

CentER 

Discussion Paper

No. 2011-091

**INFLATION EXPECTATIONS AND MONETARY POLICY  
DESIGN: EVIDENCE FROM THE LABORATORY**

By Damjan Pfajfar, Blaž Žakelj

July 17, 2011

This is a revised version of CentER Discussion Papers

No. 2009-007  
January 16, 2009

ISSN 0924-7815

# Inflation Expectations and Monetary Policy Design: Evidence from the Laboratory\*

Damjan Pfajfar<sup>†</sup>

Blaž Žakelj<sup>‡</sup>

CentER, EBC, University of Tilburg

European University Institute

July 17, 2011

## Abstract

Using laboratory experiments within a New Keynesian macro framework, we explore the formation of inflation expectations and its interaction with monetary policy design. The central question in this paper is how to design monetary policy in the environment characterized by heterogeneous expectations. Rules that use actual rather than forecasted inflation produce lower inflation variability and alleviate expectational cycles. Degree of responsiveness to deviations of inflation from its target in the Taylor rule produces nonlinear effects on inflation variability. We also provide considerable support for the existence of heterogeneity of inflation expectations and show that a significant proportion of subjects are rational in our experiment. However, most subjects rather than using a single model they tend to switch between alternative models.

**JEL:** C91, C92, E37, E52

**Key words:** Laboratory Experiments, Inflation Expectations, New Keynesian Model, Monetary Policy Design.

---

\*We would like to thank Klaus Adam, Steffan Ball, Bruno Biais, Wändi Bruine de Bruin, John Duffy, Chryssi Giannitsarou, Cars Hommes, Seppo Honkapohja, Tobias Klein, Aniol Llorente Sagner, Rosemarie Nagel, Charles Noussair, Jan Potters, Juan Manuel Puerta, Emiliano Santoro, and other participants at the Bank of England, Humbolt University Berlin, University of Amsterdam, Universitat Pompeu Fabra, Erasmus University Rotterdam, European University Institute, University of Hamburg, Catholic University in Milan, Tilburg University, 2010 New York FRB Conference on Consumer Inflation Expectations, 2008 Computational Economics Conference in Paris, 2008 ESA meetings in Lyon, 2008 Nordic Conference on Behavioral and Experimental Economics in Copenhagen and 2008 Learning Week at St. Louis FED for their comments and suggestions. Blaž Žakelj gratefully acknowledges financial support of the Max Weber Fellowship. Damjan Pfajfar gratefully acknowledges funding from a Marie Curie project PIEF-GA-2009-254956 – EXPMAC. All remaining errors are ours. Previous versions of the paper were circulated as "Experimental Evidence on Inflation Expectation Formation."

<sup>†</sup>Department of Economics, Faculty of Economics and Business Administration, P.O. Box 90153, NL-5000 LE Tilburg, Netherlands. *E-mail:* D.Pfajfar@uvt.nl. *Web:* <https://sites.google.com/site/dpfajfar/>.

<sup>‡</sup>Address: Villa La Fonte, Via delle Fontanelle 10, 50014 San Domenico di Fiesole, Italy. *Email:* [blaz.zakelj@eui.eu](mailto:blaz.zakelj@eui.eu).

# 1 Introduction

This paper discusses an experimental study on the interaction between expectations formation process and monetary policy design within a macroeconomic framework. With the development of explicit microfounded models expectations have become pivotal in the modern macroeconomic theory. Moreover, the relationship between monetary policy and expectations is crucial for promoting economic stability. Friedman's proposals (1948 and 1960) for economic stability argue in favor of simple rules as they are easier to learn and facilitate coordination of agents' beliefs. Although, several leading macroeconomists and policy makers, including the Chairman of the Federal Reserve Bernanke (2007), stress the importance of improving understanding of the relationship between economic policies (monetary policy), agents expectations, and equilibrium outcomes, macroeconomists spent little effort on empirically elucidating these issues. Laboratory experiments provide suitable environments to test these relationships as we can keep the underlying model under control and expectation formation processes are observable. The advantage of our experiment lies in the possibility to compare the aggregate dynamics of inflation and output gap and the effectiveness of monetary policy with the theoretical results. We study this question by employing several simple instrumental monetary policy rules in different treatments and examine potential implications of the design of monetary policy for forecasting inflation. Before we focus on the relationship between policy actions and the formation of inflation expectations, we establish some stylized facts about the inflation expectation formation. This paper provides substantial evidence in support of heterogeneity in the forecasting process both across subjects and time.

The experiment is repeated under different monetary policy regimes to assess how alternative conducts of monetary policy influence the expectation formation process and the degree of heterogeneity. The effectiveness of Taylor-type rules is then compared in terms of variability of inflation and inflation forecasts. We explore how different monetary policy settings anchor inflation expectations. We find that the variability of inflation is significantly affected by the degree of aggressiveness of monetary policy. Our results also suggest that instrumental rules responding to contemporaneous inflation perform better than rules responding to inflation expectations. Furthermore, the design of monetary policy significantly affects the composition of forecasting rules used in the experiment (heterogeneity) – especially the proportion of trend extrapolation rules – and thus the stability of the main macroeconomic variables. The proportion of trend extrapolation rules increases in an environment characterized by excessive inflation variability and expectational cycles and then further amplifies the cycles. As already pointed out by Marimon and Sunder (1995) the actual dynamics of an economy is a product of complex interaction between underlying stability properties of the model and agents' behavior. Thus, it is imperative to design a monetary policy that is robust to different expectation

formation mechanisms.

Central banks increasingly attribute more importance to the developments of households' inflation expectations as they signal future inflationary risks and provide useful guidance how to anchor inflation expectations. When analyzing individual responses from students of the Universitat Pompeu Fabra and Tilburg University, we find that subjects form expectations in accordance with different theoretical models.<sup>1</sup> Question of rational expectations is carefully studied and found that for approximately one third of subjects we cannot reject rationality. In the remaining sample the most popular rule seems to be trend extrapolation. A significant share of population also uses adaptive expectations, adaptive learning and sticky information type models.<sup>2</sup> Adaptive learning results are also novel as this paper represents one of the first estimations of the gain parameter. The average gain of agents that employ adaptive learning models is around 0.045. Furthermore, when we allow switching between different models, we find that adaptive learning models are the most frequently used models for forecasting inflation.

Rather than sticking to one model, switching between alternative models seems to describe subjects' behavior better. We observe that on average subjects switch every 4 periods. Therefore, this paper provides an empirical support for models that postulate endogenous switching (e.g. Brock and Hommes, 1997). Furthermore, we also show that agents use different models as on average in each period 4.5 different models are used in groups of 9 subjects. This suggests that observed heterogeneity is pervasive.

A few experimental studies investigate the expectation formation process. Learning to forecast experiments have been conducted before within a simple macroeconomic setup (e.g. Williams, 1987; Marimon, Spear, and Sunder, 1993; Evans, Honkapohja, and Marimon, 2001; Arifovic and Sargent, 2003; Adam, 2007) and also within the asset pricing framework (see Hommes, Sonnemans, Tuinstra, and van de Velden, 2005 and Anufriev and Hommes, 2011).<sup>3</sup> These studies mainly focus on the aggregate expectations formation and tend to reject the rational expectations assumption in favor of adaptive way of forming beliefs.<sup>4</sup> Some analysis of the micro expectations data is conducted by Marimon and Sunder (1995) and Bernasconi and Kirchkamp (2000) in an overlapping generations

---

<sup>1</sup>Several new theories of expectation formation, such as adaptive learning and information stickiness, have not yet been subject to rigorous empirical tests. Empirical contributions so far mostly employed aggregate data. We need to assure that agents' current information sets encompass all the information from the previous periods in order to assess these theories. Controlled laboratory environment avoids these methodological issues that are present in the survey data.

<sup>2</sup>Adaptive learning assumes that subject are acting as econometricians when forecasting, i.e. reestimating their model each time new data becomes available. See Evans and Honkapohja (2001).

<sup>3</sup>See Duffy (2008) and Hommes (2011) for a surveys on experimental macroeconomics.

<sup>4</sup>Arifovic and Sargent (2003) also find support of adaptiveness and some evidence of heterogeneity of forecasts. Arifovic and Sargent (2003) focus on the time inconsistency problem, asserting that in many cases policy makers achieve time-inconsistent optimal inflation rate, although in some treatments the economy moves towards sub-optimal (Nash) time consistent outcomes. Also Fehr and Tyran (2008) suggest that expectations of individuals are heterogeneous. They study the adjustments of nominal prices after the anticipated monetary shock.

framework. These authors estimate several different regressions in order to study inflation expectation formation and find that most subjects behave adaptively, although Bernasconi and Kirchkamp (2000) provide evidence that adaptive expectations are not of first order degree as argued in Marimon and Sunder (1995). So far, these two studies are also the only exceptions that investigate the effects different monetary policies to inflation volatility. Marimon and Sunder (1995) compare different monetary rules in the overlapping generations (OLG) framework to see their influence on the stability of inflation expectations. In particular, they focus on the comparison between Friedman's  $k$ -percent money rule and the deficit rule where the government is fixing the real deficit and finance it through the seigniorage. They find little evidence that Friedman's rule could help to coordinate agents beliefs and help to stabilize the economy. Inflation process might be even more volatile in the case when Friedman rule is announced and maintained compared to the economy where inflation target is not announced, but the rational expectations equilibrium is more stable under learning dynamics. Similar analysis is also performed in Bernasconi and Kirchkamp (2000). They argue that Friedman's money growth rule produces less inflation volatility, but higher average inflation compared to constant real deficit rule.<sup>5</sup>

Closer to our framework is the experiment by Adam (2007). He conducts experiments in a sticky price environment where inflation and output depend on expected inflation and analyzes the resulting cyclical patterns of inflation around its steady state. These cycles exhibit significant persistence and he argues that they closely resemble an restricted perception equilibrium<sup>6</sup> where subjects make forecasts with simple underparametrized rules. In our experiment we also detect cyclical behavior of inflation and output gap in some treatments, however we show that these phenomena are not only associated with underparametrization but also with heterogeneity of expectations, the design of monetary policy and (its influence on) the degree of backward-looking behavior. Recently similar setup to ours is used in Assenza, Heemeijer, Hommes, and Massaro (2011), where they focus on the analysis of switching between different rules.

This paper is organized as follows: Section 2 describes the model for experimental analysis. Section 3 outlines the experimental design. In Section 4 we focus on the analysis of individual responses while in Section 5 we analyze switching dynamics between different models. Section 6 studies the relationship between the monetary policy design and expectation formation; Section 7 concludes.

---

<sup>5</sup>The effects of monetary policy design on expectations were also examined in Hazelett and Kernen (2002) where they search for hyperinflationary paths in the laboratory.

<sup>6</sup>Restricted perception equilibrium is generally more volatile than rational expectation equilibrium (for more details see Evans and Honkapohja, 2001).

## 2 Model

In our experiment we use forward-looking sticky price New Keynesian (NK) monetary model with different monetary policy reaction functions. The advantage of the NK model is that it is widely used in policy analysis and allows us to compare our experimental results with those obtained theoretically. However, there are two implicit complications for participants. First, it requires to forecast two periods ahead. It would definitely be easier for participants to produce one period ahead forecast (sometimes called "nowcasting") as they would observe the realizations immediately after their forecasts are made. This would also enable us to simplify the analysis of individual responses, especially in the case of adaptive learning. The second complication is that the forward-looking NK models assume that agents have to forecast both inflation and output gap. We were afraid that this would represent a too difficult task for subjects. This is a considerably more difficult decision to make as we would depart from a standard macro model if we would only ask participants to forecast inflation. Nevertheless, we decided to do this experiment only with expectations of inflation as we were afraid that both issues mentioned in this paragraph would make the task too difficult for individuals. We leave the fully forward-looking NK model for future work.

The baseline framework in the NK approach is a dynamic stochastic general equilibrium model with money, nominal price rigidities, and rational expectations (RE). Lately some authors have augmented this model for adaptive learning and also for heterogeneous expectations (e.g. Branch and McGough, 2009). The model consists of a forward-looking Phillips curve (PC), an IS curve, and a monetary policy reaction function.<sup>7</sup>

In this paper we decided to focus on the reduced form of the NK model, where we can clearly elicit forecasts and study their relationship with monetary policy. Of course, there is a trade-off between using the model from "first principles" and employing a reduced form. The former has the advantage of setting the objectives (payoff function) exactly in line with microfoundations, however forecasts are difficult to elicit in this environment where subject act as producers and consumers and interact on labor and final product markets and do not explicitly provide their forecasts (for the latter approach, see Noussair, Pfajfar, and Zsiros, 2011). Therefore, an appropriate framework, for the question that we address in this paper, is the "learning to forecast" design where incentives are set in order to induce as accurate forecasts as possible.<sup>8</sup> In this framework, thus, we do not assign subjects a particular role in the economy, rather they act as "professional" forecasters.<sup>9</sup>

The information set at the time of forecasting consists of macro variables at the time

---

<sup>7</sup>Detailed derivations are in, e.g., Woodford (1996), or textbooks such as Walsh (2003) or Woodford (2003).

<sup>8</sup>The argument is similar to that in papers by Marimon and Sunder, where the same tradeoff was first recognized.

<sup>9</sup>One way to think about the relation between "professional forecasters" and consumers/firms is that these economic subjects employ professional forecasters to provide them with forecasts of inflation.

$t - 1$ , although the forecasts are made in period  $t$  for period  $t + 1$ . Mathematically we denote this as  $E_t \pi_{t+1}$ . Strictly speaking, it should be denoted as  $E_t (\pi_{t+1} | \mathcal{I}_{t-1})$ . In fact,  $E_t$  (forecast made at period  $t$  with information set  $t - 1$ ) might not be restricted to just rational expectations.

The IS curve is specified as follows:

$$y_t = -\varphi (i_t - E_t \pi_{t+1}) + y_{t-1} + g_t, \quad (1)$$

where interest rate is  $i_t$ ,  $\pi_t$  denotes inflation,  $y_t$  is output gap, and  $g_t$  is an exogenous shock. The parameter  $\varphi$  is the intertemporal elasticity of substitution in demand. We can observe that we do not have expectations of output gap in the specification. Instead, we have lagged output gap.<sup>10</sup> Compared to purely forward-looking specifications, our model might display more persistence in output gap. This is the most significant departure from otherwise standard macroeconomic model.

