Managing Self-organization of Expectations through Monetary Policy: a Macro Experiment

Tiziana Assenza[a,b] Peter Heemeijer[c]
Cars Hommes[b,d,*] Domenico Massaro[b]

October 25, 2014

[a] *Department of Economics and Finance, Università Cattolica del Sacro Cuore, Milano,*
[b] *CeNDEF, University of Amsterdam*
[c] *ABN Amro,* [d] *Tinbergen Institute, Amsterdam* [*] *corresponding author: C.H.Hommes@uva.nl*

Abstract. We use laboratory experiments to study individual expectations and aggregate macro behavior in a New Keynesian framework. Four different aggregate outcomes arise: convergence to equilibrium, explosive behavior along inflationary or deflationary spirals, persistent or dampened oscillations. A heuristics switching model, driven by relative performance, explains these patterns as emerging properties of the path-dependent self-organization process of heterogeneous expectations leading to coordination on an almost self-fulfilling rule. A more aggressive Taylor rule can manage the self-organization process adding negative feedback to the overall positive feedback system, making coordination on destabilizing trend-following expectations less likely and coordination on stabilizing adaptive expectations more likely.

JEL codes: C91, C92, D84, E52.

Keywords: Laboratory Experiments, Monetary Policy, Expectations, Heterogeneity.

Acknowledgment: An earlier version of this paper under the title "Individual expectations and aggregate macro behavior" has been presented at the Computing in Economics and Finance (CEF) Conference, Sydney, July 15-17, 2009, the LeeX International Conference on Theoretical and Experimental Macro, Barcelona, June 11-12, 2010, the Conference on Experiments in Macroeconomics and Financial Economics, Columbia, March 1, 2013, the Conference on Expectations and Macroeconomics, June 10-11, 2013, Bank of Spain, Madrid, the Conference on Expectations in Dynamic Macro Models, at the San Francisco FRB, August 8-10, 2013, the ESEM 2013 meeting, August 26-30, 2013, Gothenburg, the Workshop on Agent-based Models, November 7-8, 2013, Bordeaux, the Computing in Economics and Finance (CEF) Conference, Oslo, June 22-24, 2014, the 2^{nd} Macro, Banking and Finance Workshop, University of Rome "Tor Vergata" and in various seminars at the New York FED, the Dutch Central Bank, Bocconi, Bicocca and the University of Exeter in 2013 and 2014. We would like to thank Klaus Adam, Jess Benhabib, Jim Bullard, John Duffy, George Evans, Seppo Honkapohja, Albert Marcet, Ricardo Reis, Argia Sbordone, Shyam Sunder, John Williams and Mike Woodford and other seminar and conference participants for stimulating discussions and helpful comments. Financial support from the EU 7^{th} framework collaborative project "*POLHIA*", grant no. 225408, from the Netherlands Organization for Scientific Research (NWO), project "*Understanding Financial Instability through Complex Systems*" and from the EU 7^{th} framework collaborative projects "*Complexity Research Initiative for Systemic InstabilitieS (CRISIS)*", grant no. 288501 and "*RAStaNEWS*", grant no 320278, is gratefully acknowledged. None of the above are responsible for errors in this paper.

1 Introduction

Inflation expectations are crucial for the transmission of monetary policy. The way in which individual expectations are formed, therefore, is key in understanding how a change in the interest rate affects output and the actual inflation rate. Since the seminal papers of Muth (1961) and Lucas (1972) the rational expectations (RE) hypothesis has become the cornerstone of macroeconomic theory, with representative rational agent models dominating mainstream economics. For monetary policy analysis the most popular model is the New Keynesian (NK) framework which assumes, in its basic formulation, a representative rational agent structure (see e.g. Woodford (2003) and Gali (2008)). The standard NK model with a rational representative agent however has lost much of its appeal in the light of empirical evidence: it is clear from the data that this approach is not the most suitable to reproduce stylized facts such as the persistence of fluctuations in real activity and inflation after a shock (see e.g. Chari, Kehoe, and McGrattan (2000) and Nelson (1998)). Moreover, large movements in aggregate macro variables during the recent economic crisis have revealed serious limitations of a representative rational agent NK framework. Economists have therefore proposed a number of extensions to the standard framework by embedding potential sources of endogenous persistence. They have incorporated features such as habit formation or various adjustment costs to account for the inertia in the data (e.g. Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007)). More recently, models with financial frictions have been introduced within the NK framework, e.g. in Curdia and Woodford (2009), Curdia and Woodford (2010), Gertler and Karadi (2011), Gertler and Karadi (2013), Christiano, Motto, and Rostagno (2010) and Gilchrist, Ortiz, and Zakrajsek (2009).

A complementary route to introduce frictions within a standard NK framework focuses on modeling bounded rationality in expectations formation. In the last two decades adaptive learning has become an important alternative to modeling expectations (see e.g. Evans and Honkapohja (1998), Sargent (1999) and Evans and Honkapohja (2001)). Bullard and Mitra (2002), Evans and Honkapohja (2003), Preston (2005), Hommes and Zhu (2014) among others, introduce adaptive learning in the NK framework. Milani (2007) shows that learning can represent an important source of persistence in the economy and that some extensions which are typically needed under rational expectations to match the observed inertia become redundant under learning. More recently a number of authors have extended the NK model to include *heterogeneous expectations*, e.g. Gali and Gertler (1999), Branch and Evans (2006), De Grauwe (2010), Branch and McGough (2009, 2010), Kurz (2011), Massaro (2013), Anufriev, Assenza, Hommes, and Massaro (2013) and Cornea, Hommes, and Massaro (2012).

The empirical literature on expectations in a macro-monetary policy setting can be subdivided in work on survey data and laboratory experiments with human subjects. Mankiw, Reis, and Wolfers (2003) find evidence for heterogeneity in inflation expectations in the Michigan Survey of Consumers and argue that the data are inconsistent with rational or adaptive expectations, but may be consistent with a sticky information model. Branch (2004) estimates a simple switching model with heterogeneous expectations on survey data and provides empirical evidence for dynamic switching that depends on the relative mean squared errors of the predictors. Capistran and Timmermann (2009) show that heterogeneity of inflation expectations of professional forecasters varies over time and depends on the level and the variance of current inflation. Pfajfar and Santoro (2010) measure the degree of heterogeneity in private agents' inflation forecasts by exploring time series of percentiles from the empirical distribution of survey data. They show that heterogeneity in inflation expectations is persistent and identify three different expectations formation mechanisms: static or highly autoregressive rules, nearly rational expectations and adaptive learning with sticky information.

Macro experiments with human subjects in a controlled laboratory environment to study individual expectations and aggregate behavior have been carried out by, e.g., Marimon, Spear, and Sunder (1993), Marimon and Sunder (1994), Hommes, Sonnemans, Tuinstra, and van de Velden (2005), Adam (2007); see Duffy (2008) for an overview of macro experiments. The beauty contest or guessing game of Nagel (1995) and follow-up papers represent another interesting and early example of a macro experiment, where subjects must forecast the average forecasts of a group of individuals. Fehr and Tyran (2008) show that, in a price setting game with perfect information about best response functions and average forecasts of all individuals, the aggregate behavior depends upon the strategic environment and find fast convergence to equilibrium in the case of strategic substitutes, but slow convergence under strategic complements. An important insight from the macro experiments in Heemeijer, Hommes, Sonnemans, and Tuinstra (2009) and Bao, Hommes, Sonnemans, and Tuinstra (2012) is that macro systems with negative expectations feedback are stable and settle down to equilibrium quickly, even with very limited market information, in contrast to macro systems with positive feedback which do not converge but rather oscillate around equilibrium; see Hommes (2011) for an overview of learning to forecast macro experiments.[1] A classical example of negative expectations feedback is the cobweb framework of an agricultural market, where higher expectations lead firms to produce more and the realized market

[1]See also Schotter and Trevino (2007) for an overview of lab experiments on elicitation of beliefs in strategic games.

price to be lower. Positive feedback arises e.g. in asset markets, where higher price expectations induce higher speculative demand and therefore higher realized asset prices. Positive feedback systems may easily destabilize through coordination on almost-self-fulfilling equilibria driven by trend-extrapolating expectations (Hommes (2013)). Finally, evidence from laboratory experiments supports heterogeneity of individual expectations and switching between different forecasting rules (e.g., Anufriev and Hommes (2012)).

In this paper we use laboratory experiments with human subjects to study the individual expectations formation process within a standard NK setup. We run three different treatments, with different coefficients and inflation targets of the Taylor interest rate rules, to study the stabilizing role of monetary policy rules in managing heterogeneous expectations. We ask subjects to forecast the inflation rate and the output gap and we address the following questions:

- Which forecasting rules do individuals use in a NK macro system?
- Are expectations heterogeneous or does the system self-organize into coordination of individual forecasts on a common rule? If so, on which rule do subjects coordinate?
- Which theory of (heterogeneous) expectations and learning fits individual as well as aggregate experimental data?
- How can monetary policy affect the self-organization process of individual expectations and stabilize aggregate macro behavior?

Our paper makes three contributions. First, we run a learning to forecast macro experiment within the NK framework and we test the validity of standard monetary policy recommendations (i.e, the Taylor principle) in the experimental NK economies. The experimental part of our paper is similar in spirit to the learning to forecast experiments of Pfajfar and Zakelj (2012), but differs in two important dimensions from their experimental design. While in Pfajfar and Zakelj (2012) participants are forecasting inflation *only*, we allow agents to forecast both inflation and output gap, in accordance to the theoretical NK model. To our best knowledge, this is the first experimental economy in which fluctuations of the aggregate variables depend endogenously on the individual forecasts of *two* different variables, inflation and output gap.[2] A second crucial difference with the experi-

[2]Pfajfar and Zakelj (2012) experiments were run in May 2006 and June 2009, while most of our sessions were run in March and May 2009, with additional sessions in February-March 2013. In our experiments, to keep the task as simple as possible, forecasting of inflation and output was done by two different groups. Recently, Kryvtsov and Petersen (2013) used the design of Pfajfar and Zakelj (2012) to run similar forecasting experiments with a group of subjects forecasting both inflation and the output gap.

mental design of Pfajfar and Zakelj (2012) concerns the stochastic process of the shocks. In Pfajfar and Zakelj (2012) the shocks follow an AR(1) process, implying an autocorrelated RE solution. In such an experimental environment it is not clear whether fluctuations in inflation and output are *expectations driven* or solely *driven by exogenous shocks.* In contrast, we use small IID shocks to our experimental economy, so that the RE fundamental solution is an IID process (with very small fluctuations). Therefore any observed fluctuations in aggregate variables in our experimental economy must be *endogenously driven by individual expectations.* Our macro experiments thus addresses the important issue whether monetary policy can stabilize endogenous expectations-driven fluctuations in the economy.

The second contribution of our paper is to fit a behavioral theory of *heterogeneous expectations* to the individual as well as aggregate experimental data in a NK framework. We use the heuristic switching model of Brock and Hommes (1997), extended by Anufriev and Hommes (2012), to explain the coordination of individual expectations on a common rule and the self-organization of the NK macro system into four different types of emerging aggregate behaviors: (i) convergence to a stable RE or a non-fundamental steady state, (ii) exploding inflation and output, with either inflationary or deflationary spirals, (iii) dampened oscillations, or (iv) persistent oscillations in inflation and output.

The third contribution of our paper is to study how monetary policy can affect the self-organization process and stabilize the experimental macro economy. More precisely, we compare different experimental treatments with a weak and an aggressive interest rate Taylor rule –i.e. with different coefficients, $\phi_\pi = 1$ (where the interest rate responds one-to-one to changes in inflation) and $\phi_\pi = 1.5$ (with a stronger response to inflation)– and different inflation targets $\bar{\pi} = 2\%$ and $\bar{\pi} = 3.5\%$. It turns out that with a weak Taylor rule the economy may converge to arbitrary equilibrium levels or may lead to inflationary or deflationary spirals, with exploding inflation and output, due to coordination of individual expectations on a trend-following rule. In our experimental macro economy, a more aggressive interest rate Taylor rule can stabilize exploding inflation and output and enforce (slow) convergence of the economy to the RE outcome. Our heuristics switching model provides an intuitive explanation of how a more aggressive monetary policy can manage the self-organization process of heterogeneous expectations and stabilize inflation and output. With a weak Taylor rule, the economy exhibits strong *positive feedback*, i.e. optimistic inflation and output expectations are (almost) self-fulfilling and may easily lead to coordination of individual expectations on trend-following extrapolative expectations destabilizing the economy into explosive inflationary or deflationary spirals. A more aggressive Taylor rule adds *negative feedback* to the macroeconomic system when the nominal interest rate offsets expected inflation. Due to this extra negative feedback, the overall positive feedback in the macro system becomes weaker and trend-following forecasting strategies do not easily survive evolutionary competition in such an economy. In the presence of a more aggressive Taylor rule, the self-organization of individual expectations through evolutionary selection causes adaptive expectations to dominate trend-following strategies and the economy stabilizes. Finally, we use the insights from our behavioral heterogeneous expectations model to study an augmented Taylor rule that sets the interest rate rule not only in response to inflation but also in response to the output gap. Our behavioral model predicts that, due to the additional negative feedback in the system when the monetary policy also reacts to the output gap, the augmented Taylor rule is more likely to effectively manage the self-organization of heterogeneous expectations and stabilize the macro economy.

The paper is organized as follows. Section 2 describes the underlying NK-model framework, the different treatments, the experimental design and the experimental results. Section 3 proposes a behavioral heterogeneous expectations model explaining individual expectations and aggregate outcomes. The model explains the emergence of different types of aggregate outcomes through self-organization and coordination of individual expectations. Section 4 studies how monetary policy can manage the self-organization of heterogeneous expectations and affect the stability of the experimental macro system. Finally, Section 5 concludes.

2 The learning to forecast macro experiment

In subsection 2.1 we briefly recall the NK model and then we give a description of the treatments in the experiment. Subsection 2.2 gives an overview of the experimental design, while subsection 2.3 summarizes the main results.

2.1 The New Keynesian model

In this section we recall the monetary model with nominal rigidities that will be used in the experiment. We adopt the heterogeneous expectations version of the

NK model,[3] which is described by the following equations:

$$y_t = \bar{y}^e_{t+1} - \varphi(i_t - \bar{\pi}^e_{t+1}) + g_t\,, \tag{2.1}$$

$$\pi_t = \lambda y_t + \rho\bar{\pi}^e_{t+1} + u_t\,, \tag{2.2}$$

$$i_t = \mathrm{Max}\{\bar{\pi} + \phi_\pi(\pi_t - \bar{\pi}), 0\}\,, \tag{2.3}$$

where y_t and $\bar{y}^e_{t+1}$ are respectively the actual and average expected output gap, i_t is the nominal interest rate, π_t and $\bar{\pi}^e_{t+1}$ are respectively the actual and average expected inflation rates, $\bar{\pi}$ is the inflation target, φ, λ, ρ and ϕ_π are positive coefficients and g_t and u_t are white noise shocks. The coefficient ϕ_π measures the response of the nominal interest rate i_t to deviations of the inflation rate π_t from its target $\bar{\pi}$. Equation (2.1) is the aggregate demand in which the output gap y_t depends on the average expected output gap $\bar{y}^e_{t+1}$ and on the real interest rate $i_t - \bar{\pi}^e_{t+1}$. Equation (2.2) is the New Keynesian Phillips curve according to which the inflation rate depends on the output gap and on average expected inflation. Equation (2.3) is the monetary policy rule implemented by the monetary authority in order to keep inflation at its target value $\bar{\pi}$. Note that the interest rate rule has a zero lower bound (ZLB). The NK model is widely used in monetary policy analysis and allows us to compare our experimental results with those obtained in the theoretical literature. However the NK framework requires agents to forecast *both* inflation and the output gap. Since forecasting two variables at the same time is a complicated task for the participants in an experiment, we decided to elicit forecasts of the endogenous variables from different groups of subjects. In fact, in each experimental economy, there are two groups of participants with different tasks: one group forecasts inflation while the other forecasts the output gap.[4] The aggregate variables inflation and output gap are thus driven by individual expectations feedbacks from two different variables by two different groups. The model describing the experimental economy can be written as

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \Omega \begin{bmatrix} \varphi\bar{\pi}(\phi_\pi - 1) \\ \lambda\varphi\bar{\pi}(\phi_\pi - 1) \end{bmatrix} + \Omega \begin{bmatrix} 1 & \varphi(1 - \phi_\pi\rho) \\ \lambda & \lambda\varphi + \rho \end{bmatrix} \begin{bmatrix} \bar{y}^e_{t+1} \\ \bar{\pi}^e_{t+1} \end{bmatrix} + \Omega \begin{bmatrix} 1 & -\varphi\phi_\pi \\ \lambda & 1 \end{bmatrix} \begin{bmatrix} g_t \\ u_t \end{bmatrix}\,, \tag{2.4}$$

[3]Micro-founded NK models consistent with heterogeneous expectations have been derived by Branch and McGough (2009), Kurz (2011), Kurz, Piccillo, and Wu (2013) and Massaro (2013). System (2.1) - (2.3) corresponds to the model developed by Branch and McGough (2009) augmented with demand and supply shocks, or to the model derived in Kurz (2011) and Kurz, Piccillo, and Wu (2013) in which the error terms are interpreted respectively as the deviations of average agents' forecasts of individual future consumption (prices) from average forecast of aggregate consumption (price).

[4]In more recent experiments Kryvtsov and Petersen (2013), using the design of Pfajfar and Zakelj (2012), ask individuals to forecast both inflation and the output gap in a similar NK setting.

where $\Omega = (1 + \lambda\varphi\phi_\pi)^{-1}$, while $\bar{y}^e_{t+1} = \frac{1}{H}\sum_{i=1}^{H} y^e_{i,t+1}$ and $\bar{\pi}^e_{t+1} = \frac{1}{H}\sum_{i=1}^{H} \pi^e_{i,t+1}$ are respectively the average output gap and the average inflation predictions of the participants in the experiment. The experiment involved three treatments that explored the stabilizing properties of alternative parameterizations of the monetary policy rule. The details of the different treatments are described below.

