UNANCHORED EXPECTATIONS: SELF-REINFORCING LIQUIDITY TRAPS AND MULTIPLE STEADY STATES

JOEP LUSTENHOUWER
Otto-Friedrich-Universität Bamberg

We study a New Keynesian model with bounded rationality, where agents choose their expectations heterogeneously from a discrete choice set. The range of their set of possible expectation values can be interpreted as the anchoring of expectations. In the model, multiple locally stable steady states can arise that reflect coordination on particular expectation values. Moreover, bad shocks to the economy can trigger a self-reinforcing wave of pessimism, where the zero lower bound on the nominal interest rate becomes binding, and agents coordinate on a locally stable liquidity trap steady state. When we let the anchoring of expectations evolve endogenously, it turns out that the anchoring of expectations at the time the bad shocks hit the economy is crucial in determining whether the economy can recover from the liquidity trap. Finally, we find that a higher inflation target makes it less likely that self-reinforcing liquidity traps arise.

Keywords: Liquidity Traps, Heterogeneous Expectations, Multiple Steady States

1. INTRODUCTION

Over the last decades, central banks have become increasingly transparent in the hope to "anchor" private sector expectations [see e.g. Woodford (2005) and Dincer and Eichengreen (2014)]. The theoretical foundation is that only when participants in the economy coordinate their expectations on the desired equilibrium, this equilibrium can be reached. If, instead, private sector expectations considerably deviate from the targets, the central bank will have a hard time achieving its goals. However, being more transparent also allows more effective external scrutiny and evaluation of monetary policy [Svensson (2009)]. If monetary policy is judged not to be well executed and goals are not met, transparency may not lead to more anchored expectations. In this paper, we contribute to the understanding of the anchoring of expectations to targets communicated by the central bank, and investigate how unanchored expectations can lead to waves of optimism and pessimism and to liquidity traps.

1469-8056/19 **845**

Most of the theoretical macro literature has maintained the assumption that private agents all have the same expectations and that these are perfectly rational. However, when one considers survey data on expectations, considerable heterogeneity is found [see e.g., Carroll (2003), Mankiw et al. (2003), Andrade et al. (2013), Cole and Milani (2019)]. The same finding arises in laboratory experiments with human subjects [e.g. Assenza et al. (2014), Pfajfar and Žakelj (2018)]. Furthermore, as highlighted by, for example, Blinder et al. (2008), under rational expectations and full information and credibility, there is no explicit role for central bank communication. Assuming homogeneous rational expectations may therefore not be the best modeling approach when one wants to realistically investigate the influence of expectations and their anchoring on the economy. Moreover, expectations in survey data typically consist of round numbers. This reflects the realistic feature that people do not come up with expectations with infinitely many decimals, but instead prefer round numbers (see Curtin (2010) for a discussion of this phenomenon labeled "digit preference" with regard to the Michigan Surveys of Consumers and, e.g., Assenza et al. (2014) for the same finding in laboratory experiments with human subjects).

In this paper, we present a model where expectations are both heterogeneous and are formed in a discrete state-space. This allows us to intuitively capture the concept of anchoring of expectations. In our model, agents consider a discrete distribution of expectation values around the targets of the central bank. When this distribution of possible expectation values has a small range, we say that expectations are strongly anchored to the targets. When the range is large, and agents also consider expectation values that are far away from the targets, expectations are unanchored. In each period, each agent picks an expectation from the distribution of possible expectation values. They do this randomly, where the probability of choosing each expectation value depends on past realizations. That is, when agents that had pessimistic expectations turn out to be right (e.g. due to bad shocks to the economy), in subsequent periods, other agents are more likely to choose pessimistic expectations as well. However, there will always be heterogeneity present, with some agents expecting values close to the targets and other agents having more pessimistic or optimistic expectations.

Our first finding is that multiple steady states can arise. First of all, if the reaction of monetary policy to inflation (expectations) is not very strong, expectations can become almost self-fulfilling and coordination on one of the round expectation values that agents can choose from can occur. Secondly, agents' expectations can move close to the maximum or the minimum value that they are considering, that is, close to the most extreme values that lie in the range of possible expectation values, given the current anchoring of expectations. The latter type of steady state is especially relevant under the zero lower bound (ZLB) on the nominal interest rate. The monetary authority can then not adjust the interest rate to stimulate inflation and the output gap and has no means of stabilizing expectations. In this case, *self-reinforcing waves of pessimism* can arise.

However, since expectations are at least somewhat anchored to the targets of the central bank, agents will not adjust their expectations downward unboundedly. Instead, after observing very low inflation and output gap, the anchoring of expectations induces agents to choose expectations somewhat in the direction of the central bank's targets. These expectations, in turn, can imply that inflation and output stay as they are, so that expectations remain constant as well. The economy is then in a liquidity trap steady state. The values of expectations and of output gap and inflation in this steady state depend on the anchoring of expectations.

We first illustrate the above findings in a stylized, analytically tractable model with only three types of agents: optimists, pessimists, and fundamentalists. The latter expect inflation and output gap to be exactly at their targets, while optimist (pessimists) have somewhat higher (lower) expectations. We then turn to our benchmark model where there are many different expectation values, each spaced by 0.5%. Here the range of expectation values that agents are considering (i.e., the anchoring of expectations) is furthermore allowed to vary *endogenously* with past absolute deviations from target in the economy. That is, when inflation and output gap fluctuate closely around their targets for a long time, expectations become more anchored, while they become less anchored when inflation and output gap deviate considerably from their targets for longer periods of time (e.g., in a liquidity trap).

We show that in our benchmark model, the anchoring of expectations at the time that the economy is hit by bad shocks is crucial in determining whether the economy can recover, or instead ends up in a liquidity trap that can last for prolonged periods of time. In both cases, the bad shocks initiate a wave of pessimism where the majority of agents coordinate on the lowest expectations that they are willing to consider, given the current anchoring of expectations. If expectations are strongly enough anchored to the targets, then these most extreme expectations will not be too low, and small positive shocks to economic fundamentals can bring the economy back on a path to the target steady state. If expectations are more unanchored, the wave of pessimism becomes more severe, and larger shocks would be needed to achieve recovery. If such larger shocks do not arise, the economy will remain in the liquidity trap, which slowly becomes worse as expectations become less anchored over time. Therefore, there is an important role for the size of the exogenous shocks as well. Finally, we find that not only more anchored expectations but also a higher inflation target can prevent such a liquidity trap.

Our results are related to those of Evans et al. (2008), Eusepi (2010), and Benhabib et al. (2014), who study liquidity traps under adaptive learning. In their models deflationary spirals with rapidly decreasing inflation and output can arise. We add to this by adding the concept of anchoring of expectations. Since the expectations of our agents are anchored to the targets of the central bank, these rapid deflationary spirals do not arise. Instead, the liquidity trap initially remains

bounded and only slowly starts to get worse as expectations become less anchored over time.[1]

The self-fulfilling dynamics that arise in our model, especially at the ZLB, are also related to the literature on sunspot shocks under rational expectations [see e.g., Mertens and Ravn (2014) and Aruoba et al. (2017)]. Our approach adds to this literature by providing a mechanism where both self-fulfilling expectations and the evolution of the bounds on these expectations arise endogenously and do not rely on an exogenous sunspot shock process. This allows us to obtain more insights in the anchoring of expectations to the targets of the central bank, which is arguably an endogenous phenomenon. Moreover, where the adaptive learning and sunspot literature assume homogeneous expectations, we extend the analysis by modeling heterogeneity in expectations.

Our model is further related to the heterogeneous expectation models of Branch and McGough (2010), Massaro (2013), Gasteiger (2014), Di Bartolomeo et al. (2016), De Grauwe and Ji (2019), and Hommes and Lustenhouwer (2019*a*). These papers, however, assume that agents can choose between two expectation formation rules, rather than a distribution of heterogeneous expectations.

More closely related are the works of De Grauwe (2011), Anufriev et al. (2013), Agliari et al. (2015), Agliari et al. (2017), Pecora and Spelta (2017), and Hommes and Lustenhouwer (2019*b*). However, these papers discuss either a model with a relatively small number of different expectation values (like the 3-type model) or the large type limit specification with a continuum of expectation values. None of these papers consider the case where the range of expectation values varies endogenously, and where the number of expectation values can be large but finite. Moreover, with the exception of Hommes and Lustenhouwer (2019*b*), the above literature focuses on the desired, and in some cases optimal, specification of monetary policy when the ZLB is not binding. In contrast, our main goal is to investigate how self-reinforcing waves of pessimism can arise when conventional monetary policy can no longer be used due to the restrictions of the ZLB on the nominal interest rate. The current paper differs from Hommes and Lustenhouwer (2019*b*) by explicitly modeling digit preference and by modeling the anchoring of expectations as the endogenously varying range of expectation values that agents are considering. This leads to the existence of multiple steady states, including a locally stable liquidity trap steady state, which do not arise in Hommes and Lustenhouwer (2019*b*).

Finally, the model in this paper is closely related to Guse and Carton (2014), who consider replicator dynamic learning in a simple model of price movements. In this model, they make the choice set of firms discrete and find that multiple steady states arise with similar properties and a similar intuition as in our model.

The remainder of this paper is organized as follows. In Section 2 the economic model and general expectation formation mechanism are presented. In Section 3 we illustrate the main intuitions of our model with the stylized 3-type example, while Section 4 considers our benchmark model with an endogenously varying anchoring of expectations. Section 5 concludes.

2. MODEL SPECIFICATION

2.1. Economic Environment

We use a log-linearized New Keynesian model in line with Woodford (2003). Micro-foundations for this model when expectations are heterogeneous can be found in Hommes and Lustenhouwer (2019*a*). The derivations in that paper largely follow Kurz et al. (2013).[2] Alternative Microfoundations of the same New Keynesian model with heterogeneous expectations are presented in Branch and McGough (2009).

