) + \Lambda_1(m_t)y_{t-1} + \Lambda_2 y^{e,cb}_{t|t}, \qquad (18)
$$

where the matrices $\Lambda_0(\cdot), \Lambda_1(\cdot)$ and Λ_2 are given by

$$
\Lambda_0(m_t) = \begin{pmatrix} \Big[\frac{1-\beta}{\kappa}\big(\frac{1+m_t}{2}\big) + \frac{(N-1)\phi-N}{\sigma}\big(\frac{1-m_t}{2}\big) + \frac{\phi}{\sigma}\Big]\bar{\pi} \\ \Big[\frac{1+m_t}{2} + \frac{\kappa\big[(N-1)\phi-N\big]}{\sigma}\big(\frac{1-m_t}{2}\big) + \frac{\kappa\phi}{\sigma}\Big]\bar{\pi} \end{pmatrix},
$$

$$
\Lambda_1(m_t) = \begin{pmatrix} \frac{1-m_t}{2} & -\frac{(N-1)\phi-N}{\sigma}\big(\frac{1-m_t}{2}\big) \\ \kappa\Big(\frac{1-\beta^N}{1-\beta}\Big)\frac{1-m_t}{2} & \Big(\beta^N - \frac{\kappa\big[(N-1)\phi-N\big]}{\sigma}\Big)\frac{1-m_t}{2} \end{pmatrix} \quad \text{and} \quad \Lambda_2 = \begin{pmatrix} 0 & -\frac{\phi}{\sigma} \\ 0 & -\frac{\kappa\phi}{\sigma} \end{pmatrix}.
$$

Regarding the steady states of the system, we can postulate the following result.

Proposition 1 (Target steady state). *The dynamic system in Equation* (18) *has a steady state with* $\pi^* = \bar{\pi}, x^* = \bar{x}, i^* = \bar{r} + \bar{\pi}$ *and* $m^* = 0$.

The proof is given in the Appendix C. In the following we refer to this as the "target" steady state. Other steady states as well as (quasi) period cycles may exist in certain parameter regions, however, the analysis exceeds the scope of this paper, as we are primarily interested in the dynamics that arise due to the ZLB constraint and forward guidance.[15]

The target steady state depends on two interacting learning dynamics, the HSM and the central bank learning. Abstracting from the nonlinearity that arises from the central bank learning, the law of motion of the model economy without binding ZLB constraint, i.e. equation (18), is a convex combination of the constant predictors stemming from the credibility believers and a system of first order difference equations stemming from naive households, with the specific weight attached to each of them being governed by the different fractions. Therefore, the lower m the larger will be the impact of the lagged variables and hence the larger will be the likelihood of the economic system to be unstable. This observation follows from an eigenvalue analysis of the dynamic system. Four of the six eigenvalues of the Jacobian are zero, when evaluated at the target steady state. The remaining two eigenvalues of the reduced form Jacobian are also zero, if all private households are credibility believers and the system becomes degenerate. However, the eigenvalues increase (in absolute value) in the fraction of naive households. See Appendix D for detailed calculations.

In the following section, we discuss the expectational stability properties of the central bank's learning. To anticipate the results, we will see that, if the fractions of private households in the economy using a specific heuristic are fixed, the system under learning is stable. However, allowing the fractions to be time-varying may prevent the coefficients to converge.

4.2 Expectational stability

In this section, we derive the modified E-stability conditions for the central bank's VAR learning.[16] We start by substituting the central bank's PLM (12) into (18) to find the Actual Law of Motion (ALM):

$$\begin{aligned} y_t =& \Lambda_0(m_t) + \Lambda_1(m_t)y_{t-1} + \Lambda_2\left[A_0 + A_1 y_{t-1}\right] \\ =& \Lambda_0(m_t) + \Lambda_2 A_0 + \left[\Lambda_1(m_t) + \Lambda_2 A_1\right] y_{t-1}, \end{aligned} \tag{19}$$

The ALM can be seen as describing the stochastic process followed by the economy if the central bank's forecasts are made under the fixed rule given by the PLM (12). In other words, given both the central bank's forecasts and those of the private sector, the economy attains a temporary equilibrium. The mapping, T, from the PLM to the ALM is given by

$$T\begin{pmatrix} A_0 \\ A_1 \end{pmatrix} = \begin{pmatrix} \Lambda_0 + \Lambda_2 A_0 \\ \Lambda_1 + \Lambda_2 A_1 \end{pmatrix}, \tag{20}$$

which has a unique fixed point for any given value of the difference in fractions $m_t \equiv m$

$$A_0^* = (I - \Lambda_2)^{-1}\Lambda_0(m), \tag{21}$$

$$A_1^* = (I - \Lambda_2)^{-1}\Lambda_1(m), \tag{22}$$

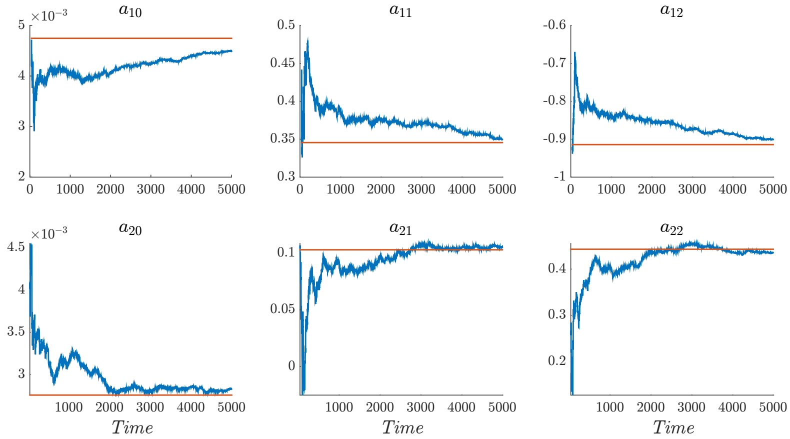
This result is summarized in the following Proposition.

Proposition 2 (Fixed point of the T-map). *The mapping, T, from the PLM* (12) *to the ALM* (19) *has a unique fixed point that depends on the difference in fractions* $m \in [-1, 1]$.

[15] See Hommes and Lustenhouwer (2015) for a detailed analysis of possible steady states in a model with Euler equation learning, i.e. $N = 1$.

[16] Appendix E provides a more detailed analysis.

Figure 1: Convergence of the central bank's learning algorithm



Note: The figure suggests the convergence of the central bank's estimated VAR(1) coefficients given by (12) under constant and equal fractions. The red line depicts the corresponding fixed point in equations (21) and (22) for $m = 0$, while the blue line gives it's time-t estimate of the coefficients following from (13) and (14).

Whether the estimated coefficients converge to their respective fixed point is determined by the following set of ordinary differential equations

$$\frac{dA_0}{d\tau} \equiv T_0(A_0) - A_0 = \Lambda_0 + \Lambda_2 A_0 - A_0,$$

and

$$\frac{dA_1}{d\tau} \equiv T_1(A_1) - A_1 = \Lambda_1 + \Lambda_2 A_1 - A_1,$$

According to the E-stability principle (see Evans and Honkapohja, 2001, p. 39), convergence under least squares learning requires the eigenvalues of $\Lambda_2 - I$ to have negative real parts. Since the eigenvalues are -1 and $-(1 + \kappa\phi/\sigma)$, with both $\kappa, \sigma > 0$, this is indeed the case. Thus, under fixed fractions, Equations (21) and (22) depict a unique and (asymptotically) E-stable fixed point of the T-map. This E-stability result is summarized in Proposition 3.

Proposition 3 (**E-stability**). *Assume the fractions of households are constant, i.e. $m_t = m$, and that the ZLB is not binding, then the central bank learns the true model given in* (19)*.*

Assuming the difference in fractions m_t is constant and equal to zero, i.e. $m = 0$, Figure 1 illustrates the convergence of the VAR coefficients to their respective fixed points, where we use our benchmark calibration as outlined in Table 1. Under time-varying fractions, however, the true model features time-varying coefficients and therefore, convergence of the central bank's estimated coefficients may not occur. However, asymptotically, a so-called restricted perception equilibrium (RPE) can be attained, in which the central bank's beliefs are confirmed. We discuss the implications of the RPE below in Section 6.2.1.

5 The ZLB and the role of forward guidance

Given our specification of the interest rate rule (3), the central bank will lower the nominal interest rate once its inflation nowcast $\pi^{e,cb}_{t|t}$ to be lower than the inflation target $\bar{\pi}$. Moreover, if the central bank expects inflation to be lower than $\bar{\pi} - \phi^{-1}(\bar{\pi} + \bar{r})$, the ZLB constraint will be binding and the central bank's conventional policy measures to stimulate the economy are exhausted. In this section, we therefore analyze the effects of central bank's unconventional policy, in particular the commitment to a lower-for-long policy and the publication of macroeconomic projections, to lift the economy out of the liquidity trap. Especially the first policy tool leads to mixed theoretical results. When agents have perfect foresight, the effects of forward guidance are highly overestimated (see e.g. Carlstrom et al., 2015; McKay et al., 2016)−named the "forward guidance puzzle" by Del Negro et al. (2012)−which can be partly attributed to the fact that the central bank is assumed to enjoy perfect credibility as well as perfect foresight (Haberis et al., 2014; García-Schmidt and Woodford, 2015). Contrary, when agents are learning (and thus are purely backward-looking) the central bank has no credibility and announcement of such policies should have no immediate effects at the time of the announcement. Thus, both assumptions on how agents form their expectation can be interpreted as limiting cases. To fill this gap, this section proposes a model in which agents are forward-looking for a finite number of periods, and in which central bank credibility evolves endogenously depending on how well the central bank performed in achieving price stability.

5.1 The model without forward guidance

As mentioned above the ZLB constraint on nominal interest rates will be binding if the central bank's expectations for current inflation are pessimistic enough, i.e. satisfy

$$\pi_{t|t}^{e,cb} \leq \bar{\pi} - \phi^{-1}(\bar{r} + \bar{\pi}), \tag{23}$$

In the following we will refer to all (x_{t-1}, π_{t-1}) combinations for which $\pi_{t|t}^{e,cb}$ satisfies (23) as the "zero lower bound region" or simply "ZLB region". Consequently, the nominal interest rate i_t will be set to zero. To derive the economic model in the ZLB region without forward guidance, however, we further have to distinguish between two cases that depend on what households using the naive heuristic expect. To be precise, the naive households' expectation rule (11) states that naive agents either expect the ZLB to be binding in the next N periods, or they expect the economy leaving the ZLB already by next period. Which of the two scenarios takes place crucially depends on period $t-1$ inflation. In fact, if π_{t-1} is sufficiently low, i.e. satisfies

$$\pi_{t-1} \leq \bar{\pi} - \phi^{-1}(\bar{\pi} + \bar{r}), \tag{24}$$

then naive agents expect the ZLB to be binding for the next N periods. It is important to see that condition (24)−other than (23)−depends solely on period $t-1$ inflation. The economic model in the ZLB region *without* forward guidance is then given by

$$\begin{aligned} x_t =& \Bigg[\left(\frac{1-\beta}{\kappa} + \frac{1}{\sigma}\right)\frac{1+m_t}{2} + \frac{(N-1)(\phi-1)\mathbb{1}_n}{\sigma}\big(\frac{1-m_t}{2}\big)\Bigg]\bar{\pi} + \big(\frac{1-m_t}{2}\big)x_{t-1} \\ &- \Bigg[\frac{(N-1)\mathbb{1}_n\phi - N}{\sigma}\big(\frac{1-m_t}{2}\big)\Bigg]\pi_{t-1} + \frac{\mathbb{1}_n\bar{r}}{\sigma} + \frac{(1-\mathbb{1}_n)\bar{r}}{\sigma}\big[N - \frac{(N-1)(1+m_t)}{2}\big], \end{aligned} \tag{25}$$

$$\begin{aligned} \pi_t =& \Bigg[\left(1+\frac{\kappa}{\sigma}\right)\frac{1+m_t}{2} + \frac{\kappa(N-1)(\phi-1)\mathbb{1}_n}{\sigma}\big(\frac{1-m_t}{2}\big)\Bigg]\bar{\pi} + \kappa\left(\frac{1-\beta^N}{1-\beta}\right)\frac{1-m_t}{2}x_{t-1} \\ &+ \Bigg[\left(\beta^N - \frac{\kappa\big[(N-1)\mathbb{1}_n\phi - N\big]}{\sigma}\right)\frac{1-m_t}{2}\Bigg]\pi_{t-1} + \frac{\kappa\mathbb{1}_n\bar{r}}{\phi} + \frac{(1-\mathbb{1}_n)\bar{r}}{\sigma}\big[N - \frac{(N-1)(1+m_t)}{2}\big], \end{aligned} \tag{26}$$

with $\mathbb{1}_n$ being an indicator function taking on value 0 if (24) holds and 1 otherwise.

Equations (25) and (26) illustrate that−conditional on (23) to hold−the central bank's own forecasts do not enter the model equations, and hence there will also be no feedback from central bank's forecasts to the true realizations. It is therefore that we are not able to discuss E-stability properties in this particular case. In fact, as long as inflation satisfies (23), equations (25) and (26) represent an independent data generating process that is estimated by the central bank through RLS. However, given time-varying fractions, the central bank's model misspecification persists.

Next, we consider the steady states of the dynamic system in the ZLB region. To clarify the role of the HSM, we will distinguish between constant fractions, i.e. $m_t = m, \forall t$ and time-varying fractions. The crucial difference is that, while under time-varying fractions the steady state difference in fractions m^* cannot exceed zero, this is not a necessary condition under fixed fractions. To see this, recognize that adaptive learners (and thus naive households) always predict correctly in steady state and since, under time-varying fractions, households choose the better performing heuristic(s), it must be that $m^* \leq 0$. In the Appendix F, we derive the steady state values of output and inflation (x^*_{zlb}, π^*_{zlb}) under the ZLB, given any fixed fraction m. Intuitively, the fraction of credibility believers must be sufficiently small for the steady state to lie inside the ZLB region, because credibility believers push the system out of the ZLB region. Under time-varying fractions, this upper threshold level for the difference in fractions m translates into a lower threshold for the intensity of choice b. Again, the intuition is that, given a large intensity of choice, households switch to the better performing heuristic, that is adaptive (or naive). Thus, if b is too low, the steady state fraction of credibility believers will be too large to support a ZLB steady state. Lastly, we find that a sufficient condition for the ZLB steady state to exist is $\phi > 1$ (see also Benhabib et al., 2001*a*,*b*). These results are collected in the following Proposition. Its proof is given in the Appendix F.

Proposition 4 (**ZLB steady state**). *Suppose $\phi > 1$. Then a ZLB steady state exists*

i. for constant fractions, if the fraction of naive agents is sufficiently large, i.e. $m \in [-1, \tilde{m}_1]$, where $\tilde{m}_1$ is given by

$$\tilde{m}_1 = \frac{-\alpha_2 + \sqrt{\alpha_2^2 - 4\alpha_1\alpha_3}}{2\alpha_1}$$

and steady state values $(x^*, \pi^*, i^*) = (x^*_{zlb}, \pi^*_{zlb}, 0)$, with $\alpha_1, \alpha_2, \alpha_3$ and (x^*_{zlb}, π^*_{zlb}) given in Appendix F.

ii. for time-varying fractions, if the intensity of choice is high enough, i.e. $b \in (\underline{b}, \infty)$ where $\underline{b}$ is given by

$$\underline{b} = \frac{2}{\Delta U^*} \tanh^{-1}(\tilde{m}_1).$$

Steady state values are then given by $(x^*, \pi^*, i^*, m^*) = (x^*_{zlb}, \pi^*_{zlb}, 0, m^*_{zlb})$.

In the Appendix G.1 we derive an expression for the eigenvalues of the ZLB steady state under switching. Unfortunately, the expression is too involved to provide analytical insights, even after some simplifying assumptions. However, a numerical analysis shows that, for our baseline calibration of Table 1, at least one of the two eigenvalue is outside the unit circle, so that the ZLB steady state is a saddle point. Thus, the ZLB creates the possibilities of deflationary spirals due to the backward-looking expectations as in Friedman (1968).

5.2 The model with forward guidance

In this section, we allow the central bank to publish its own forecasts for inflation and output and to commit to a future interest rate path. Specifically, for the latter policy, we consider the commitment to keep nominal interest rates at zero for a prolonged period of time. Let the horizon for both policies be $q^O = q^D = q$ periods. However, both these communicational policy measures are only picked up by the credibility believers.[17] To illustrate the role of central bank Delphic forward guidance in our behavioral model, consider the stochastic time series simulation presented in Figure 2. In period 93 inflation was significantly below the central bank's target $\bar{\pi}$, such that the central bank's one-period ahead inflation projections demand low nominal interest rates. Since, however, the central bank is constraint by the ZLB, it cannot lower rates sufficiently to provide the needed stimulus, i.e. condition (23) holds. In this particular scenario, backward-looking households, that is those with adaptive (dash-dotted) or naive expectations (dotted), are pessimistic about the future development of inflation. On the other hand, credibility believers (solid with crosses) expect monetary policy to quickly lift inflation back to the central bank's target. While the latter arguably appears overly optimistic, the expectations of purely backward-looking households might be too pessimistic. However, if the fractions of adaptive learners is large enough, their deflationary expectations can potentially be self-fulfilling by depressing current economic outcomes and thus giving further momentum to adaptive learner's beliefs under the dynamics of the heuristic switching model. Ultimately, this can induce a liquidity trap due to the ZLB constraint faced by the central bank. To prevent this possibility, the central bank can try to influence the public's expectations by publishing it's own forecasts. In particular, the mechanism goes via the beliefs of the credibility believers, who adopt the central bank's forecasts in case of Delphic guidance. Contrary to the credibility believers, the central bank expects inflation to slowly mean-revert to its long run equilibrium (red-dashed); a path for inflation which is arguably more reasonable.[18] Ultimately, if the central bank's projections turn out to the best ex post, trust in the central bank will increase, thereby stabilizing the system, as households switch away from the destabilizing adaptive expectations rule. If, on the other hand, credibility was initially high, adopting the central bank's projections as own forecasts de facto deterioration in aggregate expectations, thus potentially prolonging the recessionary state.