Treatments

In order to study the stabilization properties of a monetary policy rule such as the Taylor rule (2.3), we ran three different treatments.

In treatment *a* the monetary policy responds only weakly to inflation rate fluctuations and we set the reaction coefficient $\phi_\pi = 1$, i.e., the Taylor principle does not hold. Moreover we set the inflation target $\bar{\pi} = 2$.[5]

In treatment *b* the monetary policy responds aggressively to inflation and we set the reaction coefficient $\phi_\pi = 1.5$, so that the Taylor principle holds. We keep the inflation target $\bar{\pi} = 2$.

Since $\bar{\pi} = 2$ could be a focal point for the inflation target, we ran a third treatment *c* in which we keep an aggressive monetary policy with $\phi_\pi = 1.5$, but we set the inflation target to $\bar{\pi} = 3.5$ to investigate the robustness of the policy rule to alternative target values.

Table 1 summarizes all treatments implemented in the experiments. In total 216 subjects participated in the experiment in 18 experimental economies, 6 for each of the treatments *a*, *b*, and *c* with 12 subjects each. Total average earnings over all subjects were € 32.

	ϕ_π	$\bar{\pi}$	# groups	average earnings π (y) in €
Treatment *a*	1	2	6	24 (26)
Treatment *b*	1.5	2	6	33 (32)
Treatment *c*	1.5	3.5	6	30 (29)

Table 1: Experimental treatments summary

2.2 Experimental design

The experiment took place in the CREED laboratory at the University of Amsterdam, March-May 2009 and February-March 2013. Subjects were divided randomly to form two groups of six individuals, one group forecasting inflation, the other group forecasting the output gap. Most subjects are undergraduate students from

[5]Notice that when the policy parameter ϕ_π is equal to 1, the system (2.4) exhibits a continuum of equilibria.

Economics, Chemistry and Psychology. At the beginning of the session each subject can read the instructions (see Appendix B, (Translation of Dutch) Instructions for participants) on the screen, and subjects receive also a written copy. Participants are instructed about their role as forecasters and about the experimental economy. They are assumed to be employed in a private firm of professional forecasters for the key variables of the economy under scrutiny i.e. either the inflation rate or the output gap. Subjects have to forecast either inflation or the output gap two periods ahead, as forecasts for period $t+1$ are made in period t with the information set consisting of macro variables up to $t-1$, for 50 periods. We give subjects some general information about the variables that describe the economy: the output gap (y_t), the inflation rate (π_t) and the interest rate (i_t). Subjects are also informed about the expectations feedback, that realized inflation and output gap depend on (other) subjects' expectations about inflation and output gap. They also know that inflation and output gap are affected by small random shocks to the economy. Subjects did *not* know the equations of the underlying law of motion of the economy nor did they have any information about its steady states. In short, subjects did not have quantitative details, but only qualitative information about the economy, which is a standard strategy in learning to forecast experiments (see Duffy (2008) and Hommes (2011)).

The payoff function of the subjects describing their score that is later converted into Euros is given by

$$score = \frac{100}{1+f}, \tag{2.5}$$

where f is the absolute value of the forecast error expressed in percentage points. The points earned by the participants depend on how close their predictions are to the realized values of the variable they are forecasting. Information about the payoff function is given graphically as well as in table form to the participants (see Fig. 1). Notice that the prediction score increases sharply when the error decreases to 0, so that subjects have a strong incentive to forecast as accurately as they can; see also Adam (2007) and Pfajfar and Zakelj (2012), who used the same payoff function.

Absolute forecast error	0	1	2	3	4	9
Score	100	50	$\frac{33}{3}$	25	20	10

In each period individuals can observe on the left side of the screen the time series of realized inflation rate, output gap and interest rate as well as the time series of their own forecasts. The same information is displayed on the right hand

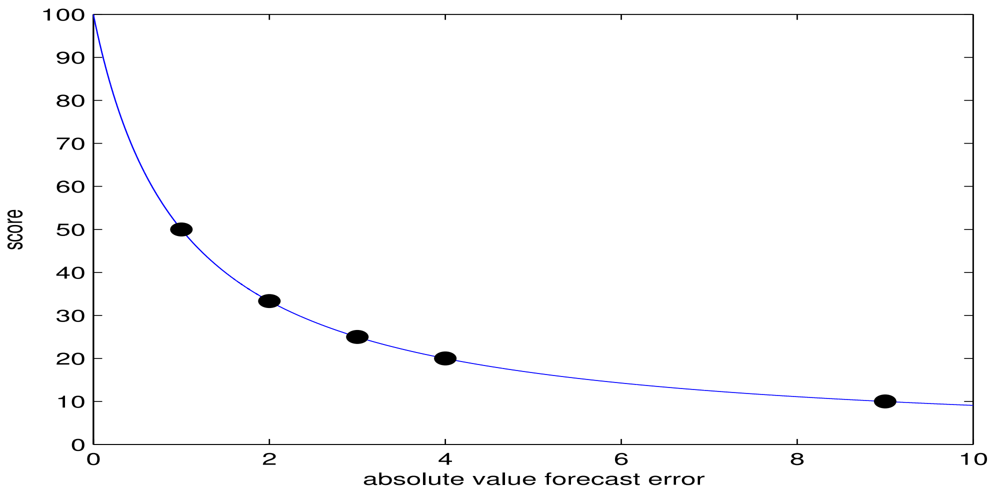


Figure 1: Payoff function

side of the screen in table form, together with subjects own predictions scores (see Fig. 2). Subjects did not have any information about the forecasts of others.

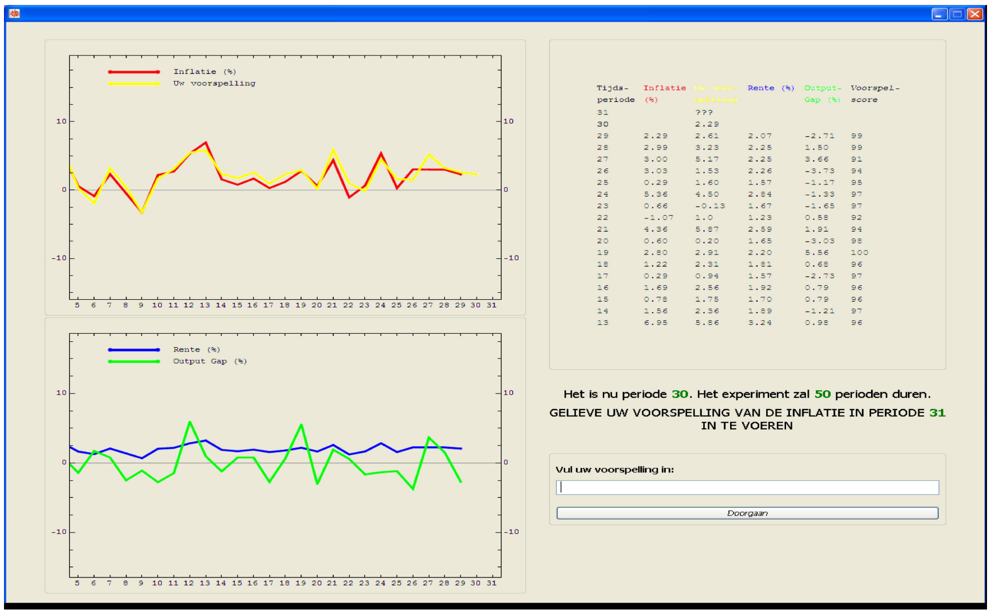


Figure 2: Computer screen for inflation forecasters with time series of inflation forecasts and realizations (top left), output gap and interest rate (bottom left) and table (top right).

2.3 Experimental results

This subsection describes the results of the experiment. We fix the parameters at the Clarida, Gali, and Gertler (2000) calibration, i.e. $\rho = 0.99$, $\varphi = 1$, and $\lambda = 0.3$.

In treatments a and b we set the inflation target to $\bar{\pi} = 2$, while in treatment c we set the inflation target to $\bar{\pi} = 3.5$. Fig. 3 depicts the behavior of the output gap, inflation and individual forecasts in six different sessions of treatment a. The dashed lines in the figure represent the RE steady states for inflation and the output gap corresponding to the inflation target of 2 and the respective output gap equilibrium level of 0.07. In this treatment we observe two different types of aggregate behavior. The first three groups show (almost) convergence to a non-fundamental steady state.[6] In the last three groups we observe an extremely unstable behavior. In group 4 the nominal interest rate hits the zero lower bound in period 11 and the economy enters a severe recession and never recovers.[7] In group 5 after period 10 we observe inflation and inflation forecast following a self-fulfilling exploding path.[8] In group 6 we observe an initial upward trend in inflation which is then reversed in period 19 and a downward inflation spiral follows. The interest rate hits the zero lower bound in period 27 and the economy enters a severe recession and never recovers.[9]

Fig. 4 shows the behavior of the output gap, inflation and individual forecasts in six different sessions of treatment b. In this treatment we observe convergence to the 2 percent fundamental steady state, although the converging paths are different. In groups 1, 4, and 6, after some (small) initial oscillations, inflation and output gap converge more or less monotonically, while in groups 2, 3, and 5 the convergence is oscillatory.

Fig. 5 shows the behavior of the output gap, inflation and individual forecasts in six different sessions of treatment c. In this treatment we observe convergence to the fundamental steady state of 3.5 percent for inflation and 0.12 for the output gap in four out of six experimental economies (with convergence being both monotonic and oscillatory), while in two groups, namely groups 2 and 6, we observe persistent oscillatory behavior of inflation and output around the steady state.

The experimental evidence presented suggests that a monetary policy that responds aggressively to deviations of the inflation rate from the target ($\phi_\pi = 1.5$) stabilizes fluctuations in inflation and output and leads the economy to the desired target. The value of the target seems to have little influence on the stabilizing

[6]Note that group 1 ends in period 26 because of a crash of one of the computers in the lab. Moreover in group 2 we observe a clear end-effect. In fact, participant 3 predicted an inflation rate of 100% in the last period, causing actual inflation to jump to about 20%.

[7]Wild oscillations following period 16 are not meaningful from an economic point of view as subjects reach the minimum value of inflation they were allowed to submit as forecast, i.e., −100%.

[8]Wild oscillations following period 36 are not meaningful from an economic point of view as subjects reach the maximum value of inflation they were allowed to submit as forecast, i.e., +1000%.

[9]Dynamics following period 37 are not meaningful from an economic point of view as subjects reach the minimum value of inflation they were allowed to submit as forecast, i.e., −100%.

properties of the monetary policy rule. On the other hand, when the interest rate rule reacts weakly to inflation fluctuations ($\phi_\pi = 1$), we observe convergence to non-fundamental steady states or exploding behavior.

Table 2 summarizes the quadratic distance of inflation and output gap from its target for all treatments. The table confirms our earlier graphical observation that a more aggressive Taylor rule stabilizes inflation and the output gap.

Group	Inflation	Output gap
a-1	0.4316	0.0444
a-2	1.2159	0.1073
a-3	3.3633	0.0508
a-4	8367.4795	2864.3735
a-5	113302.9587	25951.2481
a-6	5736.6674	8071.3761
a (median)	2870.0153	1432.2404
b-1	0.4804	0.1865
b-2	0.4366	0.2256
b-3	0.1638	0.0527
b-4	0.0183	0.0652
b-5	0.1437	0.0764
b-6	0.0254	0.0171
b (median)	0.1537	0.0708
c-1	0.2218	0.4075
c-2	0.2028	0.1894
c-3	0.0180	0.0814
c-4	0.0474	0.1106
c-5	0.1335	0.3479
c-6	0.1132	0.2168
c (median)	0.1234	0.2031

Table 2: Average quadratic difference from the target

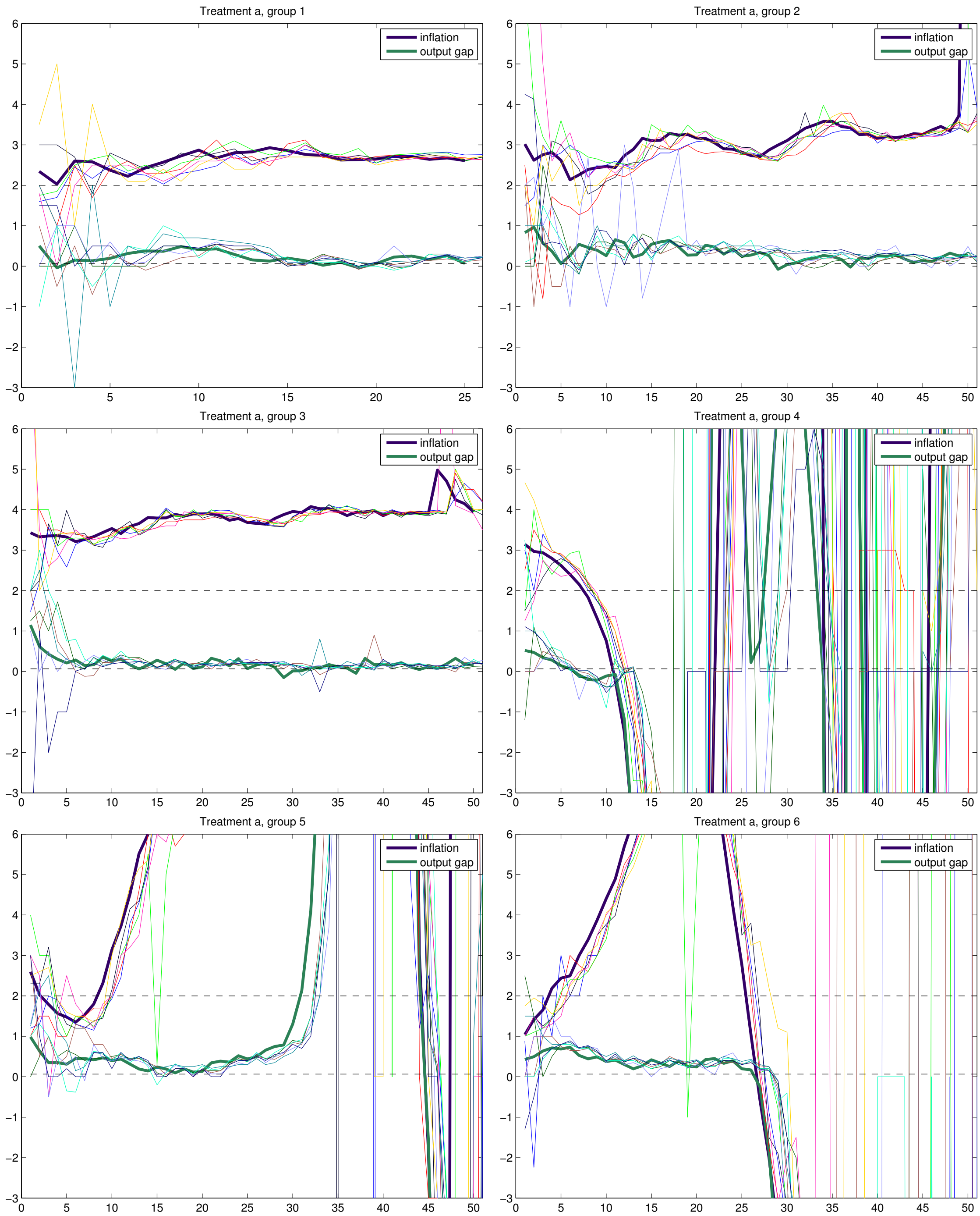


Figure 3: **Treatment** a. Blue thick line: realized inflation; green thick line: realized output gap; thin lines: individual forecasts for inflation and output gap.

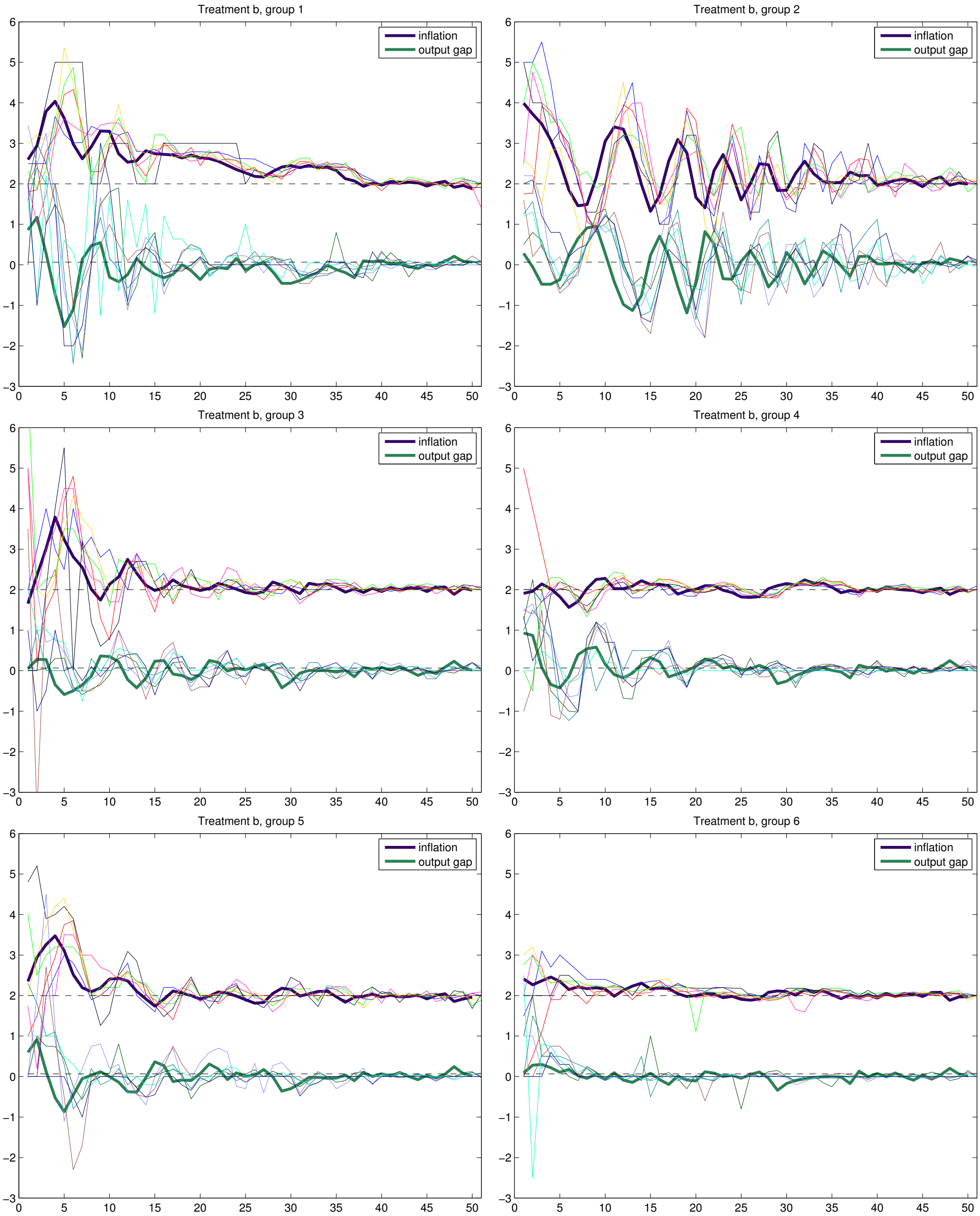


Figure 4: **Treatment** *b*. Blue thick line: realized inflation; green thick line: realized output gap; thin lines: individual forecasts for inflation and output gap.

