Endogenous Belief Switching
Revisiting the Forward Guidance Puzzle*

Mátyás Farkas[†]

August 25, 2022

*I thank Klaus Adam, James Costain, Luca Dedola, Michael Ehrmann, Péter Karádi, Michele Lenza, Alberto Martin, Bartosz Maćkowiak, Roberto Motto, Fulvio Pegoraro and Oreste Tristani for the very helpful comments and discussions.

[†]ECB, email: Matyas.Farkas@ecb.europa.eu

Abstract

Forward guidance has become an important tool for central banks as short-term rates have approached the zero lower bound. As economic theory explored the effectiveness of unconventional monetary policy, the fundamental role of expectations in macroeconomic models received renewed attention. At the same time, there is growing evidence that expectations become adaptive at the zero lower bound, yet there is little research on forward guidance in an adaptive learning environment. The present paper fills this gap, by revisiting the forward guidance puzzle under adaptive learning and highlighting the role of monetary policy for expectation formation. I propose the framework of endogenous belief switching to explain the role of learning and dynamic expectation formation, when thinking about the impact of unconventional monetary policy. Endogenous belief switching is a fundamental alternative to both rational and adaptive learning. In it expectations are determined by central bank action, and so the effectiveness of forward guidance is endogenous. Learning monetary policy implementation agents learn to form expectations about future macroeconomic variables: either by responding to pre-announced future policy rate changes, following forward looking beliefs, or neglecting them and focusing only on current conditions, forming backward looking beliefs. I endogenize belief switching using a mean squared learning transition between the two beliefs. Estimating a switching Kálmán filter (Murphy, 1998) every period agents update their beliefs about the probability that either of the regimes is the best descriptor of the economy, and thus dynamically decide to become forward or backward looking. Simulation results indicate that the effectiveness of forward guidance is non-linear. The forward guidance puzzle is nonexistent if agents are adaptive and backward looking. The puzzle emerges, if expectations are adaptive and forward looking. The framework predicts that forward guidance is highly effective in low uncertainty environments, where the model fits the data well and there is small observation error, while it can become ineffective in high uncertainty economies. In other words, agents choose to become backward looking, if there is too much noise compared to the signal of the forward guidance, and can learn to trust the central bank, if it sends a strong enough signal about its commitment.

Keywords: Unconventional monetary policies, Adaptive expectations, DSGE models
JEL-Codes: C53, C54, E43, E52, E58

Non-technical summary

Interest rate forward guidance has become an important tool for central banks in order to enhance the effectiveness of monetary policy at the zero lower bound. At the same time there is growing evidence that expectations are adaptive at the zero lower bound. Yet there is little research on the effectiveness of forward guidance in an adaptive learning environment. My paper fills this gap, by revisiting the forward guidance puzzle under adaptive learning and highlighting the role of central bank action for expectation formation.

The forward guidance puzzle states that rational expectations based dynamic stochastic general equilibrium (DSGE) models overestimate the impact of forward guidance. With increasing length of the forward guidance DSGE models not only display explosive responses, but counter-intuitive reversals as well. Economic research about the forward guidance is steadily growing, and there is a consensus that the culprit for the puzzle is the way expectations are modelled. Thus, in order to solve the puzzle, I deviate from rational expectations, by assuming adaptive learning. In my quest for a framework where announced policy actions can still have an impact, I develop endogenous belief switching, where the central bank action determines expectations. The paper's main contributions are the following:

1. I show that constant gain adaptive learning can overcome the forward guidance puzzle and that forward guidance's efficacy is dependent on the initial beliefs agents have. I demonstrate that adaptive learning beliefs can behave both as backward and forward looking. I call adaptive beliefs backward looking if agents form expectations only based on current and past conditions; and forward looking when expectations respond to announcements of future policy action. The forward guidance puzzle is overcome by backward looking adaptive expectations. I also document that the forward guidance puzzle still emerges, if expectations respond to credible announcements.

2. I propose endogenous belief switching as a novel form of expectation formation to bridge the two extremes. I argue that a model with endogenous belief switching can both solve and nest the forward guidance puzzle. It provides a fundamental alternative to both rational and adaptive expectations, delivering the insight that self-fulfilling beliefs can be result of central bank action.

3. Using this framework I study the dynamic evolution of central bank credibility, and highlight the conditions when announcing a forward guidance can gain central bank credibility, and when it can destroy it.

My first contribution is to show how adaptive learning can solve the forward guidance puzzle. Constant gain adaptive learning is well understood in the economic literature. The most important difference compared to rational expectations is that agents form expectations, i.e. beliefs, as least-squares econometricians. Thus, adaptive expectations become only a function of the information set agents consider in their regression model. Throughout the paper I distinguish between backward and forward looking adaptive beliefs. Backward looking adaptive beliefs do not respond to forward guidance, while forward looking adaptive beliefs form expectations conditional the anticipated future path of the policy rate. I argue that if no past periods of credible forward guidance have been observed, then agents do not know how to respond to forward guidance, and it will be perceived as a sequence of unanticipated monetary shock, and the anticipated peg that grants effectiveness to forward guidance cannot be learnt. The longer the forward guidance is implemented, the more the delivered path enters the information set of agents and increases the persistence of beliefs. Thus, the marginal shock needed to implement the path is diminishing in the horizon, while the forward looking anticipatory impact is not learnable as it is requires a model of beliefs that is of a different regime: where expectations are conditioned on pre-announced paths of the policy rate. I show, should adaptive learning beliefs feature the concept of anticipated news shocks, the forward guidance puzzle re-emerges.

The second contribution is the development of a novel form of expectation formations: endogenous belief switching. It is a combination of constant gain adaptive learning and regime switching Kálmán filtering. Endogenous belief switching features self-fulfilling beliefs, and multiple equilibria, as the central bank determines how expectations are formed. In this framework agents learn about the probability that either adaptive backward looking or adaptive beliefs, that also respond to anticipated shocks, are the best descriptor of the observed evolution of the economy. In short, agents behave as econometricians who interpret the world through two distinct regimes representing extreme cases, forward and backward looking adaptive beliefs. Doing so, they choose beliefs that best fit the observed economy. I build intuition for endogenous belief switching in a model with with one, and then with two dimensional state space: showing how forward guidance can make expectations forward looking.

The third contribution is the study of the dynamic evolution of central bank credibility, by modelling it as the probability attached to beliefs that respond to announced forward guidance. Using endogenous belief switching I highlight the conditions when through forward guidance the central bank can build credibility, and when it fails to do so. The model shows, that for a given size of the path announced, the longer forward guidance horizon, the longer delivery it requires to be credible. At the same time, for a given horizon a deeper forward guidance gains credibility faster. Subsequently, I discuss the signal extraction problem inherent to econometric analysis, and thus to adaptive learning. Endogenous belief switching agents try to estimate the true model, doing so they face not only the ambiguity of decomposing shocks and model dynamics, but the problem of observation uncertainty versus uncertainty about the policy shocks. By relating the observation error to unexplained and the policy shock to explained variation of the economy, I highlight the role of the central bank's model fit to the reality. If the central bank gives a weak signal, agents confuse it with noise and will not respond to it. However should the signal be large compared to noise, it can make beliefs become forward looking. From a policy perspective this implies that forward guidance can be highly effective in low uncertainty environments, while it can become ineffective under high uncertainty.

Contents

1 Introduction **1**

1.1 Forward Guidance Puzzle 3

1.2 Departures from Rational Expectations 4

1.3 Adaptive Expectations 5

2 Constant Gain Adaptive Learning **9**

2.1 The Smets Wouters Model with Constant Gain Adaptive Learning 13

3 Backward Looking Beliefs in the Smets Wouters Model under Adaptive Learning **15**

3.1 Extended Path Simulation 16

3.2 Simulation of Forward Guidance with Backward Looking Beliefs 18

4 Forward Looking Beliefs under Adaptive Learning **20**

4.1 Smets Wouters Model with Forward Looking Beliefs 22

5 Endogenous Beliefs Switching **24**

5.1 Timing Assumptions and Adaptive Learning 29

5.2 Endogenous Belief Switching as Accelerated Learning 30

5.3 Belief Switching Illustration in One Dimension 31

5.4 The Three Equation New Keynesian Model 34

5.5 Endogenous Belief Switching in Higher Dimensions 37

5.6 Central Bank and Delivering Forward Guidance: Gaining Credibility 42

5.7 Noise and Signal 45

6 Endogenous Belief Switching in the Smets Wouters Model with Noise **46**

7 Conclusion **51**

List of Figures

1 Expanded Path Simulation of Forward Guidance under AL 17
2 Forward Guidance in the SW07 Model with Backward Looking Beliefs 19
3 SW07 Model with Forward Looking Beliefs 24
4 Endogenous Belief Switching 28
5 Belief Switching Illustration - Initial Beliefs 32
6 Belief Switching Illustration - Second Period Beliefs 33
7 Forward Guidance Shock in 3EQ Model with Adaptive Learning 36
8 3EQ Model - Endogenous Belief Switching 38
9 Belief Switching Illustration - Backward to Forward Beliefs in the 1^{st} Period 40
10 Belief Switching Illustration - Forward to Backward Beliefs in the 2^{nd} Period . . . 42
11 Central Bank Credibility - Period of Belief Switching 44
12 Belief Switching Illustration - Bird's eye view - 1^{st} Period 60
13 Belief Switching Illustration - Bird's eye view - 2^{nd} Period 61

List of Tables

1 Parameter calibrations in the 3 Equation model 34
2 Regime Switching Beliefs in SW07 with Adaptive Learning 49
3 Smets Wouters Model with Adaptive Learning: Endogenous variable definitions . 64
4 Smets Wouters Model with Adaptive Learning: Exogenous variable definitions . 65
5 Smets Wouters Model with Adaptive Learning: Parameter Definitions 65
6 Smets Wouters Model with Adaptive Learning: Parameter Values 67

1 Introduction

Prior to the great recession, the literature on monetary policy had converged as setting the short term inter-temporal return of assets, the policy rate. A large literature has documented the transmission mechanism in an economy with financial, nominal, and real frictions of surprise changes in short-term interest rates onto the economy, using either VARs or DSGE models (e.g., Sims (1980), Christiano et al. (1999), Clarida et al. (1999),Christiano et al. (2005)).

While the effects of short-term interest rates in normal times is well understood, the policy rate has been constrained by the zero lower bound (ZLB) during the financial crisis in most developed economies. To overcome the disastrous theoretical consequences a constrained monetary policy meant for models, alternative tools were developed and proposed, and the term unconventional monetary policy has been coined. Major central banks have employed these new tools, such as announcements about the future path of the policy rate, forward guidance, or quantitative easing[1].

As our theoretical understanding evolved so did the interpretations of the primary channels of the policy instruments. Forward guidance has been early identified as the promising tool to overcome the ZLB. Eggertsson and Woodford (2003) argued that the central bank committing to an interest rate path, that is lower than one that would be implemented under normal circumstances, can have additional expansionary impact. The primary effect is through the credible pegging of expectations of persistent variables. Due to persistence of rational expectations this peg translates into a sizeable impact on announcement. As unconventional measures were further explored and implemented, so evolved our interpretation of unconventional monetary policy. For instance, recent understanding of quantitative easing attributes its effectiveness to the signalling channel, and the implicit forward guidance it conveys (Bhattarai et al., 2015). However the lower for longer policy is not time consistent, and thus credibly committing to it is of theoretical and practical concern (Woodford, 2012).

There is growing evidence that expectations become de-anchored from their inflation target and become adaptive at the ZLB [2]. At the same time, there is limited understanding on how forward guidance should work in an adaptive learning environment[3]. Therefore understanding

[1] Measures involving a change in the size and the composition of the central bank balance sheet.

[2] Mario Draghi highlighted that the "risk that too prolonged period of low inflation becomes embedded in inflation expectations" Draghi (2014). Inflation expectations are (perfectly) anchored if long-term inflation expectations do not respond to surprises and news. Should there be evidence for co-movement between expectations and macroeconomy, then expectations are better described as adaptive. Beechey et al. (2011) documented that long-term inflation expectations indeed respond to inflation news in the US, but found no significant effect in the euro area, while more recently Van der Cruijsen and Demertzis (2011) showed that inflation expectations correlate with past inflation for euro area countries. Dovern and Kenny (2017); Łyziak and Paloviita (2017) exploited micro data on surveys and found that past inflation correlates with euro area inflation expectations and the central bank's past performance, thus inflation expectations are endogenous to monetary policy.

[3] Cole (2015) made the first exploratory study to compare impact of forward guidance under adaptive learning

and explaining how forward guidance works in an adaptive learning environment is of high relevance.

Standard medium-scale DSGE models tend to overestimate the impact of forward guidance on the macroeconomy, a phenomenon called the “forward guidance puzzle.” The failure of macroeconomists to explain the missing effectiveness of forward guidance evoked a wave of criticism regarding the way expectation formations are modelled. It is undeniable that forward guidance has impact in reality, although the gap between its actual and theory predicted effect is sizeable enough to be termed a puzzle.

Deviations from standard rational expectations (henceforth RE) benchmark has been shown to mitigate, and in some cases overcome the forward guidance puzzle. The present paper also abandons RE and studies adaptive expectations. Furthermore, the role of dynamic belief formation is still unexplored. My work fills this gap by proposing a novel adaptive learning framework, where expectation formation is endogenous and dynamic. The framework is designed to study central bank credibility evolution in light of forward guidance policies.

The paper's main contributions are the following: First, it shows that constant gain adaptive learning can overcome the forward guidance puzzle. I define two types of constant gain adaptive expectations, backward and forward looking beliefs[4]. I present how to model forward guidance under both backward and forward looking adaptive expectations. I prove that the forward guidance puzzle does not emerge if expectations are backward looking. At the same time I show that forward looking constant gain adaptive expectations still feature the forward guidance puzzle.

Second, building on these insights, I propose endogenous belief switching, an adaptive learning framework with regime switching to bridge the two extremes. I argue that a model with endogenous belief switching expectations can both solve and nest the forward guidance puzzle. It provides a simple generalization to heterogeneous beliefs, by making the share of attentive agents endogenous. Endogenous belief switching gives rise to changing efficacy of forward guidance.

Third, using this framework I study the dynamic evolution of central bank credibility, and highlight the conditions when announcing a forward guidance can gain central bank credibility, and when it can destroy it. The endogenous belief switching proposed in the paper provides policy relevant interpretation of the prolonged zero lower bound periods observed in multiple countries. It argues that central banks can be made partially responsible for the development and severity of the liquidity trap, as they failed to provide strong enough signals to overcome the heightened uncertainty. This endogenously turn expectations to become backward looking. Furthermore the framework also implies that close to the zero lower bound, there is little room

versus rational expectations, however his approach suffers from shortcomings that will be discussed in detail.

[4]It is important to realize that adaptive expectations are not equivalent to backward looking beliefs: adaptive expectations can also be formed based on current information about the future. Thus adaptive expectations can also be forward looking. Furthermore forward looking adaptive beliefs are not rational: in response to shocks they still make forecast errors and learn. For a detailed discussion of these concepts please see Sections 1.3, 3 and 4.

left for monetary policy to give a strong contemporaneous accommodation that is necessary for its credibility, and thus forward guidance can become ineffective. It calls for rethinking of central bank communication policy as well. The model predicts that the fit of the central bank's communicated model to the data, and thus the understanding of the central bank's policy function, is crucial for central bank credibility. Without an accurate model forward guidance might have diminishing impact, or worse turn agents endogenously to be backward looking.

In summary, this paper develops a novel form of expectations, endogenous belief switching. It highlights the role of learning and dynamic expectation formation when thinking about the impact of unconventional monetary policy.

1.1 Forward Guidance Puzzle

Forward guidance works through the fact that agents understand credible announcements about future policy shocks and respond to them. In my paper I consider time dependent forward guidance, where forward guidance is an announcement and delivery of setting the short-term interest rate to a pre-announced low level, possibly to its ZLB, for a predetermined horizon[5]. At the same time it is a concept that is not without caveats, its time inconsistency and policy implications have sparked an intense debate about its efficacy. The problem of the forward guidance puzzle received large attention after the seminal paper of Del Negro et al. (2012). It states that standard DSGE models generate an excessive response to announcements of credibly committed future changes of interest rates. Carlstrom et al. (2015) show that a fully anticipated and unconditional expansionary monetary policy over a finite number of periods deliver unreasonably large responses of inflation and output. They extend the forward guidance puzzle by documenting that inflation indexation in a DSGE model produces counterintuitive reversals[6]. Therefore the forward guidance puzzle incorporates not only the increasing marginal impact of future announcements, but the inherent instability it generates as well. In what follows, I focus mostly on the first aspect of the puzzle, as the second is conditional on the first being displayed.

There are many proposals to overcome the forward guidance puzzle. McKay et al. (2016) show when agents face uninsurable income risk and borrowing constraints, the precautionary savings effect mitigates the forward guidance puzzle. Their work highlights that the power of forward guidance is highly sensitive to the assumption of complete markets. There are two reasons why uninsurable risk will mitigate the impact of future interest rate adjustments: First, households anticipate that with some probability they will face a sequence of idiosyncratic shocks that lead them to hit the borrowing constraint. Second, households face the trade-off of running down their precautionary asset holdings: they evaluate the costs of having lower buffer

[5] I do not consider state contingent forward guidance. For recent discussion of the different effectiveness of types of forward guidance impact see Ehrmann et al. (2019).

[6] Policy reversals after a forward guidance mean that the model produces a result that the impact of a policy peg can switch from highly expansionary to highly contractionary for modest changes in the length of the interest rate peg. For example, an interest rate peg of 7 periods may be strongly expansionary, but a peg of 8 periods may be sharply contractionary.

and being more exposed to idiosyncratic shocks, against taking advantage of the implicit higher inter-temporal substitution.

Caballero and Farhi (2017) argue that forward guidance at the zero lower bound has no effect at all on the real economy but has sizable effects on risk premia. In their view forward guidance works as an insurance against the liquidity trap rather than a policy instrument of credible commitment. It “promises stimulus when the economy has already recovered and the marginal utility of consumption is low”. My approach focuses on expectation determination and the implicit the credibility of forward guidance, by departing from the RE paradigm.

1.2 Departures from Rational Expectations

Assumptions about how expectations are formed are central to the efficacy of forward guidance, and their role has been only recently appreciated in the literature. Carlstrom et al. (2015) point out that sticky information models can eliminate forward guidance reversals and mitigate the forward guidance puzzle. In sticky information models agents update their beliefs in a staggered manner, given an exogenous information revelation process(Mankiw and Reis, 2002). My framework proposes to overcome this limitation by making belief formation endogenous.

Chung et al. (2015) study how information stickiness in DSGE models compares to price stickiness and what it implies for the forward guidance puzzle. They document that forward guidance creates the puzzle due to the forward-looking nature of inflation under sticky prices. They show that ”a credible promise to remain highly accommodative can lead to substantial effects on real activity and inflation. Under sticky prices [and partial indexation], these effects can be very large [...] , a promise of prolonged future accommodation raises future inflation, which leads to higher current inflation, which lowers real interest rates and raises output (which then raises inflation further).” (Chung et al., 2015, p. 35) Inflation indexation creates inertia by adding lagged inflation as an endogenous state variable to the Phillips curve. This amplification effect is absent in sticky-information models. Sticky information models are not prone to the forward guidance puzzle as expectations are a function of past expectations of current conditions, rather than of future conditions[7].

The primary caveat with sticky information models is inherent to the information revelation: ever period only an exogenous portion of firms get to update their beliefs. This implies that belief formation is only time, but not state dependent.

In regard to treating expectation formation as an information extraction,i.e. signal processing problem, my model stands closest to the rational inattention literature pioneered by Sims (2006); Mackowiak and Wiederholt (2009); Maćkowiak and Wiederholt (2015). Rational inattention provides micro foundation to the decision problem, and answers how agents optimally

[7] More specifically, in the Mankiw-Reis model inertia in the Phillips curve is due to the slow evolution of information to firms. Sticky information translates expectations to be driven by an exogenous state variable tracking the number of firms that have updated prices since the announcement of the forward guidance. For a recent summary of sticky information models, see Mankiw and Reis (2010).

allocate limited attention between signals. The result of this optimization is simple and intuitive: they best allocate attention to accurate signals. As mean squared adaptive learning agents also feature his property, rational inattention can be used to micro-found adaptive learning and the ensuing endogenous beliefs as shown by Molavi (2019). I rely on this interpretation by linking endogenous belief switching to a signal extraction problem.

Andrade et al. (2019) build on information heterogeneity to make forward guidance less effective. They propose a model with heterogeneous interpretations of forward guidance policy to study its effectiveness. Since the commitment ability of the central banks is not observable, agents can agree to disagree about the actual policy conducted.

Two type of beliefs emerge in response to forward guidance, "Odyssean agents - who believe in the commitment ability of the central bank - see the announcement as including some periods of extra accommodation, contingent to any possible realization of the length of the trap. By contrast, Delphic agents - who do not believe in the commitment ability - consider there will be no period of extra accommodation at the end of the trap." (Andrade et al., 2019, p.3.) They show that forward guidance will have counteracting forces on central bank commitment. In response to an Odyssean announcement Delphic agents become excessively pessimistic[8], believing the zero lower bound trap is longer than actual, that in turn may lead a central bank to find it optimal to abandon commitment and engage in a Delphic forward guidance. Their model studies the central bank's optimal reaction conditional on exogenous shares of heterogeneous beliefs.

Analysing the role of heterogeneous beliefs on forward guidance my work builds on a similar concept as Andrade et al. (2019), distinguishing between forward looking, and backward looking adaptive beliefs. However in contrast to existing literature, where the shares of beliefs were assumed to be exogenous, I endogenize belief formation with regime switching adaptive learning and thus enabling the central bank to affect its own credibility. Furthermore, I will assume that both beliefs feature adaptive learning and thus are not rational expectations. This way opening up a novel avenue to study central bank credibility in light of forward guidance announcements.

1.3 Adaptive Expectations

My paper is also based on the literature of adaptive learning and policy announcements. Throughout the paper I will explore two types of adaptive beliefs: backward and forward looking beliefs. Section 2 reviews constant gain adaptive learning in detail, in this section I focus on my contribution in contrast to the literature. I define backward looking adaptive beliefs, where expectations are formed based on past state variables and current exogenous shocks and do not incorporate information about announced, anticipated future policy action. Translating backward looking beliefs to forward guidance means that agents do not believe the forward

[8]And if a "sufficiently large fraction of agents are Delphic, additional periods of policy accommodation will have increasingly negative effects on current macroeconomic conditions." (Andrade et al., 2019, p.3.)

guidance, from their perspective it is not credible. In other words, agents are Delphic, or inattentive to forward guidance. Backward looking beliefs are econometric forecasts where only the past and present enters the model. Section 3 discusses how to model backward looking adaptive expectations in DSGE models[9]. Contrarily, I define forward looking adaptive beliefs, as adaptive learning expectations that are formed based on past state variables, current and anticipated future exogenous shocks. That is forward looking beliefs anticipate the credible commitment of the forward guidance. Forward looking adaptive beliefs are attentive to the forward guidance, thus form adaptive expectations conditional the central bank's commitment,i.e. are Odyssean beliefs. It is fundamental to realize that forward guidance works because agents believe it, thus it is heavily reliant on self-fulfillment of beliefs. Forward looking adaptive expectations feature the self-fulfilling efficacy of forward guidance: as forward looking adaptive beliefs can be thought of conditional forecasts given the announced forward guidance. Section 4 presents how to model forward looking adaptive expectations[10].

There is growing evidence of expectations becoming adaptive at the ZLB, and then the question emerges, which type of adaptive expectations are followed in reality, backward, forward looking or a mixture of the two?

In this paper I develop a theoretical framework where the mixture of the extremes of adaptive beliefs arises endogenously in response to monetary policy.