Log-linearized output gap (x_t) and inflation (π_t) are given by

$$x_t = \bar{E}_t x_{t+1} - \frac{1}{\sigma}(i_t - \bar{E}_t \pi_{t+1} - \bar{r}) + u_t, \qquad (1)$$

$$\pi_t = \beta \bar{E}_t \pi_{t+1} + \kappa x_t + e_t. \qquad (2)$$

Here i_t is the nominal interest rate, β is the discount factor of households, σ is their relative risk aversion, $\bar{r} = \frac{1}{\beta} - 1$ is the steady-state real interest rate, and κ is a composite parameter. $\bar{E}_t$ denotes the aggregate of individual expectations and e_t and u_t are shocks to the economy.

We assume that the nominal interest rate is set according to a forward-looking Taylor type interest rate rule[3]

$$i_t = \bar{r} + \pi^T + \phi_1(\bar{E}_t \pi_{t+1} - \pi^T) + \phi_2(\bar{E}_t x_{t+1} - x^T), \qquad (3)$$

where π^T is the central banks inflation target, and $x^T = \frac{1-\beta}{\kappa}\pi^T$ is the output gap target consistent with π^T.

Plugging (3) into (1) gives the following model:

$$x_t = (1 - \frac{\phi_2}{\sigma})\bar{E}_t x_{t+1} - \frac{\phi_1 - 1}{\sigma}(\bar{E}_t \pi_{t+1} - \pi^T) + \frac{\phi_2}{\sigma}x^T + u_t, \qquad (4)$$

$$\pi_t = \beta \bar{E}_t \pi_{t+1} + \kappa x_t + e_t. \qquad (5)$$

2.2. Expectation Formation

We assume that agents form expectations heterogeneously, with some agents expecting somewhat higher inflation and/or output gap and other agents expecting inflation and/or output gap that are lower. This heterogeneity could be caused by agents making mistakes. Alternatively, the heterogeneity can arise because some agents think they have reasons to be more optimistic or pessimistic in their predictions than is warranted by the publicly available information.

More specifically, expectations are distributed around the targets of the central bank, π^T and x^T. Therefore, we can interpret the heterogeneity in expectations as some agents trusting the central bank and expecting inflation and output gap to be equal to their targets, while other agents expect that the central bank will

not be able to exactly achieve its goals, but that inflation and output gap will be somewhat higher or lower.

We further assume that when agents that were more optimistic or pessimistic in their prediction turn out to be right, other agents will *learn* from this and adjust their expectations in the direction of the better performing agents. Agents may, for example, think that the correct agents had additional information available to them, or just had more skills in analyzing the economic environment.

We implement this way of expectation formation with a heuristic switching model as in Brock and Hommes (1997), where agents switch between simple prediction rules, or heuristics. The heuristics in our model consists of deviations from the target values of the central bank.

The fraction of agents using the heuristic with deviation $b_{z,h}$ for variable z in period t is updated according to the discrete choice model with multinomial logit probabilities [see Manski et al. (1981)] given by

$$n_t^{z,h} = \frac{e^{\omega U_{t-1}^{z,h}}}{\sum_{h=1}^{H^z} e^{\omega U_{t-1}^{z,h}}}, \qquad\qquad z = \pi, x. \qquad (6)$$

Here, H^π and H^x are the total number of prediction values used for, respectively, inflation and output gap. $U_t^{z,h}$ is the fitness measure of predictor h in period t, which will depend on past prediction errors, and ω is the intensity of choice. The intensity of choice determines how sensitive agents are with respect to relative performance of prediction values. When the intensity of choice is low, agents have a relatively high probability of choosing any forecast value that they are considering, also when this forecast did not work well in the past. When the intensity of choice is high, agents are only likely to choose forecast values that performed well in the past. In the latter case, agents' forecasts will be more coordinated, while in the former case there will be a large degree of heterogeneity.

Aggregate expectations are then given by a weighted average of the predictions of all types. This gives for, respectively, inflation and output gap:

$$\bar{E}_t \pi_{t+1} = \pi^T + \sum_{h=1}^{H^\pi} b_{\pi,h} n_t^{\pi,h}, \qquad (7)$$

and

$$\bar{E}_t x_{t+1} = x^T + \sum_{h=1}^{H^x} b_{x,h} n_t^{x,h}. \qquad (8)$$

3. AN EXAMPLE WITH THREE TYPES

Before we turn to the richer model with a large number of prediction values in Section 4, we will first illustrate some key properties of our model with a simple example where $H^x = H^\pi = 3$. There then are three types of agents for both output and inflation, which we can label **fundamentalists**, **optimists**, and **pessimists**.

Fundamentalists believe in the targets of the central bank. Their expectations are therefore given by $E_t^{fun} x_{t+1} = x^T$ and $E_t^{fun} \pi_{t+1} = \pi^T$. Then there are optimists and pessimists who expect a deviation from target of b and $-b$, respectively. Their expectations are therefore given by $E_t^{opt} x_{t+1} = x^T + b$, $E_t^{opt} \pi_{t+1} = \pi^T + b$, $E_t^{pes} x_{t+1} = x^T - b$, and $E_t^{pes} \pi_{t+1} = \pi^T - b$. The magnitude of the deviation, b, determines the range of values around the targets that agents are willing to consider as expectations. We take this as a measure of the "anchoring of expectations". If b is small, then agents will always expect values close to the targets of the central bank. We say that expectations are **strongly anchored** in this case. If, on the other hand, b is large, then agents could potentially switch to expectations that lie far from the target by becoming optimistic or pessimistic. In this case, expectations are **unanchored**. In this section, we will assume that b is fixed so that the anchoring of expectations is exogenously given. In Section 4, we will allow for endogenous evolution of the anchoring of expectations. Note that we allow an agent to be optimistic about one variable, while being pessimistic or fundamentalistic about the other, so that the fractions of agents that are optimistic and pessimistic (denoted respectively $n_t^{z,opt}$ and $n_t^{z,pes}$) may differ between the two variables $z = x, \pi$. The fraction of fundamentalists of variable z equals $n_t^{z,fun} = 1 - n_t^{z,opt} - n_t^{z,pes}$.

In this 3-type model, we will first show how our heterogeneous expectations framework with nonlinear updating can imply multiple steady states, even in an otherwise linear model (Section 3.1). Here we also illustrate how monetary policy can prevent the existence of multiple steady states, so that coordination on optimism or pessimism is no longer possible in the long run. In Section 3.2, we introduce the ZLB on the nominal interest rate and show that this drastically alters policy implications. Now the existence of a liquidity trap steady state with coordination on pessimism for both inflation and output gap can no longer be prevented by choosing adequate coefficients in the Taylor rule. Instead, such a "bad" equilibrium can only be ruled out with a sufficiently high inflation target or with sufficiently anchored expectations.

3.1. Steady States in the 3-Type Model Without ZLB

We first ignore the ZLB on the nominal interest rate and let output gap and inflation be always given by (4) and (5). Plugging in the expected deviations from target of fundamentalists, optimist, and pessimist (respectively, $b_{x,h} = -b$, $b_{x,h} = 0$, and $b_{x,h} = b$ for output, and $b_{\pi,h} = -b$, $b_{\pi,h} = 0$, and $b_{\pi,h} = b$ for inflation) in (7) and (8) results in

$$\bar{E}_t x_{t+1} = x^T + b(n_t^{x,opt} - n_t^{x,pes}), \tag{9}$$

$$\bar{E}_t \pi_{t+1} = \pi^T + b(n_t^{\pi,opt} - n_t^{\pi,pes}). \tag{10}$$

Fractions are given by (6), with $h = opt, pes, fun$. For analytical tractability, we will assume in this section that the fitness measure is given by the most recently

observed squared prediction error. That is, $U_{t-1}^{\pi,h} = -(\pi_{t-1} - E_{t-2}^h \pi_{t-1})^2$ for inflation, and $U_{t-1}^{x,h} = -(x_{t-1} - E_{t-2}^h x_{t-1})^2$ for output gap, with $h = opt, pes, fun$. In Section 4, we will allow the fitness measure to also depend on forecasting performance in earlier periods.

We now look at what steady states can exist in this nonlinear model. Proposition 1 states that a fundamental target steady state always exists. The proof of the proposition is given in Appendix A.1.

PROPOSITION 1 (**Existence fundamental steady state 3-type model**). *A fundamental steady state where inflation and output gap are equal to the targets of the central bank, and where being a fundamentalist is the best performing heuristic, always exists, independent of monetary policy and the parametrization of the model.*

While Proposition 1 implies that, in the absence of shocks, the central bank could achieve its target in every period, there might also exist additional steady states, where inflation and output gap are not equal to their targets. This possibility arises due to the nonlinear updating process of the expectation fractions (6). In such non-fundamental steady states, fundamentalism would no longer be the best performing heuristic for both variables, but, instead, optimism or pessimism would dominate for inflation and/or output gap. Since agents can be either pessimist, fundamentalist, or optimist about both inflation *and* output gap, there are a total of nine combinations of dominating heuristics that could potentially comprise a steady state. That is, beside the fundamental steady state, eight different non-fundamental steady states could potentially exist. For analytical tractability, we focus below on the limiting case where the intensity of choice goes to infinity ($\omega = \infty$). This implies that agents in every period immediately switch to the best performing heuristic and coordinate perfectly on either optimism, pessimism, or fundamentalism.[4] As shown in Appendix B, when $\omega = \infty$, all nine steady states are locally stable when they exist. Therefore, if the central bank sets the parameters in the Taylor rule such that multiple steady states exist, convergence to any of them could occur, depending on initial conditions. It is then hard (or even impossible) for the central bank to guarantee convergence to the target steady state.

As we are interested in the possibility of a liquidity trap and this is most likely to arise under pessimistic expectations, we will focus on the steady state where pessimism dominates for both inflation and output gap. In Section 3.2 we will include the ZLB on the nominal interest rate and see whether this steady state can imply a liquidity trap.

Proposition 2 states, for the case of $\omega = \infty$, under what conditions this pessimistic steady state exists. The proof of the proposition is given in Appendix A.2.