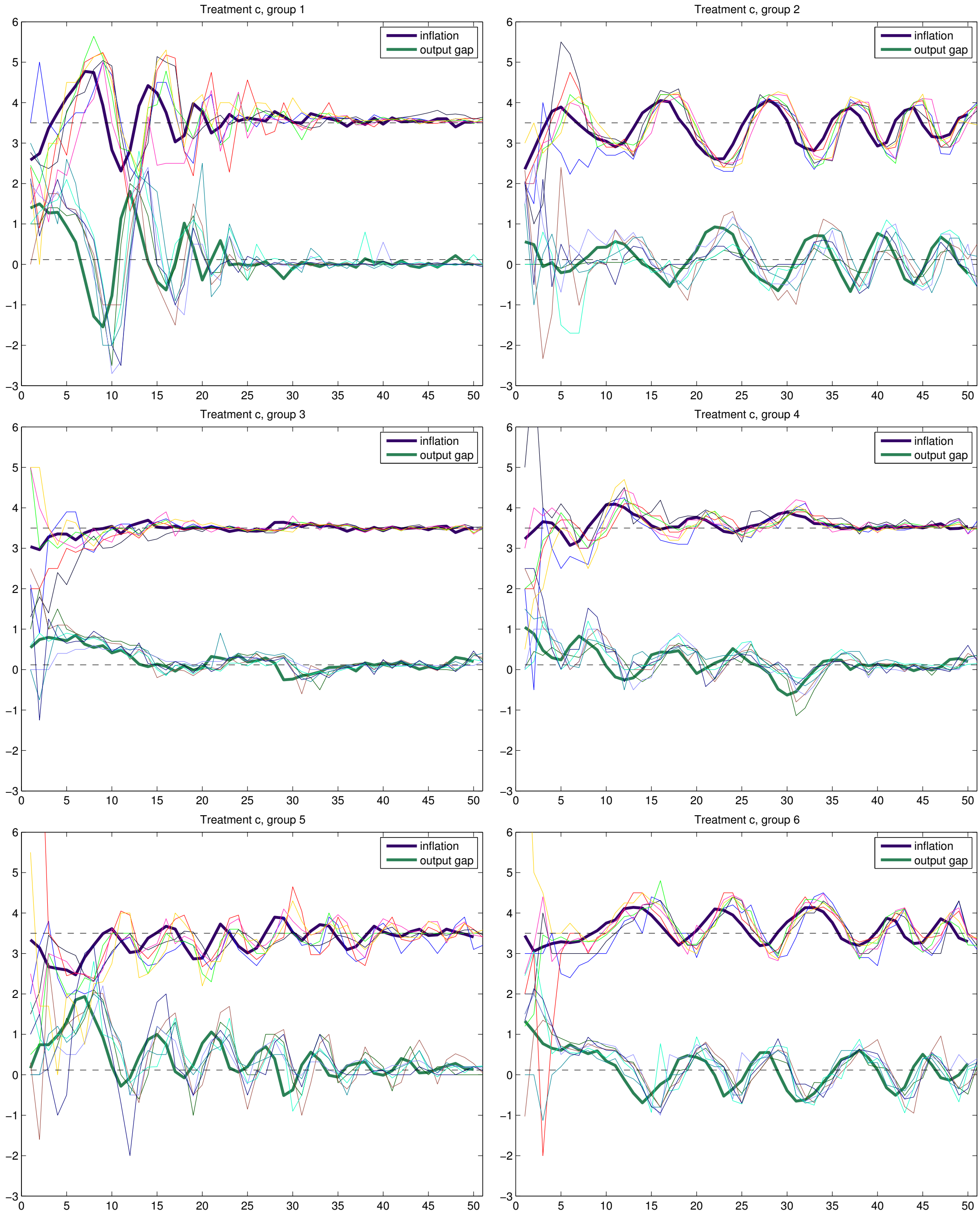


Figure 5: **Treatment** *c*. Blue thick line: realized inflation; green thick line: realized output gap; thin lines: individual forecasts for inflation and output gap.

3 Individual and aggregate behavior

The goal of this section is to characterize individual forecasting behavior and explain the emergence of the four different patterns observed in aggregate behavior of inflation and output, namely convergence to (some) equilibrium level, explosive inflationary or deflationary spirals, dampened oscillations and persistent oscillations, using a simple behavioral model of learning.

3.1 Individual forecasting behavior

The fact that different types of aggregate behavior arise in our experiments suggests that heterogeneous expectations play an important role in determining the aggregate outcomes. Indeed, a stylized fact that emerged from the investigation of individual experimental data is that there is a pervasive heterogeneity in the forecasting rules used by the subjects in the experiment. For each participant we estimated a simple linear prediction rule of the form

$$\pi^e_{h,t+1} = \alpha_1\pi_{t-1} + \alpha_2\pi_{h,t} + (1-\alpha_1-\alpha_2)\frac{1}{39}\sum_{t=12}^{50}\pi_t + \alpha_3(\pi_{t-1}-\pi_{t-2}) + \mu_t \tag{3.1}$$

$$y^e_{h,t+1} = \gamma_1 y_{t-1} + \gamma_2 y_{h,t} + (1-\gamma_1-\gamma_2)\frac{1}{39}\sum_{t=12}^{50} y_t + \gamma_3(y_{t-1}-y_{t-2}) + \nu_t\,, \tag{3.2}$$

in which $\pi^e_{h,t+1}$ and $y^e_{h,t+1}$ refer to the inflation or output gap forecast of participant h for period $t+1$. Rules (3.1) - (3.2) are referred to as *First-Order Heuristics* and can be interpreted as anchoring-and-adjustment heuristics *à la* Tversky and Kahneman (1974).[10] The first three terms in (3.1) and (3.2) are a weighted average of the latest realization of the forecasting objective, the latest own prediction and the forecasting objective's sample mean (excluding a learning phase).[11] This weighted average is the (time varying) "anchor" of the prediction, which is a zeroth order extrapolation from the available data at period t. The fourth term in (3.1) and (3.2) is a simple linear, i.e. first order, extrapolation from the two most recent realizations of the forecasting objective; this term is the "adjustment" or trend extrapolation part of the heuristic. An advantages of the FOH rule is that it simplifies to well-known rules-of-thumb for different boundary values of the parameter space. For example, the inflation prediction rule (3.1) reduces to Naive Expectations if

[10] For other applications of the FOH in modeling expectation formation, see Heemeijer, Hommes, Sonnemans, and Tuinstra (2009).

[11] In the estimation of (3.1) and (3.2) we included the sample mean of inflation resp. output, which is of course not available to the subjects at the moment of the prediction, but acts as a proxy of the equilibrium level. In the heuristic switching model of section 3.2 one of the rules will use the past observed sample average. Similar estimation results are obtained when we replace the sample average by a constant.

$\alpha_1 = 1$, $\alpha_2 = \alpha_3 = 0$; it reduces to Adaptive Expectations if $\alpha_1 + \alpha_2 = 1$, $\alpha_3 = 0$; and it reduces to the simplest Trend-Following rule if $\alpha_1 = 1$, $\alpha_2 = 0$ and $\alpha_3 > 0$. In the more flexible case $\alpha_1 + \alpha_3 = 1$, $\alpha_2 = 0$ the anchor is time varying; we will refer to this case as a *learning anchor and adjustment* (LAA) rule.

For most of the subjects a simple FOH fits individual forecasting behavior well.[12] Moreover, for quite a number of subjects the FOH reduces to one of the special cases discussed above, in particular to adaptive expectations in relatively stable groups, to a strong trend-extrapolating rule in the unstable, exploding groups, to a weak trend-following rule in dampened oscillatory groups and to a LAA rule in oscillating groups. Therefore, the heuristics switching model discussed below contains these four types of forecasting rules.

Another interesting stylized fact that emerged from the experimental data is that individual forecasting behaviors entail a learning process which takes the form of switching from one heuristic to another. Evidence of switching behavior can be found by inspecting the time series of individual forecasts. Here we report in Fig. 6 some graphical evidence of individual switching behavior.[13]

Fig. 6 shows the time series of some individual forecasts together with the realizations of the variable being forecasted. For every period t we plot the realized inflation or output gap together with the two period ahead forecast of the individual. In this way we can graphically infer how the individual prediction uses the last available observation. For example, if the time series coincide, the subject is using a naive forecasting strategy.

In Fig. 6(a) (group 2, treatment a), subject 2 strongly extrapolates changes in the output gap in the early stage of the experiment, but starting from period 18 she switches to a much weaker form of trend extrapolation.

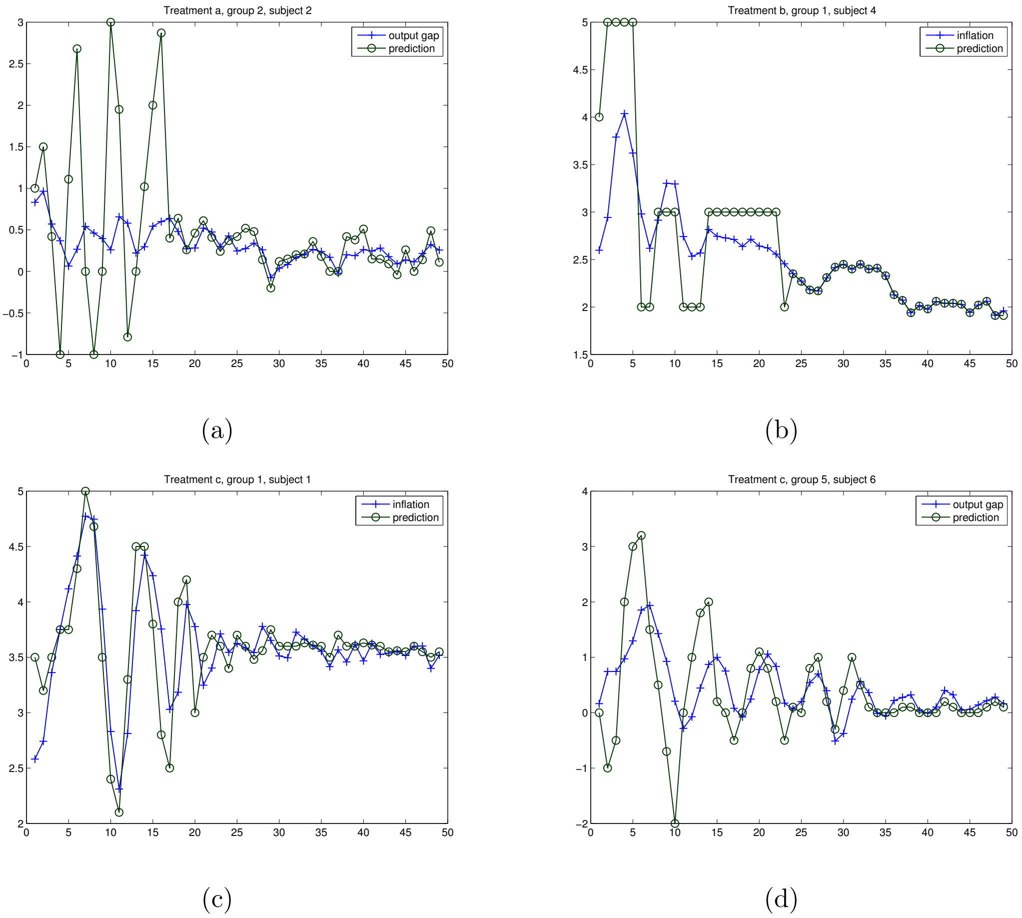
In Fig. 6(b) (group 1, treatment b), subject 4 switches between various constant predictors for inflation in the first 23 periods of the experimental session. She is in fact initially experimenting with three predictors, 2% 3% and 5%, and then switches to a naive forecasting strategy after period 23.

In Fig. 6(c) (group 1, treatment c), subject 1 is using a trend extrapolation strategy to forecast the inflation rate in the initial part of the experiment, i.e., when inflation fluctuates more. However, when oscillations dampen and inflation converges to the equilibrium level around period 30, she switches to an adaptive expectations strategy.

In Fig. 6(d) (group 5, treatment c), subject 6 starts by strongly extrapolating

[12] See Massaro (2012) for details about the estimation results.

[13] Direct evidence of switching behavior has been found in the questionnaires submitted at the end of the experiments, where participants are explicitly asked whether they changed their forecasting strategies throughout the experiment. About 42% of the participants answered that they changed forecasting strategy during the experiment.



(a) (b) (c) (d)

Figure 6: **Individual forecasting as switching between heuristics.** For every period the subject's forecast $x^e_{i,t+2}$ (green) and the variable being forecasted, x_t, with $x = \pi, y$, are reproduced.

trends in the output gap, and after period 16 she uses a much weaker form of trend extrapolation. In a later stage of the experiment, i.e., from period 35 on, the same subject switches to adaptive forecasting behavior.

3.2 Heuristics switching model (HSM)

In the light of the empirical evidence for heterogeneous expectations and individual switching behavior, we now introduce a simple model which features evolutionary selection between different forecasting heuristics in order to reproduce individual as well as aggregate experimental data.

Anufriev and Hommes (2012) developed a heuristics switching model (HSM) along the lines of Brock and Hommes (1997), to explain different types of aggregate price fluctuations –monotonic convergence, dampened oscillations and persistent oscillations– in the asset pricing experiment of Hommes, Sonnemans, Tuinstra, and van de Velden (2005). The key idea of the model is that the subjects chose between simple heuristics depending upon their relative past performance. The performance measure of a forecasting heuristic is based on its absolute forecasting error and has exactly the same form as the payoff function used in the experiments. More precisely, the *performance measure* of heuristic h up to (and including) time $t-1$ is given by

$$U_{h,t-1} = \frac{100}{1 + |x_{t-1} - x^e_{h,t-1}|} + \eta U_{h,t-2}, \tag{3.3}$$

with $x = \pi, y$. The parameter $0 \leq \eta \leq 1$ represents the *memory*, measuring the relative weight agents give to past errors of heuristic h.

Given the performance measure, the fraction of rule h is updated according to a *discrete choice model with asynchronous updating*

$$n_{h,t} = \delta n_{h,t-1} + (1-\delta)\frac{\exp(\beta U_{h,t-1})}{Z_{t-1}}, \tag{3.4}$$

where $Z_{t-1} = \sum_{h=1}^{H} \exp(\beta U_{h,t-1})$ is a normalization factor. The asynchronous updating parameter $0 \leq \delta \leq 1$ measures the inertia in the fraction of rule h, reflecting the fact that not all the participants update their rule in every period or at the same time. The parameter $\beta \geq 0$ represents the intensity of choice measuring how sensitive individuals are to differences in heuristics performances.

The evolutionary model can include an arbitrary set of heuristics. Since our goal is to explain the different observed patterns of inflation and output in the experiment, we keep the number of heuristics as small as possible and consider a model with only four forecasting rules. The four heuristics in the model are summarized in Table 3. In fact, we chose exactly the same 4-type HSM that has successfully been used by Anufriev and Hommes (2012) to explain the different price patterns observed in the asset pricing experiment in Hommes, Sonnemans, Tuinstra, and van de Velden (2005). Hommes (2011) shows that the same model also explains the positive and negative feedback experiments in Heemeijer, Hommes, Sonnemans, and Tuinstra (2009). This illustrates the robustness of the HSM across different experimental settings. The model is not very sensitive to the parameter values and for different choices of the coefficients of the four rules very similar results are obtained as those presented below.

3.3 Self-organization of heterogeneous expectations

In this section we discuss the empirical validation of the evolutionary switching model and show that the HSM fits both individual forecasting and aggregate macro behavior. We only report simulations for four groups (treatment a, group 5, treat-

Table 3: 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	$x^e_{4,t+1} = 0.5(x^{av}_{t-1} + x_{t-1}) + (x_{t-1} - x_{t-2})$

ment *b*, group 2, treatment *c*, groups 2 and 3), representative of the four different types of aggregate behaviors observed in the experiment. The results for experimental economies with analogous qualitative behavior are similar. Our HSM explains these four different macro patterns as emergent properties of the self-organization process of heterogeneous individual expectations.

One-period ahead simulations

The one-period ahead model simulations use exactly the same information as available to subjects in the experiment. The simulations are initialized by two initial values for inflation and output gap, π_1, y_1, π_2 and y_2, and equal initial weights $n_{h,\text{in}}$, $1 \leq h \leq 4$ used for periods 3 and 4. Given the values of inflation and output gap for periods 1 and 2, the heuristics forecasts can be computed. Using initial weights of the heuristics, inflation and output gap for periods 3 and 4 can be computed. Starting from period 5 the evolution according to the model's equations is well defined and at each point in time the HSM uses the same information as available to subjects in the experiment[14]. Once we fix the four forecasting heuristics, there are three "learning" parameters left in the model: β, η, and δ. We used exactly the same set of learning parameters as in Anufriev and Hommes (2012), namely $\beta = 0.4$, $\eta = 0.7$, $\delta = 0.9$, inflation and the output gap in the first two periods are fixed at the corresponding experimental data, while all heuristics have equal initial fractions, i.e., 0.25.