Aggregating across the price setting decisions of individual firms yields the linear relationship in the equation (2). Thus, the supply side of the economy is summarized in the following PC:

$$\pi_t = \beta E_t \pi_{t+1} + \lambda y_t + u_t. \quad (2)$$

The longer prices are fixed on average, i.e. the smaller is  $\lambda$ , the less sensitive inflation is to the current output gap. The parameter  $\beta$  is the subjective discount rate. The shocks  $g_t$  and  $u_t$  are unobservable to subjects and follow the following process:

$$\begin{bmatrix} g_t \\ u_t \end{bmatrix} = \Omega \begin{bmatrix} g_{t-1} \\ u_{t-1} \end{bmatrix} + \begin{bmatrix} \tilde{g}_t \\ \tilde{u}_t \end{bmatrix}; \quad \Omega = \begin{bmatrix} \kappa & 0 \\ 0 & \nu \end{bmatrix},$$

where  $0 < |\kappa| < 1$  and  $0 < |\nu| < 1$ .  $\tilde{g}_t$  and  $\tilde{u}_t$  are independent white noises,  $\tilde{g}_t \sim N(0, \sigma_g^2)$  and  $\tilde{u}_t \sim N(0, \sigma_u^2)$ . In the NK literature it is standard to assume AR(1) shocks.  $g_t$  could be justified as a government spending shock or a taste shock and standard interpretation of  $u_t$  is the technology shock. All these shocks are found to be quite persistent in the empirical literature (see e.g. Cooley and Prescott, 1995 or Ireland, 2004). In the experimental context it is important to have some exogenous unobservable component in the law of motion for endogenous variables, so that we prevent the extreme case where all agents coordinate on the forecasts identical to inflation target. If we would not have AR(1) shocks this would represent the dominant strategy. This is an especially relevant concern as we initialize the model in the rational expectations equilibrium (REE).

---

<sup>10</sup>In principle, one could argue that this specification of IS equation corresponds to the case when subjects have naive expectations on output gap or it is assumed the extreme case of habit persistence. The main reason for including lagged output gap in our specification is that we want another endogenous variable to influence the law of motion for inflation. Furthermore, we prefer that even in the case when agents have rational expectations they have to use the observed information on output gap for forecasting inflation as it enters into the perceived law of motion of the rational expectations form.

To close the model, we have to specify the interest rate rule.<sup>11</sup> We use two alternative Taylor-type rules in different treatments. Most of our attention is devoted to forward-looking reaction functions: inflation forecast targeting where interest rate is set in response to inflation expectations. We study three parametrizations of this rule and investigate how different degrees of central bank’s aggressiveness in stabilizing inflation influence inflation expectations. Next, we ask whether it is better for the central bank to respond to the current or expected inflation. Therefore, we also analyze the inflation targeting.

We start with the following interest rate rule (*Inflation Forecast Targeting*):

$$i_t = \gamma (E_t \pi_{t+1} - \bar{\pi}) + \bar{\pi}. \quad (3)$$

In this version the central bank responds to deviations of inflation from the target,  $\bar{\pi}$ . We vary  $\gamma$  in different treatments and study stability of the system under alternative reaction coefficients attached to inflation.

The second alternative specification is *Inflation Targeting*, where the monetary authority is assumed to respond to deviations of contemporaneous inflation from the inflation target:

$$i_t = \gamma (\pi_t - \bar{\pi}) + \bar{\pi}. \quad (4)$$

We use McCallum and Nelson (2004) calibration. This calibration represents one of the standard calibrations for the NK models. In order to have inflation in positive numbers for most of the periods we set the inflation target to  $\bar{\pi} = 3$ . A summary of the calibration is reported in the next table.

Insert Table 1 about here

Treatments are fully comparable as we have exactly the same shocks in all treatments. In particular,  $\kappa$  and  $\nu$  are calibrated to 0.6, while their standard deviations are 0.08.

---

<sup>11</sup>Engle-Warnick and Turdaliev (2010) and Noussair, Pfajfar, and Zsiros (2011) investigate the conduct of monetary policy in an experimental setting. Their subjects are only told to act as policymakers and to stabilize inflation. Most of the subjects control inflation relatively well and authors argue that Taylor rules provide a good description of subjects’ policy decisions.



## 3 Experiment

### 3.1 Design<sup>12</sup>

Experimental subjects participated in a simulated economy of 9 agents.<sup>13</sup> Each session of a treatment has 2 independent groups ("economies"), therefore 18 subjects participate in each session. All participants were recruited through a recruitment programs for undergraduate students at the Universitat Pompeu Fabra and University of Tilburg. Invitations to apply were sent to all of around 1300 students in a database at Pompeu Fabra (in May 2006) and to about 1200 students at Tilburg (in June 2009), except to those that already participated in one of our sessions before. There are 70 periods in each treatment. We scaled the length of each decision sequence and number of repetitions in a way that each session lasts approximately 90 to 100 minutes, including the time for reading the instructions and 5 trial periods at the beginning. The program is written in Z-Tree experimental software (Fischbacher, 2007).

Subjects are presented with a simple fictitious economy setup. As it is shown above, the economy is described with three macroeconomic variables: inflation, output gap and interest rate. Participants observe time series of these variables in a table, up to the period  $t - 1$ . 10 initial values (periods  $-9, \dots, 0$ ) are generated by the computer under the assumption of rational expectations. Subjects' task is to provide inflation forecasts for the period  $t + 1$ . The underlying model of the economy is qualitatively described to them. We explain the meaning of and the relationship between the main macroeconomic variables and inform them that their decisions have an impact on the realized output, inflation and interest rate in time  $t$ . This is a predominant strategy in the learning to forecast experiments (see Duffy, 2008, and Hommes, 2011).<sup>14</sup>

In every period  $t$ , there are two decision variables subjects have to input: *i*) prediction of the  $t + 1$  period inflation; *ii*) 95% confidence interval of their inflation prediction. In 4 out of 6 independent groups in each treatment subjects have to report the interval as a number of percentage points for which the actual inflation can be higher or lower. In the

---

<sup>12</sup>Experimental instructions can be found in Appendix C.

<sup>13</sup>Most of the learning to forecast experiments are conducted with 5-6 subjects, e.g. Hommes, Sonnemans, Tuinstra, and van de Velden (2005), Adam (2007), Fehr and Tyran (2008).

<sup>14</sup>In learning to forecast experiments it is not possible to achieve REE (Rational Expectations Equilibrium) simply by introspection. This holds even if we provide subjects with the data generating process as there exists uncertainty how other participants forecast, so subjects have to engage in a number of trial and error exercises or in other words adaptive learning. It has been analytically proven in Marcet and Sargent (1989) and further formalized in a series of papers by Evans and Honkapohja (see their book: Evans and Honkapohja, 2001) that it is enough that agents observe all relevant variables in the economy (as in our case, where they are specifically instructed that all of them might be relevant) and update their forecasts according to the adaptive learning algorithm (their errors) they will end up in the REE. This has been acknowledged also in Duffy (2008) and Hommes (2011). Kelley and Friedman (2008) provide a survey of experiments that support the theoretical result above. Examples of learning to forecast experiments are e.g., Marimon and Sunder (1993, 1994), Adam (2007) and Hommes, Sonnemans, Tuinstra, and van de Velden (2005).

other 2 groups in each treatment, subjects are simply asked for the lower and the upper bound of their inflation prediction interval.

After each period subjects receive information about the realized inflation in that period, their prediction of it, and the payoff they have gained. Subjects' payoffs depend on the accuracy of their predictions. The accuracy benchmark is the actual inflation rate computed from the underlying model on the basis of predictions made by all agents in the economy. We replace  $E_t\pi_{t+1}$  in equations (1), (2), and (3) by  $\frac{1}{K} \sum^k \pi_{t+1|t}^k$ , where  $\pi_{t+1|t}^k$  is  $k$ -subject' point forecast of inflation ( $K$  is total number of subjects in an economy). In the subsequent rounds subjects are also informed about their past forecasts. They do not observe the forecasts of other individuals and their performance. The payoff function,  $W$ , is a sum of two convex components as described below:

$$\begin{aligned} W &= W_1 + W_2, \\ W_1 &= \max \left\{ \frac{1000}{1+f} - 200, 0 \right\}, \quad W_2 = \max \left\{ \frac{1000x}{1+CI} - 200, 0 \right\}, \\ x &= \begin{cases} 1 & \text{if } CI \geq f \\ 0 & \text{if } otherwise \end{cases}, \quad f = |\pi_t - \pi_{t+1|t}^k|. \end{aligned}$$

The first,  $W_1$ , depends on their forecast errors and is designed to encourage subjects to give accurate predictions. It gives subjects a payoff if their forecast errors,  $f$ , are smaller than 4. The second,  $W_2$ , depends on the width of their confidence interval and intends to motivate subjects to think about the variance of actual inflation since it is more rewarding when it is narrower.  $CI$  is either equal to their point estimate of confidence interval or half of the difference between the upper and the lower bound. They receive a reward if their confidence intervals,  $CI$ , are not larger than  $\pm 4$  percentage points, conditional on the fact that actual inflation falls in the given interval:  $CI \geq \left| \pi_t - \pi_{t+1|t}^k \right|$ . With this setup we restrict to positive payoffs. Compared to more standard quadratic payoff functions, ours gives greater reward to more accurate predictions and incentivize them to think also about small variations of inflation, which may be important for their payoff. As potentially this experiment can produce quite different variations of inflation between different sessions it is important to keep the incentive scheme quite steep. The payoff function is non-linear, thus, we accompanied it with generous explanation and a payoff matrix on a separate sheet of paper to make sure all participants understood the incentives. Similar approach is used in Adam (2007).

Participants received detailed instructions before the experiment started. To ensure understanding of the task, we read instructions out loud and present their task descriptively along with examples. Subjects also filled in a short questionnaire after they have read the instructions and answered the questions about the procedures to make sure that all participants understood them.

## 3.2 Treatments

The experiment consists of 5 sessions (a pilot session and 4 regular sessions). Participants on average earn around €15 ( $\approx$ \$22), depending on treatment and individual performance. Every experimental session represents a different treatment, each using a different specification of monetary policy reaction function.

Insert Table 2 about here

The first three treatments, as shown in the Table 2, deal with the parametrization of the inflation forecast targeting given in equation (3). In this setup, the coefficient  $\gamma$  determines central bank' aggressiveness to deviations of inflation from its target. It is also believed that the higher the  $\gamma$  is, the stronger is the stabilizing effect of the monetary policy rule. It is of our key interest to see how subjects react to more and less aggressive interest rate policies. Moreover, we test in a controlled environment whether different slope coefficients indeed have the expected stabilization effect.

Majority of empirical findings agree that the magnitude of the slope coefficient is around 1.5. Generally, when  $\gamma > 1$  the interest rate rule is E-stable and produces a determinate outcome<sup>15</sup> (Taylor principle) under rational expectations while the one with  $\gamma \leq 1$  is E-unstable and indeterminate. When Taylor principle holds all our treatments yield determinate and E-stable REE. Initially, we planned to perform a treatment with  $\gamma < 1$  to check whether this leads to instability, however findings from the pilot treatments convinced us this is not a suitable choice as subjects quickly reached extremely high levels of inflation. This clearly leads to explosive behavior of the system, so our findings suggest that also under heterogeneous expectations Taylor principle is required to produce an E-stable and determinate outcome.<sup>16</sup>

For our first and benchmark treatment we decided to follow Taylor and chose  $\gamma = 1.5$ . Average behavior of groups in the first treatment show no convergence to target inflation, so we choose  $\gamma = 1.35$  as sufficiently different case for a comparison. Alternatively, we chose  $\gamma = 4$  as parametrization with high stabilizing effect where convergence to the target inflation should be faster. We study determinacy and E-stability of these treatments in Table 10.

In treatment 4 we focus on what measure of inflation should central banks target: the expected inflation by subjects or actual inflation. We perform a treatment using

---

<sup>15</sup>E-stability is asymptotic stability of an REE under least squares learning. Under determinacy we mean the existence of a unique dynamically stable REE. For more detailed definition see Evans and Honkapohja (2001). Proof that this is also the case in our setup can be found in Table 10.

<sup>16</sup>Moreover, under these circumstances inflation never returned to the target inflation and just kept growing. Therefore the effect of output gap on inflation never outweighs the expected inflation effect. Assenza, Heemeijer, Hommes, and Massaro (2011) perform a treatment where  $\gamma = 1$ . In their economy with i.i.d. shocks this results in a convergence to values of inflation that are different than the target value.

inflation targeting rule where central bank reacts to current inflation, with  $\gamma = 1.5$  as in our benchmark case.

## 4 Analysis of Individual Inflation Forecasts

Before we move to the analysis of the interaction between monetary policy rules and inflation expectations, we have to first establish how subjects form inflation expectations. The analysis of individual responses focuses in the first part on learning dynamics. Several learning models are simulated in order to find the best fit of each individual series on expectations. We also estimate other standard models of expectation formation including rationality tests. All these models are estimated for each individual using OLS. Reported results are with robust standard errors that, where appropriate, take into account the presence of clusters in groups (or treatments). Below we present each of these models and tests. In the discussion we determine the best performing model for each subject, based on the comparison of sum of squared errors (SSE) between competing models. In the Section 5 we dig deeper and investigate potential switching of subjects between different models.

In 4 treatments of our experiment and 24 independent groups we gathered 15,120 point forecasts of inflation from 216 subjects. The mean inflation forecast for all treatments is around 3.06% and the mean inflation is 3.02% where the inflation target is set to 3%. Standard deviations of inflation and inflation expectations vary substantially across groups. For inflation expectations the largest is 6.31 and the lowest 0.23 while for inflation the largest is 5.83 and the smallest is 0.24. Standard deviations of inflation forecasts are usually higher than standard deviations of inflation for groups with higher volatility while for groups with lower volatility this might not necessary be the case. Figure A1 in the Appendix A displays distribution of inflation forecasts in each treatment.

Insert Table 3 about here

In Figure 1<sup>17</sup> it is possible to distinguish signs of rounding effect (or digit preference). This is especially evident for the responses bellow 0 and above 6, however rounding is also present for the responses between 0 and 6. Overall, we can point out that 72% of all responses are reported to one decimal point accuracy, while 13% of them are to the accuracy of 2 decimal points. The remaining 15% of forecasts are rounded as integers. The overall share of the latter is significantly higher for the groups with higher volatility compared to the groups displaying lower volatility.