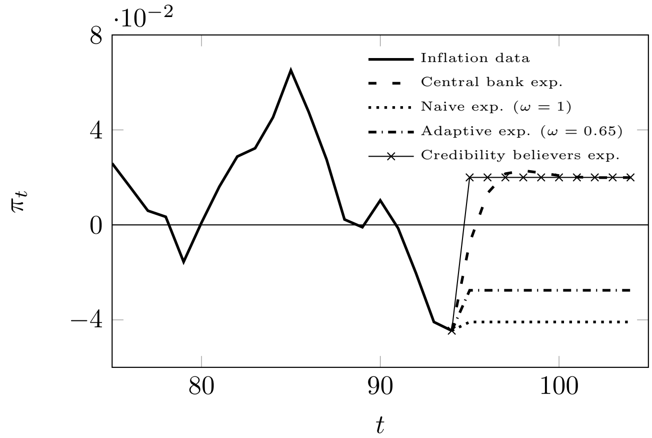
For simplicity, assume that the forward guidance horizon equals the forward-looking horizon of the agents, i.e. $q = N$.[19] Then, the economic system in the ZLB region and *with* forward guidance

[17] We abstract from any noise in the communication channel of the central bank and leave this discussion for future research. There is no doubt that including both sender as well as receiver noise into the framework could generate interesting insights. See Myatt and Wallace (2014) as an example.

[18] Looking at the relative root mean squared nowcast (forecast) errors of the central bank versus adaptive learners, i.e. $RMSE^{cb}/(RMSE^{cb} + RMSE^{AL})$ where $RMSE^j = \sqrt{\frac{\sum_i (E_t z_t - z_t)^2}{n}}$, in a total of 10,000 Monte Carlo simulation, we find the relative root mean squared nowcast error to be 33.5% (39.67%) for output (inflation), thus indicating that the central bank's VAR leads to smaller nowcast errors (on average). Similarly, for one period ahead forecasts the numbers are 43.43% and 46.02%, respectively.

[19] We relax this assumption in the Appendix I.2.

Figure 2: The role of Delphic forward guidance



Note: The figure shows the paths of the conditional inflation expectations (annualized) for the different economic agents after hitting the ZLB in a stochastic time series simulation. We use the benchmark calibration of Table 1. Adaptive and naive expectations imply pessimistic beliefs about the future path of inflation, while those of credibility believers in the absence of forward guidance are overly optimistic in that they expect inflation to be back at the target by next period. By publishing its own, arguably superior, inflation projections the central bank can coordinate public inflation expectations on the most likely outcome.

is formally given by

$$\begin{aligned} x_t =& \frac{1-m_t}{2}x_{t-1} - \frac{(N-1)\mathbb{1}_n\phi - N}{\sigma}(\frac{1-m_t}{2})\pi_{t-1} + \frac{(N-1)(\phi-1)\mathbb{1}_n}{\sigma}(\frac{1-m_t}{2})\bar{\pi} \\ &+ \frac{1+m_t}{2}x^{e,cb}_{t+N|t} + \frac{1}{\sigma}(\frac{1+m_t}{2})\sum_{j=1}^{N}\pi^{e,cb}_{t+j|t} + \frac{(1-\mathbb{1}_n)N\bar{r}}{\sigma} + \frac{\mathbb{1}_n}{\sigma}\left[N - \frac{(N-1)(1-m_t)}{2}\right]\bar{r} \end{aligned} \tag{27}$$

$$\begin{aligned} \pi_t =& \kappa\left(\frac{1-\beta^N}{1-\beta}\right)(\frac{1-m_t}{2})x_{t-1} + \frac{\kappa\mathbb{1}_n(N-1)(\phi-1)}{\sigma}(\frac{1-m_t}{2})\bar{\pi} \\ &+\left(\beta^N - \frac{\kappa[(N-1)\mathbb{1}_n\phi - N]}{\sigma}\right)(\frac{1-m_t}{2})\pi_{t-1} + \kappa(\frac{1+m_t}{2})\left(\sum_{j=1}^{N-1}\beta^j x^{e,cb}_{t+j|t} + x^{e,cb}_{t+N|t}\right) \\ &+ \frac{1+m_t}{2}\left[\frac{\kappa}{\sigma}\sum_{j=1}^{N-1}\pi^{e,cb}_{t+j|t} + (\beta^N + \frac{\kappa}{\sigma})\pi^{e,cb}_{t+N|t}\right] + \frac{(1-\mathbb{1}_n)N\kappa\bar{r}}{\sigma} + \frac{\kappa\mathbb{1}_n}{\sigma}\left[N - \frac{(N-1)(1-m_t)}{2}\right]\bar{r} \end{aligned} \tag{28}$$

which is conditional on the ZLB constraint (23) to be binding. Moreover, the condition (24) for naive households' expectations is again given by the indicator function $\mathbb{1}_n$. Rewriting equations (27) and (28) in matrix notation, we get the following expectation-feedback system

$$y_t = \Lambda_0(m_t) + \Lambda_1(m_t)\begin{pmatrix} x_{t-1} \\ \pi_{t-1} \end{pmatrix} + \Lambda_2(m_t)\begin{pmatrix} x^{e,cb}_{t+1|t} \\ \pi^{e,cb}_{t+1|t} \end{pmatrix} \ldots + \Lambda_{N+1}(m_t)\begin{pmatrix} x^{e,cb}_{t+N|t} \\ \pi^{e,cb}_{t+N|t} \end{pmatrix} \tag{29}$$

Hence, Delphic guidance extends the system (25)-(26) under no guidance by $N-1$ additional terms. The coefficient matrices Λ_j are given below. Note, that these matrices generally depend on the difference in fractions m_t.

$$\Lambda_0(m_t) = \begin{pmatrix} \frac{(N-1)(\phi-1)\mathbb{1}_n}{\sigma}\left(\frac{1-m_t}{2}\right)\bar{\pi} + \frac{(1-\mathbb{1}_n)N\bar{r}}{\sigma} + \frac{\mathbb{1}_n}{\sigma}\left[N - \frac{(N-1)(1-m_t)}{2}\right]\bar{r} \\ \frac{\kappa\mathbb{1}_n(N-1)(\phi-1)}{\sigma}\left(\frac{1-m_t}{2}\right)\bar{\pi} + \frac{(1-\mathbb{1}_n)N\kappa\bar{r}}{\sigma} + \frac{\kappa\mathbb{1}_n}{\sigma}\left[N - \frac{(N-1)(1-m_t)}{2}\right]\bar{r} \end{pmatrix}$$

$$\Lambda_1(m_t) = \begin{pmatrix} \frac{1-m_t}{2} & -\frac{(N-1)\mathbb{1}_n\phi-N}{\sigma}\left(\frac{1-m_t}{2}\right) \\ \kappa\left(\frac{2-\beta-\beta^N}{1-\beta}\right)\frac{1-m_t}{2} & \left[\beta^N - \frac{\kappa[(N-1)\mathbb{1}_n\phi-N]}{\sigma}\right]\frac{1-m_t}{2} \end{pmatrix}$$

$$\Lambda_j(m_t) = \begin{pmatrix} 0 & \frac{1+m_t}{2\sigma} \\ \kappa\beta^{j-1}\left(\frac{1+m_t}{2}\right) & \frac{\kappa}{\sigma}\left(\frac{1+m_t}{2}\right) \end{pmatrix}, \forall j = 2, ..., N$$

$$\Lambda_{N+1}(m_t) = \begin{pmatrix} \frac{1+m_t}{2} & \frac{1+m_t}{2\sigma} \\ \kappa\left(\frac{1+m_t}{2}\right) & \left(\beta^N + \frac{\kappa}{\sigma}\right)\frac{1+m_t}{2} \end{pmatrix}$$

Similar as in the exercise above we find that the matrices Λ_j are bounded given that m_t is bounded. In fact, all matrices Λ_j for $j > 1$ reduce to the null matrix if all private households are naive (i.e. $m_t = -1$) and we are back in the case of Section 5.1 without forward guidance. Intuitively, the strength of the central bank's communication channel ultimately depends on the fraction of credibility believers: if no household believes the central bank's announcements, the effects of forward guidance are nil.

Next, let us iterate the central bank's PLM (12) forward and rewrite it as

$$y^{e,cb}_{t+j|t} = \bar{A}_0 + A_1^{j+1}(y_{t-1} - \bar{A}_0), \quad j = 0, \dots q \quad \text{with} \quad \bar{A}_0 = (I - A_1)^{-1}A_0 \tag{30}$$

where $\bar{A}_0$ denotes a vector collecting the unconditional means. Substituting the central bank's forecasts into (29) we find the ALM in the ZLB region under forward guidance to be

$$\begin{aligned} y_t = \Lambda_0(m_t) + \Lambda_1(m_t)y_{t-1} + \Lambda_2(m_t)\left[\bar{A}_0 + A_1^2(y_{t-1} - \bar{A}_0)\right] \\ + \Lambda_3(m_t)\left[\bar{A}_0 + A_1^3(y_{t-1} - \bar{A}_0)\right] \\ \vdots \\ + \Lambda_{N+1}(m_t)\left[\bar{A}_0 + A_1^{N+1}(y_{t-1} - \bar{A}_0)\right] \end{aligned} \tag{31}$$

or simply

$$y_t = \left[\Lambda_0(m_t) + \sum_{j=1}^{N}\Lambda_{j+1}(m_t)\left(I - A_1^{j+1}\right)\bar{A}_0\right] + \left[\Lambda_1(m_t) + \sum_{j=1}^{N}\Lambda_{j+1}(m_t)A_1^{j+1}\right]y_{t-1} \tag{32}$$

Steady states

Let us briefly analyze the steady states of the system (29). To this end, we distinguish again between constant and time-varying fractions as in Section 5. It turns out that under the forward guidance announcements the ZLB steady state is consistent with all fractions $m \in [-1, 1]$ if $N \leq q$. To see this, recognize that for $N \leq q$ the expectations of credibility believers in the ZLB region coincide with those of the central bank. Therefore, the equilibrium with $(\pi^*_{fg}, x^*_{fg}, i^*_{fg}) = (-\bar{r}, -\frac{(1-\beta)\bar{r}}{\kappa}, 0)$ can be supported by all values of m. Under switching, equilibrium fractions are equal, given that both naive and credibility believers make no forecast errors in steady state. This result is summarized in the following Proposition with a formal proof given in the Appendix H.

Proposition 5 (ZLB steady state under forward guidance). *Suppose $\phi > 1$. Then, given $N \leq q$, the ZLB steady state under forward guidance exists*

i. under constant fractions, for all $m \in [-1, 1]$

ii. under time-varying fractions for $m^*_{fg} = 0$

In both cases, steady state levels of inflation, output and the nominal interest rate are given $(\pi^*_{fg}, x^*_{fg}, i^*_{fg}) = (-\bar{r}, -\frac{(1-\beta)\bar{r}}{\kappa}, 0)$. If $N > q$, the results of Proposition 4 apply.

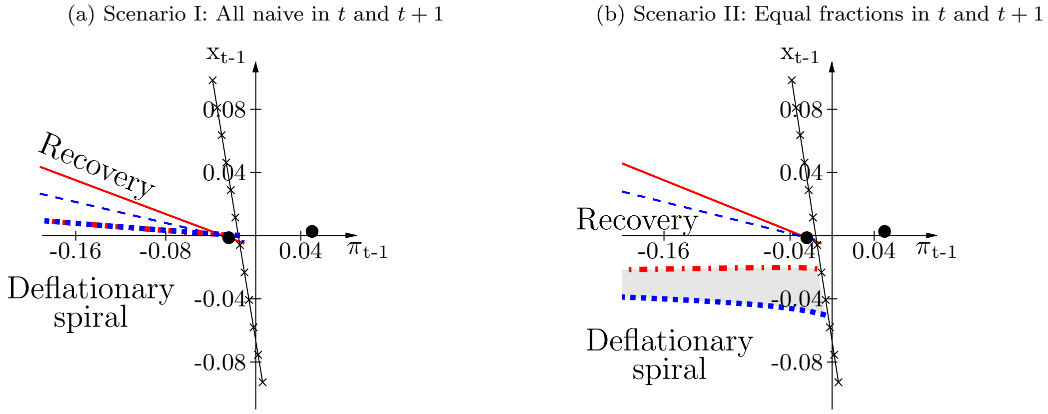
Stability and basin of attraction

The stability of the ZLB steady state under forward guidance with time-varying fractions ($m^* = 0$) is governed by the eigenvalues of the matrix $\Lambda_1(m^*) + \sum_{j=1}^{N} \Lambda_{j+1}(m^*)(A_1^*)^{j+1}$. In particular, if both eigenvalues lie inside the unit circle, the steady state is locally stable. A more detailed analysis is done in the Appendix G.2 which concludes that for the baseline calibration the steady state remains a saddle point as in the case without forward guidance.

We conclude the theoretical discussion by looking at how forward guidance affects the basin of attraction of the target steady state. That is, to evaluate the robustness of the forward guidance policies, we look at all initial conditions in the (π_{t-1}, x_{t-1})-space for which either recovery or a deflationary spiral occurs under forward guidance relative to the no guidance case. In a similar vein, Honkapohja and Mitra (2015) examine the global dynamics of New Keynesian model under learning, when the central bank targets either price level or nominal GDP. The fact that both policy regimes make monetary policy history-dependent allows the authors to implement forward guidance announcements into their model, even though households are purely backward-looking. Honkapohja and Mitra (2015) show that when the policy announcements are incorporated into the private agents' learning, the basin of attraction of the targeted steady state increases substantially, as compared to inflation targeting (the no guidance case in their specification). However, these extreme effects seem to be at least partly driven by the fact that the central bank is considered to be perfectly credible. In fact, we find that *the increase in the basin of attraction of the target steady state due to the forward guidance crucially depend on the central bank's credibility.* In other words, the success of forward guidance depends on the credibility of the central bank.

In both subplots of Figure 3, the black dots represent the target and ZLB steady states, respectively. If the central bank's learning algorithm has already converged to the fixed point described by (21) and (22) with $m^* = 0$, then the central bank will see itself constraint by the ZLB for all initial conditions to the left of the crossed black line. Furthermore, the blue-dashed and red-solid lines depict the stable manifold of the ZLB saddle point with and without forward guidance. All initial conditions that fall into the ZLB region (i.e. the region left of the crossed black line), but remain above the stable manifold will ultimately leave the ZLB region and converge to the target steady state. Contrary, those initial conditions that lie inside the ZLB region but below the stable manifold will diverge and a deflationary spiral may occur. However, there is another important nonlinearity that crucially affects the dynamics, namely the heuristic switching. Therefore, we calculated all initial conditions for which, given an infinite intensity of choice, naive households switch to the credibility believers heuristic in period $t + 2$ and thus also induces a recovery. This boundary between recovery and deflationary spiral is given by the red-dash-dotted (no forward guidance) or blue-dotted (forward guidance) line. In Figure 3, this threshold is the same with and without guidance in scenario I (all naive), while in scenario II (equal fractions) and III (65% credibility believers) the threshold that separates the recovery from the deflationary spiral region under

Figure 3: Basin of attraction of the target steady state



Note: The figure shows the basin of attraction of the target steady state. For scenario I (left panel), we assume all households to be naive in periods t and $t+1$, while for Scenario II (right panel), fractions are assumed to be equal. In each panel, the target steady state is denoted by the (right) black dot at $(\pi^*, x^*) = (\bar{\pi}, \bar{x}))$, while the (left) black dot at $(\pi^*, x^*) = (-\bar{r}, -\frac{1-\beta}{\kappa}\bar{r})$ indicates the ZLB saddle point. The black line with crosses is the ZLB condition (23), to the left of which initial conditions are such that the ZLB binds in period t. The solid red line is the stable manifold corresponding to the ZLB saddle point in the case without forward guidance. The region above this line can be interpreted as the immediate recovery region. Under forward guidance, this line becomes flatter and the blue-dashed line results. Forward guidance thus has an effect on the immediate recovery region. Assuming that $b = \infty$, households switch from naive to credibility believers heuristic in period $t+2$ for all initial conditions above the blue-dotted (forward guidance) and red-dash-dotted (no forward guidance) line, respectively, therefore inducing also convergence back to the target steady state. Forward guidance successfully increases this region for which households become credibility believers in period $t+2$ (indicated by the gray-shaded area), if the credibility of central bank was high enough in the previous periods.

forward guidance shifts stronger downwards than under no guidance. This increase in the basin of attraction of the target steady state is indicated by the gray-shaded area. As both upper panels make clear, the effectiveness of forward guidance announcements hinges on the central bank's credibility (the fraction of credibility believers). Hence, the effectiveness of forward guidance can be summarized as follows:

Result (Effectiveness of forward guidance). *Forward guidance increases the likelihood of recovery from a liquidity trap, because the basin of attraction of the target steady state is larger than without guidance. How large the impact of forward guidance is, depends on the fractions of credibility believers at the moment of the announcement. The larger the fraction of credibility believers, the more effective is forward guidance.*

6 Numerical Results

6.1 Calibration

We calibrate the model at a quarterly frequency using values common in the monetary policy literature. In particular, we set the discount factor equal to $\beta = 0.997$–a value often found in estimated DSGE models (see, e.g. Slobodyan and Wouters, 2012)–which implies an annualized steady state real interest rate of about 1.2%. Further, we follow Galí (2009) and references therein and set the inverse elasticity of intertemporal substitution $\sigma = 1$ and a Frisch elasticity of $\varphi = 1$. On the supply side, we set the elasticity of substitution between differentiated goods $\epsilon = 6$, which corresponds to a steady state mark-up of 20%. Further, we set ψ such that the slope of the Philips curve becomes $\kappa = 0.0561$ as in Galí (2009) and very close to the calibration in e.g. Eusepi and

Preston (2010). The calibration of the interest rate rule follows Taylor (1993), so we set $\phi = 1.5$, and we assume a positive inflation target $\bar{\pi}$ of 2% annualized. Moreover, we begin with a forward guidance horizon of one year, so that $q = 4$, which is broadly in line with Hanson and Stein (2015) and Swanson and Williams (2014), who argue that the Fed's forward guidance strategy operates with a roughly one-to-two year horizon.