PROPOSITION 2 (**Existence pessimistic steady state 3-type model**). *Without the ZLB on the nominal interest rate, a locally stable steady state where pessimism dominates for both inflation and output gap exists under $\omega = \infty$ if and only if $\phi_1 + \phi_2 < 1 + \frac{\sigma}{2}$.*

From Proposition 2 it follows that if the central bank does not respond to output gap (or at least not very strongly), then the pessimistic steady state can exist in the limiting case of infinite intensity of choice[5], even if $\phi_1 > 1$. That is, even if the Taylor principle is satisfied, and the fundamental equilibrium would be locally determinate under rational expectations, the additional pessimistic steady state can exist. However, when the central bank responds *more aggressively* to inflation (and/or output gap) expectations, the existence of the pessimistic steady state is ruled out.[6]

The intuition for the existence of the pessimistic steady state when monetary policy is not aggressive enough is the following. If agents are pessimists, then this induces inflation and output gap to become so low that agents have better predictive success by being pessimists than by being fundamentalists. We call such expectations **almost self-fulfilling**. When, on the other hand, monetary policy is aggressive enough, the central bank will sufficiently lower the interest rate when expectations become pessimistic, which will stimulate output and inflation, leading to realizations that are closer to the fundamental values. Pessimistic expectations then perform worse than fundamental expectations and the pessimistic steady state does not exist.

When pessimistic expectations are almost self-fulfilling, it could still be that pessimists make significant forecast errors. What is key, however, is that their forecast errors are *smaller* than those of optimists and fundamentalists and that pessimism outperforms all other available prediction values. When a larger number of expectation values is allowed (as in Section 4), forecast errors are required to be smaller for multiple steady states to exist. Below, we will also consider the case where pessimistic expectations lead to realizations that are not only low enough to make pessimistic expectations almost self-fulfilling but actually strictly lower than the expectations ($z_t < E_{t-1}z_t$). We will refer to this case as **self-reinforcing** pessimistic expectations. Without the ZLB, a self-reinforcing wave of pessimism can only arise if monetary policy does not satisfy the Taylor principle.[7] However, when we take account of the ZLB on the nominal interest rate, self-reinforcing waves of pessimism are more likely to arise, as we will show next.

3.2. A Liquidity Trap Steady State in the 3-Type Model

We now introduce the ZLB on the nominal interest rate in the above 3-type model and investigate its consequences for the existence of the pessimistic steady state. In normal times, the nominal interest rate is still given by the forward-looking Taylor rule, (3). We then say that our model is in the "positive interest rate region." When, however, the Taylor rule implies that $i_t < 0$, then i_t is instead set equal to zero. In that case, we say that the system is in the "ZLB region," and output gap and inflation reduce to

$$x_t = \bar{E}_t x_{t+1} + \frac{1}{\sigma}\bar{E}_t \pi_{t+1} + \frac{\bar{r}}{\sigma} + u_t, \tag{11}$$

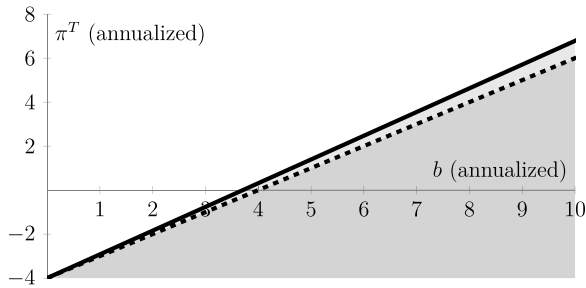$$\pi_t = \beta \bar{E}_t \pi_{t+1} + \kappa x_t + e_t. \tag{12}$$

FIGURE 1. Existence of liquidity trap steady state of Proposition 3 when $\phi_1 + \phi_2 > 1 + \frac{\sigma}{2}$, $\sigma = 0.157$, and $\beta = 0.99$, depending on the anchoring of expectations (b) and the inflation target (π^T). In the shaded area below the solid line (including both the light and the dark part) the liquidity trap steady state exists, while in the non-shaded area above the black line it is ruled out. Below the dashed line (darker shaded area) pessimistic expectations are self-reinforcing.

Proposition 3 states the conditions for existence of the pessimistic steady state under the ZLB. Its proof is given in Appendix A.3.

PROPOSITION 3 (**Existence liquidity trap steady state 3-type model**). *When all agents are pessimists, the ZLB on the nominal interest rate is binding if and only if*

$$\pi^T + \bar{r} < b(\phi_1 + \phi_2). \tag{13}$$

In that case, the pessimistic steady state of proposition 2 becomes a liquidity trap steady state, and exists (assuming $\omega = \infty$) if and only if

$$\pi^T + \bar{r} < b(1 + \frac{\sigma}{2}). \tag{14}$$

From Proposition 3 it follows that a liquidity trap steady state with pessimistic expectations and a binding ZLB can exist if expectations are *unanchored* (b large). However, its existence can then be prevented if the *inflation target* is sufficiently high. Note, furthermore, that the liquidity trap steady state can exist when a pessimistic steady state would also have existed in the absence of a ZLB on the nominal interest rate ($\phi_1 + \phi_2 < 1 + \frac{\sigma}{2}$); but that it can also exist when monetary policy is aggressive enough to rule out the existence of a pessimistic steady state without the ZLB ($\phi_1 + \phi_2 > 1 + \frac{\sigma}{2}$). In the first case, (13) is the binding condition for the existence of the liquidity trap steady state, while in the second case it is (14).

In Figure 1, we plot the condition for the existence of the liquidity trap steady state in case of $\phi_1 + \phi_2 > 1 + \frac{\sigma}{2}$ (i.e., we plot (14)) with a solid line (the meaning of the dashed line will be discussed below). Here, we parameterize the coefficient of relative risk aversion as in Woodford (1999) to $\sigma = 0.157$ and set $\beta = 0.99$. In the figure, it can be seen that, as long as pessimists expect annualized inflation to be within 3.5% of its target ($b < 3.5$), the liquidity trap steady state does not

exist even if the central bank's inflation target is zero. However, as expectations become more unanchored, a considerably higher inflation target is needed to rule out the possibility of a pessimistic liquidity trap.

Finally, we turn to the possibility of *self-reinforcing* expectations in a liquidity trap. Remember that in order for the steady state to exist, expectations must at least be almost self-fulfilling in the sense that they result in realizations of inflation and output gap that are closer to the pessimistic expectations than to the fundamentalist ones. However, we now consider whether expectations can also lead to realizations of inflation and output gap that are as low as or lower than what pessimists expected. Proposition 4, whose proof is given in Appendix A.4, states when this is the case.

PROPOSITION 4 (**A self-reinforcing liquidity trap**). *The liquidity trap steady state of Proposition 3 is self-reinforcing if and only if*

$$\pi^T + \bar{r} < b. \tag{15}$$

Condition (15) is plotted as a dashed line in Figure 1. It follows that for most of the combinations of anchoring of expectations (b) and the inflation target (π^T) where the pessimistic steady state exists, it is also self-reinforcing (the darker shaded area below the dashed line).

4. BENCHMARK MODEL

We now turn to our benchmark model with a large number of expectation values. While the 3-type model of Section 3.1 gives interesting insights (in part because of its analytical tractability), it is rather stylized in the sense that it does not allow for different degrees of optimism and pessimism. Instead, all pessimists and optimists always have the same expectations. The model in this section allows for different degrees of optimism and pessimism by introducing more possible expectation values. In addition, we will allow the number of expectation values that agents consider and their range (the anchoring of expectations) to vary over time *endogenously*.

Even though it makes sense to assume that different degrees of optimism and pessimism are allowed, some **discreteness** in expectations is empirically relevant. In practice, people do not form expectations with infinitely many decimals. Instead, humans prefer round numbers when reporting their expectations (a phenomenon labeled *digit preference*)[8], and our benchmark model matches this realism. A typical explanation for digit preference is that it is costly to come up with more exact numbers [Curtin (2010)]. We choose to model this by assuming that agents form expectations with a precision of 0.5% (annualized). The different possible expectation values will therefore be spaced by 0.5% and go in both directions of the target. Assuming a higher precision of expectations formation (e.g. 0.01%) would not change our results qualitatively.[9] Section 4.1 describes the specification of our benchmark model, Section 4.2 shows that, as in the 3-type

model, multiple steady states can exist, while Section 4.3 focuses on the ZLB on the nominal interest and the possibility of a self-fulfilling liquidity trap steady state in the benchmark model. Here we will let the anchoring of expectations vary endogenously with past deviations of inflation and output gap from their targets. We show that long-lasting liquidity traps can arise. Whether the economy can recover from such a liquidity trap very much depends on how anchored expectations were when the economy fell in the trap, as well as on the inflation target and the size of exogenous shocks.

4.1. Model with Many Expectation Values

As in Section 3.1, the model is given by Equations (4) through (8). However, instead of letting $H^x = H^\pi = 3$ in the expectation equations (6)–(8), we now allow H^x and H^π to vary and to be considerably larger. The different deviations from target that agents can choose from ($b_{\pi,h}$ and $b_{x,h}$) will be spaced by 0.5% (annualized) while extending equally to positive and negative values. For example, if $H^\pi = 41$, then the 41 possible expected deviations from target for inflation will, in annualized values, be $\{-10, -9.5, -9, ..., -0.5, 0, 0.5, ..., 9, 9.5, 10\}$. Similarly, when $H^x = 33$, the expected deviations of output gap from target will range between 8% below target and 8% above target (annualized). H^x and H^π thus determine the largest deviations from target that agents consider, similar to b in Section 3.1. We will therefore use H^x and H^π as measures of the *anchoring of expectations*. Laws of motion for the endogenous evolution of H^x and H^π are presented in Section 4.3.