Insert Figure 1 about here

---

<sup>17</sup>The full range of responses reported is between  $-13.9$  and  $24$ , however in this histogram we restrict to responses between  $-3$  and  $10$ .

However, we have to point out that survey data usually display more rounding, particularly the Michigan survey (see Curtin, 2005, Bryan and Palmqvist, 2005). Subjects in experiments are paid according to their performance and thus the accuracy of forecasts always matters. The mean of forecast errors in our experiment is 0.04 and the standard deviation is 1.23. Thus, there is only a slight positive bias of errors. Furthermore, subjects overpredict in 51.2% cases and underpredict in 48.8%. Analysis on confidence intervals can be found in our companion paper, Pfajfar and Žakelj (2011).

## 4.1 Models of Individuals Expectation Formation

We evaluate 10 models of expectation formation for each individual. Models are summarized in the Table 4. For detailed discussion and brief description of the results of each model see Appendix B.

model (eq.)	specification
AR(1) process (M1)	$\pi_{t+1 t}^k = \phi_0 + \phi_1 \pi_{t t-1}^k + \varepsilon_t$
Sticky information type (M2)	$\pi_{t+1 t}^k = \lambda_1 \eta_0 + \lambda_1 \eta_1 y_{t-1} + (1 - \lambda_1) \pi_{t t-1}^k + \varepsilon_t$
Adaptive expectations CGL (M3)	$\pi_{t+1 t}^k = \pi_{t-1 t-2}^k + \vartheta \left( \pi_{t-1} - \pi_{t-1 t-2}^k \right) + \varepsilon_t$
Adaptive expectations DGL (M4)	$\pi_{t+1 t}^k = \pi_{t-1 t-2}^k + \frac{\vartheta}{t} \left( \pi_{t-1} - \pi_{t-1 t-2}^k \right) + \varepsilon_t$
Trend extrapolation (M5)	$\pi_{t+1 t}^k - \pi_{t-1} = \tau_0 + \tau_1 (\pi_{t-1} - \pi_{t-2}) + \varepsilon_t$
General model (M6)	$\pi_{t+1 t}^k = \alpha + \gamma \pi_{t-1} + \beta y_{t-1} + \mu i_{t-1} + \zeta \pi_{t t-1}^k + \varepsilon_t$
Recursive - lagged inflation (M7)	$\pi_{t+1 t}^k = \phi_{0,t-1} + \phi_{1,t-1} \pi_{t-1} + \varepsilon_t$
Recursive - REE (M8)	$\pi_{t+1 t}^k = \phi_{0,t-1} + \phi_{1,t-1} y_{t-1} + \varepsilon_t$
Recursive - trend extrapolation (M9)	$\pi_{t+1 t}^k - \pi_{t-1} = \phi_{0,t-1} + \phi_{1,t-1} (\pi_{t-1} - \pi_{t-2}) + \varepsilon_t$
Recursive - AR(1) process (M10)	$\pi_{t+1 t}^k = \phi_{0,t-1} + \phi_{1,t-1} \pi_{t t-1}^k + \varepsilon_t$

Table 4: Models of inflation expectation formation. Note:  $\pi_t$  is inflation at time  $t$ ,  $y_t$  is output gap,  $i_t$  is interest rate, and  $\pi_{t+1|t}^k$  is  $k^{th}$  subject's inflation expectations for time  $t + 1$  made at time  $t$  (with information set  $t - 1$ ).

Model (M1) is a simple AR(1), while model (M2) represents a weighted average regression similar in formulation to sticky information model by Carroll (2003a) and adaptive expectations. In our framework we have forecasts derived under the assumption of rational expectations while Carroll (2003a) implements professional forecasters predictions. This type of models are important for forecasting, especially in our framework where some agents are backward-looking and also rational agents have to incorporate this into their forecasts. Thus we estimate the model that is stated in terms of observable variables with the restrictions on all coefficients, where  $\eta_0$  and  $\eta_1$  are REE coefficients.

In order to test for adaptive behavior, we apply different learning rules to experimental data. We first test learning on a model with constant gain updating (CGL) in model (M3), where subjects revise their expectations according to the last observed error.  $\vartheta$  is the constant gain parameter.

We also check whether they learn with a decreasing gain parameter (DGL) in model (M4). If the estimated parameter ( $\iota$  in this version) is significantly different from 0, we conclude that agents actually learn from their past mistakes with a decreasing gain over time.

Next we evaluate simple trend extrapolation rules (M5). These are pointed out as particularly important rules for expectation formation process in Hommes, Sonnemans, Tuinstra, and van de Velden (2005). Simple learning rules do not capture all macroeconomic factors that can affect inflation forecasts. Therefore we estimate a more general model of expectation formation described in model (M6).<sup>18</sup>

#### 4.1.1 Recursive Representation of Simple Learning Rules

In this subsection as in the adaptive learning literature we assume that subjects behave like econometricians, using all available information at the time of the forecast. In the following specifications, we test whether agents update their coefficients with respect to the last observed error. We use this estimation procedure for models (M7)-(M10). When agents estimate their PLMs they exploit all available information up to period  $t - 1$ . As new data become available they update their estimates according to a stochastic gradient learning (see Evans, Honkapohja, and Williams, 2010) with a constant gain. Let  $X_t$  and  $\widehat{\phi}_t$  be the following vectors:  $X_t = \begin{pmatrix} 1 & \pi_t \end{pmatrix}$  and  $\widehat{\phi}_t = \begin{pmatrix} \phi_{0,t} & \phi_{1,t} \end{pmatrix}'$ . In this version of constant gain learning (CGL) agents update coefficients according to the following rule:

$$\widehat{\phi}_t = \widehat{\phi}_{t-2} + \vartheta X'_{t-2} \left( \pi_t - X_{t-2} \widehat{\phi}_{t-2} \right). \quad (5)$$

The empirical approach consists in searching the parameter  $\vartheta$  that minimizes the sum of squared errors (SSE), i.e.  $\left( \pi_{t+1|t}^s - \pi_{t+1|t}^k \right)^2$  (see Pfajfar and Santoro, 2010 for details). The implicit problem in this approach is that we have to assume the initial values for  $\widehat{\phi}_t$  for 2 periods. Setting up the initial values is one of the main problems when we recursively estimate learning. This issue is extensively discussed in Carceles-Poveda and Giannitsarou (2007). Strictly speaking, this problem should not occur in our case since we simply try to replicate our time-series data as closely as possible. Thus, in the following recursive learning estimations, we design an exercise in order to search for the best combinations of the gain parameter and initial values to match each subjects' expectations as closely as possible. This strategy can also be considered as a testing procedure for the detection of learning dynamics for each individual. If the gain is positive under this method of initialization, then the series would exhibit learning for all other initialization methods with higher (or equal) gain.

---

<sup>18</sup>Models in groups 19-24 do not have interest rate as dependent variable as this would imply multicollinearity due to the design of monetary policy in our framework.

## 4.2 Comparison with "Classical Econometrician"

Before we discuss the best performing model for each individual we ask ourselves how would a "classical" econometrician forecast inflation in this environment.<sup>19</sup> We estimate a regression for each period in time using only the available information that is on the subjects' screens. Therefore, we estimate rolling regressions and make one-step-ahead forecasts. Similar approach is used by Branch (2004) in the survey data literature for proxying the rational expectations. He uses a trivariate VAR model and estimates it recursively. In our case, because of degrees of freedom problems, we have to resort to a univariate model (M6). For a comparison, we also recursively estimate the adaptive expectations model (M3) and a version of trend extrapolation rule without the restrictions on coefficients. In practice, this rule is then equivalent to the AR(2) model. We evaluate the general model (M6) with and without the restriction:  $\zeta = 0$ . Then we compare these forecasts with the actual realizations and compute SSE, which are presented in the table below for five competing models. Before starting the analysis it is worth pointing out that in treatments where the variance of inflation is higher also the mean SSE is higher (correlation coefficient is 0.91). In two thirds of our groups the trend extrapolation rule is performing best. However, in more stable treatments the general model can also outperform the trend extrapolation rule.

Insert Table 5 about here

This table gives us a benchmark for evaluating prediction accuracies of subjects. It is interesting to note that best performing subjects often outperform our classical econometrician (best performing model). This practically occurs in all groups except the ones comprising treatment 3 where high frequency of cycles is observed (see Figure 2). There are two potential explanations for that: first, some subjects might be rational or at least weakly rational and second, subjects might be switching between different expectation formation mechanisms. We start by investigating the first possibility while in the next section we dig deeper regarding the second possible explanation.

## 4.3 Rational Expectations

Suppose that all subjects in the economy are rational, then their perceived law of motion (PLM) is equal to the actual law of motion (ALM). There exists a unique evolutionary

---

<sup>19</sup>Of course, a more "sophisticated econometrician" could do a better job. For example, exogenous shocks are not observable in our framework, but a better econometrician could design an unobserved components model to extract information about the autoregressive shocks and then use them in these regressions. In the RE paradigm shocks play a significant part in the formation of expectations. In some treatments it is possible to observe that at least some agents extract information about the shock in the PC and at least partly use this information when forecasting. This is especially evident in treatment 4.

stable rational expectations equilibrium (REE) in our economy, which has the following form:

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = B \begin{bmatrix} 1 \\ y_{t-1} \end{bmatrix} + C \begin{bmatrix} g_{t-1} \\ u_{t-1} \end{bmatrix}, \quad (6)$$

where  $B$  is matrix of coefficient specific to each treatment that is presented in the first column of Table 10 below along with other properties of possible equilibria in this framework.  $C$  is matrix of coefficient values for exogenous variables. Note that  $\pi_{t-1}$  does not enter the REE solution. If we test this strict form of individual rationality we find that we can reject it for all subjects.<sup>20</sup>

Several econometric tests are designed to check different implications of the rationality of forecasts, therefore most of them only assess weaker forms of rationality. Standard tests commonly employed in the survey data literature<sup>21</sup> focus on the most straightforward implication of rationality: forecast efficiency. The simplest test of efficiency is a test of bias:

$$\pi_{t+1} - \pi_{t+1|t}^k = \alpha + \varepsilon_t, \quad (7)$$

By regressing expectational errors on a constant we check whether inflation expectations are centred around the right value. An additional test of rationality is presented in Mankiw, Reis, and Wolfers (2004), where it is tested whether information in forecasts are fully exploited:

$$\pi_{t+1} - \pi_{t+1|t}^k = a + (b - 1) \pi_{t+1|t}^k, \quad (8)$$

where rationality implies jointly that  $a = 0$  and  $b = 1$ . As in the test for bias, under the null of rationality these regressions are meant to have no predictive power.

So far most of the experimental and empirical studies in economics have relied on certain statistical properties to test for (weaker version of) rational expectations hypothesis as defined in (7) and (8). However, we believe that we can go one step further as we are able to control subjects' information sets and also more directly assess the RE hypothesis.<sup>22</sup> RE hypothesis postulates that rational agents use the correct distribution in predicting the variables relevant for their decisions. Note that strictly speaking this does not require the agents to know the model. Nevertheless, some discussion about the form of the solution is warranted at this point. When all agents know the macroeconomic model and behave accordingly, the form of RE solution is (6) and the actual coefficient values are presented in the first column of Table 10. However, as soon as one subject departs from this behavior, other subjects must take this departure into account when they make forecasts in order to be rational. Strictly speaking, they have to know his exact

<sup>20</sup>As it can be observed in the next subsection, this REE PLM model (M8) never outperforms other models. Therefore, individual rationality (under the assumption of homogeneous expectations) can be rejected for all our participants.

<sup>21</sup>See Pesaran (1987), and Bakhshi and Yates (1998) for a review of these methods.

<sup>22</sup>For a survey of experimental papers that assess the RE hypothesis see Kelley and Friedman (2008).



PLM, so that they can implement this information into their PLMs. This is of course very restrictive, and often impossible to expect as most of the experiments, including ours, assume that the forecasts of others are not observable. By including interest rate to their PLMs they can indirectly implement this information in treatments 1-3. Thus, in the environment of heterogeneous forecasts the REE PLM (sometimes called "social" rationality) may be of a different form than the REE PLM in the case of homogeneous rational forecasts (sometimes called individual rationality) as presented in (6).<sup>23</sup> For example, if some agents use the PLM with last observed inflation, and the (socially) rational agents are aware of that, then they have to include the last observed inflation in their PLMs as well.

As we can only estimate PLMs of individuals we have two different possibilities to test the social rationality, or more precisely RE in heterogeneous expectations environment: *i*) to use the statistical definition of rational expectations mentioned in (7) and (8), or *ii*) to develop a new test for rationality based on comparing PLMs and the ALM. The intuition behind the test comes from the minimal state variable procedures for calculating RE. We pursue this by estimating the ALM for inflation in each group and check whether the estimated coefficients of the corresponding PLM entail statistically different coefficients to the ones of ALM. The appropriate ALM should be as general as possible. This approach is also consistent with the adaptive learning view of forming expectations as it is implicitly assumed in our experimental design. We assume that the ALM is of the following form, which encompasses all the rules (forms of PLMs) that are discussed above:

$$\pi_{t+1} = \gamma_0 + \gamma_1\pi_{t-1} + \gamma_2\pi_{t-2} + \gamma_3y_{t-1} + \gamma_4i_{t-1} + \varepsilon_t, \quad (9)$$

and the corresponding correctly parameterized PLM is:

$$\pi_{t+1|t}^k = \beta_0 + \beta_1\pi_{t-1} + \beta_2\pi_{t-2} + \beta_3y_{t-1} + \beta_4i_{t-1} + \varepsilon_t. \quad (10)$$

In order that we can claim that one subject has rational expectations the estimated coefficients in both regressions should not be statistically different. To test for that we estimate the following equation:

$$\pi_{t+1} - \pi_{t+1|t}^k = \mu_0 + \mu_1\pi_{t-1} + \mu_2\pi_{t-2} + \mu_3y_{t-1} + \mu_4i_{t-1} + \varepsilon_t, \quad (11)$$

---

<sup>23</sup>In the words of Marimon and Sunder (1995): "A rational agent might not close his eyes to what he actually sees happening" (p. 116). This issue has been discussed in the experimental finance literature and was labelled by Asparouhova, Bossaerts, Shin, and Cheng (2011) as the difference between individual and social rationality. Nunes (2009) and Molnár (2007) have also discussed this in the theoretical macroeconomic literature. Nunes (2009) studies this problem in the context of forward-looking NK model and shows how to solve the model under the assumption of heterogeneous expectations. In their theoretical frameworks they assume that the information sets of individuals include other subjects' forecasts.

where  $\mu_i = \gamma_i - \beta_i$ . For a subject to forecast rationally all estimated coefficients (jointly) in the equation (11) should not be statistically significant. In the discussion below we compare these definitions of RE. Rationality is in this case "superimposed" as we classify all agents that satisfy the requirements as rational irrespective of their expectation formation mechanism.

