Monetary Policy under Behavioral Expectations: Theory and Experiment*

Cars Hommes[†] Domenico Massaro[‡] Matthias Weber[§]

November 6, 2017

Abstract

Expectations play a crucial role in modern macroeconomic models. We consider a New Keynesian framework under rational expectations and under a behavioral model of expectation formation. We show how the economy behaves in the alternative scenarios with a focus on inflation volatility. Contrary to the rational model, the behavioral model predicts that inflation volatility can be lowered if the central bank reacts to the output gap in addition to inflation. We test the opposing theoretical predictions in a learning-to-forecast experiment. The results support the behavioral model and the claim that output stabilization can lead to less volatile inflation.

JEL classification: C90, E03, E52, D84

Keywords: Experimental macroeconomics; Behavioral macroeconomics; Heterogeneous expectations; Learning-to-forecast experiment

---

*Thanks for comments and suggestions go to Arthur Schram, Mike Woodford, participants of the North American ESA meetings in Fort Lauderdale, the UCSD-Rady Workshop on Incentives and Behavior Change in Modica, the Workshop on Behavioral Macroeconomics in Amsterdam, the International Meeting on Experimental and Behavioral Social Sciences in Rome, the Maastricht Behavioral and Experimental Economics Symposium, the Workshop on Theoretical and Experimental Macroeconomics in Barcelona, the Annual Lithuanian Conference on Economic Research in Vilnius, the Computing in Economics and Finance Conference in Bordeaux, and seminar participants in Amsterdam, Mannheim, Marseille, Riga, Tilburg, and Vilnius. Financial support from the EU 7th framework collaborative project Complexity Research Initiative for Systemic InstabilitieS (CRISIS), grant no. 288501, from The Netherlands' Organisation for Scientific Research (NWO), grant no. 406-11-022, and from the Ministry of Education, Universities and Research of Italy (MIUR), program *SIR* (grant n. RBSI144KWH) are gratefully acknowledged.

[†]CeNDEF, Amsterdam School of Economics (University of Amsterdam) & Tinbergen Institute. Email: C.H.Hommes@uva.nl.

[‡]Università Cattolica del Sacro Cuore & Complexity Lab in Economics, Milan. Email: domenico.massaro@unicatt.it.

[§]CEFER, Bank of Lithuania & Faculty of Economics, Vilnius University. Email: mweber@lb.lt.

1 Introduction

Expectations play a crucial role in modern macroeconomic theory. Standard models used for scientific research and policy analysis typically assume a representative fully rational agent. However, the assumption that all agents in an economy are fully rational and able to determine the model-consistent expectation of the underlying process governing real-world economic outcomes is highly problematic. A great deal of research has shown that humans generally do not react fully rationally to the world around them. This research ranges from providing evidence for simple biases to showing the inability of humans to work with probabilities and to forecast future economic behavior (Tversky and Kahneman, 1974, and Grether and Plott, 1979, are seminal early contributions, and many have followed since; see Camerer et al., 2011 for an overview). Moreover, the claim based on evolutionary arguments that behavior deviating from the homogeneous rational expectations solution will be driven out of markets over time has not held up to scrutiny (Brock and Hommes, 1997, 1998, De Grauwe, 2012a; see also Arthur et al., 1997).

In this paper we consider a standard macroeconomic model under both rational and behavioral expectations. We examine aggregate macroeconomic behavior and policy implications arising from the alternative assumptions on expectation formation, paying particular attention to price stability. The behavioral model of expectation formation is a heuristic switching model that has been developed over a long period of time in which (mainly microeconomic) research has been conducted to investigate the question of how people form expectations and of how they adapt their ways of forming expectations when confronted with observed economic outcomes. Models of this kind perform well in describing expectation dynamics using both survey and experimental data (see Carroll, 2003, Frankel and Froot, 1987, Branch, 2004, Hommes, 2011, and Assenza et al., 2014b).

A key difference in outcomes between the macroeconomic models with rational and behavioral expectations concerns price stability, i.e. inflation volatility. Assuming rational expectations, there is a clear trade-off for a central bank between fighting inflation volatility and output gap volatility. If the central bank reacts to the output gap in addition to inflation, under rational expectations this will result in an increase of inflation volatility. The outcome is different under behavioral expectations. Starting from a situation in which the central bank does not react to the output gap at all, the central bank can simultaneously decrease inflation volatility and output gap volatility by reacting to the output gap. However, inflation volatility as a function of the extent of output gap reaction is U-shaped. This means that reacting to the output gap on top of inflation will only lower inflation volatility up to a certain level, after which inflation volatility starts to increase again.

These different outcomes regarding inflation volatility can be tested in the laboratory. We design a learning-to-forecast experiment where the only difference between treatments consists in the monetary policy rule used by the central bank. In one treatment, the central bank only reacts to inflation, while in the other it also reacts to the output gap. Our experimental results support the claim that inflation volatility can be lowered when the central bank also reacts to the output gap, in line with the predictions of the behavioral model.[1]

Our results from the behavioral model and the experimental data have clear policy implications for central banks whose sole aim is to achieve price stability, such as the European Central Bank (many other central banks, including those of New Zealand, Canada, England, and Sweden, have a hierarchical mandate with price stability as the primary objective for monetary policy). Even if these banks ultimately only care about price stability, this goal is better achieved if they also react to changes in the output gap. This is important and at odds with standard macroeconomic thinking built upon full rationality.

This paper is organized as follows. In Section 2 we describe how we model the economy and the formation of expectations. We also show the main differences between the rational and behavioral versions. In Section 3 we first describe the experimental design and the procedures. Then we show the experimental results. Section 4 concludes.

2 Theory

In this section, we first describe the underlying macroeconomic model. Then we introduce the behavioral model of expectation formation. After that, we compare the outcomes of both models and describe the economic intuition behind these outcomes.

[1] This research predominantly builds on two streams of literature. Firstly, it builds on the literature on experimental macroeconomics and learning-to-forecast experiments, e.g., Marimon and Sunder (1993), Kelley and Friedman (2002), Lei and Noussair (2002), Arifovic and Sargent (2003), Adam (2007), Heemeijer et al. (2009), Bao et al. (2012), Kryvtsov and Petersen (2013), Cornand and M'Baye (2016), Pfajfar and Zakelj (2014), Assenza et al. (2014b); see Duffy (2012), Assenza et al. (2014a), and Cornand and Heinemann (2014) for surveys (also note that learning-to-forecast experiments are closely related to beauty constest games, e.g., Nagel, 1995, Duffy and Nagel, 1997). Secondly, it builds on the literature on behavioral macroeconomics and learning in macroeconomics, e.g., Marcet and Nicolini (2003), Orphanides and Williams (2006), Branch and McGough (2009, 2010), Woodford (2010), De Grauwe (2011, 2012a,b), De Grauwe and Kaltwasser (2012), Anufriev et al. (2013), Kurz et al. (2013), Benhabib et al. (2014); see Evans and Honkapohja (2001) and Woodford (2013) for overviews.

2.1 Macroeconomic Model

The economic model we use can be described by the following aggregate New Keynesian equations:

$$
\begin{aligned}
y_t &= \bar{y}^e_{t+1} - \varphi(i_t - \bar{\pi}^e_{t+1}) + g_t && (1)\\
\pi_t &= \lambda y_t + \rho \bar{\pi}^e_{t+1} + u_t && (2)\\
i_t &= \text{Max}\{\bar{\pi} + \phi_\pi(\pi_t - \bar{\pi}) + \phi_y(y_t - \bar{y}),\, 0\}\,, && (3)
\end{aligned}
$$

where y_t and $\bar{y}^e_{t+1}$ are the actual and average expected output gap, i_t is the nominal interest rate, π_t and $\bar{\pi}^e_{t+1}$ are the actual and average expected inflation rates, g_t and u_t are exogenous disturbances and φ, λ, ρ, ϕ_π and ϕ_y are positive parameters. Equation (1) is the dynamic IS equation in which the output gap y_t depends on the average expected future output gap $\bar{y}^e_{t+1}$ and on the real interest rate $i_t - \bar{\pi}^e_{t+1}$. Equation (2) is the New Keynesian Phillips curve according to which the inflation rate depends on the output gap and on average expected future inflation. Equation (3) is the monetary policy rule implemented by the central bank describing how it reacts to deviations from the inflation target $\bar{\pi}$ and to deviations from the corresponding equilibrium level of the output gap $\bar{y} \equiv (1-\rho)\bar{\pi}/\lambda$. The coefficients ϕ_π and ϕ_y in this Taylor Rule measure how much the central bank adjusts the nominal interest rate i_t in response to deviations of the inflation rate from its target and of the output gap from its equilibrium level. As usual, the interest rate rule is subject to the zero lower bound, i.e. $i_t \geq 0$. When the zero lower bound is not binding, model (1)–(3) can be rewritten in matrix form as

$$
\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \Omega \begin{bmatrix} \varphi\bar{\pi}(\phi_\pi - 1) + \varphi\phi_y\bar{y} \\ \lambda\varphi\bar{\pi}(\phi_\pi - 1) + \lambda\varphi\phi_y\bar{y} \end{bmatrix} + \Omega \begin{bmatrix} 1 & \varphi(1-\phi_\pi\rho) \\ \lambda & \lambda\varphi + \rho + \rho\varphi\phi_y \end{bmatrix} \begin{bmatrix} \bar{y}^e_{t+1} \\ \bar{\pi}^e_{t+1} \end{bmatrix} + \Omega \begin{bmatrix} 1 & -\varphi\phi_\pi \\ \lambda & 1+\varphi\phi_y \end{bmatrix} \begin{bmatrix} g_t \\ u_t \end{bmatrix}, \quad (4)
$$

where $\Omega \equiv 1/(1 + \lambda\varphi\phi_\pi + \varphi\phi_y)$.

The economic model described by the aggregate equations (1)–(3), or equivalently by (4), is fully microfounded both under rational expectations (e.g., Woodford, 2003; Galí, 2008) and under behavioral expectations. We spell out the microfoundations for our behavioral model of expectation formation in Appendix A (these microfoundations are based on Kurz et al., 2013; for microfounded models under behavioral expectations see also Branch and McGough, 2009, and Massaro, 2013).

We also remark that, although Equations (1) and (2) are typically derived by log-linearizing around a steady state with a zero inflation rate, this does not mean that one can only consider policy rules with a zero inflation target. In fact, as argued in Woodford (2003), Equations (1) and (2) are valid approximations for the dynamics of inflation and output gap as long as the target inflation $\bar{\pi}$ in the policy rule (3) is not too large (see Appendix A for a further discussion).[2] In the remainder we will only make use of the aggregate equations presented here.

2.2 A Behavioral Model of Expectation Formation

Models with rational expectations are based on the assumption that agents have perfect information and a full understanding of the true model underlying the economy. There is, however, a large body of empirical literature documenting departures from this assumption and showing that agents use heuristics to make forecasts of future (macroeconomic) variables. This behavior is not necessarily a consequence of agents' irrationality; it can also be a "rational" response of agents who face cognitive limitations and have an imperfect understanding of the true model underlying the economy (e.g. Gigerenzer and Todd, 1999; Gigerenzer and Selten, 2002). Next, we introduce a behavioral model of expectation formation for such an environment.

Let $\mathcal{H}$ denote a set of H different heuristics used by agents to make forecasts of variable x. A generic forecasting heuristic $h \in \mathcal{H}$ based on available information at time t can be described as

$$x^e_{h,t+1} = f_h(x_{t-1}, x_{t-2} \ldots ; x^e_{h,t}, x^e_{h,t-1} \ldots). \tag{5}$$

In this paper x is either inflation π or the output gap y. Although agents can use simple rules to predict future inflation and output gap, we impose a certain discipline in the selection of such rules in order to avoid completely irrational behavior. Specifically, we introduce a selection mechanism that disciplines the choice of heuristics by agents according to a fitness criterion. This allows agents to learn from past mistakes and to choose heuristics that have performed well in the (recent) past. U_h denotes the fitness measure of a certain forecasting strategy h defined by

$$U_{h,t-1} = F(x^e_{h,t-1} - x_{t-1}) + \eta U_{h,t-2} , \tag{6}$$

where F is a generic function of the forecast error of heuristic h, and $0 \leq \eta \leq 1$ is a memory parameter measuring the relative weight agents give to past errors of heuristic h. Performance is completely determined by the most recent forecasting error if $\eta = 0$, while performance depends on all past prediction errors with exponentially declining

[2]Alternatively, one could assume a staggered price setting mechanism as in Yun (1996) where prices that are not reconsidered in any given period are automatically increased at the target rate (see e.g. García-Schmidt and Woodford, 2015 for a recent application) and log-linearize directly around the target steady state.

weights if $0 < \eta < 1$ or with equal weights if $\eta = 1$. If all agents simultaneously update the forecasting rule they use, the fraction of agents choosing rule h in each period t can be described by

$$n_{h,t} = \frac{\exp\left(\beta U_{h,t-1}\right)}{\sum_{h=1}^{H} \exp\left(\beta U_{h,t-1}\right)} . \tag{7}$$

The multinomial logit expression described in Equation (7) can be derived directly from a random utility model (see Manski and McFadden, 1981, and Brock and Hommes, 1997). The parameter $\beta \geq 0$, referred to as "intensity of choice", reflects the sensitivity of agents to selecting the optimal prediction strategy according to the fitness measure U_h.[3] If $\beta = 0$, $n_{h,t}$ is constant for all h, meaning that agents do not exhibit any willingness to learn from past performance; if $\beta = \infty$, all agents adopt the best performing heuristic with probability one. The reinforcement learning model in Equation (7) is extended in Hommes et al. (2005a) and Diks and van der Weide (2005) to include asynchronous updating in order to allow for the possibility that not all agents update their rule in every period (consistent with empirical evidence; see Hommes et al., 2005b, and Anufriev and Hommes, 2012). This yields a generalized version of Equation (7) described by

$$n_{h,t} = \delta n_{h,t-1} + (1-\delta)\frac{\exp\left(\beta U_{h,t-1}\right)}{\sum_{h=1}^{H} \exp\left(\beta U_{h,t-1}\right)} . \tag{8}$$

The parameter $0 \leq \delta \leq 1$ introduces persistence in the adoption of forecasting strategies and can be interpreted as the average fraction of individuals who, in each period, stick to their previous strategy.

In order to use this behavioral model for policy analyses or predictions, specific assumptions have to be made about the nature of agents' forecasting heuristics (in general, the set $\mathcal{H}$ may contain an arbitrary number of forecasting rules). We restrict our attention to a set of four heuristics described in Table 1.

The choice of this specific set of heuristics is motivated on empirical grounds. These heuristics were obtained and estimated as descriptions of typical individual forecasting behavior observed in Hommes et al. (2005b), Hommes et al. (2008), and Assenza et al. (2014b) building upon a rich literature on expectation formation (see Hommes, 2011, for a recent survey). Based upon the calibration in these papers, we use the parameters $\beta = 0.4$, $\delta = 0.9$, and $\eta = 0.7$.[4]

[3] Equation (7) can also be derived from an optimization problem under rational inattention (see Matějka and McKay, 2015). In this context, the parameter β is inversely related to the "shadow cost of information".

[4] Furthermore, we use the forecast error function $F(x_h^e - x) = 100 - 100/(1 + |x_h^e - x|)$, which is the function used to incentivize subjects in the experiment described in Section 3 (this incentive structure

Table 1: Set of heuristics

ADA	adaptive rule	$x^e_{1,t+1} = 0.65x_{t-1} + 0.35x^e_{1,t}$
WTR	weak trend-following rule	$x^e_{2,t+1} = x_{t-1} + 0.4(x_{t-1} - x_{t-2})$
STR	strong trend-following rule	$x^e_{3,t+1} = x_{t-1} + 1.3(x_{t-1} - x_{t-2})$
LAA	anchoring and adjustment rule	$x^e_{4,t+1} = 0.5(x^{av}_{t-1} + x_{t-1}) + (x_{t-1} - x_{t-2})$

Notes: x^{av}_{t-1} denotes the average of all observations up to time $t-1$.

2.3 Monetary Policy and Economic Behavior

2.3.1 Existence and Non-Existence of Trade-Offs

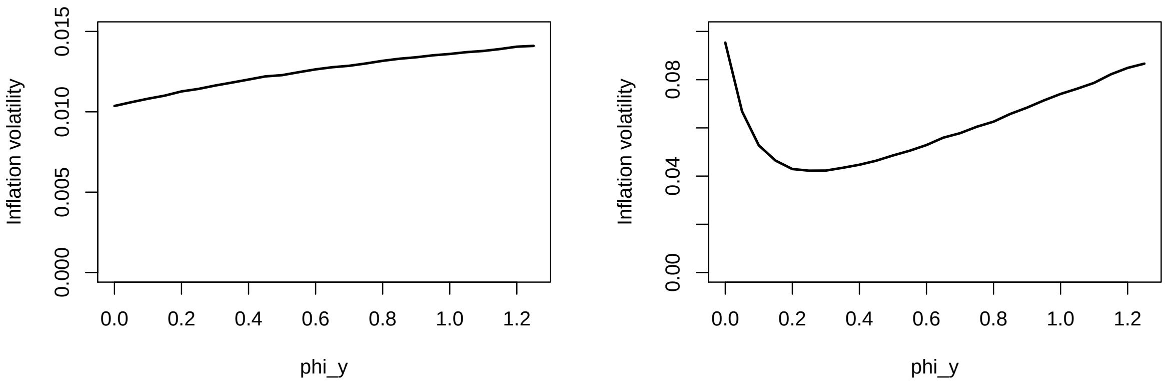
A result derived from Model (4) under rational expectations is that a policy trade-off is observed between the volatility of the output gap and the volatility of inflation. A decline in output gap volatility resulting from a more active output stabilization policy comes at the price of an increase in inflation volatility (it is reasonable to focus on volatility as for the rational and the behavioral models alike inflation and output gap are on average at their target and steady state level for reasonable values of ϕ_π and ϕ_y). This policy trade-off is described in Figure 1a, where we show the effect of ϕ_y (with which the central bank reacts to deviations of the output gap from its steady state level) on inflation volatility. Higher output stabilization, i.e. an increase in the reaction coefficient ϕ_y, comes at the price of higher inflation volatility. The immediate policy implication for a central bank whose main objective is price stability is that it is optimal to set $\phi_y = 0$, i.e. not to react to output gap fluctuations at all (cf. Galí, 2008, and Woodford, 2003).

For the simulations of this graph, the parameter ϕ_π is equal to 1.5 (different values lead to similar results, see Appendix B) and the structural parameters in Equations (1)-(3) are as estimated in Clarida et al. (2000).[5] The inflation target used for the simulations is $\bar{\pi} = 3.5$ (this is the same target that will be used in the experiment, a rationale for this value can be found in Section 3.2; the simulations yield similar results for different values of $\bar{\pi}$). This inflation target leads to a steady state level of the output gap of $\bar{y} = 0.1166667$. Inflation volatility is measured by $v(\pi) = \frac{1}{T}\sum_{t=2}^{T}(\pi_t - \pi_{t-1})^2$, with T denoting

is also used in Adam, 2007, Pfajfar and Zakelj, 2014, and Assenza et al., 2014b, among others). The simulation results in Section 2.3 are qualitatively robust to alternative specifications of the fitness metric, such as using a quadratic function.

[5] Thus, $\rho = 0.99$, $\lambda = 0.3$, and $\varphi = 1$ (for quarterly data). The shocks g_t and u_t are independent and normally distributed with standard deviation 0.1. The number of simulations for each value of ϕ_y is 10000.

the total number of periods. This measure of volatility has some properties that make it preferable to other measures of price instability (the measurement of volatility is discussed in Section 2.3.2 and in Appendix C; using alternative measures yields similar results).



(a) Rational model

(b) Behavioral model

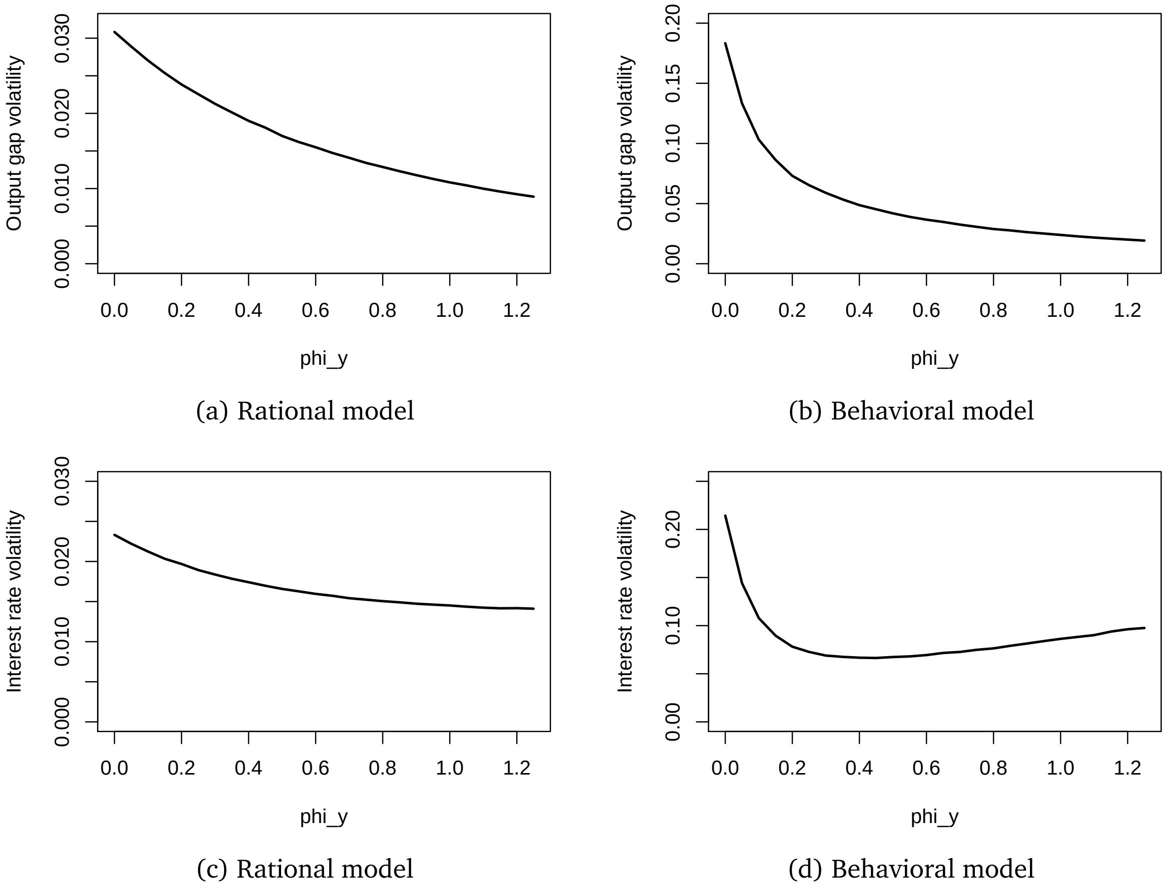
Figure 1: Inflation volatility as a function of ϕ_y for the rational and for the behavioral model

Notes: This figure shows the effect of parameter ϕ_y on inflation volatility. $\phi_\pi = 1.5$ for both sub-figures.

In Figure 1b, we show the effect of the parameter ϕ_y on inflation volatility when expectations are formed according to the behavioral model described in Section 2.2 (note that the scales in Figures 1a and 1b are different; the overall level of inflation volatility is higher under behavioral expectations than under rational expectations). In contrast to the simulation results under rational expectations, the graph of inflation volatility as a function of ϕ_y has a U-shape.[6] Thus, starting from $\phi_y = 0$, the central bank can simultaneously decrease inflation and output gap volatility by also reacting with its monetary policy to deviations of the output gap from its steady state level (in addition, reacting to the output gap would also lead to less volatile interest rates). Figure 2 depicts output gap volatility and interest rate volatility as functions of ϕ_y (as ϕ_y increases output gap volatility decreases under both rational and behavioral expectations; the interest rate decreases continuously in ϕ_y under rational expectations, while it first decreases strongly under behavioral expectations and then slowly increases again). Hence, under behavioral expectations, there is a broader scope for output stabilization.