Finally, we allow the fitness measures used in (6) to depend on the whole history of prediction errors, and not just the prediction error in the most recent period. That is,

$$U_{t-1}^{\pi,h} = -(1 - \rho)(\pi_{t-1} - E_{t-2}^h \pi_{t-1})^2 + \rho U_{t-2}^{\pi,h}, \tag{16}$$

and

$$U_{t-1}^{x,h} = -(1 - \rho)(x_{t-1} - E_{t-2}^h x_{t-1})^2 + \rho U_{t-2}^{x,h}. \tag{17}$$

Therefore, when H^x and H^π take on large values, this only implies that agents *could* choose very low or high expectation values, but not that there necessarily will be agents that do so. Whether high (or low) expectation values will be chosen depends on whether past inflation and output gap were high (low), according to (16) and (17). This point will be further illustrated in Figure 3, discussed in Section 4.3.

4.2. Multiple Steady States

Before turning to the dynamics that arise when H^x and H^π evolve endogenously, we first consider a fixed value of H^x and H^π to illustrate how almost self-fulfilling expectations can lead to the existence of multiple non-fundamental steady states
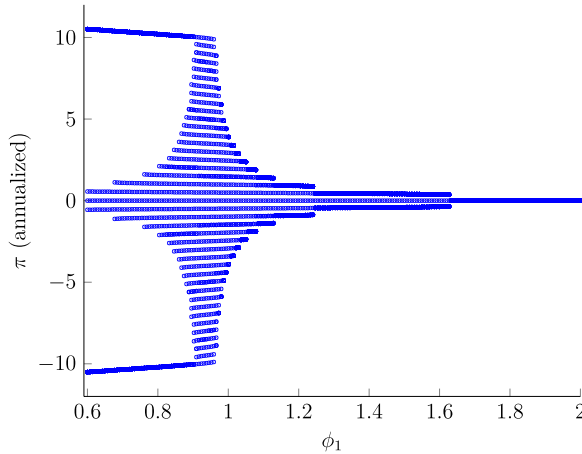
FIGURE 2. Bifurcation diagram of 41-type model in ϕ_1 for large intensity of choice ($\omega = 2 \cdot 10^6$). The blue lines represent (locally stable) steady states at the 41 different levels of inflation expectations.

where optimism or pessimism dominates, as in the 3-type model. In particular, we fix $H^x = H^\pi = 41$, so that expectations are relatively unanchored, and agents consider values up to 10% above and below the inflation and output gap targets.[10] In Section 4.3 we will look at how the ZLB on the nominal interest rate can create the possibility of a liquidity trap steady state in a specification with endogenously varying H^x and H^π.

Figure 2 plots a bifurcation diagram that is obtained by simulating the model without exogenous shocks for many different initial conditions for different values of the policy parameter ϕ_1. The intensity of choice is chosen high ($\omega = 2 \cdot 10^6$) to facilitate coordination and hence the existence of almost self-fulfilling non-fundamental steady states.

In the figure, it can be seen that for each of the 41 inflation prediction values there is a range of ϕ_1-values where this prediction comprises an *almost self-fulfilling steady state*. In particular, when ϕ_1 is around 0.95, any value of inflation that agents could expect would lead to a realization of inflation that is close to this expectation value. Agents will then again choose these optimistic or pessimistic expectations, so that the model remains in this steady state. From a policy perspective, this implies that for these relatively low values of ϕ_1, the central bank might have difficulty achieving its targets. Shocks to fundamentals of the economy (or to expectations themselves) could easily trigger *coordination on optimistic or pessimistic expectations* with corresponding high or low inflation and output gap.

When, however, the monetary authority responds more aggressively to infla-tion expectations, non-fundamental expectations will lead to realizations that are less self-fulfilling due to stronger mean reversion to the central bank's targets.

In the subsequent period, agents will then adjust their expectations and become less optimistic or pessimistic. This leads to realizations of inflation even closer to the target, which again bring expectations closer to the targets. As ϕ_1 increases, the non-fundamental steady states in Figure 2 therefore start to disappear. When ϕ_1 is larger than 1.3, only steady states very close to the central bank's targets exist, which might not be considered a problem from a policy perspective. As ϕ_1 is increased to 1.7, the fundamental steady state is unique and convergence to the targets will always occur in the long run.

When the central bank responds very weakly to inflation expectations (e.g., $\phi_1 < 0.8$), waves of optimism and pessimism become *self-reinforcing*. That is, low expectations of inflation lead to even lower realizations, which causes agents to become even more pessimistic. This process continues until the lowest possible expectations are reached (lowest blue line in Figure 2). Analogously, self-reinforcing optimistic expectation leads to convergence to the highest possible expectation values (highest blue line in Figure 2).

For lower intensity of choice, the range of policy parameters for which intermediate steady state exists becomes smaller than in Figure 2. However, when monetary policy is too weak, it will always be the case that optimistic and pessimistic expectations are self-reinforcing, implying the existence of non-fundamental steady states where the most extreme expectation values dominate.

4.3. Liquidity Traps in Many Type Models

We now consider the possibility of self-reinforcing liquidity traps in our benchmark model. As in Section 3.2, output gap and inflation evolve according to (4) and (5) when the model is in the "positive interest rate region," and according to (11) and (12) when the ZLB on the nominal interest rate is binding ("ZLB region").

When the model is in the ZLB region, there is no feedback on expectations through monetary policy, and expectations can easily become self-reinforcing. In Section 3.2 we saw that in the 3-type model this can lead to the existence of a liquidity trap steady state with coordination on pessimistic expectations. We will now investigate whether *coordination on a liquidity trap steady state with self-reinforcing expectations* can also arise in our benchmark model.

In order to realistically model dynamics under the ZLB, we let the anchoring of expectations evolve endogenously over time. We do this by assuming that expectations become *less* anchored when realizations are far away from the target for an extended period of time and become *more* anchored when realizations continue to be relatively close to the central bank's targets. We model this by letting agents adaptively update the annualized largest absolute deviation from target (LAD) that they consider, based on new realizations of absolute deviations of variables from target. That is,

$$LAD_t^z = \gamma LAD_{t-1}^z + (1 - \gamma) * 400 * |z_{t-1} - z^T|, \qquad z = \pi, x. \qquad \textbf{(18)}$$

Here, γ is a persistence parameter that determines how fast expectations can become more or less anchored, and multiplying by 400 converts the data to annualized percentage points. The time t anchoring of expectations, H_t^z, is obtained from LAD_t^z by first multiplying by 4, and then truncating (flooring) to the nearest odd integer, with a minimum of 1.[11] That is,

$$H_t^x = \max(floor_{odd}(4 * LAD_t^x), 1), \tag{19}$$

$$H_t^\pi = \max(floor_{odd}(4 * LAD_t^\pi), 1), \tag{20}$$

with

$$floor_{odd}(z) = \begin{cases} floor(z) & \text{if } floor(z) \text{ is odd} \\ floor(z) - 1 & \text{if } floor(z) \text{ is even} \end{cases}. \tag{21}$$

Note that, with the anchoring of expectations evolving according to (18)–(20), expectations would eventually become **fully anchored** in the target steady state ($H_t^x = H_t^\pi = 1$). When the model is subject to shocks, however, the economy never fully reaches this steady state, but instead inflation and output gap keep fluctuating around their targets. In this case, expectations never become fully anchored, and there will *always* be heterogeneity in expectations.[12]

4.3.1. Simulation with different initial anchoring of expectations. We now simulate the above model for different levels of the initial anchoring of expectations. With this simulation exercise we can investigate whether liquidity traps can arise, and how this depends on the anchoring of expectations.

We use the calibration of Woodford (1999) (with $\sigma = 0.157$, $\kappa = 0.024$, and $\beta = 0.99$), and set the persistence in the fitness measure to $\rho = 0.5$. Shocks to inflation and output gap are white noise and have an annualized standard deviation of 1.5%. In the interest rate rule we set $\pi^T = 0$, $\phi_1 = 1.5$, and $\phi_2 = \sigma = 0.157$. We further set $\gamma = 0.995$ so that expectations can become more or less anchored endogenously, but that this happens at a much slower rate than the switching of agents between expectation values. This reflects the intuitive assumption that expectations remain relatively anchored when shocks drive inflation and output gap far away from their targets for only a couple of periods, but that the range of expectation values that agents consider to be realistic starts increasing as variables significantly deviate from their targets for *longer* periods of time.

We calibrate the intensity of choice at $\omega = 63,500$ to let the model match expectations from survey data.[13] At this calibration, the interquartile range of the expectations distribution in the fundamental steady state is 1% (in annualized terms). Outside of this steady state the interquartile range then typically is 1.5%, and less for realized values close to the highest or lowest possible prediction value. This is in line with the findings of Mankiw et al. (2003), who show that the interquartile range of expectations in the Livingston Survey and the Survey of Professional Forecasters is around 1%.

To illustrate what this intensity of choice implies for the expectation values that agents will choose, we plot two representative cross-sectional distributions
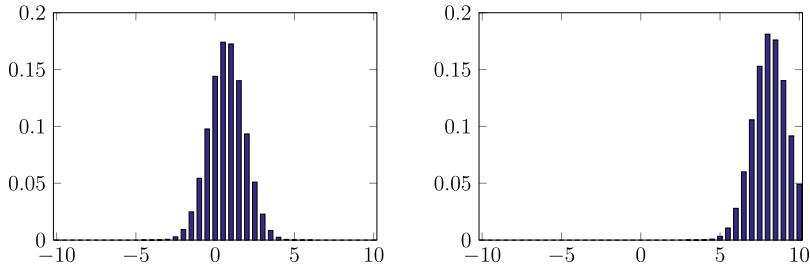
FIGURE 3. Cross-sectional distribution of inflation expectations over the different discrete expectation values at two representative time periods. In the left panel past inflation was close to the target, while in the right panel past inflation was close to the upper bound on expectation values of 10%.

of inflation expectations in Figure 3. The left panel corresponds to a point in time where past inflation was close to the target steady state. The right panel depicts the distribution in a time period where past inflation was close to the upper bound on expectations, determined by the anchoring of expectations that is here set at $H_t^x = H_t^\pi = 41$.