## 4.4 Discussion

In this section we determine which theoretical model on average best describes the behavior of each individual. We compare the SSE<sup>24</sup> of each individual for the 10 models of expectation formation that are described above. A subject is regarded to use the model which produces the lowest SSE between the model predictions and their actual predictions.

We compare 9 models of inflation expectation formation that best describe the behavior of at least 1 participant. Model (M4) is never used as it is always outperformed by other models.

Insert Table 6 about here

In Table 6 we present 6 different comparisons as we use different definitions of the RE. In comparisons 1 and 2 we define the RE based on statistical properties as defined in eq. (8) while in comparisons 3 and 4 based on theory as it is outlined above in section 4.3 in eq. (11): in comparison 1 (3) at 5% significance level and in comparison 2 (4) at 1% significance level. In comparison 6 we compare all empirical models, while in comparison 5 we exclude the general model from the set of alternative models.

We can observe that results are indeed quite similar across alternative definitions of RE, although the theoretical definition (comparisons 3 and 4) suggests a slightly higher proportion of rational subjects. One possible reason is that we estimate the model (11) under the assumption of common AR(1) errors as the experiment design embeds unobserved AR(1) shocks. Without this assumption comparisons 3 and 4 would imply 27.3% and 31.0% of rational subjects. Generally, there is evidence that in all treatments about 30 – 45% of subjects are rational and about 25 – 35% of agents simply extrapolate trend. Around 5 – 10% of agents employ adaptive expectations while the remaining 15 – 25% of subjects mostly behave in accordance with new theories of expectation formation, adaptive learning and sticky information type models. As mentioned before, most of other papers in the experimental literature stress the importance of adaptive expectations.<sup>25</sup>

---

<sup>24</sup>Results and conclusions are the same irrespectively whether we use RMSE (root mean square error),  $R^2$  or SSE as they are all monotonic transformations of each other.

<sup>25</sup>For an illustration we report results from the empirical literature on heterogeneous expectation, although these results might not be fully comparable to ours. Branch (2004) presents the results for 3 competing models of expectation formation (VAR, adaptive, and naive) estimated based on Michigan

The estimation of recursive models (adaptive learning models) deserves a special attention. For a comparison we consider comparison 1 in the Table 6. In this case, we find that the average gain of these subjects is 0.0447 with a standard deviation of 0.0537 (median is 0.0260). The standard deviation is quite high as there are a few very high values, but most of the gains fall in the range between 0.01 and 0.07. There are only a few estimates of the gain coefficient in the literature. Orphanides and Williams (2005a) suggest a gain between 0.01 – 0.04 and Milani (2007) estimates it at 0.0183, while Pfajfar and Santoro (2010) find smaller gains (around 0.00021 for a similar version of learning). Results in this paper suggest slightly higher gains than most of the above papers, but our data might be more volatile than the actual US inflation.

The availability of information is probably the main reason why our results suggest higher degree of rationality than some previous studies on the inflation expectation. We must bear in mind that subjects in our experiment have always available historical series on all relevant macroeconomic variables and their past predictions. In the real world all variables might not be readily observable or the information cost for collecting them might play an important role. The other reason for high degree of rationality is that we initialize the model under RE as explained in eq. (6). All these increase the possibility of not rejecting the assumption of rationality.

We also study the degree of heterogeneity in each treatment separately. We present comparison 1 across all treatments in Table A1 in Appendix A where we can observe that there is quite a lot of heterogeneity across treatments. We further discuss this in the next section, where we analyze switching between different rules and assess the second possible explanation why subjects might perform better than the "classical econometrician."

## 5 Switching Between Different Models

The aim of this section is to further investigate how subjects form expectations. Do they consistently use one model or do they switch between different models? There are some attempts in the literature to link the performance of forecasting rules to the share of agents using that rule. Models that explore this issue are generally labeled as rationally heterogeneous expectations models. Some examples of these models are Brock and Hommes (1997), Branch and McGough (2008) and Pfajfar (2008). Their main argument is that it is not always optimal from an utility maximization point of view to forecast rationally as this might entail some costs.

In this section we tackle the problem from a slightly different perspective as we only

---

survey data. He finds that about 48% of agents use a VAR predictor and 44% of agents behave adaptively, while the 7% are naive. Expectation formation of prices is also studied in the US beef market. Chavas (2000) estimates that 81.7% of agents are boundedly rational using simple univariate models to forecasts prices and 18.3% of agents are rational. Contrary to that, Baak (1999) finds that the proportion of rational agents is higher, i.e. about two thirds of agents are rational, while others are boundedly rational.

have 9 subjects in each group. Their information sets are different as subjects do not directly observe past forecasts of other subjects. Thus it is not possible to compare these different models of dynamic predictor selection in our setup. We rather focus on establishing some stylized facts about "unrestricted" switching on the individual basis. An alternative approach, where all agents have the same information set is investigated in Anufriev and Hommes (2008). They provide support for switching based on a version of the predictor dynamics analyzed in Hommes, Huang, and Wang (2005) and show that in an asset pricing environment the model with switching between simple heuristic rules can replicate the main results of the Hommes, Sonnemans, Tuinstra, and van de Velden (2005) experiment in terms of individual behavior and aggregate dynamics. This approach is followed also in Assenza, Heemeijer, Hommes, and Massaro (2011) where the environment is more similar to ours. We proceed this analysis somehow differently as our results above postulate that several of the employed rules are based on personal information, i.e. they include their own past forecast (which is unobservable to others) to their forecasting rule. In essence, we look at the roots of the switching behavior, where we do not impose a particular switching mechanism.

## 5.1 Unrestricted Switching

We start this analysis by determining the optimal model for each individual in each period with a recursive estimation of the models specified above. Our approach consists of recursively computing the SSE up to a period  $t$  and then compares them in a period  $t$  for each individual. This comparison is performed for all periods except for the first 4 periods. Therefore, we can determine which model best fits the actual forecasting series in each point in time and whether there is any switching observed among these models. As many models' predictions are very similar at least in some episodes, we assume that there is no switching if the model that performs best in the previous period is not outperformed in the current period by 0.1 percentage points in terms of forecasts accuracy or 0.01 in terms of SSE. The rationale behind this choice is that the majority of forecasts are reported to one decimal point accuracy and subjects are not able to differentiate between these competing models. In Table 7 we report the relative shares of the usage of each forecasting model considered along with descriptive statistics for inflation.

Insert Table 7 about here

In Table 7 it can be seen that on average around 17% of cases subjects use the general model, and in about 12% of all forecast decisions they behave in accordance with the sticky information type model. The remaining third of all forecasts are best explained with some sort of backward-looking models. In specific, around 14% of cases subjects use simple trend extrapolation rules while the remaining 20% of cases they behave in an adaptive

manner. If we compare Table 7 with Table 6 we can observe that higher proportion of all forecasts are made using one of the stochastic gradient learning algorithms when allowing for switching. Depending on the treatment, in 23 to 45% of all cases agents use these algorithms to forecast (as shown in Table 7). If we average this across groups, 36.7% of the forecast decisions are best explained with adaptive learning. This means that, on average, adaptive learning is the most popular way of forming beliefs. Compared to the results outlined above for the "average" best model, we can immediately observe that there is approximately the same proportion of backward-looking cases as there are subjects that use backward-looking rules. However, when allowing for switching there are more forecast decisions made in an adaptive way. Also model (M10) is only a predominant model for one subject, but when we allow for switching it is used on average in 15.6% of all forecasts.

In Table 7 we can also observe that there are considerable differences among treatments with respect to the average usage of forecasting models. While we postpone this discussion to the next section, we can immediately observe that in more volatile treatments (groups) there is a higher proportion of trend extrapolative models (M5) and a lower proportion of sticky information type models (M2).

Generally, we can observe that when we allow subjects to switch between different models, they are in fact using alternative models to forecast. Under this assumption, agents use between 1 and 7 different models (average number of models used for forecasting is 6.5) and they on average switch every 4 periods. However, switching is occurring less frequently in treatments 3 and 4 compared to treatments 1 and 2 (significant at 5% level with different tests of equality of medians).<sup>26</sup> Only one subject did not switch between models. Overall, these results support the idea of intrinsic heterogeneity that is theoretically modelled in Branch and Evans (2006) and Pfajfar (2008).

To further analyze the degree of heterogeneity in the data, we compute the average number of models used in each period across subjects. We find that on average 4.5 different models (between 2 and 7) are used within a group in each period. This additionally supports the above conjecture that heterogeneity is pervasive as there are no significant differences across treatments. The average number of models employed for forecasting within a group varies (in each period) only between 4.2 and 5.3. Furthermore, there is no "smoothing" employed across different subjects in the same group. We have only employed some "smoothing" within each subject as some models perform quite similarly and cannot be differentiated at one decimal point accuracy.

We also investigate the pattern (timing) of switching with panel probit and logit models (with random, and fixed effects, and population averages), where dependant variable,

---

<sup>26</sup>With e.g. Kruskal-Wallis test. Switching is occurring on average every 6.1 periods in treatment 4, 3.7 periods in treatment 3, 2.6 period in treatment 2, and 2.9 periods in treatment 1.

$z_t^k$ , is 1 when switching occurs and 0 otherwise. We estimate the following regression:

$$z_t^k = \alpha_1^k + \alpha_2 \pi_{t-1} + \alpha_3 y_{t-1} + \alpha_4 i_{t-1} + \alpha_5 (\pi_{t-1} - \pi_{t-1|t-2}^k)^2 + \varepsilon_t^k. \quad (12)$$

We find that subjects decide to switch according to the developments of inflation, output gap, and interest rate. Alternative models exhibit similar effects of the explanatory variables. The most pronounced effect expectably comes from the output gap which has a strong negative impact on the probability of switching. Positive change in inflation trend increases the probability of switching, however, higher inflation decreases it. This demonstrates that there exists a certain pattern in the structure of individual switching. There are also some differences across treatments, especially in treatment 4 the pattern of switching is different. However, treatment dummies are insignificant if we insert them to the above regression. Results are reported in Table 8.<sup>27</sup>

Insert Table 8 about here

## 6 Monetary Policy in the Presence of Heterogeneous Expectations

### 6.1 Inflation Variability and Monetary Policy

As we outlined in the introduction, a crucial question in the environment where expectations are heterogeneous is whether policy prescriptions under a rational expectations assumption remain intact in this environment. We start addressing this question by showing that different monetary policy rules produce different variability of inflation. Woodford (2003) pointed out that within a standard NK model monetary policy should minimize variance of inflation and output gap as this corresponds to maximizing utility of consumers. Therefore we start this section with the analysis of variance of inflation as the monetary authority cares only about inflation in instrumental rules under scrutiny. Tests for differences in medians across treatments where the null hypothesis that the medians are the same in all treatments is rejected at 1% significance with Kruskal-Wallis and van der Waerden tests (see Conover, 1999). Therefore, we can argue that the design of monetary policy matters in our framework. The following table shows the pair-wise comparison of median standard deviations of inflation in treatments 2, 3, and 4 with the baseline treatment 1. We report p-values of the Kruskal-Wallis test.<sup>28</sup>

---

<sup>27</sup>In this case it is not straightforward whether in the estimation procedure for the standard errors to allow for intragroup correlation or intratreatment correlation. In the main text we report standard errors that are clustered in groups and in the Table A2 in the Appendix A standard errors that are clustered in treatments.

<sup>28</sup>Other nonparametric tests perform very similarly.

Insert Table 9 about here

We can immediately observe that treatment 3 and 4 produce lower variation of inflation compared to treatment 1. In addition it is sensible to make the comparison between treatments 2 and 3. We also find that there is a significant difference between these two treatments (p-value is 0.0250). Thus, we can argue that treatments 3 and 4 produce significantly lower inflation variability than treatments 1 and 2. Now that we establish that there is a difference in variance of inflation between treatments, we further analyze the roots of these differences between and within treatments. We graph the evolution of inflation across treatments in Figure 2.<sup>29</sup>

Insert Figure 2 about here

## 6.2 Expectations, Monetary Policy and Equilibrium Outcomes

There are several potential explanations for the differences outlined above and observed cyclical behavior of inflation: stability properties of different models (monetary policy rules) that we employ in the four treatments, the forecasting behavior of subjects and the interaction between both features. We first establish the differences in properties of (temporary) equilibria under different expectation formation mechanisms and then empirically assess the relationship between monetary policy and expectations.

Insert Table 10 about here

To have an illustration of how important are expectations for the stability of the system we simulate our treatments with different forecasting rules under the assumption of homogeneous expectations (see Figures A4 and A5). We can immediately observe that adaptive expectations (M3; with a gain coefficient higher than 1) and trend extrapolation rules (M5) can lead to pronounced cyclical variability of inflation. It is also possible to observe that treatments 2 and 4 perform better than 1 and 3 in stabilizing those expectation formation mechanisms. However, the evidence is reversed with respect to "stable" expectation formation mechanisms. To further understand the stability of the system under different expectation formation mechanisms in different treatments we analyze determinacy and E-stability (see Table 10) under several expectation formation mechanisms that were empirically studied in the previous sections.<sup>30</sup> There are several

---

<sup>29</sup>Detailed Figures with the evolution of inflation and inflation forecasts in each treatment are reported in Appendix A (Figures A2 and A3).