We performed tests of the equality of observed and simulated mean and variance to all groups in treatments b and c (i.e. 12 groups in total; excluding the unstable, non-stationary treatment a)[15]. The results are as follows:

- 1% confidence level:

[14]Massaro (2012) also investigates 50-period ahead simulations of the HSM and shows that they reproduce qualitatively the different types of behaviour, i.e. monotonic convergence, dampened oscillations, persistent oscillations and exploding inflationary/deflationary spirals.

[15]We performed the tests on the equality of observed and simulated mean and variance on a sample that goes from period 5 to the end of the experimental session in order to minimize the impact of the initial conditions.

  - Mean: we do not reject the null (equality) in 22/24 cases (92%);
  - Variance: we do not reject the null (equality) in 19/22 cases (86%).

- 5% confidence level:
  - Mean: we do not reject the null (equality) in 19/24 cases (79%);
  - Variance: we do not reject the null (equality) in 16/22 cases[16] (73%).

The HSM is thus capable to match the first two moments, the mean and the variance, of the aggregate variables in most experimental groups.

Four types of macro behavior

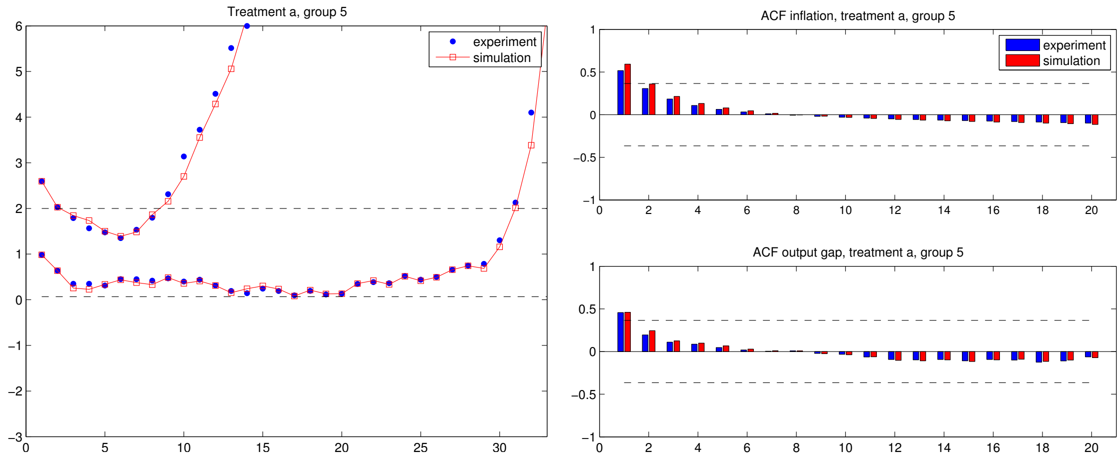
Figs. 7-10 illustrate that the model is able to reproduce qualitatively all four different aggregate patterns observed in the experiment, which are, explosive inflationary or deflationary paths, convergence to (some) equilibrium, dampened oscillations and persistent oscillations. Figs. 7-10 are organized as follows:

- top left panels: experimental and simulated HSM aggregate data for inflation and output gap;
- middle panels: six individual forecasts for inflation and output gap from experiment (left) and four simulated heuristics of HSM;
- top right panel: autocorrelation function (ACF) of average individual forecasts in experiment and average forecasts of four heuristics in HSM;
- bottom panels: evolution of the fractions of each of the 4 forecasting heuristics for inflation (left) and output gap (right).

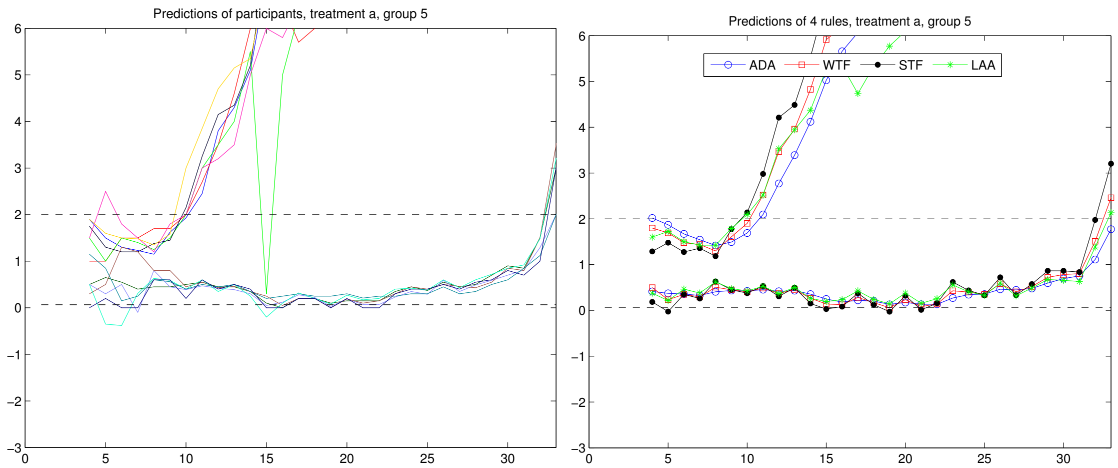
These figures show that the HSM fits both aggregate and individual behavior. The one-step ahead simulations closely track each of the four aggregate patterns of inflation and output, exploding behavior (Fig. 7), convergence to a stable steady state (Fig. 8), dampened oscillations (Fig. 9) and persistent oscillations (Fig. 10). Moreover, the subjects' individual forecasts are coordinated in a similar way as the forecasts of the 4 heuristics. The fact that the ACF of the average forecast in the experiments is very similar to the ACF of the simulated average forecasts, weighted by the fractions (3.4) in the HSM, supports this observation.

Finally, let us look at the evolution of the fractions (3.4) in more detail. The dominating strategy to a large extent determines the aggregate outcome. In different groups different heuristics are taking the lead after starting from a uniform

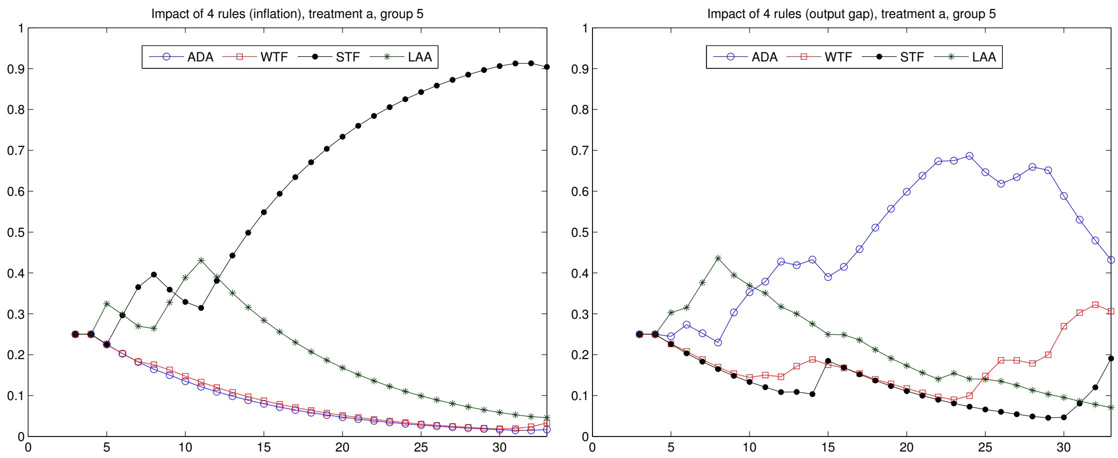
[16]In two cases we were not able to perform the test on the variance due to non-stationary.



(a)

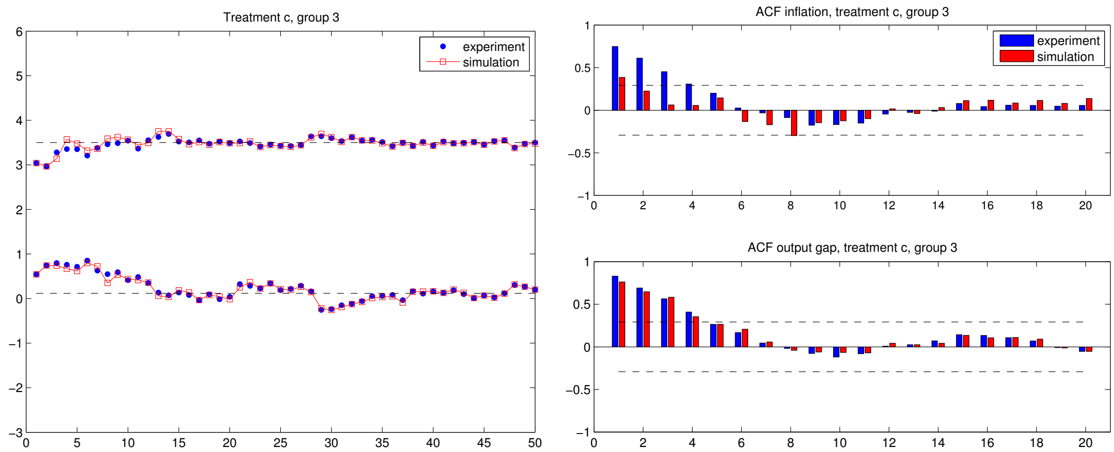


(b)

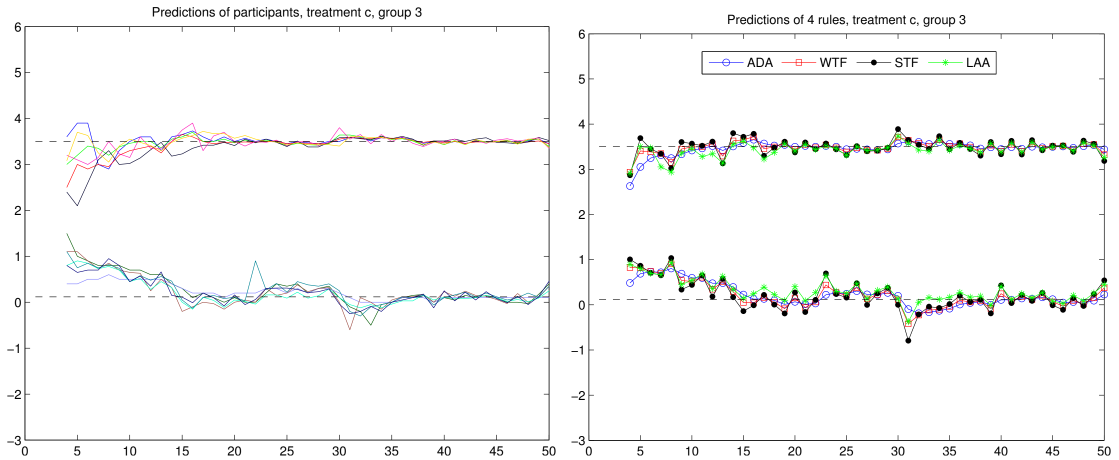


(c)

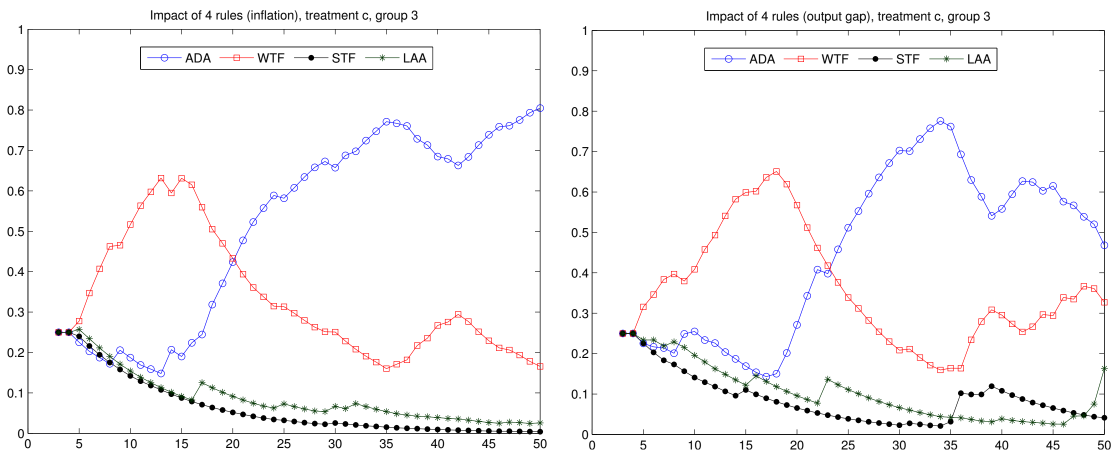
Figure 7: Experimental data (blue points) and one-period ahead heuristics switching model (HSM) simulations (red lines) for Treatment a, group 5, with exploding inflation and output gap (top left panel); Middle panels show individual predictions in experiment (left) and predictions of 4 HSM heuristics (right); Top right panel compares ACFs of average individual forecasts and average expectations in HSM for inflation (top) and output gap (bottom); Bottom panel shows fractions of the 4 heuristics for forecasting inflation (left) and output gap (right). Coordination on strong trend-following rule explains explosive dynamics.



(a)

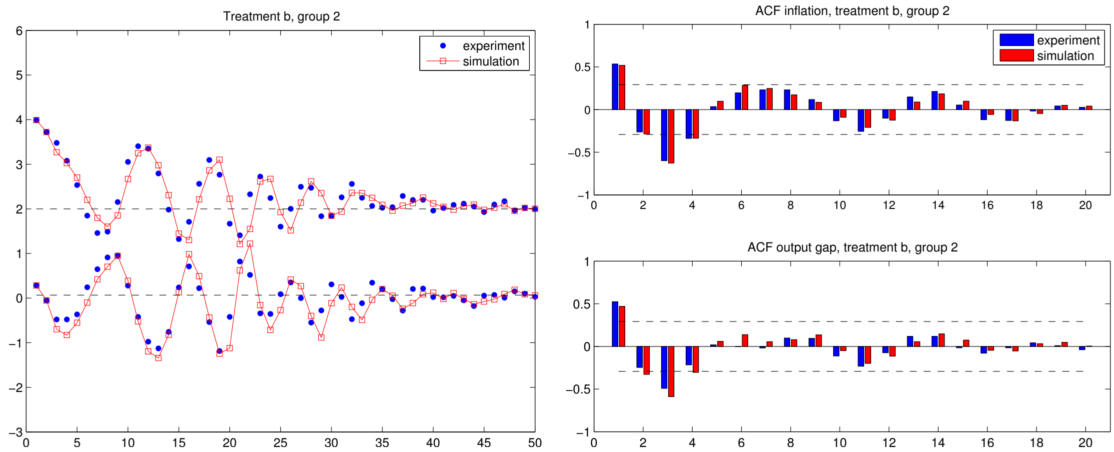


(b)

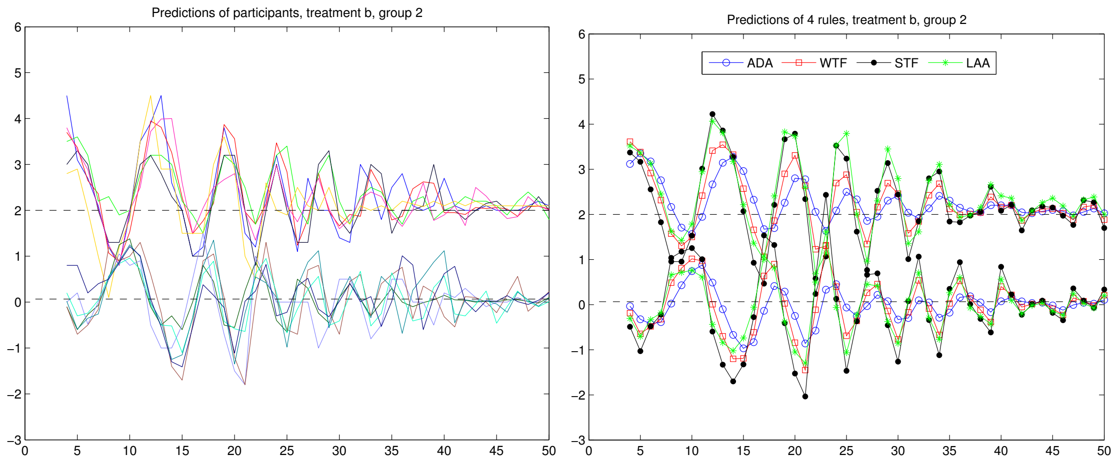


(c)

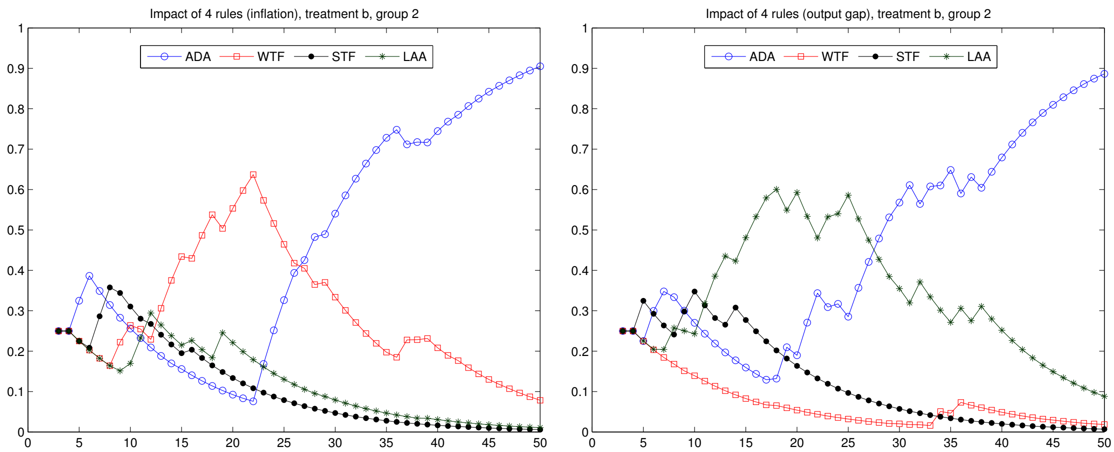
Figure 8: Experimental data (blue points) and one-period ahead heuristics switching model (HSM) simulations (red lines) for Treatment c, group 3, with stable inflation and output gap (top left panel); Middle panels show individual predictions in experiment (left) and predictions of 4 HSM heuristics (right); Top right panel compares ACFs of average individual forecasts and average expectations in HSM for inflation (top) and output gap (bottom); Bottom panel shows fractions of the 4 heuristics for forecasting inflation (left) and output gap (right). Coordination on adaptive expectations explains convergence of inflation and output to RE steady state.