In the figure, it can be seen that considerable heterogeneity is always present for this calibration of the intensity of choice and that expectations in principle follow a normal distribution. However, when expectations come close to their highest or lowest possible values, given the current anchoring of expectations, the distribution of expectations becomes skewed.

Figure 4 shows time series of inflation and output gap (top two panels) for two different values of initial anchoring. H_t^π and H_t^x evolve over time and are plotted in the middle two panels. Green curves depict the case where the anchoring of expectations start out at $H_0^\pi = H_0^x = 41$, while the purple curves depict the case of initial anchoring of $H_0^\pi = H_0^x = 33$. That is, either agents initially consider expectation values up to 10% above and 10% below the target, or expectations are somewhat more anchored, and agents only consider values between -8% and 8% around the target. We deliberately take two initial conditions that are relatively close to each other and are both relatively unanchored, in order to illustrate that a small difference in initial anchoring of expectations can have a large impact on subsequent dynamics. Finally, the nominal interest rate is plotted in the bottom panel of Figure 4.

In Figure 4, it can be seen that the economy starts out with low inflation and high output gap.[14] Subsequently, both variables revert back to their target and stay close to it for a while. This induces inflation expectations to become *more anchored* (LAD_t^π and H_t^π decrease), as can be seen in the middle left panel.

Around period 20, some bad shocks hit the economy, which drive down inflation. Almost self-fulfilling expectations then induce a wave of pessimism about inflation and optimism about output gap. However, at some point (after period 25), the low inflation leads to a binding ZLB, and the nominal interest rate is set to zero. As inflation falls even more, the *real interest rate* becomes higher and
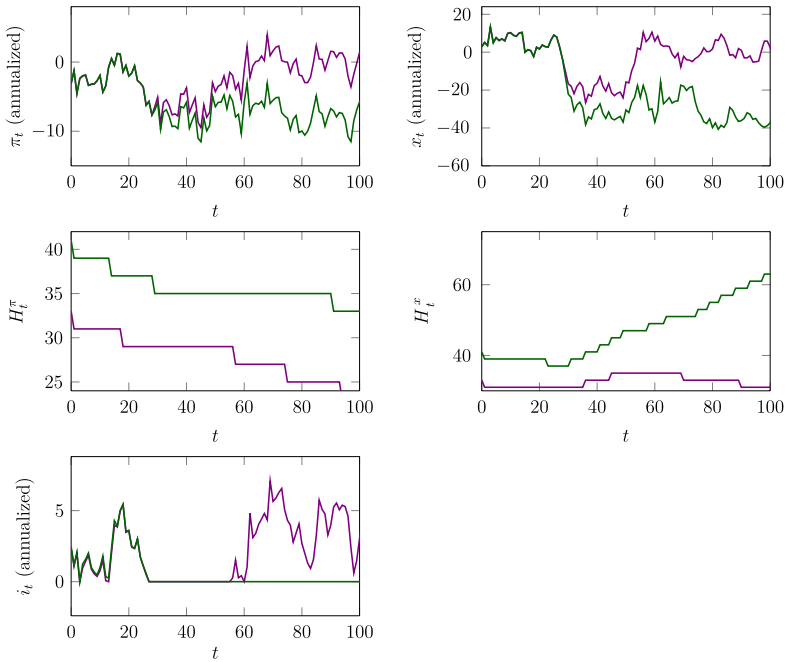
FIGURE 4. Simulated time series of model with endogenous anchoring of expectations and ZLB. The purple and green curves represent time series with different initial anchoring of expectations, H_0^π and H_0^x.

higher, leading output gap to start falling as well. Since there is no feedback from monetary policy anymore, pessimistic expectations now are self-reinforcing.

Inflation and output gap keep falling until output gap expectations start to hit their lowest possible value (around period 30). The economy now is stuck in a *liquidity trap steady state*, both in case of the purple and of the green curve. In the case of less anchored expectations (the green curves), output gap expectations can fall more however, leading to lower realizations of output gap and inflation than in the case of more anchored initial expectations (purple curves). As the liquidity trap with low inflation and output gap continues to last, output gap expectations start becoming less anchored (LAD_t^x and H_t^x increase), as can be seen in the middle right panel of Figure 4.

Then, around period 50, favorable shocks start to hit the economy, and this is where the two curves really start to differ. In the case of *more* anchored expectations (purple curves), where inflation and output gap have not become too low yet, these shocks are enough to get the economy out of the liquidity trap. Inflation and output and their expectations now become high enough for the model to revert back to the positive interest rate region, from which monetary policy can achieve convergence back to the target. As this happens, expectations become more and more anchored (middle panels).
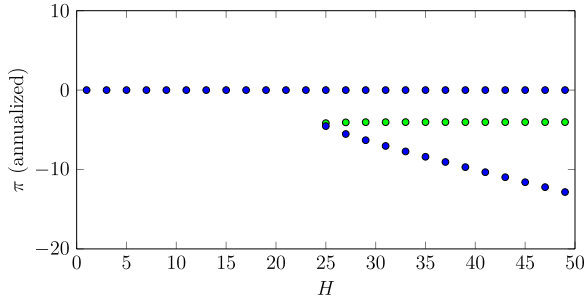
FIGURE 5. Stable target steady state (top blue dots) and stable liquidity trap steady state (bottom blue dots), for different levels of the anchoring of expectations. The unstable steady state in the middle (green dots) separates their basins of attraction.

In case of *less* anchored expectations (green curves), the favorable shocks also increase inflation and output gap. However, here the liquidity trap already was too deep, so that the ZLB remains binding, and both variables quickly start to fall again. During the remaining periods, the economy keeps fluctuation around the liquidity trap steady state, which has lower and lower inflation and output gap values as output gap expectations keep becoming more unanchored (middle right panel). It therefore becomes more and more difficult for favorable shocks, or policy interventions, to get the economy out of this liquidity trap.

The key policy insight that can be obtained from this simulation exercise is the following. Since the initial anchoring of expectations at the time bad shocks hit the economy is crucial for eventual recovery of a liquidity trap, the central bank must make sure that this initial anchoring is strong enough. It can do so by making sure that in the years before the bad shocks hit, inflation and output gap are stabilized close to their targets. Since the ZLB was not yet a binding constraint during these periods, this can be achieved with adequate conventional monetary policy.

4.3.2. Liquidity traps and anchoring of expectations. To gain more insight in the role of the anchoring of expectations and to provide robustness to the above simulation exercise, we now turn to a bifurcation analysis. Figure 5 plots the steady-state values of both the target steady state (upper blue dots) and the liquidity trap steady state (lower blue dots), for different fixed values of H_t^π and H_t^x. Additionally, we plot the unstable steady state (green dots), lying between these two steady states. This unstable steady state separates the **basin of attraction** of the target steady state from that of the liquidity trap steady state.[15] That is, in the absence of shocks, inflation and output gap on one side of the unstable steady state imply convergence to the target steady state, while initial inflation and output gap on the other side of the unstable steady state imply convergence to the liquidity trap steady state.

To keep the graphical representation clear, we do not show the output gap value of the steady state and put its inflation value on the vertical axis of Figure 5. Additionally, we vary H_t^π and H_t^x together and show on the horizontal axis $H = H_t^\pi = H_t^x$. This allows us to keep the plot two-dimensional, without loss of generality.

The reason that it is insightful to look at the existence and basins of attractions of steady states for fixed values of H_t^π and H_t^x is that (by construction) the anchoring of expectations evolves considerably slower than inflation and output gap dynamics, and is not immediately effected by shocks. In the short run, the dynamics of our system can therefore be closely approximated by the dynamics that would arise for fixed anchoring of expectations. In the medium- to long run, one subsequently needs to consider that the basins of attractions of the different steady states will change as H starts to change. This can be done by considering shifts on the horizontal axis of Figure 5.

It can be seen in Figure 5 that for strong anchoring of expectations (low H), the target steady state is unique, and a persistent liquidity trap cannot arise. However, as expectations become less anchored, the liquidity trap steady state comes into existence ($H = 25$). It then still lies relatively close to the target steady state, and, more importantly, very close to the unstable steady state. This implies that small favorable shocks can bring inflation out of the liquidity trap, and above the unstable steady state, from which the system will converge to the target steady state. As expectations become even more unanchored, the liquidity trap steady state has a lower inflation value and lies farther away from the unstable steady state. Recovery due to favorable shocks is less likely in this case. This is exactly what we saw in Figure 4. In the purple time series, inflation was still close to the basin of attraction of the target steady state (i.e. close to the unstable steady state), and favorable shocks led to recovery. In the green time series, inflation was too far away from the unstable steady state, and favorable shocks did not bring the economy back to the basin of attraction of the target steady state.

The unstable steady state in Figure 5 lies very close to the negative of the steady state real interest rate: $-\bar{r}$. This is no coincidence, and the intuition is that the (log-linearized) period t real interest rate (given by $i_t - \bar{r} - \bar{E}_t \pi_{t+1}$) is positive under a binding ZLB if and only if $\bar{E}_t \pi_{t+1} < -\bar{r}$.[16] It follows from equation (1) that when this is the case, output gap realizations will be lower than output gap expectations. It then follows from (2) that for reasonable calibrations of κ and β, the same holds for inflation. This implies that only if inflation expectations are low enough compared to the steady-state real interest rate, pessimistic expectations will be *self-reinforcing* and hence lead to convergence to the liquidity trap steady state. A necessary condition for inflation expectations to become low enough is that H is large enough, so that the lowest possible inflation value that agents consider is lower than $-\bar{r}$. If expectations are more anchored (lower H) and agents do not consider inflation values below $-\bar{r}$, then $\bar{E}_t \pi_{t+1} < -\bar{r}$ can never hold and the liquidity trap steady state *does not exist*. Hence, if the central bank manages to anchor expectations strongly enough to its targets (by conducting aggressive
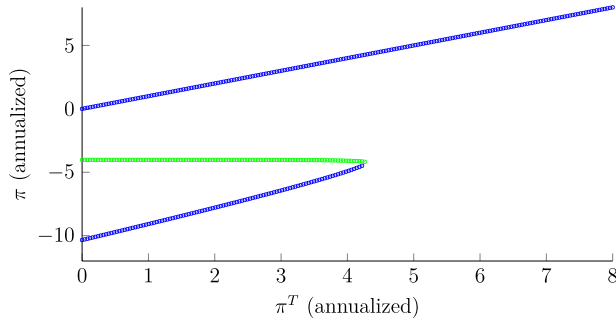
FIGURE 6. Bifurcation diagram in π^T for $H_t^{\pi} = H_t^x = 41$. The upper blue curve represents the fundamental steady state, and the lower blue curve the liquidity trap steady state. The green curve depicts the unstable steady state that separates their basins of attraction.

enough monetary policy) during normal times, it can *exclude* to possibility of bad shocks driving the economy into a self-reinforcing liquidity trap.