My work is also motivated by ample empirical evidence. Ehrmann (2015) provided the first international analysis on the fact that expectations turn adaptive at the zero lower bound by documenting that under persistently low inflation, inflation expectations become more dependent on lagged inflation. Ehrmann (2015) studied Consensus Economics inflation forecasts for ten inflation targeting countries and ran panel regressions on a very rich data set. His paper found that under persistently weak inflation, expectations are not as well anchored as otherwise: "they tend to become more backward looking; disagreement across forecasters increases." (Ehrmann, 2015, p.244.) Furthermore there is micro-evidence that inflation expectations are adaptive, and are influenced by "the ex post historical track record of the central bank" (Dovern and Kenny, 2017, p.4.). Dovern and Kenny (2017) show on the panel of SPF that higher moments, i.e. varaince and skewness of long term inflation expectations have increased following the Great Recession, as the short term rate hit the ZLB. Their results also document how SPF forecasters "update their assessment of long-run inflation uncertainty in response to macroeconomic developments. Factors which influence this assessment include the volatility in recent inflation rates and perceptions of increased inflation uncertainty at shorter horizons." (Dovern and Kenny, 2017, p.4.) Both results support the argument that expectations are adaptive, where past inflation forecast errors are weighted with their perceived uncertainty to update beliefs.

[9] Throughout the paper I will use the term backward looking adaptive expectations interchangeably to backward looking beliefs.

[10] In what follows, I will use forward looking beliefs to denote forward looking adaptive expectations. Please keep in mind that forward looking beliefs are still a specification of adaptive learning expectations and are not RE!

Carvalho et al. (2019) build on this insight and propose a DSGE with regime switching to provide estimates to the degree of de-anchoring, in form of adaptiveness of expectations. They document that expectations indeed are better described with a constant gained adaptive learning structure than RE at the ZLB. Their approach focuses on estimating the constant gain learning parameter. For a set of euro area countries they document a partial de-anchoring of expectations at the ZLB, first after the great recession and more importantly recently.

Mitra et al. (2012) examined the effects of the fiscal authority giving guidance on the future course of government purchases and taxes in an adaptive learning environment. Their results show that a change in announced fiscal policy leads to different effects depending on the assumed expectation formation. Adaptive learning based expectations results in output multipliers that match empirical data more than the output multipliers of and RE.

Wieland (2008) introduced adaptive learning and endogenous indexation in the New-Keynesian Phillips curve and studied disinflation under inflation targeting policies. He shows that if adaptive learning price-setting firms revise their degree of persistence, it also lowers the cost of disinflation.

Eusepi and Preston (2010) study the link between adaptive learning and central bank communication strategies. Improved central bank communication, such as communicating the monetary policy rule and the variables within the rule, can lead to increased macroeconomic stability, highlighting the learn-ability of a monetary policy rule.

Cole (2015) studies the impact of adaptive learning on forward guidance, when the latter is modelled as anticipated news shock. He shows that adaptive learning aggravates the forward guidance puzzle. Furthermore he illustrates how adaptive learning agents fail to understand the nature of forward guidance and will perceive a contemporaneous shock every period that implements the path.

Cole (2015) correctly identifies that forward guidance under adaptive learning, given inaccurate initial beliefs, will lead to deviations from the actual path, and thus to learning. However his work suffers from multiple shortcomings. First, he assumes all agents understand the forward guidance and respond to it. Second, even though agents are forward looking they act as backward looking agents. Cole models beliefs being forward looking, yet agents only perceive the current period implementation of the path, when learning.

I argue that this is inconsistent. If beliefs are indeed forward looking they should be re-anchored to the full path every period, if they deviate from it. Contemporaneous monetary policy shock is only one part of the forward guidance, the bulk of the policy is in the future. In other words, the assumption that only current monetary policy shock is perceived abstracts from the impact of expectations that the promised path will be implemented[11]. Thus in order to maintain forward looking expectations the perceived monetary policy shock has to be not only on contemporaneous response, but it has to account for the anticipated future shocks as well,

[11] A current period monetary policy shock will make expectations deviate from the pre-announced forward guidance path, as it creates anticipated response.

i.e. every period a new set of perceived anticipated shocks have to be observed. This way expectations can be re-anchored to the path.

I argue that forward looking beliefs will require to solve the model forward as well. This way forward looking agents will be able to find expectations for the full set of monetary and forward guidance shocks that correctly re-anchor expectations to the (remaining) path of the forward guidance. This is a very important observation[12]. In fact my distinction between backward and forward looking beliefs comes from the way each perceives the forward guidance: backward looking beliefs only see the implementation of the forward guidance path as a sequence of unanticipated monetary policy shocks. While forward looking expectations understand the concept of anticipated news shocks, and should beliefs need updating, re-anchor their expectations every period to the whole path.

By modelling beliefs with endogenous belief switching I not only show how central bank credibility can dynamically respond to central bank actions, but provide a tractable and simple alternative to heterogeneous beliefs. Exploiting the mean squared econometric interpretation of adaptive beliefs, I motivate the use of the switching Kálmán filter for updating a priori beliefs about the share of beliefs, and endogenize the belief formation. My work therefore provides a novel extension to adaptive learning towards heterogeneous beliefs, incorporating them in a unified framework[13].

The remainder of the paper is organized as follows. Section 2 reviews constant gain adaptive learning. In Section 3, I present the Smets Wouters model (henceforth SW07 model) with constant gain adaptive learning based on Slobodyan and Wouters (2012) and explain how the backward looking beliefs can overcome the forward guidance puzzle. In Section 4, I introduce how anticipated news shocks can be implemented in the adaptive learning framework and show how the forward guidance puzzle emerges.

In Section 5, I develop the concept of endogenous belief switching based on the switching Kálmán Filter to combine the two extremes of expectations and introduce dynamic belief switching. I illustrate endogenous belief switching in models with one and two dimensional state space. Doing so I present a small scale three equation New Keynesian model that combines forward, and backward looking beliefs with endogenous belief switching. This toy model is then used to study how the length and the size of forward guidance influences central bank credibility. Subsequently I return to the SW07 model and highlight the role of signal to noise for endogenous expectation determination. Finally, Section 7 provides concluding remarks.

[12] Cole (2015) implicitly causes beliefs to act as backward looking although upon announcement they are assumed to respond to news shocks, and thus feature forward looking characteristics.

[13] For references on endogenous regime switching DSGE models see Section 5

2 Constant Gain Adaptive Learning

Deviations from RE via learning in the real business cycles literature dates back to Kydland and Prescott (1982). Learning in macroeconomics is usually modelled as a signal extraction problem, where agents need to gather information from a noisy signal and form beliefs about the dynamic nature of the economy. Marcet and Sargent (1989) describe how agents' perceived law of motion (PLM) affects the actual law of motion (ALM) of the economy. They introduce least squares learning and derive conditions for convergence of the expectations, and show how they relate to the RE equilibrium.

Adaptive learning has gained popularity ever since, receiving its first textbook treatment by Evans and Honkapohja (2012). Bullard and Mitra (2002) evaluate how recursive learning influences monetary policy rules, while Preston (2005) studies the impact of long-term expectation determination on monetary policy design. Both approaches discuss the expectation stability (E-stability) of the model given monetary policy rules. From the policy perspective the shared conclusion is to have a policy design which delivers learnable RE equilibria in the limit. Goy et al. (2018) studies adaptive beliefs and endogenous heterogeneous expectations' response to forward guidance under regime switching using a heuristic switching model (Brock and Hommes, 1997; Hommes and Lustenhouwer, 2019). The two regimes assumed are either adaptive learning expectations or credibility believers. Switching is driven by sum of squared forecast errors, where the central bank itself has a separate forecasting model to publish forecasts where the inflation target and RE equilibrium is satisfied. In contrast, I assume no information friction beyond adaptive expectations. In endogenous belief switching the regimes will be determined by the likelihood of the belief, given the past and present. In other words the DSGE estimation using a Kálmán filter delivers that switching occurs not due to forecast errors, but due to the fit of the (updated) model to the past and present. Generally a fit of a DSGE is indeed driven by (one step ahead forecast, prediction) filtering errors, but these errors are evaluated, weighted by the state covariance matrix. In my model, unlike Goy et al. (2018), I do not have N-period limits to expectation formation, and both the central bank and the private agents share the same forecasting rule. The central bank controls the size of the monetary and forward guidance shocks. Agents can learn the effectiveness of forward guidance as the economic response to the central bank action makes forward looking beliefs more likely.

In what follows I will employ the notation used in Dynare (Juillard, 2001) to introduce the structural, reduced and state space representation of the dynamic stochastic general equilibrium (henceforth DSGE) model.

A DSGE's solution after after a linearization around the steady state can be written in the structural form as follows:

$$A_0 \begin{bmatrix} y_{t-1} \\ w_{t-1} \end{bmatrix} + A_1 \begin{bmatrix} y_t \\ w_t \end{bmatrix} + A_2 E_t \left[y_{t+1} \right] + B \varepsilon_t = const, \tag{1}$$

where y_t stands for the endogenous state variable,e.g. output or inflation, w_t denotes the exogenous state variables, e.g. technology process, and ε_t is the realization of the exogenous shock, while A_0 is the backward solution of the DSGE, A_1 the contemporaneous response,A_2 is the forward solution given rational expectations. B captures the impact of the contemporaneous exogenous shocks on state variables, while the right hand side $const$, constant accounts for endogenous drifts. A DSGE's RE solution can be then obtained in many ways, either using Sims' method, or Binder and Pesaran's algorithm (Binder and Pesaran, 1995). I initialize beliefs at their RE counterpart, and to solve for the initialization,i.e. RE equilibrium, I use Dynare (Juillard, 2001), which is broadly based on gensys published in Sims (2002).

The RE solutions, in order to ensure the uniqueness, imposes the central idea from optimal control: if there exist explosive expectations dynamics, then in order to have a unique stable solution, exogenous shocks have to span all explosive dimensions. In mathematical terms it means that the explosive roots should be in the space spanned by the structural shocks(Sims, 2002). This assumption is highly restrictive[14]. Nevertheless it is the asymptotic limit to which adaptive learning beliefs converge.

Given uniqueness of the RE solution, the DSGE model can be written in the reduced form:

$$\begin{bmatrix} y_t \\ w_t \end{bmatrix} = \mu + T \begin{bmatrix} y_{t-1} \\ w_{t-1} \end{bmatrix} + R\varepsilon_t, \tag{2}$$

where the reduced form matrices T and R are non-linear functions of the structural parameters. The intercept, μ, can be non-zero for observables that are not demeaned[15].

Finally, based on the reduced from, one can write the state space model representation of the DSGE, that is the reduced form, augmented with the observation equation:

$$y_t^{state} = \mathbf{F} y_{t-1}^{state} + \mathbf{w}_t, \tag{3}$$

$$X_t^{obs} = \mathbf{H} y_t^{state} + \mathbf{u}_t \tag{4}$$

Where y_t^{state} is the state vector with dynamic evolution $\mathbf{F}$, $\mathbf{w}_t$ is the exogenous state disturbance, with a covariance matrix Q_t. The state covariance matrix is sometimes also referred to as long-run variance of the states or mean squared error matrix of the states and it is equal to the square of R,i.e. $R*R'$, from the reduced form representation.

$\mathbf{H}$ is the emission matrix that selects the observable states of the model[16]. Finally vector $\mathbf{u}_t$ is the noise of the observation equation, with covariance matrix U_t. Its relative size to the state disturbance determines the signal to noise ratio of the DSGE.

Under adaptive learning, as in Slobodyan and Wouters (2012), Marcet and Sargent (1989) and Evans and Honkapohja (2012), agents forecast the values of the forward variables as

[14]Deviating from it introduces multiple equilibria, where the hyperspace in which the degree of freedom exists is spanned the the sunspot shock(s).

[15]Note that the RE solution is time invariant.

[16]This matrix will compress the state space of the DSGE into an subspace where observables are measured.

a econometricians: replacing expectations with forecasts that are linear function of the state variables. The adaptive learning literature calls the state space that enables agents to form expectations asymptotically equivalent to RE the minimum state variable space(MSV), that is a linear vector space spanned by the endogenous state variables forming the RE solution. In this paper I abandon RE assumption and assume that agents form expectations about the forward looking variables with the help of a linear function of state variables. Furthermore, I assume that agents can observe the MSV state thus can learn the RE equilibrium equivalent beliefs over time. This observation is assumed for the time being to be exact, and will be relaxed in chapter 6, when discussing the relationship between signal and noise.

The equation that describes the agents' expectations is the PLM:

$$y_t^f = \alpha_{t-1} + \beta_{t-1}' \begin{bmatrix} y_{t-1} \\ w_t \end{bmatrix}. \tag{5}$$

Where $\alpha_{t-1}, \beta_{t-1}$ are representing beliefs of the means squared econometrician. The belief parameters are dated time $t-1$ to represent that expectations are formed based on predetermined states y_{t-1} and current shocks w_t and are usually collected into the belief matrix $\Phi_{t|t-1}$.

$$y_t^f = \alpha_{t-1} + \beta_{t-1}' \begin{bmatrix} y_{t-1} \\ w_t \end{bmatrix} = \Phi_{t-1} \cdot Z_t \tag{6}$$

The variables based on which expectations are formed, i.e. the regressors in the econometricians' model are usually denoted as $Z_t = [1, y_{t-1}, w_t]'$.

Subjective beliefs are thus time dependent and are completely described by the information set available to the agents. Z_t is the state variables observable in period t, based on which beliefs are formed, these are previous states y_{t-1}, and current shocks w_t[17]; $\Phi_{t|t-1}$ is the belief matrix ; and $R_{t|t-1}$ [18] is the beliefs about the accuracy of the states, the mean squared forecast errors before learning, i.e. updating. Once expectations are formed the stochastic shocks realize and the ALM is determined.Given the expectation errors agents update beliefs in a mean squared learning process.

Standard adaptive learning is an expectation formation process, that can be described as a sequence of events.(Evans and Honkapohja, 2012) The timing of the events is the following:

1. At the beginning of period t, the agents inherit the beliefs formed in the previous period: $\Phi_{t|t-1}, R_{t|t-1}$. Agents observe the shocks ε_t and the states of the previous period $[y_{t-1}, w_{t-1}]'$.

2. Agents form expectations based on their previous period beliefs $\Phi_{t|t-1}, R_{t|t-1}$ and the in-

[17]Forward looking agents will have to form expectations about more shocks that are in the MSV as the backward looking agents.

[18]As you will see, forward looking beliefs will have belief matrices that are of larger dimension than backward looking!

formation set Z_t.

3. The current state is determined as the solution to the reduced form of the DSGE. The economy evolves given the reduced form solution implied by the belief matrix $\Phi_{t|t-1}$ and the mean-squared-error matrix of the states $R_{t|t-1}$. In other words, the solution of the model given beliefs constitute the ALM:

$$\begin{bmatrix} y_t \\ w_t \end{bmatrix}^{ALM} = \mu(\Phi_{t|t-1}, R_{t|t-1}) + T(\Phi_{t|t-1}, R_{t|t-1}) \begin{bmatrix} y_{t-1} \\ w_{t-1} \end{bmatrix} + R_{t|t-1}\varepsilon_t. \tag{7}$$

   Note that the both the steady state as well as the policy function is time dependent. Both μ and T are the function of the beliefs $\Phi_{t|t-1}, R_{t-1|t}$. Furthermore information captured in Z_t enters in two way, first in form of the past state variables, y_{t-1}, multiplied by the policy function $T(\Phi_{t|t-1}, R_{t|t-1})$ and second in the form of the exogenous variables w_t. The latter as, first, the impact of past exogenous variables w_{t-1}, through the policy function and, second, as the impact of current shocks through the contemporaneous response in $R_{t|t-1}\varepsilon_t$.

4. Current states $[y_t, w_t]^{ALM'}$ are then revealed. Agents update their beliefs, given the errors they made following the learning of equations:

$$\Phi_{t|t} = \Phi_{t|t-1} + \tau R_{t|t}^{-1} Z_t (y_t^{ALM} - y_t^f) \tag{8}$$

$$R_{t|t} = R_{t|t-1} + \tau (Z_t Z_t' - R_{t|t-1}). \tag{9}$$

I build the endogenous belief switching model around this standard, a well understood learning framework. In what follows next, I present the SW07 model with adaptive learning and discuss what forward and backward looking beliefs imply for its solution.

In a nutshell the Taylor rule under backward looking beliefs will only feature i.i.d. monetary policy shocks, while forward looking beliefs will include policy shocks that follow an integrated moving average process of the order of the forward guidance length.

Beliefs under both regimes will be spanned by the same MSV space, yet will be based on different information sets (Z_t). Forward looking agents will form expectations knowing the anticipated news shocks, that is based on a larger belief matrix Φ_t, considering a larger shock variance space, R_t. In the next section, I will argue that the baseline SW07 did not feature anticipated news shocks in the Taylor rule, and thus is consistent with backward looking beliefs. I will argue that the extended path simulation is the right approach to simulate forward guidance when adaptive beliefs are backward looking. Then I will introduce forward looking beliefs, by incorporating anticipated news shocks.

2.1 The Smets Wouters Model with Constant Gain Adaptive Learning

The Smets and Wouters (2007) model is a medium scale DSGE model estimated for the US economy, following the work of Christiano et al. (2005) it contains both nominal an real frictions. Households maximize expected utility over an infinite horizon, given their habit formation. Labour services are aggregated by a union facing nominal Calvo wage rigidities. Households make consumption and savings decisions given investment adjustments costs. Capital faces capital utilization costs that will affects its use of intensity. Intermediate firms produce differentiated goods using labour and capital as input, and face Calvo type nominal rigidities. The both wage and product pricing is subject to partial indexation to lagged inflation. Together with marginal costs that are dependent on real wages, rental rate of capital and an exogenous technology this results in a Phillips Curve that is both backward and forward looking. Therefore inflation dynamics have both an expectation and a lag term. Monetary policy is described by a Taylor type rule, with interest rate smoothing in the reaction to inflation- and output gap. Importantly for the current application and forward guidance, the baseline model does not feature anticipated news shocks about the future interest rate shocks, i.e. agents do not form beliefs nor have information about future monetary policy shocks. Therefore the baseline model's minimum state variable RE solution will represent the initial beliefs of the backward looking agents. Recall, backward looking beliefs do not consider forward guidance credible, and are inattentive, do not incorporate pre-announced future monetary policy shocks into their expectations, thus they will not respond to signals about the path, neither in size nor in length. The output gap is defined against the flex price equilibrium level of output[19]. The model I use is the original SW model based on the AL toolbox of the Macroeconomic Model Database (MMB), developed by Segei Slobodyan Wieland et al. (2012) adapted for forward guidance exercise.

The model contains 44 endogenous[20] state variables, of which 13 are forward looking[21]. For notational simplicity I denote all the state variables y_t and collect them into a single vector.

Furthermore the model is driven by seven shocks: The neutral and investment-specific technology shock, risk premium shock, exogenous spending shock and monetary policy shock are AR(1) processes, while price and wage mark-up disturbances are ARMA(1,1). The vector w_t collects all seven exogenous variables, as well as the lagged innovations ε_{t-1} for the mark-up shocks.

In contrast to Slobodyan and Wouters (2012), who employ Kálmán filter learning, I assume

[19]This is in contrast to the adaptive learning SW type model presented in Slobodyan and Wouters (2012), where output gap is just a deviation of output from the neutral. Their reason to drop the flex price equilibrium is to considerably reduce the number of forward variables agents have to solve.

[20]The number of endogenous state variables are larger than in the original SW model, as the Macroeconomic Model Database adds fiscal variables, and joint variables to the system, thus the model has the following state variables: a, b, c, cf, eg, epinfma, ewma, fispol, g, inflation, inflationq, inflationql, inflationql2, inflationqls, interest, inve, invef, k, kf, kp, kpf, lab, labf, mc, ms, output, outputgap, pinf, pinf4, pk, pkf, qs, r, rk, rkf, rrf, spinf, sw, w, wf, y, yf, zcap, zcapf.

[21]The forward looking variables are the following: c, cf, inve, invef, lab, labf, pinf, inflationq, pk, pkf, rk, rkf, w.

constant gain adaptive learning[22]. The constant gain adaptive learning can be interpreted as perpetual learning. (Eusepi and Preston, 2011). The gain parameter being constant implies that the rate at which older observations are discounted follows a power function. This also means that expectations forget the past fast and pay more attention to recent forecast errors. The power of decay being fixed is in contrast to the variance weighted approach of optimal control based, e.g. Kálmán filter, learning. Constant gain has advantages in terms of interpretation of the policy function.

Furthermore I assume that information set agents use in forecasting spans the space of the baseline SW model's expectations under RE. I introduce a deviation from RE by including the constant (μ in 2 in the beliefs[23]. Similar to Slobodyan and Wouters (2012), I initiate beliefs at the policy function that is consistent with the MSV RE equilibrium solution. The deviations of the PLM from the ALM will arise due to realization of stochastic shocks. Since expectations are formed using linear quadratic econometric models given observables and feature constant term in them, it is instructive to check if expectations are stable. To check stability, I follow the practice in literature and solve the model given beliefs and simulate the development of the economy, and I discard all PLMs that are explosive or inaccurate. Due to the linear approximation of the solution around the RE steady state the solution is only locally accurate, therefore for both stability and accuracy purposes, I only consider PLMs that remain in a eight standard deviation neighbourhood of the steady state. Should expectations become unstable[24], I replace them with RE equivalent counterparts.

Recall that agents form expectations about the thirteen forward looking variables, seven of which are of the sticky price equilibrium: consumption, investment, hours worked, wages, inflation, real capital prices and real return of capital, and six are capture the flex price counterpart[25]. Furthermore the the MSV representation of the model is based on 17 variables[26].

The Equations 8 and 9 tracks the agents' beliefs over time, if both the observables and the initial beliefs are known. Similar to any state space model, the results are very sensitive to the initialization, Slobodyan and Wouters (2012) explore the role of initialization in detail. As discussed before, I select the initial beliefs for both the policy function, Φ, and the state disturbance covariance matrix R, that are compatible with the RE equilibrium. This choice is motivated by two arguments, first the constant gain adaptive learning will converge to the RE

[22]Kálmán filter learning produces faster adaptation of expectations, than constant gain learning. It has been shown that constant gain learning, represents as special case of least squares learning, that in turn is a special calibrated version of a hidden factor model (Molavi, 2019). Furthermore constant gain algorithm and the Kálmán filter have the same asymptotic behavior in large samples, however their transitional dynamics display differences due to learning.

[23]This constant allows to have trending beliefs, and can be thought of as biased expectations about the steady state of the model. This assumption furthermore relaxes the common deterministic growth rate assumptions for the beliefs.

[24]In the examples considered below, it is never the case.

[25]I have one less state variable for the flex price as one of the prices is normalised to be one.

[26]The MSV I follow is that of the SW07 with adaptive learning in the MMB, it has the following states: mc, cf, inve, invef, kp, kpf, inflationq, w, a, b, epinfma, ewma, g, qs, spinf, sw, interest.

solution, thus in the steady state RE is an accurate description of what mean beliefs would be.[27] Second, having this starting beliefs make agents' forecasting to be uniquely determined, nesting beliefs that would be observed in a RE equilibrium.

This also implies, that an economy with zero learning will maintain the RE equilibrium dynamics, the RE expectations without responding to central bank announcements, i.e. inattentive beliefs. In summary the baseline SW07 model considered here has 44 state variables. That is y is a 44×1 vector. The beliefs are described by $13 \times (1+17)$ by the Φ_t matrix. Where the first dimension is the number of forward looking states 13, and the second dimension of Z_t: the constant plus MSV variables used to form the beliefs. The mean square error matrix R_t is a $(1+17) \times (1+17)$ squared matrix. Thus for the beliefs the following holds:

$$\underset{18\times 13}{\Phi_{t|t}} = \underset{18\times 13}{\Phi_{t|t-1}} + \underset{1\times 1}{\tau}\, \underset{18\times 18}{R_{t|t}^{-1}}\, \underset{18\times 1}{Z_t}\, (\underset{1\times 13}{y_t^{ALM}} - \underset{1\times 13}{y_t^{f}}) \tag{10}$$

$$\underset{18\times 18}{R_{t|t}} = \underset{18\times 18}{R_{t|t-1}} + \underset{1\times 1}{\tau}\, (\underset{18\times 1}{Z_t}\, \underset{1\times 18}{Z_t'} - \underset{18\times 18}{R_{t|t-1}}). \tag{11}$$

3 Backward Looking Beliefs in the Smets Wouters Model under Adaptive Learning

In this section I discuss how forward guidance should be implemented if agents have not experienced forward guidance before and the model they follow is the baseline, original Smets Wouters model. I call expectations of this case backward looking expectations, i.e. agents have inattentive beliefs. In what follows, I argue that the extended path simulation (Fair and Taylor, 1983) is the appropriate method to solve for the dynamics of the forward guidance given the beliefs that the original Smets Wouters model has been written down with, as the implementation of the path will be perceived by the backward looking agents as a sequence of unanticipated monetary policy shock. Therefore the amplification of the responses will not occur as the forward guidance horizon increases and the forward guidance will not emerge. It must be noted that this result hinges upon the assumption about the initial beliefs, and the state variables in the information set of the agents. In what follows I assume, that agents do not form expectations based on anticipated news about future monetary policy, they only form expectations about shocks that were originally considered in the SW model: neutral and investment-specific technology shock, risk premium shock, exogenous spending shock and monetary policy shock. Furthermore agents have information beyond the current period about the one period ahead price and wage markup shocks, as well, so the model is consistent with the baseline SW model. In solving for the impact of the forward guidance shocks there are primary two approaches: First, is the use of deterministic simulations to find the time dependent

[27] Consider the economy being at the stochastic steady state before the forward guidance announcement. Agents use constant gain AL to form expectations, which are centred around the steady state. Therefore the agents learn the RE equilibrium dynamics.

policy function as used by Lindé and Trabandt (2018). Second, by incorporating anticipated news shock, that are known in period 0. I argue that neither is a viable solutions for backward looking beliefs.