For our behavioral parameters, however, there is less consensus about the calibration. For this reason, Section I in the Appendix provides a robustness analysis with respect to the crucial parameters. We choose a forward-looking horizon of the private households of $N = 4$, which corresponds to a one year planning horizon. When asked to report their (financial) planning horizons, roughly 60-70% of survey respondents in the Survey of Consumer Finances or the Health and Retirement Study (HRS) in the US indicated to have planning horizons ranging from a few month up to a few years (see e.g. Rodriguez de Rubio, 2015; Dow Jr and Jin, 2013).[20] Further, we set the intensity of choice $b = 2500$.[21] With regards the gain parameter ω in the adaptive expectations rule, we follow Anufriev and Hommes (2012*a*,*b*) and set it equal to 0.65. This coefficient was found to predict individual forecasting behavior in learning-to-forecast experiments reasonably well.

Concerning the demand and supply shock, we first follow De Grauwe (2012) and set the persistence in both processes to zero, i.e. $\rho_x = \rho_\pi = 0$ and standard deviations equal to $\sigma_x = \sigma_\pi = 0.005$. This calibration is used in Sections 6.2.1 and 6.2.2. In fact, the model produces reasonable degrees of persistence in key macro-variables without the need of highly persistent shocks (see De Grauwe, 2012). However, to bring the model to the ZLB, large disturbance terms are needed. To circumvent such an assumption, we instead assume autocorrelated shocks with $\rho_x = \rho_\pi = 0.5$ and reduce the standard deviations to $\sigma_x = \sigma_\pi = 0.003$ in Section 6.2.3, which presents our main results.

The benchmark calibration of the main model parameters is summarized in Table 1. If the calibration differs from the benchmark case it will be made explicit. The simulations below are chosen to illustrate our key findings.

6.2 Simulation-based policy analysis

In this section, we illustrate our findings of the numerical analysis. The presented behavioral macro model allows for a rich set of dynamics, so that we restrict ourselves and only present the main numerical results. In a nutshell, our model predicts the following. First, the combination of real-time learning, model misspecification and imperfect credibility of the central bank gives rise to policy mistakes that can result in periods of high inflation. Second, forward guidance announcements are effective in preventing the economy from entering a deflationary spiral. Notably, those announcements of the Odyssean type can decrease the probability of deflationary spirals by around 18 percentage points.

[20] Figure 7 shows the eigenvalues of the model as a function of N, indicating that for too large values the canonical NKM considered here becomes unstable under N-step Euler equation learning.

[21] Note that the calibration of the intensity of choice crucially depends on both the definition and the unit of measurement of the fitness measure. In our simulations a 1% deviation of inflation from steady state is measured as 0.01, which corresponds a squared forecast error of 0.0001. For instance De Grauwe (2011) sets $b = 1$, which corresponds to $b = 10,000$ in our framework. Similarly, Anufriev and Hommes (2012*a*) set $b = 0.4$, but in an asset pricing environment. Hommes and Lustenhouwer (2015) therefore argue that an intensity of choice of 40,000 should not be considered as exceptionally large. Since we, however, define the fitness measure as the sum of squared forecast errors resulting from output, inflation and interest rate expectations, a lower value of b results in a similar degree of sensitivity.

Table 1: Parameter calibration

Parameter	Value	Description	Source
β	.997	Discount factor	Slobodyan and Wouters (2012)
N	4	Forward-looking horizon of private households	
$1/\sigma$	1	Elasticity of intertemporal substitution	Galí (2009)
φ	1	Frisch elasticity of labor supply	Galí (2009)
ϵ	6	Elasticity of substitution between differentiated goods	Galí (2009)
κ	.0561	Slope of the NKPC	Galí (2009)
ϕ	1.5	Central bank's inflation response	Taylor (1993)
$\bar{\pi}$	.02	Inflation target (annualized)	
$\tilde{\pi}$	.08	Inflation tolerance (annualized)	
q	4	Forward guidance horizon	Swanson and Williams (2014)
b	2500	Intensity of choice	
ω	.65	Gain parameter of adaptive learners	Anufriev and Hommes (2012*a*,*b*)
ρ_x	0 or .5	Persistence parameter in demand shock	De Grauwe (2012)
ρ_π	0 or .5	Persistence parameter in supply shock	De Grauwe (2012)
σ_x	.003 or .005	Std. deviation in demand shock	De Grauwe (2012)
σ_π	.003 or .005	Std. deviation in supply shock	De Grauwe (2012)

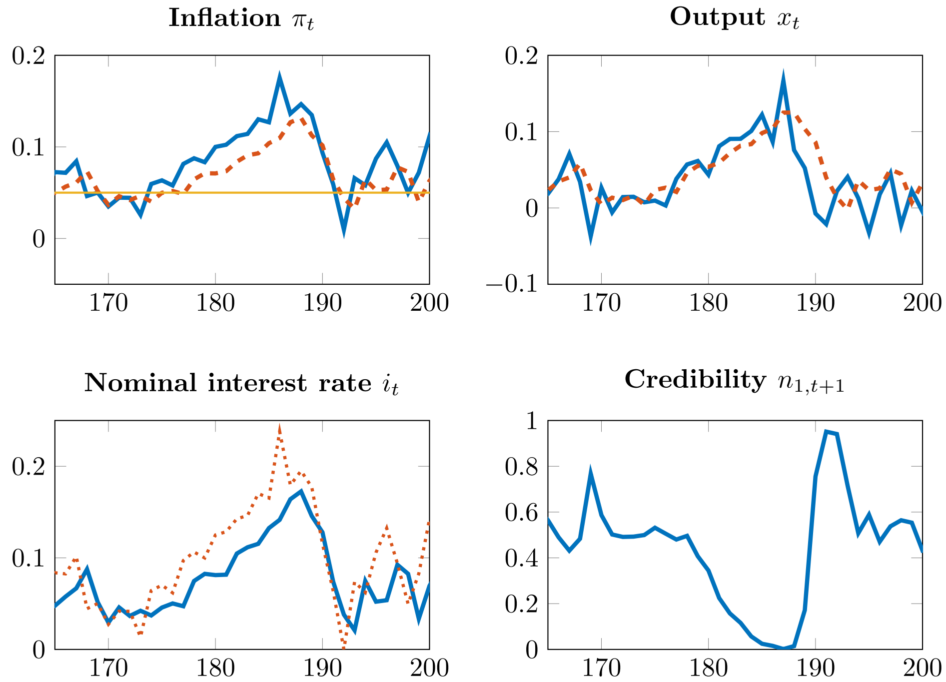
6.2.1 Model misspecification and policy mistakes

Before going to the analysis of forward guidance, we want to clarify the role of the central bank's model misspecification. It turns out that this informational friction can create−or at least amplify−booms and busts. In particular, if the central bank (systematically) underestimates inflation over several quarters, it will pursue an interest rate policy that is not anti-inflationary enough, and vice versa. Figure 4 illustrates such a scenario, in which the economy overheats due to an inflationary interest rate policy. The upper panels depict time series for inflation and output (both solid blue lines), respectively, together with the central bank's real time estimates (red-dashed). Evident from the left panel, the central bank underestimates inflation (the red-dashed line is permanently below the blue-solid line) and thus inadvertently pursues an interest rate policy that is not anti-inflationary enough. The latter can be seen in the lower-left panel, which plots both the interest rate actually set by the central bank that is consistent with its expectations (blue-solid line), as well as the nominal interest rate that would have been set if the central bank had observed current inflation (red dotted). This full-information benchmark is equivalent to rational expectations.

Arguably, these policy mistakes can occur for two reasons. Either the central bank finds itself in an early stage of learning (i.e. the central bank's VAR coefficients have not yet converged) and therefore forecast may be inaccurate or, secondly, more households coordinate on adaptive expectations, which is not realized by the central bank due to model misspecification. In both cases the drift of inflation and/or output induced by the initial policy mistake leads to further coordination away from the credibility believer heuristic, which leads to self-fulfilling expectations that amplify the drift. That is, these periods with a too weak response on inflation (due to underestimation of inflation, but not due to a too small response coefficient ϕ) are characterized by a decline in central bank credibility (see lower right panel in Figure 4).[22] This result is in line with Lubik and Matthes (2016) who study the Fed's interest rate policy during the high-inflation period of 1970s (also known as the Great Inflation). Using Bayesian estimation methods the authors calibrate a stylized New Keynesian model to argue that the Great Inflation is the

[22] In longer stochastic time series simulations, we find that these booms and busts can also occur if the central bank's VAR coefficients seemed to have nearly converged. In other words, the model misspecification of the central bank due to the presence of heterogeneous households allows for policy mistakes at any point in time. However, the simulations also suggest that a larger autocorrelation in shocks increases the forecast performance and accuracy of the central bank's VAR model, as does a lower standard deviation in disturbance terms.

Figure 4: Model misspecification and policy mistakes



Note: The figure shows that when the central bank underestimates inflation (true inflation given by the blue line versus central bank's inflation nowcast red-dashed in the upper left panel), it inadvertently pursues an interest rate policy that is not anti-inflationary enough. The latter can be seen by comparing the nominal interest rate that results from the central bank's real-time nowcasts (blue solid line) versus the nominal interest rate set under full information, i.e. rational expectations, given by the red dotted line in the left lower panel. During this period, the central bank's credibility decreases (bottom right panel), which leads to a period of rising output and inflation. Shocks are assumed to be *iid* with zero autocorrelation.

result of equilibrium indeterminacy, inadvertently arising from an optimizing central bank that was constraint by uncertainty about the structure of the economy and measurement error in real-time data. Both these constraints are also present in our paper, although we interpret the structural uncertainty as the central bank not being aware of the time-variation in aggregate expectations. The latter is broadly in line with Mankiw et al. (2003), who document substantial disagreement in inflation expectations over time.

6.2.2 Constant gain learning preventing policy mistakes

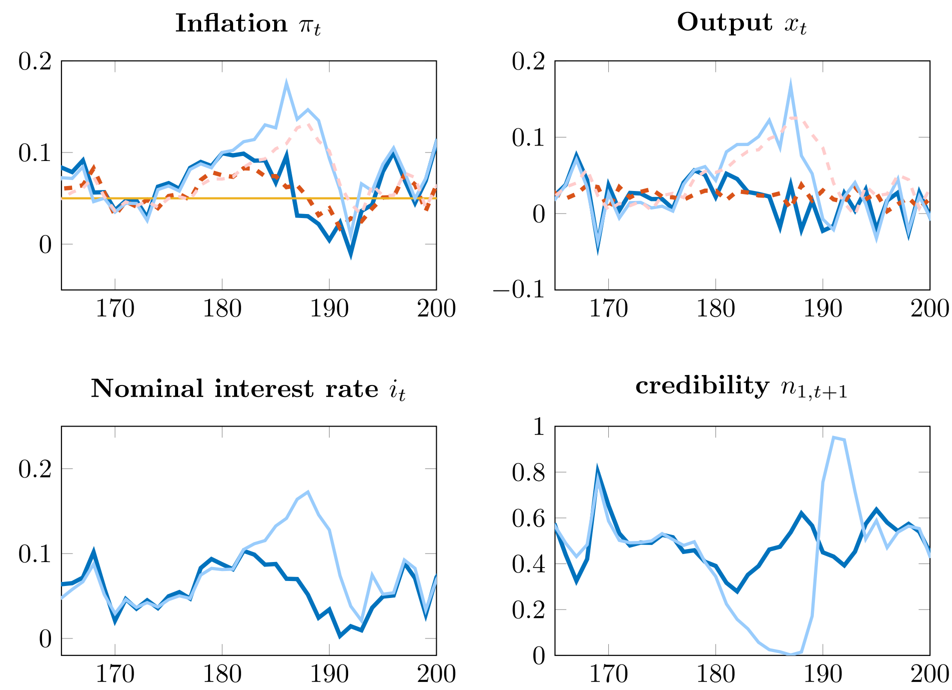
In this subsection we aim to answer the question of how important the central bank's model misspecification is in driving the policy mistakes. First, recall that this misspecification is coming from the fact that the central bank believes the aggregate law of motion to have time-invariant coefficients. In other words, the central bank assumes that the economy is not subject to time switches in the sentiment of economic agents, here modeled by the time-variation in the distribution of aggregate expectations of inflation, output and nominal interest rate. To relax this assumption, we allow the central bank to estimate its VAR model using time-varying parameters (i.e. a constant gain learning (CGL) algorithm).[23] In this case, the central bank is potentially able to recognize drifts in inflation earlier than under decreasing gain learning (DGL), and thus might pursue policies

[23] Constant gain learning is similar to a rolling window estimation, since past observations receive a geometrically declining weight. In our numerical simulations we set $\gamma = 0.02$, however, the results are robust to different choices of the constant gain parameter, which still can be considered reasonable (see, e.g. Branch and Evans, 2006).

which are more anti-inflationary.

In what follows, we present a counter-factual analysis to the simulation presented in Figure 4. Using the same random seed and holding everything else equal, we change the gain coefficient from the decreasing gain $\gamma_t = \frac{1}{t}$ to a constant gain of $\gamma_t = \gamma$. The resulting time series under CGL are presented in Figure 5. For better comparison, we also plot the simulation results under DGL in lighter colors, again with dashed lines representing the central bank's real-time estimates. Evident from Figure 5, the central bank predicts slightly higher inflation around and after period 177 under CGL, better capturing the drift of inflation away from its target level due to a larger fraction of adaptive learners. The resulting (slightly) higher nominal interest rates are then putting a halt on inflation and output from further increasing, which boosts the central bank's credibility and prevents further households from switching to the adaptive expectations rule. Therefore, we conclude that CGL$-$by putting more weight on recent observations$-$reduces the chance of the aforementioned policy mistakes.

Figure 5: Constant vs. decreasing gain in the central bank's learning algorithm

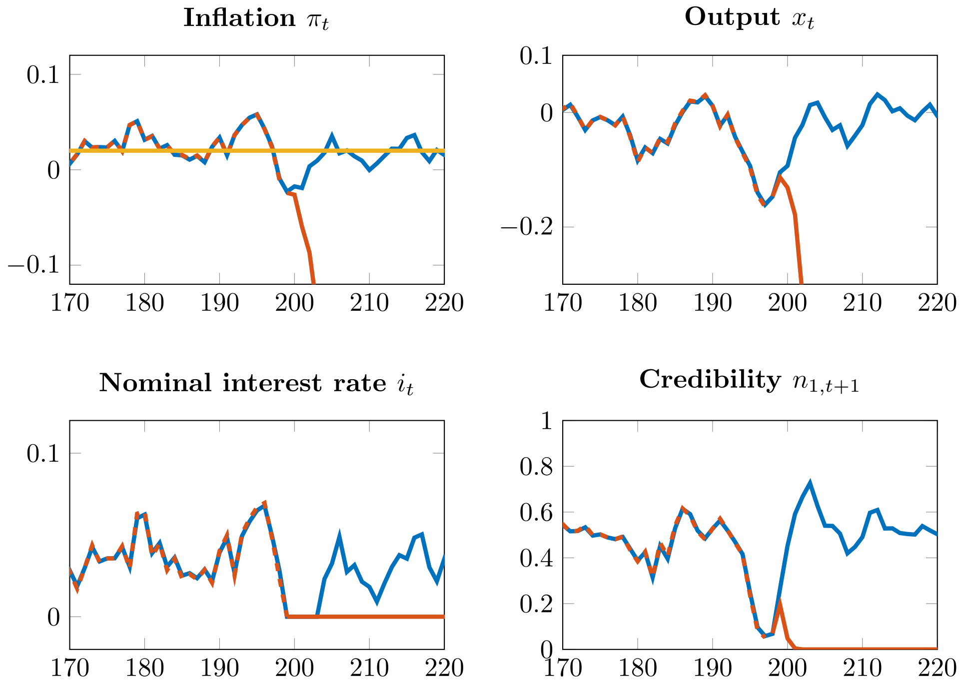


Note: The figure shows the counter-factual to Figure 4 by assuming that the monetary authority uses a constant gain (stronger colors) instead of a decreasing gain (lighter colors) in its updating equation, thereby allowing for time-variation in its PLM. It is therefore able to better detect time-variation in aggregate expectations which improves the nowcasting performance and reduces policy mistakes.