4.3.3. Increasing the inflation target. Since agent's expectations are distributed around the targets, another way the central bank could reduce the likelihood (or even exclude the possibility) of a self-reinforcing liquidity trap is by choosing a higher inflation target. For a given anchoring of expectations, the *most pessimistic expectations* that agents consider would then be higher. This would make it less likely that $\bar{E}_t \pi_{t+1} < -\bar{r}$ will hold, and that the economy enters a liquidity trap.

In Figure 6, we plot the target steady state (top blue curve) and the liquidity trap steady state (bottom blue curve) as a function of the inflation target (π^T), for the case of $H = 41$. Note that, as the inflation target is chosen higher, the target steady state has an equivalently higher inflation level as well, making the top blue curve upward sloping. Since agents consider 10% below the inflation target as their lowest possible value, the bottom blue line is also upward sloping. Figure 6 also plots the unstable steady state (green curve) separating the basin of attraction of the liquidity trap from that of the target steady state.

In Figure 6, it can be seen that for a zero inflation target, the liquidity trap steady state is associated with a quite low level of inflation and lies far away from the basin of attraction of the target steady state. As the inflation target is increased, the liquidity trap steady state comes closer to the unstable steady state (green curve), making it more likely that shocks bring the economy from the liquidity trap steady state to the basin of attraction of the target steady state. As the inflation target is increased even more, the liquidity trap steady state disappears altogether, and convergence to the target steady state *always* occurs. Note, however, that if we would let H vary endogenously as in Figure 4, a high inflation target can only temporarily exclude the possibility of a liquidity trap. If, for some reason (e.g., bad shocks), inflation realizations remain low for a long period of time, expectations would become *less anchored*, and agents might consider expectation

values that satisfy $\bar{E}_t \pi_{t+1} < -\bar{r}$, even when the inflation target is high. In that case, the liquidity trap steady state would exist again.

The above result is similar to findings under rational expectations, where liquidity traps can be prevented by choosing the inflation target high enough, so that large shocks in the model do not lead to a severely binding ZLB and a high real interest rate.[17] Our model, however, emphasizes the role of expectations rather than that of exogenous shocks. Low expectations might come about because of large shocks but could also arise slowly because of declining confidence after a period where small negative shocks were relatively more prevalent.

4.3.4. Shock size. Besides the anchoring of expectations and the inflation target, another important determinant of emerging dynamics is the size of the shocks that hit the economy. When shocks are very small, the ZLB will never become binding and the model never leaves the basin of attraction of the target steady state, even when expectations are unanchored and the inflation target is low. In this case, expectations will become (and remain) strongly anchored to the targets of the central bank.

When on the other hand, the economy is continuously hit by larger shocks, the model will switch back and forth between extended periods of time near the target steady state and extended periods of time near the liquidity trap steady state, similar to the purple curve in Figure 4. In that case, expectations become less anchored during liquidity trap episodes but then become more anchored again during episodes near the target steady state.

If, however, the shocks that hit the economy are *very* large, then even when the model is near the target steady state, fluctuations in output and inflation are large, so that expectations never become very anchored (not even when monetary policy responds aggressive to inflation expectations). As a consequence, the cumulative effect of time spent around the liquidity trap steady state leads expectations to become slowly less and less anchored over time. This is illustrated in Figure 7, which plots a time series simulation as in Figure 4, but now over 4000 periods and with shocks with an annualized standard deviation of 8%. Moreover, the initial anchoring in this simulation is set at $H_0^\pi = H_0^x = 3$ to illustrate that under such large shocks expectations endogenously become unanchored (ultimately unboundedly so), even if they initially are very anchored to the central bank's targets.

As can be seen in the figure, in the first 700 periods expectations become more and more unanchored. Moreover, because the liquidity trap steady state still lies relatively close to the basin of attraction of the target steady state (relative to the shock sizes), the economy continuously switches between the ZLB region and the positive interest region but does not stay at the ZLB for longer periods of time.

Next, the anchoring of expectations roughly stabilizes at $H_t^\pi = 40$ and $H_t^x = 150$. With these unanchored expectations, the liquidity trap steady state lies far enough away from the basin of attraction of the target steady state for prolonged liquidity traps to arise. As mentioned above, expectations become less anchored
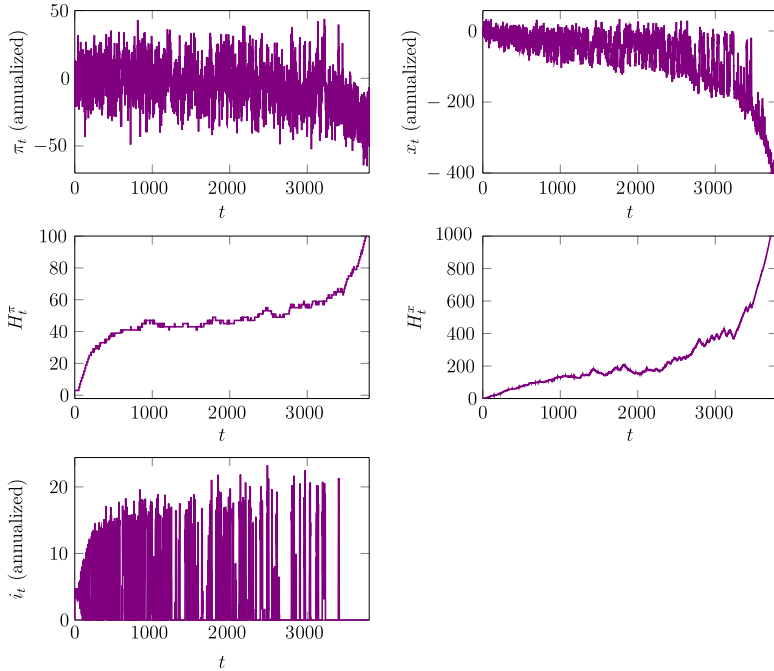
FIGURE 7. Simulated time series of model with endogenous anchoring of expectations and ZLB in case of very large shocks (annualized standard deviation of 8%).

during such liquidity trap episodes but then more anchored again during episodes where the model fluctuates around the target steady state.

However, for this simulation with very large shocks, expectations become increasingly unanchored from approximately period 2500 onward as longer liquidity traps arise. These more unanchored expectations increase the likelihood of further long-lasting liquidity traps, reinforcing the decrease in anchoring of expectations. Eventually expectations become so unanchored that the liquidity trap steady state lies too far from the basin of attraction of the target steady state for even a series of large positive shocks to bring the economy back to the target steady state, and a deflationary spiral arises.

5. CONCLUSION

We presented a model where expectations are heterogeneous. The range of expectations that agents are considering in this model can be interpreted as the *anchoring* of expectations. We showed that multiple steady states can arise. First of all, non-fundamental steady states can arise because of the discrete choice set of expectation values that is consistent with reported expectations found in survey data and in laboratory experiments. Multiple steady states then reflect

coordination on a round number that is possible when expectations are almost self-fulfilling. This can be prevented with a more aggressive response of monetary policy to inflation expectations, which drives inflation and output gap faster toward their targets for a given distribution of private sector expectations.

Secondly, multiple steady states can arise due to the anchoring of expectations. In particular, when monetary policy is restricted by the ZLB and is unable to stimulate output gap and inflation, pessimistic expectations can become self-reinforcing. This leads the distribution of agent's expectations to move toward the lowest possible inflation and output gap values that agents are currently considering. When this happens, expectations start to become less anchored, so that agents consider a wider range of values around the target. The liquidity trap steady state then shifts downward. This makes it less likely that favorable shocks or policy interventions bring the economy back to the basin of attraction of the fundamental steady state, and hence makes it less likely that the economy can recover from the liquidity trap.

We further showed that whether the economy can recover from a liquidity trap crucially depends on how unanchored expectations were when the bad shocks hit, and that a small difference in initial anchoring of expectations can make a large qualitative difference for subsequent aggregate dynamics. If shocks are, however, very large, then it is unavoidable that expectations eventually become very unanchored and the economy ends up in a deflationary spiral. Finally, it is found that the central bank can reduce the likelihood of the occurrence of self-reinforcing liquidity traps by increasing its inflation target.

NOTES

1. Evans et al. (2016) and Arifovic et al. (2018) also find coordination on a liquidity trap steady state rather than a deflationary spiral. This is, however, not because of anchored expectations. Instead, Evans et al. (2016) assume an exogenous lower bound on inflation and consumption, while Arifovic et al. (2018) assume a learning process that makes the well-known ZLB steady state discussed by Benhabib et al. (2001a, 2001b) locally stable.

2. In the derivations in Hommes and Lustenhouwer (2019a), the aggregation of individual decisions in terms of aggregates only [which is not done in Kurz et al. (2013)] hinges on a specific property of the way we model heterogeneous expectations. With the Heuristic switching model of Brock and Hommes (1997), in any period there is heterogeneity in expectations, with different fractions of agents using different heuristics. However, these fractions are updated in each period according to a probability distribution that depends on the relative past performance of each expectation formation heuristic. With this updating process, each agent has in each period the same probability of following a particular heuristic as all other agents, independent of the heuristic the particular agent followed in the previous period. By assuming that agents are aware of this, their expectations about their own future consumption coincide with their expectation about the future consumption of any other agent, and therefore with their expectations about aggregate consumption, which is crucial for aggregation. See Hommes and Lustenhouwer (2019a) for details.