Deterministic simulations are a theory consistent way to model forward guidance. They abstract from the effects of future uncertainty by assuming that the agents can form perfect foresight, thus making future shocks perfectly known, anticipated, i.e. deterministic. In practice it boils down to a finite period backward iteration of the model solution conditional on forward guidance path of the state variables. The deterministic simulation starts by assuming the return of the model to steady state in finite time. Then it will solve the system of nonlinear equations backward, iteratively to find the time dependent sate transition dynamics, i.e. time dependent policy function of the model. Perfect foresight under adaptive learning is not a reasonable approach. Deterministic, perfect foresight will translate to no expectations errors, and thus no learning. This implies that the model dynamics will not be updated and will result in simple backward solution from the terminal steady state. The dynamics at the final, steady state will determine the impact of forward guidance. Note however that adaptive learning models under no uncertainty will coincide with the RE solution. Based on this one can easily see how one recovers the forward guidance puzzle.

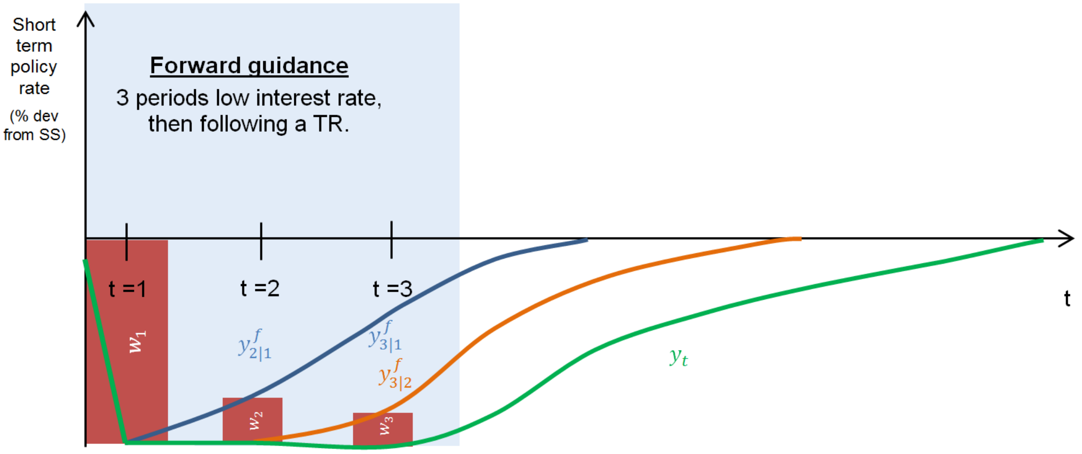
Furthermore deterministic simulations require a mapping of the reduced from DSGE solution into a dynamic factor representation, which is not unique, if one has to allow for expectation errors. This second drawback is only considered in the light of non uniqueness and sunspot equilibria. Giacomini (2013) argues that the mapping of the DSGE back to a VAR representation hinges upon the identification of the state transition dynamics. Franchi and Paruolo (2015) study the existence of the MSV and present necessary and sufficient conditions for the VAR representation and thus the deterministic simulation to exist and be unique. I will discuss how forward guidance under adaptive learning should be simulated using anticipated news shocks in Section 4 in detail. Here I will only stress, that they represent forward looking beliefs. It is natural to start with backward looking beliefs, as the baseline SW model was not designed to feature anticipated news shocks about monetary policy. That is, the SW model specified a belief structure, where anticipated news shocks only entered markup shocks, and even there with a short, one period horizon. To consider anticipated news shock one has to write down a different model, where agents' information set includes news shocks about monetary policy.

3.1 Extended Path Simulation

In what follows I illustrate that constant gain adaptive learning (AL) of the minimum state variable space can overcome the forward guidance puzzle, as amplification of the the impact will not occur as the forward guidance horizon increases. To simulate the forward guidance's impulse response function (IRF) I will employ the extended path simulation for the stochastic model[28]. I

[28]For a quick refresher on extended path simulations see Lawrence Christiano's lecture at Gerzensee in 2014 (Christiano, 2014).

use the definition of the impulse response function, that is the evolution of the economy to a forward guidance shock, a combination of monetary policy and anticipated news shocks, setting all other shocks zero. The extended path simulation requires to know expectations of agents. Given the adaptive, backward looking nature of inattentive beliefs proceed as though they have certainty equivalence, acting as though they believe forecasts are certain to occur. The main ingredient is the information that is relevant for expectation formation. With backward looking beliefs only the current and already realized interest rates will matter for the expectations formations, the announced horizon or path not. Expectations will not respond to forward guidance. The only impact will stem from the central bank's current period action, that will be interpreted as unanticipated monetary policy shock. Implementing the forward guidance will thus turn into a sequence of unanticipated shocks of diminishing magnitudes. Figure 1 illustrates forward guidance given adaptive beliefs and learning of this sort.



Notes: Figure is only illustrative. It shows how forward guidance under backward looking beliefs is perceived. The figure shows the PLM dynamics, i.e. the impulse responses of the short-term interest rate. The forward guidance is a time dependent forward guidance. It constitutes of three periods of low interest rate, and a subsequent implementation of the Taylor rule. (Source: Author's calculations)

Consider that the central bank announces forward guidance in the SW model. As discussed before, beliefs are initialized at the RE equilibrium, the do not feature anticipated news shocks and thus are inattentive to the forward guidance announcement. Agents' beliefs are a linear function of the past state and current exogenous variables. They only learn about the present impact of the forward guidance. They will perceive the first period of the forward guidance as an unanticipated monetary policy shock. This unanticipated monetary policy shock is represented by w_1 on Figure 1.

As the next period arrives, agents will follow their policy function and move the economy along the IRF of the unanticipated monetary policy shock of the previous period, as the IRF represents the expectations of the agents if no further shocks hit the economy. These PLMs are shown on Figure 1 with $y^f_{h|t}$. The ALM is represented by y_t. By local stability of the DSGE's solution under adaptive learning the next period's state will deviate from the preannounced path, and therefore the central bank will have to act. The central bank by enforcing the forward guidance, will surprise the agents by a monetary shock that eliminates difference of the forward guidance path and PLM y, represented w_2. This unanticipated monetary accommodation will create its own IRFs. Together with the previous shock it will mean that the PLM will be more persistent than that of the RE response describing the first shock, as now two shocks of the path are in the information set of the agents. This implies that the next period the monetary policy accommodation w_3 will have to be even smaller, not only due to the cumulative impact of the past shocks, but due to the adaptive learning as well. Formally this means that conditional on first period information expectations forecasted h periods ahead will be the IRF of the RE solution to an unanticipated shock: $y^f_{h|1,w_t,\tau} = y^f_{h|w_t,RE,\tau}$, where the shock's contemporaneous impact is $R_{1|1,\tau}$. Note the dependence of the beliefs on τ, it highlights that the larger the constant gain parameter in the learning the more more persistence it will create. This in turn will lead to higher long-run variance,i.e. mean squared error matrix of the states compared to pure RE solution.

The change of the persistence of the dynamics, i.e. the change of the policy function will be driven by the expectation error the agents make compared to the forward guidance. In the first period, they suffer the expectation error to the full extent of the monetary policy shock, however as time progresses, the agents learn the new path and adjust expectations even more, by how much it will be driven by their gain parameter. This diminishing impact of the shock is why the forward guidance does not occur. The marginal impact of an additional horizon of forward guidance in the period of announcement is zero, while in the period of implementation it is also smaller and smaller.

3.2 Simulation of Forward Guidance with Backward Looking Beliefs

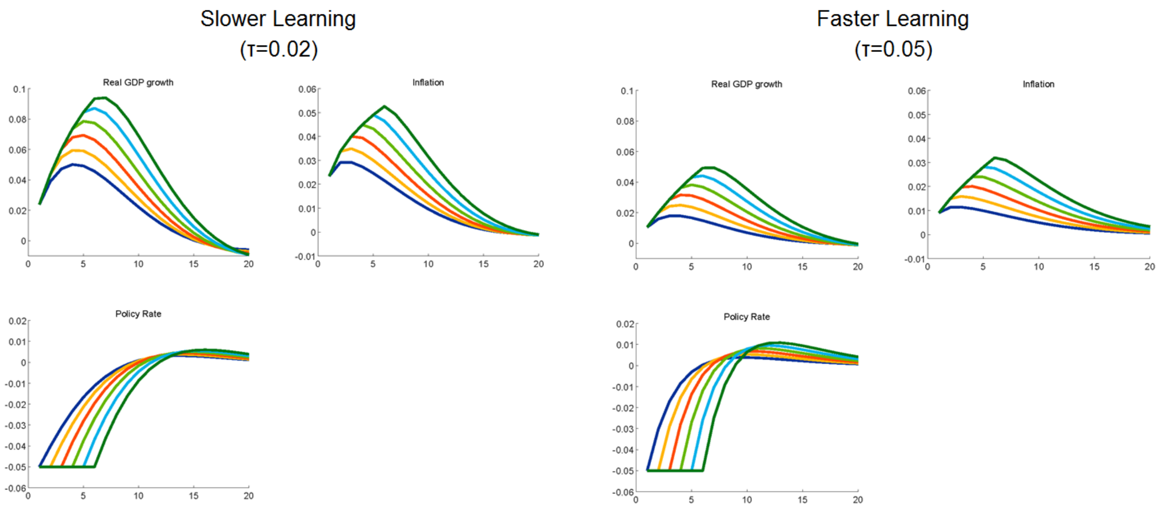
Figure 2 shows the IRFs of selected variables to a forward guidance of setting the interest rate to -0.05 percent per quarter for 1-6 horizon under different degree of learning. A larger τ means more adaptive expectations and more learning "away" from the RE dynamics. Similarly a lower τ translates to slower learning, less adaptive expectation, agents stick more their RE dynamic beliefs. The blue color represents one horizon, yellow two, red three, green four, light blue five and dark green six periods of low interest rates. The IRFs are plotting the ALM computed using extended path simulation of the forward guidance in the SW model[29].

Figure 2 illustrates that the size of response of the economy to unanticipated monetary policy shocks depends on the degree of attentiveness. The IRFs show that the forward guidance puzzle will not emerge if agents are inattentive.

Recall that initial beliefs are the same for the right and the left panel. The amplitudes are still different. If agents update their beliefs more frequently the structural shock of the monetary

[29] The parameters are that of the original model, for details please consider Smets and Wouters (2007).

Figure 2: Forward Guidance in the SW07 Model with Backward Looking Beliefs



Notes: Forward guidance of setting the interest rate at -0.05 (quarterly rate) for 1-6 horizon, then following the model's Taylor rule. The blue color represents one horizon, yellow two, red three, green four, light blue five and dark green six periods of low interest rates. The model was solved using the AL tools of the MMB. The model is initiated in the RE SS as in Slobodyan/Wouters (2012). A larger τ means more adaptive expectations and more learning "away" from the RE dynamics. Similarly a lower τ translates to slower learning, less adaptive expectation, agents stick more their RE dynamic beliefs. (Source: Author's calculations)

policy path translates to smaller impact on the state variables. A reason for this is the overall larger volatility of the model if agents change, i.e. update their expectations more. Interest rate variance is 10.8 for a higher τ, while 7.6 for a lower τ.

Although beliefs are updated with a constant gain, there is a signal extraction process involved in learning. Recall that the belief matrix is updated as the state forecast errors are regressed onto the states. If the model varies more around the steady state, the signal is less precise. In other words, the forecast errors are compared to the asymptotic variation of the states, and the larger the forecast error on state with a given means squared error matrix the more the coefficient loading on the respective state updates. Similarly the higher the uncertainty in the state given the forecast error, the less the beliefs in that dimension get updated. Furthermore the constant gain gives a weighting of the forecast errors when updating, the more recent information is given a higher weight. The size of the mean squared forecast error matrix will scale up and down the initial response, while it will also play a role how slow beliefs adapt to forecast errors made. Overall, we see that forward guidance puzzle does not emerge if agents are backward looking. The forward guidance will be perceived as a sequence of unanticipated monetary policy shocks. The degree of learning will determine the overall impact and the the persistent of the economy to forward guidance. In general the central bank has reduced control over the economy under adaptive, backward looking beliefs, as agents will respond less, but more persistent to a given shock than agents with fully ration expectations [30].

[30] Evans and Honkapohja (2012) show that setting τ close to ∞ the beliefs converge asymptotically to the RE

4 Forward Looking Beliefs under Adaptive Learning

In this section I present the approach to model forward looking adaptive beliefs given a credibly committed forward guidance. The optimal commitment to forward guidance was first analyzed by Eggertsson and Woodford (2003). It is also widely understood that the theoretically compelling model of the optimal forward guidance policy is dynamically inconsistent(Eggertsson, 2006). The optimal commitment means promising real interest rate below the efficient rate of interest rate, than it would be necessary in the absence of the shock. However when the time comes for the central bank to deliver on this promise, it has little to no incentive to do so. The central bank would rather reoptimize and implement a stance that is closer to the efficient level. This creates the dynamic inconsistency of forward guidance and the deflationary bias at the zero lower bound (Eggertsson, 2006). Woodford (2012) also recognizes this when he discusses how to implement the lower for longer commitment of forward guidance in practice. The only way to overcome the dynamic inconsistency for the central bank is to make a credible commitment: no matter what, it will stick to the plan of lower for longer. To describe such a credible action Andrade et al. (2019) compare the central bank's promise to Odysseus' journey when passing the Island of the Sirens. The Greek hero ordered his sailors to tie him firmly to the ship's mast, creating a literally binding commitment. When he is firmly tied, and had his men put the beeswax in their ears, was his ship safe to pass the Island. This way the central bank can commit herself not to be the charmed by the benefits of deviating from a preset forward guidance path, similar to song of the sirens and pull the ship off the preset course. By creating an Odyssean commitment the central bank can, in theory, create a credible case for forward guidance and achieve the optimal outcome. Gertler (2017) studied credible forward guidance in a hybrid expectations model, where agents had to learn about the projected deviations from the Taylor rule due to forward guidance, by learning about the trend inflation, and thus the implicit central bank inflation target. Gertler (2017) assumes learning about the steady state dynamics, in contrast I argue: should expectations already consider credibly anticipated news shocks, i.e. known future deviations from the Taylor rule, then they will only learn about the variance of these shocks as they realize, while the steady state of the economy remains unchanged. Therefore, given a credible commitment to forward guidance, if initial beliefs understood it, the puzzle will re-emerge. Anticipated news shocks enable to study the economic agents' response to information about anticipated events, shocks. Anticipated news shocks have been widely used in the DSGE literature[31]. In practice they are implemented as lagged structural shocks at the ergodic dynamics and are meant to capture the information content shock of future policy actions. The implementation nests the insight that agents' response to the information about (credible) future shocks today can be compared to past responses to current period shocks that have been announced in the past, and thus were fully and correctly

equilibrium. Note as the model is initialized at the RE equivalent, backward looking beliefs, setting τ to zero, will switch off learning, and deliver RE as well.

[31] See for example:Christiano et al. (2014), Del Negro et al. (2012) and Laséen and Svensson (2011).

anticipated. This necessitates that the model should be featuring anticipated news shocks already in the ergodic steady state. Anticipated new shocks clearly change the DSGE's structure, as in contrast to a model without them where the equilibrium is a VARMA model, anticipated news shock create an integration in the moving average terms turning the equilibrium model into a VARIMA.

As argued before, the SW07 model in its baseline version did not include anticipated news shocks about future monetary policy actions. Therefore in order to incorporate news shocks the baseline model's Taylor rule has to be expanded with the moving average representation of news shocks ($\sum_{l=1}^{L} \varepsilon_{t-l}^{R,FG,l}$), and thus the DSGE's structure has to change.

$$r_t = \rho r_{t-1} + (1-\rho)(\theta_\pi \pi_t + \theta_x x_t) + \varepsilon_t^R + \sum_{l=1}^{L} \varepsilon_{t-l}^{R,FG,l} \tag{12}$$

Anticipated news shocks in this representation are announcement by the central bank in period $t-l$ that the interest rate will change l periods later, i.e. in period t. They can be implemented in a DSGE with the help of new auxiliary variables $f_{k,t}$, that capture the cumulative impact of the forward guidance for periods beyond k and have a recursive definition, as proposed by Laséen and Svensson (2011):

$$f_{k,t} = f_{k+1,t-1} + \varepsilon_t^{R,FG,k}. \tag{13}$$

If the central bank communicates guidance on the interest rate for L periods ahead, there would be $1,2,3...L$ forward guidance shocks that affect the monetary policy rule in period t, and thus the Taylor Rule can be written as:

$$r_t = \rho r_{t-1} + (1-\rho)(\theta_\pi \pi_t + \theta_x x_t) + \varepsilon_t^R + f_t, \tag{14}$$

$$f_t = f_{1,t-1}, \tag{15}$$

$$f_{1,t} = f_{2,t-1} + \varepsilon_t^{R,FG,1}, \tag{16}$$

$$f_{2,t} = f_{3,t-1} + \varepsilon_t^{R,FG,2}, \tag{17}$$

$$f_{3,t} = f_{4,t-1} + \varepsilon_t^{R,FG,3}, \tag{18}$$

$$\vdots \tag{19}$$

$$f_{L,t} = \varepsilon_t^{R,FG,L}, \tag{20}$$

Note the missing index for the auxiliary variable f_t as it is the cumulated forward guidance shock in period t entering the Taylor rule. Thus f_t is the sum of all news shocks revealed in period t: $f_t = \sum_{l=1}^{L} \varepsilon_{t-l}^{R,FG,l}$. Note that the one period ahead news shock is $\varepsilon_t^{R,FG,1}$, it enters the Taylor rule with a lag, while all subsequent periods are denoted by: $\varepsilon_t^{R,FG,2}, \varepsilon_t^{R,FG,3}, ..., \varepsilon_t^{R,FG,L}$.

4.1 Smets Wouters Model with Forward Looking Beliefs

Forward looking beliefs need additional exogenous state variables to characterize the solution. It is important to discuss the span of the model space, as the additional shocks do enlarge the state space. However they can be spanned given a solution by other shocks. When building the econometric model for expectations the agents use this notion. They eliminate the MA polinomial, and build a model that is in the space spanned by the backward looking belief's MSV. Although the forward guidance shocks are predetermined, and show up as a larger error matrix in the reduced form, but the MSV state space of the state space representation is still the same. The easiest to think about the auxiliary variables as representations of the expected, endogenous yield curve, that account for anticipated news shock, that can be expressed as a result of a combination of other shock generating them. Assuming the yield curve is separately observable, the agents could fully back out the implied shocks. However with endogenous belief switching, I treat expectations to be endogenous, not predetermined, and thus unobservable. This means that the emission matrix H will not select the forward looking auxiliary variables. The model variables therefore will be described by:

$$\{y_{SW07+L}\} = \{y_{SW07}\} \cup \{f_t\} = \{y_{SW07}\} \cup \{f_{1,t}, \ldots f_{L,t}\} \tag{21}$$

This implies that the RE solution will incorporate anticipated news shocks. Furthermore once the model has to been solved, resulting reduced form policy function denoted by μ, T in Equation 2 and structural shock covariance matrix R will be larger as well under forward looking beliefs. With forward looking beliefs all information about the forward guidance will be known in period t. Initializing beliefs at the RE equilibrium, with news, i.e. forward guidance shocks, means that adaptive learning agents will be responding on impact to the forward guidance path shock, but beyond the first shock, there will be no updating and learning, since as time advances, the economy would follow their PLM, and agents would not make any expectation error anymore. This means that there will be learning only in period t, when agents are surprised, by the central bank's announcement. As monetary policy is fully credible and all action is anticipated learning will be only on impact, even under adaptive learning beliefs will not dynamically evolve, unless other shocks hit the economy. Thus the SW07 model with anticipated news shock will have $44+L$ endogenous variables, and $7+L$ exogenous processes, where L stands for the auxiliary variables needed to capture the horizon of anticipated news shocks. Note that forward news shocks are in the state space of the short-term interest rates. Therefore similarly to the baseline model the version with forward looking expectations can be spanned in the same minimum state variable space. Expectations will also be formed about the same 13 endogenous state variables as before[32]. Implying that the model the belief matrix that Φ_t will have a dimension of $(1+17+L) \times 13$.

[32]Realize the fact that although news shocks reveal future information, they are not forward looking variables, but lagged structural shocks!

To summarize the SW07 model with anticipated news shocks of horizon L features $44+L$ state variables. That is the extended state space y with the auxiliary variables of $f_1, f_2, ..., f_L$ is a $44+L\times 1$ vector. The beliefs are described by $13\times(1+17+L)$ by the Φ_t matrix. Where the first dimension is the number of forward looking states 13, and the second is the dimension of Z_t: the constant plus MSV variables and the auxiliary forward guidance states, in the space of the short-term rate used to form beliefs. The mean square error matrix R_t is a $(1+17+L)\times(1+17+L)$ squared matrix. Thus the belief matrix in terms of dimension has the following format:

$$\underset{(18+L)\times 13}{\Phi_{t|t}} = \underset{(18+L)\times 13}{\Phi_{t|t-1}} + \underset{1\times 1}{\tau} \underset{(18+L)\times(18+L)}{R^{-1}_{t|t}} \underset{(18+L)\times 1}{Z_t} (\underset{1\times 13}{y^{ALM}_t} - \underset{1\times 13}{y^{f}_t}) \tag{22}$$

$$\underset{(18+L)\times(18+L)}{R_{t|t}} = \underset{(18+L)\times(18+L)}{R_{t|t-1}} + \underset{1\times 1}{\tau} (\underset{(18+L)\times 1}{Z_t} \underset{1\times(18+L)}{Z'_t} - \underset{(18+L)\times(18+L)}{R_{t|t-1}}). \tag{23}$$

Figure 3 shows the IRFs of selected variables to a forward guidance of setting the interest rate to -0.05 percent per quarter for 1-6 horizon. As before blue color represents one horizon, yellow two, red three, green four, light blue five and dark green six periods of low interest rates. As all information becomes available in period t and there is learning on impact only and updating only regarding the long-run variance. The IRFs are plotting the ALM that coincides with the PLM formed aver period.