6.2.3 Effectiveness of forward guidance

This section returns to our main research question: What are the effects of forward guidance if households are heterogeneous and boundedly rational? To anticipate the results, we find that forward guidance decreases the likelihood of deflationary spirals and thereby stabilizes the economy. For illustration, Figure 6 compares two stochastic time series simulation in which we allow for forward guidance (blue) or not (red). In the simulation without forward guidance (red) the economy enters a deflationary spiral, while under forward guidance (blue) the economy is able to recover (blue). The crucial point is the pronounced increase in central bank credibility shortly after the

Figure 6: Deflationary spirals and the power of forward guidance



Note: The figure shows stochastic time series simulations in which the ZLB starts binding from period 199 and the economy enters a deflationary spiral in case of no forward guidance (red), while under Delphic and Odyssean forward guidance the economy can escape the liquidity trap (blue).

forward guidance announcement (see bottom right panel of Figure 6). The reason is twofold. First, the publication of the central bank's own projections for inflation and output turned out to be more accurate than the private sector's own (adaptive) beliefs, which increases the bank's credibility ex post. This switching away from the adaptive expectations heuristic puts a halt to the diverging and self-fulfilling process in deflation and falling output. In fact, if successful, Delphic guidance induces a self-fulfilling recovery. Second, the announced lower-for-long policy additionally lowers the aggregate expected future real interest rate and thus increases aggregate demand. Lastly, the trust in the central bank also increases ex post once households observe that the bank sticked to the announced lower-for-long policy.

To quantitatively assess the role of both Delphic and/or Odyssean forward guidance in preventing deflationary spirals, we run a total of 10,000 Monte Carlo simulations with a length of 30 years (120 quarters) and compute the likelihood of a deflationary spiral under four different scenarios: no guidance, both types of guidance, only Delphic or only Odyssean guidance. Additionally, an initialization period for the central bank's learning algorithm of 10 years (40 quarters) is simulated.[24] A deflationary spiral is then defined as a liquidity trap with diverging inflation and/or output with values in excess of -20 percent quarterly. Results are both qualitatively and quantitatively robust to other values. We also present moments for the US economy over the last 30 years (i.e. from 1988Q1 until 2017Q4), because our calibration assumes an inflation target of 2% and a relatively low level for the steady state real rate ($\frac{1}{\beta} - 1$).

Besides looking at the likelihood of deflationary spiral, we also conduct a welfare analysis based on an *ad hoc* loss function, that is quadratic in deviations of output and inflation from their steady

[24] In the initialization period the central bank is assumed to have full information, i.e. central bank perfectly observes current period inflation.

state levels, following Evans and Honkapohja (2003, 2006) and Gasteiger (2014, 2017). Formally, this quadratic loss function $\mathbb{L}$ is given by

$$\mathbb{L} = -\left[(\pi_t - \bar{\pi})^2 + \alpha_x (x_t - \bar{x})^2\right], \tag{33}$$

where $\alpha_x = \frac{\kappa}{\epsilon}$.[25] Since we do not consider optimal forward guidance policy, and therefore, are not interested in minimizing expected future welfare losses, we instead consider *ex post* welfare that resulted from a given policy. Thus, we take the non-discounted sum over the simulation horizon of T periods to find the following welfare criterion

$$\mathbb{W} = -\sum_{t=1}^{T} \left[(\pi_t - \pi)^2 + \alpha_x (x_t - x)^2\right]. \tag{34}$$

The results are summarized in Table 2. It shows that both Delphic and Odyssean forward guidance jointly can decrease the likelihood of deflationary spirals from 19.91 to 9.04 percent, i.e. a reduction of roughly 11 percentage points.[26] In comparison, the U.S. federal funds rate has been pinned close to zero for more than one-fifth of the last 30 years. However, Table 2 also reveals a lower-for-long policy alone might be even more effective by reducing the likelihood even further to 8.48 percent. Looking at the variation in both output and inflation, forward guidance clearly stabilizing effect. While Odyssean guidance appears to reduce the standard deviation of both variables the least, Delphic guidance seems to be crucial in lowering the variation. This observation translates one-to-one into implications for average welfare. The combination of Delphic and Odyssean guidance ranks first, followed by Delphic and only then by Odyssean guidance in welfare-terms under our benchmark calibration.

Table 2: Forward guidance and the likelihood of deflationary spirals

	US Data 1988Q1 - 2017Q4	*without* forward guidance	Delphic guidance	Odyssean guidance	Both
Likelihood (in %)	23.33	19.91	10.48	8.48	9.04
avg. SD x_t (in %)	1.93	2.41	1.71	1.81	1.77
avg. SD π_t (in %)	0.87	0.92	0.68	0.69	0.66
avg. Welfare $\mathbb{W} \cdot 100$		-0.3686	-0.3522	-0.3562	-0.3515

Note: In row one, the table presents the relative share of 10,000 Monte Carlo simulations in which a deflationary spiral occurred under different policies. The lower the probability, the less likely it is for the economy to be locked in a liquidity trap. The second and third rows present the average standard deviations of output and inflation, while the last row shows the average of our measure of ex post welfare, with averages taken over all simulations. The calibration is given in Table 1. With regards the data, we use output gap data from the Congressional Budget Office (CBO), while data for the federal funds rate to calculate the % of ZLB periods and PCE inflation are taken from the Federal Reserve Economic Data (FRED).

In our model, forward guidance is acting through two channels. First, Delphic guidance can help to coordinate aggregate expectations if the central bank's own projections are more accurate

[25] This quadratic loss function can be derived from a second-order Taylor approximation to the level of expected utility of the representative household in an equilibrium with a given monetary policy. See Rotemberg and Woodford (1999) and Woodford (1999) for more details. Such an ad hoc welfare function might understate the true welfare losses due to heterogeneity in expectations, and an alternative therefore would be to utilize a model-consistent loss function as derived by Di Bartolomeo et al. (2016). However, Gasteiger (2017) discusses practical reasons in favor of such an ad hoc loss function used here in Chapter 3.2.

[26] The qualitative results are robust to other threshold values. For instance, imposing a threshold level of either -25%, -10% or -5%, we find the likelihood of a deflationary spiral under both types of guidance to change from 9.04% to 8.99%, 9.21% or 10.55%, respectively. Evidently, a lower threshold leads to a marginally higher likelihood. The same holds true if you look at the policies separately. Decreasing the threshold level to e.g. -5% increases the likelihood of a spiral under Odyssean guidance only by a mere 2 basis points to 9.06%.

than those of the adaptive learners. The idea is that, once the central bank publishes its (potentially more accurate) projections, households observe these forecasts, which can ex post increase its credibility and more households coordinate on the credibility believers rule. However, if credibility was already large at the moment of the announcement, this channel can also lower aggregate inflation expectations, given that in a recession credibility believers forecasts are somewhat more optimistic than those of the central bank. This channel relates to the signaling channel of monetary policy in Melosi (2017) in that households adjust their expectation in the direction of new information. Second, Odyssean guidance can reduce the expected real interest rates of the credibility believers, thereby increasing demand through the standard Euler equation channel. However, the effectiveness of guidance depends crucially on the central bank's credibility.

To further sharpen this result, we looked at the effectiveness of an interest rate peg, i.e. a lower-for-long policy, without the explicit forward guidance announcement. In this case, the likelihood of a deflationary spiral even increases to 26.58%, suggesting that the announcement of such a policy is crucial.[27]

7 Conclusion

In this paper, we provide a model with N-step horizon Euler equation learning and endogenous central bank credibility to study the effectiveness of forward guidance under both heterogeneous and boundedly rational expectations. In particular, we model heterogeneity in expectations across households, but also with respect to the central bank. Private households switch between an adaptive expectations rule and a credibility believer heuristic. The central bank, on the other hand, estimates a VAR model, which it updates recursively.

In our baseline model, the central bank uses recursive least squares with a decreasing gain, thereby the time-variation in aggregate expectations. As a result of this model misspecification, the central bank learns the true model only if the fractions of households using a specific forecasting heuristic converge, i.e. the distribution of aggregate expectations becomes time-invariant. Otherwise, the economy settles down in a restricted perception equilibrium with time-varying heterogeneous beliefs.

Our results suggest that the combination of real-time learning, model misspecification and imperfect credibility of the central bank gives rise to policy mistakes that can result in periods of high inflation, potentially explaining periods such as the Great inflation in the 1980s in the United States. These policy mistakes are, however, not because of too weak (or too strong) responses to inflation, but due to the aforementioned combination of reasons. Under constant gain learning (CGL), however, the central bank allows for time-varying parameters and thus better account for the time-variation in aggregate expectations, which ultimately reduces policy mistakes.

The key novelty of our paper is that we make the effectiveness of forward guidance depending on the central bank's credibility, which itself is endogenously determined by how well the central bank performed in achieving its target in the recent past. That is, an endogenous fraction of private households incorporate the central bank's guidance announcements into their expectation formation, while the rest of the population forms expectations in an adaptive fashion.

Forward guidance is then acting through two channels. First, Delphic guidance can help to

[27] We also looked at the effectiveness of Delphic guidance when only a subset of forecasts are published. Intuitively, Delphic guidance becomes less powerful when only inflation and nominal interest rate projections are communicated with the likelihood of spirals increasing from 10.48% to 12.62%.

coordinate aggregate expectations if the central bank's own projections are more accurate than those of the adaptive learners. The idea is that, once the central bank publishes its (potentially more accurate) projections, households observe the forecasts of the central bank, which can ex post increase its credibility and more households coordinate on the stabilizing credibility believers rule. However, if credibility was still large, this effect can also lower aggregate inflation expectations, given that in a recession credibility believers forecasts are somewhat more optimistic than those of the central bank. Second, both Delphic and Odyssean guidance policies can reduce the expected real interest rates of the credibility believers, thereby increasing demand through the standard Euler equation channel. However, the effectiveness of guidance depends crucially on the central bank's credibility.

Using Monte Carlo simulations, we find that both Delphic and Odyssean forward guidance are effective in lowering the likelihood of deflationary spirals. Specifically, Delphic and Odyssean forward guidance jointly decrease the likelihood of a deflationary spiral by 11 percentage points. Notably, a Odyssean-style lower-for-long policy alone can decrease the probability of a deflationary spiral even more, however at the cost of increased macroeconomic volatility and thus lower welfare. Delphic guidance alone, on the other hand, can reduce the probability by only ca. 9.5 percentage points, yet is crucial in reducing inflation and output volatility, making the a mix of both preferred in welfare-terms. We leave the assessment of the optimal forward guidance mix and length for future research.

References

Adam, Klaus (2007), 'Experimental evidence on the persistence of output and inflation*', *The Economic Journal* **117**(520), 603–636.

Ahrens, Steffen, Joep Lustenhouwer and Michele Tettamanzi (2016), The stabilizing role of forward guidance: A macro experiment, Technical report.

Andrade, Philippe, Gaballo Gaetano, Eric Mengus and Benoit Mojon (2015), Forward guidance and heterogeneous beliefs, SSRN Working Paper Series No. 2677644.

Anufriev, Mikhail and Cars Hommes (2012*a*), 'Evolution of market heuristics', *The Knowledge Engineering Review* **27**(2), 255–271.

Anufriev, Mikhail and Cars Hommes (2012*b*), 'Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments', *American Economic Journal: Microeconomics* **4**(4), 35–64.

Aoki, Kosuke and Kalin Nikolov (2006), Rule-based monetary policy under central bank learning, *in* 'NBER International Seminar on Macroeconomics 2004', The MIT Press, pp. 145–195.

Arifovic, Jasmina and Luba Petersen (2015), Escaping expectations-driven liquidity traps: Experimental evidence. Manuscript, Simon Fraser University.

Assenza, Tiziana, Peter Heemeijer, Cars H Hommes and Domenico Massaro (2013), Individual expectations and aggregate macro behavior, Tinbergen Institute Discussion Paper 13-016/II.

Benhabib, Jess, Stephanie Schmitt-Grohé and Martin Uribe (2001*a*), 'Monetary policy and multiple equilibria', *American Economic Review* **91**(1), 167–186.

Benhabib, Jess, Stephanie Schmitt-Grohé and Martin Uribe (2001*b*), 'The perils of taylor rules', *Journal of Economic Theory* **96**(1), 40–69.

Bernanke, Ben S, Laubach Thomas Mishkin Frederic S and Adam S Posen (2001), *Inflation targeting: lessons from the international experience*, Princeton University Press.

Bilbiie, Florin O (2016), Optimal forward guidance, CEPR Discussion Paper No. DP11251.

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

Branch, William A (2007), 'Sticky information and model uncertainty in survey data on inflation expectations', *Journal of Economic Dynamics and Control* **31**(1), 245–276.

Branch, William A and Bruce McGough (2009), 'A new keynesian model with heterogeneous expectations', *Journal of Economic Dynamics and Control* **33**(5), 1036–1051.

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

Branch, William, George W Evans and Bruce McGough (2012), Finite horizon learning, CDMA, Working Paper Series CDMA12/04.

Brock, William A and Cars H Hommes (1997), 'A rational route to randomness', *Econometrica: Journal of the Econometric Society* **65**, 1059–1095.

Busetti, Fabio, Davide Delle Monache, Andrea Gerali and Alberto Locarno (2017), Trust, but verify. de-anchoring of inflation expectations under learning and heterogeneity, Technical report, ECB Working Paper No. 1994.

Caballero, Ricardo J. and Emmanuel Farhi (2013), A model of the safe asset mechanism (sam): Safety traps and economic policy, NBER Working Paper No. 18737.

Campbell, Jeffrey, Jonas Fisher, Alejandro Justiniano and Leonardo Melosi (2016), Forward guidance and macroeconomic outcomes since the financial crisis, *in* 'NBER Macroeconomics Annual 2016, Volume 31', University of Chicago Press.

Campbell, Jeffrey R, Charles L Evans, Jonas DM Fisher, Alejandro Justiniano, Charles W Calomiris and Michael Woodford (2012), 'Macroeconomic effects of federal reserve forward guidance [with comments and discussion]', *Brookings Papers on Economic Activity* pp. 1–80.

Carlstrom, Charles T, Timothy S Fuerst and Matthias Paustian (2015), 'Inflation and output in new keynesian models with a transient interest rate peg', *Journal of Monetary Economics* **76**, 230–243.

Carroll, Christopher D (2003), 'Macroeconomic expectations of households and professional forecasters', *The Quarterly Journal of Economics* pp. 269–298.

Cochrane, John H. (2013), The new-keynesian liquidity trap, NBER Working Paper No 19476.

Cole, S (2015), Learning and the effectiveness of central bank forward guidance, MPRA Working Paper No 65207.

Cornea-Madeira, Adriana, Cars Hommes and Domenico Massaro (2017), 'Behavioral heterogeneity in us inflation dynamics', *Journal of Business & Economic Statistics* pp. 1–13.

De Graeve, Ferre, Pelin Ilbas and Raf Wouters (2014), 'Forward guidance and long term interest rates: Inspecting the mechanism'.

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

De Grauwe, Paul (2012), *Lectures on Behavioral Macroeconomics*, Princeton University Press.

Del Negro, Marco, Marc Giannoni and Christina Patterson (2012), The forward guidance puzzle, Federal Reserve Bank of New York Staff Reports 574.

Demertzis, Maria, Massimiliano Giuseppe Marcellino and Nicola Viegi (2008), A measure for credibility: tracking us monetary developments, CEPR Discussion Paper No. DP7036.

Di Bartolomeo, Giovanni, Marco Di Pietro and Bianca Giannini (2016), 'Optimal monetary policy in a new keynesian model with heterogeneous expectations', *Journal of Economic Dynamics and Control* **73**, 373–387.

Dow Jr, James P and Yanbo Jin (2013), 'The determination of individual financial planning horizons', *Southwestern Economic Review* **40**, 137–149.

Eggertson, Gauti B. and Michael Woodford (2003), 'The zero bound on interest rates and opimtal monetary policy', *Brookings Papers on Economic Activity* pp. 139–211.

Eusepi, Stefano (2005), Central bank transparency under model uncertainty, Technical report, FRB of New York Staff Report No. 199.

Eusepi, Stefano (2010), 'Central bank communication and the liquidity trap', *Journal of Money, Credit and Banking* **42**(2-3), 373–397.

Eusepi, Stefano and Bruce Preston (2010), 'Central bank communication and expectations stabilization', *American Economic Journal: Macroeconomics* **2**(3), 235–271.

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

Evans, George W, Eran Guse and Seppo Honkapohja (2008), 'Liquidity traps, learning and stagnation', *European Economic Review* **52**(8), 1438–1463.

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

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

Evans, George W and Seppo Honkapohja (2006), 'Monetary policy, expectations and commitment*', *The Scandinavian Journal of Economics* **108**(1), 15–38.

Faust, Jon and Eric M Leeper (2005), 'Forecasts and inflation reports: An evaluation', *Washington, DC: Federal Reserve Board* .

Ferrero, Giuseppe and Alessandro Secchi (2009), 'The announcement of monetary policy intentions', *Bank of Italy Temi di Discussione (Working Paper) No* **720**.

Ferrero, Giuseppe and Alessandro Secchi (2011), Central bank's macroeconomic projections and learning, Technical Report No. 72.

Friedman, Milton (1968), 'The role of monetary policy', *The American economic review* **58**(1), 1–17.

Gabaix, Xavier (2016), A behavioral new keynesian model, NBER Working Paper No. 22954.

Galí, Jordi (2009), *Monetary Policy, inflation, and the Business Cycle: An introduction to the new Keynesian Framework*, Princeton University Press.

García-Schmidt, Mariana and Michael Woodford (2015), Are low interest rates deflationary? a paradox of perfect-foresight analysis, NBER Working Paper No. 21614.

Gasteiger, Emanuel (2014), 'Heterogeneous expectations, optimal monetary policy, and the merit of policy inertia', *Journal of Money, Credit and Banking* **46**(7), 1535–1554.