Now we turn to the intuition of these results. Considering the outcome simulated with rational expectations (Figure 1a), one may be tempted to believe that the following sim-

[6]The starting values used for the simulations of the behavioral model are $\pi_{start} = 3.0$ and $y_{start} = 0.5$, Appendix B provides graphs for different starting values, which are also U-shaped. The initial fraction of agents using any of the four heuristics is 0.25.



(a) Rational model

(b) Behavioral model

(c) Rational model

(d) Behavioral model

Figure 2: Output gap and interest rate volatility as functions of ϕ_y for the rational and for the behavioral model

Notes: This figure shows the effect of parameter ϕ_y on output gap and interest rate volatility. $\phi_\pi = 1.5$ for all sub-figures.

ple rule is correct: “If there are two variables, targeting one variable will always come at the expense of the other variable”. In general, this is not the case, however. The intuition is slightly more complex. Homogeneous rational expectations are strictly forward looking and in this model always equal to the inflation target and the corresponding steady state level of the output gap, respectively (assuming that $\phi_\pi + \phi_y(1-\rho)/\lambda > 1$, which ensures a determinate model solution, see e.g. Woodford, 2003). These expectations do not depend in any way on the current level of inflation and output gap or on any past behavior. It is precisely via the dependence of expectations on (past) actual variables that reacting to the output gap can also pay off in terms of inflation volatility. To illustrate this, imagine that inflation and output gap are constant at $\bar{\pi}$ and $\bar{y}$, respectively, and that a combination of shocks arrive in one period that would lead (without any reaction by the central bank) to inflation staying constant and the output gap rising above the steady state level. Should the central bank react to this shock if it only cares about inflation? The rational expectations answer would be "no"; inflation is at its target and in the next period one would (assuming no further shocks) again be at the inflation target and the steady state level of the output gap, because expectations do not react to the past. However, under behavioral expectations, what happens now matters for the future. If there is some adaptive or trend-following behavior, a higher output gap now will lead agents to revise their expectations of the future output gap upwards, leading to a higher realized output gap in the future, which will in turn lead to upward pressure on inflation. Therefore, it can be beneficial for the central bank to curb the increase of the output gap now (at the expense of slightly lower inflation now) in order to reduce the upward pressure on inflation in the future. However, if the monetary authority puts too much weight on output gap stabilization, the ensuing fluctuations in inflation dominate the stabilization bonus provided by less volatile output, leading to higher inflation volatility.

2.3.2 Robustness and Measurement of Inflation Volatility

The simulation results are qualitatively robust to a wide variety of changes. This includes changes in all parameters of the macroeconomic model. It also includes changes in the parameters of the behavioral model of expectation formation. More interestingly, the results are also robust to other models of behavioral expectation formation, such as a heuristic switching model with fewer and simpler heuristics or adaptive expectations without any switching involved; the results are even robust to using the behavioral switching model as we describe it with an additional heuristic of forecasting the central bank's inflation target). Such variations are shown in Appendix B. While the results are qualitatively robust to these changes within this macroeconomic framework (which is the most standard framework for macroeconomic policy analysis), it is possible in other macroeconomic frameworks to reverse the results obtained by rational expectations. That is, it is possible to obtain a reduction of inflation volatility by increasing ϕ_y; an example are models that only include shocks to the aggregate demand equation (1) such as technology shocks, preference shocks or variations in government purchases, but do not include shocks to the short-run aggregate supply relationship (2) (see Woodford, 2003). In such frameworks, behavioral expectations are an additional reason why inflation volatility decreases when also targeting the output gap.[7]

[7]Results similar to ours are obtained, for example, in a different macroeconomic model when employing simplistic behavioral rules of expectation formation (De Grauwe, 2011, 2012a). A non-monotonic relationship between inflation and output gap volatility can also arise in sticky information economies in which the degree of attentiveness or the rate at which agents update their information is endogenized (Branch et al., 2009).

We focus on inflation volatility as measured by $v(\pi) = \frac{1}{T}\sum_{t=2}^{T}(\pi_t - \pi_{t-1})^2$ for the simulations of the theoretical model (and for the predictions for the experiment). This measure has advantages over alternative measures of price instability. For some economists, the mean squared deviation from the target springs to mind as a measure. However, the measure we use has a few intuitive advantages over the mean squared deviation. For example, the mean squared deviation does not distinguish between erratic behavior around the target with decreasing distance from the target and slow convergence if the absolute distance to the target is always equal. The differences between these two measures and other simple measures are discussed in more detail in Appendix C. Note, however, that we obtain similar results when using different measures. When using, for example, the mean squared deviation from the target, while the shapes in the simulations still persist, differences become smaller, i.e. the curve becomes flatter. The same holds for our experimental results, which are described in the next section: the results go in the same direction but are not quite as strong (though the mean squared deviation from the target in the "inflation targeting only" treatment is still more than 20% above that in the "inflation and output gap targeting" treatment).

One of the reasons why the mean squared deviation from the target may be popular among economists is that it constitutes a welfare criterion under homogeneous expectations. However, as shown in Di Bartolomeo et al. (2016), it is not an appropriate welfare criterion when agents have heterogeneous expectations. In this case, price dispersion arises not only because of the staggered price setting mechanism but also because of the heterogeneity of prices set by reoptimizing firms in the Calvo lottery, which depends on the heterogeneity of firms' expectations of future inflation. In our behavioral model, this heterogeneity increases in the relative changes in inflation. We refrain from using the precise welfare criterion as it depends on more than inflation alone. While welfare criteria derived in particular models may have influenced the fact that price stability is now the sole aim of many central banks, these central banks now have the mandate to achieve price stability and not the aim of maximizing a model-dependent welfare criterion. In addition, using a composite welfare measure would reduce the clarity and readability of the paper. Note that both simulation results and experimental results are similar when considering the precise welfare criterion.

3 Experiment

The only task for subjects in the experiment is to forecast inflation and output gap. These forecasts are then used to calculate subsequent realizations. The model underlying the experimental economy is the macroeconomic model described in Section 2.1 (with the same calibration of macroeconomic parameters as before). Before we describe the experiment in more detail, we now explain the treatments and hypotheses. The design of the experiment and the hypotheses can be motivated with the theory described in Section 2.

3.1 Treatments and Hypotheses

There are two treatments, $T1$ ("inflation targeting only") and $T2$ ("inflation and output gap targeting"). The only difference between the treatments lies in the Taylor rule describing monetary policy. In $T1$, the parameters of the Taylor rule are $\phi_\pi = 1.5$ and $\phi_y = 0$, whereas they are $\phi_\pi = 1.5$ and $\phi_y = 0.5$ in $T2$. That is, the only difference between the treatments is that in $T1$ the central bank only targets inflation, whereas it targets the output gap in addition to inflation in $T2$.

We are interested in testing the null-hypothesis (which can be derived from the rational expectations model in Section 2) that inflation volatility in $T1$ is less or equal to inflation volatility in $T2$ against the alternative hypothesis (which can be derived from the behavioral model) that inflation volatility is greater in $T1$ than in $T2$. Figure 3 summarizes these hypotheses.[8]

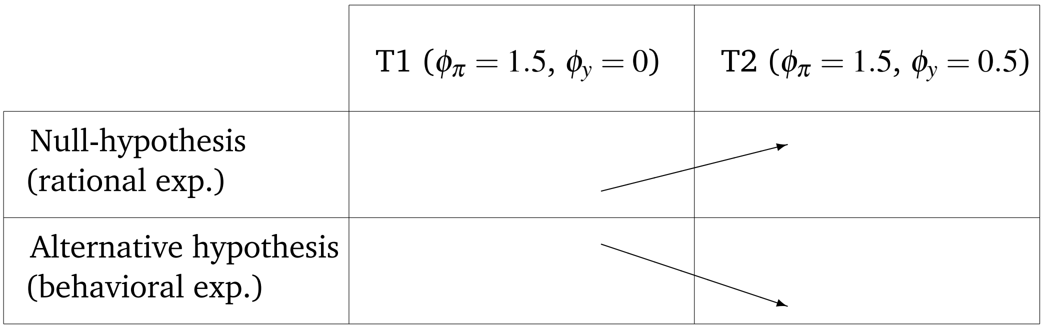


Figure 3: Hypotheses about inflation volatility

In the experiment, the number of subjects per experimental economy is six. Evidence from other experiments indicates that four to six subjects are enough to justify the use of the competitive equilibrium as equilibrium concept (see, e.g., Huck et al., 2004).

[8]The experiment can be seen as a controlled investigation of the outcomes of different monetary policies but also as a test between the rational and the behavioral models. While some people may argue that the best test of the models is to compare subjects' forecasts to the model predictions (in which the behavioral model does much better), others might question such a comparison on the ground that it is a within-treatment comparison; the directionally different hypotheses in our experiment make it a cleaner test (in laboratory experiments, the comparative statics of treatment comparisons are generally considered to be most robust and relevant; see Schram, 2005, or Falk and Heckman, 2009).

Note, however, that also in a game theoretic analysis the unique Nash equilibrium is forecasting $\bar{\pi}$ and $\bar{y}$.

3.2 Course of Events and Implementation

The design is a between-subjects design with within session randomization. In the beginning, all participants are divided into groups (experimental economies) of six. Subjects only interact with other subjects in their group, without knowing who they are. Subjects are asked to make forecasts of inflation and output gap. The average forecasts of all subjects in one group are then used to calculate the realizations of inflation and output gap according to model equations (1)–(3) (only the average forecasts $\bar{\pi}^e_{t+1}$ and $\bar{y}^e_{t+1}$ are needed to calculate the realizations π_t and y_t). When making their forecasts for period $t+1$, the information subjects can see on their screen (as numbers and partly also in graphs) is the following: all realizations of inflation, output gap, and interest rate up to period $t-1$, their own forecasts of inflation and output gap up to period t and their scores stating how close their past forecasts were to realized values up to period $t-1$ (these scores determine the payments). As subjects are only informed about realizations up to period $t-1$, their forecasts for period $t+1$ are effectively two-period-ahead forecasts. Figure 4 shows a screenshot of the experiment (a larger version of the same screenshot can be found in Appendix E).

The inflation target of the central bank in the experiment is $\bar{\pi} = 3.5$. This target is chosen for two reasons. First, it is distant from the zero lower bound, which is desirable as we do not wish to investigate behavior in a liquidity trap. Second, it is different from focal points such as 2% or 2.5%, which are standard inflation targets in the real world. We avoid these focal points so that learning can be observed in the experiment. Our theory and experiment concern feedback from the monetary policy rule to deviations of inflation and output gap from their target and steady state levels. Laboratory subjects are very heterogeneous, and if most of them start out with their forecasts extremely close to the target already, the feedback plays a smaller role in comparison to subjects' heterogeneity and mistakes.[9]

Subjects' payments depend on their forecasting performance. Whether a participant is paid for inflation forecasting or output gap forecasting is determined randomly at the end of the experiment. The total scores for inflation and output gap forecasting are the sums of the respective forecasting scores over all periods. This score is for subject i's

[9] While we are convinced that the experimental results would go in the same direction with an inflation target of 2%, we expect that one would need many more subjects to detect these results.

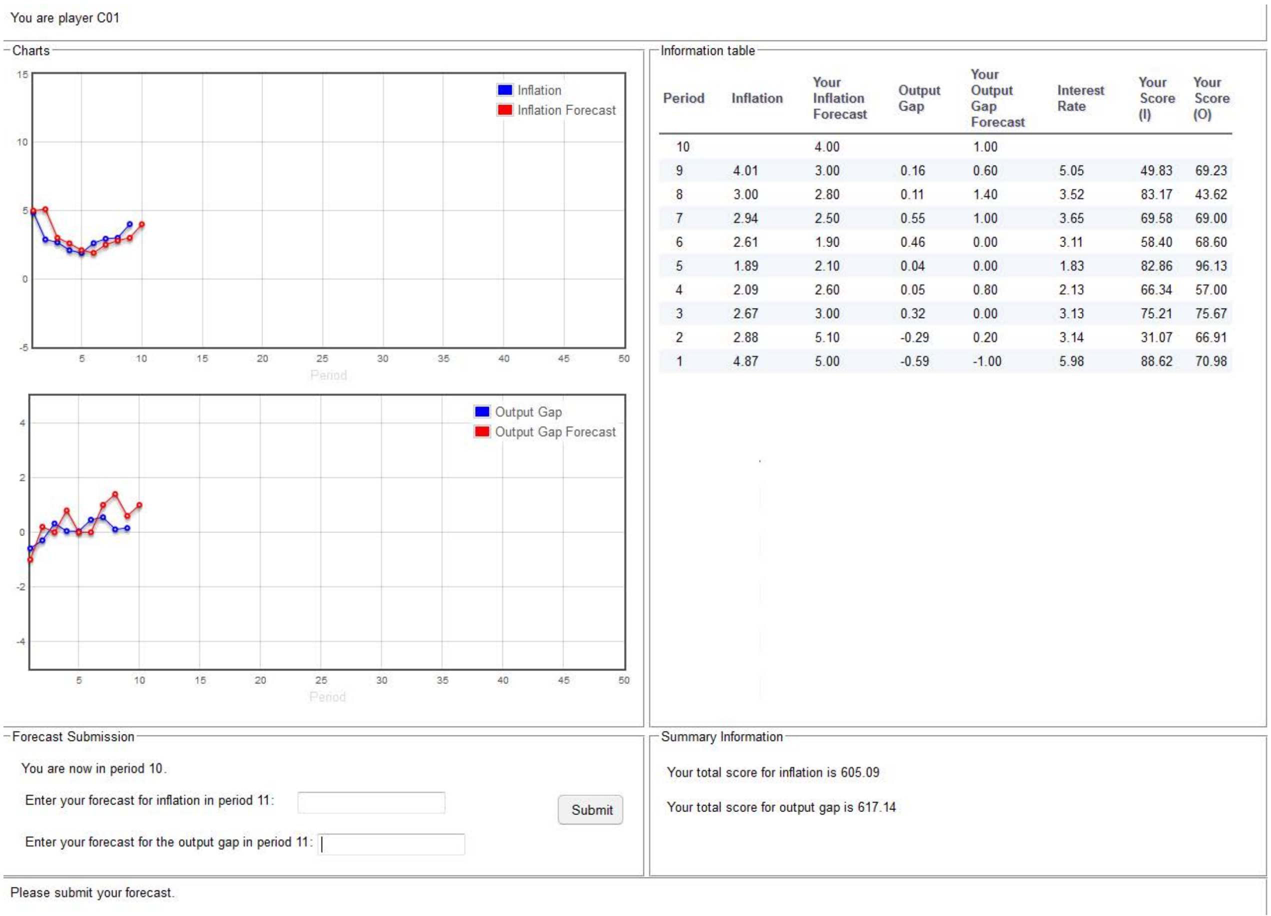


Figure 4: Screenshot

inflation forecast in period t equal to $100/(1+|\pi^e_{t,i}-\pi_t|)$, where $\pi^e_{t,i}$ denotes subject i's forecast for period t and π_t the realized value of this period. The score for output gap forecasting is calculated analogously. This means that subjects' payments decrease with the distance of the realizations from their forecasts.

In the instructions, subjects receive a qualitative description of the economy that includes an explanation of the mechanisms that govern the model equations. Concerning monetary policy, subjects in both treatments are only told that the central bank decreases the interest rate if it wants to increase inflation or output gap, and that it increases the interest rate if it wants to decrease inflation or output gap.[10] Except for the precise formulation of the equations of the macroeconomic model, the instructions contain full information about the experiment (i.e. on the number of subjects per group, payments, etc.). The complete instructions can be found in Appendix D.

[10]As the experiment uses two-period-ahead forecasts, after reading the instructions subjects are asked to enter forecasts for periods 1 and 2 simultaneously. Subjects therefore receive some indication of reasonable values by being told in the instructions that in economies similar to the one at hand inflation has historically been between −5% and 10% and the output gap between −5% and 5%.

The experiment was programmed in Java and conducted at the CREED laboratory at the University of Amsterdam. The experiment was conducted with 258 subjects recruited from the CREED subject pool (43 groups of six subjects each, distributed over thirteen sessions). After each session, participants filled out a short questionnaire. Participants were primarily undergraduate students; the average age was slightly above 22 years. About half of the participants were female, about two-thirds were majoring in economics or business, and about half were Dutch. During the experiment, 'points' were used as currency. These points were exchanged for euros at the end of each session at an exchange rate of 0.75 euros per 100 points. The experiment lasted around two hours, and participants earned on average about 30 euros. The series of error terms used in the model equations (g_t and u_t in equations 1 and 2) differed across groups within each treatment, but the sets of noise series used in the two treatments were the same.[11]

3.3 Results

There are data of 43 different groups, 21 in $T1$ and 22 in $T2$. The groups' actions do not influence one another in any way; thus the observations at the group level are statistically independent. The data for all groups separately including all individual forecasts can be found in Appendix E.

3.3.1 Inflation

Figure 5 gives an overview of inflation in all experimental economies, separately for $T1$ and $T2$. Each line corresponds to the inflation in one experimental economy, tracked over all 50 periods of the experiment. Almost all economies are close to the inflation target after 50 periods, and for the economies with inflation still oscillating around the target the amplitude of these oscillations is decreasing. That many economies are

[11] Before conducting the experiment, two pilot sessions were conducted (with a total of six groups). The pilot sessions differ from the actual experiment as follows: the error terms added to the model equations had a larger standard deviation, a different inflation target was used, and subjects in the pilot did not receive any information on the number of participants in each group. For two of the groups, a different combination of parameters for the Taylor rule was used.

We excluded two of the groups from the analysis (including these two groups, the experiment was conducted with 270 subjects). One of the groups was excluded because of a very large typo (a forecast of 30 instead of 3.0; the corresponding participant notified us about this typo in the post-experiment questionnaire). The other group was excluded due to a severe misunderstanding on the part of one subject, who systematically stayed very far from the actual realizations (thereby also losing a lot of money). Our conclusions do not change if we include these groups in our analysis. The realizations and forecasts of inflation and output gap for these two groups are shown in Figure 30, Appendix E.

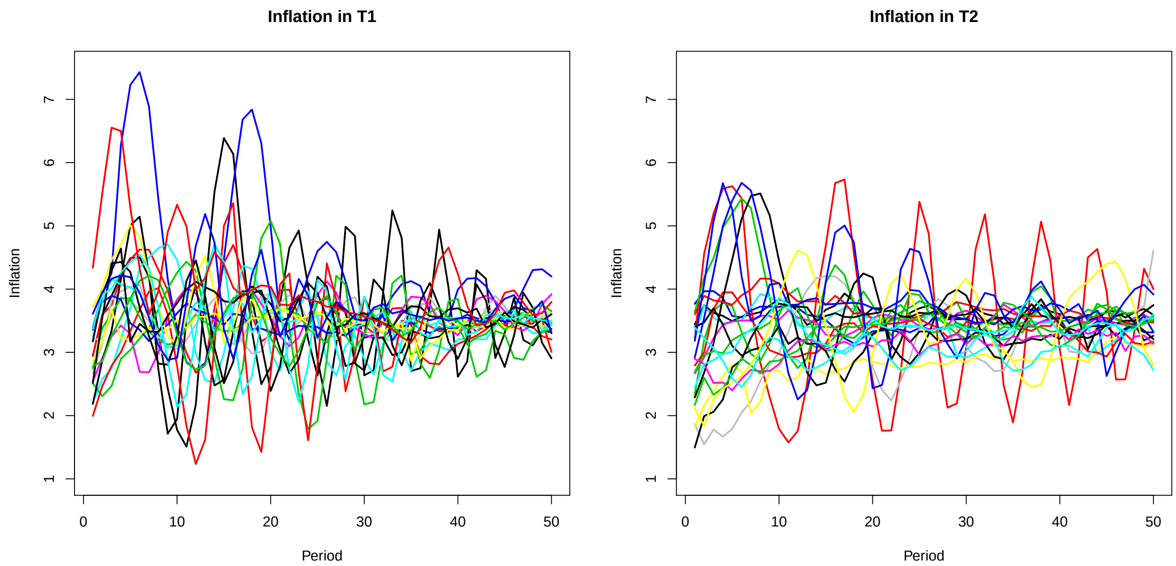


Figure 5: Realized inflation for all groups in both treatments

Notes: Each line represents realized inflation in one economy. On the horizontal axis is the number of periods (1 to 50), on the vertical axis inflation in percent (from 1 to 7.5).

converging to the steady state over the course of the experiment is not necessarily surprising, as there are 50 periods without any changes to the underlying model (cf. Pfajfar and Zakelj, 2014, and Assenza et al., 2014b).

It is easy to see from this figure that groups are very heterogeneous. A few groups exhibit much larger volatility than the other groups. Inflation in many groups in both treatments is within one percentage point from the inflation target in most or all periods. Nevertheless, one can see that on average inflation fluctuates a bit less in $T2$ than in $T1$. In particular, when looking at the many groups that stay within roughly one percentage point from the target, one can see that there is more up-and-down movement of the lines in $T1$ than in $T2$ (although there is one more observation in $T2$). This is also what one can see when one follows single lines from period 1 to 50; the lines of most groups in $T2$ are flatter than the lines of most groups in $T1$.

A first look at these data thus suggests that inflation is indeed less volatile in $T2$ when the output gap is also targeted than in $T1$, as predicted by the behavioral model. While it may be difficult to distinguish between the lines in such a densely populated graph, the following data analysis does not rely on good eyes. Note that while inflation volatility is different between the treatments, inflation generally fluctuates around its target: the mean of inflation over all 50 periods is between 3.13 and 4.33 in $T1$ and between 2.79 and 3.82 in $T2$ for all groups.

We now turn to more detail about inflation volatility in the experiment. As in Section 2.3, we use $v(\pi) = \frac{1}{T}\sum_{t=2}^{T}(\pi_t - \pi_{t-1})^2$ as measure of inflation volatility (see Section 2.3.2 and Appendix C for a discussion). The volatility in all experimental economies can be seen in Figure 6 where the empirical cumulative distribution functions (ECDFs) are drawn, for groups in both treatments (for each value on the horizontal axis, the ECDF shows on the vertical axis the fraction of groups in each treatment with inflation volatility less or equal to this value; the colored dots represent the observations). It can easily be seen that inflation volatility is lower in $T2$ than in $T1$. In fact, the whole ECDF of observations in $T2$ lies to the left of the ECDF of observations in $T1$ (the single one high value in $T2$, i.e. the rightmost blue dot, corresponds to the oscillating red line in the right graph of Figure 5).[12]

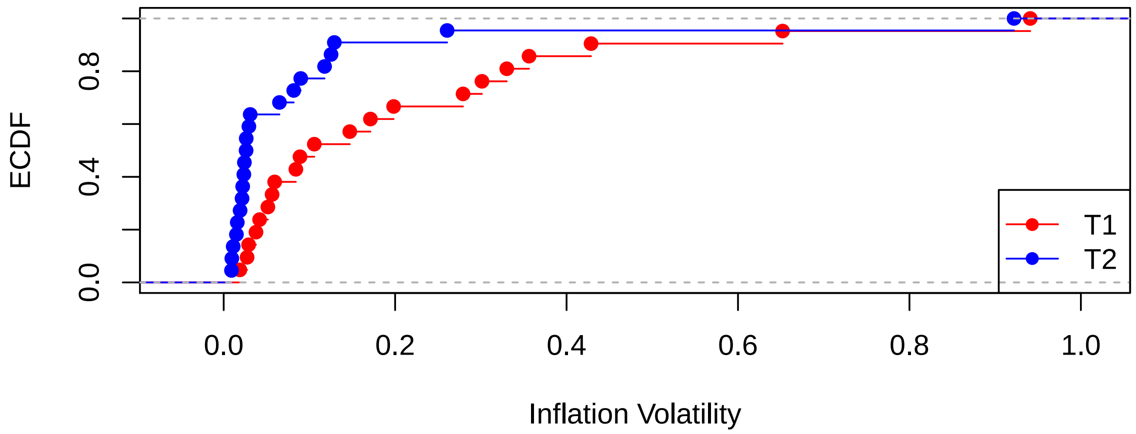


Figure 6: Empirical distribution functions of inflation volatility

Notes: For each value on the horizontal axis, the fraction of observations with inflation volatility less or equal to this value (i.e. the ECDF) is shown on the vertical axis, separately for $T1$ and $T2$.