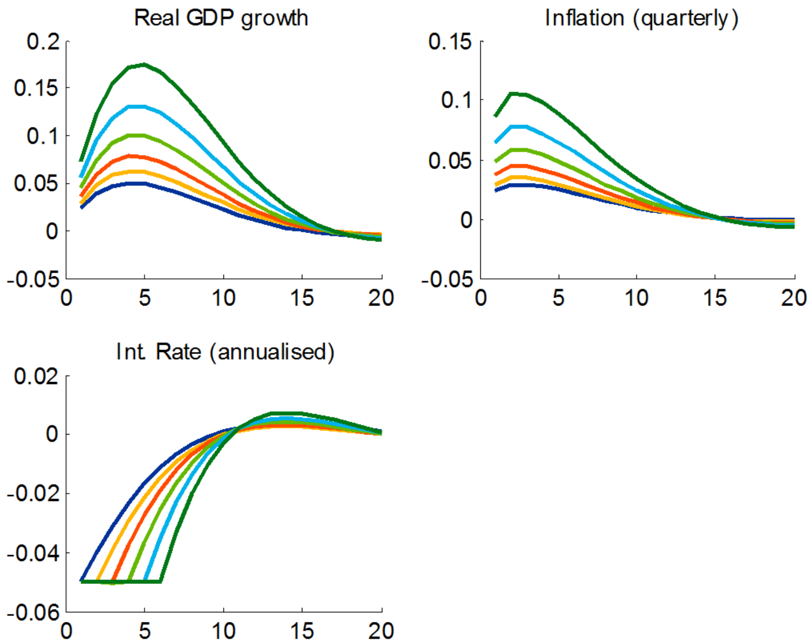
To account for adaptive learning the model has to be solved forward based on the current period beliefs. Depending on the size of the shock the agents will learn not only the actual path implemented, but update the long-run state covariance matrix, the mean squared error matrix R_t. Based on the updated beliefs Φ_t the agents will make expectations about the collection of shocks that implements the forward guidance path. The idea for the solution to be forward looking is to ensure that the anticipated path of the short-term interest rate, described by the PLM, has to coincide ever period with that announced by the central bank. To solve the model forward, I developed a non-linear solution in order to find the period t subjective shocks: both current period monetary $\varepsilon^R_{t|t}$ and the sequence of anticipated news shocks $\sum_{l=1}^{L}\varepsilon^{R,FG,l}_{t-l}$ that implement the forward guidance path. Formally, it is a simulated method of moments estimation to match expected interest rate path to the forward guidance. This non-unique mapping ensures that under period t expectations the monetary policy shock and anticipated news shock implement a forward guidance of L periods at the path announced by the central bank.

The methodological innovation of the paper is present here. Instead of solving the model backward for the implied policy function, as deterministic simulation would do form a steady state in the future; I will solve the beliefs forward to the find the belief consistent sequence of shocks (monetary policy and anticipated news), that implements the forward guidance path. The approach is similar in spirit to conditional forecasts of Waggoner and Zha (1999) applied to DSGEs.

The solver starts with a guess on monetary and anticipated news shocks, calculates the PLM, and iterates until the forward guidance path matches the PLM, by altering the shocks.

This implies that every period the forward looking agents re-anchor their expectations to the path, anticipating a new sequence of shocks [33].

Figure 3: SW07 Model with Forward Looking Beliefs



Notes: Forward guidance of setting the interest rate at -0.05 (quarterly rate) for 1-6 horizon, then following the model's Taylor rule. The blue color represents one horizon, yellow two, red three, green four, light blue five and dark green six periods of low interest rates. The model was solved using the AL tools of the MMB. The model is initiated in the RE steady state as in Slobodyan/Wouters (2012). (Source: Author's calculations)

Figure 3 illustrates that even with adaptive learning forward looking belief with news shocks is still subject to the forward guidance puzzle. Most of the learning takes place initially as the announcements is made, however as it is implemented the PLM being anchored at the path will ensure that, unless other shocks hit the economy, no more learning is needed. Should the ALM not follow the PLM of the forward looking beliefs, as it will be the case under endogenous belief switching in later sections, learning and re-anchoring of expectations become relevant.

5 Endogenous Beliefs Switching

In previous sections I have established how forward guidance works if agents do or do not form beliefs about anticipated shocks. I have argued that the baseline DSGE models, like the SW07 model, (usually) do not feature anticipated news shocks, and thus standard models with

[33] As eluded before Cole (2015) only estimated the current period shock and did not ensure PLMs also match the path every period.

constant gain adaptive learning beliefs will not feature the forward guidance puzzle. Furthermore, even though backward looking agents will perceive the forward guidance as a sequence of unanticipated monetary policy shock, due to learning they will respond more and more persistently to it, as it gets implemented. Therefore the idea naturally emerges, how could one model adaptive beliefs that can learn to be forward looking, attentive? How can one design a model of belief switching that is tractable and combines existing understanding? Endogenous belief switching is the answer, that combines endogenous regime switching DSGE with adaptive learning.

Endogenous regime switching has been proposed to solve models with occasionally binding constraints when agents have rational expectations by Binning and Maih (2017). Lansing (2018) used endogenous regime switching at the zero lower bound. By allowing for expectations to feature regime switching between two local equilibria, labelled the "targeted" and "deflationary" regimes, he studied the role of central bank anchoring. In his model the expectations were a result of model averaging, and switching was driven by the root mean squared forecast error of beliefs given a logistic probability. Similar to his model, endogenous belief switching is a mixed expectations equilibrium, where mixing is not following a logistic rule, but using a Kálmán filter to estimate the perceived likelihood of the belief's state representation factor model. Note while Lansing (2018) focuses on learning about the long-run equilibrium conditions, my framework studies adaptive learning about the policy rule and the resulting dynamic structure of the economy.

Carvalho et al. (2019) propose to link long-run inflation expectations to a state-contingent short-run forecast errors using adaptive learning agents. Their model is built on constant gain learning about the long-term inflation expectations. Due to nominal rigidities in price setting shifting subjective beliefs become self fulfilling. I extend this literature by incorporating learning based on constant gain adaptive expectations about the model dynamics rather than the long-run inflation target[34].

I start by showing that for endogenous belief switching two ingredients are needed. First, beliefs should be time-varying. This is ensured by adaptive learning. Second, switching of beliefs should take place based on likelihood, i.e. current period information of agents. To capture the second point, I consider belief formation, with symmetric equilibrium of information sets. Agents have symmetric information, and observe the MSV that spans beliefs under both regimes. Each agent has a private, symmetric beliefs, in form of an a priori probability that the expectation of other agents' are either backward or forward looking. Given symmetry of information the agents can be represented in a representative agent framework. In settling for a representative agent framework, I also avoid the problematics of heterogeneous beliefs: simultaneity and higher order expectations.

Endogenous belief switching features self-referentiality. Beliefs become self-fulfilling,that is

[34] Noe that both Lansing (2018) and Carvalho et al. (2019) focus on the long-run beliefs, my approach is closer to Chang et al. (2018) who focus on two distinct regimes differentiated by their dynamics.

in absence of the shocks the a priori regime is inherently more likely to prevail, compared to the alternative that perceives shocks to explain the dynamics of the economy. Should the economy be hit by shocks that cannot be attributed to the a priori beliefs, beliefs will change. Endogenous belief switching nests the problem of learning both the shock process and the dynamic nature of the economy makes extraction of information supporting switching complex.

The easiest to relate to it, to recognise that agents with endogenous belief switching behave as time varying coefficient estimating econometricians. They try to identify the best forecast rule for the environment, not only following a least squares learning but combining the two distinctly different regimes, solutions of the DSGE given forward or backward looking beliefs.

Under backward looking beliefs they consider the economy to be a simple VAR of a given length, while under forward looking beliefs the their beliefs are following a VARIMA process, where the moving average term, i.e. the subjective shock process, follows integration of the order the length of the forward guidance. This creates a problem, as every period the agents understand that the model dynamics can be of higher order of integration in the moving average root process[35] As in Lansing (2018); Bullard and Duffy (2004) I assume agents are aware of the two local beliefs, and the resulting self-fulfilled equilibria of the forward and backward looking beliefs[36].

Every period agents form expectations knowing both regimes, however not knowing the true Taylor rule, that is a choice of the central bank. The endogeneity denotes that the equilibrium is jointly determined by agents learning the form of the Taylor rule and the central bank's decision to choose either commitment, and indicating forward looking beliefs and a Taylor rule with anticipated news shocks, or control only the contemporaneous short-term rate, resulting in backward looking beliefs and a Taylor rule only featuring contemporaneous monetary policy shocks. In my paper I only consider a central bank that is committed to delivering the promised forward guidance.

Agents are fully aware of the potential implication of forward guidance, in that they evaluate the model's response under both alternatives. They form optimal expectations given regimes and their two PLMs will be then combined using probability weights. This combination is allowed as the SW07 model is linear Gaussian model. They then update their a priori beliefs given the expectation errors and the model's actual law of motion[37]. The updating is using the regime switching Kálmán filter on the state space of the beliefs of backward looking agents, they can do that as discussed in Section 4.1.

[35] Recall that even a single VARMA process needs the MA polynomial to be inverted using Wold's theorem, given quite strict conditions, into a VAR(∞). Thus adaptive backward looking beliefs, characterised by iid monetary policy shocks, and constant gain learning can never learn to behave as forward looking agents, as the constant gain approach enforces finite memory, that a VAR(∞) violates. This means that regimes might converge in beliefs to be close to each other, yet they will be always two distinct regimes.

[36] In contrast to Bullard (2010) there is no concern about the possibility of being stuck in a liquidity, deflationary trap, as the economies implied by the beliefs feature the same asymptotic steady states

[37] Regime switching DSGEs have been already explored in the literature see for instance Sims et al. (2008), Farmer et al. (2009), Farmer et al. (2011), Bianchi (2012) and more recently Maih (2015) and Borağan Aruoba et al. (2017). However neither approach considered a belief formation on that incorporated endogenous belief switching.

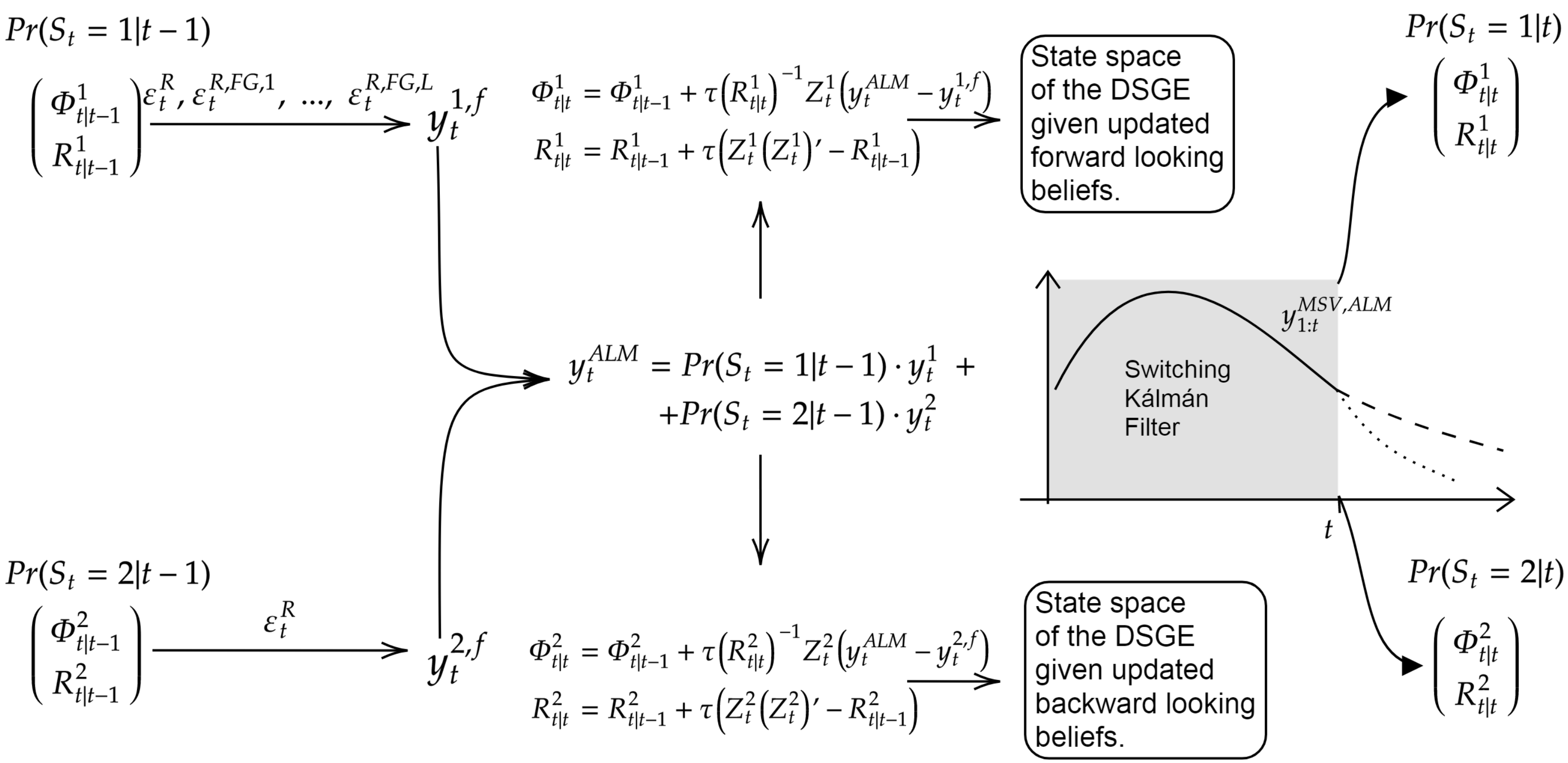
Evans and Honkapohja (2012) call this approach the ”restricted perceptions equilibrium”. Agents are allowed to form expectations that are only optimal for a limited class of self-confirming, self-fulfilling beliefs, i.e. PLM (Evans and Honkapohja, 2012, p.57). Note however that endogenous belief switching is not a misspecification on the agent's PLM about the true economy. On the contrary there is no limited information of the agents, they are econometricians, and given a belief they can learn, asymptotically the true ALM of the model. To build intuition for the concept of endogenous belief switching, let us consider a simple example: As stated before the adaptive learning agents behave as econometricians [38]

Assume for an extreme case, that there was no precedent of forward guidance and anticipated news shock being employed in the past. Then that means agents would learn to exclude the yield curve from their econometric model, as it would have no significant explanatory power beyond the MSV and the true dynamics of the ALM. In other words, even if yields were to be used as regressors to form beliefs, asymptotically they would be excluded from the PLM regressors, i.e. beliefs would converge on a coefficient of zero. Therefore the true model of the economy will be perfectly learned by agents running the VAR regressions. Now consider the alternative regime, where the past has been characterized by forward guidance commitments, that were also credible anticipated. Then the agents would learn the true dynamics of the model, where not only current but past monetary policy shocks matter. The past shocks will translate into VARIMA models with integration on the MA terms. The stability of such models is the concern f the forward guidance puzzle. Therefore it is natural and reasonable the assume that econometricians would reject a model where the local stability of the economy is at question. Should the model of forward looking beliefs generate a PLM that is making very bad predictions, the beliefs will switch to a stable, VAR model.

The process is summarized on Figure 4. It shows how initial beliefs, on the left, will be updated through learning, in the middle, to give rise to state representation of the regimes and allow for belief switching based on the MSV IRF, on the right. To understand the role of learning and of belief switching let us walk through the diagram.

[38] If their belief is backward looking they consider iid monetary policy shocks, if their beliefs are forward looking, they feature anticipated news shocks. A practical example for anticipated news shock from an econometrician's prospective would be the predictive power of past yield curve information. Realize if the central bank follows a Taylor rule with iid innovation, then the coefficients responding to the yeilds would converge to zero.

Figure 4: Endogenous Belief Switching



(Source: Author’s illustration)

5.1 Timing Assumptions and Adaptive Learning

In my model the agents are aware of the global structure of the economy, yet specific timing assumptions are needed to avoid the simultaneity of expectations and infinite higher order beliefs.

I assume that the agents start out every period with an a priori belief about the probability they assigns to the two beliefs $Pr(S_t = i|t-1)$. The agents inherit this knowledge from the previous period in form of belief matrices under both regimes $\Phi^i_{t|t-1}$ and $R^i_{t|t-1}$. The agent then observe the stochastic shock(s) depending on the information set of the regime. For forward looking beliefs, they observe the current monetary policy shock and the sequence of anticipated news shocks implementing the forward guidance, $\varepsilon^R_t, \varepsilon^{R,FG,1}_t, ..., \varepsilon^{R,FG,L}_t$. For backward looking beliefs, she only observes the current monetary policy shock ε^R_t. Given beliefs and the shocks, the agents form the PLM for both beliefs: $y^{i,f}_t$. It means that they have to respond first, and can only learn afterwards. The agent perfectly understands that the two beliefs are distinct, however he has to respond to the shock with his a priori information, and is allowed to learn afterwards. This timing restriction is a common property of adaptive learning models, it ensures that expectations are not contemporaneous, i.e. agents do not observe y_t, that would lead to simultaneity problems. Furthermore I assume E-stability of the actual law of motion dynamics. (Evans and Honkapohja, 2012, pp.237-263.)

I make this important assumption about the timing learning and decisions to avoid infinite regress and expectation determination problems of heterogeneous beliefs.

Using the a priori probabilities the agents weight the regimes and act accordingly. Weighting the beliefs with their a priori probabilities is a complex combination of best responses. Therefore it is the best subjective response the agents can make given the information and assumption about learning.

This gives rise to the ALM of the economy:

$$y^{ALM}_t = Pr(S_t = 1|t-1) \cdot y^1_t + Pr(S_t = 2|t-1) \cdot y^2_t \tag{24}$$

Note that no observation of the states is necessary until this stage. The ALM and both of the PLMs are known by the agents. Subsequently the agents are allowed to update both beliefs using 8 and 9. Here is where learning is taking place: as long as PLMs of the two beliefs are different expectation errors will arise and lead to updating. However if one of the regimes is active only the other will get updated. This will lead to an eventual convergence of beliefs. This convergence of beliefs is the reason, I build endogenous belief switching on a constant gain adaptive learning framework.

Finally given the updated beliefs the DSGE is cast into the state space representation. These state representations of the beliefs are used to update the a priori belief, which regime fits better the observable MSV state of the IRF:

$$y_t^i = \mathbf{F}(\Phi_{t|t}^i, R_{t|t}^i)y_{t-1}^i + \mathbf{w^i}_t, \tag{25}$$

where y_t^i is the state given belief i, $\mathbf{F}$ is the state-transition matrix, that is dependent on the updated beliefs, and finally the state noise $\mathbf{w^i}_t$ that is assumed to be normally distributed with mean zero and covariance matrix $Q_t^i = R_{t|t}^i \Sigma \left(R_{t|t}^i\right)'$. Where Σ is a diagonal matrix with the size of the exogenous shocks, providing the scaling of their variances. Therefore the state noise is multivariate normal: $\mathbf{w^i}_t \sim N\left(0, Q_t^i\right)$.

The goal of the agents is to make sense of the economy. To that end they will filter the observed MSV states starting from the onset of the monetary policy action agent to adapt their beliefs regarding the shares. This requires them to filter the whole path of the economy, and try to reinterpret past states of the economy in light of the current beliefs. This enables them to reinterpret a consistent delivery of the forward guidance from being a sequence of backward looking shocks into a credible forward guidance.

The agents switch beliefs by evaluating the likelihood of the perceived shocks' probability in light of the observed states of the MSV using the switching Kálmán filter (Murphy, 1998). For the mathematical details on the switching Kálmán filter please see Appendix 7. The intuition is the following, they use their most current beliefs to reinterpret the past and figure out which belief is more consistent with what they observed. This creates a stability of beliefs, that is required. Without shocks hitting the economy agents should have little reason to switch the narrative of the past and present. If no shocks hit, then the beliefs are not updated, then the best description of the past is the one that generated it.

5.2 Endogenous Belief Switching as Accelerated Learning

Belief switching enables the acceleration of learning. Recall that backward looking beliefs adaptive learn the forward guidance implementation as a sequence of contemporaneous monetary policy shock, and become more persistent. However there is relevant surprise to beliefs, and thus expectation errors in the initial period only, if beliefs are forward looking and get anchored at the path implied by the forward guidance. Endogenous belief switching enables to jump between backward and forward looking expectations. The switch takes place based on the probability allocated to either of the beliefs. The representative agent chooses the belief to follow that has the highest (posterior) probability to have generated the observed path. This probability has a natural interpretation: it captures the confidence of the agents that the central bank forward guidance commitment will be followed, thus it is a measure of central bank credibility. In contrast to Kálmán filter learning, where the speed of adaptation to errors is faster compared to constant gain learning, nonetheless they only allow for smooth evolution expectations. Endogenous belief switching creates a discontinuous change in the model, which creates a non-linear effectiveness of the forward guidance: nesting the puzzle and its solution at the same time. If beliefs are backward looking forward guidance has no effect, if they are forward looking, the forward guidance puzzle emerges. Belief switching means that the agents not only behave as an econometricians, when updating their PLMs, but then behaves similarly when considering the two extreme cases, being either forward looking or backward looking to decide for either or a mixture of the two. This decision is based on the likelihood that the observable model was generated by to either of the models consistent with the beliefs. The evolution of this probability enables a more dynamic change, switching.

Indeed, if the economy is allowed to evolve infinitely, due to adaptive learning the distance between backward and forward looking beliefs in the MSV space would almost but vanish as they would converge to one dynamics. But since the two beliefs feature different shocks they will be only observationally equivalent, but still distinct, and thus will not be the same. The existence and properties of asymptotic equilibrium to which the beliefs converge is beyond the scope of this paper and will be explored in future research[39]. In this respect one can interpret the initial conditions of each belief, being centered at their RE equilibrium, as an off equilibrium state. Realize however that depending on the initial share of beliefs and the specifics of filtering, either can be an absorbing equilibrium. Furthermore, with constant gain parameters in the range of the empirical estimates, only 2% of the variance scaled errors will be used to update beliefs, making them more persistent.Since both backward and forward looking belief share the same steady state and are E-stable, the only way to set them apart in the steady state is by observing the long-run variance of the states. In what follows I abstract from this concern. Switching of beliefs takes place based on the probability that the deviation from the steady state the observed states show fits either model better.

5.3 Belief Switching Illustration in One Dimension

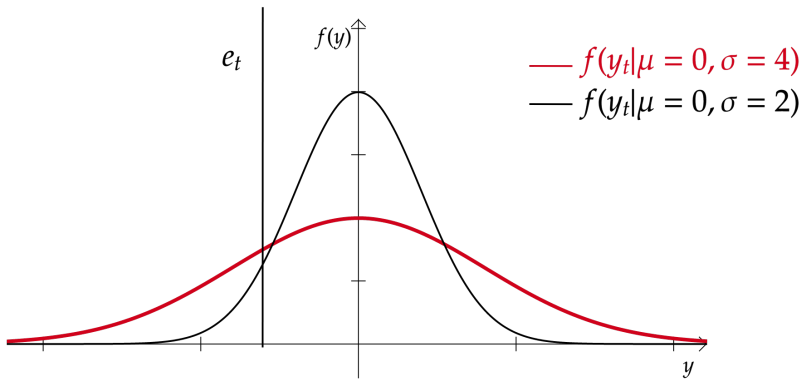
To illustrate how adaptive learning with endogenous belief switching works let us build a simple example. First, we need to answer what is the minimal number of states that is required to make switching possible? If the MSV is univariate, that is, the only state variable is enough to characterize the economy, the only state we can have should capture monetary policy. Denote this state as the short-term rate. The problem of only one dimension is that agents have no way to tell during forward guidance the two regimes apart, since the only observable is the short-term rate. Recall, that the central bank implements forward guidance by setting the short-term interest rate to a pre-announced path for a given number of periods. Since, this state variable follows the same path under both beliefs, there is only room to switch beliefs on announcement. As every period is the same irrespective of the beliefs. The short-term rate has the same path under both regimes.

[39]Simulations indicate that without shocks the backward looking beliefs, due to featuring less uncertainty are the absorbing regime. However depending on the assumption of the signal to noise ratio, alternative equilibrium can arise: e.g. if the observation error and thus noise around the model are large, there will be no distinction between the beliefs in terms of probability and no endogenous switching will take place. Then the asymptotic equilibrium will be centered around the equilibrium where adaptive beliefs make no errors given an ALM of 50% of each beliefs.

For the purpose of simplicity, assume that a priori beliefs are backward looking. Then all it suffices is to calculate the likelihood function of the states given switching takes place. Assume then that the model with forward looking beliefs is filtered as a normal with mean zero and 4 standard deviation denoted as $f(y_t|\mu = 0, \sigma = 4)$. Similarly the agent has knowledge about the state distribution, beliefs remain backward looking. Backward looking beliefs feature less structural shocks, less sources of uncertainty and are thus display a smaller variance than forward looking beliefs. Assume, that the backward looking beliefs regime has the same mean and half the variance, e.g. 2 standard deviation, $f(y_t|\mu = 0, \sigma = 2)$.