<sup>30</sup>Where possible, we compute the equivalent of the MSV REE solution of the model under a particular expectation formation mechanism. Where this is not possible (optimal coefficients would be 0), we approach this from the perspective of a temporary equilibrium, as in the case of naive expectations. Similar approach would have to be used if we would calculate properties for sticky information type model and adaptive expectation, however here it is not straightforward how to choose the parameter values (and as one can see in Figures A4 and A5 the results will depend on these assumptions).

potential theoretical reasons why we might observe excessive volatility: indeterminacy, E-instability, underparametrization of the PLM, rules with non-optimal coefficients, and adaptive learning with constant gain. Let us point out that in order the system is locally E-stable all the eigenvalues of the t-map associated with a particular solution (see Evans and Honkapohja, 2001) have to be less than 1. We can observe that this is never the case for the trend extrapolation rule (M5), naive expectations, and the general model (M6). Furthermore, complex eigenvalues suggest that convergence is cyclical (and the higher the eigenvalues are the higher the variability will be) as pointed out in Marimon and Sunder (1995). Another source of instability could be indeterminacy. We see that this arises when expectations are formed using trend extrapolation rules (except in treatment 3) and using the general model (M6). However if we look at the Figures A4 and A5, we see that under trend extrapolation rule treatment 3 is potentially performing worst. On the one hand, we confirm the results from Marimon and Sunder (1995) that stability of the system provides a good explanation of the inflation volatility, but only with respect to the stable expectation formation mechanisms. On the other hand, for adaptive expectations (M3) and trend extrapolation rules (M5) the link between the underlying stability of the model and inflation variability might well be reversed. This will be detailed in the following subsections. Generally, we can conclude that the stability and determinacy of the system crucially depend on the expectation formation mechanism. The system that is E-stable and determinate under RE might not be under different expectation formation mechanisms.

### 6.3 Inflation Variability and Expectations

One could argue that potentially there are differences across treatments in both the proportion of subjects using a particular forecasting rule and also in parameter estimates associated with a particular rule. We first briefly assess the latter while majority of this section will then be devoted to the former explanation.

As we saw above, parameter estimates of certain rules, especially trend extrapolation and adaptive expectations, can produce very different equilibrium dynamics. First, it would be interesting to see whether the average coefficient of trend extrapolation rule ( $\tau_1$ ) in M5 is higher in more volatile treatments. Indeed, it is the highest in treatment 1 (0.53) and the lowest in treatments 3 (0.38). Also sticky information type models (M2) exhibit significant differences across treatments. Groups in treatment 3 had the highest average  $\lambda_1$  (0.37), while those in treatment 2 had the lowest (0.11). Therefore in treatment 3 these expectation rules produce a more stabilizing effect than in treatment 2. Similar evidence to that for trend extrapolation rules is also found for adaptive expectations (M3) models. These rules with a gain coefficient larger than 1 represent another threat for stability. Here we can conclude that these differences are important (and some-



times statistically significant), however the proportion of agents using a particular rule is potentially even more important to explain the observed results. The second issue that is worth to investigate is the relationship between the gain parameter in adaptive learning PLMs and the stability of the system as in e.g. Marcet and Nicolini (2003). The theory suggests that constant gain learning produces higher variability of the underlying series compared to decreasing gain learning. Furthermore, the variability increases in the level of the (constant) gain parameter. Therefore, one could expect that average gains would be higher in more volatile treatments. When empirically assessing this conjecture, we find evidence of higher average (and median) gains for more stable treatments (3 and 4) compared to more volatile treatments (1 and 2). Therefore we could not find support for this conjecture. This is most likely because adaptive learning becomes more important when "fine tuning" expectations. When the underlying series becomes too volatile subjects tend to switch to backward looking expectations, especially trend extrapolation rules. We can conclude that constant gain learning does not represent an important source of the difference between volatilities in different treatments.

As it was mentioned before, the proportion of especially trend extrapolation subjects plays a particularly important role for the stability of the system. We can observe that there is a considerable degree of heterogeneity across treatments (see Table A1). Even more, results from Table 7 suggest that differences in the proportion of trend extrapolation subjects can explain the differences in variability between groups in the same treatment. The results are intuitive as we find that there is a strong correlation between the variability of inflation and the degree of trend extrapolation behavior. We further test these conjectures regarding the relationship between the variability and the proportion of different groups of subjects with cross-sectional and panel data regressions. With former, while estimating the below regression, we find that especially the increasing proportion of trend extrapolation behavior is increasing the variance. Also the increasing proportion of CGL adaptive expectations rules is increasing the variance as most of the estimated coefficients  $\vartheta$  in equation (M3) are higher than 1 while the proportion of recursive learning (M10) and also sticky information rules (M2) is reducing it. We estimate the following regressions:

$$sd_s = \eta_0 + \eta_1 p_{js} + \varepsilon_s, \quad (13)$$

where  $sd_s$  is a standard deviation of inflation in group  $s$ , and  $p_{js}$  is the proportion of agents using  $j$ -th model for forecasting in group  $s$ . The set of alternative models is the same as in Table 6 above. Regression results are reported in Table 11, both with robust and clustered standard errors. Initially, we added treatment dummies to the above regression; however they were insignificant in almost all cases.<sup>31</sup> We have to point out that all estimated coefficients (that are significantly different than 0) have the expected signs.

---

<sup>31</sup>Thus we decided to report them only for the regression (14). See Table A2.

Insert Table 11 about here

These results are confirmed also with the system GMM estimator of Blundell and Bond (1998) for dynamic panels. To construct the panel we compute the  $sd_{s,t}$ , standard deviation from the first period up to period  $t$ . Using the switching analysis we similarly compute  $p_{js,t}$ , the share of model  $j$  in group  $s$  up to the period  $t$ . We estimate the following model:

$$sd_{s,t} = \eta_0 + \eta_L sd_{s,t-1} + \sum_j \eta_j p_{js,t} + \varepsilon_{st}. \quad (14)$$

Results are reported in Table 12. Coefficient on the proportion of the general model (M6) gives inconclusive results in the above regressions. Therefore it is difficult to say from this analysis what is the effect of the "popularity" of general model (M6) to the stability of inflation. Although these agents use all relevant information to forecast inflation simulation exercise shows that at low values of  $\gamma$  this forecasting model (if used exclusively) will result in a high variability of inflation (see Figure A6), which confirms the analysis in Table 10. Furthermore, as explained above, if one uses the general model (M6) for forecasting the model exhibits indeterminacy, i.e. there might be a multiple equilibria problem.

Insert Table 12 about here

Results in Tables 11 and 12 show that the proportion of subjects that use trend extrapolation rules exert a strong effect to the standard deviation of inflation, thus robust to different techniques used in these tables. The proportion of these agents probably plays the most important role for the stability of inflation. It also helps us to explain the differences among groups within the same treatment. Generally, we note that the group with lower proportion of trend extrapolation rules is more stable compared to other groups in the same treatment. To further study treatment effects we add treatment dummies to the specification in eq. (14). Result are reported in Table A2 in Appendix, where we can observe that treatment 4 (inflation targeting) produces additional effects to those explained above, which lead to significantly lower variability of inflation compared to treatment 1.

## 6.4 Discussion

The relationship between the underlying model and the expectation formation has been recently studied in Heemeijer, Hommes, Sonnemans, and Tuinstra (2009). They compare experimental results in positive and negative expectations feedback models.<sup>32</sup> In a positive expectations system, e.g. asset pricing model, they observe a similar aggregate

---

<sup>32</sup>Also Fehr and Tyran (2008) compare the two environments, although in a different context.

behavior to ours and note that when there is a stronger positive feedback more agents resort to backward-looking, especially trend following rules. In our case, by changing the monetary policy, we augment the degree of positive feedback from inflation expectations to current inflation. Therefore, the design of monetary policy is important for the prevailing expectation formation mechanism and vice versa, as can be seen if we compare results within the same treatment.

However, this is only a part of the story in our experiment. We expected that the treatment 2 where monetary authority does not react too strongly to inflation expectations ( $\gamma = 1.35$ ) performs better regarding the stability of inflation than the benchmark treatment, although the theory under rational expectations suggests that higher  $\gamma$  leads to lower variability of inflation. This is not confirmed in our analysis above as the median standard deviation is not statistically different than in treatment 1. This might be due to expectations of cycles by some individuals in groups 4 and 5 of this treatment and extensive use of strong trend extrapolation rules at the beginning of the experiment. In order to study the relationship between  $\gamma$  and the variance of inflation under different expectation formation mechanisms we design simulation exercises that exactly replicate the design, parameterization and shocks employed in the experiment, we only vary  $\gamma$  and study the variability of inflation (see Figure A6). On the one hand, when all subjects have rational expectations we confirm the theory that higher  $\gamma$  leads to lower variability of inflation; similar evidence holds for other expectation formation mechanisms that we labelled above as stable, although the decreasing variability might sometimes be non-monotonic. On the other hand, several expectation formation mechanisms produce a U-shaped behavior of the inflation variance. Especially trend extrapolation rules will lead to the U-shaped behavior and eventually higher variability when increasing  $\gamma$ . The minimum variability of inflation with sticky information and trend extrapolation rule is achieved at  $\gamma = 1.1$  (see Figure A6). For naive expectations the minimum is around  $\gamma = 3$  (non-monotonic U-shaped). This can be also observed from Figures A4 and A5. As already mentioned in section 6.2, more stable models under REE have the potential to produce more volatile inflation when unstable expectation formation mechanisms prevail. Therefore, the relationship between the variability of inflation and different rules is nontrivial and the question whether treatment 2 should produce lower variability compared to treatment 1 depends particularly on the proportions of alternative rules used. Based on simulation results and observed behavior of individuals we can argue that in the presence of heterogeneous expectations instrumental rules that are less aggressive have the potential to produce lower variability of inflation, however there is the risk that, e.g. after a shock, the amplitude of inflation increases significantly as monetary policy is not aggressive enough. Thus, one could argue that non-linear Taylor-type rules would perform best in this environment, although the literature in monetary economics has not devoted much attention to this type of instrumental rules.

As we have seen so far, the expectational feedback is not the only source of instability in our underlying model.<sup>33</sup> Treatment 3 produces lower variability of inflation compared to treatment 2, but in the former case the frequency of cycles is significantly higher as the monetary authority is (too) strongly responding to deviations from inflation target. After some threshold of response to inflation forecast (depends on the proportion of agents using each rule) the resulting amplitude of the inflation variability decreases, while the frequency of cycles increases. The latter makes it more difficult to forecast and more participants resort to simpler underparametrized rules. Using simulations explained above we can identify two effects of increasing  $\gamma$  on the variability of inflation: *i*) it always increases the frequency of cycles irrespective of the expectation formation mechanism and *ii*) it increases or decreases the amplitude of the cycle depending on the level of  $\gamma$  and the expectation formation mechanism. It can produce non-monotonic or even U-shaped responses of variability, except for rational expectations where it decreases monotonically (see Figure A6).

Also treatment 4 produces lower variance of inflation than the benchmark treatment as it can be observed in Table 9. Responding to contemporaneous inflation (as in treatment 4) turns out to be a better practice for a central bank in our experiment compared to responding to inflation expectations,<sup>34</sup> although the stability properties of the two models under different expectation formation mechanisms (see Table 10) are quite similar (and treatment 1 might be even in a slight advantage). Moreover, this treatment resembles quite closely the behavior of survey forecasts, as there are periods when subjects systematically overpredict inflation (low and stable inflation) and underpredict inflation (when inflation is high). This is evident in Figure A3 in the appendix A. In this treatment there is the highest proportion of biased agents and also results from the general model suggest similar behavior of these agents to the results obtained in the survey data literature. Moreover, if we compare the means of inflation forecasts in treatments 1 and 4 we find that the mean of inflation forecasts of groups in treatment 4 is significantly higher than the mean of inflation forecasts of groups in treatment 1 (at 10% significance with Kruskal-Wallis test). Also average inflation in treatment 4 is higher (3.10 in treatment 4 compared to 3.00 in treatment 1), however the difference is statistically insignificant with nonparametric tests. Comparison between treatments 1 and 4 implies that significantly lower standard deviation of inflation (and inflation forecasts) for treatment 4 (see Table 9) comes at a "cost" of higher inflation expectations (and possibly inflation). This result is similar to Bernasconi and Kirchkamp (2000) as they suggest Friedman's money growth rule produces less inflation volatility, but higher average inflation compared to constant

<sup>33</sup>Generally, as  $\gamma$  is increasing the positive feedback is decreasing. If we compare treatments 1 and 4, we find that expectational feedback is lower in treatment 4.

<sup>34</sup>Pfajfar and Santoro (2008) and Muto (2008) reach similar conclusion in different versions of the NK model: Muto (2008) in case when agents learn from central banks' forecasts, while Pfajfar and Santoro (2008) when they introduce the cost channel and capital market imperfections.

real deficit rule. We can also observe that the variability of inflation is generally lower than the variability of inflation expectations. Intuitively, this provides an explanation to the result that responding to current inflation stabilizes the system in a more efficient way compared to reacting to expected inflation. In booms monetary policy overreacts less in treatment 4 than in the case when interest rate is set to respond to expected inflation (in presence of trend extrapolating subjects). At the root of this pattern is that trend extrapolating subjects do not observe the informational content of output gap and do not predict the change in the growth rate of inflation when approaching the top of the boom. They still expect that inflation will accelerate as in the last few periods. Then, if the monetary authority is reacting with respect to the expected inflation, they do not change the stance of monetary policy in time. The economy is pushed in the recession where the backward-looking agents underpredict inflation and the recession is more severe than if all agents were rational. The whole process repeats in the next cycle. Again, we have to point out that the causality goes in both directions as the proportion of trend extrapolating subjects depends on the design of monetary policy (degree of aggressiveness) and also the stability of the economy is influenced by the proportion of subjects using these rules.