In order to test the statistical significance of this finding, we use a Wilcoxon rank-sum test. We test the null-hypothesis that inflation volatility is less or equal in $T1$ than in $T2$ against the alternative hypothesis that inflation volatility is lower in $T2$.[13] This test rejects the null-hypothesis ($p < 10^{-3}$). The advantage of the Wilcoxon rank-sum test is that it makes very unrestrictive assumptions on the underlying data. Note, however, that the results are robust to employing different tests.[14]

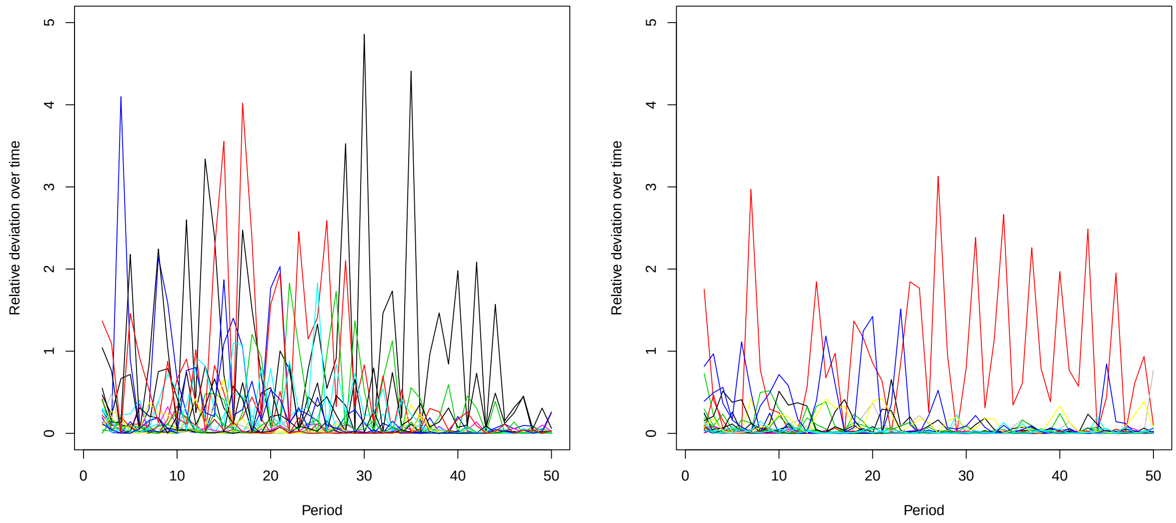
In addition to looking at the volatility, it is also possible to look at the squared dif-

[12]The ECDFs of other measures of price instability look similar to the one in Figure 6 and can be found in Appendix F (Figure 32).

[13]Strictly speaking, the Wilcoxon rank-sum test tests the null-hypothesis that the distribution shifts to the right (from $T1$ to $T2$) or that it does not change.

[14]The data are not normally distributed, but the logarithms of the data look rather close to a normal distribution (and are statistically not significantly different from it, according to a Kolmogorov-Smirnov test). A t-test on the logarithms of the data also rejects the null-hypothesis ($p = 0.002$).

ferences from period to period (without aggregating over all periods). Thus, for each group and each period t, one can compute $(\pi_t - \pi_{t-1})^2$. Figure 7 shows all of these differences per group, separately for $T1$ and $T2$. Figure 7 shows that there are multiple groups with relatively large and very large jumps in inflation in $T1$ while there are only two groups in $T2$ with jumps in inflation that can be considerate relatively large (shown with a red and blue line). Across the board, this graph shows less smooth movements of inflation in $T1$ than in $T2$.



(a) T1

(b) T2

Figure 7: Squared differences per group across all periods

Notes: This figure shows for each group the squared difference in inflation from period to period, i.e. $(\pi_t - \pi_{t-1})^2$. Each line has the same color as the line for the same group's inflation in Figure 5.

3.3.2 Output Gap and Interest Rate

Figure 8 shows the output gap in all experimental economies. Here, the differences are even larger; the output gap is much more volatile in $T1$ than in $T2$. This was to be expected, as both models predict that the output gap is more stable when it is also targeted by the central bank. The mean of the output gap is between -0.12 and 0.70 in $T1$ and between -0.03 and 0.66 in $T2$. A Wilcoxon rank-sum test rejects the null-hypothesis that output gap volatility is less or equal in $T1$ than in $T2$ ($p < 10^{-4}$).

Similarly, Figure 9 shows the interest rates in all groups. In addition, it shows a horizontal line at zero. As one can see in these graphs, the zero lower bound is never hit (it is almost hit in one group in $T1$, but the lowest interest rate in this group is still

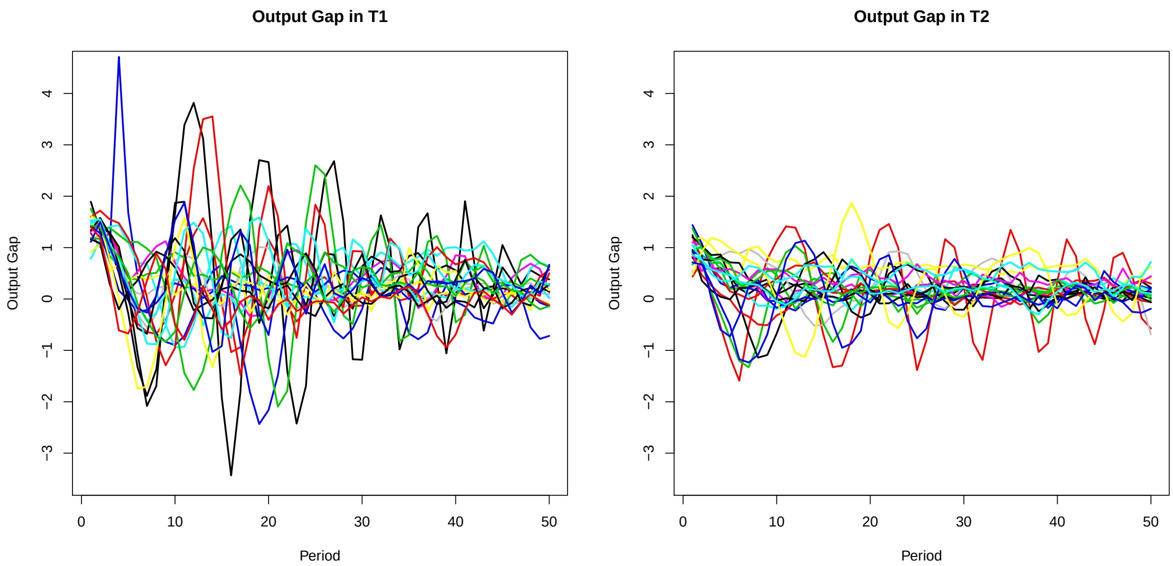


Figure 8: Realized output gap for all groups in both treatments

Notes: Each line represents the realized output gap in one economy. On the horizontal axis is the number of periods (1 to 50), on the vertical axis output gap in percent (from −3.5 to 4.5). Each line has the same color as the line for the same group's inflation in Figure 5.

slightly above zero). The mean of the interest rate is between 2.94 and 4.74 in $T1$ and between 2.70 and 3.92 in $T2$. This figure shows that the interest rate is much smoother in $T2$ than in $T1$. Note that this smoothness is achieved without interest rate smoothing in the Taylor Rule. Thus, reacting to changes in the output gap on top of inflation not only decreases inflation volatility and output gap volatility simultaneously but also leads to a less volatile interest rate. This can be seen as an additional reason for central banks to react to the output gap on top of inflation (a smooth interest rate may not be included in the mandate of a central bank, but in practice central bankers care about it; for a discussion see Srour, 2001). These differences are also statistically significant: a Wilcoxon rank-sum test rejects the null-hypothesis that interest rate volatility is less or equal in $T1$ than in $T2$ with a p-value of less than 10^{-4}.[15]

[15] In an experiment on the effects of communicating the inflation target, Cornand and M'Baye (2016) also have treatments with different Taylor rules. Between their two treatments most closely related to our work, not only the output gap reaction coefficient is changed but also the inflation reaction coefficient. Looking at their results from our viewpoint, they find no differences in inflation or interest rate variation between the treatments, while they have a marginally significant result that output gap variation is lower when the central bank also reacts to the output gap (while simultaneously reacting less to inflation). Low statistical power in their experiment with four observations per treatment could explain these differences (or, alternatively, that the inflation reaction coefficient is altered with the output gap coefficient).

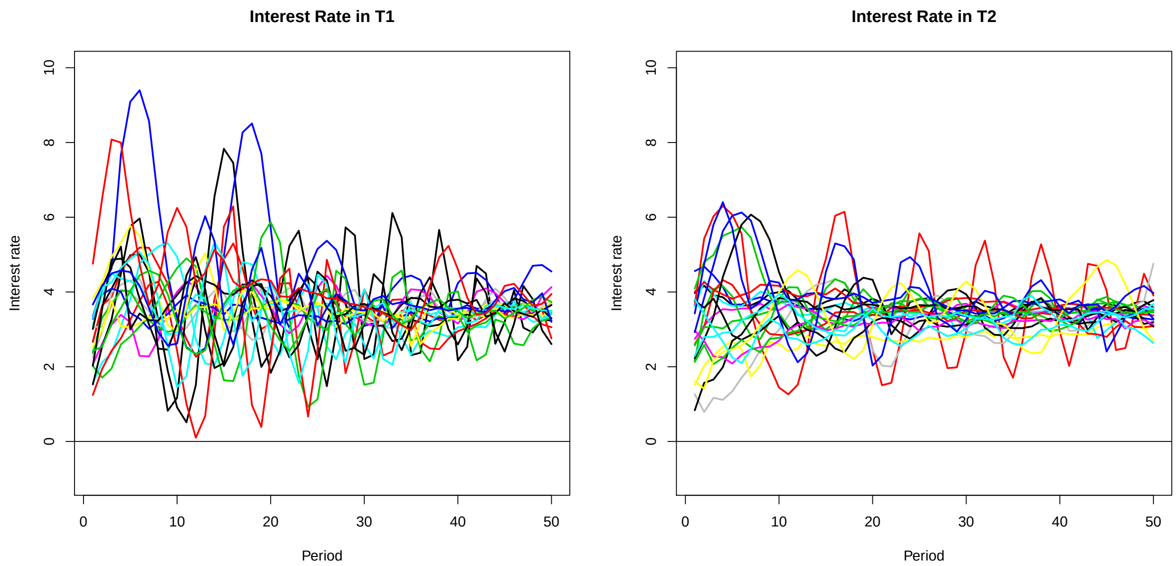


Figure 9: Interest rate for all groups in both treatments

Notes: Each line represents the interest rate in one economy. On the horizontal axis is the number of periods (1 to 50), on the vertical axis the interest rate in percent. Each line has the same color as the line for the same group's inflation in Figure 5. A thin horizontal line is added to indicate the zero lower bound.

3.3.3 Subjects' Forecasts and Models of Expectation Formation

After having analyzed the economic outcomes, we now examine the performance of the heuristic switching model used to derive the predictions in the experiment. Does this model accurately describe subjects' forecasts? Or does the rational expectation solution or one of the heuristics alone predict subjects' forecasts better than the switching model? Table 2 shows how well these models of expectation formation predict subjects' forecasts. We report the prediction performance of the heuristic switching model (HSM), the performance of the homogeneous rational agent solution (RE) and the performance of the four heuristics involved in the switching model without any switching: adaptive expectations (ADA), weak trend-following (WTR), strong trend-following (STR), and the learning, anchoring and adjustment rule (LAA). The calibration is the same as in Section 2. In each group, we derive two-period-ahead forecasts of the models and calculate squared prediction errors with respect to this group's average forecast.[16] The table shows the averages across all periods and all groups in a treatment.

[16]Computing the prediction error as the squared deviation of the prediction from the theoretical model with a continuum of agents from a group's average forecast is conservative in the sense that the prediction error is then in general greater than it would be if calculated at the individual level. Computing it for each subject separately would yield lower error terms. However, this would come at the expense of more degrees of freedom (because one would use the minimal distance from any of the simple heuristics). We prefer to put our model at a slight disadvantage over raising doubts about whether the comparison is justified.

Table 2: Mean squared prediction errors

	Inflation $T1$	Output gap $T1$	Inflation $T2$	Output gap $T2$
HSM	0.072	0.141	0.040	0.022
RE	0.541	0.753	0.422	0.222
ADA	0.254	0.399	0.168	0.095
WTR	0.106	0.193	0.063	0.037
STR	0.246	0.415	0.088	0.068
LAA	0.107	0.180	0.063	0.037

Notes: This table shows mean squared errors of two-period-ahead predictions from different models of expectation formation. The mean is taken over all periods and all groups, separately for $T1$ and $T2$.

The first thing one can see in the table is that, across the board, the switching model performs much better than rational expectations. Also evident in the table is that the rational expectation solution is a worse predictor in all cases than any of the four investigated heuristics standing alone. Furthermore, the switching model is a better predictor in all cases than any of the four heuristics alone. In general, the differences are considerable. The switching model does much better than most of the other models. There are two heuristics that do very well when employed alone: the weak trend-following rule and the anchoring and adjustment rule. Nevertheless, the switching model predicts all four forecasts better than these heuristics. The prediction errors of these two best-performing heuristics when employed alone are always at least 25% greater than the prediction errors of the switching model. Most of the differences discussed above are statistically highly significant.[17] When considering a different error measure which puts less weight on the (potentially few) largest deviations and more weight on the many small deviations, the results are similar (see Appendix F.2).

Moreover, it is noticeable when looking at Table 2 that prediction errors of all models are smaller in $T2$ than in $T1$. This can be explained by the fact that the realizations of the variables are more volatile in $T1$ than in $T2$. More volatile realizations and more volatile forecasts naturally go together. When looking at the data, groups' average forecasts are indeed more volatile in $T1$ than in $T2$. Inflation forecast volatility is 0.281

[17]For the pairwise comparisons of the heuristic switching model with the other models, two-sided Wilcoxon rank-sum tests yield the following results. The difference between the HSM and the homogeneous rational agent solution is significant with $p < 10^{-4}$ and $p < 10^{-5}$ for inflation and output gap forecasting, respectively. The differences between the HSM and adaptive expectations or strong trend-following are statistically significant for both inflation and output-gap forecasting ($p = 0.006$ and 0.011 for the comparisons with ADA and $p < 10^{-3}$ and $p = 0.001$ for the comparisons with STR). The comparisons with the weak trend-following rule and the anchoring and adjustment rule are not statistically significant ($p = 0.383$ and $p = 0.211$ for WTR and 0.146 and 0.139 for LAA).

in $T1$ and 0.134 in $T2$, output gap forecast volatility is 0.464 in $T1$ and 0.096 in T2. These differences are statistically significant when tested with two-sided Wilcoxon rank-sum tests (the p-values are 0.006 for inflation and $< 10^{-3}$ for the output gap). It is not surprising that the models have a harder time accurately describing subjects' forecasts when these are more volatile.

The comparison of prediction errors supports the switching model. In the following, we consider the model fitted to the experimental data. This is useful to understand which heuristics are employed and whether there are patterns concerning the use of the heuristics over time. Figure 10 shows the fractions of the heuristics over time for inflation and output gap in $T1$ and $T2$ (the lines represent averages across groups).

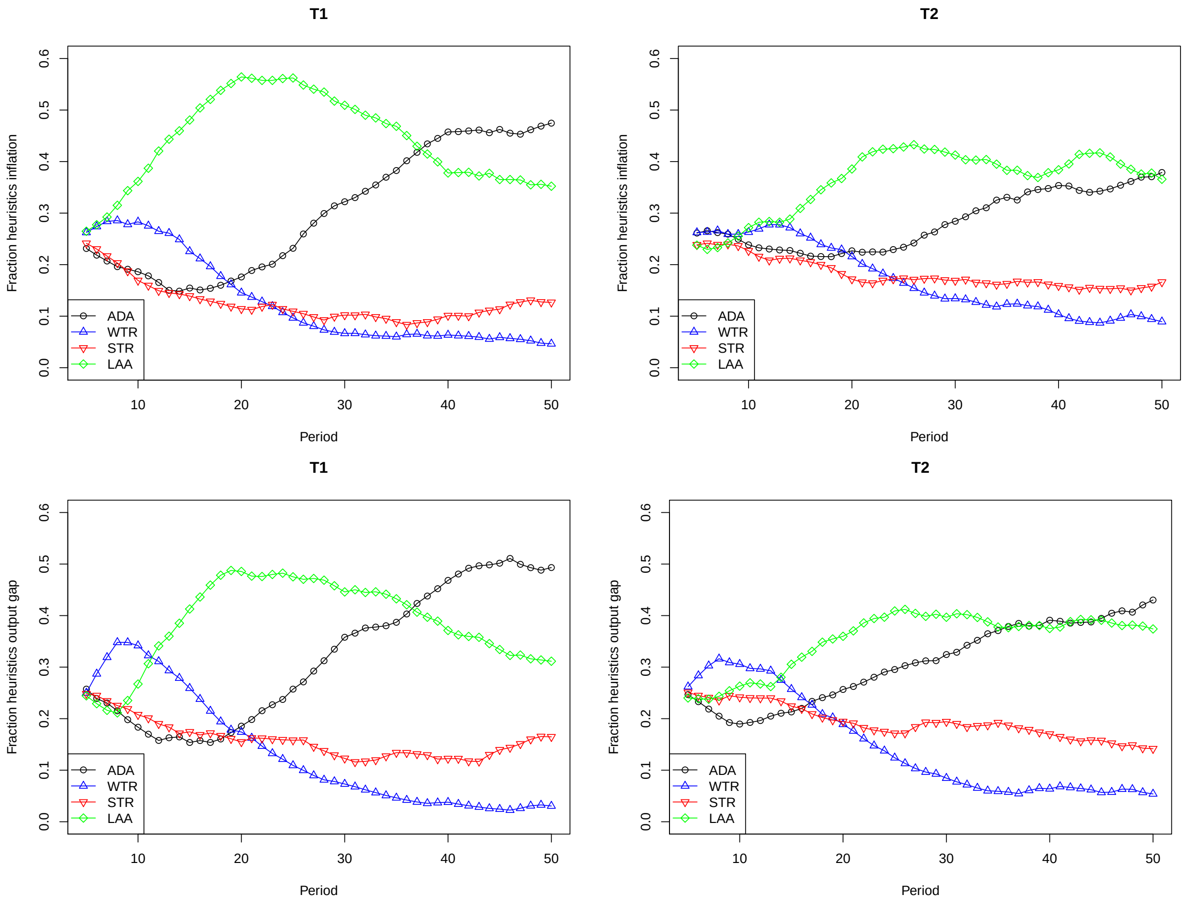


Figure 10: Heuristics employed in inflation and output gap forecasting

Notes: This figure shows averages across groups of the fractions of heuristics employed over time in the model fitted to the experimental data. On the left side are the fractions of inflation and output gap for $T1$. $T2$ is on the right.

One can see that all heuristics have some support in the experiment. One can also see that the graphs of inflation and output gap forecasting in $T1$ are very similar. The same holds for the corresponding graphs in $T2$. This suggests that while subjects learn and update over time the way that they form expectations, they form expectations on inflation and output gap in similar ways, and change how they form expectations on inflation and output gap in similar ways. This is not self-evident: it could well have been the case that subjects rely more on trend extrapolation for one variable while behaving adaptively when forecasting the other variable. However, we do not see such behavior.

Regarding the use of the heuristics themselves, the adaptive rule and the anchoring and adjustment rule are used more often than the trend-following rules. The use of the adaptive rule increases over the course of the experiment, partially explaining the learning observed in the experiment. Furthermore, the trend-following rules are used less and less as the experiment proceeds (for inflation and output gap alike in both treatments; there are some small upward movements toward the end of the experiment in $T1$, however). This also contributes to the stability of inflation and output gap in the second half of the experiment, as the trend-following rules are destabilizing. The use of the anchoring and adjustment rule follows a less-clear pattern. It increases strongly in the beginning in $T1$ and decreases again thereafter. In $T2$ the use of this rule increases more slowly; afterwards, it levels off. This rule, which has two components, has a less clear-cut interpretation than the other rules. The first component is destabilizing, taking into account last trends rather than predicting a return to the anchor immediately, while the second component, the anchor itself, is very stabilizing, as the long-run averages are very close to the inflation target and the steady state of the output gap. It is interesting to see that, overall, relatively little use is made of the weak trend-following rule. While this rule alone predicts group level aggregates of forecasts rather well (Table 2), when looking at it from the point of view of the heuristic switching model, it seems that this is only the case because it gives a relatively good approximation of the prevailing mixes of the whole set of heuristics.

4 Concluding Remarks

We have conducted a learning-to-forecast experiment to test the predictions of a macroeconomic model with behavioral expectations. This behavioral model yields results that differ from those of the same macroeconomic model based on rational expectations. Namely, the behavioral model yields that inflation volatility can be reduced if the central bank reacts to the output gap on top of inflation. The predictions of the behavioral model are supported by the outcomes of our experiment, in which the only treatment variation consists in a modification of the central bank's monetary policy reaction function.

These results are relevant for monetary policy analysis. They show a different relationship between inflation and output-gap than is usually assumed. The policy implications are particularly straightforward for central banks that aim at price stability alone, such as, for example, the ECB; these banks should react to the output gap even if they are ultimately only interested in price stability.

References

Adam, K. (2007). Experimental Evidence on the Persistence of Output and Inflation. *The Economic Journal*, 117(520):603–636.

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

Anufriev, M. and Hommes, C. H. (2012). Evolutionary selection of individual expectations and aggregate outcomes. *American Economic Journal: Microeconomics*, 4:35–64.

Arifovic, J. and Sargent, T. (2003). Laboratory experiments with an expectational phillips curve. In *Evolution and Procedures in Central Banking*, pages 23–55. Cambridge University Press.

Arthur, W. B., Holland, J. H., LeBaron, B., Palmer, R. G., and Tayler, P. (1997). Asset pricing under endogenous expectations in an artificial stock market. In Arthur, W. B., Durlauf, S., and Lane, D., editors, *The Economy as an Evolving Complex System II*. Addison-Wesley.

Assenza, T., Bao, T., Hommes, C., and Massaro, D. (2014a). Experiments on expectations in macroeconomics and finance. In Duffy, J., editor, *Experiments in Macroeconomics*, volume 17 of *Research in Experimental Economics*.

Assenza, T., Heemeijer, P., Hommes, C., and Massaro, D. (2014b). Managing self-organization of expectations through monetary policy: a macro experiment. CeNDEF Working Paper 14-07, University of Amsterdam.

Bao, M., Hommes, C., Sonnemans, J., and Tuinstra, J. (2012). Individual expectations, limited rationality and aggregate outcomes. *Journal of Economic Dynamics and Control*, 36:1101–1120.

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

Branch, W. (2004). The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations. *The Economic Journal*, 114(497):592–621.

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

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

Branch, W. A., Carlson, J., Evans, G. W., and McGough, B. (2009). Monetary policy, endogenous inattention and the volatility trade-off. *The Economic Journal*, 119(534):123–157.

Brock, W. A. and Hommes, C. H. (1997). A rational route to randomness. *Econometrica*, 65(5):1059–1095.

Brock, W. A. and Hommes, C. H. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 22:1235–1274.

Camerer, C. F., Loewenstein, G., and Rabin, M. (2011). *Advances in behavioral economics*. Princeton University Press.

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

Clarida, R., Gali, J., and Gertler, M. (2000). Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory. *Quarterly Journal of Economics*, 115(1):147–180.

Cornand, C. and Heinemann, F. (2014). Experiments on monetary policy and central banking. In Duffy, J., editor, *Experiments in Macroeconomics*, volume 17 of *Research in Experimental Economics*, pages 167–227.

Cornand, C. and M'Baye, C. K. (2016). Does inflation targeting matter? An experimental investigation. *Macroeconomic Dynamics (forthcoming)*.

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

De Grauwe, P. (2012a). Booms and busts in economic activity: A behavioral explanation. *Journal of Economic Behavior & Organization*, 83(3):484–501.

De Grauwe, P. (2012b). *Lectures on behavioral macroeconomics*. Princeton University Press, New Jersey.

De Grauwe, P. and Kaltwasser, P. R. (2012). Animal spirits in the foreign exchange market. *Journal of Economic Dynamics and Control*, 36(8):1176–1192.

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

Diks, C. and van der Weide, R. (2005). Herding, a-synchronous updating and heterogeneity in memory in a cbs. *Journal of Economic Dynamics and Control*, 29:741–763.