Figure 5 displays the state densities given initial backward beliefs: red is the sate distribution if switching to forward looking beliefs takes place. It displays a more dispersed short-term rate. The black line is the state distribution with backward looking beliefs. Assume the central bank

Figure 5: Belief Switching Illustration - Initial Beliefs

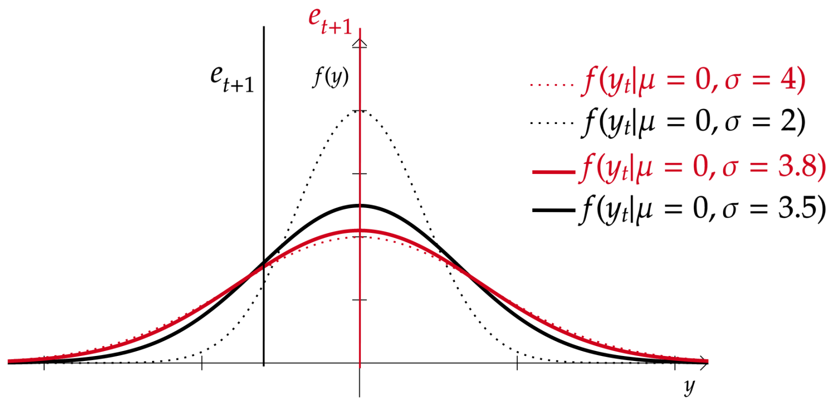


(Source: Author's illustration)

announces forward guidance for two periods. Since the model's observable dimension is one, and no monetary policy action was expected, the current shock is perceived almost as the same movement in the short-term rate, settig it to the path.

Evaluating the likelihood in this case is simple, it means to find the model that has a larger density given a shock. If the size of the perceived monetary policy shock,e_t results in a state of the economy where the density of the belief switching to forward looking is larger, agents will become centered at that point. Note as the subsequent period of the forward guidance is implemented, there is little room for belief switching, as the only one dimension is fairly restrictive in terms of shape of the model likelihood. Figure 6 illustrates what happens in the next period, the dashed lines are the beliefs of the previous period, i.e. the PLMs of the previous, and the solid lines are the PLMs of the current period after learning, but before evaluating switching. Assume that after learning the resulting in a standard deviation of $3,5$ for backward, and 3.8 for forward looking model.

Figure 6: Belief Switching Illustration - Second Period Beliefs



(Source: Author's illustration)

With adaptive learning not only the perception of the shocks' will change, but more importantly the state variances as well. As it is seen on Figure 6, in the next period the two distributions get closer to each other. Forward looking beliefs are only updated little, as the realized shock was smaller than the standard error of the interest rate and thus the variance shrinks. The effect of learning is larger on backward looking beliefs. As the expectation adjust more, as there has been a large forecast error made. Since switching took place, and the PLM for the forward looking is markedly different, due to the forward guidance puzzle, all expectation errors shifts the model mean and variances significantly more. I abstract from the mean shift, as it is difficult to tell which way and how much the bell curve moves, but what matters, is that the two distribution will get closer to each-other as shown on Figure 6. Since beliefs switched to forward looking, they get updated less, while backward looking beliefs get updated more.

As now the forward guidance gets implemented, the shock perceived under the beliefs is also change. Under forward looking beliefs, no more monetary policy shock is present, shown with e_{t+1} in red. With backward looking the implementation of the path is just another current period monetary policy shock, shown in black on Figure 6. In this case based on the current densities alone, keeping in mind that the a priori was forward looking, expectations remain forward looking. Without learning there is only selection of beliefs by the central bank's signal, and there is little room for dynamic evolution. What learning enables is the convergence of the beliefs, to the one that is the better fit. Switching enables the determination of expectations towards which beliefs should converge, and depending on perceived shocks it enables jumps not only in the first, but in later periods as well, thus enabling to study policy actions, like the recurring extension of forward guidance. Therein lies the necessity for the two components of endogenous belief switching: adaptive learning and regime switching, the former enables the evolution of the two regimes over time, slow convergence, the latter enables the faster adjustment of expectations in response to policy.

5.4 The Three Equation New Keynesian Model

Now that we have built intuition for belief switching in one dimension, let us make the next small step towards a general equilibrium framework. Consider the adaptive learning version of the standard three equation New Keynesian textbook (henceforth 3EQ) model following (Ravenna and Walsh, 2006). It is a small model that re-creates the trade-off monetary policy faces between stabilizing the inflation rate versus the output gap due. The trade-off emerges due a cost component dependent on short-term rates in firms' marginal cost. The model in its simplest version consists of the following equations:

1. IS Curve
$$x_t = E_t[x_{t+1}] - \frac{1}{\sigma}(r_t - E_t[\pi_{t+1}]) + \varepsilon_t^{IS} \tag{26}$$

2. Phillips Curve
$$\pi_t = \beta E_t[\pi_{t+1}] + \kappa(\sigma + \eta)x_t + \kappa r_t \tag{27}$$

3. Taylor Rule
$$r_t = \rho r_{t-1} + (1-\rho)(\theta_\pi \pi_t + \theta_x x_t) + \varepsilon_t^R + \sum_{l=1}^{L} \varepsilon_{t-l}^{R,FG,l} \tag{28}$$

Where anticipated monetary shocks only exists if the agent forms beliefs about them. Introducing adaptive learning to this model means replace the expectations with forward looking variables predicted using constant gain learning:

$$E_t[y_{t+1}] = E_t \begin{bmatrix} x_{t+1} \\ \pi_{t+1} \end{bmatrix} = y_t^f = \hat{\alpha}_{t-1} + \hat{\beta}_{t-1}' \begin{bmatrix} x_{t-1} \\ \pi_{t-1} \\ r_{t-1} \end{bmatrix} = \hat{\Phi}_{t-1} \cdot Z_t \tag{29}$$

Consider the calibration of the parameters:

Table 1: Parameter calibrations in the 3 Equation model

Parameter	Value	Description
β	0.99	Discount factor
η	1	Frisch elasticity
κ	0.0858	Slope of the PC
ρ	0.900	Interest rate smoothing
θ_π	0.150	Inflation response
θ_x	0.013	Output gap response
$\sigma_{\varepsilon^{IS}}$	1	Standard error of IS curve shock

Table 1 – Continued

Parameter	Value	Description
σ_{ε^R}	1	Standard error of Monetary Policy shock
$\sigma_{\varepsilon^{R,FG,l}}$	1	Standard error of l period ahead forward guidance shock

Given this parametrization the model's MSV has two dimensions, i.e. there are two pre-determined variables.A potential problem that might arise is the non-uniqueness of the state representation of the dynamic system. In macroeconomic models this is overcome by timing assumptions, that clearly separate model variables to states and controls. Mapping macroeconomic RE models to adaptive learning representation requires these assumptions to be clearly defined. I resort to the MSV implied by the timing used by Ravenna and Walsh (2006), and leave the discussion of transfer function representation to future research. Therefore the two states describing the system is the interest rate and output gap, conveniently mapped by the two structural shocks.

Table 7 shows the impulse response function of the 3EQ model to a forward guidance shock of different horizons. The left panel illustrates the impulse response if beliefs are backward looking and the agent has a learning coefficient of $\tau = 0.02$. The right panel displays the impulse response of the model variables to the forward guidance shock with forward looking beliefs and adaptive learning. As the model does not have strong persistence, the forward guidance is relatively short lived in both cases. The left panel shows that the marginal impact of the sequence unanticipated monetary policy shocks decline, and thus the marginal effect of implementing the path vanes. This diminishing marginal effect on other variables explains the flattened response of output and inflation during the implementation of the forward guidance path.

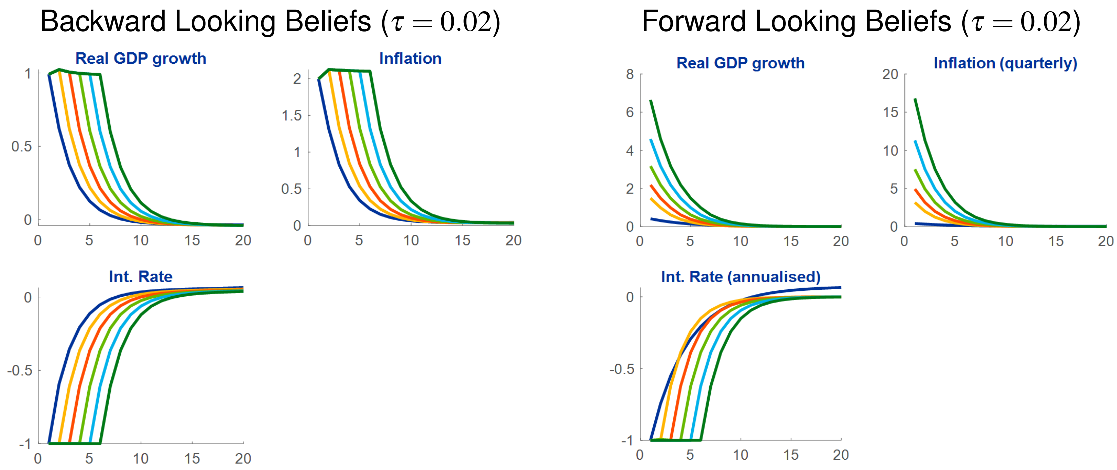
Given backward looking beliefs the long-run covariance matrix of the states is the following:

$$Q^2 = \begin{bmatrix} 1.1462 & 0.6465 \\ 0.6465 & 0.7318 \end{bmatrix}, \tag{30}$$

where the first dimension is the short-term interest rate and the second is the output gap. In terms of scale, the state covariance matrix Q is small. The scale can be thought of as the measure of the state dispersion, best characterised by Q matrix's mean singular value, i.e. trace per dimension being 0.9390, that is smaller than 1.

The right panel displays the forward guidance puzzle. Notice the increasing marginal effect on impact for both output and inflation, as the length increases so does the impact of forward guidance.

Figure 7: Forward Guidance Shock in 3EQ Model with Adaptive Learning



Notes: Forward guidance of setting the interest rate at -1 for 1-6 horizon, then following the model's Taylor rule. The blue color represents one horizon, yellow two, red three, green four, light blue five and dark green six periods of low interest rates. The model was solved using the AL tools of the MMB. The model is initiated in the RE solution. For further reference see: Wieland et al. (2012)
(Source: Author's calculations)

The long-run covariance matrix of the forward looking beliefs is the following:

$$Q^1 = \begin{bmatrix} 2.1272 & 0.9927 \\ 0.9927 & 0.8580 \end{bmatrix} \tag{31}$$

As expected, the interest rate has larger variance than before, while the correlation between the two sates is also increased. The scale of the state covariance matrix in terms of its mean singular values is also higher 1.4926. This tells, that forward looking beliefs generate a state space with higher dispersion.

Note the different dynamics of the unanticipated monetary policy shock seen as the different IRFs in blue between the two panels. It has smaller impact under forward looking beliefs. The source of the difference is due to adaptive learning. As the forward looking model has larger variance, agents learn less from the forecast error of given size compared to being backward looking, where the same size of forecast error makes them update their beliefs more.

This is further supported by the observation, that learning accelerates as the amplitude of the impulse response increases. A higher amplitude results in larger forecast errors, and as learning is constant gains weights on least squares updating, a higher forecast error given a variance translate into accelerated learning.

Having seen the two extremes, let us turn to endogenous belief switching.

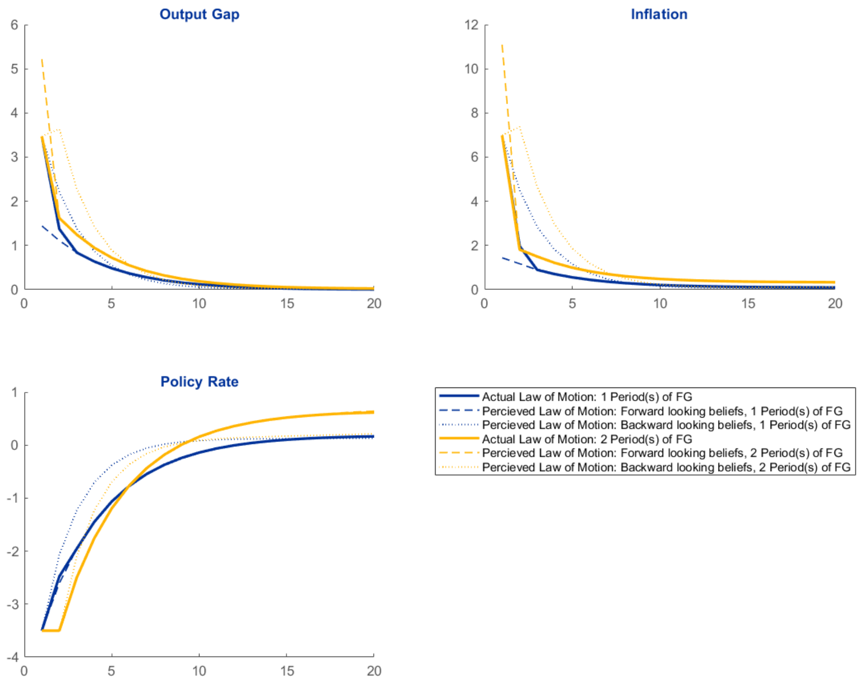
5.5 Endogenous Belief Switching in Higher Dimensions

The following simple illustration is meant to show, how backward looking agents can switch to be forward looking. It highlights the role of belief switching in response to the central bank's forward guidance. The following example will give the case of switching even if no periods of forward guidance has been observed before.

Consider an economy, where the a priori share of forward looking agents is 0. In this backward looking environment the central bank announces the forward guidance of 2 periods setting interest rate lower than the equilibrium, by 3.5$\%$, i.e. three and a half times the standard deviation of the average monetary policy shock for the next two periods. It is important for the sake of illustration that the size of the shock is chosen to be very large in comparison the the average monetary policy shock. This will enable switching as it will be, by design, very unlikely that such a huge response could have been caused by a single monetary policy shock. In contrast, if beliefs are forward looking then the variability of the states is larger, thus a size of this shock will be more probable and command a larger likelihood.

Figure 8 shows the IRFs of all three variables to a forward guidance of setting the interest rate to -3.5 percent for 1 and 2 horizon.

Figure 8: 3EQ Model - Endogenous Belief Switching



Notes: The Figure shows the IRF of the state variables in the three equation New Keynesian model to a forward guidance shock when initial beliefs are backward looking. Forward guidance shock means setting the interest rate at -3.5 (quarterly rate) for 1 and 2 horizons, then following the model's Taylor rule. The blue color represents one horizon, yellow two horizon of low interest rates. The dotted line represent the perceived law of motion for current period under backward looking beliefs. The dashed lines show the perceived law of motion of variables under forward looking beliefs. (Source: Author's calculations)

As before blue color represents one horizon, yellow two horizon forward guidance of low interest rates. The solid lines show the ALM of the model, that is the weighted average of the PLM under each regime. Backward looking beliefs' PLM is also displayed in dotted lines, while the forward looking beliefs' PLM is shown with a dashed line. Recall, initial beliefs are backward looking: since beliefs switch only after having responded to the shock, the ALM of the initial impact is overlapping with the PLM of the backward beliefs. In period 0, when the central bank makes the announcement both beliefs will generate a PLM. Since backward looking beliefs are only surprised in the realization of the shock, not only forward, but the backward looking beliefs will also make forecast errors, learn and updated. Neither would have anticipated such large monetary policy accommodation. Given initial, RE beliefs the agents form the PLMs. The PLM of backward beliefs will be the ALM. In response to this forward looking beliefs change from the RE to a larger variance:

$$Q_{RE}^{1} = \begin{bmatrix} 2.1272 & 0.9927 \\ 0.9927 & 0.8580 \end{bmatrix} \rightarrow Q_{1,AL}^{1} = \begin{bmatrix} 2.7303 & 0.7438 \\ 0.7438 & 0.9607 \end{bmatrix} \tag{32}$$

Similarly backward looking beliefs will learn the larger short-term rate variance, as on average only a size of 1 would be compatible with the RE beliefs, while a zero shock in the IS curve is also not anticipated:

$$Q^2_{RE} = \begin{bmatrix} 1.1462 & 0.6465 \\ 0.6465 & 0.7318 \end{bmatrix} \rightarrow Q^2_{1,AL} = \begin{bmatrix} 1.3425 & 0.4519 \\ 0.4519 & 0.9247 \end{bmatrix} \tag{33}$$

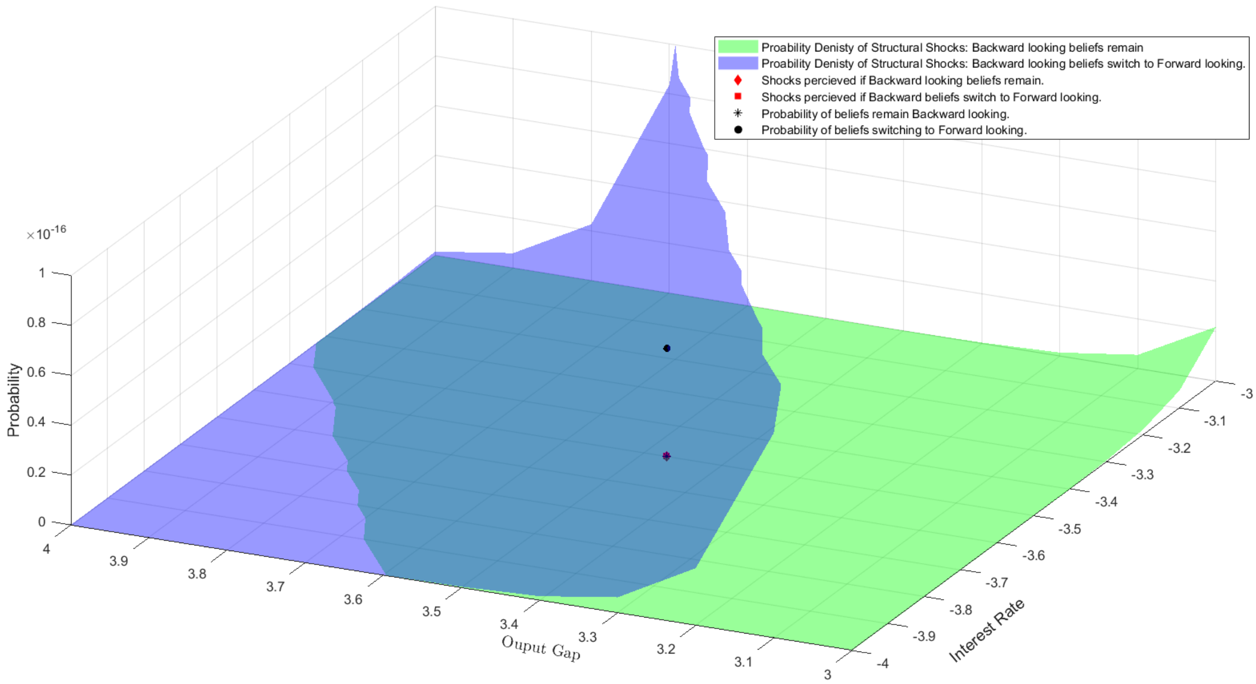
Once the ALM is realized, and learning took place, the agents will update their a priori beliefs about the probability of the regimes, i.e. the probability attached to a credible forward guidance. In doing so they will back out the probability that the response could have been a result of backward and forward looking behaviours, respectively. Finding that such a large movement in the states could have been only generated by an economy where the central bank can credibly commit to forward guidance, the agents choose to switch to forward looking. Thus the next period the PLM of forward looking will be the ALM of the economy. Figure 8 shows this as the dashed lines, forward PLM, becomes overlapping with the solid line, ALM, revealing the the dotted line, the backward looking beliefs. Since no further shocks hit the economy,i.e. as long as the forward guidance is not deviated from, forward looking beliefs prevail.

To understand the switching in detail, let us reconstruct the figures representing the state probabilities under both regimes. Figure 9 shows the initial, updated belief distribution of the MSV under both regimes in the first period [40]. It maps any combination of the state space to a probability. The green surface is that of the initial beliefs remaining backward looking, while the blue surface represents the probability of the states given backward beliefs switch to forward looking. Note that both have zero mean, i.e. the steady states are shared. Furthermore forward looking beliefs result in larger current interest rate state variation compared to backward looking beliefs in the MSV. Therefore the blue multivariate normal state distribution is more spread out. Recall that the agents start out with backward looking beliefs, and thus the economy follows the PLM of the backward looking beliefs upon a shock resulting in a drop -3.5 of the interest rate and an increase of output gap to 3.47. This is the ALM of the model. Note that no prior anticipated shocks were announced, thus any movement in the states is seen as a shock. Therefore the red rhombus, showing the perceived shocks if beliefs remain and the red square showing the perceived shocks if beliefs switch are overlapping. However the likelihood that shocks originated from either of the regimes is different. The likelihood that backward looking beliefs could have generated this shock is $7.3279 * 10^{-23}$ [41] in other words, it is perceived as almost impossible that backward looking beliefs could have generated such a large shock.

[40] For a birds eye picture on the state covariance please see Figure 12 in Appendix 7.

[41] To compute the likelihood that backward looking beliefs prevail the agents compute $L^{2(2)} = \sqrt{det(Q^2_{1,AL})} \cdot exp^{-\frac{1}{2}\Sigma e_1^{2(2)} Q^2_{1,AL} e_1^{2(2)}}$, where $e_1^{2(2)} = [-3.53.47]$ is the perceived shock in period 1 if beliefs remain backward looking, and $Q^2_{1,AL}$ is the updated covariance matrix of the states in period 1 given backward looking beliefs.

Figure 9: Belief Switching Illustration - Backward to Forward Beliefs in the 1st Period



(Source: Author's illustration)

On the other hand the likelihood of forward looking beliefs could have resulted this state is $4.3484*10^{-17}$ [42], magnitudes larger. This results in a log-likelihood ratio of the two beliefs 26.78 in support of the forward looking beliefs. It is a very strong evidence in favor of forward looking model being a better characterization of the actual economy. As a result backward looking beliefs switch to forward looking.

At the beginning of the second period adaptive learning takes place. This time the economy follows the path of PLM of forward looking beliefs, characterized by a fast decay. Thus the covariance matrix shrinks for both beliefs: neither the IS equation shock, nor additional forward guidance shocks materialize. Both beliefs get updated, forward looking beliefs learn to have a

[42] The likelihood that backward looking beliefs switch to forward looking is given by $L^{1(2)} = \sqrt{det(Q^1_{1,AL})} \cdot exp^{-\frac{1}{2}\Sigma e_1^{1(2)} Q^1_{1,AL} e_1^{1(2)}}$, following the notation in the Appendix on Switching Kálmán filter, where $e_1^{1(2)} = [-3.53.47]$ is the perceived shock in period 1 if beliefs switch forward looking, and $Q^2_{1,AL}$ is the updated covariance matrix of the states in period 1 if beliefs switch forward looking. Note that forward looking beliefs evaluates the likelihood of the state in the MSV of backward looking, i.e. in a subspace of two variables. The DSGE under forward looking beliefs has one additional auxiliary state to capture the anticipated shocks, it is however not observable, thus the DSGE estimation involves an emission matrix H^1 that eliminates this auxiliary state. Therefore the likelihood of the model is only measured in the observable state space of only two variables.

more correlated state distribution, and smaller variance, than before:

$$Q^1_{1,AL} = \begin{bmatrix} 2.7303 & 0.7438 \\ 0.7438 & 0.9607 \end{bmatrix} \rightarrow Q^1_{2,AL} = \begin{bmatrix} 2.1174 & 0.9993 \\ 0.9993 & 0.8663 \end{bmatrix} \tag{34}$$

Similarly backward looking beliefs will adapt to the new ALM, making sates more correlated, and variances slightly less dispersed:

$$Q^2_{1,AL} = \begin{bmatrix} 1.3425 & 0.4519 \\ 0.4519 & 0.9247 \end{bmatrix} \rightarrow Q^2_{2,AL} = \begin{bmatrix} 1.0914 & 0.6662 \\ 0.6662 & 0.7616 \end{bmatrix} \tag{35}$$

Therefore in the second period the agent will have forward looking beliefs, considering the forward guidance fully credible. Now the agents will have to evaluate the probability if it makes sense to switch back to backward looking beliefs. Figure 10 shows the probability distributions of either remaining forward looking or switching[43]. As Figure 8 has shown, the ALM in the second period is overlapping with the forward looking PLM. Therefore, due to being anticipated, forward looking beliefs have made close to zero error[44]. They see the anticipated implementation of the -3.5 interest rate path and perceive it as almost no error on both the interest rate and output. This is shown on Figure 10 with the red rhombus close to zero.