Adam (2007) obtains a similar dynamic pattern of inflation and inflation expectations, especially to our treatment 3. He argues that the cause for observed behavior is the subjects' reliance to simpler underparametrized rules for forecasting inflation. Thus, he characterizes the dynamics of inflation as a restricted perception equilibrium, as inflation exhibits excessive volatility around its REE. Our paper supports his findings as some agents do not take into account output gap when forecasting. However, we also show that the volatility of inflation depends on the way monetary policy is designed and conducted. We argue that the proportion of backward-looking subjects plays an important role, especially those that use trend extrapolation rules.<sup>35</sup>

## 7 Conclusion

In this paper we present results from a macroeconomic experiment where subjects are asked to forecast inflation. The underlying model of the economy is a simplified version of the standard New Keynesian model that is commonly used for the analysis of monetary policy. The focus is on the formation of inflation expectations and monetary policy design. In different treatments we employ various modifications of Taylor-type instrumental rules. We study the potential effectiveness of alternative monetary policy designs in a framework

---

<sup>35</sup>Also several asset pricing experiments have observed the dynamics of aggregate price exhibiting bubbles (see eg. Smith, Suchanek, and Williams, 1988 and Hommes, Sonnemans, Tuinstra, and van de Velden, 2005). Even more, Lei, Noussair, and Plott (2001) show that this can occur also in an environment where speculation is not possible. They conclude that this occurs due to systematic errors in decisions.

characterized by heterogeneous expectations. In all our treatments we observe cyclical behavior of inflation and output gap around their steady states. We find that some monetary policy designs more effectively stabilize and anchor the process of inflation expectations and thus result in lower inflation variability. However, these prescriptions are quite different to those derived under homogeneous rational agents. We find that the variability of inflation is significantly lower in the inflation targeting treatment and in the inflation forecasting treatment with a strong reaction to deviations of inflation forecasts from inflation target compared to other inflation forecasting treatments with weaker reactions. When we employ current inflation in the monetary policy rule, we observe that this dynamics of inflation expectations most closely resembles the behavior of survey data. This setup results in lower inflation variability than the setup where expected inflation is used in the policy rule. There are several reasons for that: first, the expectational feedback of the former model is lower compared to the latter one; second, lower observed variability of inflation compared to inflation expectations in conjunction with the first reason further decreases the resulting variability of inflation and output gap. Thus, we reduce the amplitude of expectational cycles. However, in this interplay between inflation variability, monetary policy, and inflation expectations the causality does not go only in one direction. We can also point out that the underlying process of inflation expectation formation depends on the way monetary policy is conducted. The proportion of backward-looking agents, especially trend extrapolating subjects, plays a crucial role for the stability of the system. In some environments it is more difficult to forecast rationally and thus more subjects resort to simpler backward-looking rules. Thus, we show that monetary policy also influences the expectation formation process.

Our results suggest that there is a significant and relatively large share of agents who behave consistently with the assumptions of rationality. The remaining agents use some version of adaptive expectations, adaptive learning or trend extrapolation rules. Most importantly, we find that expectations are heterogeneous both across subjects and across time for individual subjects. Most subjects, in fact, tend to switch between different rules. When we take into account this possibility, adaptive learning models become more important as this mechanism for forecasting is used in more than 1/3 of all forecasting decisions. This paper is one of the first empirical contributions to postulate that these models represent one of the most popular ways of forecasting inflation. Furthermore, as pointed above, both the proportions of subjects using a particular rule and the estimated coefficients in this rule are dependent on the monetary policy.

Our economy is represented by a simplified version of a commonly used model in macroeconomics, where a model is already a simplification of the reality so we have to take this into account when making policy prescriptions. Nevertheless, from our analysis it comes clear that the policy-makers should ask themselves what they can do to help consumers to anchor their expectations (or learn to form more accurate forecasts) and

design monetary policy accordingly. This is a relevant policy question that triggered quite some attention in the literature. However it is crucial to understand the interplay between monetary policy and expectations in order give answer to the above question as it is already argued in Friedman (1948 and 1960) and recently recognized also by Bernanke (2007). Our paper attempts to shed light to the understanding of this relationship that has been to a large extent unexplored in the literature so far. We argue that rules that respond to observed inflation are less likely to trigger expectational cycles and "push" agents into forming expectations in a backward-looking fashion, especially into trend extrapolation mechanism, which even further stretches the expectational cycles. We also demonstrate that the slope coefficient in the Taylor rule that reacts to deviations of inflation from its target plays a crucial role in the heterogeneous expectations environment. In some cases it might even be beneficial that the policy makers do not react too strongly to deviations of inflations, although the Taylor principle should always be satisfied.

## References

- ADAM, K. (2007): “Experimental Evidence on the Persistence of Output and Inflation,” *Economic Journal*, 117(520), 603–636.
- ANUFRIEV, M., AND C. HOMMES (2008): “Evolutionary Selection of Individual Expectations and Aggregate Outcomes,” Mimeo, University of Amsterdam.
- (2011): “Evolution of Market Heuristics,” *Knowledge Engineering Review*, forthcoming.
- ARIFOVIC, J., AND T. J. SARGENT (2003): “Laboratory Experiments with an Expectational Phillips Curve,” in *Evolution and Procedures in Central Banking*, ed. by D. E. Altig, and B. D. Smith. Cambridge University Press, Cambridge.
- ASPAROUHOVA, E., P. BOSSAERTS, W. SHIN, AND M. CHENG (2011): “Market Bubbles and Crashes as an Expression of Tension between Social and Individual Rationality: An Experiment,” Mimeo, California Institute of Technology.
- ASSENZA, T., P. HEEMEIJER, C. HOMMES, AND D. MASSARO (2011): “Individual Expectations and Aggregate Macro Behavior,” CeNDEF Working Papers 11-01, Universiteit van Amsterdam, Center for Nonlinear Dynamics in Economics and Finance.
- BAAK, S. J. (1999): “Tests for bounded rationality with a linear dynamic model distorted by heterogeneous expectations,” *Journal of Economic Dynamics and Control*, 23(9-10), 1517–1543.
- BAKHSI, H., AND A. YATES (1998): “Are U.K. Inflation Expectations Rational?,” Discussion Paper 81, Bank of England Working Paper.
- BERLEMANN, M., AND F. NELSON (2005): “Forecasting Inflation via Experimental Stock Markets Some Results from Pilot Markets,” Ifo Working Paper Series Ifo Working Paper No. 10, Ifo Institute for Economic Research at the University of Munich.
- BERNANKE, B. (2007): “Inflation Expectations and Inflation Forecasting,” Discussion paper, Speech at National Bureau of Economic Research Summer Institute, Cambridge, Massachusetts.
- BERNASCONI, M., AND O. KIRCHKAMP (2000): “Why do monetary policies matter? An experimental study of saving and inflation in an overlapping generations model,” *Journal of Monetary Economics*, 46(2), 315–343.
- BLUNDELL, R., AND S. BOND (1998): “Initial conditions and moment restrictions in dynamic panel data models,” *Journal of Econometrics*, 87(1), 115–143.



- BRANCH, W. A. (2004): “The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations,” *Economic Journal*, 114(497), 592–621.
- BRANCH, W. A., AND G. W. EVANS (2006): “Intrinsic heterogeneity in expectation formation,” *Journal of Economic Theory*, 127(1), 264–295.
- BRANCH, W. A., AND B. MCGOUGH (2008): “Replicator Dynamics in a Cobweb Model with Rationally Heterogeneous Expectations,” *Journal of Economic Behavior and Organization*, 65(2), 224–244.
- (2009): “A New Keynesian model with heterogeneous expectations,” *Journal of Economic Dynamics and Control*, 33(5), 1036–1051.
- BROCK, W. A., AND C. H. HOMMES (1997): “A Rational Route to Randomness,” *Econometrica*, 65(5), 1059–1096.
- BRYAN, M. F., AND S. PALMQVIST (2005): “Testing Near-Rationality using Detailed Survey Data,” Working Paper Series 183, Sveriges Riksbank.
- CARCELES-POVEDA, E., AND C. GIANNITSAROU (2007): “Adaptive learning in practice,” *Journal of Economic Dynamics and Control*, 31(8), 2659–2697.
- CARROLL, C. D. (2003a): “Macroeconomic Expectations Of Households and Professional Forecasters,” *The Quarterly Journal of Economics*, 118(1), 269–298.
- (2003b): “The Epidemiology of Macroeconomic Expectations,” NBER Working Papers 8695, National Bureau of Economic Research, Inc.
- CHAVAS, J.-P. (2000): “On information and market dynamics: The case of the U.S. beef market,” *Journal of Economic Dynamics and Control*, 24(5-7), 833–853.
- CONOVER, W. J. (1999): *Practical Nonparameteric Statistics*. Wiley, 3rd edn.
- COOLEY, T. F., AND E. C. PRESCOTT (1995): “Economic Growth and Business Cycles,” in *Frontiers of Business Cycle Research*, ed. by T. F. Cooley. Princeton University Press, Princeton.
- CURTIN, R. (2005): “Inflation Expectations: Theoretical Models and Empirical Tests,” Mimeo, University of Michigan.
- DUFFY, J. (2008): “Macroeconomics: A Survey of Laboratory Research,” Working Papers 334, University of Pittsburgh, Department of Economics.
- ENGLE-WARNICK, J., AND N. TURDALIEV (2010): “An experimental test of Taylor-type rules with inexperienced central bankers,” *Experimental Economics*, 13(2), 146–166.

- EVANS, G. W., AND S. HONKAPOHJA (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press.
- EVANS, G. W., S. HONKAPOHJA, AND R. MARIMON (2001): “Convergence in monetary inflation models with heterogeneous learning rules,” *Macroeconomic Dynamics*, 5(1), 1–31.
- EVANS, G. W., S. HONKAPOHJA, AND N. WILLIAMS (2010): “Generalized Stochastic Gradient Learning,” *International Economic Review*, 51(1), 237–262.
- 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.
- FISCHBACHER, U. (2007): “z-Tree: Zurich toolbox for ready-made economic experiments,” *Experimental Economics*, 10(2), 171–178.
- FRIEDMAN, M. (1948): “A Monetary and Fiscal Framework for Economic Stability,” *The American Economic Review*, 38(3), 245–264.
- (1960): *A Program for Monetary Stability*. Fordham University Press, New York.
- HAZELETT, D., AND A. KERNEN (2002): “Hyperinflation and seigniorage in an experimental overlapping generations economy,” Mimeo, Whitman College.
- HEEMEIJER, P., C. HOMMES, J. SONNEMANS, AND J. TUINSTRA (2009): “Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation,” *Journal of Economic Dynamics and Control*, 33(5), 1052–1072.
- HOMMES, C. (2011): “The heterogeneous expectations hypothesis: Some evidence from the lab,” *Journal of Economic Dynamics and Control*, 35(1), 1–24.
- HOMMES, C., H. HUANG, AND D. WANG (2005): “A robust rational route to randomness in a simple asset pricing model,” *Journal of Economic Dynamics and Control*, 29(6), 1043–1072.
- HOMMES, C., 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.
- IRELAND, P. N. (2004): “Technology Shocks in the New Keynesian Model,” *The Review of Economics and Statistics*, 86(4), 923–936.

- KELLEY, H., AND D. FRIEDMAN (2008): “Learning to Forecast Rationally,” in *Handbook of Experimental Economics Results*, ed. by C. R. Plott, and V. L. Smith, vol. 1, chap. 35, pp. 303–310. Elsevier.
- LEI, V., C. N. NOUSSAIR, AND C. R. PLOTT (2001): “Nonspeculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality,” *Econometrica*, 69(4), 831–59.
- MANKIW, N. G., R. REIS, AND J. WOLFERS (2004): “Disagreement about Inflation Expectations,” *NBER Macroeconomics Annual 2003*, 18, 209–248.
- MAR CET, A., AND J. P. NICOLINI (2003): “Recurrent Hyperinflations and Learning,” *American Economic Review*, 93(5), 1476–1498.
- MAR CET, A., AND T. J. SARGENT (1989): “Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information,” *Journal of Political Economy*, 97(6), 1306–22.
- 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 (1993): “Indeterminacy of Equilibria in a Hyperinflationary World: Experimental Evidence,” *Econometrica*, 61(5), 1073–107.
- (1994): “Expectations and Learning under Alternative Monetary Regimes: An Experimental Approach,” *Economic Theory*, 4(1), 131–62.
- (1995): “Does a Constant Money Growth Rule Help Stabilize Inflation: Experimental Evidence,” *Carnegie - Rochester Conference Series on Public Policy*, 45, 111–156.
- MCCALLUM, B. T., AND E. NELSON (2004): “Timeless Perspective vs. Discretionary Monetary Policy in Forward-Looking Models,” *Federal Reserve Bank of St. Louis Review*, 86(2), 43–56.
- MILANI, F. (2007): “Expectations, learning and macroeconomic persistence,” *Journal of Monetary Economics*, 54(7), 2065–2082.
- MOLNÁR, K. (2007): “Learning with Expert Advice,” *Journal of the European Economic Association*, 5(2-3), 420–432.
- MUTO, I. (2008): “Monetary Policy and Learning from the Central Bank’s Forecast,” IMES Discussion Paper Series 08-E-01, Institute for Monetary and Economic Studies, Bank of Japan.