Duffy, J. (2012). Macroeconomics: A survey of laboratory research. Working Paper 334, University of Pittsburgh.

Duffy, J. and Nagel, R. (1997). On the robustness of behaviour in experimental 'beauty contest' games. *The Economic Journal*, 107(445):1684–1700.

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

Falk, A. and Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science*, 326(5952):535–538.

Frankel, J. and Froot, K. (1987). Using survey data to test standard propositions regarding exchange rate expectations. *American Economic Review*, 77(1):133–153.

Galí, J. (2008). *Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton University Press.

García-Schmidt, M. and Woodford, M. (2015). Are low interest rates deflationary? A paradox of perfect-foresight analysis. *NBER Working Paper 21614*.

Gigerenzer, G. and Selten, R. (2002). *Bounded rationality: The adaptive toolbox*. Mit Press.

Gigerenzer, G. and Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press, USA.

Grether, D. and Plott, C. (1979). Economic theory of choice and the preference reversal phenomenon. *American Economic Review*, 69(4):623–638.

Heemeijer, P., Hommes, C. H., Sonnemans, J., and Tuinstra, J. (2009). Price stability and volatility in markets with positive and negative expectations feedbacks: An experimental investigation. *Journal of Economic Dynamics and Control*, 33:1052–1072.

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

Hommes, C. H., Huang, H., and Wang, D. (2005a). A robust rational route to randomness in a simple asset pricing model. *Journal of Economic Dynamics and Control*, 29:1043–1072.

Hommes, C. H., Sonnemans, J., Tuinstra, J., and van de Velden, H. (2005b). Coordination of expectations in asset pricing experiments. *Review of Financial Studies*, 18(3):955–980.

Hommes, C. H., Sonnemans, J., Tuinstra, J., and van de Velden, H. (2008). Expectations and bubbles in asset pricing experiments. *Journal of Economic Behavior and Organization*, 67(1):116–133.

Huck, S., Normann, H., and Oechssler, J. (2004). Two are few and four are many: number effects in experimental oligopolies. *Journal of Economic Behavior and Organization*, 53:435–446.

Kelley, H. and Friedman, D. (2002). Learning to forecast price. *Economic Inquiry*, 40(4):556–573.

Kryvtsov, O. and Petersen, L. (2013). Expectations and monetary policy: Experimental evidence. Working Papers 13-44, Bank of Canada.

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

Lei, V. and Noussair, C. N. (2002). An experimental test of an optimal growth model. *American Economic Review*, pages 549–570.

Manski, C. and McFadden, D. (1981). Alternative estimators and sample designs for discrete choice analysis. In Manski, C. and McFadden, D., editors, *Structural Analysis of Discrete Data*. MIT Press.

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

Marimon, R. and Sunder, S. (1993). Indeterminacy of equilibria in a hyperinflationary world: Experimental evidence. *Econometrica*, 61:1073–1107.

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

Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–298.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review*, 85(5):1313–1326.

Orphanides, A. and Williams, J. C. (2006). Monetary policy with imperfect knowledge. *Journal of the European Economic Association*, 4(2-3):366–375.

Pfajfar, D. and Zakelj, B. (2014). Experimental evidence on inflation expectations formation. *Journal of Economic Dynamics and Control*, 44:147–168.

Schram, A. (2005). Artificiality: The tension between internal and external validity in economic experiments. *Journal of Economic Methodology*, 12(2):225–237.

Srour, G. (2001). Why do central banks smooth interest rates? *Bank of Canada Working Paper 2001-17*.

Tversky, A. and Kahneman, D. (1974). Judgement under uncertainty: Heuristics and biases. *Science*, 185:1124–1130.

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

Woodford, M. (2010). Robustly optimal monetary policy with near-rational expectations. *American Economic Review*, 100(1):274–303.

Woodford, M. (2013). Macroeconomic Analysis Without the Rational Expectations Hypothesis. *Annual Review of Economics*, 5(1):303–346.

Yun, T. (1996). Nominal price rigidity, money supply endogeneity, and business cycles. *Journal of Monetary Economics*, 37(2):345–370.

A Appendix (for Online Publication): Microfoundations of the Behavioral Macroeconomic Model

The following derivation follows the work of Kurz et al. (2013). The economy is populated by a continuum of households-producers indexed by j. Agents are identical except for the fact that they may have different expectations about future macroeconomic variables. Household j thus chooses consumption C_t^j, labor L_t^j, and bond holdings B_t^j to maximize

$$E_t^j \sum_{\tau=0}^{\infty} \rho^\tau \left(\frac{(C_{t+\tau}^j)^{1-\sigma}}{1-\sigma} - \frac{(L_{t+\tau}^j)^{1+\eta}}{1+\eta} - \frac{\tilde{\tau}_b}{2} \left(\frac{B_{t+\tau}^j}{P_{t+\tau}} \right)^2 \right) ,$$

subject to the budget constraint

$$C_t^j + \frac{B_t^j}{P_t} = \frac{W_t}{P_t} L_t^j + \left(\frac{B_{t-1}^j R_{t-1}}{P_{t-1}} \right) \frac{P_{t-1}}{P_t} + \frac{T_t^j}{P_t} .$$

E_t^j denotes the subjective expectations of household j, ρ is the discount factor, W_t is the nominal wage, R_t is gross interest, P_t is the aggregate price level, and T_t^j are lump sum transfers including profit from firms. We assume that B_0^j is given and that there is no aggregate debt. As in Kurz et al. (2013) we include a penalty term $\tilde{\tau}_b$ in the utility function in place of institutional constraints to limit borrowing (with sufficiently small values of $\tilde{\tau}_b$ solutions with explosive borrowing are not equilibria). The first order conditions are given by

$$\tilde{\tau}_b \frac{B_t^j}{P_t} + (C_t^j)^{-\sigma} = E_t^j \left(\rho (C_{t+1}^j)^{-\sigma} \frac{R_t}{(P_{t+1}/P_t)} \right) \tag{A.1}$$

$$(C_t^j)^{-\sigma} \frac{W_t}{P_t} = (L_t^j)^\eta . \tag{A.2}$$

For a generic variable X_t we denote the steady-state value by $\bar{X}$ and define $\hat{x}_t = (X_t - \bar{X})/\bar{X}$, while for bond holdings we define $\hat{b}_t = B_t/(P_t \bar{Y})$ (with a steady-state value of zero). Denoting gross inflation P_t/P_{t-1} as Π_t and log-linearizing Equation (A.1) around a zero inflation steady state we get

$$\hat{c}_t^j = E_t^j \hat{c}_{t+1}^j - \sigma^{-1}(\hat{R}_t - E_t^j \hat{\Pi}_{t+1}) + \tau_b \hat{b}_t^j ,$$

where $\tau_b \equiv \tilde{\tau}_b (\bar{Y})^{1+\sigma}$ (using the steady-state relations $\bar{C}^j = \bar{Y}^j = \bar{C} = \bar{Y}$). It is worth remarking at this point that we log-linearize the system around a zero inflation steady state for the sake of algebraic simplicity. However, as argued in Woodford (2003) and further discussed below, this does not imply that we can only consider policies under which the inflation target is zero, as long as the target inflation rate is not too large. Rewriting the individual consumption function above as

$$\hat{c}_t^j = E_t^j \hat{c}_{t+1} + (E_t^j \hat{c}_{t+1}^j - E_t^j \hat{c}_{t+1}) - \sigma^{-1}(\hat{R}_t - E_t^j \hat{\Pi}_{t+1}) + \tau_b \hat{b}_t^j \ ,$$

where $\hat{c}_t = \int \hat{c}_t^j dj$ is aggregate consumption and using both the aggregate market clearing condition $\hat{c}_t = \hat{y}_t$ and the fact that $\hat{b}_t = 0$, we can aggregate the individual consumption functions to get

$$\hat{y}_t = \bar{E}_t \hat{y}_{t+1} - \sigma^{-1}(\hat{R}_t - \bar{E}_t \hat{\Pi}_{t+1}) + \Phi_t(\hat{c}) \ , \tag{A.3}$$

where $\bar{E}_t$ is the aggregate expectation operator defined as $\bar{E}_t(x_{t+1}) = \int E_t^j x_{t+1} dj$ for a generic variable x and the term $\Phi_t(\hat{c}) = \int (E_t^j \hat{c}_{t+1}^j - E_t^j \hat{c}_{t+1}) dj$ denotes the difference between the average expectation of individual consumption and average consumption.

We now turn to the supply side of the economy. Final consumption of household j is composed by intermediate goods, indexed by i and produced by a continuum of monopolistically competitive firms so that

$$C_t^j = \left(\int_0^1 (C_{it}^j)^{\frac{\theta-1}{\theta}} di \right)^{\frac{\theta}{\theta-1}} \ ,$$

with $\theta > 1$. Individual demand of good i is therefore given by

$$C_{it}^j = \left(\frac{P_{it}}{P_t} \right)^{-\theta} C_t^j \ ,$$

where P_{it} is the price of good i and P_t denotes the aggregate price level defined as

$$P_t = \left(\int_0^1 P_{it}^{1-\theta} di \right)^{\frac{1}{1-\theta}} \ .$$

Aggregating demand for each good i over households and using the aggregate market clearing condition $C_t = Y_t$, we get

$$C_{it} = \left(\frac{P_{it}}{P_t} \right)^{-\theta} Y_t \ .$$

Each firm has a linear production technology using labor as only input

$$Y_{it} = A_t N_{it} \ , \tag{A.4}$$

where A_t is the aggregate productivity. Given the production function we can write the expression for real marginal costs as

$$mc_t = \frac{W_t}{A_t P_t} , \tag{A.5}$$

so that individual real profits can be expressed as

$$\Psi_{it} = \left(\left(\frac{P_{it}}{P_t} \right)^{1-\theta} - mc_t \left(\frac{P_{it}}{P_t} \right)^{-\theta} \right) Y_t .$$

We assume a staggered price setting as in the Calvo model, where only a fraction $1-\omega$ of prices are readjusted in every period. Moreover, we consider a scenario in which households have equal ownership shares in all firms (so that income effects of random price adjustments are removed), though each household j manages only one firm (i.e., makes price decisions for only one firm). Since each firm produces a single good and is managed by a single household j (with subjective expectations j), we can, without loss of generality, use a single index, say j, to denote the produced good and the subjective expectations. A firm j adjusting its price in period t maximizes (given its subjective expectations) the present discounted value of profits in all future states prior to the next price readjustment

$$E_t^j \sum_{\tau=0}^{\infty} (\omega\rho)^\tau \left(\frac{C_{t+\tau}^j}{C_t^j} \right)^{-\sigma} \left(\left(\frac{P_{jt}}{P_{t+\tau}} \right)^{1-\theta} - mc_{t+\tau} \left(\frac{P_{jt}}{P_{t+\tau}} \right)^{-\theta} \right) Y_{t+\tau} .$$

In this expression, $\rho^\tau (C_{t+\tau}^j / C_t^j)^{-\sigma}$ is the stochastic discount factor of household j managing the firm. Defining $q_{jt}^* = P_{jt}^* / P_t$ as the optimal price set by firm j relative to the aggregate price level, we can write the first order condition as

$$q_{jt}^* = \frac{\theta}{\theta - 1} \frac{E_t^j \sum_{\tau=0}^{\infty} (\omega\rho)^\tau \left(C_{t+\tau}^j \right)^{-\sigma} Y_{t+\tau} \left(\frac{P_{t+\tau}}{P_t} \right)^{\theta} mc_{t+\tau}}{E_t^j \sum_{\tau=0}^{\infty} (\omega\rho)^\tau \left(C_{t+\tau}^j \right)^{-\sigma} Y_{t+\tau} \left(\frac{P_{t+\tau}}{P_t} \right)^{\theta-1}} . \tag{A.6}$$

Log-linearizing Equation (A.6) and using the steady-state relation $\overline{mc} = (\theta - 1)/\theta$ we get the individual pricing rule

$$\hat{q}_{jt}^* = (1 - \omega\rho) E_t^j \sum_{\tau=0}^{\infty} (\omega\rho)^\tau (\widehat{mc}_{t+\tau} + \hat{\Pi}_{t+\tau}) . \tag{A.7}$$

Assuming, as standard in the literature, that the law of iterated expectations holds at the individual level (see e.g. Evans and Honkapohja, 2001, Branch and McGough, 2009, and Kurz et al., 2013), we can rewrite Equation (A.7) as

$$\hat{q}^*_{jt} = (1-\omega\rho)\widehat{mc}_t + \omega\rho E^j_t(\hat{q}^*_{jt+1} + \hat{\Pi}_{t+1}) . \tag{A.8}$$

Given the Calvo pricing scheme, in each period only a set of firms $S_t \in [0,1]$ of measure $1-\omega$ adjust prices, while a set $S^c_t \in [0,1]$ of measure ω do not adjust. We assume that the sample of firms allowed to adjust prices in each period is selected independently across agents, so that the distribution of subjective expectations is the same for firms that adjust prices and for those that do not. Using the aggregate price definition we can then write

$$P^{1-\theta}_t = \int_{S_t} (P^*_{jt})^{1-\theta} dj + \int_{S^c_t} (P_{jt-1})^{1-\theta} dj ,$$

which can be rewritten as

$$1 = (1-\omega) \int (q^*_{jt})^{1-\theta} dj + \omega(\Pi_t)^{\theta-1} .$$

Log-linearizing the above relation we get

$$\hat{\Pi}_t = \frac{1-\omega}{\omega} \int \hat{q}^*_{jt} dj . \tag{A.9}$$

Denoting $\hat{q}_t = \int \hat{q}^*_{jt} dj$ and integrating Equation (A.8) on both sides we get

$$\hat{q}_t = (1-\omega\rho)\widehat{mc}_t + \omega\rho \int E^j_t(\hat{q}^*_{jt+1} + \hat{\Pi}_{t+1}) dj ,$$

which can be rewritten as

$$\hat{q}_t = (1-\omega\rho)\widehat{mc}_t + \omega\rho \int E^j_t(\hat{q}^*_{jt+1} + \hat{q}_{t+1} - \hat{q}_{t+1} + \hat{\Pi}_{t+1}) dj .$$

Recalling from Equation (A.9) that $\hat{q}_t = \omega/(1-\omega)\hat{\Pi}_t$ and substituting it in the equation above we get

$$\hat{\Pi}_t = \frac{(1-\omega)(1-\omega\rho)}{\omega}\widehat{mc}_t + \rho\bar{E}_t\hat{\Pi}_{t+1} + \Phi_t(\hat{q}) , \tag{A.10}$$

where again $\bar{E}_t$ is the aggregate expectation operator and $\Phi_t(\hat{q}) = \rho(1-\omega)\int(E^j_t\hat{q}^*_{jt+1} - E^j_t\hat{q}_{t+1})dj$ denotes the difference between the average expectation of the individual price and the average price.

Log-linearizing Equations (A.2), (A.4), and (A.5) and combining them with market clearing yields the following expression for real marginal costs as a function of output and productivity:

$$\widehat{mc}_t = (\sigma+\eta)\hat{y}_t - (1+\eta)\hat{a}_t \; . \tag{A.11}$$

Equation (A.11) implies a natural level of output under flexible prices given by

$$\hat{y}^n_t = \frac{1+\eta}{\sigma+\eta}\hat{a}_t \; .$$

Plugging Equation (A.11) into Equation (A.10) and defining the output gap as $y_t = \hat{y}_t - \hat{y}^n_t$ results in

$$\hat{\Pi}_t = \lambda y_t + \rho \bar{E}_t \hat{\Pi}_{t+1} + \Phi_t(\hat{q}) \; , \tag{A.12}$$

where $\lambda = (1-\omega)(1-\omega\rho)(\sigma+\eta)/\omega$. Rewriting Equation (A.3) in terms of the output gap yields

$$y_t = \bar{E}_t y_{t+1} - \sigma^{-1}(\hat{R}_t - \bar{E}_t\hat{\Pi}_{t+1}) + \Phi_t(\hat{c}) + g_t \; , \tag{A.13}$$

where $g_t = (\hat{a}_{t+1} - \hat{a}_t)(1+\eta)/(\sigma+\eta)$. Equations (A.12) and (A.13) describe a New Keynesian Phillips curve and a dynamic IS relation expressed in a general form which is consistent with arbitrary subjective expectations.

Next, we simplify Equations (A.12) and (A.13) given the behavioral model of expectation formation outlined in Section 2.2. The behavioral model assumes that agents deviate from fully rational behavior by using the described heuristics to forecast future output gap and inflation. We assume that using the heuristics for these forecasts is the only source of irrationality. More precisely, we assume that agents are not irrational when forming expectations about their own future consumption relative to average consumption and the price set by the firm managed by them relative to the average price. This implies that the terms $\Phi_t(\hat{c})$ and $\Phi_t(\hat{q})$ are equal to zero.[18] We can therefore rewrite the aggregate demand and supply equations as

$$\begin{aligned} y_t &= \bar{E}_t y_{t+1} - \sigma^{-1}(\hat{R}_t - \bar{E}_t\hat{\Pi}_{t+1}) + g_t && \text{(A.14)} \\ \hat{\Pi}_t &= \lambda y_t + \rho\bar{E}_t\hat{\Pi}_{t+1} \; . && \text{(A.15)} \end{aligned}$$

Defining π_t as the inflation rate and $i_t \equiv \log(1+yield_t) - \gamma \simeq yield_t - \gamma$, where $yield_t$ denotes the yield on the one period bond and $\gamma \equiv -\log\rho$, we can write Equations (A.14)

[18]These terms also drop out if one assumes instead that agents' expectations of the average future consumption across all agents and of the average price set by all firms equal their expectations of their own future consumption and of the price set by the firm managed by them, respectively.

and (A.15) as

$$
\begin{aligned}
y_t &= \bar{E}_t y_{t+1} - \varphi(i_t - \bar{E}_t \pi_{t+1}) + g_t && \text{(A.16)}\\
\pi_t &= \lambda y_t + \rho \bar{E}_t \pi_{t+1} + u_t \ , && \text{(A.17)}
\end{aligned}
$$

with $\varphi \equiv \sigma^{-1}$ and with a cost-push shock u_t added to the aggregate supply relation. We close the model with a monetary policy rule of the form

$$
i_t = \bar{\pi} + \phi_\pi(\pi_t - \bar{\pi}) + \phi_y(y_t - \bar{y}) \tag{A.18}
$$

when not at the zero lower bound, where $\bar{\pi}$ is the inflation target and $\bar{y} \equiv (1-\rho)\bar{\pi}/\lambda$ is the steady state level of the output gap consistent with the inflation target $\bar{\pi}$.

As argued in Woodford (2003), the fact that the New Keynesian equations above have been log-linearized around a zero inflation steady state does not mean that one can consider only policy rules that involve a target inflation rate of zero. In fact, Equations (A.16) and (A.17) are valid approximations as long as the target inflation is not too large.[19]

The economy can thus be described by Equations (A.16) and (A.17) together with the monetary policy rule in Equation (A.18), potentially subject to the zero lower bound. These equations correspond to Equations (1)-(3) in the main text. When the zero lower bound is not binding, the model can be written in matrix form as

$$
\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \Omega \begin{bmatrix} \varphi\bar{\pi}(\phi_\pi - 1) + \varphi\phi_y\bar{y} \\ \lambda\varphi\bar{\pi}(\phi_\pi - 1) + \lambda\varphi\phi_y\bar{y} \end{bmatrix} + \Omega \begin{bmatrix} 1 & \varphi(1-\phi_\pi\rho) \\ \lambda & \lambda\varphi + \rho + \rho\varphi\phi_y \end{bmatrix} \begin{bmatrix} \bar{y}^e_{t+1} \\ \bar{\pi}^e_{t+1} \end{bmatrix} + \Omega \begin{bmatrix} 1 & -\varphi\phi_\pi \\ \lambda & 1+\varphi\phi_y \end{bmatrix} \begin{bmatrix} g_t \\ u_t \end{bmatrix} ,
$$

where $\Omega \equiv 1/(1+\lambda\varphi\phi_\pi + \varphi\phi_y)$ and $\bar{x}^e_{t+1} \equiv \bar{E}_t x_{t+1}$ denotes average expectation about a generic variable x. The system above describes the law of motion of the output gap and inflation as a function of agents' average expectations on output gap and inflation (the above matrix equation is identical to Equation (4) in the main text).

[19] More precisely, for average inflation rates of order ν (where ν is an expansion parameter characterizing monetary policy such that the average inflation rate is zero for policies with $\nu = 0$), the error in the characterisation of the dynamics of aggregate variables is of order $\mathcal{O}(\|\nu, \xi\|^2)$, where $\|\xi\|$ is a bound on the size of the disturbances in the model (see Woodford, 2003, for details).

B Appendix (for Online Publication): Additional Graphs from Simulations of the Macroeconomic Model

B.1 Changes in the Parameters of the Macroeconomic Model

Figure 11 shows inflation volatility as a function of the output gap reaction coefficient ϕ_y for the model assuming rational expectations, similarly to Figure 1a. The graph now shows multiple coefficients of ϕ_π simultaneously (from top to bottom the lines correspond to ϕ_π-values of 1.4, 1.5, 1.6, and 1.7). Figure 12 shows the same graph for the behavioral model (again the lines correspond to ϕ_π-values of 1.4, 1.5, 1.6, and 1.7, from top to bottom).

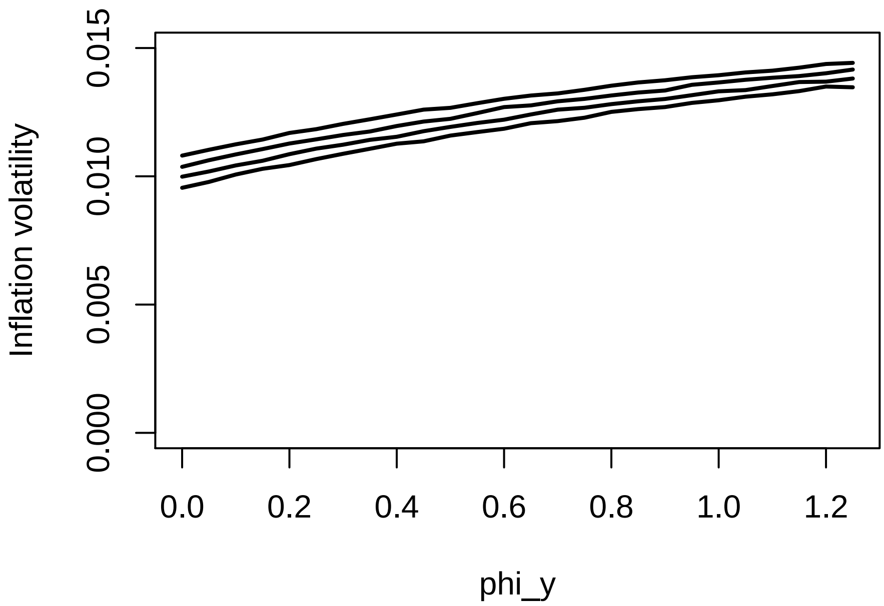


Figure 11: Inflation volatility as a function of ϕ_y for the rational model for different values of ϕ_π (from 1.4 (top line) to 1.7)

Figures 13-16 show inflation volatility as a function of the output gap reaction coefficient similar to Figure 1 for different parameter values of the macroeconomic model. Parameters which are not specifically mentioned in a graph are fixed to the same standard values as used for the graphs in Section 2.3. Under rational expectations, inflation volatility increases monotonously in the output gap reaction coefficient in all of the graphs, similarly to Figure 1a. Under behavioral expectations the U-shape arises in all graphs similarly to Figure 1b.