In contrast when the agent considers backward looking beliefs, she recognizes, that from that perspective an additional monetary policy shock would be needed to implement the -3.5. In fact from a backward looking belief the forward guidance ALM is only achievable when the shock on interest rate is -1.8854, while on output gap it is -0.7997 [45]. Figure 10 illustrates this with the red square. Since the size of the perceived shock is smaller, and thus more probable, under forward looking beliefs, the agent decides not to switch and remain forward looking.

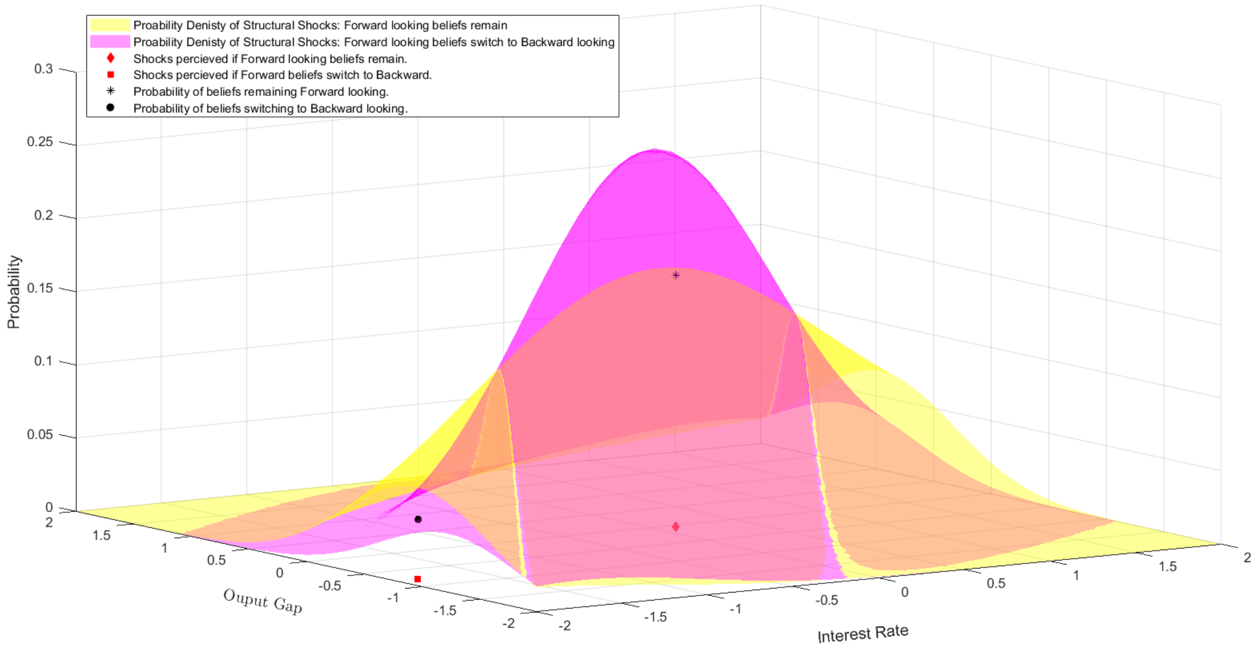
Figure 10 shows this as the black star of remaining lies higher than the black dot of switching.

[43]For another perspective of the same chart please see Figure 13 in Appendix 7.

[44]The errors made are due to the learning of the first period. Since the first period made forward looking beliefs closer to backward looking ones, they now perceive a shock that compensates.

[45]This combination of shock is solely do to a monetary policy shock of a size -1.8854.

Figure 10: Belief Switching Illustration - Forward to Backward Beliefs in the 2nd Period



(Source: Author's illustration)

In this simple example, a forward guidance of large size has made beliefs switch to be forward looking. The central bank built credibility by the size of the accommodation it made. It illustrates, that a large size forward guidance will make agent's beliefs forward looking. However there is another aspect to it, the length of the forward guidance. A longer forward guidance might be seen less credible than a shorter. To understand if this could be the case, the next section will explore how the central bank can determine beliefs, and make backward looking, inattentive beliefs forward looking, and thus a forward guidance credible.

5.6 Central Bank and Delivering Forward Guidance: Gaining Credibility

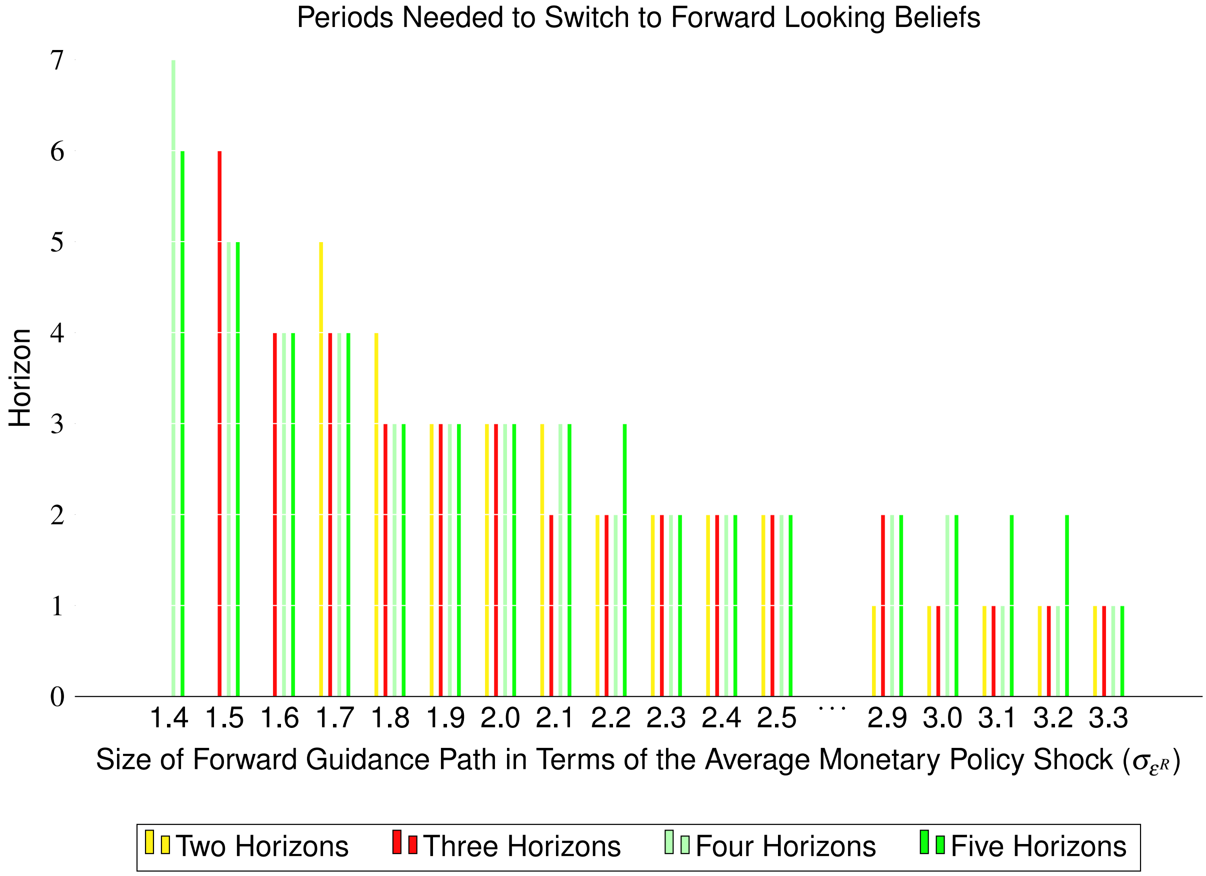
Endogenous belief switching enables to study the credibility of forward guidance from a new perspective. With endogenous belief switching a large shock, compared to expectations can make beliefs switch. However a longer forward guidance might require a stronger commitment and a larger signal as the distance of the backward and forward states grow with the horizon. This sections studies how delivering forward guidance can create credibility. As seen before the size of the current period interest rate cut, and thus the size of the accommodation of the path is of crucial importance when backward looking agents consider to believe the forward guidance. A larger accommodation today, measured in relation to the average size of the monetary policy shock, can make backward looking beliefs switch to forward looking. In other words, an agent facing a very large shock, will reconsider that it might be induced by not only the current monetary policy but by credible announcements of future policy action. Therefore the central bank by announcing a large interest cut today, and moving the whole economy by a lot, can signal to agents facing endogenous belief switching that the forward looking beliefs are better suited to describe the path of the economy.

To study when such belief switching will take place consider that the central bank can decide to implement a forward guidance of any size and length. Agents start out having purely backward looking beliefs, and are assumed to observe the economy as is, without measurement error. Under the initial beliefs the central bank has no credibility when announcing forward guidance, i.e. the puzzle does not emerge. Assume that the central bank would like to build a case for credible forward guidance and thus make beliefs forward looking. This could be desirable as it creates another tool to control the business cycle. Future announcements of the interest rate can control of the economy better. What size of interest rate path is needed in order to make beliefs endogenously switch to being forward looking? Does the credibility of forward guidance depend on its length? How long of a forward guidance, given size is needed to make beliefs forward looking?

Figure 11 answers these questions. It shows the points where belief switching take place as a function of the length and the size of the forward guidance. The economy changes nature when the forward looking beliefs become more likely and beliefs switch[46].

[46]As a reminder, the average size of monetary policy shock in the economy is of size 1. Similarly future forward guidance shocks have the same size. The forward looking beliefs feature forward guidance up to 5 horizon ahead.

Figure 11: Central Bank Credibility - Period of Belief Switching



Notes: The chart shows the period when beliefs switch from backward to forward looking as a function of the size and length of the forward guidance. Backward looking beliefs do not respond to future interest rate changes, while forward looking beliefs are of a model with up-to five periods of forward guidance. The size of the path is measured in terms of the standard deviation of the average monetary policy shock. The yellow bar indicates two periods forward guidance, where the possibility of future forward guidance up to 5 horizons is considered. The red, light green and green show forward guidance of three, four and five horizons respectively.
(Source: Author's calculation)

A shorter forward guidance requires larger signal to switch beliefs. A two periods of forward guidance can make beliefs forward looking if the size of the shock is large enough, i.e. larger than 1.7 times the average size of monetary accommodation. This is displayed on the chart 11 with yellow bar that shows that lower than 1.7 there is no switching. If the size of the monetary policy shock is two times the standard deviation of an average shock beliefs switch 3 periods after the announcement. For a 2.2 times shock it only requires 2 periods, while for a 3.0 times shock beliefs switch in the first period, i.e. similar to the case explained in Section 5.5.

On the other hand a longer forward guidance of five horizon, shown in dark green, requires smaller shocks for the switch to take place. Already a shock of 1.4 times the average monetary policy shock will make case for forward looking beliefs, albeit only after the successful delivery of the promised path. If the size of the shock increases beliefs switch earlier. Looking at the chart one interesting pattern emerges.

Intuitively it is harder to believe a longer forward guidance than a shorter one. Endogenous belief switching delivers this result. The difference in credibility is due to the different size of variance different horizon forward guidance imply. A longer forward guidance commends a larger state dispersion, compared to a shorter, therefore it requires a larger shock to make it credible. For instance it takes a 3.3 times shock to make five horizon forward guidance credible in the first period, whereas for a four horizon forward guidance a shock of a size 3.1 is enough, for three horizon a shock of size 3.0 suffices.

Overall the figure shows that a larger forward guidance shock can make beliefs switch earlier, for any given length of the forward guidance. As we have seen, a larger shock represents a larger signal and thus can make beliefs switch endogenously earlier in the implementation of the path. Given the size of monetary policy accommodation, a longer forward guidance requires a longer delivery of the promised path to gain credibility.

However larger variance per se does not necessarily imply forward looking behaviour. The relevant aspect is not the state dispersion, but the part of the variation that can be attributed to the signal. This ties endogenous belief switching to the signal extraction problem of the regime switching Kálmán filtering. Larger uncertainty can make filtering out the actual shock more difficult. So far we had the assumption, that the model is observed without errors, however in practice it is rarely the case that the model is a perfectly observed, and thus an accurate descriptor of the economy. The next section will illustrate how the signal to noise ratio will influence belief switching.

5.7 Noise and Signal

In the unconventional monetary policy literature the signalling effect of monetary policy is widely understood. With regards to learning this becomes even more important, as learning involves an estimation of the true model. The Switching Kálmán filter naturally gives rise to the signal extraction problem, in what follows I explain the relevance of the observation error and provide an intuition that generates practical interpretation of the noise and policy signal. The relation of noise to signal will play an important role in central bank policy transmission. If the central bank makes a too little action, gives a weak signal, agents confuse it with noise and will not respond to it. However should the signal be large compared to noise, the central bank can make beliefs become forward looking.

When the agent filters the observables, he translates the DSGE solution to the observables space of the MSV using the observation equation:

$$y_t^{MSV} = \mathbf{H^i} y_t^i + \mathbf{u}_t, \tag{36}$$

where $\mathbf{H^i}$ is the so called observation equation matrix that selects the MSV states of the model. Denote $\mathbf{u}$ the measurement error with a covariance matrix that is $\mathbf{U}$[47]. The traditional interpretation of the observation error is measurement noise. It tells that there is an inherent imprecision to the observations that cannot be eliminated, and thus they have to be taken into account when designing a linear filter. I propose an econometric analogy of the measurement noise that is applicable here, and helps the intuition. Under general conditions one can decompose the total variance of any observable into conditional, explained variance and residual variation. Formally:

$$TSS = ESS + RSS, \tag{37}$$

where TSS stands for total variance, while ESS is the explained component of the variance, while RSS is the unexplained, residual variance. Then the fit of the model is the share of total variation explained, $R^2 = 1 - \frac{RSS}{TSS}$. The statistical name of fraction of variance unexplained

Consider that MSV states are the observables as in Equation 36. Think of the term $\mathbf{H^i} y_t^i$ as explanatory variables, that tells how much of the MSV total variation is due to central bank action. Another interpretation is the decomposition of the IRF, i.e. the movement of the states, to the explained and unexplained factors. Explained movements are those compatible with the model, be it a sequence of unanticipated monetary shocks, or a credible forward guidance. The unexplained, residual errors capture all sources of uncertainty that are beyond the scope of the model. As in optimal control, the relative size of variation explained by the model to the noise plays a crucial role for the results. Therefore it is best to think of the observation error, $\mathbf{u}$, in terms of the DSGE models (long-run) variances of the MSV states.

If $\mathbf{U}$ is small compared to the total state variation, then the changes in observables will be driven by the monetary policy, and a large fraction of it can be explained by the model. In other words the DSGE model has fit the data well.

While if observation errors are perceived to be large, it will result in the total variation being driven by the unexplained observation errors. The noise will be high compared to the signal, and thus the DSGE will not have a good fit to the data[48].

6 Endogenous Belief Switching in the Smets Wouters Model with Noise

In the previous section we have introduced endogenous belief switching, built intuition for it and highlighted the role of signal to noise. In the macro-econometric literature, when estimating a DSGE, it is usually assumed that the model is describing the data perfectly, that has no observation error. However when studying the signalling of central bank future actions, this assumption becomes highly restrictive. This section aims to relax this assumption and an-

[47] The measurement error covariance is usually denoted by $\mathbf{R}$. In order not to confuse it with the mean squared error matrix of the beliefs, I changed the notation.

[48] In practice when estimating DSGE models, $\mathbf{U}$ is usually set to a very small number.

swer the question, how endogenously switching beliefs evolve in respone to forward guidance commitment if there is unexplained source of uncertainty in the data?

Recall that endogenous belief switching from backward to forward looking needed a large and accurate signal. But the accuracy of the signal was assured by assumption. However in practice a central bank might make an announcement, that is interpreted with noise and inaccurately. This inaccuracy can be linked to the explained variance, R^2, of a model: depending on the fit of the model the agents might understand the signal differently. A signal can be accurate because it either trumps the noise of the observation, or because it is in an environment where the model is a good explanation of the data.

The charts collected in Table 2 show the IRFs to a forward guidance announcement of the SW07 model with regime switching beliefs as a function of initial share of attentive beliefs, columns, and the relative fit of the model, rows. Forward guidance is a time dependent announcement, i.e. the central bank gives guidance about the short-term interest rate for 1, 2, ..., 6 periods ahead. The initial shares vary on the probability scale from 0% to 50% up until 100%. Where an a priori 0% probability of attentive beliefs tells that the initial beliefs are backward looking. 50% represents a middle ground, where the agent is equally unsure about either being backward or forward looking. Finally, 100% represents fully forward looking belief initialization.

The model fit is measured as fraction of variance explained by the IRF due to the forward guidance[49]. The first row of the table represents the case when the signal to noise ratio very small. This is achieved by setting the observation error covariance matrix large in comparison to the long-run state variance. In this case the DSGE model of either regime has a very low fit, $\simeq 0$. In practice this represents a highly uncertain environment, where the model does not tell anything about the data. If there is little information in the data, agents do not learn anything when filtering the data with the regime switching Kálmán filter. If there is close to no information in the observables, then agents will not be able to update their a priori beliefs and will stick to their initial beliefs. This is shown in the charts as initial beliefs remain and no switching takes place.

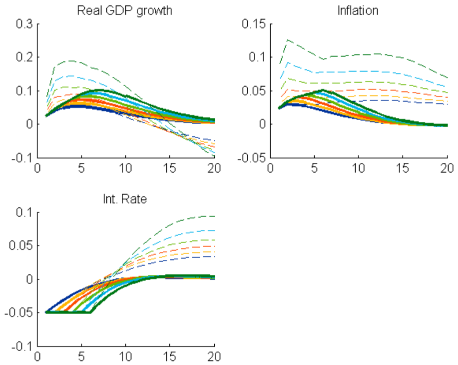
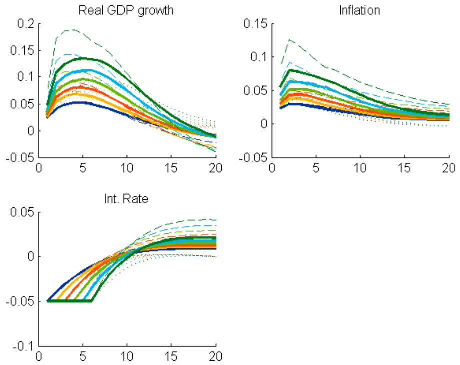
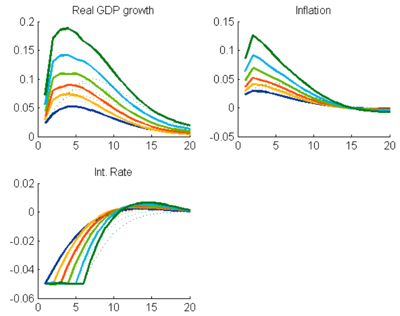
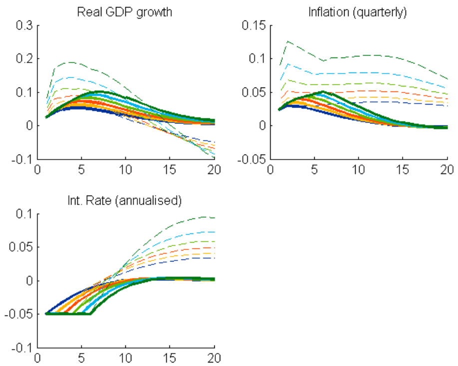
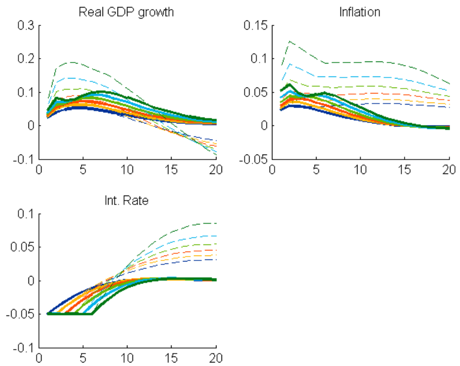
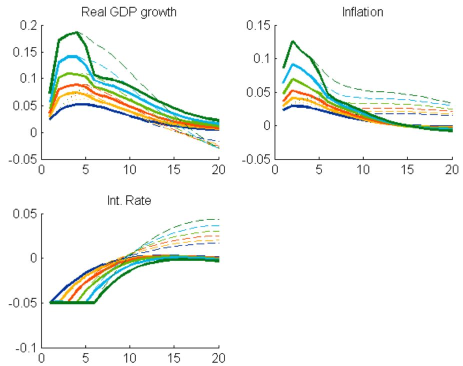
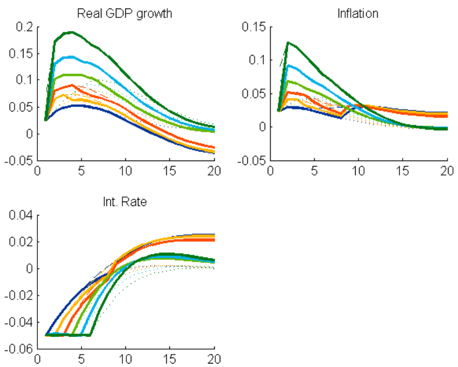
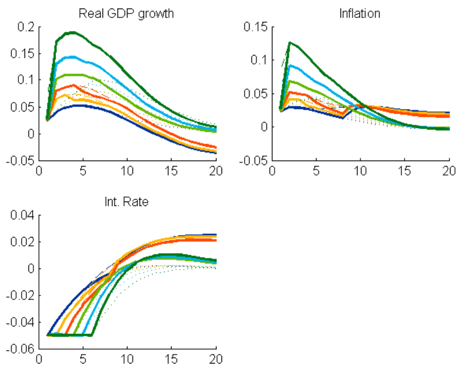
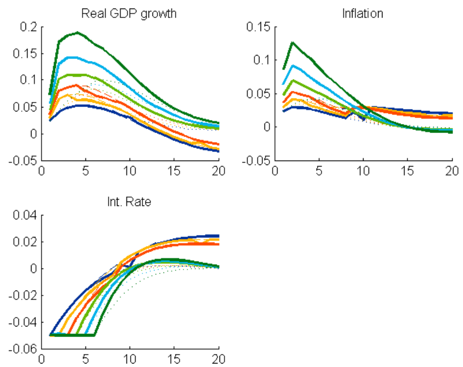
The second row shows the case when the scale of the observation error compared to the sate variance is similar in measure. A model fit of 50 $\%$ can be interpreted that half of the observed variation in the data is attributed to the observation error, i.e. unexplained variance and half of it is considered to be due to the central bank announcing the forward guidance. In this case beliefs will converge on being backward looking. Forward looking beliefs which are subject to the forward guidance puzzle, will not prevail, beliefs will converge to being backward looking. This means that even with initial beliefs with both 50$\%$ and 100$\%$ forward looking the central bank cannot commit to the forward guidance, and will loose credibility. This case can be thought of as the high uncertainty environment, where on one hand the forward puzzle will

[49] Recall that the standard error of the monetary policy shock and the respective forward guidance shocks' standard deviation has been normalized. I follow the practice of the MMB and normalize standard deviation of all shocks entering the Taylor rule to to 1% in annual, 1bp in quarterly terms.

not prevail, on the other hand, the central bank cannot credibly commit to a forward guidance, as expectations converge to the equilibrium where they are backward looking.

The last row shows the already discussed case, where low volatility of the observation error matrix enables switching from backward beliefs to forward beliefs. In an economy with close to zero observation error, the central bank can create credibility, by giving a strong signal about its forward guidance commitment. The case with 0% a priori beliefs illustrates how a short forward guidance will lose credibility over the medium term. See how the yellow and red ALMs switch back to backward looking in the third and forth period respectively. regain it as time evolves. A long forward guidance on the other hand can be seen as most credible, as it does not show belief switching.

Table 2: Regime Switching Beliefs in SW07 with Adaptive Learning

<table>
<tr><th>Model fit</th><th colspan="3">Initial a priori probability of forward looking beliefs $Pr(S_0 = 1)$</th></tr>
<tr><th></th><th>0 %</th><th>50 %</th><th>100 %</th></tr>
<tr><td>$\simeq$ 0 %</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>50 %</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
<tr><td>$\simeq$ 100 %</td>
<td>
</td>
<td>
</td>
<td>
</td></tr>
</table>

Overall, table 2 shows that the credibility of forward guidance depends on the capacity of the central bank to signal its commitment. If the economy is observed with large unexplained observation errors, endogenous belief switching will not take place, as there will be little signal compared to noise.