(a)

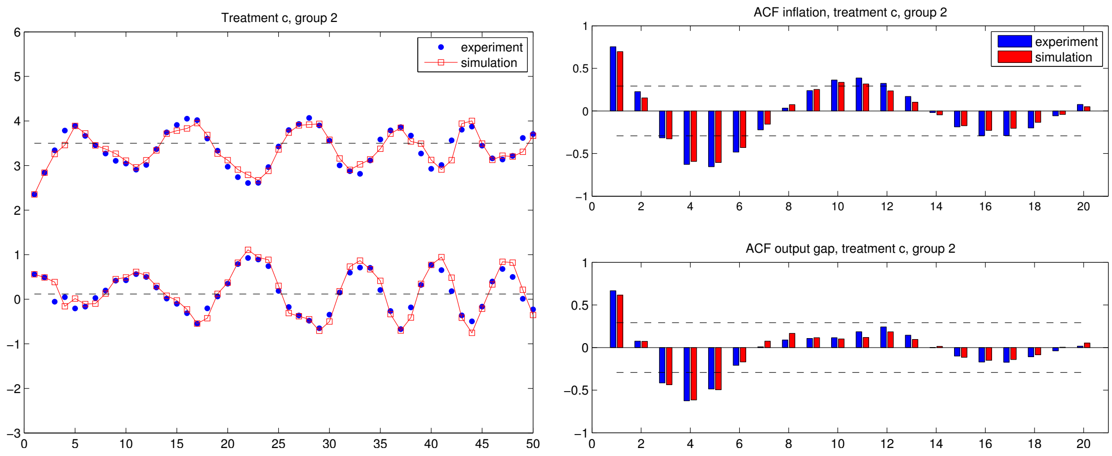


(b)

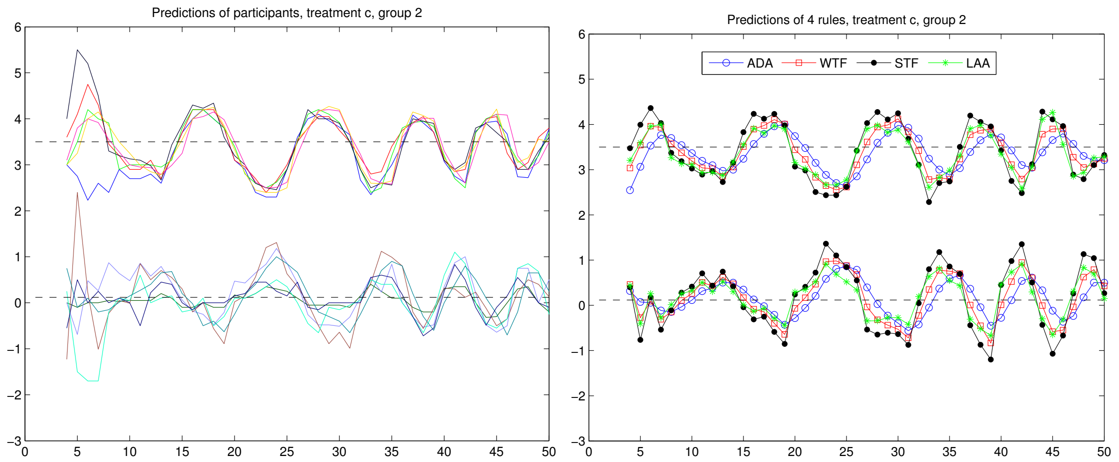


(c)

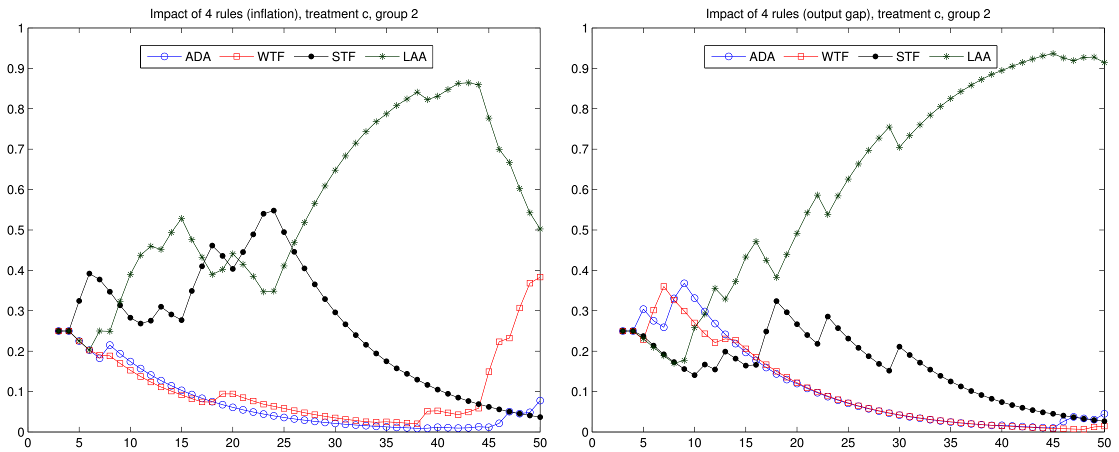
Figure 9: Experimental data (blue points) and one-period ahead heuristics switching model (HSM) simulations (red lines) for Treatment b, group 2, with dampening oscillations in inflation and output gap (top left panel); Middle panels show individual predictions in experiment (left) and predictions of 4 HSM rules (right); Top right panel compares ACFs of average individual forecasts and average expectations in HSM for inflation (top) and output gap (bottom); Bottom panel shows fractions of the 4 rules for forecasting inflation (left) and output gap (right). In the first 20-25 periods the weak trend-following (WTR) rule dominates inflation forecasting and the learning anchor and adjustment (LAA) trend-following rule dominates output gap forecasting. After period 20 the share of adaptive expectations rapidly increases and coordination on adaptive expectations explains dampened oscillations.



(a)



(b)



(c)

Figure 10: Experimental data (blue points) and one-period ahead heuristics switching model (HSM) simulations (red lines) for Treatment c, group 2, with persistent oscillations in inflation and output (top left panel); Middle panels show individual predictions in experiment (left) and predictions of 4 HSM rules (right); Top right panel compares ACFs of average individual forecasts and average expectations in HSM for inflation (top) and output gap (bottom); Bottom panel shows fractions of the 4 rules for forecasting inflation (left) and output gap (right). In the first 20 periods the weak (WTR) and strong trend-following (STR) rules have a relatively large impact due to weak or strong trends in inflation and output. After period 25 the share of the learning anchor and adjustment (LAA) rule dominates and coordination on LAA explains persistent oscillations.

distribution. As we will see, the learning process may self-organize into coordination on one of the four rules which then determines (long run) aggregate behavior.

In treatment *a* group 5 (Fig. 7) inflation follows an upward explosive path starting in the early phase of the experiment, amplified by a sharp increase in the share of the strong trend-following (STR) rule which, by the end of the simulation (period 33), reaches a share of around 90%.[17] Interestingly, the output gap follows a different path. In fact, output is more or less stable until period 28, with adaptive expectations (ADA) dominating. However, the sharp increase of the output gap in the late stages of the simulation, caused by rising inflation expectations, leads subjects to coordinate on an upward trend, resulting in a decline of the ADA rule and in an increase of the impacts of the weak and strong trend-following rules (WTF and STF). An explosive inflationary path and exploding output are thus explained by coordination of individual expectations on strong trend-following rules.

Fig. 8 (bottom panels) describes the dynamics of the fractions of different heuristics in treatment c, group 4, converging to the 3.5% inflation target. Both inflation and output gap start at some distance from the target equilibrium and there is an initial phase in which both aggregate variables follow an adjustment path towards equilibrium, characterized by an (almost) monotonic increase (decrease) of inflation (output gap). In this phase the weak trend (WTF) rule takes the lead. However, as soon as the aggregate variables reach the target equilibrium level, the share of WTF declines and the adaptive expectations (ADA) rule dominates. Convergence to the (RE) steady state is thus explained by coordination of individual expectations on adaptive expectations.

Fig. 9 depicts the dynamics of different forecasting rules for treatment *b* group 2, characterized by dampened oscillations. The one-step ahead simulation shows a rich evolutionary competition among heuristics. In the initial part of the experiment, starting in period 7, the strong trend-following (STF) rule matches the strong decline in the inflation rate and its share increases to about 35%. However the fraction of the STF rule starts to decrease after it misses the first turning point. After the initial phase of strong trend in inflation, oscillations slowly dampen and the impact of the weak trend (WTF) rule starts to rise. Around period 25, when inflation oscillations have dampened, adaptive expectations (ADA) dominates the other heuristics. The evolutionary selection dynamics are somewhat different for the output gap predictors. Oscillations of the output gap are more frequent and this implies a relatively worse forecasting performance of the pure trend extrapolating heuristics which tend to overshoot more often. In the middle phase of the experimental session the learning anchor and adjustment (LAA), which performs

[17] Group 5 of treatment *a* has been simulated for 33 periods due to the observed explosive behavior in the late phases of the experiment.

better in forecasting turning points, dominates output gap forecasting with a share over 50%. However, with dampening oscillations the impact of the LAA rule gradually decreases and adaptive expectations (ADA) starts increasing after period 25 and dominates in the last 10 periods with a rising share of more than 80%.

Finally, Fig. 10 illustrates the dynamics of different forecasting rules for the experimental groups characterized by persistent oscillations. In treatment c group 2 both inflation and the output gap oscillate around the target level for the entire experimental session. The heuristics switching model explains the sustained oscillations by coordination of most agents on a LAA rule. In fact, in the presence of cyclical oscillations, the purely extrapolative rules (WTF and STF) tend to overshoot the trend reversal. On the other hand, the LAA rule uses an anchor which is given by a weighted average of the sample mean and the last observation, predicting mean reversion towards equilibrium and making smaller forecast errors at the turning points of the trend. For inflation, the LAA rule dominates reaching a peak share of about 85% in period 43 and slowly decreases afterwards as the amplitude of oscillations decreases, while for the output gap, which presents stable oscillations until the end of the session, the share of the LAA rule is still higher than 90% at the end of the simulation.

The one-step ahead simulations show that the proportion of agents using (strong) trend extrapolation rules plays an important role for the stability of aggregate variables. Groups with a lower fraction of trend extrapolation rules are more stable than groups with a higher proportion of trend following behavior. Instead, having more agents that follow adaptive expectations schemes has a stabilizing effect on aggregate dynamics, while oscillatory behavior is associated with anchoring and adjustment heuristics. Interestingly, Pfajfar and Zakelj (2012) reach a similar conclusion and note that a higher proportion of trend extrapolation increases the standard deviation of inflation while having more agents behaving according to adaptive expectations decreases the standard deviation of inflation.

Forecasting performance

Table 4 compares the MSE of the one-step ahead predictions in the experimental groups for 9 different models: the rational expectation prediction (RE), six homogeneous expectations models (naive expectations, fixed anchor and adjustment (AA) rule,[18] and each of the four heuristics of the switching model), the switching model with benchmark parameters $\beta = 0.4$, $\eta = 0.7$, and $\delta = 0.9$, and the "best"

[18]In the AA rule we consider the full sample mean, which is a proxy of the equilibrium level, as an anchor. In the LAA rule instead we use the sample average of all the previous realizations that are available at every point in time as an anchor.

Table 4: MSE over periods 5 – 50 of the one-period ahead forecast.

Model	3a-1	3a-2	3a-3	3a-4	3a-5	3a-6	3b-1	3b-2	3b-3	3b-4	3b-5	3b-6	3c-1	3c-2	3c-3	3c-4	3c-5	3c-6
RE	*	*	*	*	*	*	0.4892	0.4629	0.1251	0.0470	0.1049	0.0190	0.5061	0.3533	0.0462	0.1148	0.4307	0.2841
naive	0.0085	0.0252	0.0315	24.2498	2741.5945	1.4855	0.0881	0.2416	0.0347	0.0170	0.0250	0.0051	0.1749	0.1390	0.0037	0.0262	0.0928	0.0793
AA	0.0196	0.0656	0.0770	19.9063	3022.5834	4.3801	0.0971	0.1229	0.0577	0.0162	0.0316	0.0202	0.1153	0.0465	0.0265	0.0179	0.0836	0.0457
ADA	0.0405	0.0403	0.0628	43.7808	3865.9582	4.1680	0.1672	0.3465	0.0681	0.0382	0.0644	0.0169	0.3598	0.2791	0.0256	0.0662	0.1506	0.1648
WTF	0.0097	0.0257	0.0445	12.0997	1888.6095	0.5089	0.0690	0.1723	0.0382	0.0132	0.0185	0.0109	0.1098	0.0831	0.0078	0.0140	0.0587	0.0427
STF	0.0333	0.0649	0.1273	4.7511	568.3589	2.3153	0.1359	0.3761	0.1332	0.0403	0.0602	0.0429	0.2867	0.1106	0.0347	0.0267	0.1178	0.0522
LAA	0.0167	0.0488	0.0764	23.2407	3146.9096	4.8301	0.0803	0.1375	0.0559	0.0162	0.0286	0.0191	0.1118	0.0472	0.0197	0.0119	0.0526	0.0397
4 rules (benchmark)	0.0120	0.0228	0.0323	6.3159	698.6560	0.2512	0.0548	0.1399	0.0391	0.0096	0.0121	0.0054	0.1042	0.0369	0.0049	0.0100	0.0387	0.0205
4 rules (best fit)	0.0094	0.0225	0.0300	0.6250	568.3979	0.2457	0.0539	0.1295	0.0376	0.0094	0.0111	0.0037	0.0742	0.0318	0.0046	0.0084	0.0370	0.0199
β	5.6000	10.0000	1.1000	10.0000	7.9000	0.1000	0.1000	6.2000	10.0000	6.3000	10.0000	0.8000	0.3000	0.4000	0.9000	0.1000	1.8000	0.1000
η	0.2000	0.5000	0.9000	0.3000	0.6000	0.8000	0.8000	0.3000	0.7000	0.7000	0.5000	1.0000	0.3000	1.0000	0.4000	0.6000	0.8000	0.8000
δ	0.7000	0.9000	0.7000	0.0000	0.0000	0.9000	0.9000	0.8000	0.9000	0.9000	0.9000	0.5000	0.7000	0.9000	0.8000	0.9000	0.9000	0.9000

Note that the squared prediction errors refer to the sum of the errors relative to inflation and to output gap. The MSE of the RE model for treatment a has not been computed due to indeterminacy. Moreover, the MSE for groups 1, 2, 4, 5, 6 of treatment a has been computed respectively for 25, 48, 15, 33, and 30 periods due to either ending effects or anomalous explosive behavior.

switching model fitted by means of a grid search in the parameters space.[19] The table shows that RE generally performs poorly and that, independent of the aggregate macro pattern of a particular group, the HSM is always the best or close to the best forecasting model. The fact that the best fitted HSM only performs slightly better than the benchmark HSM shows the robustness of the forecasting performance of the HSM with respect to the parameter choices.

4 Monetary policy and macroeconomic stability

In this section we evaluate the impact of monetary policy on the stability of the experimental macro system. Our experiment shows that a more aggressive interest rate policy, i.e., an increase of the coefficient ϕ_π from 1 to 1.5, stabilizes the macro economy, at least in the long run. The Learning-to-Forecast Experiments (LtFEs) and the heuristics switching model (HSM) fitted to these experimental data offer a simple and intuitive explanation of the stabilizing mechanism of a more aggressive monetary policy in managing the self-organization process of heterogeneous expectations of boundedly rational agents.

Earlier LtFEs Heemeijer, Hommes, Sonnemans, and Tuinstra (2009) and Bao, Hommes, Sonnemans, and Tuinstra (2012) have shown that, within a simple univariate environment, the type of expectations feedback –positive or negative– is crucial for macroeconomic stability. Positive (negative) expectations feedback means that the realized aggregate variable increases (decreases) when the average forecast increases (decreases). Positive feedback occurs e.g. in demand driven speculative asset markets, when more optimistic expectations induce higher asset demand and therefore higher asset prices. Negative feedback arises, e.g., in supply driven commodity markets, when more optimistic beliefs lead suppliers to produce more and, as a consequence, realized market clearing prices decrease. The LtFEs in Heemeijer, Hommes, Sonnemans, and Tuinstra (2009) and Bao, Hommes, Sonnemans, and Tuinstra (2012) show that negative feedback experimental markets are rather stable and converge quickly to the (unique) RE steady state. In contrast, positive feedback markets are rather unstable and typically do not converge, but fluctuate persistently around the RE steady state.

Anufriev and Hommes (2012) and Hommes (2013) fit a HSM to these positive and negative feedback experiments and provide an intuitive explanation of why positive feedback markets fluctuate. Under positive feedback, trend-following strategies perform rather well and, when the positive feedback is sufficiently strong,

[19] Massaro (2012) also investigates out-of-sample forecasting and shows e.g. that the benchmark 4-type HSM with 3 parameters outperforms a linear AR(2) model for both inflation and output gap with 6 parameters.

coordination of individual expectations on a trend-following strategy amplifies fluctuations leading to almost self-fulfilling equilibria very different from the RE steady state. In contrast, under negative feedback, trend-following strategies perform poorly and are driven out of the market by adaptive expectations, stabilizing price oscillations. Coordination on trend-following strategies thus destabilize, while coordination on adaptive expectations stabilizes the macrosystem. This is particularly true for near unit root (univariate) positive feedback systems, where coordination on trend-following rules seems more likely.[20]

The NK macroeconomic setting with two endogenous variables, inflation and the output gap, and a policy variable, the interest rate, is however much more complicated than these simple univariate systems and exhibits a mixture of positive and negative feedback. Inflation responds positively to output and inflation expectations according to the NKPC (Eq. 2.2). Output on the other hand, according to the IS curve (2.1), responds positively to output expectations and inflation expectations, but negatively to the nominal interest rate. Moreover, this negative feedback from the nominal interest rate to output becomes stronger when the coefficient ϕ_π of the Taylor interest rate rule increases. A sufficiently aggressive interest rate Taylor rule thus adds negative feedback to the macro-system and a more aggressive monetary policy is therefore potentially stabilizing.