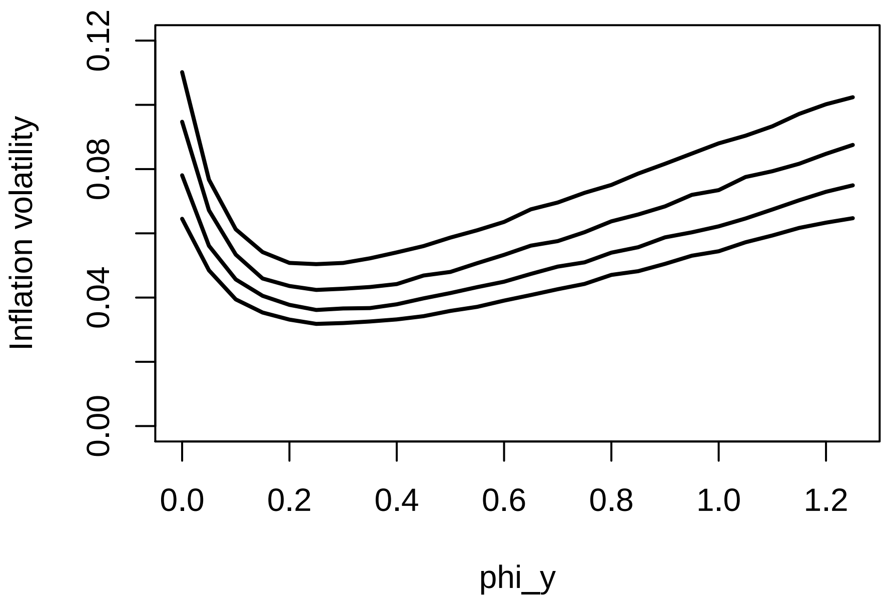
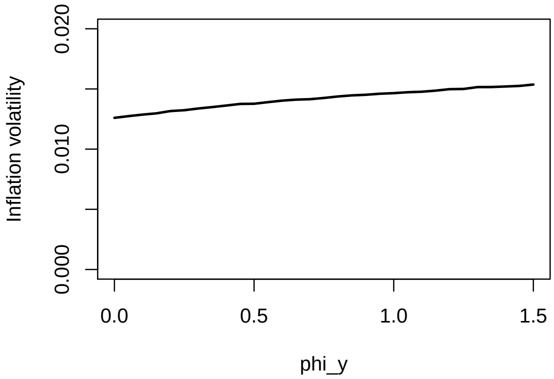


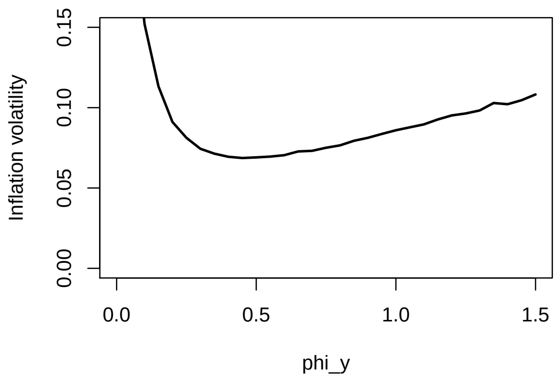
Figure 12: Inflation volatility as a function of ϕ_y for the behavioral model for different values of ϕ_π (from 1.4 (top line) to 1.7)

B.2 Results with Different Behavioral Models of Expectation Formation

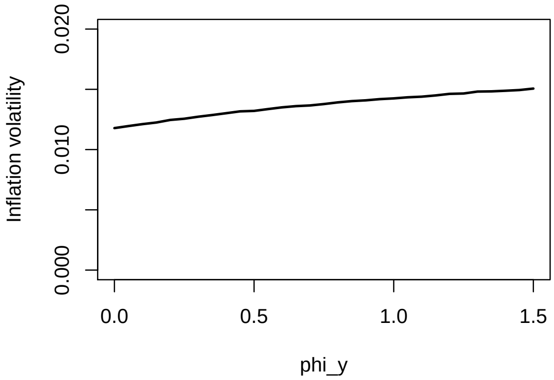
The results are robust to variations of the parameters of the behavioral model of expectation formation we employ. Furthermore, the results are qualitatively the same for a wide variety of other behavioral mechanisms. We show two examples here. Figure 17 shows inflation, output gap, and interest volatility as a function of the output gap reaction coefficient. Expectations are not formed according to the main heuristic switching model described in Section 2.2, but according to two simpler models of behavioral expectation formation. On the left side of this figure, it is assumed that agents use a heuristic switching model similar to the one described before but including only two very simple heuristics, naive expectations which always forecast the last observation and a trend-following rule with trend-following coefficient one. On the right side, the graphs show the results from naive expectations alone (thus without any switching). Here as well, the results look similar to the ones in Figures 1 and 2.



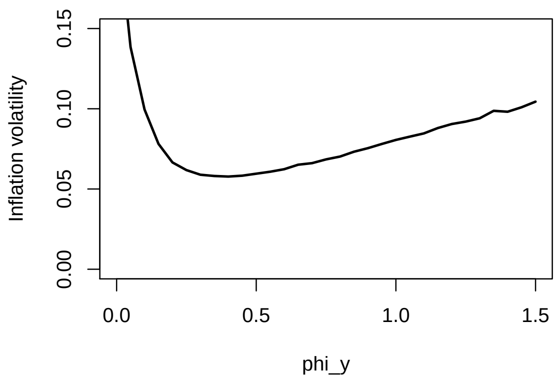
(a) $\varphi = 0.7$, RE



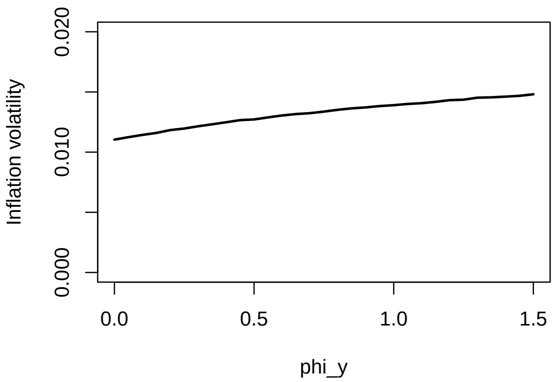
(b) $\varphi = 0.7$, BE



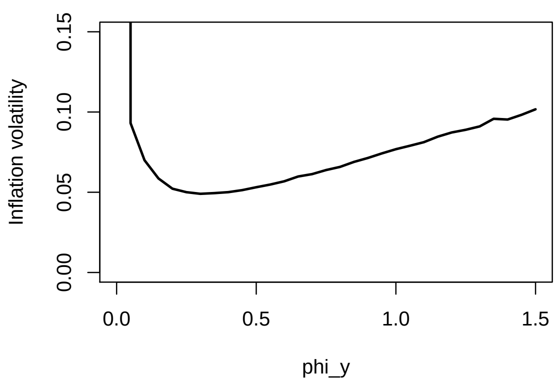
(c) $\varphi = 0.8$, RE



(d) $\varphi = 0.8$, BE



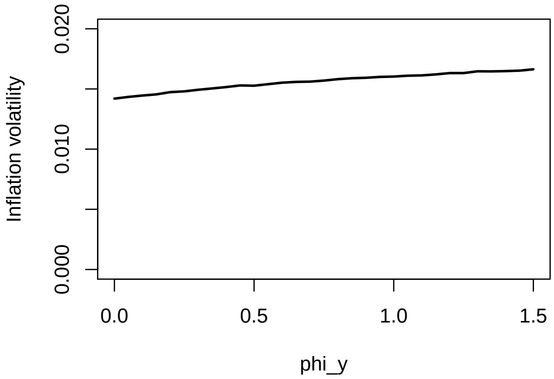
(e) $\varphi = 0.9$, RE



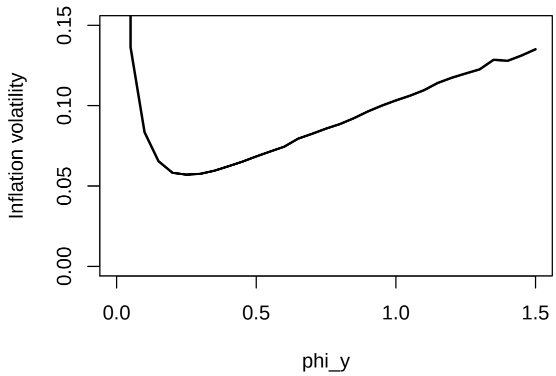
(f) $\varphi = 0.9$, BE

Figure 13: Inflation volatility for different values of φ

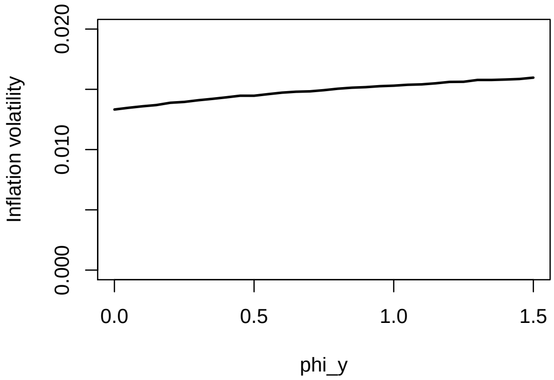
Notes: This figure shows the effect of parameter ϕ_y on inflation volatility for different values of φ ($\phi_\pi = 1.5$ throughout).



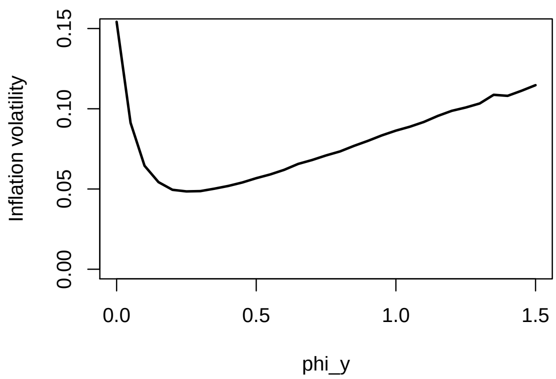
(a) $\lambda = 0.2$, RE



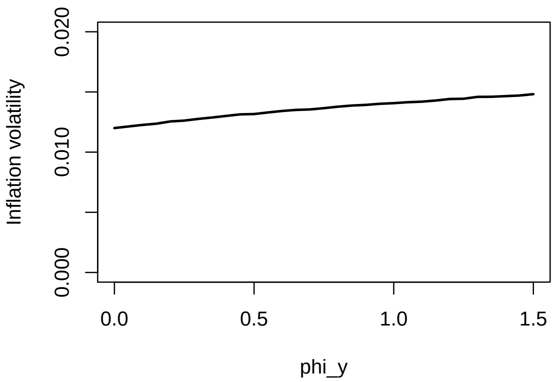
(b) $\lambda = 0.2$, BE



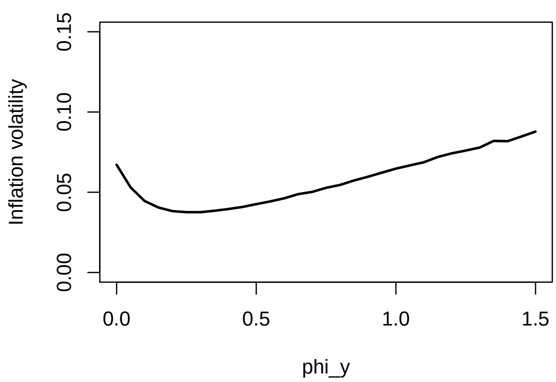
(c) $\lambda = 0.25$, RE



(d) $\lambda = 0.25$, BE



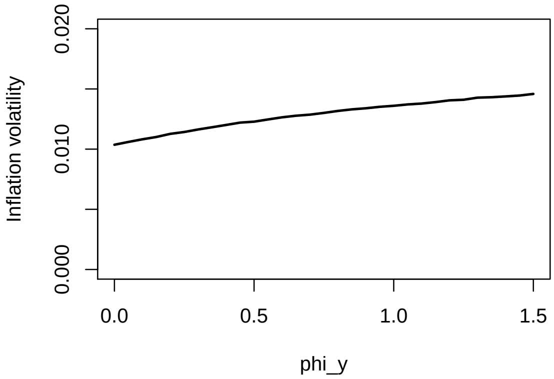
(e) $\lambda = 0.35$, RE



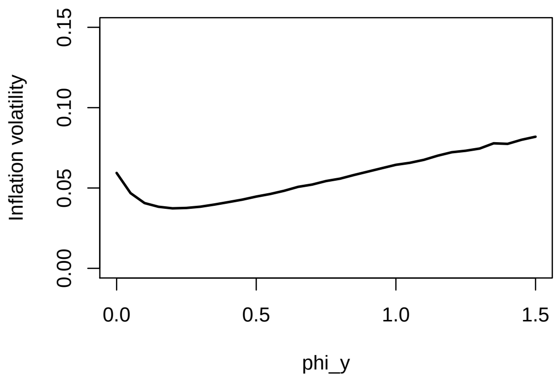
(f) $\lambda = 0.35$, BE

Figure 14: Inflation volatility for different values of λ

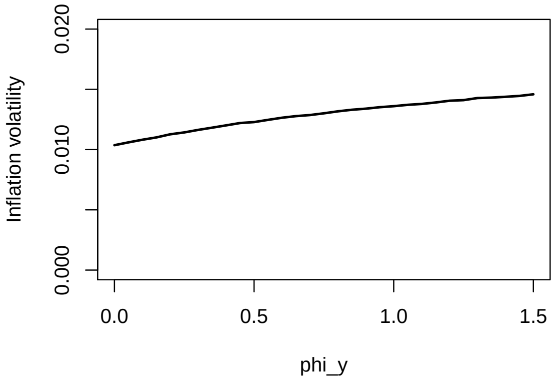
Notes: This figure shows the effect of parameter ϕ_y on inflation volatility for different values of λ ($\phi_\pi = 1.5$ throughout).



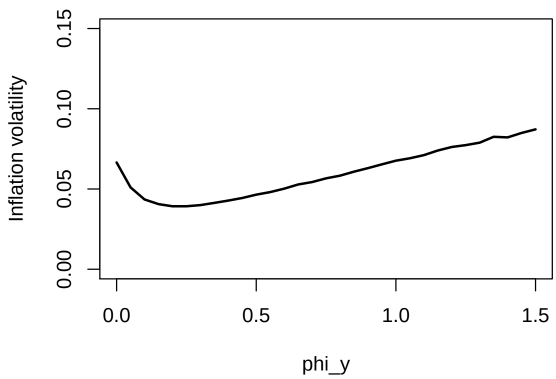
(a) $\rho = 0.96$, RE



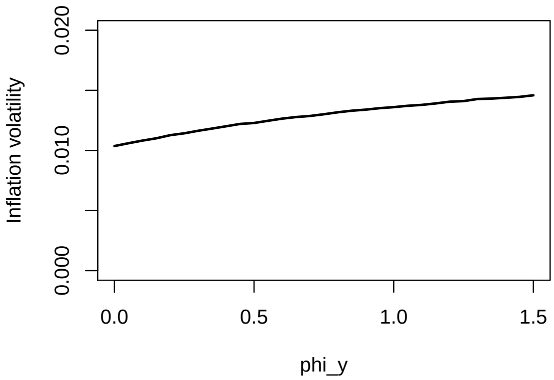
(b) $\rho = 0.96$, BE



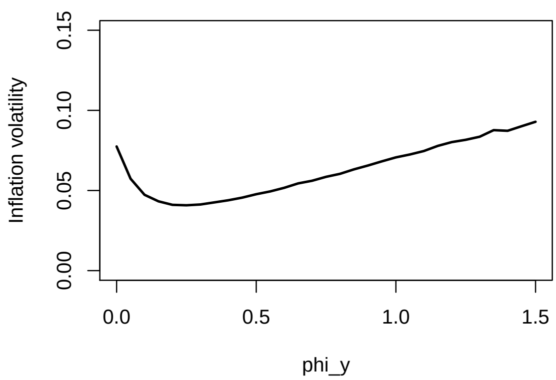
(c) $\rho = 0.97$, RE



(d) $\rho = 0.97$, BE



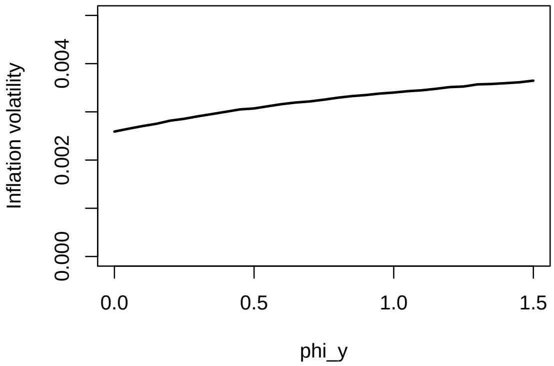
(e) $\rho = 0.98$, RE



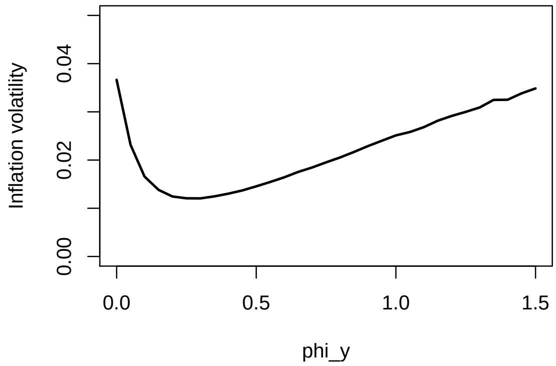
(f) $\rho = 0.98$, BE

Figure 15: Inflation volatility for different values of ρ

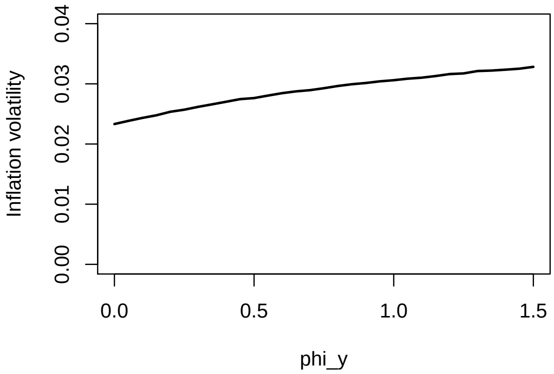
Notes: This figure shows the effect of parameter ϕ_y on inflation volatility for different values of ρ ($\phi_\pi = 1.5$ throughout).



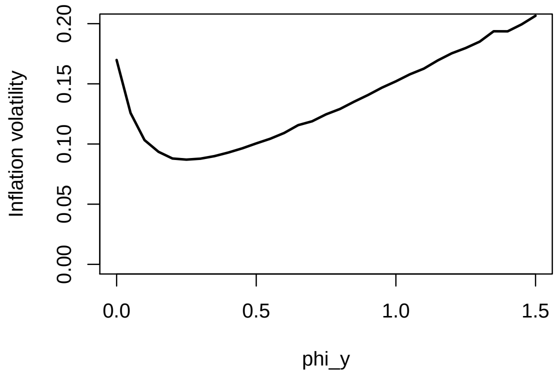
(a) $sd = 0.05$, RE



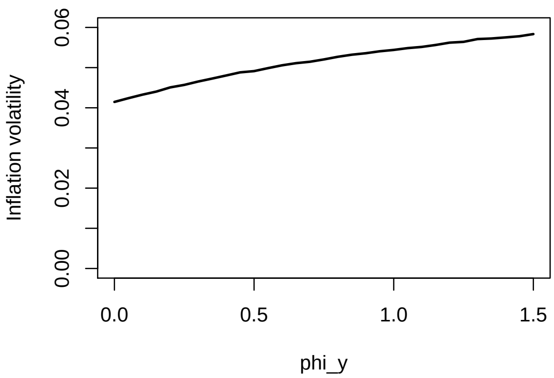
(b) $sd = 0.05$, BE



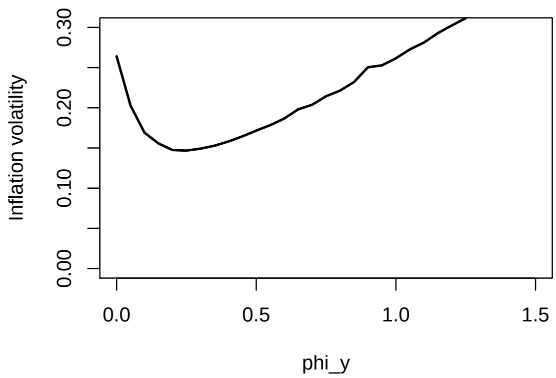
(c) $sd = 0.15$, RE



(d) $sd = 0.15$, BE



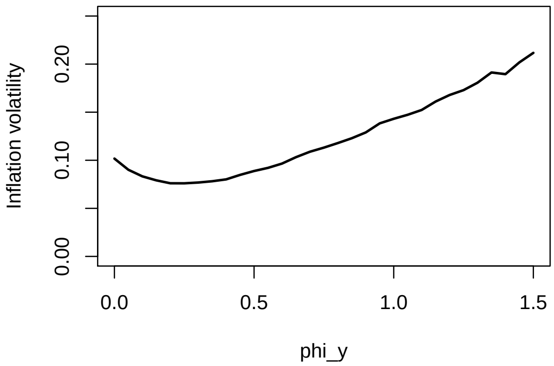
(e) $sd = 0.20$, RE



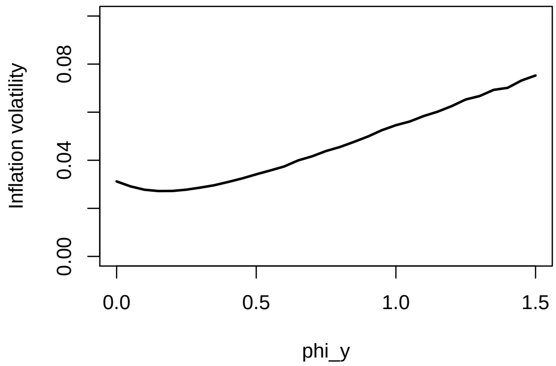
(f) $sd = 0.20$, BE

Figure 16: Inflation volatility for different values of the standard deviation (of both demand and supply shocks)

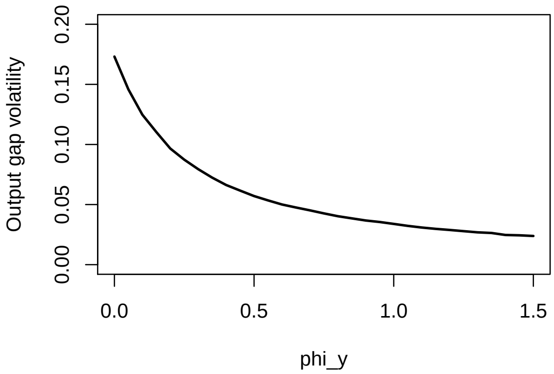
Notes: This figure shows the effect of parameter ϕ_y on inflation volatility for different values of the standard deviation of demand and supply shocks ($\phi_\pi = 1.5$ throughout).



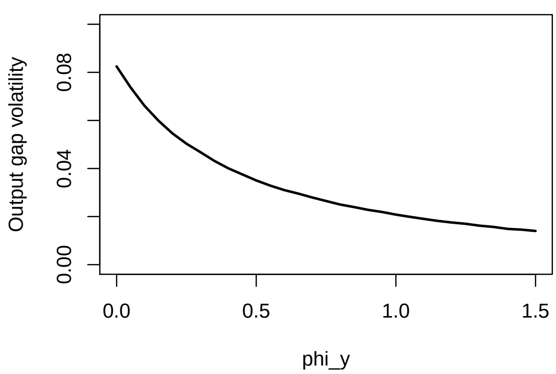
(a) HSM naive + trend



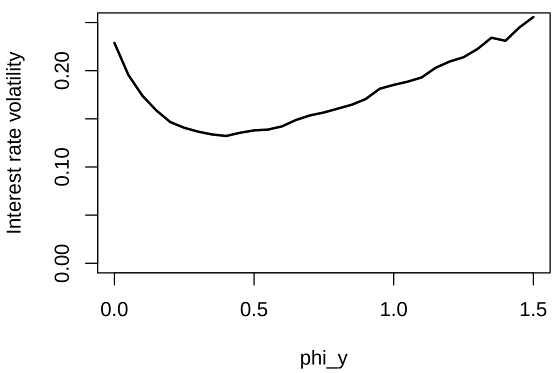
(b) Naive expectations



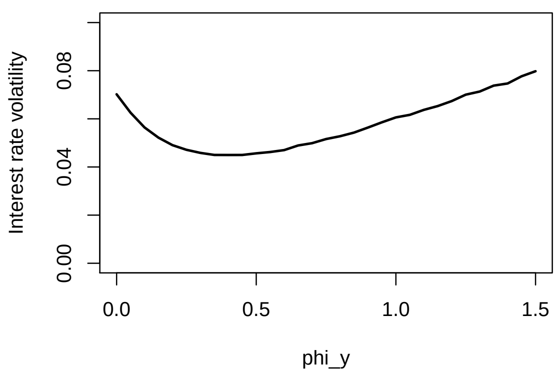
(c) HSM naive + trend



(d) Naive expectations



(e) HSM naive + trend



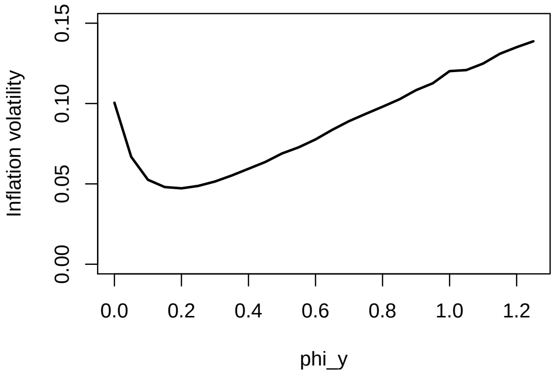
(f) Naive expectations

Figure 17: Inflation, output gap, and interest rate volatility for a simple HSM of expectation formation (with switching between naive expectations and trend-following) and for naive expectations

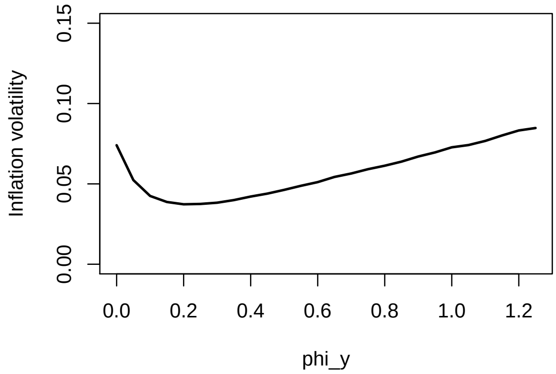
Notes: This figure shows the effect of parameter ϕ_y on inflation, output gap, and interest rate volatility for alternative models of expectation formation ($\phi_\pi = 1.5$ throughout).