If there is ample uncertainty, or if the DSGE model communicated by the central bank is an imperfect filter for the economy, i.e. the volatility of the observation error matrix is high forward guidance will not be credible. In this case backward looking beliefs become the equilibrium expectations.

Finally, in a very a low uncertainty environment, when the DSGE is a good description of the economy endogenous belief switching from backward to forward looking expectations can take place. In an economy where there is little unexplained uncertainty compared to the central banks signal, the central bank can gain credibility. Due to the lower observation uncertainty in filtering the central bank signal agents will allocate higher probability to the forward looking expectations.

From a policy perspective, forward guidance can be highly effective in low uncertainty environments, while can become ineffective in high uncertainty economies. Due to endogenous belief switching expectations may become backward looking if there is too much uncertainty compared to the signal the central bank makes. In a liquidity trap, where there is little room to move the short-term rate down, forward guidance may harm central bank credibility, or even create a credibility trap, where no matter the central bank action, expectations become adaptive. While, if the central bank can give a strong signal, agents can learn to be forward looking, this case calls for strong comprehensive measures supporting forward guidance.

The policy relevance of these results is wide reaching. The framework tells that a central bank can be partially the reason why the liquidity trap at the zero lower bound becomes so hard to defeat. If a central bank is conservative and fails to provide strong guidance, it might find itself having drifted so close to the ZLB, that switching backward beliefs to forward looking ones becomes impossible to achieve, as there might be not enough room to give a signal large enough to invoke a switch of beliefs. In such an environment the central bank can only make forward guidance credible by activating other measures as well. Should it fail to do so, it might reinforce expectations to stay adaptive and become responsible for the inefficient impact of forward guidance announcements and the prevailing liquidity trap.

Japan could be seen as a country that tried but failed to escape the ZLB. Endogenous belief switching explains, why monetary policy in Japan failed: it shows that due to perceived inaction why the credibility of forward guidance waned, and agents became endogenously backward looking, as documented by Hogen and Okuma (2018). But it is not only the fault of the Bank of Japan (BoJ), another reason could have been why the BoJ failed to provide a strong enough signal, that it was impossible to do so. It would have had to overcome the heightened uncertainty Japanese agents perceived as Japan entered a recession. Endogenous belief switching predicts, that central bank commitment might be harder to believe if there is heightened uncertainty.

A timely solution could have been a rethinking of the communication, such that the BoJ's explained its model to the public. This could have made the central bank's model fit the economy better, and thus reduce the residual, noise, and uncertainty. Another option would have been to embrace a wide range of unconventional measures and abandon conservative approach to surprise agents with unprecedented monetary policy accommodation, and thus turn them forward looking. In a sense, the yield curve control of the BoJ enacts a stronger signalling device to anchor expectations more firmly to the forward guidance commitment, and could be seen as the right step to regaining central bank control of expectations.

7 Conclusion

Standard medium-scale DSGE models overstate the effect of forward guidance on the macroeconomy, a contradiction called the forward guidance puzzle. In search for an alternative to rational expectations, that can account for the missing effectiveness of pre-announced monetary policy actions, researchers started to re-think the role played by expectations of future variables. Multiple solutions to the forward guidance puzzle were proposed, first by distinguishing how forward guidance works, either being a time dependent or state dependent forward guidance, economists granularized the picture. It has been shown that even in standard models, non-linearities and precautionary savings can remedy the forward guidance puzzle. Second, by pointing out that the efficacy of forward guidance originates from the modelling of staggered prices, sticky information was shown to solve the puzzle. Finally, heterogeneous beliefs, where central bank credibility was the division of beliefs, was shown to be a solution as well. This paper contributes to the latter stream of literature introducing endogenous belief switching. I build endogenous belief switching on constant gain adaptive learning and allow belief switching on the perceived share of agents who are forward looking. By making the share of attentive beliefs endogenous, and enabling beliefs to switch dynamically between backward and forward looking, I highlight the role of central banks for expectation determination, and dynamic evolution of their credibility. If expectations can be influenced by central bank action, so can the effectiveness of forward guidance.

First, my paper shows that constant gain adaptive learning can overcome the forward guidance puzzle if agents are adaptive and backward looking. I also document that the forward guidance puzzle still emerges, if expectations are adaptive learning forward looking. To enable the effectiveness of forward guidance and to study the dynamic evolution of central bank credibility, I propose a novel form of expectation formation: endogenous belief switching, that is a combination of constant gain adaptive learning and regime switching Kálmán filtering.

Focusing on the efficacy of forward guidance the paper shows that not only the size, but the promised length of forward guidance also affects how beliefs respond. Under endogenous belief switching agents learn the monetary policy implementation and optimally decide to form expectations about future macroeconomic variables: either by responding to pre-announced future policy rate changes, having adaptive forward looking beliefs, or by focusing only on current conditions and being backward looking.

I show that a forward guidance of a large size and long horizon is the most likely to make expectations switch to forward looking, albeit requiring a persistent delivery of promises for this to happen. On the other hand, central banks need to give larger signals in support for a short horizon forward guidance, as it is more likely to become falsely interpreted as a sequence of unanticipated shocks, and thus not build central bank credibility.

The paper illustrates using simple examples how endogenous belief switching occurs. First in a model of only one dimension, where switching takes place only upon announcement, then in a small scale New Keynesian model with a two dimensional state space.

I argue that endogenous belief switching resembles rational inattention, as it relies on optimal filtering of information, highlighting the role of uncertainty for expectation determination. Based on the Smets Wouters model I study how uncertainty impacts the credibility of forward guidance. Depending on the noise to signal ratio, I find that credibility can be built only if the central bank gives a signal that is well understood, either because it is so large, that agents can easily filter it, or because it is provided in an environment where there is little uncertainty.

Endogenous belief switching opens new possibilities to analyze unconventional monetary policy and its role for expectations, and enables research to rethink the conduct of monetary policy from a novel perspective.

References

Andrade, P., G. Gaballo, E. Mengus, and B. Mojon (2019, July). Forward guidance and heterogeneous beliefs. *American Economic Journal: Macroeconomics 11*(3), 1–29.

Beechey, M. J., B. K. Johannsen, and A. T. Levin (2011). Are long-run inflation expectations anchored more firmly in the euro area than in the united states? *American Economic Journal: Macroeconomics 3*(2), 104–29.

Bhattarai, S., G. B. Eggertsson, and B. Gafarov (2015, July). Time Consistency and the Duration of Government Debt: A Signalling Theory of Quantitative Easing. NBER Working Papers 21336, National Bureau of Economic Research, Inc.

Bianchi, F. (2012). Regime switches, agents' beliefs, and post-world war ii us macroeconomic dynamics. *Review of Economic studies 80*(2), 463–490.

Binder, M. and M. H. Pesaran (1995). Multivariate rational expectations models and macroeconometric modelling: A review and some new results. *Handbook of applied econometrics 1*, 139–187.

Binning, A. and J. Maih (2017). Modelling occasionally binding constraints using regime-switching.

Borağan Aruoba, S., P. Cuba-Borda, and F. Schorfheide (2017). Macroeconomic dynamics near the zlb: A tale of two countries. *The Review of Economic Studies 85*(1), 87–118.

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

Bullard, J. (2010). Seven faces of" the peril". *Federal Reserve Bank of St. Louis Review 92*(September/October 2010).

Bullard, J. and J. Duffy (2004). Learning and structural change in macroeconomic data. *FRB of St. Louis Working Paper No*.

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

Caballero, R. and E. Farhi (2017, 02). The Safety Trap. *The Review of Economic Studies 85*(1), 223–274.

Carlstrom, C. T., T. S. Fuerst, and M. Paustian (2015). Inflation and output in new keynesian models with a transient interest rate peg. *Journal of Monetary Economics 76*, 230 – 243.

Carvalho, C., S. Eusepi, E. Moench, and B. Preston (2019, February). Anchored inflation expectations. Working paper, CEPR.

Chang, Y., J. Maih, and F. Tan (2018). State space models with endogenous regime switching.

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

Christiano, L. J. (2014). The extended path method. *Gerzensee Lecture*.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (1999). Chapter 2 monetary policy shocks: What have we learned and to what end? *1*, 65 – 148.

Christiano, L. J., R. Motto, and M. Rostagno (2014). Risk shocks. *American Economic Review 104*(1), 27–65.

Chung, H., E. Herbst, and M. T. Kiley (2015). Effective monetary policy strategies in new keynesian models: A reexamination. *NBER Macroeconomics Annual 29*(1), 289–344.

Clarida, R., J. Gali, and M. Gertler (1999, December). The science of monetary policy: A new keynesian perspective. *Journal of Economic Literature 37*(4), 1661–1707.

Cole, S. (2015). Learning and the effectiveness of central bank forward guidance. Technical report, University Library of Munich, Germany.

Del Negro, M., M. P. Giannoni, and C. Patterson (2012). The forward guidance puzzle. *FRB of New York Staff Report* (574).

Dovern, J. and G. Kenny (2017). Anchoring inflation expectations in unconventional times: Micro evidence for the euro area. Forthcoming in International Journal of Central Banking.

Draghi, M. (2014, Nov). Frankfurt european banking congress. Speech by Mario Draghi at the Frankfurt European Banking Congress, Frankfurt am Main,21 November 2014.

Eggertsson, G. (2006). The deflation bias and committing to being irresponsible. *Journal of Money, Credit and Banking 38*(2), 283–321.

Eggertsson, G. B. and M. Woodford (2003). The zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity 2003*(1), 139–211.

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

Ehrmann, M., G. Gaballo, P. Hoffmann, and G. Strasser (2019). How to signal the future path of interest rates? The international evidence on forward guidance. *Research Bulletin 61*.

Eusepi, S. and B. Preston (2010). Central bank communication and expectations stabilization. *American Economic Journal: Macroeconomics 2*(3), 235–71.

Eusepi, S. and B. Preston (2011). Expectations, learning, and business cycle fluctuations. *The American Economic Review 101*(6), 2844–2872.

Evans, G. W. and S. Honkapohja (2012). *Learning and expectations in macroeconomics*. Princeton University Press.

Fair, R. C. and J. B. Taylor (1983). Solution and maximum likelihood estimation of dynamic nonlinear rational expectations models. *Econometrica 51*(4), 1169–1185.

Farmer, R. E., D. F. Waggoner, and T. Zha (2009). Understanding markov-switching rational expectations models. *Journal of Economic theory 144*(5), 1849–1867.

Farmer, R. E., D. F. Waggoner, and T. Zha (2011). Minimal state variable solutions to markov-switching rational expectations models. *Journal of Economic Dynamics and Control 35*(12), 2150 – 2166. Frontiers in Structural Macroeconomic Modeling.

Franchi, M. and P. Paruolo (2015). Minimality of state space solutions of dsge models and existence conditions for their var representation. *Computational Economics 46*(4), 613–626.

Gertler, M. (2017). Rethinking the power of forward guidance: Lessons from japan. Technical report, National Bureau of Economic Research.

Giacomini, R. (2013). The relationship between dsge and var models. In *VAR Models in Macroeconomics–New Developments and Applications: Essays in Honor of Christopher A. Sims*, pp. 1–25. Emerald Group Publishing Limited.

Goy, G., C. H. Hommes, and K. Mavromatis (2018). Forward guidance and the role of central bank credibility under heterogeneous beliefs.

Hogen, Y. and R. Okuma (2018). The anchoring of inflation expectations in japan: A learning-approach perspective. *Bank of Japan Working Paper Series 18-E-8*, 43.

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

Juillard, M. (2001, April). DYNARE: A program for the simulation of rational expectation models. Computing in Economics and Finance 2001 213, Society for Computational Economics.

Kydland, F. and E. Prescott (1982). Time to build and aggregate fluctuations. *Econometrica 50*(6), 1345–70.

Lansing, K. J. (2018). Endogenous regime switching near the zero lower bound. Federal Reserve Bank of San Francisco.

Laséen, S. and L. E. Svensson (2011). Anticipated alternative instrument-rate paths in policy simulations. *CEPR Discussion Paper No. DP8176*.

Lindé, J. and M. Trabandt (2018). Should we use linearized models to calculate fiscal multipliers? *Journal of Applied Econometrics 33*(7), 937–965.

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

Mackowiak, B. and M. Wiederholt (2009). Optimal sticky prices under rational inattention. *American Economic Review 99*(3), 769–803.

Maih, J. (2015). Efficient perturbation methods for solving regime-switching dsge models. *Norges Bank Working Paper 1— 2015*.

Mankiw, N. G. and R. Reis (2002). Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics 117*(4), 1295–1328.

Mankiw, N. G. and R. Reis (2010). Imperfect information and aggregate supply. In *Handbook of monetary economics*, Volume 3, pp. 183–229. Elsevier.

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

Maćkowiak, B. and M. Wiederholt (2015, 08). Business Cycle Dynamics under Rational Inattention. *The Review of Economic Studies 82*(4), 1502–1532.

McKay, A., E. Nakamura, and J. Steinsson (2016). The power of forward guidance revisited. *American Economic Review 106*(10), 3133–58.

Mitra, K., G. W. Evans, and S. Honkapohja (2012). Fiscal policy and learning. *Bank of Finland Research Discussion Paper* (5).

Molavi, P. (2019). Macroeconomics with learning and misspecification: A general theory and applications. MIT, Job Market Paper, available at: `https://economics.mit.edu/files/16326`.

Murphy, K. P. (1998). Switching kalman filters. *DEC/Compaq Cambridge Research Labs Tech. Report 98*(10).

Preston, B. (2005). Learning about monetary policy rules when long-horizon expectations matter. *International Journal of Central Banking 1*(2).

Ravenna, F. and C. E. Walsh (2006). Optimal monetary policy with the cost channel. *Journal of Monetary Economics 53*(2), 199 – 216.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica 48*(1), 1–48.

Sims, C. A. (2002). Solving linear rational expectations models. *Computational economics 20*(1), 1–20.

Sims, C. A. (2006). Rational inattention: Beyond the linear-quadratic case. *American Economic Review 96*(2), 158–163.

Sims, C. A., D. F. Waggoner, and T. Zha (2008). Methods for inference in large multiple-equation markov-switching models. *Journal of Econometrics 146*(2), 255–274.

Slobodyan, S. and R. Wouters (2012). Learning in an estimated medium-scale dsge model. *Journal of Economic Dynamics and control 36*(1), 26–46.

Smets, F. and R. Wouters (2007). Shocks and frictions in US business cycles: a Bayesian DSGE approach. *American Economic Review 97*(3), 586–606.

Van der Cruijsen, C. and M. Demertzis (2011). How anchored are inflation expectations in emu countries? *Economic Modelling 28*(1-2), 281–298.

Waggoner, D. F. and T. Zha (1999). Conditional forecasts in dynamic multivariate models. *Review of Economics and Statistics 81*(4), 639–651.

Wieland, V. (2008). Learning, endogenous indexation and disinflation in the new-keynesian model. In *Monetary Policy under Uncertainty and Learning*, Volume 13, Chapter 11, pp. 413–450.

Wieland, V., T. Cwik, G. J. Müller, S. Schmidt, and M. Wolters (2012). A new comparative approach to macroeconomic modeling and policy analysis. *Journal of Economic Behavior & Organization 83*(3), 523–541.

Woodford, M. (2012). Methods of policy accommodation at the interest-rate lower bound. The Changing Policy Landscape: 2012 Jackson Hole symposium. Federal Reserve Bank of Kansas City.

Appendices

Switching Kálmán filter

In what follows, I will discuss the main steps needed to derive updating of the a priori state probabilities. Inference is filtering only - the probability distribution of a switch happening at time t depends only on past data, i.e. $1:t$. For a full discussion of the switching Kálmán filter please consult Murphy (1998).

Let us define a notation: y is the states, while y^{MSV} the observations. Denote the regimes with i.

Consider the state representation of the DSGE with observable noise on the MSV states:

$$y_t^i = \mathbf{F^i} y_{t-1}^i + \mathbf{w^i}_t, \tag{38}$$

$$y_t^{MSV} = \mathbf{H^i} y_t^i + \mathbf{u}_t \tag{39}$$

Where y_t^i is the state vector given belief i, either forward of backward looking, $\mathbf{w^i}_t$ is the exogenous state disturbance. Denote its covariance matrix with Q_t^i. $\mathbf{H^i}$ is the emission matrix that selects the MSV states of the model. Furthermore $\mathbf{u}$ is the measurement noise with a covariance matrix that is usually denoted by $\mathbf{R}$. Not to confuse the mean squared error matrix under the beliefs, and I will express the measurement error covariance matrix with $\mathbf{U}$.

Furthermore one needs to specify an exogenous transition probability from state matrix Z $i \rightarrowtail j$. In the example I assume a highly persistent exogenous state transition probability of the form:

$$Z = \begin{pmatrix} 0.9999 & 0.0001 \\ 0.0001 & 0.9999 \end{pmatrix} \tag{40}$$

Introducing the notation:

$$y_{t|t}^{i(j)} = E\left[y_t | y_{1:t}^{MSV}, S_{t-1} = i, S_t = j\right] \tag{41}$$

Notice that the superscript in the brackets is the switching of regime from i to j in period t. The Equation 41 tells, what the value of the full state is given the (full) history of the MSV states, if it switches from regime i to j.

The switching Kálmán filter pass will be the following: First, the state distribution is inherited. It is all possible combination of states $y_{t|t-1}^{i(j)}$ and their respective covariance matrix based on information from $t-1$. The indices of states are looped over before progressing to the next step of the filter. The notation below exemplifies the filter as conditional on being in i switching to the next regime j. If the index is the same, then no regime switch takes place, if it is different it represents switching. As in the Kálmán filter the first step is called the prediction:

$$y^{i(j)}_{t|t-1} = \mathbf{F^j} y^i_{t-1}, \tag{42}$$
$$Q^{i(j)}_{t|t-1} = \mathbf{F^j} Q^i_{t-1} (\mathbf{F^j})' + Q^j_{t-1}. \tag{43}$$

Then, we compute the Kálmán gain given switching:

$$K^{i(j)} = Q^{i(j)}_{t|t-1} (\mathbf{H^j})' (\mathbf{H^j} Q^{i(j)}_{t|t-1} (\mathbf{H^j})' + \mathbf{U}) \tag{44}$$

Using the gain update the one can generate the nowcast, i.e. posterior of the state and state covariance matrix given information t:

$$y^{i(j)}_{t|t} = y^{i(j)}_{t|t-1} + K^{i(j)} (y^{MSV}_t - \mathbf{H^j} y^{i(j)}_{t|t-1}); \tag{45}$$
$$Q^{i(j)}_{t|t} = (I - K^{i(j)} \mathbf{H^j}) Q^{i(j)}_{t|t-1}; \tag{46}$$

With the nowcast, the likelihood of data given $S_t = j$ and $S_{t-1} = i$ can be computed that is the object of my application of the filter:

$$e^{i(j)}_t = y^{MSV}_t - \mathbf{H^j} y^{i(j)}_{t|t-1}, \tag{47}$$
$$L^{i(j)}_t = \sqrt{det(\mathbf{H^j} Q^{i(j)}_{t|t-1} \mathbf{H^j}' + \mathbf{U})} \cdot exp^{-\frac{1}{2}\Sigma\left(e^{i(j)}_t \left(\mathbf{H^j} Q^{i(j)}_{t|t-1} \mathbf{H^j}' + \mathbf{U}\right) e^{i(j)}_t\right)} \tag{48}$$

Finally one can update the a priori probabilities $Pr(S_t = i|t-1)$ using the following algorithm for all $i, j \in \{1,2\}$ and all t:

$$Pr(S_t = j|t, S_{t-1} = i) = \frac{L^{i(j)}_t Z(i,j) Pr(S_t = i|t-1)}{\sum_{i\in\{1,2\}} \sum_{j\in\{1,2\}} L^{i(j)}_t Z(i,j) Pr(S_t = i|t-1)} \tag{49}$$
$$Pr(S_t = j|t) = \sum_{i\in\{1,2\}} Pr(S_t = j|t, S_{t-1} = i) \tag{50}$$

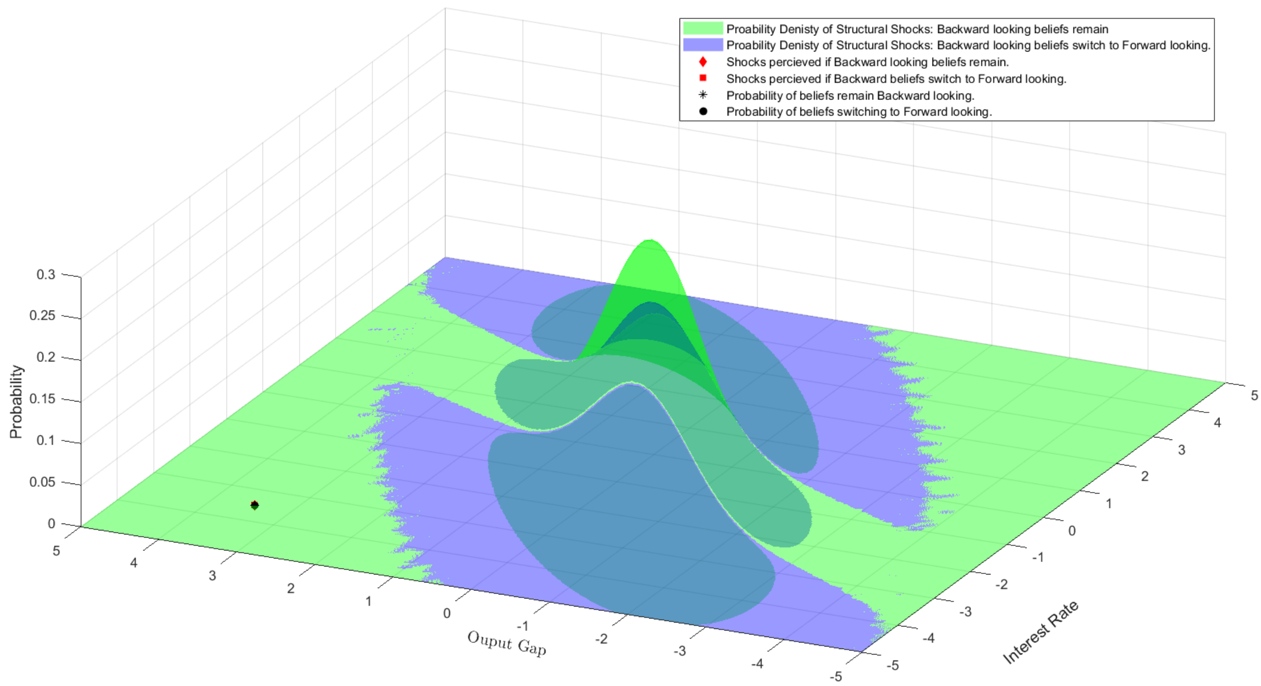
The final collapsing step assures that states are merged from across the regimes with the weighted probabilities:

$$y^j_{t|t} = \sum_{i\in\{1,2\}} y^{i(j)}_{t|t} \cdot \frac{Pr(S_t = j|t, S_{t-1} = i)}{Pr(S_t = j|t)}, \tag{51}$$
$$Q^j_{t|t} = \frac{Pr(S_t = j|t, S_{t-1} = i)}{Pr(S_t = j|t)} \left(Q^j_{t-1} + \left(y^{i(j)}_{t|t} - y^j_{t|t}\right)\left(y^{i(j)}_{t|t} - y^j_{t|t}\right)'\right). \tag{52}$$

There are two points needed to recognise. First, is the Markov assumption on the exogenous state state transition matrix Z, and the role it plays. It scales the likelihood and regulates switching. This is important as switching Kálmán filters have been documented to show instability of regimes and display way too many jumps. However over-regularizing the switching, and imposing an identity matrix, eliminates changing of regimes entirely. Therefore having a reasonable yet persistent exogenous regime dynamics is preferred. This is implemented with the calibration of entries in Z. Second, is the role of the observation error covariance matrix, $\mathbf{U}$. It is added to the (observation space compressed) state variance matrix, when computing the likelihood. That is the variation of the data is either driven by the model or the observation. It's relative size and possible correlation structure compared to that of the state variances is key in determining switching. I assume only the scale, expressed in terms of the average trace of the state covariance matrix, varies, the correlation structure does not.