Gasteiger, Emanuel (2017), Optimal constrained interest-rate rules under heterogeneous expectations, unpublished manuscript.

Haberis, Alex, Richard Harrison and Matt Waldron (2014), Transitory interest-rate pegs under imperfect credibility, Technical report.

Hanson, Samuel G and Jeremy C Stein (2015), 'Monetary policy and long-term real rates', *Journal of Financial Economics* **115**(3), 429–448.

Hommes, Cars (2011), 'The heterogeneous expectations hypothesis: Some evidence from the lab', *Journal of Economic Dynamics and Control* **35**(1), 1–24.

Hommes, Cars and Joep Lustenhouwer (2015), Inflation targeting and liquidity traps under endogenous credibility, CeNDEF Working Paper No. 15-03.

Hommes, Cars, Joep Sonnemans, Jan Tuinstra and Henk Van de Velden (2005), 'Coordination of expectations in asset pricing experiments', *Review of Financial Studies* **18**(3), 955–980.

Honkapohja, Seppo and Kaushik Mitra (2005), 'Performance of monetary policy with internal central bank forecasting', *Journal of Economic Dynamics and Control* **29**(4), 627–658.

Honkapohja, Seppo and Kaushik Mitra (2006), 'Learning Stability in Economies with Heterogeneous Agents', *Review of Economic Dynamics* **9**(2), 284–309.

Honkapohja, Seppo and Kaushik Mitra (2014), 'Targeting nominal gdp or prices: Guidance and expectation dynamics'.

Honkapohja, Seppo and Kaushik Mitra (2015), 'Comparing inflation and price-level targeting: The role of forward guidance and transparency', *The Manchester School* **83**(S2), 27–59.

Hubert, Paul (2014), 'Fomc forecasts as a focal point for private expectations', *Journal of Money, Credit and Banking* **46**(7), 1381–1420.

Hubert, Paul (2015*a*), 'Do central bank forecasts influence private agents? forecasting performance versus signals', *Journal of Money, Credit and Banking* **47**(4), 771–789.

Hubert, Paul (2015*b*), 'The influence and policy signalling role of fomc forecasts', *Oxford Bulletin of Economics and Statistics* **77**(5), 655–680.

Kiley, Michael T (2016), 'Monetary policy statements, treasury yields, and private yields: before and after the zero lower bound', *Finance Research Letters* **18**, 285–290.

King, Robert G (2010), 'Discussion of "limitations on the effectiveness of forward guidance at the zero lower bound"', *International Journal of Central Banking* **6**(1), 191–203.

Laubach, Thomas and John C Williams (2016), 'Measuring the natural rate of interest redux', *Business Economics* **51**(2), 57–67.

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

Mankiw, N Gregory, Ricardo Reis and Justin Wolfers (2003), 'Disagreement about inflation expectations', *NBER macroeconomics annual* **18**, 209–248.

McKay, Alisdair, Emi Nakamura and Jón Steinsson (2016), 'The power of forward guidance revisited', *American Economic Review* **106**(10), 3133–58.

Melosi, Leonardo (2017), 'Signaling effects of monetary policy', *The Review of Economic Studies* **84**(2), 853–884.

Moessner, Richhild, David-Jan Jansen and Jakob de Haan (2017), 'Communication about future policy rates in theory and practice: A survey', *Journal of Economic Surveys* **31**(3), 678–711.

Myatt, David P and Chris Wallace (2014), 'Central bank communication design in a lucas-phelps economy', *Journal of Monetary Economics* **63**, 64–79.

Orphanides, Athanasios (2001), 'Monetary policy rules based on real-time data', *American Economic Review* **91**(4), 964–985.

Orphanides, Athanasios and John C Williams (2005), 'Imperfect knowledge, inflation expectations, and monetary policy', *The Inflation-Targeting Debate* **32**, 201.

Pfajfar, D and B Zakelj (n.d.), Inflation expectations and monetary policy design: Evidence from the laboratory, CentER Discussion Paper No. 2011-091.

Pfajfar, Damjan (2009), Heterogeneous expectations in macroeconomics, PhD thesis, University of Cambridge.

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

Preston, Bruce (2006), 'Adaptive learning, forecast-based instrument rules and monetary policy', *Journal of Monetary Economics* **53**(3), 507–535.

Raskin, Matthew (2013), The effects of the federal reserve's date-based forward guidance, FRB Finance and Economics Discussion Series 2013-37.

Rodriguez de Rubio, Alicia (2015), 'Factors associated with households' planning horizons for making saving and spending decisions', *Family and Consumer Sciences Research Journal* **43**(3), 284–292.

Rotemberg, Julio J (1982), 'Monopolistic price adjustment and aggregate output', *The Review of Economic Studies* **49**(4), 517–531.

Rotemberg, Julio J and Michael Woodford (1999), Interest rate rules in an estimated sticky emodel, *in* 'Monetary policy rules', University of Chicago Press, pp. 57–126.

Sargent, Thomas J (1999), *The conquest of American inflation*, Princeton University Press.

Simon, Herbert A (1984), 'On the behavioral and rational foundations of economic dynamics', *Journal of Economic Behavior & Organization* **5**(1), 35–55.

Slobodyan, Sergey and Raf Wouters (2012), 'Learning in a medium-scale dsge model with expectations based on small forecasting models', *American Economic Journal: Macroeconomics* **4**(2), 65–101.

Swanson, Eric T and John C Williams (2014), 'Measuring the effect of the zero lower bound on medium-and longer-term interest rates', *American Economic Review* **104**(10).

Taylor, John B (1993), Discretion versus policy rules in practice, *in* 'Carnegie-Rochester conference series on public policy', Vol. 39, Elsevier, pp. 195–214.

Woodford, Michael (1999), 'Optimal monetary policy inertia', *The Manchester School* **67**(s1), 1–35.

A Microfoundations

A.1 Households

The model features a continuum of infinitely lived households that differ only in the way they form expectations. In particular, households can either follow an backward-looking adaptive expectation rule (so called *adaptive learners*) or they can follow the central bank's announcements (the so-called *credibility believers*). In each period, household $i \in [0,1]$ solves an intratemporal problem, by choosing consumption over a continuum of differentiated goods $j \in [0,1]$ with price $P_t(j)$ to minimize expenditures. The elasticity of substitution between different goods is ϵ, so that households choose

$$C_t^i(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\epsilon} C_t^i$$

with aggregate price level P_t

$$P_t = \left(\int_0^1 P_t(j)^{1-\epsilon} dj\right)^{\frac{1}{1-\epsilon}},$$

and total consumption of the household C_t^i

$$C_t^i = \left(\int_0^1 (C_t^i(j))^{\frac{1-\epsilon}{\epsilon}} dj\right)^{\frac{\epsilon}{1-\epsilon}}.$$

Then, household i maximizes her utility function of the following form

$$\tilde{E}_t \sum_{s=t}^{\infty} \beta^s \left[\frac{(C_s^i)^{1-\sigma}}{1-\sigma} - \frac{(H_s^i)^{1+\varphi}}{1+\varphi}\right],$$

subject to a flow-budget constraint

$$P_t C_t^i + B_t^i \leq W_t H_t^i + (1+i_{t-1}) B_{t-1}^i + P_t \int_0^1 \Xi(j) dj),$$

by choosing consumption C_t^i, labor H_t^i and nominal bonds B_t^i, which yields, besides a labor supply trade-off, the standard equation

$$\frac{W_t}{P_t} = (H_t^i)^{\varphi} (C_t^i)^{-\sigma}, \tag{A.1}$$

$$(C_t^i)^{-\sigma} = \beta \tilde{E}_t \left[\frac{(1+i_t)(C_{t+1}^i)^{-\sigma}}{\Pi_{t+1}}\right], \tag{A.2}$$

where W_t is the nominal wage rate, $\int_0^1 \Xi(j) dj$ are aggregate real profits and $\Pi_t = \frac{P_t}{P_{t-1}}$ is gross inflation. Log-linearization of individual i's Euler equation (A.2) yields

$$c_t^i = \tilde{E}_t^i[c_{t+1}^i] - \frac{1}{\sigma}\left(i_t - \tilde{E}_t^i \pi_{t+1} - \bar{r}\right), \tag{A.3}$$

where small letters indicate percentage deviations from steady state, i.e. $y_t = \frac{Y_t - Y}{Y} \approx \log(Y_t/Y)$. Since households are boundedly rational, they are not able to fully optimize over an infinite horizon. Instead, we follow Branch et al. (2012) and assume that our boundedly rational households use N-step Euler equation learning, implying that households use the marginal costs versus marginal benefits trade-off of the Euler equation (A.2) to make decisions given their budget constraint, and given their subjective forecasts of aggregate variables. Assuming that the law of iterated expectations holds at the individual level, we can iterate (A.3) forward to express household i's consumption decision as

$$c_t^i = \tilde{E}_t^i[c_\infty^i] - \frac{1}{\sigma}\tilde{E}_t^i \sum_{j=0}^{\infty} \left(i_{t+j} - \tilde{E}_t^i \pi_{t+j+1} - \bar{r}\right)$$

Next, we make the following assumptions. First, households know that market clearing must hold, i.e. that $x_t = c_t = \int c_t^l dl$, where x_t denotes total production and we index the continuum of households $l \in [0,1]$ to distinguish from a particular household i. Second, households expect market clearing to also hold at the end of some horizon N, i.e. $\tilde{E}_t^i x_{t+N} = \tilde{E}_t^i c_{t+N} = \tilde{E}_t^i \int c_{t+N}^l dl$. Third, we assume agents to know that other agents consumption satisfies their individual Euler equations. Then, we can write the last equation as

$$\tilde{E}_t^i x_{t+N} = \tilde{E}_t^i \int \tilde{E}_{t+1}^l c_\infty^l dl - \tilde{E}_t^i \int \tilde{E}_{t+1}^l \frac{1}{\sigma} \sum_{j=N}^{\infty} \left(i_{t+j} - \tilde{E}_t^i \pi_{t+j+1} - \bar{r}\right)$$

As in Branch and McGough (2009), we assume that the law of iterated expectations also holds at the aggregate level, i.e. that $E_t^i E_{t+k}^l y_{t+k} = E_t^i y_{t+k}$ for any variable y_t. Then, the above expression reduces to

$$\tilde{E}_t^i x_{t+N} = \tilde{E}_t^i \int \tilde{E}_{t+1}^l c_\infty^l dl - \tilde{E}_t^i \frac{1}{\sigma} \sum_{j=N}^{\infty} \left(i_{t+j} - \tilde{E}_t^i \pi_{t+j+1} - \bar{r}\right)$$

Households use this relation to form expectations about the sum of future real interest rates behind their horizon N. Using above we can rewrite iterated Euler equation as

$$\begin{aligned} c_t^i =& \tilde{E}_t^i[c_\infty^i] - \frac{1}{\sigma}\tilde{E}_t^i \sum_{j=0}^{\infty} \left(i_{t+j} - \tilde{E}_t^i \pi_{t+j+1} - \bar{r}\right) \\ =& \tilde{E}_t^i[c_\infty^i] - \frac{1}{\sigma}\tilde{E}_t^i \sum_{j=0}^{N-1} \left(i_{t+j} - \tilde{E}_t^i \pi_{t+j+1} - \bar{r}\right) - \frac{1}{\sigma}\tilde{E}_t^i \sum_{j=N}^{\infty} \left(i_{t+j} - \tilde{E}_t^i \pi_{t+j+1} - \bar{r}\right) \\ =& \tilde{E}_t^i[c_\infty^i] - \tilde{E}_t^i \int \tilde{E}_{t+1}^l c_\infty^l dl + \tilde{E}_t^i x_{t+N} - \frac{1}{\sigma}\tilde{E}_t^i \sum_{k=0}^{N-1} \left(i_{t+j} - \tilde{E}_t^i \pi_{t+j+1} - \bar{r}\right) \end{aligned}$$

Aggregating the consumption decision over all households i gives

$$c_t = \tilde{E}_t^i[c_\infty^i] - \tilde{E}_t \int \tilde{E}_{t+1}^l c_\infty^l dl + \tilde{E}_t x_{t+N} - \frac{1}{\sigma}\tilde{E}_t \sum_{j=0}^{N-1} \left(i_{t+j} - \pi_{t+j+1} - \bar{r}\right),$$

where $\tilde{E}_t y_{t+1} = \int \tilde{E}_t^i [y_{t+1}] = n_{1,t}\tilde{E}_{1,t}[y_{t+1}] + (1 - n_{1,t})\tilde{E}_{2,t}[y_{t+1}]$. Finally, we follow Branch and McGough (2009) and assume that agents agree on terminal wealth and consumption, such that $\tilde{E}_t^i[c_\infty^i] - \tilde{E}_t \int \tilde{E}_{t+1}^l c_\infty^l dl = 0$. Using market clearing we then find the aggregate IS curve

$$x_t = \tilde{E}_t y_{t+N} - \frac{1}{\sigma}\tilde{E}_t \sum_{j=0}^{N-1} \left(i_{t+j} - \pi_{t+j+1} - \bar{r}\right) + e_t,$$

which resembles equation (1) in the paper.

A.2 Firms

The model features a continuum of monopolistic competitive firms that produce final differentiated goods. Each firm is run by a household and therefore follows the same heuristic for predicting future variables as the particular household. We assume nominal price rigidities, following Rotemberg (1982), such that each monopolistic firm j faces a quadratic cost of adjusting nominal prices, measured in terms of the final good, and given by

$$\frac{\psi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)}-1\right)^2 X_t,$$

where ψ measures the degree of nominal price rigidity. Moreover, each firm has a linear production technology of the form

$$X_t(j) = A_t H_t(j),$$

where A_t captures aggregate productivity following a stationary AR(1) process (potentially with zero auto-correlation). Cost-minimization implies that real marginal cost is $MC_t = \frac{w_t}{A_t}$, where $w_t = W_t/P_t$ is the real wage, common to all firms. Nominal flow profit for producer j is given by:

$$\Phi_t = P_t(j)X_t(j) - MC_t X_t(j) - \frac{\psi}{2}\left(\frac{P_t(j)}{P_{t-1}(j)}-1\right)^2 X_t.$$

The problem of firm j involves maximizing future discounted profits

$$\max \sum_{s=0}^{\infty} Q^j_{t,t+s}\Phi_{t+s},$$

subject to the demand constraint

$$X_{t+s}(j) = \left(\frac{P_t(j)}{P_{t+s}}\right)^{-\epsilon} C_{t+s},$$

where $Q^j_{t,t+s} = \beta^s \left(\frac{C^j_{t+s}}{C^j_t}\right)^{-\sigma} \frac{P_t}{P_{t+s}}$ is the stochastic discount factor of the household that runs firm j. After substituting the constraint, the first order condition is

$$(\epsilon-1)X_t(j)+\psi\frac{P_t}{P_{t-1}(j)}\big(\Pi_t(j)-1\big)X_t = \epsilon MC_t X_t(j)\frac{P_t}{P_t(j)}+\psi\tilde{E}^j_t\left[Q^j_{t,t+1}X_{t+1}\frac{P_{t+1}}{P_t(j)}\Pi_{t+1}(j)\big(\Pi_{t+1}(j)-1\big)\right],$$

where $\Pi_t(j)$ is gross-inflation of the good produced by firm j. Multiplying by $\frac{P_t(j)}{P_t X_t}$ and substituting for household j's stochastic discount factor, yields

$$(\epsilon-1)\frac{P_t(j)X_t(j)}{P_t X_t}+\psi\Pi_t(j)\big(\Pi_t(j)-1\big) = \epsilon MC_t\frac{X_t(j)}{X_t}+\psi\beta\tilde{E}^j_t\left[\left(\frac{C^j_{t+1}}{C^j_t}\right)^{-\sigma}\frac{X_{t+1}}{X_t}\Pi_{t+1}(j)\big(\Pi_{t+1}(j)-1\big)\right].$$

Log-linearization around a zero-inflation steady state leaves us with

$$\pi_t(j) = \frac{\epsilon-1}{\psi}\left[p_t - p_t(j) + mc_t + \frac{\beta\psi}{\epsilon mc}\tilde{E}^j_t\pi_{t+1}(j)\right]. \tag{A.4}$$

As above, households (and thus firms) are assumed to form expectations only for N periods. Therefore, iterating the above expression N times (assuming−as above−that the law of iterated expectations holds at the individual level), we can express current price inflation of the good produced by firm j as

$$\pi_t(j) = \tilde{E}_t^j \pi_t(j) + \kappa \tilde{E}_t^j \sum_{s=0}^{N-1} \beta^s \big[p_{t+s} - p_{t+s}(j) + mc_{t+s}\big], \tag{A.5}$$

where we defined $\kappa \equiv \frac{\epsilon-1}{\psi}$ and used the fact that steady state marginal costs $mc = \frac{\epsilon-1}{\epsilon}$. Finally, aggregating over all firms j, we assume that agents are aware of the switching and assign equal probabilities to both heuristics in the future such that $\int_0^1 \tilde{E}_t^j \pi_{t+s}(j)dj = \tilde{E}_t \pi_{t+s}$ for $s = 1, \dots, N$. Then, using the definition of the aggregate price level $p_t = \int_0^1 p_t(j)dj$ and substituting for the marginal cost term mc_t, we find

$$\pi_t = \int_0^1 \pi_t(j)dj = \beta^N \tilde{E}_t \pi_{t+N} + \kappa \tilde{E}_t \sum_{s=0}^{N-1} \beta^s x_{t+s},$$

which yields the N-step ahead Philips curve (2).