B.3 Results with Different Starting Values of Output Gap and Inflation Forecasts

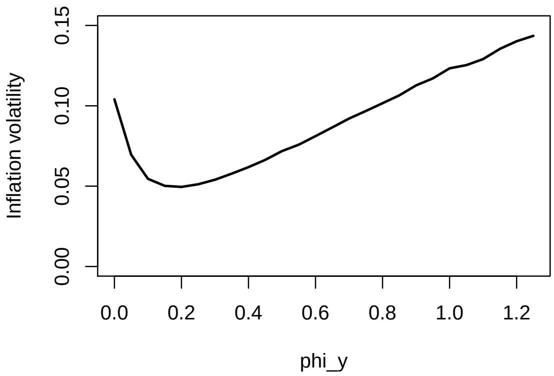
Figures 18 and 19 show graphs similar to Figure 1b for different combinations of starting values of inflation and output gap (i.e. inflation and output gap are set to these starting values in the first two periods). In all cases the U-shape arises similarly to Figure 1b.



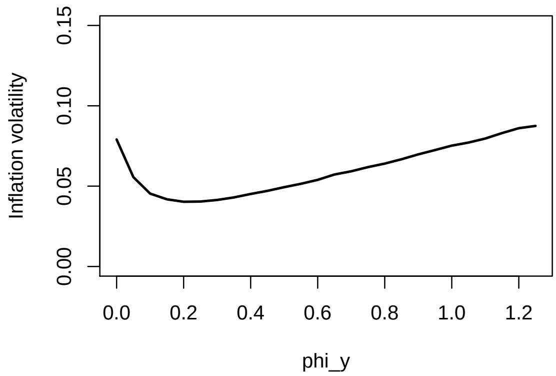
(a) $\pi_{start} = 2.5$, $y_{start} = -0.5$



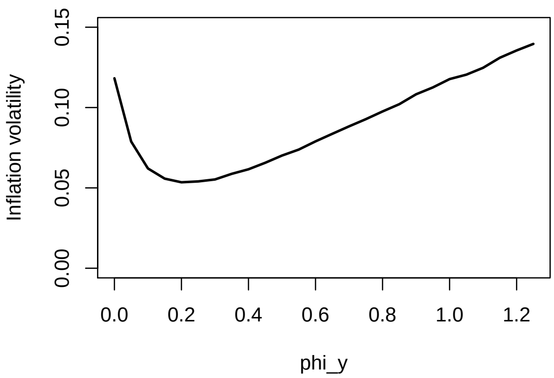
(b) $\pi_{start} = 3.0$, $y_{start} = -0.5$



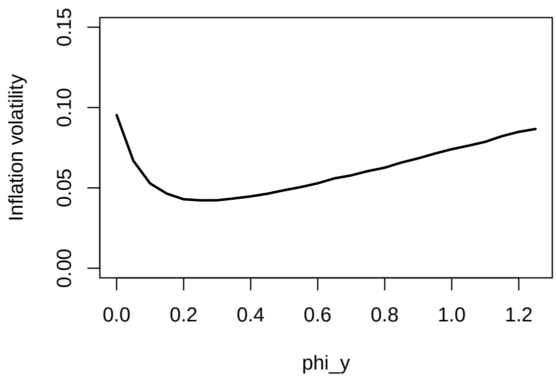
(c) $\pi_{start} = 2.5$, $y_{start} = 0$



(d) $\pi_{start} = 3.0$, $y_{start} = 0$



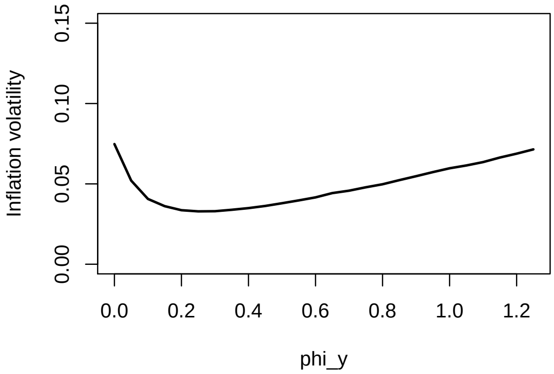
(e) $\pi_{start} = 2.5$, $y_{start} = 0.5$



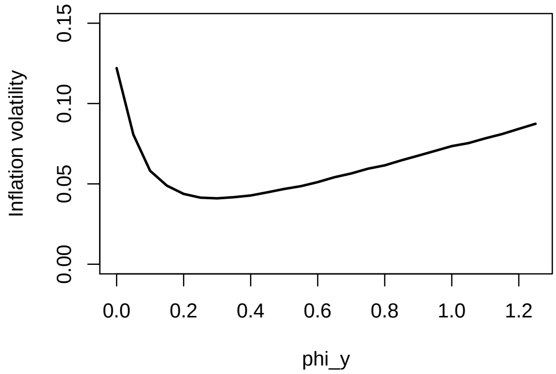
(f) $\pi_{start} = 3.0$, $y_{start} = 0.5$

Figure 18: Inflation volatility in the behavioral model for different starting values

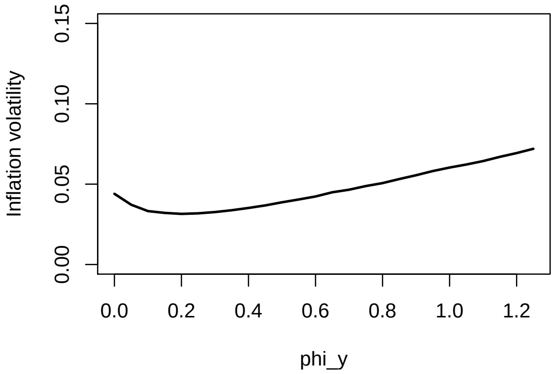
Notes: This figure shows the effect of parameter ϕ_y on inflation volatility for different starting values of y and π ($\phi_\pi = 1.5$ throughout).



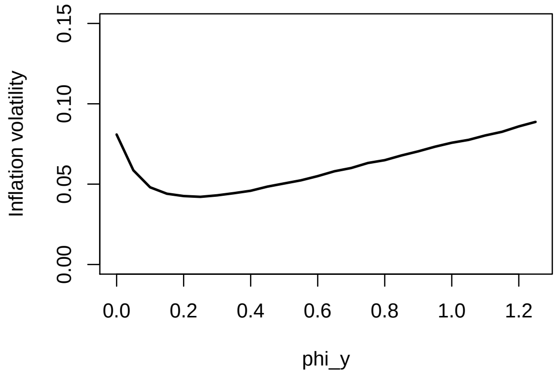
(a) $\pi_{start} = 3.5$, $y_{start} = -0.5$



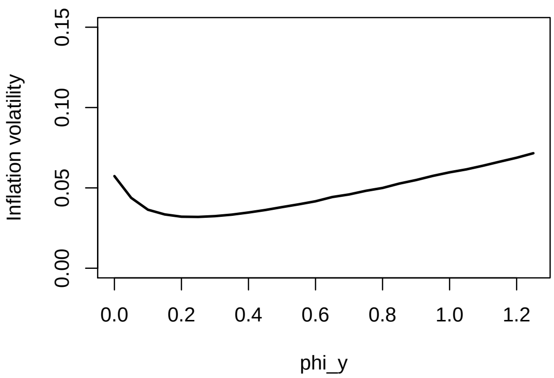
(b) $\pi_{start} = 4.0$, $y_{start} = -0.5$



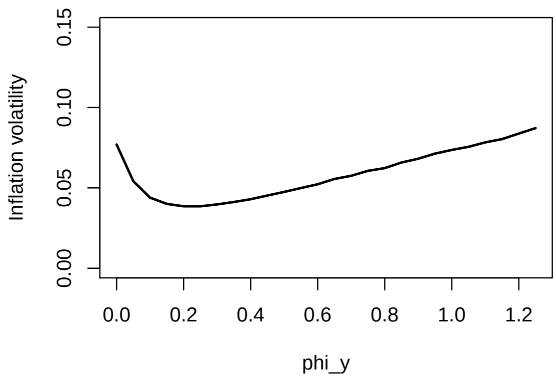
(c) $\pi_{start} = 3.5$, $y_{start} = 0$



(d) $\pi_{start} = 4.0$, $y_{start} = 0$



(e) $\pi_{start} = 3.5$, $y_{start} = 0.5$



(f) $\pi_{start} = 4.0$, $y_{start} = 0.5$

Figure 19: Inflation volatility in the behavioral model for different starting values

Notes: This figure shows the effect of parameter ϕ_y on inflation volatility for different starting values of y and π ($\phi_\pi = 1.5$ throughout).

C Appendix (for Online Publication): Discussion of the Measurement of Volatility

In general, different simple measures of price instability, i.e. of volatility, dispersion, or distance from the target are possible. We discuss mainly two of them here. The first one is the measure that we use, $v(\pi) = \frac{1}{T}\sum_{t=2}^{T}(\pi_t - \pi_{t-1})^2$, which is sometimes also referred to as relative deviation (equivalently, one could of course take $v_1(\pi) = \frac{1}{T-1}\sum_{t=2}^{T}(\pi_t - \pi_{t-1})^2$ or even $v_2(\pi) = \sum_{t=2}^{T}(\pi_t - \pi_{t-1})^2$ if the number of periods is fixed). The second one is the mean squared deviation from the target, $msd(\pi) = \frac{1}{T}\sum_{t=2}^{T}(\pi_t - \bar{\pi})^2$. Other alternatives that one could use are the so called absolute deviation, $ad(\pi) = \frac{1}{T}\sum_{t=2}^{T}|\pi_t - \pi_{t-1}|$ and the standard deviation, $sd(\pi) = \frac{1}{T}\sum_{t=2}^{T}(\pi_t - \pi^{av})^2$, where π^{av} is the average of inflation in a group taken over the whole time period. We do not discuss these measures here in detail; in general, the absolute deviation shares many features with $v(\cdot)$, the standard deviation shares many features with $msd(\cdot)$.

The measures $v(\cdot)$ and $msd(\cdot)$, are different in the following ways. The mean squared deviation from the target exclusively takes into account the distance to the target, not whether or not this distance is positive or negative. Figure 20 illustrates this with two example time series.

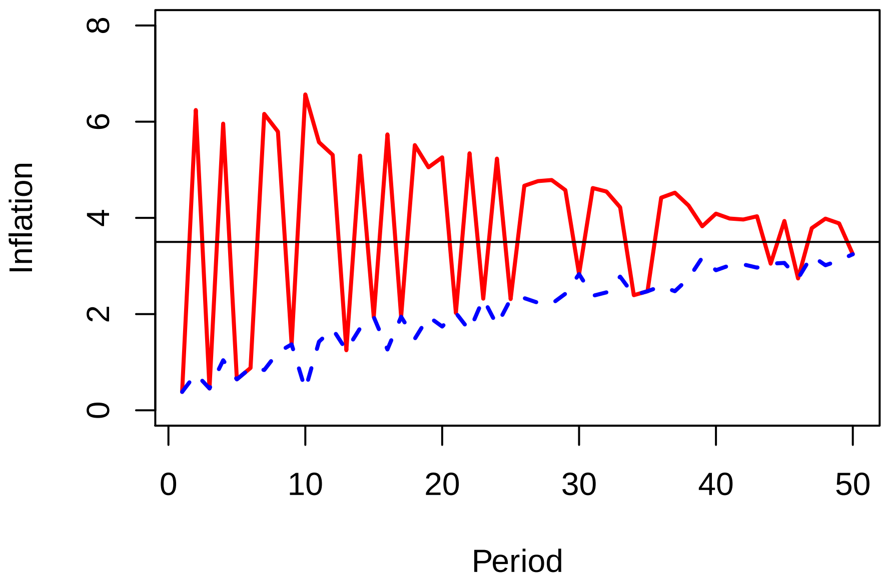


Figure 20: Example time series

The solid red line and the dashed blue line have exactly the same distance from the target in each period. However, it seems clear that the red line is much more volatile than the blue line, which converges slowly but nicely to the target. Any policy maker would prefer inflation as shown by the blue line over inflation as shown by the red line. However, $msd(\cdot)$ does not differentiate between these lines (while $v(\cdot)$ does).

There are other examples one can use to illustrate the differences between the measures. Imagine for example inflation staying constant for the first half of a time span at one percentage point below the target and then changing once and staying constant at one percentage point above the target. $msd(\cdot)$ does not distinguish between this very stable series and a series which randomly jumps back and forth between one percentage point below and one above the target (being at either value half of the time). $v(\cdot)$ distinguishes between these time series. $v(\cdot)$ is also not a perfect measure, however. For example if one were to compare inflation represented by two horizontal lines of which one is close to the target while the other is relatively far from the target, $v(\cdot)$ does not distinguish between these lines, while $msd(\cdot)$ does.

From a practitioner's or policy maker's point of view, which measure to use can thus depend on what kind of dynamics are present. For example if there are a lot of inflation time series which are relatively constant on one side of the target while some of these observations are close and some far from the target, $msd(\cdot)$ looks like a better measure. If one sees both erratic behavior or oscillations partly below and partly above the target and slow convergence, $v(\cdot)$ is the better measure. The latter case is exactly what we observe in the experiment. Inflation mainly oscillates around the target with mean values close to the target with some observations converging gradually to the target. From this point of view $v(\cdot)$ is clearly to be preferred.

D Appendix (for Online Publication): Instructions in the Experiment

Subjects in the experiment received the following instructions (as subjects only received qualitative information on the model governing the experimental economy the instructions are the same for both treatments):

Instructions

Welcome to this experiment! The experiment is anonymous, the data from your choices will only be linked to your station ID, not to your name. You will be paid privately at the end, after all participants have finished the experiment. After the main part of the experiment and before the payment you will be asked to fill out a short questionnaire. On your desk you will find a calculator and scratch paper, which you can use during the experiment.

During the experiment you are not allowed to use your mobile phone. You are also not allowed to communicate with other participants. If you have a question at any time, please raise your hand and someone will come to your desk.

General information and experimental economy

All participants will be randomly divided into groups of six people. The group composition will not change during the experiment. You and all other participants will take the roles of statistical research bureaus making predictions of inflation and the so-called "output gap". The experiment consists of 50 periods in total. In each period you will be asked to predict inflation and output gap for the next period. The economy you are participating in is described by three variables: inflation π_t, output gap y_t and interest rate i_t. The subscript t indicates the period the experiment is in. In total there are 50 periods, so t increases during the experiment from 1 to 50.

Inflation

Inflation measures the percentage change in the price level of the economy. In each period, inflation depends on inflation predictions of the statistical research bureaus in the economy (a group of six participants in this experiment), on actual output gap and on a random term. There is a positive relation between the actual inflation and both inflation predictions and actual output gap. This means for example that if the inflation predictions of the research bureaus increase, then actual inflation will also increase (everything else equal). In economies similar to this one, inflation has historically been between -5% and 10%.

Output gap

The output gap measures the percentage difference between the Gross Domestic Product (GDP) and the natural GDP. The GDP is the value of all goods produced during a period in the economy. The natural GDP is the value the total production would have if prices in the economy were fully flexible. If the output gap is positive (negative), the economy therefore produces more (less) than the natural GDP. In each period the output gap depends on inflation predictions and output gap predictions of the statistical bureaus, on the interest rate and on a random term. There is a positive relation between the output gap and inflation predictions and also between the output gap and output gap predictions. There is a negative relation between the output gap and the interest rate. In economies similar to this one, the output gap has historically been between -5% and 5%.

Interest Rate

The interest rate measures the price of borrowing money and is determined by the central bank. If the central bank wants to increase inflation or output gap it decreases the interest rate, if it wants to decrease inflation or output gap it increases the interest rate.

Prediction task

Your task in each period of the experiment is to predict inflation and output gap in the next period. When the experiment starts, you have to predict inflation and output gap for the first two periods, i.e. π_1^e and π_2^e, and y_1^e and y_2^e. The superscript e indicates that these are predictions. When all participants have made their predictions for the first two periods, the actual inflation (π_1), the actual output gap (y_1) and the interest rate (i_1) for period 1 are announced. Then period 2 of the experiment begins. In period 2 you make inflation and output gap predictions for period 3 (π_3^e and y_3^e). When all participants have made their predictions for period 3, inflation (π_2), output gap (y_2), and interest rate (i_2) for period 2 are announced. This process repeats itself for 50 periods.

Thus, in a certain period t when you make predictions of inflation and output gap in period $t+1$, the following information is available to you:

- Values of actual inflation, output gap and interest rate up to period $t-1$;
- Your predictions up to period t;

- Your prediction scores up to period $t-1$.

Payments

Your payment will depend on the accuracy of your predictions. You will be paid either for predicting inflation or for predicting the output gap. The accuracy of your predictions is measured by the absolute distance between your prediction and the actual values (this distance is the prediction error). For each period the prediction error is calculated as soon as the actual values are known; you subsequently get a prediction score that decreases as the prediction error increases. The table below gives the relation between the prediction error and the prediction score. The prediction error is calculated in the same way for inflation and output gap.

Prediction error	0	1	2	3	4	9
Score	100	50	33.33	25	20	10

Example: If (for a certain period) you predict an inflation of 2%, and the actual inflation turns out to be 3%, then you make an absolute error of $3\% - 2\% = 1\%$. Therefore you get a prediction score of 50. If you predict an inflation of 1%, and the actual inflation turns out to be negative 2% (i.e. -2%), you make a prediction error of $1\% - (-2\%) = 3\%$. Then you get a prediction score of 25. For a perfect prediction, with a prediction error of zero, you get a prediction score of 100. The figure below shows the relation between your prediction score (vertical axis) and your prediction error (horizontal axis). Points in the graph correspond to the prediction scores in the previous table.

[Figure 21 appears here in the experimental instructions.]

At the end of the experiment, you will have two total scores, one for inflation predictions and one for output gap predictions. These total scores simply consist of the sum of all prediction scores you got during the experiment, separately for inflation and output gap predictions. **When the experiment has ended, one of the two total scores will be randomly selected for payment.**

Your final payment will consist of 0.75 **euro for each** 100 **points in the selected total score (**200 **points therefore equals** 1.50 **euro). This will be the only payment from this experiment, i.e. you will not receive a show-up fee on top of it.**

Computer interface

The computer interface will be mainly self-explanatory. The top right part of the screen will show you all of the information available up to the period that you are in (in period t, i.e. when you are asked to make your prediction for period $t+1$, this will be actual inflation, output gap, and interest rate until period $t-1$, your predictions until period t, and the prediction scores arising from your predictions until period $t-1$ for both inflation (I) and output gap (O)). The top left part of the screen will show you the information on inflation and output gap in graphs. The axis of a graph shows values in percentage points (i.e. 3 corresponds to 3%). **Note that the values on the vertical axes may change during the experiment and that they are different between the two graphs – the values will be such that it is comfortable for you to read the graphs.**

In the bottom left part of the screen you will be asked to enter your predictions. When submitting your prediction, use a decimal point if necessary (not a comma). For example, if you want to submit a prediction of 2.5% **type "2.5"; for a prediction of** -1.75% **type "-1.75".** The sum of the prediction scores over the different periods are shown in the bottom right of the screen, separately for your inflation and output gap predictions.

At the bottom of the screen there is a status bar telling you when you can enter your predictions and when you have to wait for other participants.

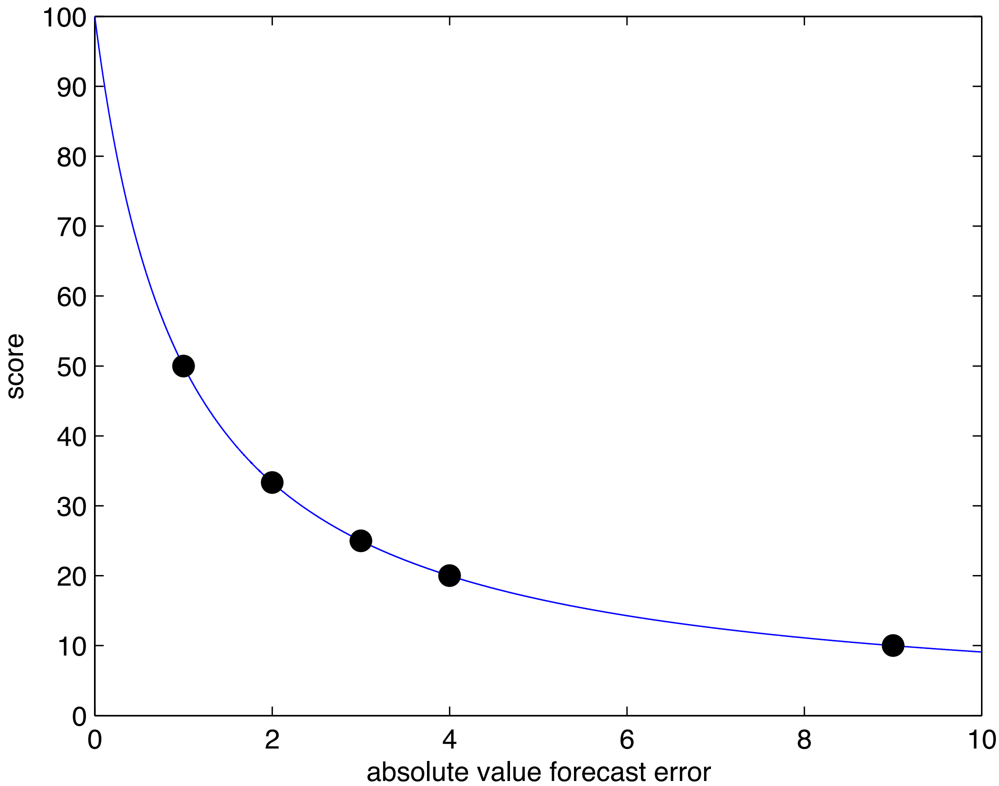


Figure 21: Relation score and forecast error (not labeled in the instructions)

E Appendix (for Online Publication): Graphs of the Experimental Data by Group and Screenshot

E.1 Realizations and Forecasts of Inflation and Output Gap in All Groups

Figures 22 to 30 show the realizations and forecasts of inflation and output gap. Each graph corresponds to one group of six people (one experimental economy). The thick black line shows the realization of inflation, the thin dashed black lines show the inflation forecasts of the six individuals in the group. The thick gray line shows the realization of the output gap and the thin dashed gray lines show the output gap forecasts of all individuals in a group. On the horizontal axis are the periods (from 1 to 50), on the vertical axis are the values of inflation and output gap in percent (the numbers on the vertical axis reach from -3 to 8). The upper red line corresponds to the steady state value of inflation ($\bar{\pi} = 3.5$), the lower red line corresponds to the steady state value of the output gap ($\bar{y} = 0.1166667$). Figures 22 to 25 show all groups of treatment $T1$, Figures 26 to 29 show the groups of treatment $T2$. Figure 30 shows the two groups (from $T2$) that have been excluded from the analysis as explained in Footnote 11.