- NOUSSAIR, C. N., D. PFAJFAR, AND J. ZSIROS (2011): “Frictions, persistence, and central bank policy in an experimental dynamic stochastic general equilibrium economy,” Discussion Paper 2011-030, Tilburg University, Center for Economic Research.
- NUNES, R. (2009): “Learning The Inflation Target,” *Macroeconomic Dynamics*, 13(02), 167–188.
- ORPHANIDES, A., AND J. C. WILLIAMS (2005a): “The decline of activist stabilization policy: Natural rate misperceptions, learning, and expectations,” *Journal of Economic Dynamics and Control*, 29(11), 1927–1950.
- PESARAN, M. H. (1987): *The Limits to Rational Expectations*. Basil Blackwell, Oxford, reprinted with corrections 1989 edn.
- PFAJFAR, D. (2008): “Formation of Rationally Heterogeneous Expectations,” Mimeo, University of Tilburg.
- PFAJFAR, D., AND E. SANTORO (2008): “Credit Market Distortions, Asset Prices and Monetary Policy,” Cambridge Working Papers in Economics 0825, Faculty of Economics, University of Cambridge.
- (2010): “Heterogeneity, learning and information stickiness in inflation expectations,” *Journal of Economic Behavior & Organization*, 75(3), 426–444.
- PFAJFAR, D., AND B. ŽAKELJ (2011): “Uncertainty and Disagreement in Forecasting Inflation: Evidence from the Laboratory,” Mimeo, University of Tilburg.
- ROOS, M. W., AND W. J. LUHAN (2008): “Are Expectations Formed by the Anchoring-and-adjustment Heuristic? An Experimental Investigation,” Ruhr Economic Papers 0054, Rheinisch-Westfälisches Institut für Wirtschaftsforschung.
- SMITH, V. L., G. L. SUCHANEK, AND A. W. WILLIAMS (1988): “Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets,” *Econometrica*, 56(5), 1119–51.
- WALSH, C. E. (2003): *Monetary Theory and Policy*. The MIT Press, second edn.
- WILLIAMS, A. W. (1987): “The Formation of Price Forecasts in Experimental Markets,” *Journal of Money, Credit and Banking*, 19(1), 1–18.
- WOODFORD, M. (1996): “Loan Commitments and Optimal Monetary Policy,” *Journal of Monetary Economics*, 37(3), 573–605.
- (2003): *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

# Tables and Figures

$\beta = 0.99$	$\bar{\pi} = 3$	$\nu = 0.6$
$\varphi = 0.164$	$\lambda = 0.3$	$\kappa = 0.6$

Table 1: McCallum-Nelson Calibration

Treatment	Groups	Taylor rule (equation)	Parameters
Inflation forecast targeting (1)	1-6	Forward looking (3)	$\gamma = 1.5$
Inflation forecast targeting (2)	7-12	Forward looking (3)	$\gamma = 1.35$
Inflation forecast targeting (3)	13-18	Forward looking (3)	$\gamma = 4$
Inflation targeting (4)	19-24	Contemporaneous (4)	$\gamma = 1.5$

Table 2: Treatments

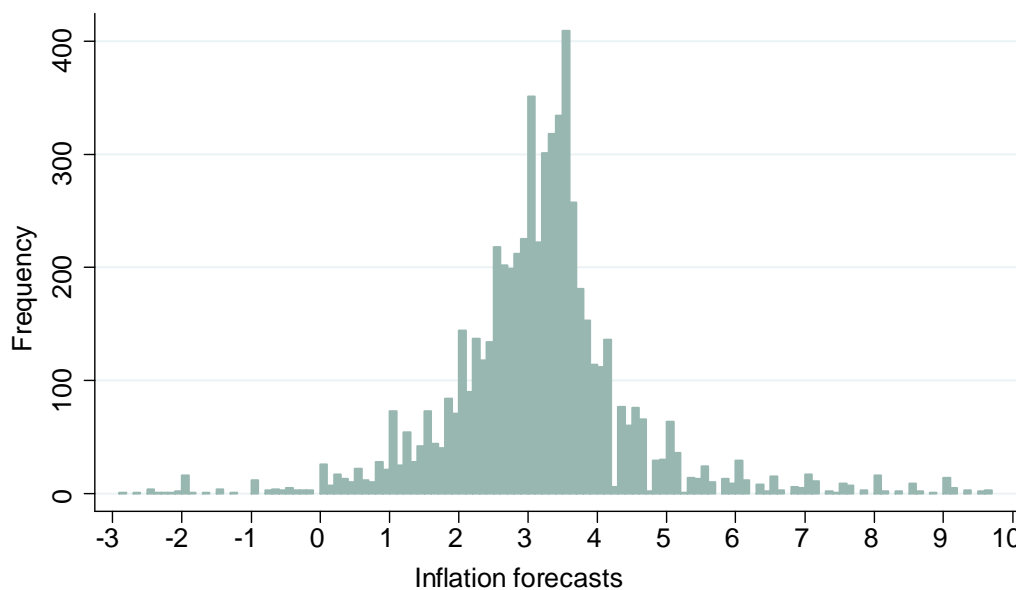


Figure 1: Histogram of inflation forecasts for all treatments.

Statistics by groups	Treatment 1						Treatment 2						Treatment 3						Treatment 4						
	Inflation forecast targeting, $\gamma=1.5$						Inflation forecast targeting, $\gamma=1.35$						Inflation forecast targeting, $\gamma=4.0$						Inflation targeting, $\gamma=1.5$						
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	Avg
Mean	2.94	3.00	3.04	3.01	3.12	3.14	3.11	3.09	3.12	3.18	2.72	3.04	3.02	3.03	3.01	3.00	3.00	3.00	3.12	3.29	3.07	3.05	3.10	3.15	3.06
StdDev	6.31	3.48	2.02	0.73	1.12	0.93	0.75	1.87	0.49	5.77	3.76	0.86	0.57	1.07	0.26	0.30	0.30	0.23	0.38	0.89	0.48	0.36	0.54	1.42	2.20
Min	-13.9	-6.10	-2.50	0.40	0.30	0.50	1.00	-0.70	0.20	-12.0	-8.80	0.52	1.70	0.00	2.00	1.20	2.10	2.40	2.30	1.00	1.60	2.30	0.50	0.00	-13.9
Max	24.0	52.0	7.50	3.98	5.40	5.20	4.50	9.50	4.20	16.1	10.5	4.50	4.76	6.90	3.80	4.50	4.00	3.70	4.20	5.20	4.00	3.90	4.40	7.00	52.0
	Inflation																								
Mean	2.85	2.88	2.92	3.00	3.13	3.12	3.12	3.09	3.13	3.02	2.52	3.03	3.01	3.02	2.99	3.00	2.99	3.01	3.09	3.23	3.05	3.05	3.09	3.11	3.02
StdDev	5.83	2.89	1.95	0.75	1.09	0.90	0.76	1.81	0.51	5.50	3.56	0.88	0.51	0.94	0.24	0.26	0.31	0.24	0.39	0.81	0.48	0.38	0.51	1.28	1.37
Min	-9.53	-5.27	-0.84	0.67	0.81	1.16	1.26	0.06	1.84	-9.04	-6.74	0.80	2.00	0.97	2.49	2.41	2.48	2.51	2.40	1.77	1.88	2.46	1.77	0.68	-9.53
Max	16.7	10.5	6.51	3.89	5.01	4.78	4.38	7.42	3.98	12.6	8.17	4.13	3.84	5.28	3.44	3.46	3.74	3.56	3.78	4.49	3.62	3.70	4.00	5.46	16.7

Table 3: Preliminary statistics

Sum of square errors	Treatment 1						Treatment 2						Treatment 3						Treatment 4						
	Inflation forecast targeting, $\gamma=1.5$						Inflation forecast targeting, $\gamma=1.35$						Inflation forecast targeting, $\gamma=4.0$						Inflation targeting, $\gamma=1.5$						
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	Avg
Subjects min	524	112	29.6	3.92	9.87	8.83	4.89	22.8	3.54	133	155	4.26	7.47	40.9	2.48	3.56	3.77	2.87	2.72	10.9	2.21	2.12	3.55	25.9	-
Subjects max	2354	1812	83.0	14.2	27.2	32.6	37.5	76.4	14.4	908	474	12.7	30.8	80.6	10.3	15.1	7.12	4.97	4.84	59.5	8.70	5.70	16.3	120	-
Subjects mean	1050	352	59.8	6.07	18.4	16.7	10.0	40.8	6.32	522	219	6.02	15.4	61.0	5.38	5.85	5.73	4.28	3.46	24.2	3.78	3.42	6.67	55.4	-
Sticky info. (M2)	2110	1317	270	33.8	61.2	40.4	38.1	268.1	14.4	1720	1017	40.4	11.5	32.3	3.12	4.04	5.39	3.25	8.92	77.7	14.8	9.88	16.4	222	305
Gen. mod. (M6), $\zeta=0$	881	355	67.8	5.84	15.7	13.8	6.75	59.3	5.83	451	315	6.70	8.40	16.5	2.37	3.21	3.63	2.23	3.77	34.0	3.30	3.59	5.87	60.2	97
Trend ext. (M5)	558	184	27.0	5.73	9.59	9.31	7.77	23.9	5.83	260	158	7.00	7.81	18.6	2.08	2.46	2.55	1.96	3.55	12.2	3.69	3.29	4.77	27.3	56
General model (M6)	755	310	54.2	6.15	15.2	13.2	6.89	49.1	4.82	445	246	5.67	6.67	13.6	2.52	2.44	3.07	1.99	2.59	22.4	2.85	3.76	6.48	5.3e <sup>8</sup>	88
Adaptive exp. (M3)	973	210	67.8	6.15	21.3	15.2	8.63	65.2	5.68	805	313	8.04	12.8	53.6	4.36	5.78	5.77	3.41	3.67	21.7	3.71	3.30	5.80	61.7	112

Table 5: Comparison between subjects and Classical Econometrician

model (eq.)	Comparison					
	1	2	3	4	5	6
Rational expectations: Stat (8)	28.7	42.1	-	-	-	-
Rational expectations: Theory (11)	-	-	40.7	44.9	-	-
AR(1) process (M1)	0.5	0.5	0.5	0.5	0.5	0.5
Sticky information type (M2)	6.5	5.6	4.2	3.2	10.2	6.5
Adaptive expectations (M3)	7.4	5.1	4.2	4.2	11.6	9.3
Trend extrapolation (M5)	30.1	25.5	28.2	26.9	36.6	26.9
Recursive - lagged inflation (M7)	11.6	7.9	8.8	8.3	21.8	9.3
Recursive - REE (M8)	2.8	2.3	2.8	1.9	4.2	1.4
Recursive - AR(1) process (M10)	0.5	0.5	0.5	0.5	0.5	
Recursive - trend extrapolation (M9)	12.0	10.6	10.2	9.7	14.8	12.0
General model (M6), $\zeta = 0$	-	-	-	-	-	34.3

Table 6: Inflation expectation formation (percent of subjects)

by group	Treatment 1						Treatment 2						Treatment 3						Treatment 4						
	Inflation forecast targeting, $\gamma=1.5$						Inflation forecast targeting, $\gamma=1.35$						Inflation forecast targeting, $\gamma=4.0$						Inflation targeting, $\gamma=1.5$						
	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	Avg
Gen. mod. (M6), $\zeta=0$	17.7	18.5	24.6	16.2	14.8	11.4	13.3	14.5	14.3	18.5	19.5	16.8	12.1	16.3	27.9	18.4	10.3	26.4	11.4	15.7	13.6	22.2	12.8	19	16.9
Sticky info. type (M2)	12.3	6.4	8.2	5.9	10.1	14.3	11.6	10.8	13.6	9.6	9.9	8.4	14.8	11.1	19.2	23.9	15.8	18.2	9.4	8.9	12.6	8.6	15.8	11.1	12.1
Ad. exp. CGL (M3)	16	12.6	13	15.8	15	11.4	16.7	13.8	10.8	17.8	14.3	18.2	15.2	14.3	12.5	11.1	12.6	5.1	11.1	15.8	17.5	11.4	11.6	20.4	13.9
Ad. exp. DGL (M4)	7.2	5.4	4.7	8.4	2.7	5.9	8.2	6.2	9.6	5.6	5.4	6.6	6.7	4.9	5.6	8.8	8.8	7.6	7.7	5.2	5.6	6.4	5.1	4	6.3
Trend extr. (M5)	23.6	20	14.1	15	16.2	11.8	15.5	13.8	13.8	16.5	20.9	13.6	13.3	19.2	6.1	10.4	12.6	6.4	7.6	13.1	10.1	12.5	11.8	18.4	14
Rec. AR(1) (M10)	7.9	12.1	11.1	19	19.2	12.8	19.2	9.4	20	8.4	8.8	18.9	11.1	10.8	12.5	13.5	22.2	18.7	26.1	16	29.1	19.7	19.7	7.1	15.6
Rec. trend extr. (M9)	15.3	24.9	24.2	19.7	22.1	32.3	15.5	31.5	17.8	23.6	21.2	17.5	26.8	23.4	16.3	14	17.7	17.7	26.6	25.3	11.4	19.2	23.2	20	21.1
inflation min	-9.53	-5.27	-0.84	0.67	0.81	1.16	1.26	0.06	1.84	-9.04	-6.74	0.8	2	0.97	2.49	2.41	2.48	2.51	2.4	1.77	1.88	2.46	1.77	0.68	-9.53
inflation max	16.68	10.46	6.51	3.89	5.01	4.78	4.38	7.42	3.98	12.56	8.17	4.13	3.84	5.28	3.44	3.46	3.74	3.56	3.78	4.49	3.62	3.7	4	5.46	16.68
inflation s.d.	5.83	2.89	1.95	0.75	1.09	0.9	0.76	1.81	0.51	5.5	3.56	0.88	0.51	0.94	0.24	0.26	0.31	0.24	0.39	0.81	0.48	0.38	0.51	1.28	2.05

Table 7: Inflation expectation formation (percent of all cases)



	Probit RE	Probit PA	Logit RE	Logit PA	Logit FE
Cons.	-0.2502*** (0.0836)	-0.2210*** (0.0749)	-0.4139*** (0.1449)	-0.3552*** (0.1188)	
$ \pi_{t-1} - \pi_{t-2} $	0.0422 (0.0293)	0.0402 (0.0247)	0.0661 (0.0482)	0.0639* (0.0388)	0.0545 (0.0354)
$\pi_{t-1}$	-0.0568*** (0.0219)	-0.0533*** (0.0190)	-0.0919*** (0.0345)	-0.0857*** (0.0302)	-0.076** (0.0383)
$y_{t-1}$	-0.1702*** (0.0391)	-0.1596*** (0.0381)	-0.2747*** (0.0674)	-0.2577*** (0.0623)	-0.2540*** (0.0591)
$i_{t-1}$	0.0440** (0.0181)	0.0415** (0.0161)	0.0715** (0.0286)	0.0670*** (0.0254)	0.0575** (0.0275)
$\left(\pi_{t-1} - \pi_{t-1 t-2}^k\right)^2$	0.0061 (0.0171)	0.006 (0.0143)	0.011 (0.0248)	0.0099 (0.0260)	0.0089 (0.0359)
$\ln(\sigma^2)$ (panel)	-1.5874*** (0.1996)		-0.5814*** (0.2064)		
$\sigma$ (panel)	0.4522*** (0.0441)		0.7478*** (0.0783)		
$\rho$ (panel)	0.1670*** (0.0270)		0.1453*** (0.0256)		
N	14040	14040	14040	14040	13975
Groups	216	216	216	216	215
Obs per Group	65	65	65	65	65
Wald $\chi^2(9)$	34.0	31.8	31.2	32.6	36.2

Table 8: Determinants of switching behavior. Notes: RE stands for random effects, PA population averages, while FE is for fixed effects model. Standard errors in parentheses. \*/\*\*/\*\* denotes significance at 10/5/1 percent level. Standard errors are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in groups.