3. Note that the nominal interest rate is here not defined in terms of deviation from steady state (which would give $\hat{i}_t = i_t - \bar{r}$). This makes the ZLB analysis more intuitive.

4. The exception to perfect coordination arises in the knife edge case where fundamentalists have exactly the same fitness as optimists (pessimists). In this case, half of the agents will become

fundamentalists, while the other half become optimists (pessimists). These knife edge cases could also comprise steady states. However, in Appendix B it is shown that such steady states are always unstable. The intuition for the existence of these knife edge unstable steady states is that they separate the basins of attraction of coexisting locally stable steady states. That is, initial conditions to one side of the unstable steady state imply convergence to one locally stable steady state, while initial conditions to the other side of the unstable steady state imply convergence to another locally stable steady state.

5. For *finite* intensity of choice there would not be perfect coordination, implying that during a wave of pessimism also some agents are fundamentalists and optimists. These agents would put upward pressure on inflation and output gap, driving realizations further away from the expectations of pessimists, making them less self-fulfilling. In this case, the pessimistic steady state can still exist, but the condition for its existence would become more stringent. This is illustrated in Lustenhouwer (2017).

6. The other seven non-fundamental steady states can also be ruled out with an aggressive enough response to inflation and/or output gap. Conditions for existence of these steady states can be found in Lustenhouwer (2017). In particular, the condition for existence of an optimistic steady state is exactly the same as the one given in Proposition 2, since (abstracting from the ZLB) the model is completely symmetric.

7. It would require $\phi_1 + \phi_2 < 1$.

8. See, for example, Baker (1992), Edouard and Senthilselvan (1997), and Curtin (2010).

9. Having expectations that are spaced closer together would make conditions for existence of multiple steady states more stringent. There would, however, continue to exist a range of policy parameters for which expectations are sufficiently self-fulfilling for multiple steady states to exist. In the limit where expectation values are infinitely close to each other (the large type limit), this range of policy parameters is reduced to a single bifurcation point (see Hommes and Lustenhouwer (2019b)). Similar findings in a different model are obtained by Guse and Carton (2014).

10. Note that the qualitative results of this subsection do not depend on the value at which we fix H^x and H^π. For less anchored expectations, extra steady states exist around $\phi_1 = 0.95$ in addition to the ones plotted in Figure 2. For more anchored expectations, less steady states exist here. Moreover, as ϕ_1 becomes lower, the two most extreme values will always comprise steady states, and only the *values* of these steady states depend on the values of H^x and H^π, as illustrated in Section 4.3.

11. It is multiplied by 4, because an extra percentage point in LAD_t^z implies a percentage point in the positive and a percentage point in the negative direction, and hence four steps of 0.5%. Truncating to the nearest *odd* integer ensures that symmetry in expectation values is maintained.

12. Also note that the fractions of agents expecting different expectation values evolve according to a probability distribution, where in any period each agent has a chance of choosing a particular value. It is therefore not necessarily the case that there are individual agents who are either constantly too optimistic or constantly too pessimistic and hence make persistent forecast errors. Instead, some agents will be too optimistic for a while and then too pessimistic, while other agents are first too pessimistic and then too optimistic.

13. Note that the magnitude of the intensity of choice depends on the units of measurement of the data. Since a 1% deviation of quarterly inflation from steady state is measured as 0.01 and results in a squared forecast error of 0.0001, an intensity of choice of 63, 500 should not be considered particularly large. Moreover, as discussed in Hommes and Lustenhouwer (2019b) where the intensity of choice is of similar magnitude, this level of the intensity of the choice is, controlling for the units of measurement in the fitness measure, comparable to the empirical estimates of Cornea-Madeira et al. (2019).

14. We first run our model for 50 periods to initialize it (this is necessary for the fitness measures). The initial levels of inflation and output gap are therefore random.

15. More precisely, the unstable steady state is a saddle-point, and the basins of attraction of the two stable steady states are separated by the stable manifold of the saddle-point steady state.

16. In Hommes and Lustenhouwer (2019*b*), we analytically show for a model specification where there is a continuum of expectation values that the steady-state inflation value exactly becomes $-\bar{r}$ as expectations become unanchored.

17. See, for example, Ball (2013).

REFERENCES

Agliari, A., D. Massaro, N. Pecora and A. Spelta (2017) Inflation targeting, recursive inattentiveness, and heterogeneous beliefs. *Journal of Money, Credit and Banking* 49(7), 1587–1619.

Agliari, A., N. Pecora and A. Spelta (2015) Coexistence of equilibria in a new keynesian model with heterogeneous beliefs. *Chaos, Solitons & Fractals* 79, 83–95.

Andrade, P., R. K. Crump, S. Eusepi, E. Moench, et al. (2013) Noisy information and fundamental disagreement. *Staff Reports, Federal Reserve Bank of New York* 655.

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

Arifovic, J., S. Schmitt-Grohé and M. Uribe (2018) Learning to live in a liquidity trap. *Journal of Economic Dynamics and Control* 89, 120–136.

Aruoba, Borağan, 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.

Assenza, T., P. Heemeijer, C. Hommes and D. Massaro (2014) Managing self-organization of expectations through monetary policy: A macro experiment, Technical report, CeNDEF Working Paper, University of Amsterdam.

Baker, M. (1992) Digit preference in CPS unemployment data. *Economics Letters* 39(1), 117–121.

Ball, L. M. (2013) The case for four percent inflation. *Central Bank Review* 13(2), 17.

Benhabib, J., G. W. Evans and S. Honkapohja (2014) Liquidity traps and expectation dynamics: Fiscal stimulus or fiscal austerity? *Journal of Economic Dynamics and Control* 45, 220–238.

Benhabib, J., S. Schmitt-Grohé and M. Uribe (2001a) Monetary policy and multiple equilibria. *American Economic Review* 91(1), 167–186.

Benhabib, J., S. Schmitt-Grohé and M. Uribe (2001b) The perils of taylor rules. *Journal of Economic Theory* 96(1), 40–69.

Blinder, A. S., M. Ehrmann, M. Fratzscher, J. De Haan and D.-J. Jansen (2008) Central bank communication and monetary policy: A survey of theory and evidence. *Journal of Economic Literature* 46(4), 910–945.

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

Branch, W. A. and B. McGough (2010) Dynamic predictor selection in a new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control* 34(8), 1492–1508.

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

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

Cole, S. J. and F. Milani (2019) The misspecification of expectations in new keynesian models: A DSGE-VAR approach. *Macroeconomic Dynamics* 23(3), 974–1007.

Cornea-Madeira, A., C. Hommes and D. Massaro (2019) Behavioral heterogeneity in US inflation dynamics. *Journal of Business & Economic Statistics* 37(2), 288–300.

Curtin, Richard (2010), Inflation expectations and empirical tests: Theoretical models and empirical tests. In: *Inflation Expectations*, pp. 52–79. New York: Routledge.

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

De Grauwe, P. and Y. Ji (2019) Inflation targets and the zero lower bound in a behavioral macroeconomic model. *Economica* 86(342), 262–299.

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

Dincer, N. N. and B. Eichengreen (2014) Central bank transparency and independence: Updates and new measures. *International Journal of Central Banking* 10(1), 189–259.

Edouard, L. and A. Senthilselvan (1997) Observer error and birthweight: Digit preference in recording. *Public Health* 111(2), 77–79.

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

Evans, G. W., E. Guse and S. Honkapohja (2008) Liquidity traps, learning and stagnation. *European Economic Review* 52(8), 1438–1463.

Evans, G. W., S. Honkapohja and K. Mitra (2016) Expectations, stagnation and fiscal policy. Bank of Finland Research Discussion Paper 25.

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

Guse, E. A. and J. Carton (2014) Replicator dynamic learning in Muth's model of price movements. *Macroeconomic Dynamics* 18(3), 573–592.

Hommes, C. H. and J. Lustenhouwer (2019a) Inflation targeting and liquidity traps under endogenous credibility. *Journal of Monetary Economics, forthcoming*.

Hommes, C. H. and J. Lustenhouwer (2019b) Managing unanchored, heterogeneous expectations and liquidity traps. *Journal of Economic Dynamics and Control* 101, 1–16.

Kurz, M., G. Piccillo and H. Wu (2013) Modeling diverse expectations in an aggregated new keynesian model. *Journal of Economic Dynamics and Control* 37(8), 1403–1433.

Lustenhouwer, J. (2017) Monetary and fiscal policy under bounded rationality and heterogeneous expectations, Technical report, Ph.D. dissertation at the University of Amsterdam.

Mankiw, N. G., R. Reis and J. Wolfers (2003) Disagreement about inflation expectations, Technical report, National Bureau of Economic Research.

Manski, C. F., D. McFadden, et al. (1981) *Structural Analysis of Discrete Data with Econometric Applications*. Mit Press Cambridge, MA.

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

Mertens, K. R. S. M. and M. O. Ravn (2014) Fiscal policy in an expectations-driven liquidity trap. *Review of Economic Studies* 81(4), 1637–1667.

Pecora, N. and A. Spelta (2017) Managing monetary policy in a new keynesian model with many beliefs types. *Economics Letters* 150, 53–58.

Pfajfar, D. and B. Žakelj (2018) Inflation expectations and monetary policy design: Evidence from the laboratory. *Macroeconomic Dynamics* 22(4), 1035–1075.

Svensson, L. E. O. (2009) Transparency under flexible inflation targeting: Experiences and challenges. No. 7213. CEPR Discussion Papers.

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

Woodford, M. (2003) *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton: Princeton University Press.

Woodford, M. (2005) Central bank communication and policy effectiveness, Technical report, National Bureau of Economic Research.