E.2 Screenshot

Figure 31 shows a screenshot (a larger version of the screenshot already used in Figure 4.

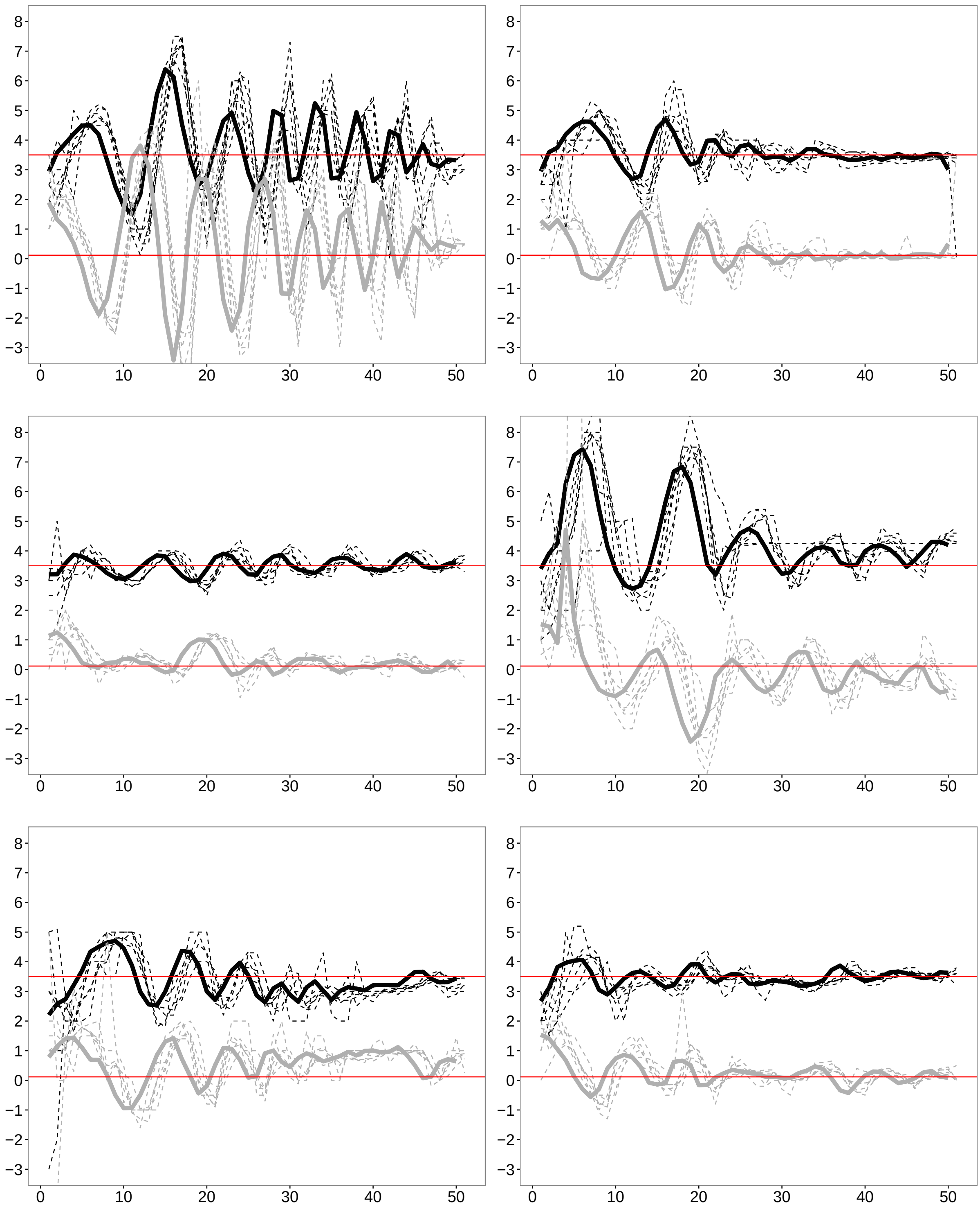


Figure 22: Realizations and forecasts of inflation and output gap ($T1$, groups $1-6$)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.

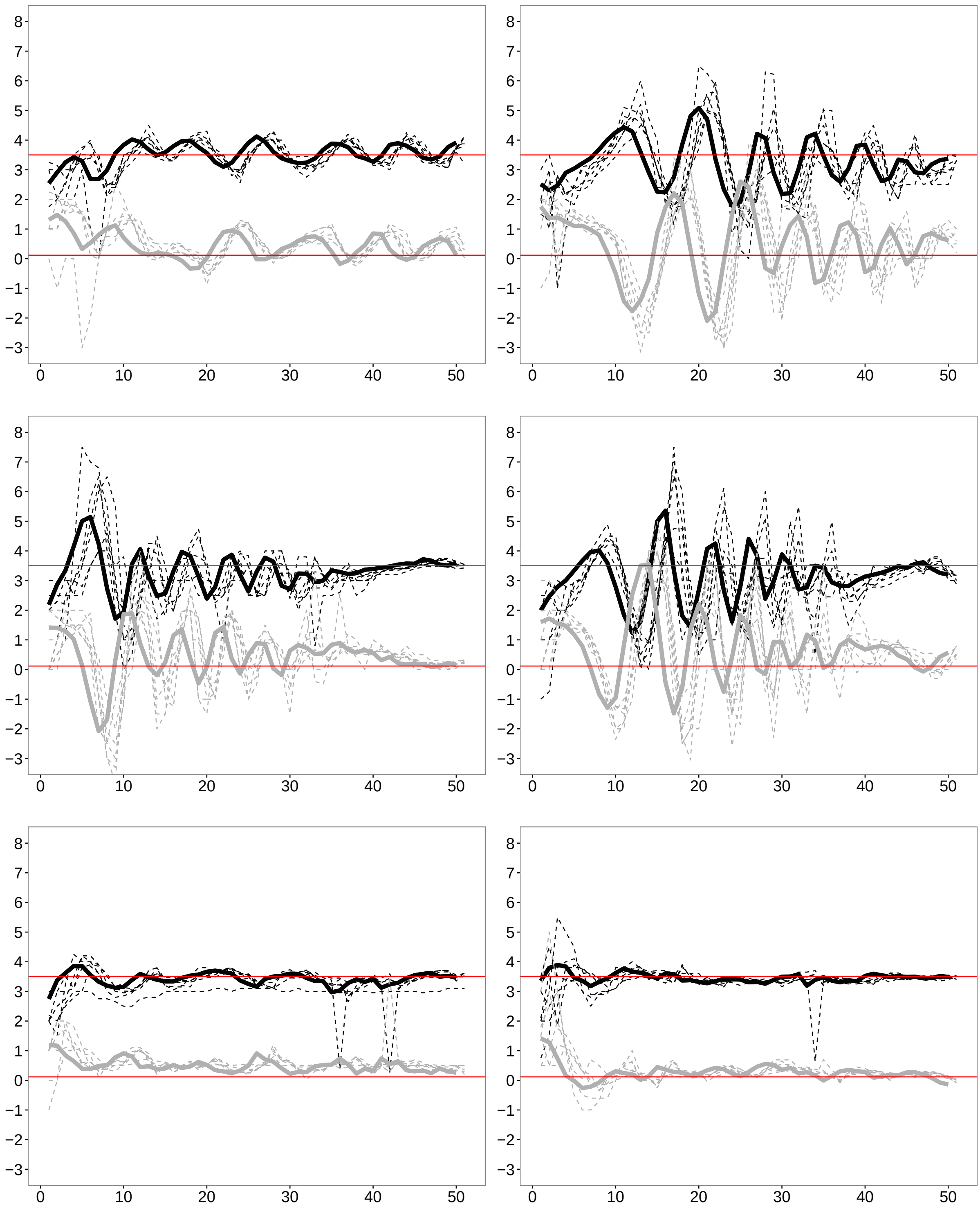


Figure 23: Realizations and forecasts of inflation and output gap ($T1$, groups $7-12$)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.

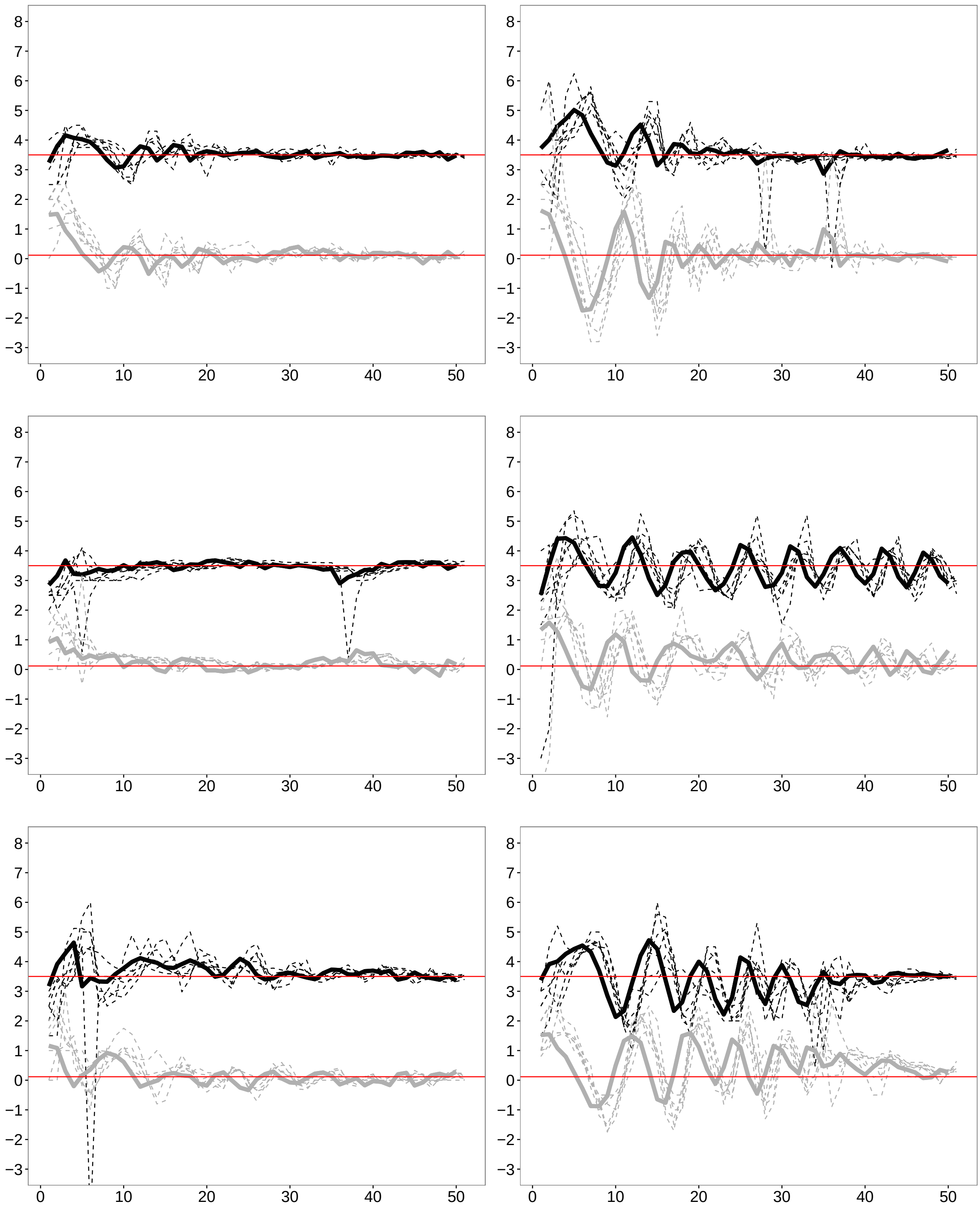


Figure 24: Realizations and forecasts of inflation and output gap ($T1$, groups $13-18$)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.

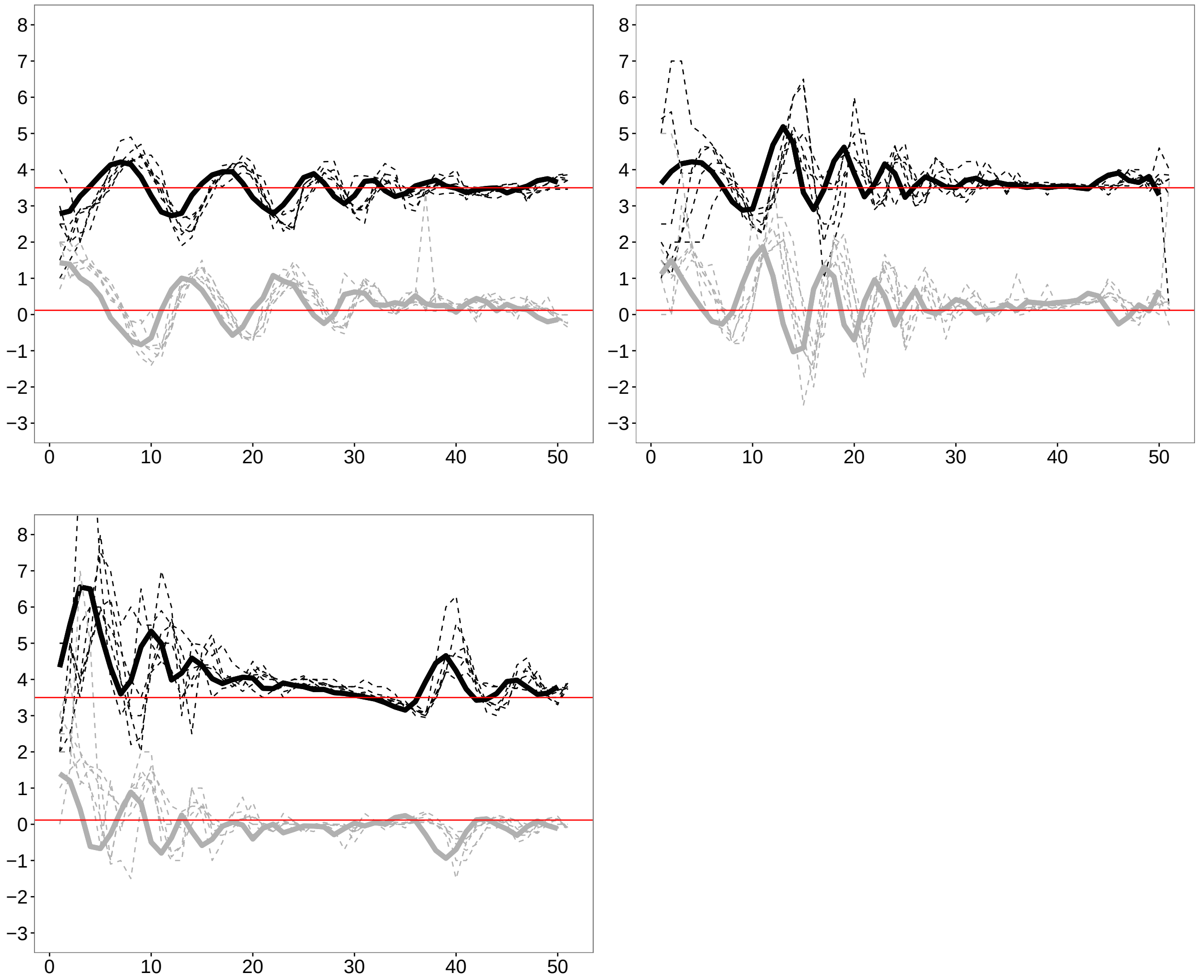


Figure 25: Realizations and forecasts of inflation and output gap ($T1$, groups $19-21$)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.

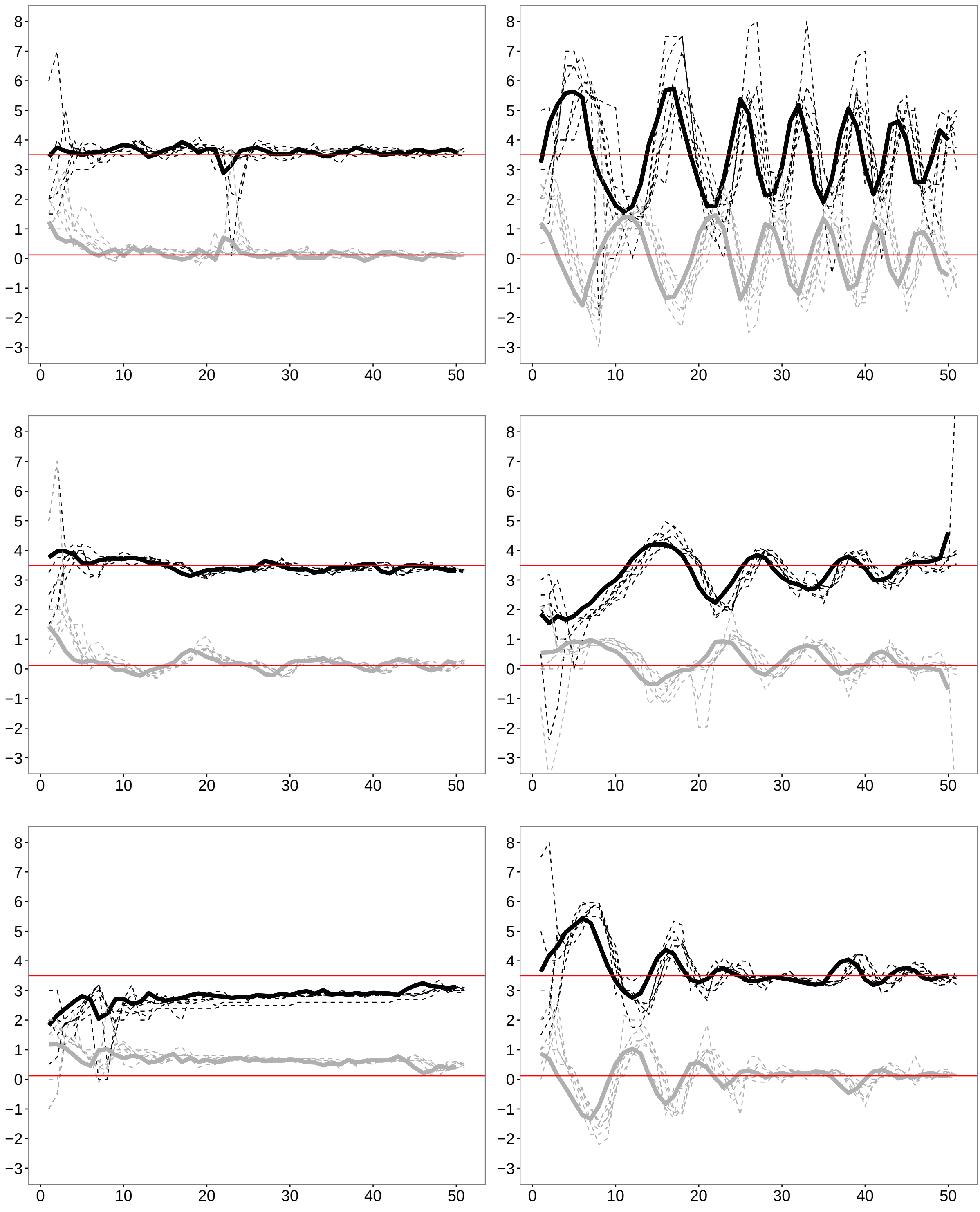


Figure 26: Realizations and forecasts of inflation and output gap ($T2$, groups $1-6$)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.

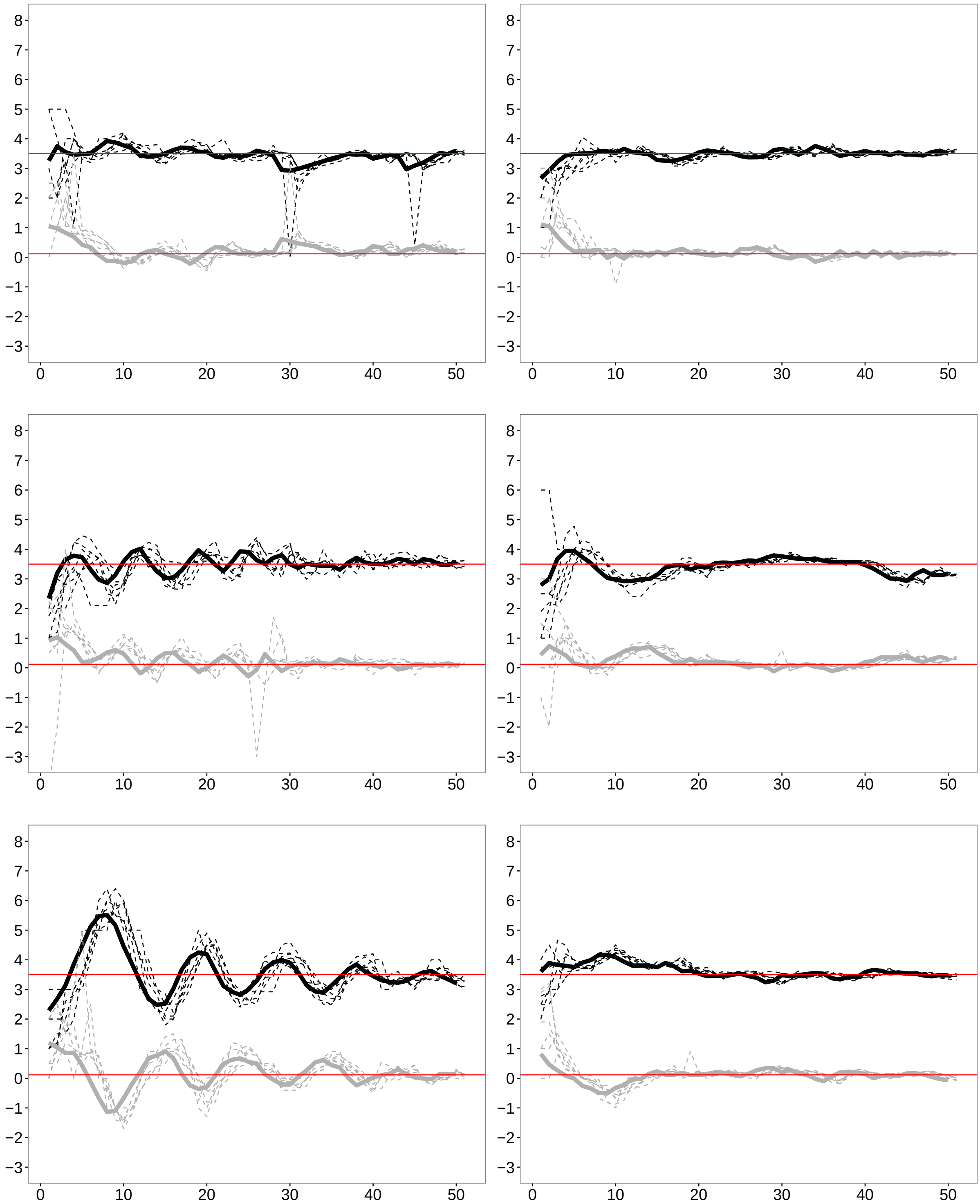


Figure 27: Realizations and forecasts of inflation and output gap ($T2$, groups $7-12$)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.