How aggressive should monetary policy be to manage expectations of boundedly rational agents and stabilize the economy? Stated differently, how aggressive should the interest rate rule be to prevent trend-following expectation rules to survive in the competition between heterogeneous forecasting rules? To get some first intuition on this question assume that agents coordinate inflation expectations on the simple forecasting rule

$$\pi^e_{t+1} = \bar{\pi} + \gamma(\pi_{t-1} - \bar{\pi}). \tag{4.1}$$

According to this rule, if the coefficient $\gamma > 1$, agents expect inflation to move away from its target level $\bar{\pi}$. Substituting the forecasting rule (4.1) and the interest rate Taylor rule (2.3) into (2.1) yields an implied IS curve for output

$$y_t = \bar{y}^e_{t+1} - \varphi[\phi_\pi(\pi_t - \bar{\pi}) - \gamma(\pi_{t-1} - \bar{\pi})]. \tag{4.2}$$

It is then immediately clear that, in order for the feedback from the real interest

[20] In the positive feedback systems in Heemeijer, Hommes, Sonnemans, and Tuinstra (2009) and Bao, Hommes, Sonnemans, and Tuinstra (2012) the eigenvalues of the positive and negative feedback systems are $\lambda = +0.95$ (equal to a discount factor close to 1) and $\lambda = -0.95$.

rate rule on output to be negative we must have

$$\phi_\pi(\pi_t - \bar{\pi}) > \gamma(\pi_{t-1} - \bar{\pi}).$$

Stated differently, in order for the monetary policy to add negative feedback to the macro-system the interest rate response of the CB to the current deviation from the inflation target must at least offset the expected extrapolation of the deviation of inflation from its target.

This intuitive explanation, however, only takes inflation expectations into account. To be more precise about how monetary policy can stabilize the macro system, both inflation and output expectations need to be taken into account. Consider therefore the representation $(y_t, \pi_t) = F(y^e_{t+1}, \pi^e_{t+1})$ of the NK-model in (2.4), describing output gap and inflation as a linear function of both output gap and inflation expectations. The coefficients in the matrix

$$\begin{bmatrix} 1 & \varphi(1-\phi_\pi\rho) \\ \lambda & \lambda\varphi+\rho \end{bmatrix} \tag{4.3}$$

determine the sign of the feedback of expectations on realizations. Clearly, if $\phi_\pi < 1/\rho \approx 1$, then all feedback coefficients are positive. Hence, when the monetary policy is weak, i.e., $\phi_\pi \leq 1$, both inflation expectations and output gap expectations have a positive impact on both the realizations of inflation and of the output gap. In such an environment, higher (lower) forecasts of future inflation and future output gap yield higher (lower) realizations. This kind of positive expectations feedback system favors coordination of individuals on trend following behavior. In such an environment, when a majority of individuals uses a trend-following strategy, other individuals have an incentive to use such strategy too, thus reinforcing trends in inflation and output gap. As shown in the experiments and simulations of the HSM model for $\phi_\pi = 1$ in the previous section, this is indeed what we observed in the experimental economies characterized by explosive behavior. On the other hand, when monetary policy is more aggressive, i.e., $\phi_\pi = 1.5$, there is negative feedback from inflation expectations to realizations of the output gap. In order to explain the stabilizing features of an aggressive monetary policy, we take as an example the experimental economy 2, treatment *b*. Fig. 11 plots the correlations among endogenous variables and their respective future expectations at different leads and lags. In this experimental session we observe an initial decreasing trend in both inflation and the output gap. For inflation, this leads to initial coordination on trend following behavior in inflation forecasts (see Fig. 4), resulting in reinforcement of the trend in actual inflation, which lasts until period 7, due to the positive feedback between inflation expectations and inflation realizations (see Fig. 11(f), positive correlations at leads 0 and 1). For the output gap, a reversal of the trend occurs in period 5, and this is due to the negative feedback between inflation expectations, which are increasingly revised downwards, and output gap realizations (see Fig. 11(c), negative correlations at lags 0, 1 and 2). The recovery and increase of the output gap eventually leads to upward revisions of future output gap expectations (see Fig. 11(b), positive correlations at leads 0 and 1) and this change in output gap expectations in turn has a positive feedback on realized inflation (see Fig. 11(e), positive correlations at leads 1 and 2), yielding a reversal of the downward trend in inflation. This feedback mechanism repeats and leads to oscillatory behavior in output and inflation. Whether these oscillations are stable or unstable depends on the eigenvalues of the macro system.

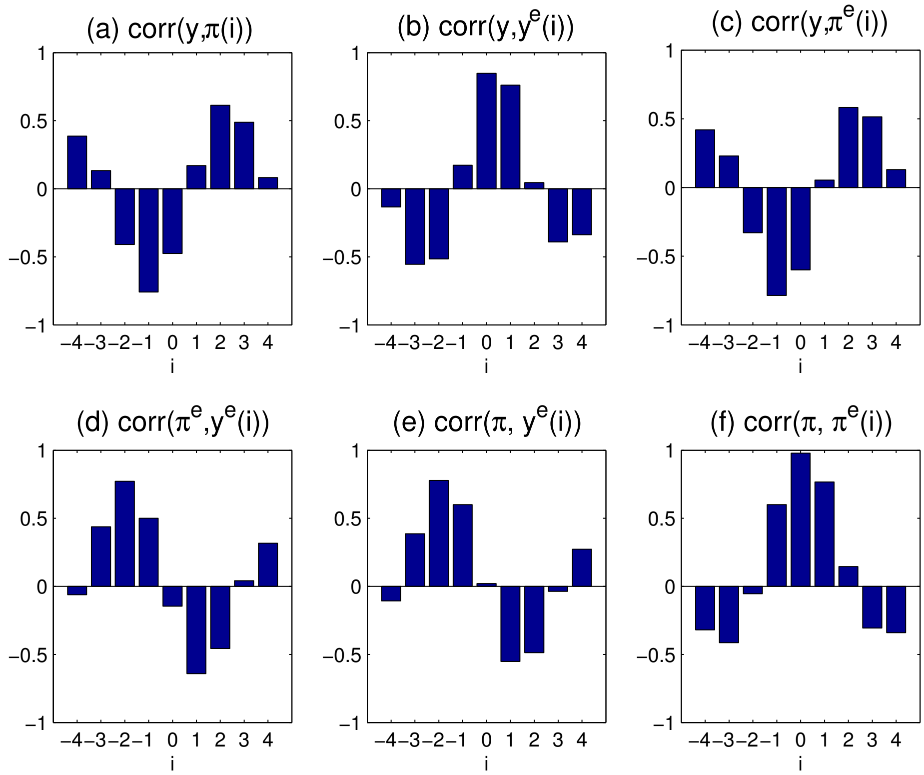


Figure 11: Correlation analysis, treatment b, group 2

Managing the positive feedback

To make this analysis more precise, we need to look at the eigenvalues of the overall macro system. Fig. 12 (left panel) plots the absolute values of the eigenvalues of system (2.4) as a function of the monetary policy parameter ϕ_π. As ϕ_π increases,

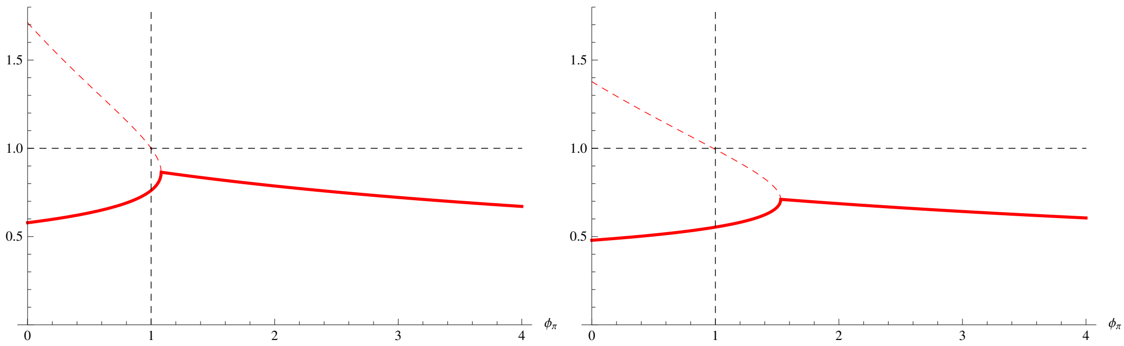


Figure 12: Absolute value of eigenvalues of system (2.4) as a function of ϕ_π. The left panel corresponds to the simple Taylor rule (2.3) that only targets inflation, while the right panel corresponds to the augmented Taylor rule (4.4) responding to both inflation and output gap. The augmented Taylor rule has smaller absolute eigenvalues and is therefore more effective in stabilizing the economy.

the absolute value of the largest eigenvalue decreases, so that a more aggressive Taylor rule weakens the positive feedback in the system. For $\phi_\pi < 1$ both eigenvalues are real, one inside and one outside the unit circle, so that the macro system is unstable and exploding (under naive expectations). Monetary policy is then too weak to prevent coordination of expectations on exploding trend-following behavior. For $\phi_\pi = 1$ the largest eigenvalue coincides with 1. The system therefore has a continuum of steady states, that is, a continuum of perfectly self-fulfilling RE equilibria. As our experiments show, in such an environment the weak Taylor rule does not prevent coordination of expectations on steady states different form the target or on exploding trend-following behavior. For $\phi_\pi > 1$ both eigenvalues are inside the unit circle and become complex for $\phi_\pi \geq 1.1$. It is important to note that, as long as the system is close to having an eigenvalue equal to 1, it has *almost self-fulfilling* equilibria. In such an environment coordination on almost self-fulfilling trend-following expectations may arise leading to persistent oscillations in output and inflation. For $\phi_\pi = 1.5$, as in treatments b and c of our experiment, the largest absolute eigenvalue has reduced to approximately 0.85. Our experiment shows that even for $\phi_\pi = 1.5$ almost self-fulfilling equilibria arise in the form of dampened oscillations and, in the case of an inflation target $\bar{\pi} = 3.5\%$, also persistent oscillations (see Fig. 5, treatment c, groups 2 and 6). In general, in order to stabilize a macro system monetary policy has to be strong enough to push the eigenvalues of the macro system well within the unit circle to weaken the positive feedback, such that coordination on trend-following behavior is not sustainable because destabilizing trend-following strategies are driven out by stabilizing adaptive expectations in the self-organization process of evolutionary competition between heterogeneous forecasting rules.

We now discuss another testable prediction of our behavioral model using a different Taylor rule for monetary policy. It has been argued that in order to stabilize the economy, the CB should not only target inflation, but also the output gap, through an augmented Taylor rule

$$i_t = \bar{\pi} + \phi_\pi(\pi_t - \bar{\pi}) + \phi_y(y_t - \bar{y}). \tag{4.4}$$

Frequently used policy parameters are $\phi_\pi = 1.5$ and $\phi_y = 0.5$. Fig. 12 (right panel) plots the absolute values of the eigenvalues of system (2.4) as a function of the monetary policy parameter ϕ_π for the augmented Taylor rule (4.4). The augmented Taylor rule adds more negative feedback to the macro system and therefore overall weakens the positive feedback more. For example, compared to the simple Taylor rule (i.e. $\phi_y = 0$), for $\phi_\pi = 1.5$ the largest absolute eigenvalue drops from 0.85 to 0.75 for the augmented Taylor rule (with $\phi_y = 0.5$).[21] Our model thus predicts that the augmented Taylor rule is more effective in managing heterogeneous expectations.

Managing coordination on trend-following behavior

We take the analysis of how the CB can manage the self-organization process of heterogeneous expectations and prevent coordination on destabilizing trend-following expectations one step further. The stability analysis in Figure 12 is valid only under the assumption of homogeneous naive expectations. This gives useful insights into how an increase of the policy parameter ϕ_π affects the eigenvalues and stability of the macro system and the strength of the positive feedback. Our experimental results however show that instability and fluctuations in inflation and output gap arise due to coordination on trend-following strategies of the form

$$x^e_{t+1} = x_{t-1} + g(x_{t-1} - x_{t-2}). \tag{4.5}$$

In the 4-type HSM fitted to the experimental data, both the WTR and STR are exactly of this form, with $g = 0.4$ resp. $g = 1.3$, and the LAA rule is also a trend-following rule with $g = 1$ (with a more flexible anchor). In order to stabilize the economy, the CB should prevent coordination on destabilizing trend-following rules. How large then should the policy parameter ϕ_π be to avoid almost self-fulfilling coordination on the trend-following rule (4.5)?

Figure 13 plots the absolute value of the eigenvalues of the macro system under a homogeneous trend-following rule (4.5) for both inflation and output gap

[21] A straightforward computation shows that the absolute eigenvalues are < 1 for $\phi_\pi = 1 - (1 - \rho)\phi_y/\lambda \approx 0.9833$.

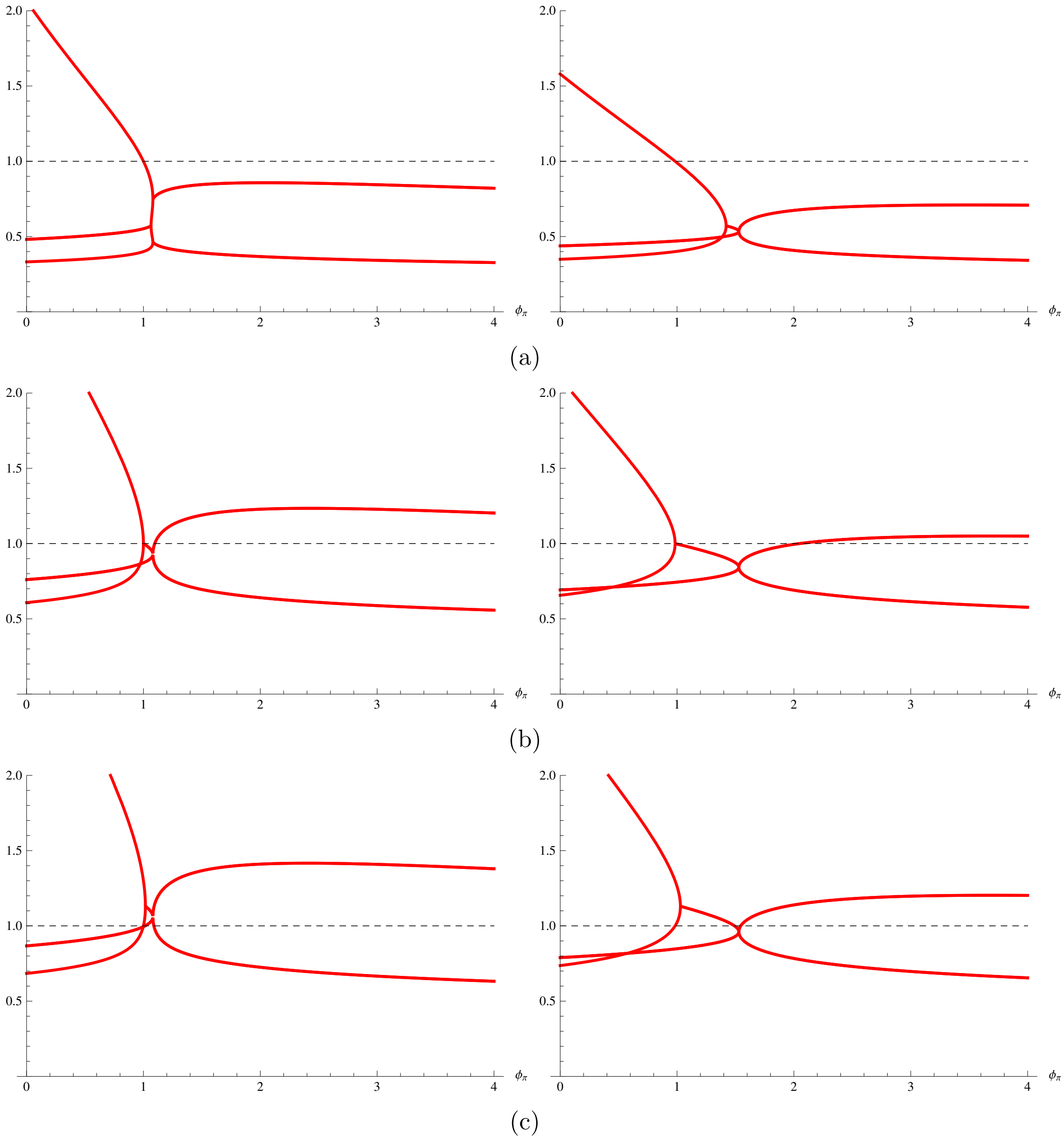


Figure 13: Absolute values of the eigenvalues as a function of the policy parameter ϕ_π, when all subjects coordinate on the trend-following (4.5), with $g = 0.4$ (top panels), $g = 1$ (middle panels) and $g = 1.3$ (top panels). Left panels correspond to the simple Taylor rule (2.3) that only targets inflation, while right panels illustrate the augmented Taylor rule (4.4) responding to both inflation and output gap. The augmented Taylor rule is more effective in preventing coordination on trend-following behavior and stabilizing the economy.

expectations.[22] The left panels correspond to the simple Taylor rule (2.3), and the right panels to the augmented Taylor rule (4.4). Two important observations can be made: (1) the dependence of the largest absolute eigenvalue on the policy parameter ϕ_π is *non-monotonic*, and (2) the augmented Taylor is more effective in stabilizing trend-following behavior, as the largest absolute eigenvalue decreases.