B Rational expectations equilibrium

In the absence of autocorrelated shocks the model under rational expectations of both private agents and the central bank coincides with the model under perfect foresight (at least under determinacy). Therefore, we substitute $E_t z_{t+j} = z_t = z^*$ for $z \in \{x, \pi, i\}$ and $j = 1, \dots N$ into the model described by (1) and (2), as well as $\pi^{e,cb} = \pi^*$ into the central bank's policy rule to find

$$\begin{cases} x^* &= x^* - \frac{N}{\sigma}\left(i^* - \pi^* - \bar{r}\right) \\ \pi^* &= \beta^N \pi^* + \frac{1-\beta^N}{1-\beta}\kappa x^* \\ i^* &= \bar{r} + \bar{\pi} + \phi(\pi^* - \bar{\pi}) \end{cases} \Leftrightarrow \begin{cases} x^* &= \frac{1-\beta}{\kappa}\bar{\pi} \\ \pi^* &= \bar{\pi} \\ i^* &= \bar{\pi} + \bar{r} \end{cases} \tag{B.1}$$

That is, the rational expectations equilibrium is characterized by inflation at target $\bar{\pi}$, nominal interest rates at $\bar{\pi} + \bar{r}$ and output being positive, but small, under standard calibration.

C Proof of Proposition 1

Proof. Evaluate (18) in steady state to find that any steady state must satisfy

$$\Bigg\{1 + \frac{\kappa[(N-1)\phi - N]}{\sigma}\left(\frac{1-\beta^N}{1-\beta}\right)\frac{(1-m)^2}{1+m} + \frac{\kappa\phi}{\sigma}\left(\frac{1-\beta^N}{1-\beta}\right)\left(\frac{1-m}{1+m}\right) + \frac{\kappa\phi}{\sigma} - \left(\beta^N - \frac{\kappa[(N-1)\phi - N]}{\sigma}\right)\frac{1-m}{2}\Bigg\}\Big[\pi^* - \bar{\pi}\Big] = 0 \tag{C.1}$$

Evidently, one solution to this equation is $\pi^* = \bar{\pi}$ with $x^* = \frac{1-\beta}{\kappa}\bar{\pi}$, which holds for any $m \in [-1, 1]$. Since both heuristics, predict equally well in this steady state, i.e. both make no prediction error, only $m^* = 0$ is consistent with the heuristic switching model. However, other steady states may exist if the term in the curly brackets equals zero. $\square$

D Asymptotic stationarity: Eigenvalues of the dynamic system in normal times

In this section, we derive the eigenvalues of the system in normal times, described by (19), as a function of the steady state level of the difference in fractions m^*. The corresponding reduced-form Jacobian matrix evaluated in the target steady state (i.e. with $\pi^* = \bar{\pi}, i^* = \bar{r} + \bar{\pi}$ and $x^* = \frac{1-\beta}{\kappa}$) but with general fractions (instead of $m^* = 0$) is given by

$$J = \begin{bmatrix} \frac{1-m^*}{2} - \frac{\phi a^*_{21}}{\sigma} & -\frac{(N-1)\phi-N}{\sigma}\left(\frac{1-m^*}{2}\right) - \frac{\phi a^*_{22}}{\sigma} \\ \kappa\left(\frac{1-\beta^N}{1-\beta}\right)\frac{1-m^*}{2} - \frac{\kappa\phi a^*_{21}}{\sigma} & \left(\beta^N - \frac{\kappa[(N-1)\phi-N]}{\sigma}\right)\frac{1-m^*}{2} - \frac{\kappa\phi a^*_{22}}{\sigma} \end{bmatrix}$$

using equations (E.6) and (E.7) we find the characteristic equation

$$\lambda^2 - \frac{1-m^*}{2}(1+\mu_1-\mu_2)\lambda + \left(\frac{1-m^*}{2}\right)^2[\mu_1+\mu_3] = 0$$

where

$$\begin{aligned}
\mu_1 =& \frac{\sigma}{\sigma+\kappa\phi}\left(\beta^N - \frac{\kappa[(N-1)\phi-N]}{\sigma}\right) \\
\mu_2 =& \frac{\kappa\phi}{\sigma+\kappa\phi}\left(\frac{1-\beta^N}{1-\beta}\right) \\
\mu_3 =& \left(\beta^N - \frac{\kappa[(N-1)\phi-N]}{\sigma}\right)\left[1+(\frac{\kappa\phi}{\sigma+\kappa\phi})^2\right] + \frac{\sigma}{\sigma+\kappa\phi}\left(\frac{1-\beta^N}{1-\beta}\right)\frac{\kappa[(N-1)\phi-N]}{\sigma}
\end{aligned}$$

Then, we eigenvalues as a function of the difference in fractions m_t immediately follow

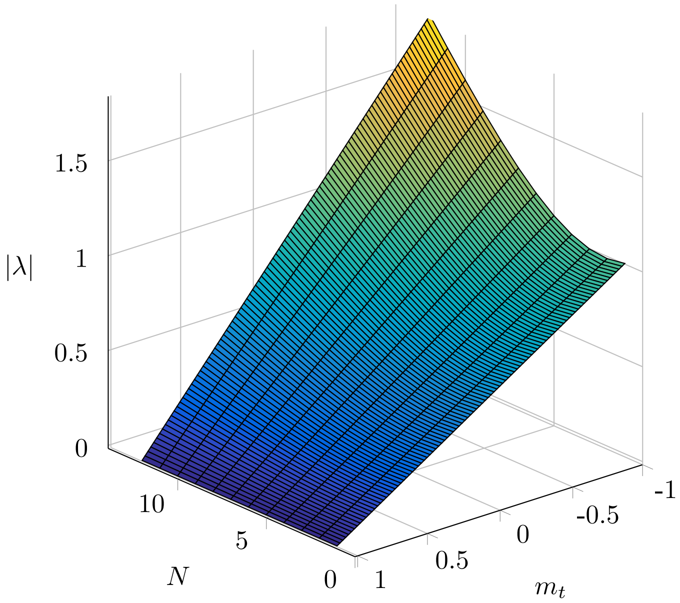
$$\lambda_{1/2}(m^*) = \frac{1}{2}(\frac{1-m^*}{2})\left[(1+\mu_1-\mu_2) \pm \sqrt{(1+\mu_1-\mu_2)^2 - 4\mu_3}\right] \tag{D.1}$$

Clearly, if all agents are credibility believers ($m^* = 1$) both eigenvalues are zero and the system becomes degenerate. However, for an increasing fraction of naive agents (decreasing m^*) the eigenvalues will increase in absolute value and potentially cross the unit circle. In the target steady state we have $m^* = 0$ (i.e. both heuristics perform equally well), so that−for N not too large−both eigenvalues remain inside the unit circle. This is illustrated in Figure 7, where we plot the absolute eigenvalue for $N \in \{1, \ldots, 12\}$ and $m \in (-1, 1)$.

E E-stability in more detail

In this section, we derive the results of Propositions 2 and 3 in more detail. To begin, let us stag both columns of the coefficient matrix A on top of each other to form a 6×1 vector denoted by $\mathbf{a}$ (formally vec$A = \mathbf{a}$) and define the T-map as the mapping from the PLM to the ALM, which

Figure 7: Absolute eigenvalue $|\lambda_1|$ as a function of N and m.



Note: This figure shows the absolute eigenvalue of dynamic system (19) as a function of the difference in fractions m_t and the forward-looking horizon of private households N. For a larger fraction of backward looking households (i.e. decreasing m_t), but also for larger forward-looking horizon N, the eigenvalues increase in absolute value, possibly crossing the unit circle for too large values.

describes the evolution of the RLS estimator $\mathbf{a}$. This mapping can be written as

$$T(\mathbf{a}) = \quad T\begin{pmatrix} a_{10} \\ a_{11} \\ a_{12} \\ a_{20} \\ a_{21} \\ a_{22} \end{pmatrix} = \begin{pmatrix} \left(\frac{1-\beta}{\kappa}\left(\frac{1+m_t}{2}\right) + \frac{(N-1)\phi-N}{\sigma}\left(\frac{1-m_t}{2}\right) + \frac{\phi}{\sigma}\right)\bar{\pi} - \frac{\phi a_{20}}{\sigma} \\ \frac{1-m_t}{2} - \frac{\phi a_{21}}{\sigma} \\ -\frac{(N-1)\phi-N}{\sigma}\left(\frac{1-m_t}{2}\right) - \frac{\phi a_{22}}{\sigma} \\ \left(\frac{1+m_t}{2} + \frac{\kappa\left[(N-1)\phi-N\right]}{\sigma}\left(\frac{1-m_t}{2}\right) + \frac{\kappa\phi}{\sigma}\right)\bar{\pi} - \frac{\kappa\phi a_{20}}{\sigma} \\ \kappa\left(\frac{1-\beta^N}{1-\beta}\right)\frac{1-m_t}{2} - \frac{\kappa\phi a_{21}}{\sigma} \\ \left(\beta^N - \frac{\kappa\left[(N-1)\phi-N\right]}{\sigma}\right)\frac{1-m_t}{2} - \frac{\kappa\phi a_{22}}{\sigma} \end{pmatrix} \tag{E.1}$$

After straight forward algebra we find the fixed point to this mapping that satisfies $T(\mathbf{a}^*) = \mathbf{a}^*$ to be dependent on the difference in fraction m and given by

$$a^*_{10}(m) = \left(\frac{1-\beta}{\kappa}(\frac{1+m}{2}) + \frac{(N-1)\phi - N}{\sigma}(\frac{1-m}{2}) + \frac{\phi}{\sigma}\right)\bar{\pi} - \frac{\phi a^*_{20}}{\sigma} \tag{E.2}$$

$$a^*_{11}(m) = \frac{1-m}{2} - \frac{\phi a^*_{21}}{\sigma} \tag{E.3}$$

$$a^*_{12}(m) = -\frac{(N-1)\phi - N}{\sigma}\left(\frac{1-m}{2}\right) - \frac{\phi a^*_{22}}{\sigma} \tag{E.4}$$

$$a^*_{20}(m) = \frac{\sigma}{\sigma + \kappa\phi}\left(\frac{1+m}{2} + \frac{\kappa\big[(N-1)\phi - N\big]}{\sigma}(\frac{1-m}{2}) + \frac{\kappa\phi}{\sigma}\right)\bar{\pi} \tag{E.5}$$

$$a^*_{21}(m) = \frac{\kappa\sigma}{\sigma + \kappa\phi}\left(\frac{1-\beta^N}{1-\beta}\right)\frac{1-m}{2} \tag{E.6}$$

$$a^*_{22}(m) = \frac{\sigma}{\sigma + \kappa\phi}\left(\beta^N - \frac{\kappa\big[(N-1)\phi - N\big]}{\sigma}\right)\frac{1-m}{2} \tag{E.7}$$

where the fixed points for a^*_{10}, a^*_{11} and a^*_{12} depend on those of a^*_{20}, a^*_{21} and a^*_{22}, respectively, which are defined by (E.5)-(E.7). Thus, we have a unique fixed point of the T-map depending on the difference in fractions $m \in [-1, 1]$, which leads to Propositions 2.

To determine whether the estimated coefficients converge to their respective fixed point for any of the given steady states, consider the following differential equation

$$\frac{d}{d\tau}(\mathbf{a}) = T(\mathbf{a}) - \mathbf{a} \tag{E.8}$$

where τ describes notional or artificial time. Component-by-component we get

$$\frac{da_{10}}{d\tau} = \left(\frac{1-\beta}{\kappa}(\frac{1+m}{2}) + \frac{(N-1)\phi - N}{\sigma}(\frac{1-m}{2}) + \frac{\phi}{\sigma}\right)\bar{\pi} - \frac{\phi a_{20}}{\sigma} - a_{10} \tag{E.9}$$

$$\frac{da_{11}}{d\tau} = \frac{1-m}{2} - \frac{\phi a_{21}}{\sigma} - a_{11} \tag{E.10}$$

$$\frac{da_{12}}{d\tau} = -\frac{\big[(N-1)\phi - N\big]}{\sigma}\frac{1-m}{2} - \frac{\phi a_{22}}{\sigma} - a_{12} \tag{E.11}$$

$$\frac{da_{20}}{d\tau} = \left(\frac{1+m}{2} + \frac{\kappa\big[(N-1)\phi - N\big]}{\sigma}(\frac{1-m}{2}) + \frac{\kappa\phi}{\sigma}\right)\bar{\pi} - \left(1 + \frac{\kappa\phi}{\sigma}\right)a_{20} \tag{E.12}$$

$$\frac{da_{21}}{d\tau} = \left(\frac{1-\beta^N}{1-\beta}\right)\frac{1-m}{2} - \left(1 + \frac{\kappa\phi}{\sigma}\right)a_{21} \tag{E.13}$$

$$\frac{da_{22}}{d\tau} = \left(\beta^N - \frac{\kappa\big[(N-1)\phi - N\big]}{\sigma}\right)\frac{1-m}{2} - \left(1 + \frac{\kappa\phi}{\sigma}\right)a_{22} \tag{E.14}$$

Now, note that the system described by (E.9)-(E.14) is linear and the equations for (a_{10}, a_{20}), (a_{11}, a_{21}) and (a_{12}, a_{22}) are independent. We rewrite (E.8) to get $\frac{d}{d\tau}(a) = M_0 + M_1\mathbf{a}$ with M_1

being a corresponding 6×6 matrix given by[28]

$$M_1 = \begin{bmatrix} -1 & 0 & 0 & -\frac{\phi}{\sigma} & 0 & 0 \\ 0 & -1 & 0 & 0 & -\frac{\phi}{\sigma} & 0 \\ 0 & 0 & -1 & 0 & 0 & -\frac{\phi}{\sigma} \\ 0 & 0 & 0 & -\left(1+\frac{\kappa\phi}{\sigma}\right) & 0 & 0 \\ 0 & 0 & 0 & 0 & -\left(1+\frac{\kappa\phi}{\sigma}\right) & 0 \\ 0 & 0 & 0 & 0 & 0 & -\left(1+\frac{\kappa\phi}{\sigma}\right) \end{bmatrix}$$

The 6×1 vector M_0 depends in general on the variable m_t and is therefore time-varying. Yet, since m_t is bounded, so will M_0. Further, the matrix M_1 is independent of m_t with three eigenvalues taking on the value -1 and the other three being equal to $-(1+\frac{\kappa\phi}{\sigma})$. Since all parameters are positive, all eigenvalues are negative and thus the differential equation is locally stable. This is summarized in Propositions 3.

F Proof of Proposition 4

Proof. We begin with assuming constant fractions, i.e. $m_t = m$. with $m \in (-1, 1)$. Under this assumption and condition (23), which implies that $\mathbb{1}_n = 0$ in steady state, the dynamic system in (25) and (26) evaluated in steady state reduces to

$$x^* = \left(\frac{1-\beta}{\kappa} + \frac{1}{\sigma}\right)\bar{\pi} + \frac{N}{\sigma}\left(\frac{1-m}{1+m}\right)\pi^* + \left[\frac{2N}{1+m} - (N-1)\right]\frac{\bar{r}}{\sigma}$$
$$\pi^* = \frac{1+m}{2}\left(1+\frac{\kappa}{\sigma}\right)\bar{\pi} + \kappa\left(\frac{1-\beta^N}{1-\beta}\right)\frac{1-m}{2}x^* + \frac{1-m}{2}\left(\beta^N + \frac{\kappa N}{\sigma}\right)\pi^* + \frac{\kappa}{\sigma}\left[N - \frac{(N-1)(1+m)}{2}\right]\bar{r}$$

from which we can derive the following steady state expression for inflation π^*_{zlb} and output x^*_{zlb}, respectively, given any constant fractions $m \in (-1, 1)$

$$\begin{aligned} \pi^*_{zlb} =& \frac{(1+m)^2\left(1+\frac{\kappa}{\sigma}\right) + (1-\beta^N)\left(1+\frac{\kappa}{\sigma(1-\beta)}\right)(1-m^2)}{2(1+m) - \frac{\kappa N}{\sigma}\left(\frac{1-\beta^N}{1-\beta}\right)(1-m)^2 - \left(\beta^N + \frac{\kappa N}{\sigma}\right)(1-m^2)}\bar{\pi} \\ &+ \frac{\frac{\kappa}{\sigma}\left[2N - (N-1)(1+m)^2 + (1-m^2)\left(\frac{1-\beta^N}{1-\beta}\right)\left(1+\frac{2N}{1+m} - N\right)\right]}{2(1+m) - \frac{\kappa N}{\sigma}\left(\frac{1-\beta^N}{1-\beta}\right)(1-m)^2 - \left(\beta^N + \frac{\kappa N}{\sigma}\right)(1-m^2)}\bar{\pi} \end{aligned} \tag{F.1}$$

$$x^*_{zlb} = \left(\frac{1-\beta}{\kappa} + \frac{1}{\sigma}\right)\bar{\pi} + \left[\frac{2N}{1+m} - (N-1)\right]\frac{\bar{r}}{\sigma} + \frac{N}{\sigma}\left(\frac{1-m}{1+m}\right)\pi^*_{zlb} \tag{F.2}$$

and for which π^*_{zlb} has a vertical asymptote at

$$\tilde{m}_1 = \frac{-\alpha_2 + \sqrt{\alpha_2^2 - 4\alpha_1\alpha_3}}{2\alpha_1}$$

[28] Alternatively, we can look at the stability of the three subsystems (a_{10}, a_{20}), (a_{11}, a_{21}) and (a_{12}, a_{22}) as in Honkapohja and Mitra (2005), using the results of Honkapohja and Mitra (2006).

where

$$\alpha_1 = \beta^N - \frac{\kappa N}{\sigma}(\frac{\beta - \beta^N}{1-\beta}) \gtrless 0$$
$$\alpha_2 = 2\left[1 + \frac{\kappa N}{\sigma}(\frac{1-\beta^N}{1-\beta})\right] > 0$$
$$\alpha_3 = 2 - \beta^N - \frac{\kappa N}{\sigma}(\frac{2-\beta-\beta^N}{1-\beta}) \gtrless 0$$

and $\tilde{m}_2 \notin [-1, 1]$. Next, recognize that the numerator of (F.1) is always positive. The denominator, however, is negative for $m \in (-1, \tilde{m}_1)$ and π^*_{zlb} approaches $-\infty$ for $m \to \tilde{m}_1^-$, while it is positive for $m \in (\tilde{m}_1, 1)$ and π^*_{zlb} approaches ∞ for $m \to \tilde{m}_1^+$. The latter is in contradiction with condition (23), since π^*_{zlb} needs to be smaller than zero for the ZLB steady state to exist, given any positive fraction of credibility believers. If $m = -1$, then the steady state satisfies $\pi^*_{zlb} = -\bar{r}, i^*_{zlb} = 0$ and $x^*_{zlb} = \frac{-(1-\beta)\bar{r}}{\kappa}$, while for $m = 1$, the system is again degenerate and not consistent with the ZLB condition. Hence, we conclude that a steady state with $\pi^*_{zlb} \in (-\infty, 0]$ in the ZLB region exists for $m \in [-1, \tilde{m}_1)$.