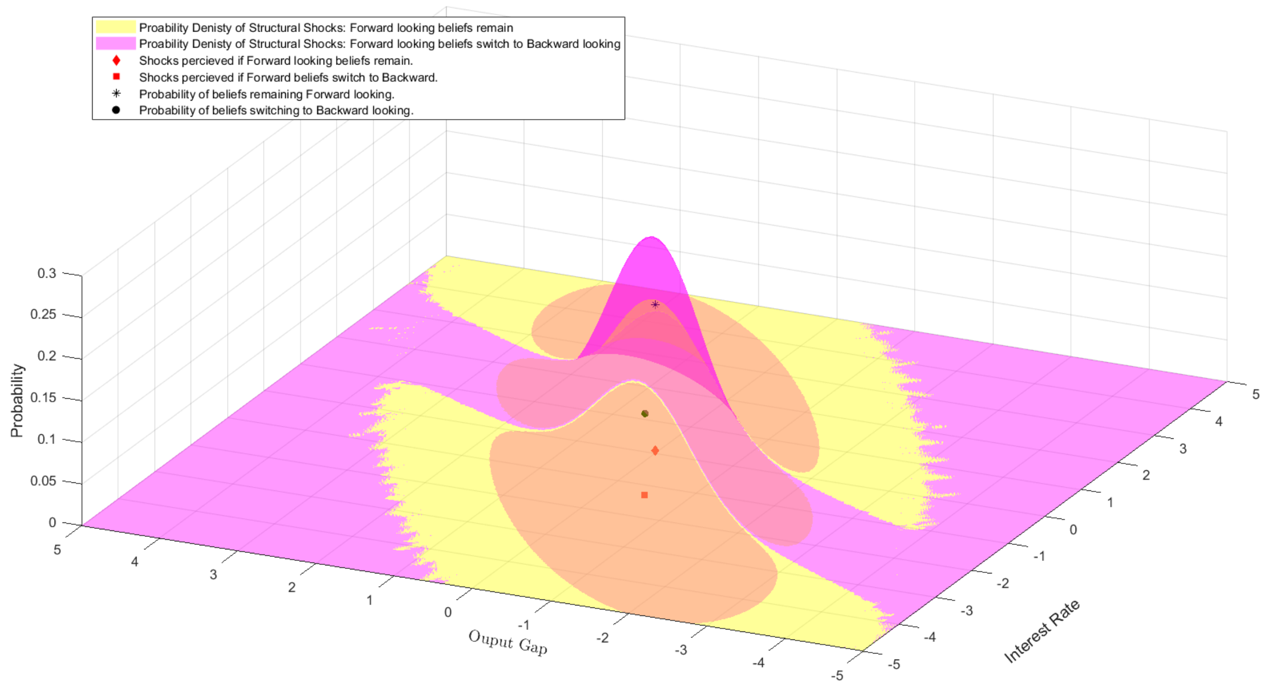
Additional Figures

Figure 12: Belief Switching Illustration - Bird's eye view - 1$^{\text{st}}$ Period



(Source: Author's calculation)

Figure 13: Belief Switching Illustration - Bird's eye view - 2nd Period



(Source: Author's calculation)

Dynamic equations

$$i_t = r_t 4 \quad (53)$$

$$\pi^{ann}{}_t = \pi^{ann}{}_t \quad (54)$$

$$\pi^q{}_t = 4\,\pi_t \quad (55)$$

$$y^{gap}{}_t = y_t - y^{flex}{}_t \quad (56)$$

$$y_t = y_t \quad (57)$$

$$g_t = \varepsilon^{gov}{}_t \quad (58)$$

$$\pi_{t-1_t} = \pi^q{}_{t-1} \quad (59)$$

$$\pi_{t-2_t} = \pi_{t-1_{t-1}} \quad (60)$$

$$\pi_{t-3_t} = \pi_{t-1_t} + \pi_{t-2_t} \quad (61)$$

$$i_t = \rho\, i_{t-1} + \theta_\pi \pi^q{}_t \; + \theta_x y_t - \sigma_r\, \varepsilon^R{}_t \tag{62}$$

$$g_t = coffispol\, \pi_{tt} \tag{63}$$

$$\varepsilon_{at} = \alpha\, r^{k,flex}{}_t + (1-\alpha)\, w^{flex}{}_t \tag{64}$$

$$z^{flex}{}_t = r^{k,flex}{}_t \,\frac{1}{\frac{\psi}{1-\psi}} \tag{65}$$

$$r^{k,flex}{}_t = w^{flex}{}_t + l^{flex}{}_t - k^{s,flex}{}_t \tag{66}$$

$$k^{s,flex}{}_t = z^{flex}{}_t + k^{flex}{}_{t-1} \tag{67}$$

$$i^{flex}{}_t = \frac{1}{1+\bar{\beta}\,\gamma_c}\left(i^{flex}{}_{t-1} + \bar{\beta}\,\gamma_c\, i^{flex}{}_{t+1} + \frac{1}{\gamma_c{}^2\,\varphi}\, q^{flex}{}_t\right) + \varepsilon^i{}_t \tag{68}$$

$$q^{flex}{}_t = \left(-r^{flex}{}_t\right) + c_2 * \varepsilon^b_{t\;t}\,\frac{1}{\frac{1-\frac{\lambda}{\gamma_c}}{\sigma_c\left(1+\frac{\lambda}{\gamma_c}\right)}} + \frac{\bar{r}^k}{\bar{r}^k+1-\delta}\, r^{k,flex}{}_{t+1} + \frac{1-\delta}{\bar{r}^k+1-\delta}\, q^{flex}{}_{t+1} \tag{69}$$

$$\begin{aligned} c^{flex}{}_t = {} & c_2 * \varepsilon^b_{t\;t} + \frac{\frac{\lambda}{\gamma_c}}{1+\frac{\lambda}{\gamma_c}}\, c^{flex}{}_{t-1} + \frac{1}{1+\frac{\lambda}{\gamma_c}}\, c^{flex}{}_{t+1} \\ & + \frac{(\sigma_c - 1)\; 1/\phi_w * (1-\alpha)/\alpha * \bar{r}^k * k_{ss}/y_{ss}}{\sigma_c\left(1+\frac{\lambda}{\gamma_c}\right)}\left(l^{flex}{}_t - l^{flex}{}_{t+1}\right) - r^{flex}{}_t\,\frac{1-\frac{\lambda}{\gamma_c}}{\sigma_c\left(1+\frac{\lambda}{\gamma_c}\right)} \end{aligned} \tag{70}$$

$$y^{flex}{}_t = c^{flex}{}_t\, c_{ss}/y_{ss} + i^{flex}{}_t\, i_{ss}/y_{ss} + \varepsilon^g{}_t + z^{flex}{}_t\, \bar{r}^k * k_{ss}/y_{ss} \tag{71}$$

$$y^{flex}{}_t = \phi_p\left(\varepsilon_{at} + \alpha\, k^{s,flex}{}_t + (1-\alpha)\, l^{flex}{}_t\right) \tag{72}$$

$$w^{flex}{}_t = l^{flex}{}_t\, \sigma_l + c^{flex}{}_t\,\frac{1}{1-\frac{\lambda}{\gamma_c}} - c^{flex}{}_{t-1}\,\frac{\frac{\lambda}{\gamma_c}}{1-\frac{\lambda}{\gamma_c}} \tag{73}$$

$$k^{flex}{}_t = k^{flex}{}_{t-1}\,(1-\bar{i}/k) + i^{flex}{}_t\,\bar{i}/k + \varepsilon^i{}_t\,\gamma_c{}^2\,\varphi\,\bar{i}/k \tag{74}$$

$$\mu_{p_t} = \alpha\, r^k{}_t + (1-\alpha)\, w_t - \varepsilon_{at} \tag{75}$$

$$z_t = \frac{1}{\frac{\psi}{1-\psi}}\, r^k{}_t \tag{76}$$

$$ {r^{k}}_{t} = w_{t}+l_{t}-{k^{s}}_{t} \tag{77} $$

$$ {k^{s}}_{t} = z_{t}+k_{t-1} \tag{78} $$

$$ i_{t} = {\varepsilon^{i}}_{t}+\frac{1}{1+\bar{\beta}\,\gamma_{c}}\,\left(i_{t-1}+\bar{\beta}\,\gamma_{c}\,i_{t+1}+\frac{1}{{\gamma_{c}}^{2}\,\varphi}\,q_{t}\right) \tag{79} $$

$$ q_{t} = c_{2}*{\varepsilon_{t}^{b}}_{t}\,\frac{1}{\frac{1-\frac{\lambda}{\gamma_{c}}}{\sigma_{c}\,\left(1+\frac{\lambda}{\gamma_{c}}\right)}}+\left(-r_{t}\right)+\pi_{t+1}+\frac{\bar{r}^{k}}{\bar{r}^{k}+1-\delta}\,{r^{k}}_{t+1}+\frac{1-\delta}{\bar{r}^{k}+1-\delta}\,q_{t+1} \tag{80} $$

$$ \begin{aligned} c_{t} = {} & c_{2}*{\varepsilon_{t}^{b}}_{t}+\frac{\frac{\lambda}{\gamma_{c}}}{1+\frac{\lambda}{\gamma_{c}}}\,c_{t-1}+\frac{1}{1+\frac{\lambda}{\gamma_{c}}}\,c_{t+1} \\ & +\frac{\left(\sigma_{c}-1\right)\,1/\phi_{w}*\left(1-\alpha\right)/\alpha*\bar{r}^{k}*k_{ss}/y_{ss}}{\sigma_{c}\,\left(1+\frac{\lambda}{\gamma_{c}}\right)}\,\left(l_{t}-l_{t+1}\right)-\frac{1-\frac{\lambda}{\gamma_{c}}}{\sigma_{c}\,\left(1+\frac{\lambda}{\gamma_{c}}\right)}\,\left(r_{t}-\pi_{t+1}\right) \end{aligned} \tag{81} $$

$$ y_{t} = {\varepsilon^{g}}_{t}+c_{ss}/y_{ss}\,c_{t}+i_{ss}/y_{ss}\,i_{t}+\bar{r}^{k}*k_{ss}/y_{ss}\,z_{t} \tag{82} $$

$$ y_{t} = \phi_{p}\,\left(\varepsilon_{at}+\alpha\,{k^{s}}_{t}+\left(1-\alpha\right)\,l_{t}\right) \tag{83} $$

$$ \pi_{t} = \frac{1}{1+\bar{\beta}\,\gamma_{c}\,\iota_{p}}\,\left(\bar{\beta}\,\gamma_{c}\,\pi_{t+1}+\iota_{p}\,\pi_{t-1}+{\mu_{p}}_{t}\,\frac{\frac{\left(1-\xi_{p}\right)\,\left(1-\bar{\beta}\,\gamma_{c}\,\xi_{p}\right)}{\xi_{p}}}{1+\left(\phi_{p}-1\right)\,\varepsilon_{p}}\right)+{\varepsilon^{p}}_{t} \tag{84} $$

$$ \begin{aligned} w_{t} = {} & \frac{1}{1+\bar{\beta}\,\gamma_{c}}\,w_{t-1}+\frac{\bar{\beta}\,\gamma_{c}}{1+\bar{\beta}\,\gamma_{c}}\,w_{t+1}+\pi_{t-1}\,\frac{\iota_{w}}{1+\bar{\beta}\,\gamma_{c}}-\pi_{t}\,\frac{1+\bar{\beta}\,\gamma_{c}\,\iota_{w}}{1+\bar{\beta}\,\gamma_{c}}+\pi_{t+1}\,\frac{\bar{\beta}\,\gamma_{c}}{1+\bar{\beta}\,\gamma_{c}} \\ & +\frac{\left(1-\xi_{w}\right)\,\left(1-\bar{\beta}\,\gamma_{c}\,\xi_{w}\right)}{\left(1+\bar{\beta}\,\gamma_{c}\right)\,\xi_{w}}\,\frac{1}{1+\left(\phi_{w}-1\right)\,\varepsilon_{w}}\,\left(\sigma_{l}\,l_{t}+\frac{1}{1-\frac{\lambda}{\gamma_{c}}}\,c_{t}-\frac{\frac{\lambda}{\gamma_{c}}}{1-\frac{\lambda}{\gamma_{c}}}\,c_{t-1}-w_{t}\right)+{\varepsilon^{w}}_{t} \end{aligned} \tag{85} $$

$$ \varepsilon_{at} = \rho_{a}\,\varepsilon_{at-1}+{\varepsilon^{a}}_{t} \tag{86} $$

$$ c_{2}*{\varepsilon_{t}^{b}}_{t} = \rho_{b}\,c_{2}*{\varepsilon_{t}^{b}}_{t-1}+{\varepsilon^{b}}_{t} \tag{87} $$

$$ {\varepsilon^{g}}_{t} = {\varepsilon^{gov}}_{t}+\rho_{g}\,{\varepsilon^{g}}_{t-1}+{\varepsilon^{a}}_{t}\,\rho_{ga} \tag{88} $$

$$ {\varepsilon^{i}}_{t} = \rho_{i}\,{\varepsilon^{i}}_{t-1}+{\varepsilon^{R}}_{t} \tag{89} $$

$$ {\varepsilon^{r}}_{t} = \rho_{r}\,{\varepsilon^{r}}_{t-1}+{\varepsilon^{m}}_{t} \tag{90} $$

$$ {\varepsilon^{p}}_{t} = \rho_{p}\,{\varepsilon^{p}}_{t-1}+{\varepsilon^{p,aux}}_{t}-\mu_{p}\,{\varepsilon^{p,aux}}_{t-1} \tag{91} $$

$$ {\varepsilon^{p,aux}}_{t} = {\varepsilon^{p}}_{t} \tag{92} $$

$$ {\varepsilon^{w}}_{t} = \rho_{w}\,{\varepsilon^{w}}_{t-1}+{\varepsilon^{w,aux}}_{t}-\mu_{w}\,{\varepsilon^{w,aux}}_{t-1} \tag{93} $$

$$ {\varepsilon^{w,aux}}_{t} = {\varepsilon^{w}}_{t} \tag{94} $$

$$ k_{t} = \left(1-\bar{i}/k\right)\,k_{t-1}+\bar{i}/k\,i_{t}+{\varepsilon^{i}}_{t}\,\varphi\,{\gamma_{c}}^{2}\,\bar{i}/k \tag{95} $$

$$ {\pi^{ann}}_{t} = 0.25\,\left({\pi_{t-2}}_{t-1}+{\pi_{t-2}}_{t}+{\pi^{q}}_{t}+{\pi_{t-1}}_{t}\right) \tag{96} $$

Baseline Smets Wouters Model variable declaration with adaptive learning

Table 3: Smets Wouters Model with Adaptive Learning: Endogenous variable definitions

Variable	LaTeX	Description
ewma	$\varepsilon^{w,aux}$	Auxiliary wage markup moving average variable
epinfma	$\varepsilon^{p,aux}$	Auxiliary price markup moving average variable
zcapf	z^{flex}	Capital utilization rate flex price economy
rkf	$r^{k,flex}$	rental rate of capital flex price economy
kf	$k^{s,flex}$	Capital services flex price economy
pkf	q^{flex}	real value of existing capital stock flex price economy
cf	c^{flex}	Consumption flex price economy
invef	i^{flex}	Investment flex price economy
yf	y^{flex}	Output flex price economy
labf	l^{flex}	hours worked flex price economy
wf	w^{flex}	real wage flex price economy
rrf	r^{flex}	real interest rate flex price economy
mc	μ_p	gross price markup
zcap	z	Capital utilization rate
rk	r^k	rental rate of capital
k	k^s	Capital services
pk	q	real value of existing capital stock
c	c	Consumption
inve	i	Investment
y	y	Output
lab	l	hours worked
pinf	π	Inflation
w	w	real wage
r	r	nominal interest rate
a	ε_a	productivity process
b	$c_2 * \varepsilon_t^b$	Scaled risk premium shock
g	ε^g	Exogenous spending
qs	ε^i	Investment-specific technology
ms	ε^r	Monetary policy shock process
spinf	ε^p	Price markup shock process
sw	ε^w	Wage markup shock process
kpf	k^{flex}	Capital stock flex price economy
kp	k	Capital stock
pinf4	π^{ann}	MMB: Annualized inflation

Table 3 – Continued

Variable	LaTeX	Description
eg	ε^{gov}	MMB: Government speding shock
interest	i	MMB: Common variable - Annualized nominal interest rate
inflation	π^{ann}	MMB: Common variable - Annualized inflation
inflationq	π^{q}	MMB: Common variable - Quarterly inflation
outputgap	y^{gap}	MMB: Common variable - Output gap[
output	y	MMB: Common variable - Output
fispol	g	MMB: Common variable - Government speding
inflationql	π_{t-1}	MMB: Common variable - Lagged Inflation
inflationql2	π_{t-2}	MMB: Common variable - 2x Lagged Inflation
inflationqls	π_{t-3}	MMB: Common variable - 3x Lagged Inflation

Table 4: Smets Wouters Model with Adaptive Learning: Exogenous variable definitions

Variable	LaTeX	Description
ea	ε^{a}	productivity shock
eb	ε^{b}	Investment-specific technology shock
eqs	ε^{R}	Investment-specific technology shock
em	ε^{m}	Monetary policy shock
epinf	ε^{p}	Price markup shock
ew	ε^{w}	Wage markup shock
interest_	ε^{R}	MMB: Common variable - MP shock
fiscal_	π_{t}	MMB: Common variable - Fiscal shock

Table 5: Smets Wouters Model with Adaptive Learning: Parameter Definitions

Variable	LaTeX	Description
cofintintb1	ρ	Taylor rule interest rate smoothing
cofintinf0	θ_{π}	Taylor rule inflation feedback
cofintoutp	θ_{x}	Taylor rule ouput gap feedback
std_r_	σ_{r}	Taylor rule monetary policy shock size (set to 0.01)
curvw	ε_{w}	Curvature Kimball aggregator wages
cgy	ρ_{ga}	Feedback technology on exogenous spending
curvp	ε_{p}	Curvature Kimball aggregator prices

Table 5 – Continued

Variable	LaTeX	Description
`constelab`	$\bar{l}$	steady state hours
`constepinf`	$\bar{\pi}$	steady state inflation rate
`constebeta`	$100(\beta^{-1}-1)$	time preference rate in percent
`cmaw`	μ_w	coefficient on MA term wage markup
`cmap`	μ_p	coefficient on MA term price markup
`calfa`	α	capital share
`czcap`	ψ	capacity utilization cost
`csadjcost`	φ	investment adjustment cost
`ctou`	δ	depreciation rate
`csigma`	σ_c	risk aversion
`chabb`	λ	external habit degree
`cfc`	ϕ_p	fixed cost share
`cindw`	ι_w	Indexation to past wages
`cprobw`	ξ_w	Calvo parameter wages
`cindp`	ι_p	Indexation to past prices
`cprobp`	ξ_p	Calvo parameter prices
`csigl`	σ_l	Frisch elasticity
`clandaw`	ϕ_w	Gross markup wages
`crpi`	r_π	Taylor rule inflation feedback
`crdy`	$r_{\Delta y}$	Taylor rule output growth feedback
`cry`	r_y	Taylor rule output level feedback
`crr`	ρ	interest rate persistence
`crhoa`	ρ_a	persistence productivity shock
`crhob`	ρ_b	persistence risk premium shock
`crhog`	ρ_g	persistence spending shock
`crhoqs`	ρ_i	persistence risk premium shock
`crhoms`	ρ_r	persistence monetary policy shock
`crhopinf`	ρ_p	persistence price markup shock
`crhow`	ρ_w	persistence wage markup shock
`ctrend`	$\bar{\gamma}$	net growth rate in percent
`cg`	$\frac{\bar{g}}{\bar{y}}$	steady state exogenous spending share
`cgamma`	γ_c	BGP growth rate of quarterly GDP
`clandap`	$\phi_{p,ss}$	SS fixed cost share
`cbetabar`	$\bar{\beta}$	SS SDF
`cr`	$\bar{r}*$	SS Real rate
`cpie`	$\bar{\pi}$	SS inflation rate
`crk`	$\bar{r}^k$	SS rental rate of capital

Table 5 – Continued

Variable	LaTeX	Description
`cw`	$\bar{w}$	SS real wage
`cikbar`	$\bar{i}/k$	SS investment rate
`cik`	$\bar{i}_{ss}/k_{ss}$	SS investment rate
`clk`	i_{ss}/k_{ss}	BGP detrended SS investment rate
`cky`	k_{ss}/y_{ss}	SS capital-output ratio
`ciy`	i_{ss}/y_{ss}	SS investment-output ratio
`ccy`	c_{ss}/y_{ss}	SS consumption-output ratio
`crkky`	$\bar{r}^k * k_{ss}/y_{ss}$	SS capital income share of output
`cwhlc`	$1/\phi_w * (1-\alpha)/\alpha * \bar{r}^k * k_{ss}/y_{ss}$	SS wage income share of output
`cwly`	$1 - \bar{r}^k * k_{ss}/y_{ss}$	SS wage share
`conster`	$(\bar{r} * -1) * 100$	SS r* in percentage points

Smets Wouters Parameters calibration

Table 6: Smets Wouters Model with Adaptive Learning: Parameter Values

Parameter	Value	Description
ρ	0.900	Interest rate smoothing
θ_π	0.150	Inflation response
θ_x	0.013	Output gap response
σ_r	0.01	MP shock size (normalized to 0.01)
ε_w	10.000	Curvature Kimball aggregator wages
ρ_{ga}	0.519	Feedback technology on exogenous spending
ε_p	10.000	Curvature Kimball aggregator prices
$\bar{l}$	0.551	steady state hours
$\bar{\pi}$	0.787	steady state inflation rate
$100(\beta^{-1}-1)$	0.166	time preference rate in percent
μ_w	0.850	coefficient on MA term wage markup
μ_p	0.701	coefficient on MA term price markup
α	0.190	capital share
ψ	0.546	capacity utilization cost
φ	5.761	investment adjustment cost
δ	0.025	depreciation rate
σ_c	1.381	risk aversion
λ	0.713	external habit degree

Table 6 – Continued

Parameter	Value	Description
ϕ_p	1.606	fixed cost share
ι_w	0.585	Indexation to past wages
ξ_w	0.706	Calvo parameter wages
ι_p	0.243	Indexation to past prices
ξ_p	0.652	Calvo parameter prices
σ_l	1.838	Frisch elasticity
ϕ_w	1.500	Gross markup wages
r_π	2.044	Taylor rule inflation feedback
$r_{\Delta y}$	0.225	Taylor rule output growth feedback
r_y	0.088	Taylor rule output level feedback
ρ	0.810	interest rate persistence
ρ_a	0.958	persistence productivity shock
ρ_b	0.219	persistence risk premium shock
ρ_g	0.977	persistence spending shock
ρ_i	0.711	persistence risk premium shock
ρ_r	0.148	persistence monetary policy shock
ρ_p	0.889	persistence price markup shock
ρ_w	0.969	persistence wage markup shock
$\bar{\gamma}$	0.431	net growth rate in percent
$\frac{\bar{g}}{\bar{y}}$	0.180	steady state exogenous spending share
γ_c	1.004	BGP growth rate based on quarterly trend growth rate to GDP
$\phi_{p,ss}$	1.606	SS fixed cost share
$\bar{\beta}$	0.992	SS SDF
$\bar{r}*$	1.016	SS Real rate
$\bar{\pi}$	1.008	SS inflation rate
$\bar{r}^k$	0.033	SS rental rate of capital
$\bar{w}$	0.682	SS real wage
$\bar{i}/k$	0.029	SS investment rate
$\bar{i}_{ss}/k_{ss}$	0.029	SS investment rate
i_{ss}/k_{ss}	0.204	BGP detrended SS investment rate
k_{ss}/y_{ss}	5.827	SS capital-output ratio
i_{ss}/y_{ss}	0.171	SS investment-output ratio
c_{ss}/y_{ss}	0.649	SS consumption-output ratio
$\bar{r}^k * k_{ss}/y_{ss}$	0.190	SS capital income share of output
$1/\phi_w * (1-\alpha)/\alpha * \bar{r}^k * k_{ss}/y_{ss}$	0.832	SS wage income share of output
$1 - \bar{r}^k * k_{ss}/y_{ss}$	0.810	SS wage share

Table 6 – Continued

Parameter	Value	Description
$(\bar{r}*-1)*100$	1.555	SS r* in percentage points