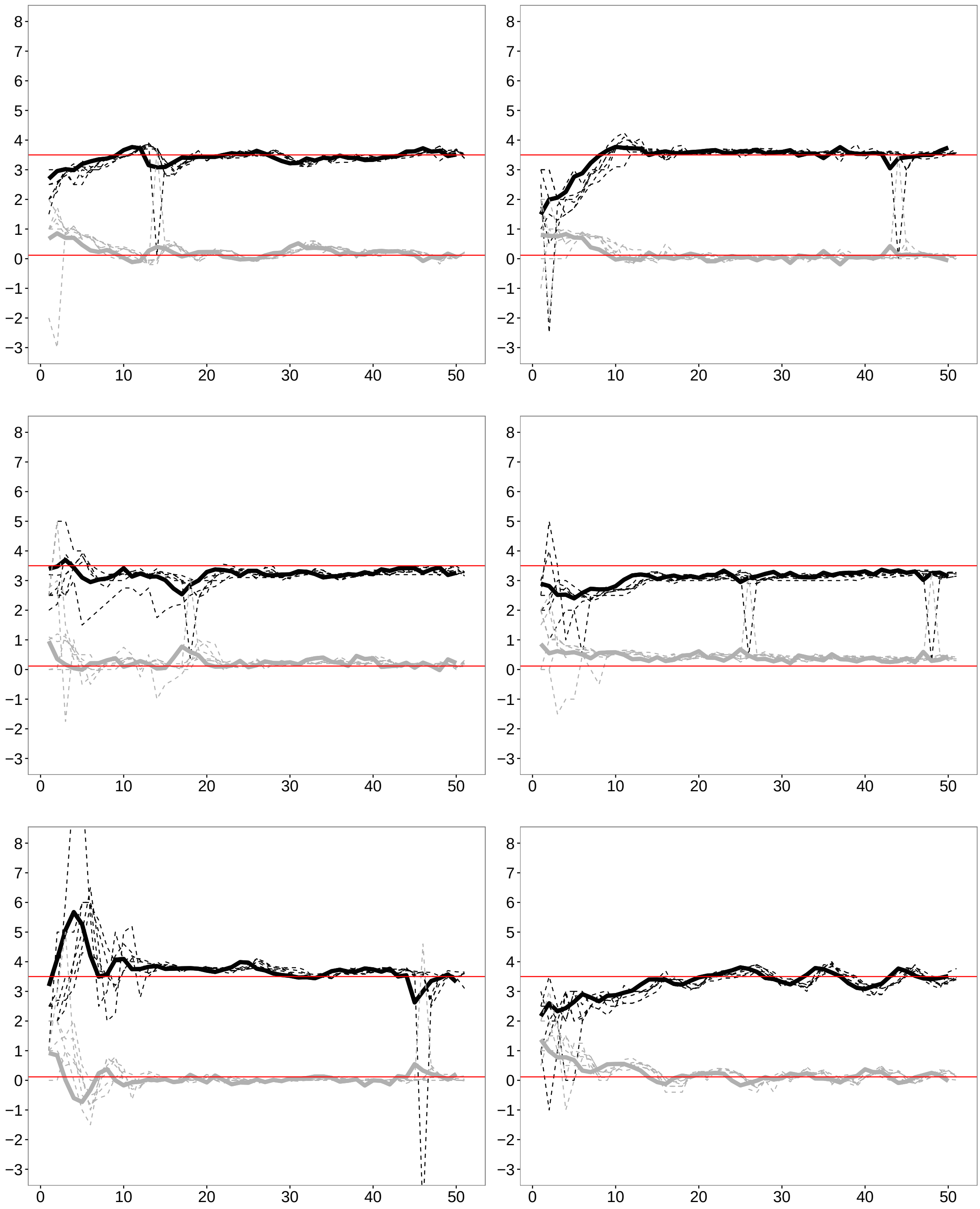


Figure 28: Realizations and forecasts of inflation and output gap ($T2$, groups $13-18$)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.

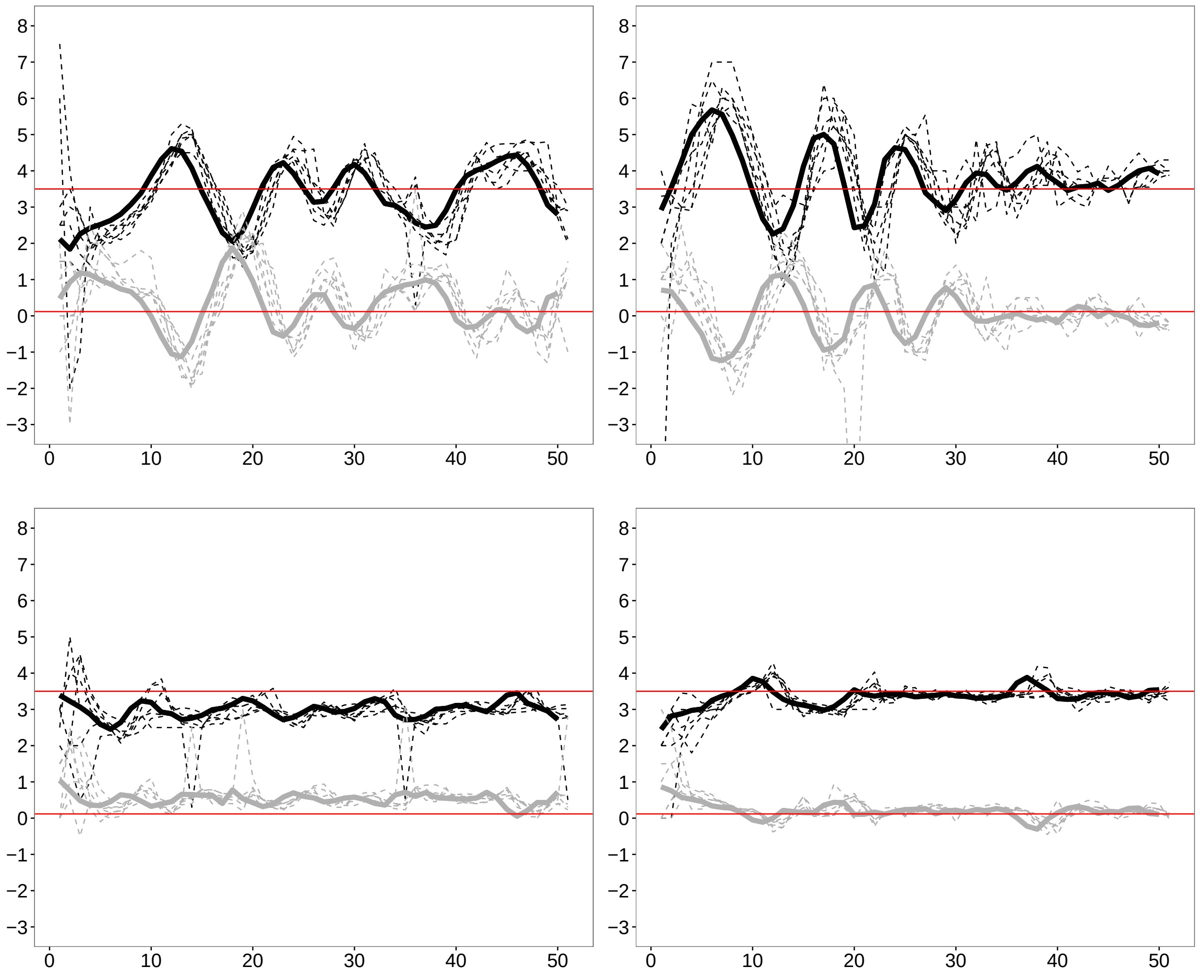


Figure 29: Realizations and forecasts of inflation and output gap ($T2$, groups $19-22$)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.

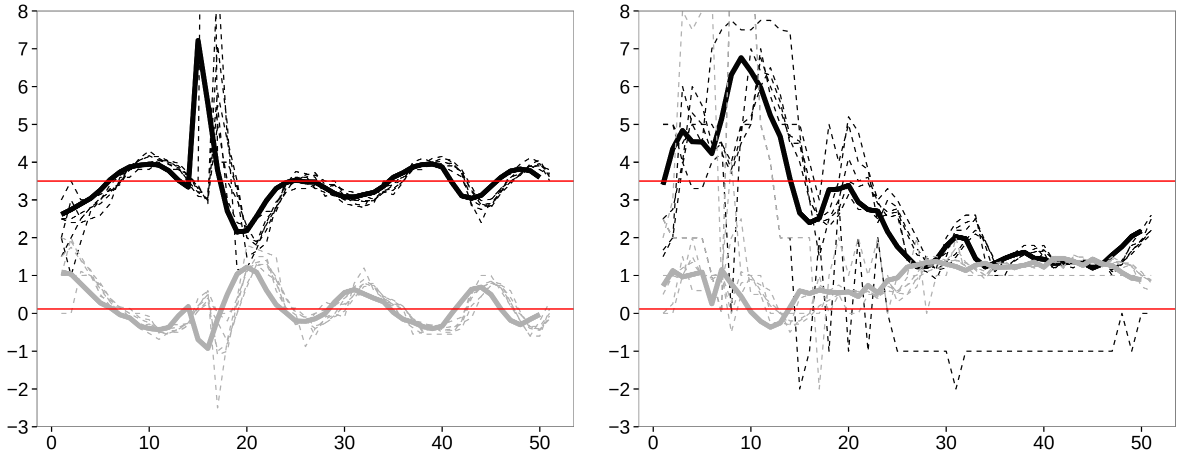
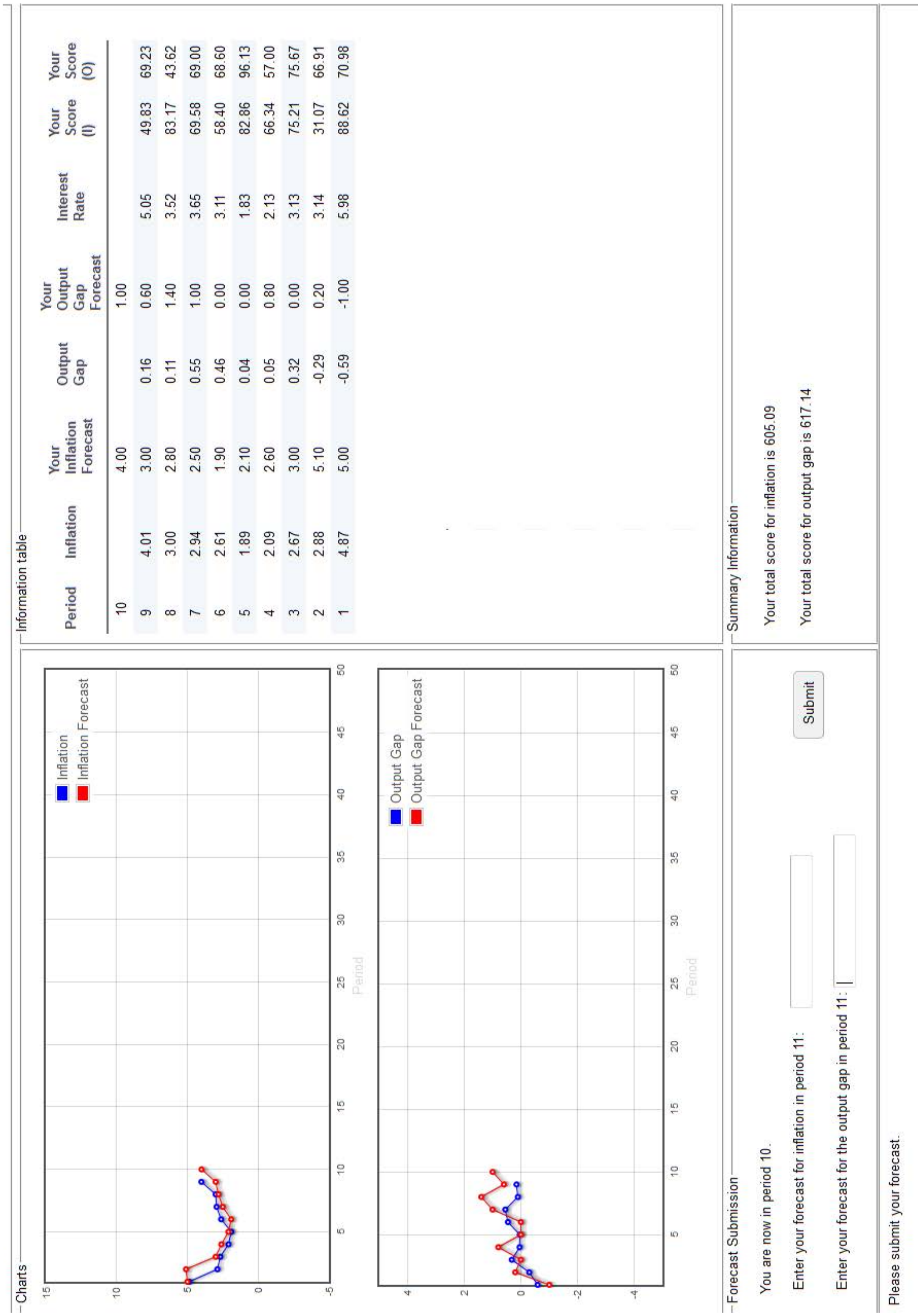


Figure 30: Realizations and forecasts of inflation and output gap (excluded groups)

Notes: Each of the graphs corresponds to one group and shows realized inflation (thick black line), individual inflation forecasts (dashed black lines), realized output gap (thick gray line), and individual output gap forecasts (dashed gray lines) over the 50 periods of the experiment.



Period	Inflation	Your Inflation Forecast	Output Gap	Your Output Gap Forecast	Interest Rate	Your Score (I)	Your Score (O)
10		4.00		1.00			
9	4.01	3.00	0.16	0.60	5.05	49.83	69.23
8	3.00	2.80	0.11	1.40	3.52	83.17	43.62
7	2.94	2.50	0.55	1.00	3.65	69.58	69.00
6	2.61	1.90	0.46	0.00	3.11	58.40	68.60
5	1.89	2.10	0.04	0.00	1.83	82.86	96.13
4	2.09	2.60	0.05	0.80	2.13	66.34	57.00
3	2.67	3.00	0.32	0.00	3.13	75.21	75.67
2	2.88	5.10	-0.29	0.20	3.14	31.07	66.91
1	4.87	5.00	-0.59	-1.00	5.98	88.62	70.98

Figure 31: Screenshot

F Appendix (for Online Publication): Additional Graphs and Data Analysis

F.1 Inflation in the Experiment

Figure 32 shows the empirical cumulative distribution functions of price instability when employing different measures. The first graph shows the absolute deviation, $ad(\pi) = \frac{1}{T}\sum_{t=2}^{T}|\pi_t - \pi_{t-1}|$. The second graph shows the means squared deviation from the target, $msd(\pi) = \frac{1}{T}\sum_{t=2}^{T}(\pi_t - \bar{\pi})^2$, and the third graph shows the standard deviation, $sd(\pi) = \frac{1}{T}\sum_{t=2}^{T}(\pi_t - \pi^{av})^2$. The average absolute deviations are 0.304 in $T1$ and 0.188 in $T2$. For the mean squared deviations, the values are 0.402 in $T1$ and 0.317 in $T2$, and for the standard deviation 0.510 in $T1$ and 0.419 in $T2$.

Figure 33 shows for the other measures what Figure 7 shows for the standard measure of inflation volatility, i.e. it tracks their components over time. The first row of graphs depicts the components of the absolute deviation over time, $|\pi_t - \pi_{t-1}|$ ($T1$ on the left, $T2$ on the right). The middle row shows the squared deviation from the target $(\pi_t - \bar{\pi})^2$ and the bottom row shows the squared deviation from the average across all time periods, $(\pi_t - \pi^{av})^2$.

F.2 Prediction Performance of the Behavioral Model

The mean squared error can potentially be criticized as a measure of prediction accuracy, as it puts a lot of weight on the largest observations (possibly outliers), while it hardly puts any weight on all of the observations which are close to the realizations. Different measures can be used if one wants to avoid that. We show here the results for measuring the prediction error with $F(x_h^e - x) = 100 - 100/(1 + |x_h^e - x|)$. This function has a kink at zero, so that small prediction errors are not ignored. The function has a supremum of 100, thus no matter how far a single prediction is from the realization the prediction error will be less than 100 (this is the same function used in the experiment to incentivize subjects, but here it just used as a prediction error). The results are similar to those reported in the paper when considering the mean squared error (Table 3).

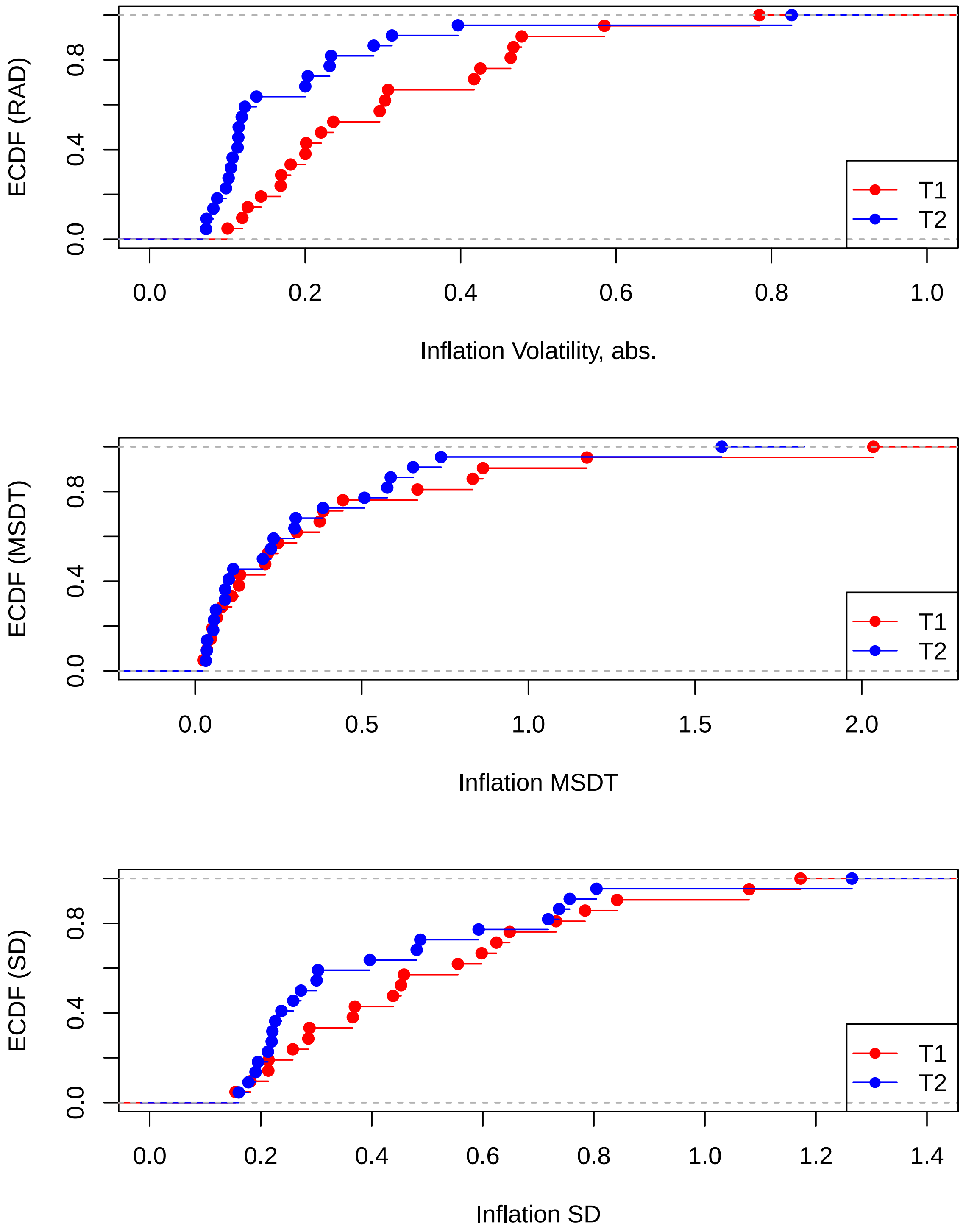


Figure 32: Empirical distribution functions of various measures

Notes: This graph shows the ECDFs for three different measures. From top to bottom: Absolute deviation, mean squared deviation from target, standard deviation. For each value on the horizontal axis, the fraction of observations with the respective measure less or equal to this value (i.e. the ECDF) is shown on the vertical axis, separately for $T1$ and $T2$.

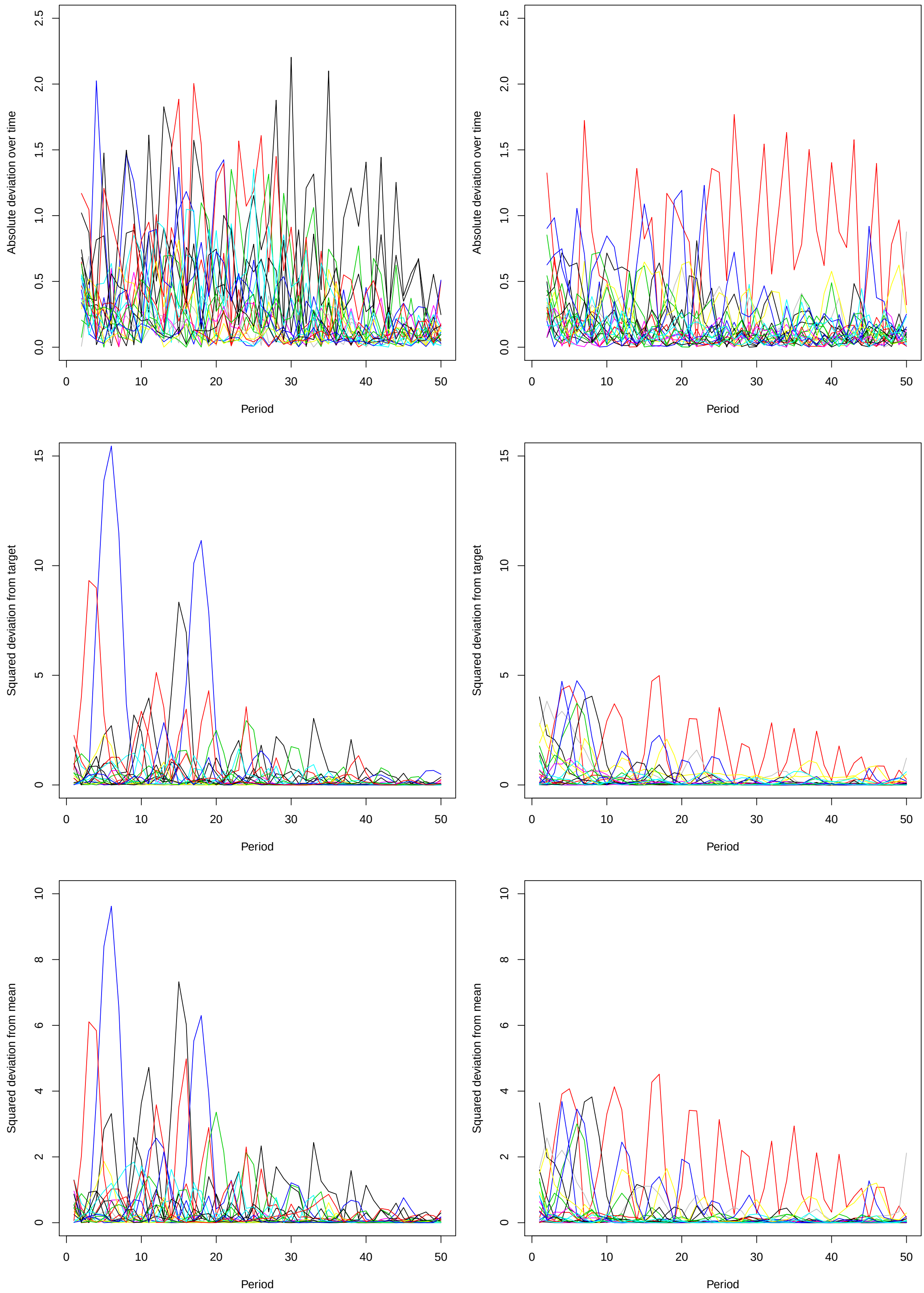


Figure 33: Inflation volatility/dispersion over time for different measures

Notes: This figure shows for each group the component of different measures of inflation volatility/dispersion. $T1$ is on the left, $T2$ on the right. From top to bottom: $|\pi_t - \pi_{t-1}|$, $(\pi_t - \bar{\pi})^2$, and $(\pi_t - \pi^{av})^2$. The colors of the groups are as in Figure 5.

Table 3: Prediction errors according to the additional measure

	Inflation T1	Output gap T1	Inflation T2	Output gap T2
HSM	584.8	598.0	409.4	337.3
RE	1192.2	1343.0	1145.9	949.6
ADA	936.6	994.0	674.6	564.3
WTR	714.3	737.2	493.8	424.5
STR	918.4	1025.4	609.6	624.4
LAA	739.8	757.6	592.9	502.7