Lastly, consider the case of time-varying fractions. As in the case of fixed fractions, we have a unique ZLB steady state under the HSM. However, the interval for m translates into an interval for the intensity of choice b, because m itself is endogenous. To see this, rewrite the difference in fractions as

$$m_t = n_{1,t} - n_{2,t} = \frac{\exp bU_{1,t-1} - \exp bU_{2,t-1}}{\exp bU_{1,t-1} + \exp bU_{2,t-1}} = \tanh\left(\frac{b}{2}\Delta U_{t-1}\right)$$

Evaluating this equation at the ZLB steady state and noting that naive agents do not make any forecast errors, we find

$$m^* = \tanh(\frac{b}{2}\Delta U^*) \tag{F.3}$$

with

$$\Delta U^* = -(\pi^*_{zlb} - \bar{\pi})^2 - (x^*_{zlb} - \bar{x})^2 - (-\bar{r} - \bar{\pi})^2 < 0$$

Solving (F.3) for the intensity of choice, and recognizing that $\tanh^{-1}$ is a strictly increasing function on the interval $(-1, 1)$, the upper bound $\tilde{m}_1$ translates into a lower bound $\underline{b}$ on the intensity of choice

$$\underline{b} = \frac{2}{\Delta U^*} \tanh^{-1}(\tilde{m}_1)$$

In other words, for too low values of the intensity of choice, i.e. $b \to 0$, the fractions become constant and ZLB steady exists only if $m \in [-1, \tilde{m}_1)$. Contrary, if $b \to \infty$, then $\tilde{m}_1 \to -1$ and there is a unique ZLB steady state with all agents being naive. Lastly, note that if $\tilde{m}_1 \geq 0$ (which is the case for our baseline calibration) the lower bound will satisfy $\underline{b} \leq 0$ such that the the ZLB steady state always exists.

Finally, for the this candidate steady state to be consistent with the ZLB region, we must additionally have that the ZLB constraint is binding, i.e. equation (23) must hold. Therefore, suppose $b = \infty$ (or equivalently $m^* - = 1$). Then, $\pi^*_{zlb} = -\bar{r}$ and the ZLB condition reduces to

$$-\bar{r} \leq \bar{\pi} - \phi^{-1}(\bar{r} + \bar{\pi}) \qquad \Leftrightarrow \qquad \phi > 1$$

That is, for the ZLB steady state to exist in the all-naive case, we need the Taylor principle to hold. Next, recognize that $\frac{\partial \pi^*_{zlb}}{\partial m} < 0$ so that for $b < \infty$ (and equivalently $m^* > -1$), the ZLB condition will be satisfied in steady state if the Taylor principle holds. □

G Eigenvalues of the dynamic system in the ZLB region

G.1 Eigenvalues of the dynamic system without forward guidance

In this section, we derive the eigenvalues of the system in the ZLB region. In the absence of forward guidance, the model is described by (25) and (26). For analytical purpose, we assume an infinite intensity of choice $b = \infty$, which implies that all households immediately switch to the best performing heuristic. Under this assumption, four eigenvalues are equal to zero and the corresponding reduced-form Jacobian matrix evaluated in the ZLB steady state $(\pi^*_{zlb}, x^*_{zlb}, i^*_{zlb})$ with $m^* = -1$ is given by

$$J = \begin{bmatrix} 1 & \frac{N}{\sigma} \\ \kappa\left(\frac{1-\beta^N}{1-\beta}\right) & \beta^N - \frac{\kappa N]}{\sigma} \end{bmatrix}$$

with the corresponding eigenvalues

$$\lambda_{1/2} = \frac{1}{2}\left(1 + \beta^N + \frac{\kappa N}{\sigma}\right) \pm \frac{1}{2}\sqrt{\left(1 + \beta^N + \frac{\kappa N}{\sigma}\right)^2 - 4\left(\beta^N - \frac{\kappa N}{\sigma}\left(\frac{\beta - \beta^N}{1-\beta}\right)\right)} \tag{G.1}$$

Figure 8 illustrates that for our baseline calibration of Table 1, at least one eigenvalue remains outside the unit circle, so that the ZLB steady state is a saddle point. Moreover, for large enough N, the ZLB steady state even becomes an unstable node.

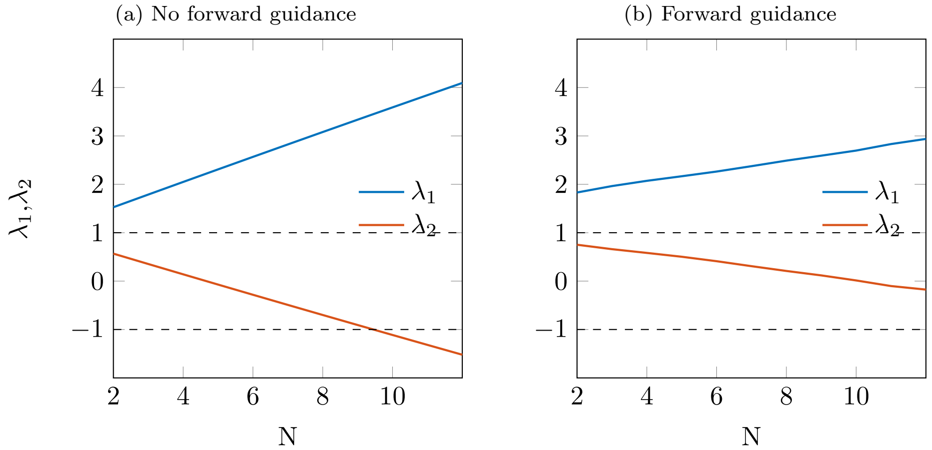
G.2 Eigenvalues of the dynamic system with forward guidance

A crucial question is, how does the eigenvalues change if we allow for forward guidance policy? As mentioned earlier, asymptotic stationarity of the system now depends on the eigenvalues of the term $\Lambda_1(m^*) + \sum_{j=1}^{N} \Lambda_{j+1}(m^*)(A_1^*)^{j+1}$. No analytical expression could be derived. The right panel of Figure 8 therefore presents the numerical results given our benchmark calibration. Forward guidance reduces the slope with which they increase in N, yet does not change asymptotic stationarity properties of the model.

H Proof of Proposition 5

Proof. As in the proof to Proposition 4, we begin with assuming fixed fractions $m \in [-1, 1]$. Further, imposing steady state and assuming condition (23) is satisfied, the unique solution to the dynamic system (27) and (28) is given by $(\pi^*_{fg}, x^*_{fg}, i^*_{fg}) = (-\bar{r}, -\frac{(1-\beta)\bar{r}}{\kappa}, 0)$. That is, for $N \leq q$ the ZLB steady state under constant fractions exists independent of the share of credibility believers.

Figure 8: Largest eigenvalues λ_1 and λ_2 of the ZLB steady state ($m = -1$) as a function of N.



Note: This figure shows both eigenvalues of the reduced form Jacobian of the dynamic system in the ZLB region, with (right panel) or without forward guidance (left panel) for different values of the forward-looking parameter N. Other parameter values taken from the benchmark calibration in Table 1. Evident from the Figure, at least one of the two eigenvalues is outside the unit circle, implying that the underlying steady state is a saddle point. However, for N too large, both eigenvalues turn negative and the steady state becomes a source. Forward guidance reduces the slope with which they increase in N, yet does not change asymptotic stationarity properties of the model.

Formally, we have

$$\begin{cases} x^* = \frac{1-m}{2}x^* + \frac{N}{\sigma}\left(\frac{1-m}{2}\right)\pi^* + \frac{1+m}{2}x^* + \frac{N}{\sigma}\left(\frac{1+m}{2}\right)\pi + \frac{N\bar{r}}{\sigma} \\ \pi^* = \kappa\left(\frac{1-\beta^N}{1-\beta}\right)\left(\frac{1-m}{2}\right)x^* + \left(\beta^N + \frac{N}{\sigma}\right)\left(\frac{1-m}{2}\right)\pi^* + \frac{\kappa(2-\beta^N-\beta)}{1-\beta}\left(\frac{1+m}{2}\right)x^* + \frac{\kappa N\bar{r}}{\sigma} \end{cases} \Leftrightarrow \begin{cases} \pi^*_{fg} = -\bar{r} \\ x^*_{fg} = -\frac{(1-\beta)\bar{r}}{\kappa} \end{cases}$$

Under time-varying fractions, the steady state necessarily features $m^* = 0$. The reason is that, through the forward guidance announcements, even the credibility believers predict correctly in the ZLB steady state. Thus, if both types of heuristics perform equally well, fractions must be equal for any $b > 0$.

As before, condition (23) needs to be satisfied, which now implies $\phi > 1$ for all $b > 0$ or $m^* \in [-1, 1]$. However, given the feedback of central bank expectations to the real economy, we additionally need a fix point of the T-map to exist. Formally, we need A_0 and A_1 to satisfy

$$A_0 = \Lambda_0(m^*) + \sum_{j=1}^{N} \Lambda_{j+1}(m^*)\left(I - A_1^{j+1}\right)(I - A_1)^{-1}A_0$$

$$A_1 = \Lambda_1(m^*) + \sum_{j=1}^{N} \Lambda_{j+1}(m^*)(A_1^*)^{j+1}$$

□

I Robustness Analysis

In this section, we are going to change some of the coefficients which are either crucial for our results, or for which no common values exist. Specifically, we discuss the effects of different calibrations for the forward-looking horizons N and the intensities of choice b in Section I.1, while we show that our main result is robust to different length of the forward guidance horizon q or forward-looking horizons N in Section I.2. Finally, we relax the assumption that households know the functional form and parameterization of the policy function in the Section I.3. We then redo our Monte Carlo simulations from Table 2 and show that forward guidance remains effective, however, given the limited information Odyssean guidance appears to increase volatility even more, while Delphic guidance becomes more powerful and thus welfare maximizing. As discussed earlier, this assumption ultimately only affects the expectations of the adaptive learners, as credibility believers expect variables to be at their target and otherwise adopt the central bank's forecasts.

I.1 Different parameterization

One of our main assumptions is that private households have a longer, but finite horizon for which they form explicit expectations as proposed by Branch et al. (2012). Generally, the stability of the dynamic system in normal times depends crucially on N. As shown in Appendix D for the special case with $\omega = 1$, an increase in N will, de facto, lead to an increase in the eigenvalues of the system. Nearly self-fulfilling expectations and slow convergence to the target steady state would be the result. Below we vary N and show that (i) deflationary spirals become more likely when N is larger, but also (ii) forward guidance with longer horizons becomes much more powerful (see, Table 4). The other results remain qualitatively the same. In Appendix J, we show that infinite horizon learning does not satisfy asymptotic stationarity and that, therefore, we have to rely on N-step Euler equation learning.

Also the intensity of choice is a parameter for which no consensus among behavioral economists exists. As already discussed in the Footnote 21 the calibration of b crucially depends on both the definition and the unit of measurement of the fitness measure U_t. However, the model dynamics depend on the size of the intensity of choice in a non-trivial way. Indeed, we can think of two opposing effects resulting from an increase in the intensity of choice. Firstly, a lower intensity of choice enlarges the region of initial values for which recovery occurs, because when output and inflation are very low for some periods, there will still be a significant fraction of private households who believe in the central bank's ability to move the economy back into the target steady state and thereby exercise an upward pressure on both inflation and output. Contrary, a lower intensity of choice can be destabilizing, if the economy is in a liquidity trap, with inflation and output falling. Then, even when at some point adaptive expectations perform worse, a lower intensity of choice will lead to the central bank credibility rising very slowly and thus potentially preventing the system to converge back to the target steady state.

I.2 Different forward guidance horizons and the role of N

Above we assumed that the forward guidance horizon of the central bank q equals the forward-looking horizon of the agents N. In this section, we will show that changing forward guidance horizons has implications for the steady states, but not on the learnability results. Note, that forward guidance announcements too far in the future−those that exceed the private households' forward-looking horizon−are not taken into account. Therefore, we focus on those cases in which $q < N$. Equation (29) becomes

$$y_t \equiv \begin{pmatrix} x_t \\ \pi_t \end{pmatrix} = \tilde{\Lambda}_0(m_t) + \Lambda_1(m_t) \begin{pmatrix} x_{t-1} \\ \pi_{t-1} \end{pmatrix} + \Lambda_2(m_t) \begin{pmatrix} x^{e,cb}_{t+1|t} \\ \pi^{e,cb}_{t+1|t} \end{pmatrix} + \ldots + \Lambda_{q+1}(m_t) \begin{pmatrix} x^{e,cb}_{t+q|t} \\ \pi^{e,cb}_{t+q|t} \end{pmatrix} \tag{I.1}$$

where matrices $\Lambda_1(m_t), \dots, \Lambda_q(m_t)$ remain unchanged, but the vector $\tilde{\Lambda}_0(m_t)$ becomes

$$\tilde{\Lambda}_0(m_t) = \begin{pmatrix} \left[\left(\frac{1-\beta}{\kappa} + \sigma^{-1}\right)\left(\frac{1+m_t}{2}\right) + \frac{\mathbb{1}_n(N-1)(\phi-1)}{\sigma}\right]\bar{\pi} + \left[\left(\frac{1-m_t}{2}\right)[\mathbb{1}_n - (1-\mathbb{1}_n)N] - \left(\frac{1+m_t}{2}\right)(q+1)\right]\frac{\bar{r}}{\sigma} \\ \left[\left(1 + \beta^{q+1} - \beta + \frac{\kappa}{\sigma}\right)\left(\frac{1+m_t}{2}\right) + \frac{\kappa\mathbb{1}_n[(N-1)(\phi-1)]}{\sigma}\right]\bar{\pi} + \left[\left(\frac{1-m_t}{2}\right)[\mathbb{1}_n - (1-\mathbb{1}_n)N] - \left(\frac{1+m_t}{2}\right)(q+1)\right]\frac{\kappa\bar{r}}{\sigma} \end{pmatrix}$$

Therefore, the ALM is

$$y_t = \left[\tilde{\Lambda}_0 + \sum_{j=1}^{q} \Lambda_{j+1}\left(I - A_1^{j+1}\right)\bar{A}_0\right] + \left[\Lambda_1 + \sum_{j=1}^{q} \Lambda_{j+1} A_1^{j+1}\right] y_{t-1} \tag{I.2}$$

The system's local stability properties are again determined by the matrix expression in the latter square brackets and Proposition 4 applies.

In the simulations, we find that the length of the guidance horizon is crucial. In particular, we do the same experiment as in Table 2, but use a guidance horizon of $q = 2$ instead of $q = 4$ and otherwise baseline calibration. The results are presented in Table 3, which indicate that guidance with a longer horizon leads to a lower likelihood of a deflationary spiral. Otherwise, the qualitative results are similar to those found with $q = 4$. The same applies when changing the forward-lookingness parameter N. Table 4 shows the results for $N = 6$ and $q = 2, 4$ or 6.

Table 3: Likelihood of deflationary spirals for different guidance horizon $q = 2$

	US Data 1988Q1 - 2017Q4	*without* forward guidance	Delphic guidance	Odyssean guidance	Both
Likelihood (in %)	23.33	19.91	12.22	11.41	11.78
avg. SD x_t (in %)	1.93	2.41	1.69	1.82	1.69
avg. SD π_t (in %)	0.87	0.92	0.69	0.73	0.68
avg. Welfare W · 100		-0.3689	-0.3458	-0.359	-0.3454

Note: The table presents the relative share of 10,000 Monte Carlo simulations in which a deflationary spiral occurred under different policies. The forward-looking parameter N is unchanged, but the forward guidance horizon is reduced to $q = 2$. The lower the likelihood, the less likely it is for the economy to be locked in a liquidity trap. The second and third rows present the average standard deviations of output and inflation, while the last row shows the average of our measure of ex post welfare, with averages taken over all simulations. All other parameters are calibrated as in Table 1. With regards the data, we use output gap data from the Congressional Budget Office (CBO), while data for the federal funds rate to calculate the % of ZLB periods and PCE inflation are taken from the Federal Reserve Economic Data (FRED).