Consider the simple Taylor rule (left panels). For the weak trend-following rule ($g = 0.4$; top left panel) the system is locally stable for all $\phi_\pi > 1$. Hence, weak trend-followers can not destabilize the macrosystem. For intermediate trend-followers ($g = 1$) the system is only (locally) stable for a very small interval $\phi_\pi \in [1, 1.15]$, while for strong trend-followers ($g = 1.3$), the system is unstable for all $\phi_\pi \in [0, 4]$. Hence, even when the Taylor principle holds, i.e. $\phi_\pi = 1.5 > 1$, coordination of individual expectations on intermediate trend-followers ($g = 1$) can destabilize the system. This seems to be what is happening in treatment c, group 2 (see Fig. 10), where the fraction of the intermediate trend-following LAA rule steadily rises until $85 - 90\%$ around period 40 for both inflation and output forecasting.

For the augmented Taylor rule the absolute value of the largest eigenvalue becomes smaller (compare the left to the right panels in Figure 13). Hence, targeting both inflation and output has a larger stabilizing effect upon the economy. In particular, for a range of parameter values, $1 < \phi_\pi < 2$, coordination on intermediate trend-following behavior ($g = 1$; middle panel) can *not* destabilize the economy.[23] Our analysis thus predicts that the augmented Taylor rule (4.4) does a much better job in managing the self-organization process of heterogeneous expectations and preventing coordination on destabilizing trend-following behavior.

5 Concluding Remarks

In this paper we use laboratory experiments with human subjects to study individual expectations, aggregate macro behavior and the role of monetary policy within a standard New Keynesian framework. We ran three different experimental treatments depending on the monetary policy rule, with a weak respectively strong Taylor interest rate rule, and different inflation targets. Under a weak Taylor rule, the experimental macro economy is unstable, with output and inflation either converging to arbitrary steady state levels different from target or exhibiting exploding inflationary or deflationary spirals. Under a strong Taylor rule, that responds more than one-to-one to inflation, the macro economy is more stable and is more likely

[22]The system is derived in Appendix A.

[23]The optimal policy parameter value minimizing the absolute value of the largest eigenvalue is $\phi_\pi \approx 1.6$.

to converge to the RE stable inflation target level, but inflation and output may also fluctuate persistently over time.

A behavioral heterogeneous expectations switching model explains both individual forecasting behavior and aggregate macro outcomes observed in the laboratory experiments. The heuristics switching model provides a simple and intuitive explanation of how the different macro patterns emerge out of a self-organization process of heterogeneous expectations driven by their relative forecasting performance. The self-organization process leads to coordination on a common rule, but exhibits path-dependence and four different aggregate outcomes can emerge. Convergence to some equilibrium level is explained by coordination on adaptive expectations, inflationary or deflationary spirals arise due to coordination on strongly extrapolating trend-following rules, persistent oscillations arise after coordination on an anchor and adjustment trend-following rule and, finally, dampened oscillations and convergence arise when initially dominating (weak) trend-following rules are at the end driven out by adaptive expectations.

In addition to a theoretical foundation of the discrete choice model with asynchronous updating in Eq. (3.4), it is worthwhile to discuss an intuitive behavioural interpretation of the HSM model. How would boundedly rational subjects in the experiment be able to compute the fitness measure of the four forecasting strategies? The four forecasting rules –adaptive, weak trend, strong trend or anchor and adjustment– predict different types of behaviour. Adaptive expectations predicts a weighted average between the last observation and the last forecast, the trend-following rules predict weak respectively strong extrapolation of the last observed price change from the last observed price, while the anchor and adjustment rule extrapolates from an anchor giving more weight to average prices, implying a mean-reverting prediction when the last observed price has moved far away from the average price level. Even without making exact computations, based on the graphical representations of the time series of inflation and/or output gap and their individual forecasts, subjects should be able to decide approximately which of the four rules –adaptive, weak trend, strong trend or anchor and adjustment– performed better in the last period. The HSM describing gradual updating of the four strategies based on their recent performance may be seen as a quantitative model of this intuitive reasoning of boundedly rational subjects. Since this intuitive reasoning described by (3.4) is relatively simple, a population of subjects may coordinate on such intuitive and (almost) self-fulfilling reasoning as observed in the experiments.

Our experiments and HSM explain behaviour in a simple lab environment. An important finding is that, even in such a simple environment, subjects may coordinate on non-rational almost self-fulfilling equilibria with temporary or permanent fluctuations or even exploding inflationary or deflationary spirals. In the real economy of course professional forecasters may use much more complicated forecasting models, e.g., advanced VAR models. How then should one model a macro economy with advanced professional forecasters? Clearly one can not include all these complicated forecasting methods into a simple, tractable model and one has to rely on parsimonious, relatively simple forecasting models, if one wants to model aggregate behaviour including professional forecasters. Branch (2004) analyzed survey data of household's expectations of inflation and estimates a Brock-Hommes switching model similar to ours, with three rules, naive, adaptive expectations and a VAR model. All three models appear to be relevant and their fractions fluctuate considerably over time. Furthermore, Stock and Watson (2007) have argued that simple univariate forecasting rules may outperform advanced VAR models in forecasting. Cornea, Hommes, and Massaro (2012) recently estimated a simple New Keynesian HSM where agents choose between a naive expectations rule (consistent with a random walk belief of inflation) and a VAR-model (including the rate of change of unit labor costs and of labor share) and show that the 2-type HSM nicely fits quarterly U.S. inflation data 1960:Q1–2010:Q4. Some parsimony of forecasting models is important in modeling boundedly rational behavior of consumers, firms and professional forecasters. Our HSM may be viewed as a first approximation of boundedly rational forecasting behaviour and can be easily extended with more sophisticated forecasting rules such as VAR models.

How can monetary policy manage the self-organization process of heterogeneous expectations, prevent coordination on destabilizing expectations and enforce a stable outcome? In our experimental economy, the implementation of a monetary policy that reacts aggressively to deviations of inflation from its target leads the economy to the desired stable outcome, at least in the long run. Our behavioral model provides a simple explanation of why a more aggressive monetary policy is stabilizing. Under a weak Taylor interest rate rule, the NK macro framework exhibits strong positive expectations feedback. In such an environment individual expectations may easily coordinate on a (strong) trend-following rule, destabilizing the economy into inflationary or deflationary spirals. A more aggressive Taylor rule adds negative feedback to the macro system, as a higher interest rate offsets high inflation expectations and decreases output. In such a macro environment the overall positive feedback is weaker and coordination on trend-following strategies is less likely, because the weaker positive feedback makes them perform relatively worse. Coordination on adaptive expectations is more likely in a macro environment with only weak positive feedback, stabilizing the economy. Our behavioral model predicts that to manage the self-organization process of heterogeneous expectations an augmented Taylor rule, that responds to both inflation and output gap, is the most effective, because it adds even more negative feedback to the macro system and makes coordination on even weakly trend-following rules less likely.

Our analysis uses the standard NK framework as a starting point. Future work should apply our behavioral HSM in an extended NK framework, e.g. by taking asset prices and financial frictions into account, and to study the relations between positive feedback and the effectiveness of monetary and fiscal policies in managing heterogeneous expectations and prevent coordination on destabilizing expectations.

References

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

Anufriev, M., T. Assenza, C. H. Hommes, and D. Massaro (2013): "Interest Rate Rules and Macroeconomic Stability under Heterogeneous Expectations," *Macroeconomic Dynamics*, 17, 1574–1604.

Anufriev, M., and C. H. Hommes (2012): "Evolutionary Selection of Individual Expectations and Aggregate Outcomes," *American Economic Journal: Microeconomics*, 4, 35–64.

Bao, M., C. Hommes, J. Sonnemans, and J. Tuinstra (2012): "Individual Expectations, Limited Rationality and Aggregate Outcomes," *Journal of Economic Dynamics and Control*, 36, 1101–1120.

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 G. Evans (2006): "Intrinsic Heterogeneity in Expectation Formation," *Journal of Economic Theory*, 127, 264–295.

Branch, W., and B. McGough (2009): "A New Keynesian Model with Heterogeneous Expectations," *Journal of Economic Dynamics and Control*, 33, 1036–1051.

——— (2010): "Dynamic Predictors Selection in a New Keynesian Model with Heterogeneous Expectations," *Journal of Economic Dynamics and Control*, 34(8), 1492–1508.

Brock, W. A., and C. H. Hommes (1997): "A Rational Route to Randomness," *Econometrica*, 65(5), 1059–1095.

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

Capistran, C., and A. Timmermann (2009): "Disagreement and biases in inflation expectations," *Journal of Money, Credit and Banking*, 41, 365–396.

Chari, V., P. Kehoe, and E. McGrattan (2000): "Sticky price models of the business cycle: Can the contract multiplier solve the persistence problem?," *Econometrica*, 68, 1151–1179.

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

Christiano, L., R. Motto, and M. Rostagno (2010): "Financial Factors in Economic Fluctuations," Working Paper Series 2010-1192, European Central Bank.

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

Cornea, A., C. Hommes, and D. Massaro (2012): "Behavioural Heterogeneity in U.S. Inflation Dynamics," CeNDEF Working Paper 2012-3, University of Amsterdam.

Curdia, V., and M. Woodford (2009): "Credit Frictions and Optimal Monetary Policy," BIS Working Papers 278, Bank for International Settlements.

——— (2010): "Credit Spreads and Monetary Policy," *Journal of Money, Credit and Banking*, 42(s1), 3–35.

De Grauwe, P. (2010): "Animal Spirits and Monetary Policy," *Economic Theory*.

Duffy, J. (2008): "Experimental Macroeconomics," in *New Palgrave Dictionary of Economics*, ed. by S. Durlauf, and L. Blume. New York: Palgrave Macmillan, forthcoming.

Evans, G., and S. Honkapohja (1998): "Economic Dynamics With Learning: New Stability Results," *Review of Economic Studies*, 65, 23–44.

——— (2003): "Expectations and the Stability Problem for Optimal Monetary Policy," *Review of Economic Studies*, 70, 807–824.

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

Fehr, E., and J.-R. Tyran (2008): "Limited rationality and strategic interaction: the impact of the strategic environment on nominal inertia," *Econometrica*, 76(2), 353–394.

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

Gali, J., and M. Gertler (1999): "Inflation Dynamics: A Structural Econometric Analysis," *Journal of Monetary Economics*, 44, 195–222.

Gertler, M., and P. Karadi (2011): "A Model of Unconventional Monetary Policy," *Journal of Monetary Economics*, 58(1), 17–34.

——— (2013): "QE 1 vs. 2 vs. 3. . . : A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool," *International Journal of Central Banking*, 9(1), 5–53.

Gilchrist, S., A. Ortiz, and E. Zakrajsek (2009): "Credit Risk and the Macroeconomy: Evidence from an Estimated DSGE Model," mimeo.

Heemeijer, P., C. H. Hommes, J. Sonnemans, and J. Tuinstra (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.

——— (2013): "Reflexivity, expectations feedback and almost self-fulfilling equilibria: economic theory, empirical evidence and laboratory experiments," *Journal of Economic Methodology*, 20(4), 406–419.

Hommes, C. H., J. Sonnemans, J. Tuinstra, and H. van de Velden (2005): "Coordination of Expectations in Asset Pricing Experiments," *Review of Financial Studies*, 18(3), 955–980.

Hommes, C. H., and M. Zhu (2014): "Behavioral Learning Equilibria," *Journal of Economic Theory*, 150, 778–814.

Kryvtsov, O., and L. Petersen (2013): "Expectations and Monetary Policy: Experimental Evidence," Working Papers 13-44, Bank of Canada.

Kurz, M. (2011): "A New Keynesian Model with Diverse Beliefs," Discussion paper, Stanford University.

Kurz, M., G. Piccillo, and H. Wu (2013): "Modeling Diverse Expectations In An Aggregated New Keynesian Model," *Journal of Economic Dynamics and Control*, 37, 1403–1433.

Lucas, R. E. (1972): "Expectations and the Neutrality of Money," *Journal of Economic Theory*, 4, 103–124.

Mankiw, N., R. Reis, and J. Wolfers (2003): "Disagreement about inflation expectations," in *NBER Macroeconomics Annual*, ed. by M. Gertler, and K. Rogoff.

Marimon, R., S. E. Spear, and S. Sunder (1993): "Expectationally Driven Market Volatility: An Experimental Study," *Journal of Economic Theory*, 61(1), 74–103.

Marimon, R., and S. Sunder (1994): "Expectations and Learning Under Alternative Monetary Regimes: An Experimental Study," *Economic Theory*, 4, 131–162.

Massaro, D. (2012): "Bounded Rationality and Heterogeneous Expectations in Macroeconomics," Phd thesis, University of Amsterdam.

——— (2013): "Heterogeneous Expectations in Monetary DSGE Models," *Journal of Economic Dynamics and Control*, 37(3), 680–692.

Milani, F. (2007): "Expectations, Learning and Macroeconomic Persistence," *Journal of Monetary Economics*, 54, 2065–2082.

Muth, J. F. (1961): "Rational Expectations and the Theory of Price Movements," *Econometrica*, 29(3), 315–335.

Nagel, R. (1995): "Unraveling in Guessing Games: an Experimental Study," *American Economic Review*, 85, 1313–1326.

Nelson, E. (1998): "Sluggish inflation and optimization models of the business cycle," *Journal of Monetary Economics*, 42, 303–322.

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

Pfajfar, D., and B. Zakelj (2012): "Inflation Expectations and Monetary Policy Design: Evidence from the Laboratory," mimeo, Tilburg University.

Preston, B. (2005): "Learning about Monetary Policy Rules when Long-Horizon Forecasts Matter," *International Journal of Central Banking*, 1(2), 81–126.

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

Schotter, A., and I. Trevino (2007): "Belief Elicitation in the Lab," *Annual Review of Economics*, 6, forthcoming.

Smets, F., and R. Wouters (2007): "Shocks and Frictions in US Business Cycle: a Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–607.

Stock, J. H., and M. W. Watson (2007): "Why Has U.S. Inflation Become Harder to Forecast?," *Journal of Money, Credit and Banking*, 39, 3–33.

Tversky, A., and D. Kahneman (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.

Appendix For Online Publication

A Derivation of NK model under trend following behavior

Let us rewrite system (2.4)

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

in the more compact form

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = \mathbf{A} + \mathbf{B} \begin{bmatrix} \bar{y}^e_{t+1} \\ \bar{\pi}^e_{t+1} \end{bmatrix} + \text{noise} . \tag{A.1}$$

If we assume homogeneous trend following behavior of the form

$$x^e_{t+1} = x_{t-1} + g(x_{t-1} - x_{t-2})$$

with $x \in \{y, \pi\}$, we can rewrite system (A.1) as

$$\begin{bmatrix} y_t \\ \pi_t \\ z_t \\ w_t \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{21} \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} b_{11} + b_{11}g & b_{12} + b_{12}g & -b_{11}g & -b_{12}g \\ b_{21} + b_{21}g & b_{22} + b_{22}g & -b_{21}g & -b_{22}g \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} \\ \pi_{t-1} \\ z_{t-1} \\ w_{t-1} \end{bmatrix} + \text{noise} , \tag{A.2}$$

where $z_t = y_{t-1}, w_t = \pi_{t-1}$ and $a_{i,j}$, b_{ij} are respectively entries of the matrices $\mathbf{A}$ and $\mathbf{B}$.

B (Translation of Dutch) Instructions for participants (inflation forecasters)

Set-up of the experiment

You are participating in an experiment on economic decision-making. You will be rewarded based on the decisions you make during the experiment. The experiment will be preceded by several pages of instructions that will explain how it works. When the experiment has ended, you will be asked to answer some questions about how it went.

- The whole experiment, including the instructions and the questionnaire, is computerized. Therefore you do not have to submit the paper on your desk. Instead, you can use it to make notes.
- There is a calculator on your desk. If necessary, you can use it during the experiment.
- If you have any question during the experiment, please raise your hand, then someone will come to assist you.

General information about the experiment

In the experiment, *statistical research bureaus* make predictions about the *inflation* and the so-called *"output gap"* in the economy. A limited amount of research bureaus is active in the economy. You are a research bureau that makes predictions about *inflation*. This experiment consists of *50 periods* in total. In each period you will be asked to predict the inflation; your *reward after the experiment has ended* is based on the accuracy of your predictions.

In the following instructions you will get more information about the *economy* you are in, about the way in which *making predictions* works during the experiment, and about the way in which your *reward* is calculated. Also, the *computer program* used during the experiment will be explained.

Information about the economy (part 1 of 2)

The economy you are participating in is described by *three variables*: the inflation π_t, the output gap y_t and the 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 through 50.

The *inflation* measures the percentage change in the price level of the economy. In each period, inflation depends on the *inflation predictions* and *output gap predictions* of the statistical research bureaus, and on *minor price shocks*. There is a *positive relation* between the actual inflation and both the inflation predictions and output gap predictions of the research bureaus. This means for example that if the inflation prediction of a research bureaus increases, then actual inflation will also increase (assuming that the other predictions and the price shock remain equal). The minor price shocks have an equal chance of influencing inflation positively or negatively.

Information about the economy (part 2 of 2)

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 would be 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 the *inflation predictions* and *output gap predictions* of the statistical bureaus, on the *interest rate* and on *minor economic shocks*. There is a *positive relation* between the output gap and the inflation predictions and output gap predictions, and a *negative relation* between the output gap and the interest rate. The minor economic shocks have an equal chance of influencing the output gap positively or negatively.

The *interest rate* measures the price of borrowing money and is determined by the *central bank*. There is a *positive relation* between the interest rate and the inflation.

Information about making predictions

Your task, in each period of the experiment, consists in predicting the *inflation in the next period*. Inflation has been historically between -5% and 15%. When the experiment starts, you have to predict the inflation for the first two periods, i.e. π_1^e and π_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 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 an inflation prediction for period 3 (π_3^e). When all participants have made their predictions for period 3, the inflation (π_2), the output gap (y_2) and the interest rate (i_2) for period 2 are announced. This process repeats for 50 periods. Therefore, when at a certain period t you make a prediction of the inflation in period $t+1$ (π_{t+1}^e), the following information is available:

- Values of the actual *inflation, output gap and interest rate* up to and including period $t-1$;
- Your *predictions* up to and including period t;
- Your *prediction scores* up to and including period $t-1$.