APPENDIX A: PROOF OF PROPOSITIONS

APPENDIX A.1: Proof of Proposition 1

Fitness measures are given by $U_{t-1}^{\pi,h} = -(\pi_{t-1} - E_{t-2}^h \pi_{t-1})^2$ for inflation, and $U_{t-1}^{x,h} = -(x_{t-1} - E_{t-2}^h x_{t-1})^2$ for output gap. Plugging in the expectations of the three types in these fitness measures, it follows that when $\pi_{t-1} = \pi^T$ and $x_{t-1} = x^T$, optimists and pessimists have identical fitness ($U_{t-1}^{x,opt} = U_{t-1}^{x,pes}$ and $U_{t-1}^{\pi,opt} = U_{t-1}^{\pi,pes}$). It then follows from (6) that in this case $n_t^{x,opt} = n_t^{x,pes}$ and $n_t^{\pi,opt} = n_t^{\pi,pes}$, so that (9) and (10) reduce to $\bar{E}_t x_{t+1} = x^T$

and $\bar{E}_t \pi_{t+1} = \pi^T$. Plugging these expectations in (4) and (5), realized inflation and output gap will, in the absence of shocks, be equal to $\pi_t = \pi^T$ and $x_t = x^T$. This will then be true in any period, so that the targets of the central bank comprise a steady state.

APPENDIX A.2: Proof of Proposition 2

In a steady state where all agents are pessimists, (4) and (5) reduce to

$$x_t = (1 - \frac{\phi_2}{\sigma})(x^T - b) - \frac{\phi_1 - 1}{\sigma}(\pi^T - b - \pi^T) + \frac{\phi_2}{\sigma}x^T, \tag{A.1}$$

$$\pi_t = \beta \bar{E}_t(\pi^T - b) + \kappa x_t. \tag{A.2}$$

Assume $\omega = \infty$. Then the steady state exists if and only if both output gap and inflation are such that in the next period again all agents are pessimists about both variables. This requires that both $x_t < x^T - \frac{b}{2}$ and $\pi_t < \pi^T - \frac{b}{2}$. Rewriting (A.1), the condition for output gap reduces to

$$x_t = x^T + b(\frac{\phi_2}{\sigma} - 1 + \frac{\phi_1 - 1}{\sigma}) < x^T - \frac{b}{2}, \tag{A.3}$$

which gives

$$\phi_1 < 1 + \frac{\sigma}{2} - \phi_2. \tag{A.4}$$

If this condition is satisfied (and thereby $x_t < x^T - \frac{b}{2}$), then it follows from (A.2) that

$$\pi_t < \beta \bar{E}_t(\pi^T - b) + \kappa(x^T - \frac{b}{2}). \tag{A.5}$$

This implies that for $2\beta + \kappa > 1$, $\pi_t < \pi^T - \frac{b}{2}$ is automatically satisfied. Since κ is positive, and it is reasonable to consider only calibrations with the discount factor satisfying $\beta > \frac{1}{2}$, we can conclude that when (A.2) holds, both $x_t < x^T - \frac{b}{2}$ and $\pi_t < \pi^T - \frac{b}{2}$, so that the pessimistic steady state exists. If (A.2) does not hold, then output gap expectations (and possibly also inflation expectations) will not stay pessimistic, so that the pessimistic steady state does not exist. Local stability of the steady state under $\omega = \infty$ is proven in Appendix B.

APPENDIX A.3: Proof of Proposition 3

In a steady state where all agents are pessimists, (11) and (12) reduce to

$$x_t = (x^T - b) + \frac{1}{\sigma}(\pi^T - b) + \frac{\bar{r}}{\sigma}, \tag{A.6}$$

$$\pi_t = \beta(\pi^T - b) + \kappa x_t. \tag{A.7}$$

Analogously to Appendix A.2, it follows that under these equations $x_t < x^T - \frac{b}{2}$ if and only if

$$\pi^T + \bar{r} < b(1 + \frac{\sigma}{2}), \tag{A.8}$$

and that $\pi_t < \pi^T - \frac{b}{2}$ then holds as well (assuming $2\beta + \kappa > 1$).

The steady state is, however, only consistent with equations (11) and (12) if it lies in the ZLB region of the model. This is the case if and only if

$$i_t = \pi^T + \bar{r} + \phi_1(\pi^T - b - \pi^T) + \phi_2(x^T - b - x^T) < 0, \qquad \text{(A.9)}$$

or equivalently

$$\pi^T + \bar{r} < b(\phi_1 + \phi_2). \qquad \text{(A.10)}$$

APPENDIX A.4: Proof of Proposition 4

Inflation and output gap are again given by (A.7) and (A.6). Now it is required, however, that $x_t < x^T - b$ and that $\pi_t < \pi^T - b$. The first inequality reduces to $\pi^T + \bar{r} < b$. If we assume that $\beta + \kappa > 1$, which will hold for most reasonable calibrations, $\pi_t < \pi^T - b$ then holds as well.

APPENDIX B: STABILITY 3-TYPE MODEL

The local stability properties of the 3-type model depend on the derivative of expectations with respect to lagged values of inflation and output gap. In particular, the Jacobian of the model is given by

$$\begin{pmatrix} (1 - \frac{\phi_2}{\sigma}) \frac{\partial \bar{E}_t x_{t+1}}{\partial x_{t-1}} & -\frac{\phi_1 - 1}{\sigma} \frac{\partial \bar{E}_t \pi_{t+1}}{\partial \pi_{t-1}} \\ \kappa(1 - \frac{\phi_2}{\sigma}) \frac{\partial \bar{E}_t x_{t+1}}{\partial x_{t-1}} & (\beta - \kappa \frac{\phi_1 - 1}{\sigma}) \frac{\partial \bar{E}_t \pi_{t+1}}{\partial \pi_{t-1}} \end{pmatrix},$$

Next, we use (9) and (10) to write

$$\frac{\partial \bar{E}_t x_{t+1}}{\partial x_{t-1}} = b \left(\frac{\partial n_t^{x,opt}}{\partial x_{t-1}} - \frac{\partial n_t^{x,pes}}{\partial x_{t-1}} \right), \qquad \text{(B.1)}$$

and

$$\frac{\partial \bar{E}_t \pi_{t+1}}{\partial \pi_{t-1}} = b \left(\frac{\partial n_t^{\pi,opt}}{\partial \pi_{t-1}} - \frac{\partial n_t^{\pi,pes}}{\partial \pi_{t-1}} \right). \qquad \text{(B.2)}$$

The derivative of $n_t^{x,opt}$ with respect to x_{t-1} is given by

$$\frac{(2\omega b - 2\omega x_{t-1})e^{-\omega(x_{t-1} - x^T - b)^2}}{e^{-\omega(x_{t-1} - x^T - b)^2} + e^{-\omega(x_{t-1} - x^T + b)^2} + e^{-\omega(x_{t-1} - x^T)^2}}$$
$$- \frac{e^{-\omega(x_{t-1} - x^T - b)^2}(2\omega b - 2\omega x_{t-1})e^{-\omega(x_{t-1} - x^T - b)^2}}{(e^{-\omega(x_{t-1} - x^T - b)^2} + e^{-\omega(x_{t-1} - x^T + b)^2} + e^{-\omega(x_{t-1} - x^T)^2})^2}$$
$$+ \frac{e^{-\omega(x_{t-1} - b)^2}((2\omega b + 2\omega x_{t-1})e^{-\omega(x_{t-1} - x^T + b)^2} + 2\omega x_{t-1}e^{-\omega(x_{t-1} - x^T)^2})}{(e^{-\omega(x_{t-1} - x^T - b)^2} + e^{-\omega(x_{t-1} - x^T + b)^2} + e^{-\omega(x_{t-1} - x^T)^2})^2}. \qquad \text{(B.3)}$$

This can be written as

$$\frac{\partial n_t^{x,opt}}{\partial x_{t-1}} = 2\omega b n_t^{x,opt}(1 - n_t^{x,opt} + n_t^{x,pes}). \qquad \text{(B.4)}$$

Similarly, the derivative of $n_t^{x,pes}$ with respect to x_{t-1} can be written as

$$\frac{\partial n_t^{x,pes}}{\partial x_{t-1}} = -2\omega b n_t^{x,pes}(1 - n_t^{x,pes} + n_t^{x,opt}). \tag{B.5}$$

The derivatives of $n_t^{\pi,opt}$ and $n_t^{\pi,pes}$ with respect to π_t are obtained by replacing x_{t-1} by π_{t-1} and x^T by π^T in (B.3) and by replacing $n_t^{x,opt}$ and $n_t^{x,pes}$ by $n_t^{\pi,opt}$ and $n_t^{\pi,pes}$ in (B.4) and (B.5).

When the intensity of choice goes to infinity, there are two possibilities for the expectations of each variable in steady state. Either all agents adhere to one heuristic, because this heuristic performs best in the steady state, or half of the agents adheres to one heuristic, while the other half adheres to another heuristic. In the second case, the two heuristics must perform equally well in steady state.

It follows from (B.4) and (B.5) that in a steady state where all agents adhere to one heuristic for both inflation and output gap (not necessarily the same heuristic), we have

$$\frac{\partial n_t^{x,opt}}{\partial x_{t-1}} = \frac{\partial n_t^{\pi,opt}}{\partial \pi_{t-1}} = \frac{\partial n_t^{x,pes}}{\partial x_{t-1}} = \frac{\partial n_t^{\pi,pes}}{\partial \pi_{t-1}} = 0. \tag{B.6}$$

This holds because either $(1 - n_t^{x,opt} + n_t^{x,pes})$ or $n_t^{x,opt}$ and $n_t^{x,pes}$ go to zero exponentially as ω goes to infinity (and analogously for inflation). This implies that $\frac{\partial \bar{E}_t x_{t+1}}{\partial x_{t-1}} = \frac{\partial \bar{E}_t \pi_{t+1}}{\partial \pi_{t-1}} = 0$ so that the Jacobian and both its eigenvalues reduce to zeros. Hence, all these steady state are locally stable.

For steady states where two heuristics about one variable perform equally well, at least one eigenvalue goes to infinity in absolute value when we let ω go to infinity so that these steady states are always unstable.