Although we do not want to take a stance on the optimal horizon of forward guidance in this paper, the results suggest that longer guidance horizons generally decrease the likelihood of deflationary spirals, yet with increasing macroeconomic volatility under Odyssean guidance.[29]

I.3 Assumption on policy rule is known

In the main model households know the functional form and the parameters of the central bank's policy rule. This allows us to derive some analytical expressions and stability conditions of the model due to a reduced state space. Although this seems to be a stark assumption, we argue that in fact, only the expectations of the adaptive learners are affected and that quantitatively, this assumption does not play a significant role. To show this, we redo our Monte Carlo simulations presented in Table 2 under the following specification for adaptive learners interest rate

[29] For a discussion on optimal forward guidance horizon see, e.g. Bilbiie (2016).

Table 4: Guidance and the likelihood of deflationary spirals: $N = 6$ and various q

	Guidance horizon	US Data 1988Q1 - 2017Q4	*without* forward guidance	Delphic guidance	Odyssean guidance	Both
Likelihood (in %)		23.33	26.00	19.3	13.45	18.35
avg. SD x_t (in %)	$q = 2$	1.93	3.24	2.56	2.14	2.5
avg. SD π_t (in %)		0.87	1.13	0.92	0.79	0.9
avg. Welfare $\mathbb{W} \cdot 100$			-0.4033	-0.3827	-0.3806	-0.3818
Likelihood (in %)				12.61	8.67	11.05
avg. SD x_t (in %)	$q = 4$			2.09	1.92	2.11
avg. SD π_t (in %)				0.77	0.71	0.75
avg. Welfare $\mathbb{W} \cdot 100$				-0.3793	-0.3719	-0.3782
Likelihood (in %)				10.51	6.5	7.15
avg. SD x_t (in %)	$q = 6$			1.92	1.9	2.04
avg. SD π_t (in %)				0.7	0.73	0.74
avg. Welfare $\mathbb{W} \cdot 100$				-0.3784	-0.3682	-0.3772

Note: The table presents the likelihood of deflationary spirals in 10,000 Monte Carlo simulations under different policies. The forward-looking parameter $N = 6$, and we vary the forward guidance horizon from $q = 2$ to $q = 6$. The lower the likelihood, the less likely it is for the economy to be locked in a liquidity trap. The table also presents the average standard deviations of output and inflation as well as the average ex post welfare (with averages taken over all simulations) for each policy and guidance horizon. All other parameters are calibrated as in Table 1. With regards the data, we use output gap data from the Congressional Budget Office (CBO), while data for the federal funds rate to calculate the % of ZLB periods and PCE inflation are taken from the Federal Reserve Economic Data (FRED).

expectations

$$\tilde{E}_{2,t} i_{t+j} = \tilde{E}_{2,t-1} i_t + \omega(i_{t-1} - \tilde{E}_{2,t-1} i_t).$$

The counterfactual is presented in Table 5. The results again confirm that forward guidance can reduce the likelihood of deflationary spirals considerably. Moreover, the results also confirm earlier studies indicating that knowing the policy rule of the central bank has a stabilizing effect (see e.g. Eusepi, 2005; Eusepi and Preston, 2010). Moreover, it appears that if the policy rule is unknown to agents, Delphic guidance is more effective in reducing the likelihood of deflationary spirals, without increasing the macroeconomic volatility. Intuitively, when the policy rule is unknown to agents, providing conditional forecasts for inflation, output and especially the nominal interest rates is even more stabilizing.

Table 5: The likelihood of deflationary spirals if policy function unknown

	US Data 1988Q1 - 2017Q4	*without* forward guidance	Delphic guidance	Odyssean guidance	Both
Likelihood (in %)	23.33	30.97	12.38	30.37	14.15
avg. SD x_t (in %)	1.93	3.61	2.03	5.04	3.14
avg. SD π_t (in %)	0.87	1.37	0.87	2.04	1.29
avg. Welfare $\mathbb{W} \cdot 100$		-0.4048	-0.722	-0.4348	-0.385

Note: The table presents the relative share of 10,000 Monte Carlo simulations in which a deflationary spiral occurred for different policies under the assumption that the central bank's policy function is unknown to the households. That is, households neither know its functional form nor its parameterization. The calibration is given in Table 1. The lower the likelihood, the less likely it is for the economy to be locked in a liquidity trap. The second and third rows present the average standard deviations of output and inflation, while the last row shows the average of our measure of ex post welfare, with averages taken over all simulations. With regards the data, we use output gap data from the Congressional Budget Office (CBO), while data for the federal funds rate to calculate the % of ZLB periods and PCE inflation are taken from the Federal Reserve Economic Data (FRED).

J Infinite horizon learning version

In the model presented in this paper households make consumption decisions to satisfy their N-step ahead perceived Euler equations following Branch et al. (2012). Regarding the choice of the parameter N, we provide a detailed analysis of the implications for stability and show that our qualitative results remain robust to different calibrations. In this section, we illustrate that the alternative expectation formation process, i.e. the infinite horizon approach of Eusepi and Preston (2018), is not asymptotically stationary when combined with the heuristic switching model.[30] Consider the following aggregate demand and supply equation under arbitrary beliefs[31]

$$\begin{aligned} x_t =& \tilde{E}_t \sum_{T=t}^{\infty} \beta^{T-t}[(1-\beta)x_{T+1} - \sigma^{-1}(i_t - \pi_{T+1} - \bar{r})] + e_t, \\ \pi_t =& \frac{1-\gamma_1}{\gamma_1} \tilde{E}_t \sum_{T=t}^{\infty} (\gamma_1\beta)^{T-t}[(1-\gamma_1\beta)x_T + \gamma_1\beta\pi_{T+1}] + u_t, \end{aligned}$$

where γ_1 is a composite parameter defined as $\gamma_1 = \frac{1}{2}\left(1+\kappa+\beta-\sqrt{(1+\kappa+\beta)^2-4\beta}\right)$. All other variables and parameters are defined the same way as in the main model. Substituting private household expectations and considering the case with a non-binding ZLB constraint, the model becomes

$$\begin{aligned} x_t =& \left(\frac{1-m_t}{2} - \frac{\phi}{\sigma}a_{21}\right)x_{t-1} - \left[\frac{\phi\beta-1}{\sigma(1-\beta)}\left(\frac{1-m_t}{2}\right) + \frac{\phi}{\sigma}a_{22}\right]\pi_{t-1} \\ &+ \left[\frac{1+m_t}{2}\left(\frac{1-\beta}{\kappa} + \sigma^{-1}\right) + \left(1 + \frac{\beta}{1-\beta}\left(\frac{1-m_t}{2}\right)\right)\frac{\phi-1}{\sigma}\right]\bar{\pi} - \frac{\phi}{\sigma}a_{20} \qquad \text{(J.1)} \\ \pi_t =& \left[\left(\kappa + (1-\gamma_1)\beta\right)\frac{1-m_t}{2} - \frac{\kappa\phi}{\sigma}a_{21}\right]x_{t-1} + \left[\left(\frac{\gamma_1\beta}{1-\beta\gamma_1} - \frac{\kappa(\beta\phi-1)}{\sigma(1-\beta)}\right)\frac{1-m_t}{2} - \frac{\phi\kappa}{\sigma}a_{22}\right]\pi_{t-1} \\ &+ \left\{\left[\frac{\kappa}{\sigma} + \frac{1-\beta}{\kappa}\left[\kappa + (1-\gamma_1)\beta\right] + \frac{\gamma_1\beta}{1-\beta\gamma_1}\right]\frac{1+m_t}{2} + \frac{\kappa(\phi-1)}{\sigma}\left[1 + \frac{\beta}{1-\beta}\left(\frac{1-m_t}{2}\right)\right]\right\}\bar{\pi} \qquad \text{(J.2)} \\ &- \frac{\kappa\phi}{\sigma}a_{20} \end{aligned}$$

To consider asymptotic stationarity, evaluate the Jacobian at the target steady state and recognize that the Jacobian reduces to the following 2×2 matrix

$$J = \begin{bmatrix} \frac{1}{2} - \frac{\phi}{\sigma}a_{21}^* & -\frac{\phi\beta-1}{2\sigma(1-\beta)} - \frac{\phi}{\sigma}a_{22}^* \\ \frac{\kappa+(1-\gamma_1)\beta}{2} - \frac{\kappa\phi}{\sigma}a_{21}^* & \frac{1}{2}\left(\frac{\gamma_1\beta}{1-\beta\gamma_1} - \frac{\kappa(\beta\phi-1)}{\sigma(1-\beta)}\right) - \frac{\phi\kappa}{\sigma}a_{22}^* \end{bmatrix}, \qquad \text{(J.3)}$$

where

$$a_{21}^* = \frac{\sigma[\kappa+(1-\gamma_1)\beta]}{2(\sigma+\kappa\phi)}, \quad \text{and} \quad a_{22}^* = \frac{\sigma}{2(\sigma+\kappa\phi)}\left[\frac{\gamma_1\beta}{1-\gamma_1\beta} - \frac{\kappa(\beta\phi-1)}{\sigma(1-\beta)}\right].$$

Unfortunately, the analytical expression of the eigenvalues does not allow for any meaningful interpretation. It is therefore that we again have to compute the eigenvalues numerically. In Figure 9 we vary several key parameters and show the absolute value of the largest eigenvalues for

[30] Asymptotic stationarity here means that when the central bank's forecasting model has converged to its fixed point and private households make no forecast errors in the deterministic steady state.

[31] The model resembles the one presented in Eusepi and Preston (2010) with Rotemberg pricing, but for general intertemporal elasiticty of substitution σ. The detailed derivations are omitted here but can be requested from the authors.

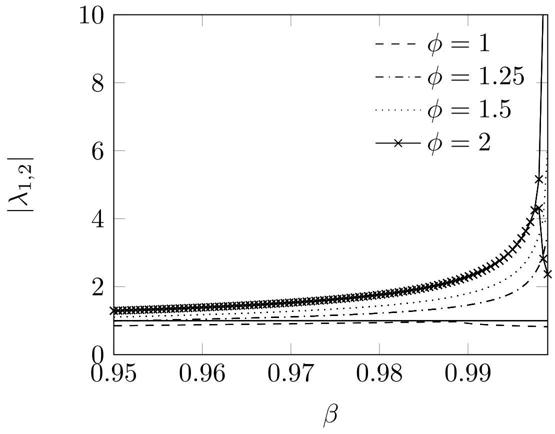
different values for the inflation response $\phi = \{1, 1.25, 1.5, 2\}$. Other parameters follow from the baseline calibration Table 1.

Evident from the Figure 9, only for low values of β and ϕ the system is asymptotically stable. These calibration, however, imply steady state real interest rates in the area of 4-5%, which is far away from the mid-point estimates of the steady state real interest rates in the recent years.[32] We find that varying these other parameters, in the range that is considered empirically plausible, does not affect stationarity considerably and usually imply a asymptotically non-stationary model. Figure 9 also plots the absolute value of the largest eigenvalue for different values of the price adjustment parameter ψ or the inverse elasticity of substitution σ. Generally, all three exercises indicate that that for standard values inflation response ϕ and a discount factor consistent with a low steady state real interest rate, the model with heuristic switching under infinite horizon learning is not asymptotically stable.

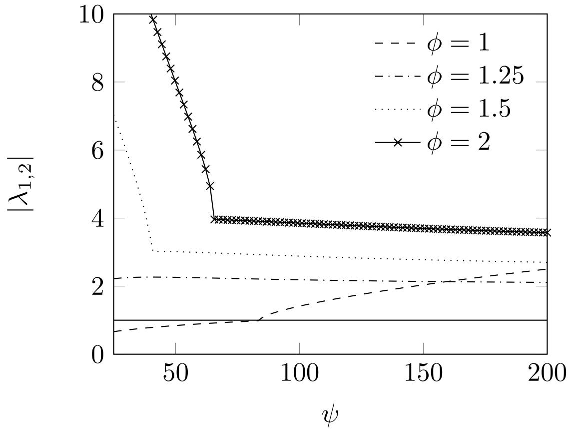
[32] For instance, Laubach and Williams (2016) indicate the natural rate in the United States fell to close to zero during the crisis and has remained there into 2016.

Figure 9: Largest eigenvalues λ_1 and λ_2 of the infinite horizon learning model

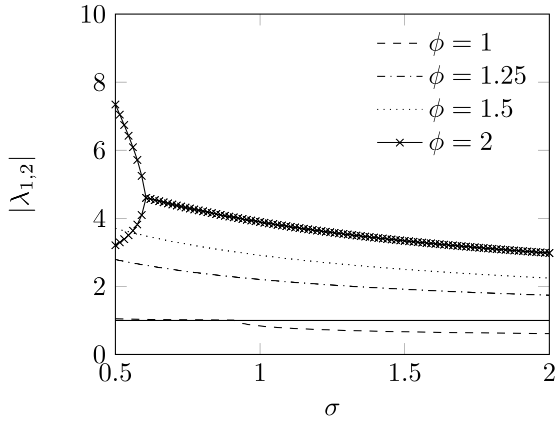
(a) Discount factor rigidity



(b) Price adjustment cost



(c) Elasticity of substituion



Note: The Figure shows the absolute value of the largest eigenvalue of the reduced form Jacobian (J.3) given different inflation response coefficients ϕ. In panel (a) we vary the discount factor β, in panel (b) we vary the price adjustment parameter ψ and finally in panel (c) we vary the inverse elasticity of substitution σ.

Previous DNB Working Papers in 2018

No. 583 **Dorinth van Dijk, David Geltner and Alex van de Minne**, Revisiting supply and demand indexes in real estate
No. 584 **Jasper de Jong**, The effect of fiscal announcements on interest spreads: Evidence from the Netherlands
No. 585 **Nicole Jonker**, What drives bitcoin adoption by retailers?
No. 586 **Martijn Boermans and Robert Vermeulen,** Quantitative easing and preferred habitat investors in the euro area bond market
No. 587 **Dennis Bonam, Jakob de Haan and Duncan van Limbergen,** Time-varying wage Phillips curves in the euro area with a new measure for labor market slack
No. 588 **Sebastiaan Pool**, Mortgage debt and shadow banks
No. 589 **David-Jan Jansen**, The international spillovers of the 2010 U.S. flash crash
No. 590 **Martijn Boermans and Viacheslav Keshkov**, The impact of the ECB asset purchases on the European bond market structure: Granular evidence on ownership concentration
No. 591 **Katalin Bodnár, Ludmila Fadejeva, Marco Hoeberichts, Mario Izquierdo Peinado, Christophe Jadeau and Eliana Viviano**, Credit shocks and the European labour market
No. 592 **Anouk Levels, René de Sousa van Stralen, Sînziana Kroon Petrescu and Iman van Lelyveld**, CDS market structure and risk flows: the Dutch case
No. 593 **Laurence Deborgies Sanches and Marno Verbeek**, Basel methodological heterogeneity and banking system stability: The case of the Netherlands
No. 594 **Andrea Colciago, Anna Samarina and Jakob de Haan**, Central bank policies and income and wealth inequality: A survey
No. 595 **Ilja Boelaars and Roel Mehlkopf**, Optimal risk-sharing in pension funds when stock and labor markets are co-integrated
No. 596 **Julia Körding and Beatrice Scheubel**, Liquidity regulation, the central bank and the money market
No. 597 **Guido Ascari, Paolo Bonomolo and Hedibert Lopes**, Walk on the wild side: Multiplicative sunspots and temporarily unstable paths
No. 598 **Jon Frost and René van Stralen,** Macroprudential policy and income inequality
No. 599 **Sinziana Kroon and Iman van Lelyveld**, Counterparty credit risk and the effectiveness of banking regulation
No. 600 **Leo de Haan and Jan Kakes**, European banks after the global financial crisis: Peak accumulated losses, twin crises and business models
No. 601 **Bahar Öztürk, Dorinth van Dijk, Frank van Hoenselaar and Sander Burgers**, The relation between supply constraints and house price dynamics in the Netherlands
No. 602 **Ian Koetsier and Jacob Bikker**, Herding behavior of Dutch pension funds in asset class investments
No. 603 **Dirk Broeders and Leo de Haan,** Benchmark selection and performance
No. 604 **Melanie de Waal, Floor Rink, Janka Stoker and Dennis Veltrop**, How internal and external supervision impact the dynamics between boards and Top Management Teams and TMT reflexivity
No. 605 **Clemens Bonner, Eward Brouwer and Iman van Lelyveld,** Drivers of market liquidity - Regulation, monetary policy or new players?
No. 606 **Tanja Artiga Gonzalez, Iman van Lelyveld and Katarina Lucivjanska,** Pension fund equity performance: Patience, activity or both?
No. 607 **Jasper de Jong en Niels Gilbert,** Fiscal discipline in EMU? Testing the effectiveness of the Excessive Deficit Procedure
No. 608 **Carin van der Cruijsen, Maurice Doll and Frank van Hoenselaar,** Trust in other people and the usage of peer platform markets
No. 609 **Jon Frost, Patty Duijm, Clemens Bonner, Leo de Haan and Jakob de Haan,** Spillovers of monetary policy across borders: International lending of Dutch banks, insurers and pension funds
No. 610 **Randall Hanegraaf, Nicole Jonker, Steven Mandley and Jelle Miedema**, Life cycle assessment of cash payments

No. 611 **Carin van der Cruijsen and Joris Knober,** Ctrl+C Ctrl+pay: Do people mirror payment behaviour of their peers?
No. 612 **Rob Bauer, Matteo Bonetti and Dirk Broeders,** Pension funds interconnections and herd behavior
No. 613 **Kai Schindelhauer and Chen Zhou** Value-at-Risk prediction using option-implied risk measures

DeNederlandscheBank

EUROSYSTEEM

De Nederlandsche Bank N.V.
Postbus 98, 1000 AB Amsterdam
020 524 91 11
dnb.nl