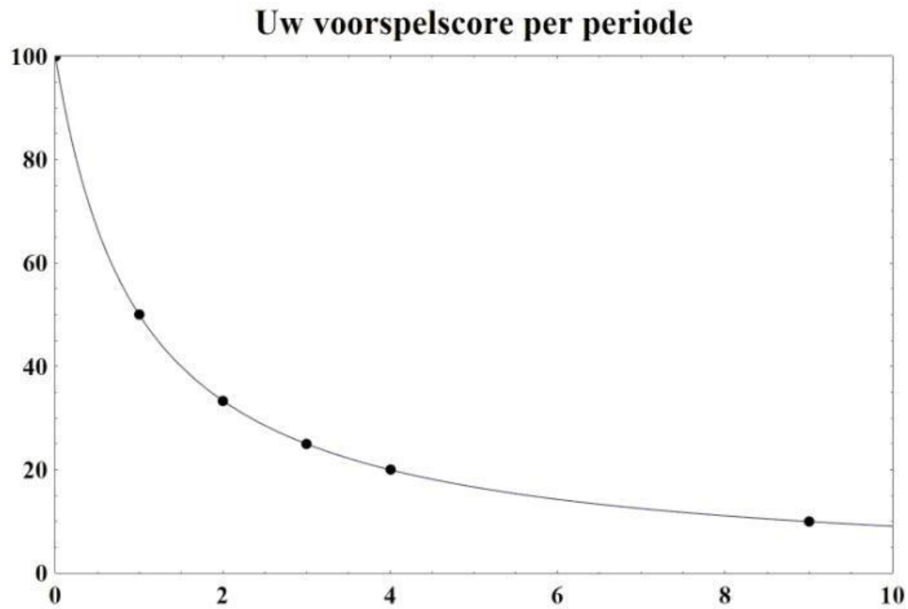
Information about your reward (part 1 of 2)

Your *reward after the experiment has ended* increases with the accuracy of your predictions. Your accuracy is measured by the *absolute error* between your inflation predictions and the true inflation. For each period this absolute error is calculated as soon as the true value of inflation is known; you subsequently get a *prediction score* that decreases as the absolute error increases. The *table below* gives the relation between the absolute predictions error and the prediction score. If at a certain period you predict for example an inflation of 2%, and the true 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 realized inflation turns out to be −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.

Absolute prediction error	0	1	2	3	4	9
Score	100	50	33/3	25	20	10

Information about your reward (part 2 of 2)

The figure below shows the relation between your *prediction score* (vertical axis) and your *prediction error* (horizontal axis). Notice that your prediction score *decreases more slowly* as your prediction error increases. Points in the graph correspond to the prediction scores in the previous table.

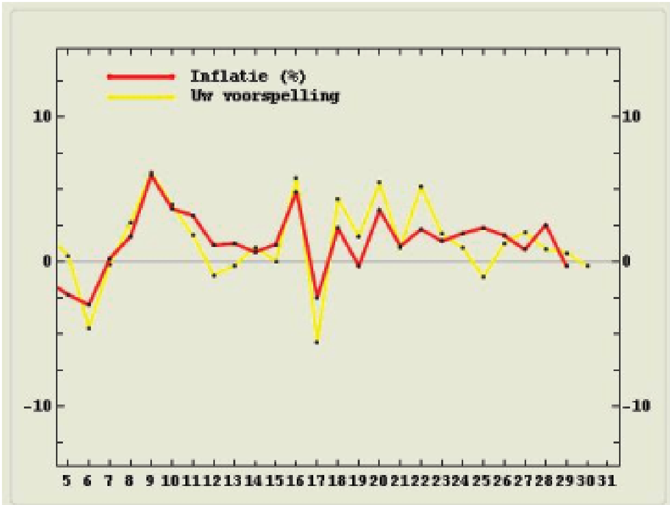


Your *total score* at the end of the experiment consists simply of the sum of all prediction scores you got during the experiment. During the experiment, your scores are shown on your computer screen. When the experiment has ended, you are shown an overview of your prediction scores, followed by the resulting total score. Your *final reward* consists of

0.75 *euro-cent* for each point in your total score (200 points therefore equals 1.50 euro). Additionally, you will receive a *show up fee* of 5 euro.

Information about the computer program (part 1 of 3)

Below you see an example of the *left upper part* of the computer screen during the experiment. It consists of a *graphical representation* of the inflation (red series) and your predictions of it (yellow series). On the horizontal axis are the *time periods*; the vertical axis is in percentages. In the *imaginary situation* depicted in the graph, the experiment is in period 30 and you predict the inflation in period 31 (the experiment lasts for 50 periods). Notice that the graph only shows results of *at most the last 25 periods* and that the *next period* is always on the right hand side.



The *left bottom part* of the computer screen also contains a graph. In this graph the *output gap* and the *interest rate* are shown in the same way as in the above graph.

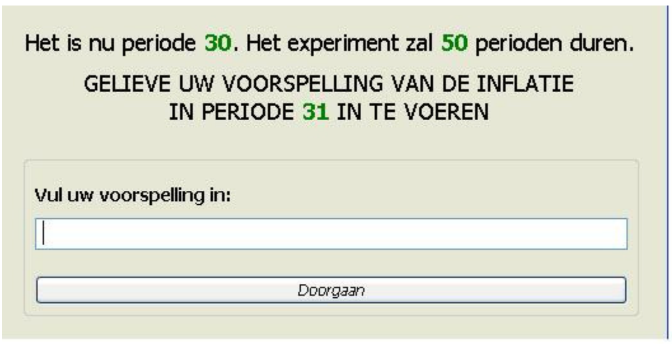
Information about the computer program (part 2 of 3)

Below you see an example of *the right upper part* of the computer screen during the experiment. It consists of a *table* containing information about the results of the experiment in at most the last 25 periods. This information is supplemental to the graphs in the left part of the screen. The *first column* of the table shows the time period (the next period, 31 in the example, is always at the top). The *second and third columns* respectively show the inflation and your predictions of it. The *fourth column* gives the output gap and the *fifth column* the interest rate. Finally, the *sixth column* gives your prediction score for each period separately. Notice that you can use the sheet of paper on your desk to save data longer than 25 periods.

Tijds-periode	Inflatie (%)	Uw voor-spelling	Output gap (%)	Rente (%)	Voorspel-score
31		???			
30		-0.32			
29	-0.32	0.63	2.15	1.42	56
28	2.54	0.88	2.09	2.13	44
27	0.83	2.04	-7.31	1.71	75
26	1.81	1.28	1.08	1.95	60
25	2.34	-1.09	-2.12	2.09	24
24	1.95	0.99	-1.33	1.99	50
23	1.41	1.9	4.1	1.85	96
22	2.24	5.21	1.39	2.06	24
21	1.07	0.96	-3.92	1.77	55
20	3.56	5.44	5.07	2.39	25
19	-0.3	1.7	-1.89	1.42	78
18	2.29	4.33	1.31	2.07	31
17	-2.49	-5.55	-0.85	0.88	13
16	4.83	5.82	-0.76	2.71	24
15	1.2	0.04	-0.94	1.8	36
14	0.66	0.99	-1.24	1.67	60
13	1.27	-0.29	-0.43	1.82	32
12	1.13	-0.91	-0.18	1.78	27
11	3.21	1.84	3.96	2.3	68
10	3.65	3.91	-1.24	2.41	40
9	5.97	6.13	-4.45	2.99	24
8	1.74	2.71	1.58	1.93	56
7	0.23	-0.14	-3.86	1.56	37
6	-2.98	-4.58	-0.07	0.75	16
5	-2.28	0.43	2.37	0.93	67

Information about the computer program (part 3 of 3)

Below you see an example of the *bottom part* of the computer screen during the experiment. In each period you are asked to submit your inflation prediction in the next period (below *Submit you prediction*). When submitting your prediction, use the *decimal point* if necessary. For example, if you want to submit a prediction of 2.5%, type "2.5"; for a prediction of −1.75%, type " − 1.75". Notice that your predictions and the true inflation in the experiment are rounded to two decimals. Moreover, prediction scores are rounded to integers.



(Translation of Dutch) Instructions for participants (output gap forecasters)

Set-up of the experiment

You are participating in an experiment on economic decision-making. You will be rewarded based on the decisions you make during the experiment. The experiment will be preceded by several pages of instructions that will explain how it works. When the experiment has ended, you will be asked to answer some questions about how it went.

- The whole experiment, including the instructions and the questionnaire, is computerized. Therefore you do not have to submit the paper on your desk. Instead, you can use it to make notes.
- There is a calculator on your desk. If necessary, you can use it during the experiment.
- If you have any question during the experiment, please raise your hand, then someone will come to assist you.

General information about the experiment

In the experiment, *statistical research bureaus* make predictions about the *inflation* and the so-called *"output gap"* in the economy. A limited amount of research bureaus is active in the economy. You are a research bureau that makes predictions about *inflation.* This experiment consists of *50 periods* in total. In each period you will be asked to predict the inflation; your *reward after the experiment has ended* is based on the accuracy of your predictions.

In the following instructions you will get more information about the *economy* you are in, about the way in which *making predictions* works during the experiment, and about the way in which your *reward* is calculated. Also, the *computer program* used during the experiment will be explained.

Information about the economy (part 1 of 2)

The economy you are participating in is described by *three variables*: the inflation π_t, the output gap y_t and the 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 through 50.

The *inflation* measures the percentage change in the price level of the economy. In each period, inflation depends on the *inflation predictions* and *output gap predictions* of the statistical research bureaus, and on *minor price shocks.* There is a *positive relation* between the actual inflation and both the inflation predictions and output gap predictions of the research bureaus. This means for example that if the inflation prediction of a research bureaus increases, then actual inflation will also increase (assuming that the other predictions and the price shock remain equal). The minor price shocks have an equal chance of influencing inflation positively or negatively.

Information about the economy (part 2 of 2)

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 would be 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 the *inflation predictions* and *output gap predictions* of the statistical bureaus, on the *interest rate* and on *minor economic shocks*. There is a *positive relation* between the output gap and the inflation predictions and output gap predictions, and a *negative relation* between the output gap and the interest rate. The minor economic shocks have an equal chance of influencing the output gap positively or negatively.

The *interest rate* measures the price of borrowing money and is determined by the *central bank*. There is a *positive relation* between the interest rate and the inflation.

Information about making predictions

Your task, in each period of the experiment, consists in predicting the *output gap in the next period*. Inflation has been historically between -5% and 5%. When the experiment starts, you have to predict the output gap for the first two periods, i.e. 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 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 an output gap prediction for period 3 (y_3^e). When all participants have made their predictions for period 3, the inflation (π_2), the output gap (y_2) and the interest rate (i_2) for period 2 are announced. This process repeats for 50 periods. Therefore, when at a certain period t you make a prediction of the inflation in period $t+1$ (y_{t+1}^e), the following information is available:

- Values of the actual *inflation, output gap and interest rate* up to and including period $t-1$;
- Your *predictions* up to and including period t;
- Your *prediction scores* up to and including period $t-1$.

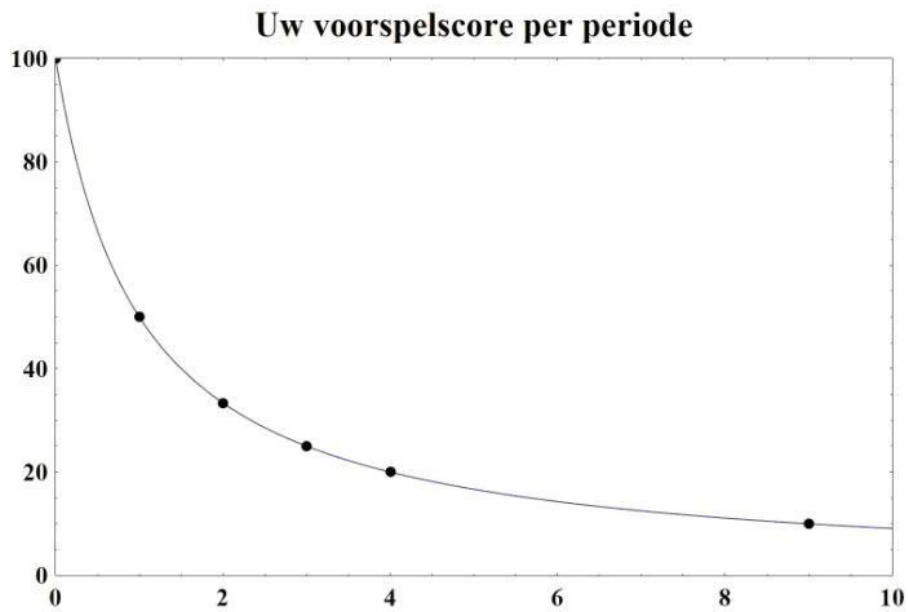
Information about your reward (part 1 of 2)

Your *reward after the experiment has ended* increases with the accuracy of your predictions. Your accuracy is measured by the *absolute error* between your inflation predictions and the true inflation. For each period this absolute error is calculated as soon as the true value of inflation is known; you subsequently get a *prediction score* that decreases as the absolute error increases. The *table below* gives the relation between the absolute predictions error and the prediction score. If at a certain period you predict for example an inflation of 2%, and the true 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 realized inflation turns out to be -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.

Absolute prediction error	0	1	2	3	4	9
Score	100	50	33/3	25	20	10

Information about your reward (part 2 of 2)

The figure below shows the relation between your *prediction score* (vertical axis) and your *prediction error* (horizontal axis). Notice that your prediction score *decreases more slowly* as your prediction error increases. Points in the graph correspond to the prediction scores in the previous table.

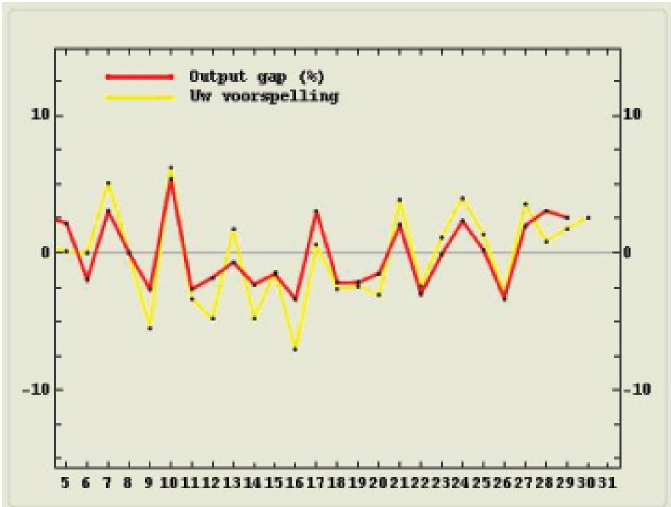


Your *total score* at the end of the experiment consists simply of the sum of all prediction scores you got during the experiment. During the experiment, your scores are shown on your computer screen. When the experiment has ended, you are shown an overview of your prediction scores, followed by the resulting total score. Your *final reward* consists of

0.75 *euro-cent* for each point in your total score (200 points therefore equals 1.50 euro). Additionally, you will receive a *show up fee* of 5 euro.

Information about the computer program (part 1 of 3)

Below you see an example of the *left upper part* of the computer screen during the experiment. It consists of a *graphical representation* of the output gap (red series) and your predictions of it (yellow series). On the horizontal axis are the *time periods*; the vertical axis is in percentages. In the *imaginary situation* depicted in the graph, the experiment is in period 30 and you predict the output gap in period 31 (the experiment lasts for 50 periods). Notice that the graph only shows results of *at most the last 25 periods* and that the *next period* is always on the right hand side.



The *left bottom part* of the computer screen also contains a graph. In this graph the *inflation* and the *interest rate* are shown in the same way as in the above graph.

Information about the computer program (part 2 of 3)

Below you see an example of *the right upper part* of the computer screen during the experiment. It consists of a *table* containing information about the results of the experiment in at most the last 25 periods. This information is supplemental to the graphs in the left part of the screen. The *first column* of the table shows the time period (the next period, 31 in the example, is always at the top). The *second and third columns* respectively show the output gap and your predictions of it. The *fourth column* gives the inflation and the *fifth column* the interest rate. Finally, the *sixth column* gives your prediction score for each period separately. Notice that you can use the sheet of paper on your desk to save data longer than 25 periods.

Tijds-periode	Output gap (%)	Uw Voor-spelling	Inflatie (%)	Rente (%)	Voorspel-score
31		???			
30		2.59			
29	2.59	1.78	3.43	2.36	55
28	3.08	0.85	3.44	2.36	31
27	1.98	3.6	2.42	2.11	38
26	-3.3	-2.98	1.38	1.84	76
25	0.24	1.37	4.47	2.62	47
24	2.31	4.02	2.15	2.04	37
23	-0.11	1.15	-0.92	1.27	44
22	-2.99	-2.43	1.	1.75	64
21	2.02	3.87	1.86	1.97	35
20	-1.48	-3.07	0.35	1.59	39
19	-2.14	-2.39	0.48	1.62	80
18	-2.18	-2.59	1.42	1.85	71
17	3.06	0.58	4.01	2.5	29
16	-3.39	-7.06	2.04	2.01	21
15	-1.53	-1.43	4.51	2.63	91
14	-2.28	-4.8	1.44	1.86	28
13	-0.67	1.74	2.78	2.2	29
12	-1.79	-4.8	1.64	1.91	25
11	-2.64	-3.34	1.78	1.95	59
10	5.41	6.25	2.02	2.01	54
9	-2.67	-5.48	1.81	1.95	26
8	0.02	-0.02	0.63	1.66	96
7	3.03	5.15	2.71	2.18	32
6	-1.91	0	-0.7	1.32	34
5	2.17	0.1	-0.76	1.31	32

Information about the computer program (part 3 of 3)

Below you see an example of the *bottom part* of the computer screen during the experiment. In each period you are asked to submit your output gap prediction in the next period (below *Submit you prediction*). When submitting your prediction, use the *decimal point* if necessary. For example, if you want to submit a prediction of 2.5%, type "2.5"; for a prediction of -1.75%, type " -1.75". Notice that your predictions and the true inflation in the experiment are rounded to two decimals. Moreover, prediction scores are rounded to integers.