Treatment	Groups	Comparison with Treatment 1 (p-value)
Inflation forc. targ. $\gamma = 1.5$	1 – 6	–
Inflation forc. targ. $\gamma = 1.35$	7 – 12	0.6310
Inflation forc. targ. $\gamma = 4$	13 – 18	0.0104
Inflation targeting $\gamma = 1.5$	19 – 24	0.0250

Table 9: Pairwise comparison of standard deviations of inflation using Kruskal-Wallis test



s.e.: method	Gen. mod. (M6), $\zeta=0$		Sticky info. (M2)		ADE CGL (M3)		ADE DGL (M4)		Trend ext. (M5)		Rec. AR(1) (M10)		Rec. trend. extr. (M9)	
	robust	cluster	robust	cluster	robust	cluster	robust	cluster	robust	cluster	robust	cluster	robust	cluster
$p_{js}$	4,6285 (-5.675)	4,6285 (-7.324)	-12.2631*** (-4.354)	-12.2631* (-4.569)	16.6381* (-8.920)	16.6381* (-6.612)	-18.7886 (-12.449)	-18.7886 (-17.302)	25.2361*** (-6.227)	25.2361*** (-7.832)	-15.6069*** (-5.392)	-15.6069 (-7.622)	2,5229 (-5.545)	2,5229 (-6.258)
cons	0,5816 (-0.8462)	0,5816 (-0.8648)	2,8509*** (-0.7462)	2,8509** (-0.8232)	-0,9502 (-1.041)	-0,9502 (-0.6043)	2,5567*** (-0.8706)	2,5567 (-1.268)	-2,1698*** (-0.7605)	-2,1698 (-0.9535)	3,7924*** (-1.043)	3,7924* (-1.504)	0,8321 (-1.311)	0,8321 (-1.586)
Observations	24		24		24		24		24		24		24	
R <sup>2</sup>	0,0192		0,1115		0,1173		0,0407		0,4989		0,3525		0,0072	

Table 11: Relation of standard deviation to certain behavioral types as dened in Table 7. Notes: Standard errors in parentheses. \*/\*\*/\*\* denotes significance at 10/5/1 percent. Standard errors take into account potential presence of clusters in treatments.

	reg1	reg2	reg3	reg4
$sd_{s,t}$	1.0147*** (0.0085)	1.0121*** (0.0073)	1.0121*** (0.0069)	1.0099*** (0.0066)
Gen. mod. (M6), $\zeta = 0$	0.0018*** (0.0007)	0.001 (0.0013)	0.0031* (0.0017)	
Sticky info. (M2)	-0.0029* (0.0016)	-0.0039 (0.0025)	-0.0018 (0.0019)	-0.0043** (0.0020)
ADE DGL (M4)	-0.0023** (0.0009)	-0.0030** (0.0013)	-0.0008 (0.0015)	-0.0027** (0.0014)
Trend Ext. (M5)	0.0067*** (0.0015)	0.0055*** (0.0018)	0.0077*** (0.0023)	0.0055*** (0.0014)
ADE CGL (M3)		-0.0011 (0.0018)	0.001 (0.0015)	
Recursive V1 (M7)		-0.0021 (0.0025)		-0.0025 (0.0018)
Recursive V4 (M9)			0.0021 (0.0025)	
cons	-0.0759* (0.0417)	0.0219 (0.1378)	-0.1895 (0.1449)	0.0373 (0.0556)
N	1560	1560	1560	1560
$\chi^2$	67328.4	54449.2	65883.1	79094.9

Table 12: Decision model's influence on standard deviation of inflation. Notes: Estimations are conducted using system GMM estimator of Blundell and Bond (1998) for dynamic panels. Standard errors in parentheses. \*/\*\*/\*\* denotes significance at 10/5/1 percent level. Standard errors are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in treatments.

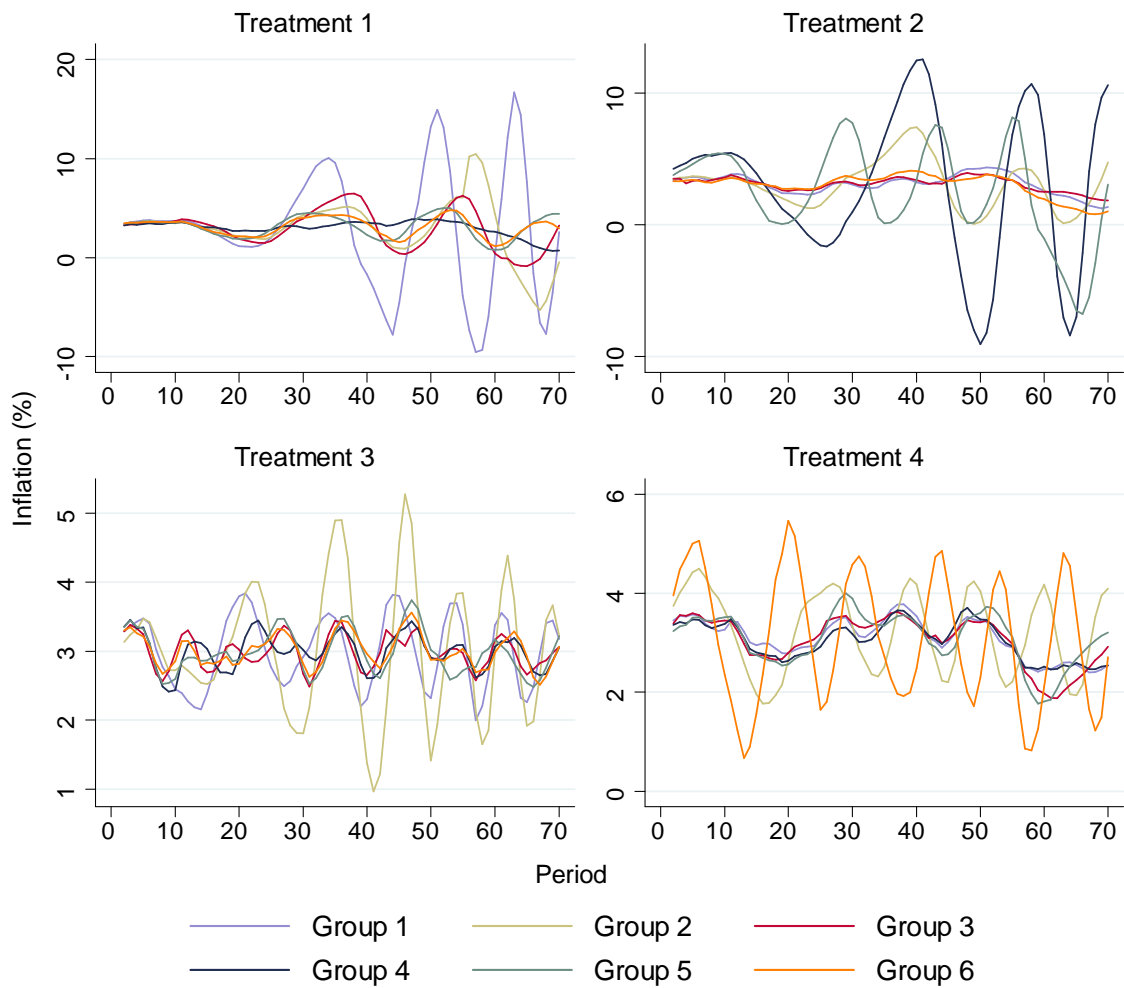


Figure 2: Group comparison of expected inflation (average subject prediction) and realized inflation by treatment. Treatment 1 has inflation forecast targeting (IFT) with  $\gamma = 1.5$ . Treatment 2 has IFT with  $\gamma = 1.35$ . Treatment 3 has IFT with  $\gamma = 4$ . Treatment 4 has inflation targeting with  $\gamma = 1.5$ .

# A Additional Tables and Figures

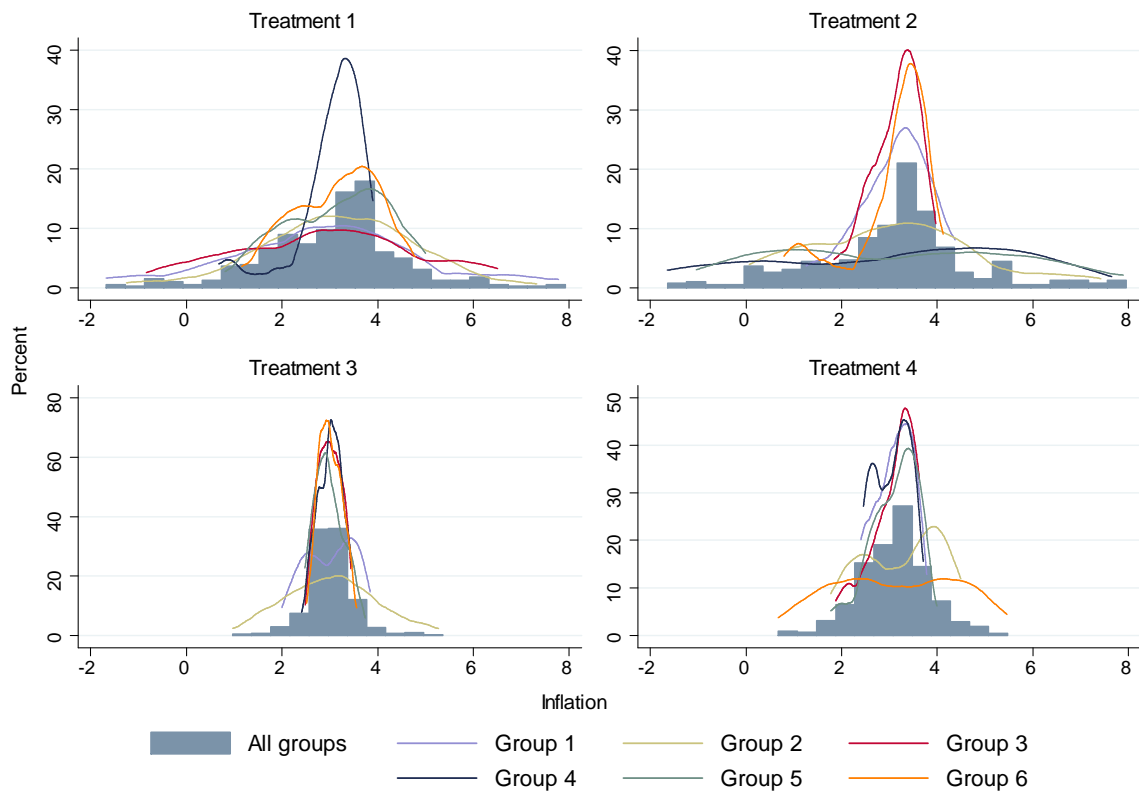


Figure A1: Histogram of individual inflation forecasts for the six independent groups in each treatment and combined

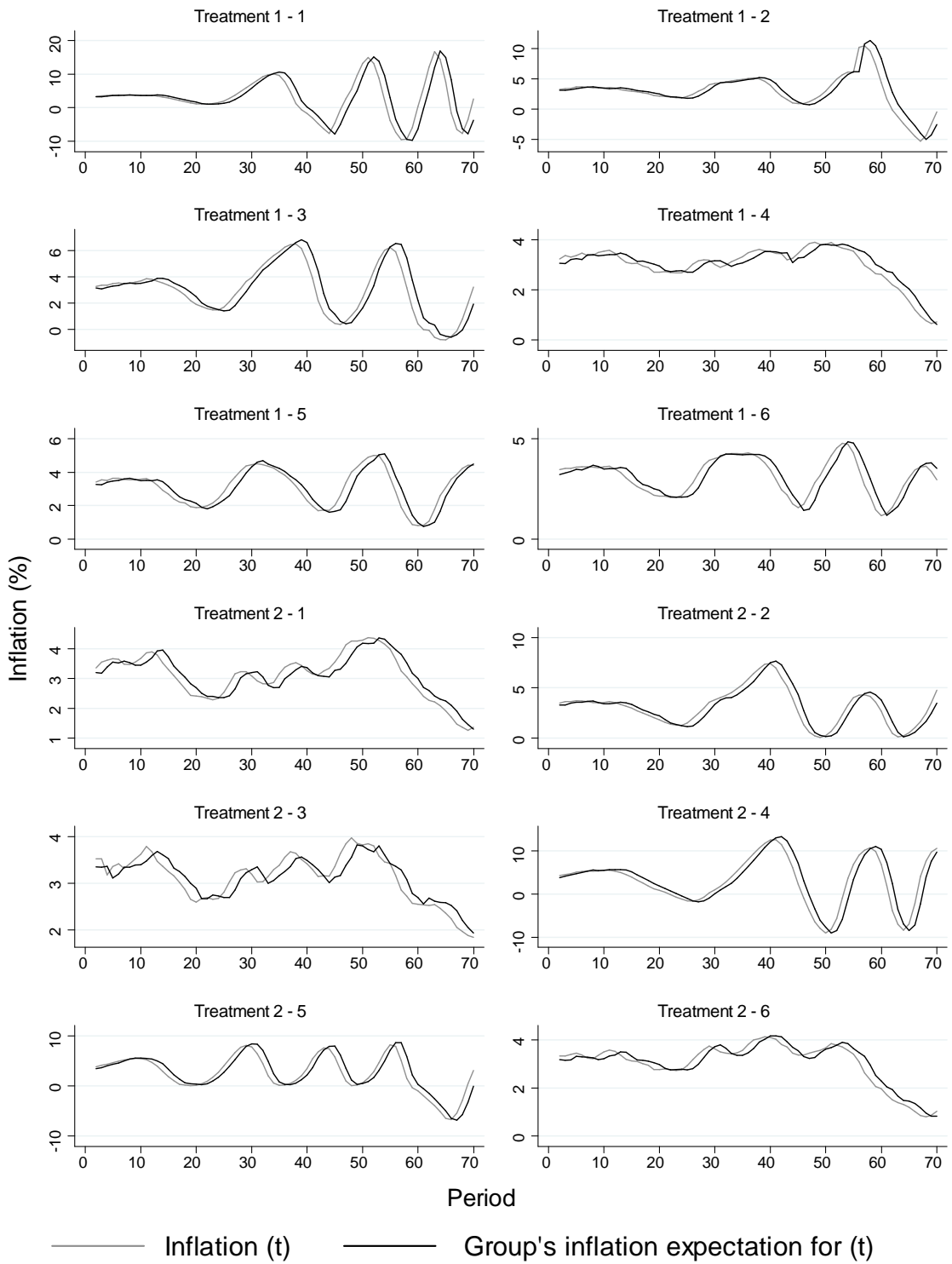


Figure A2: Inflation and inflation expectations per group, Part 1

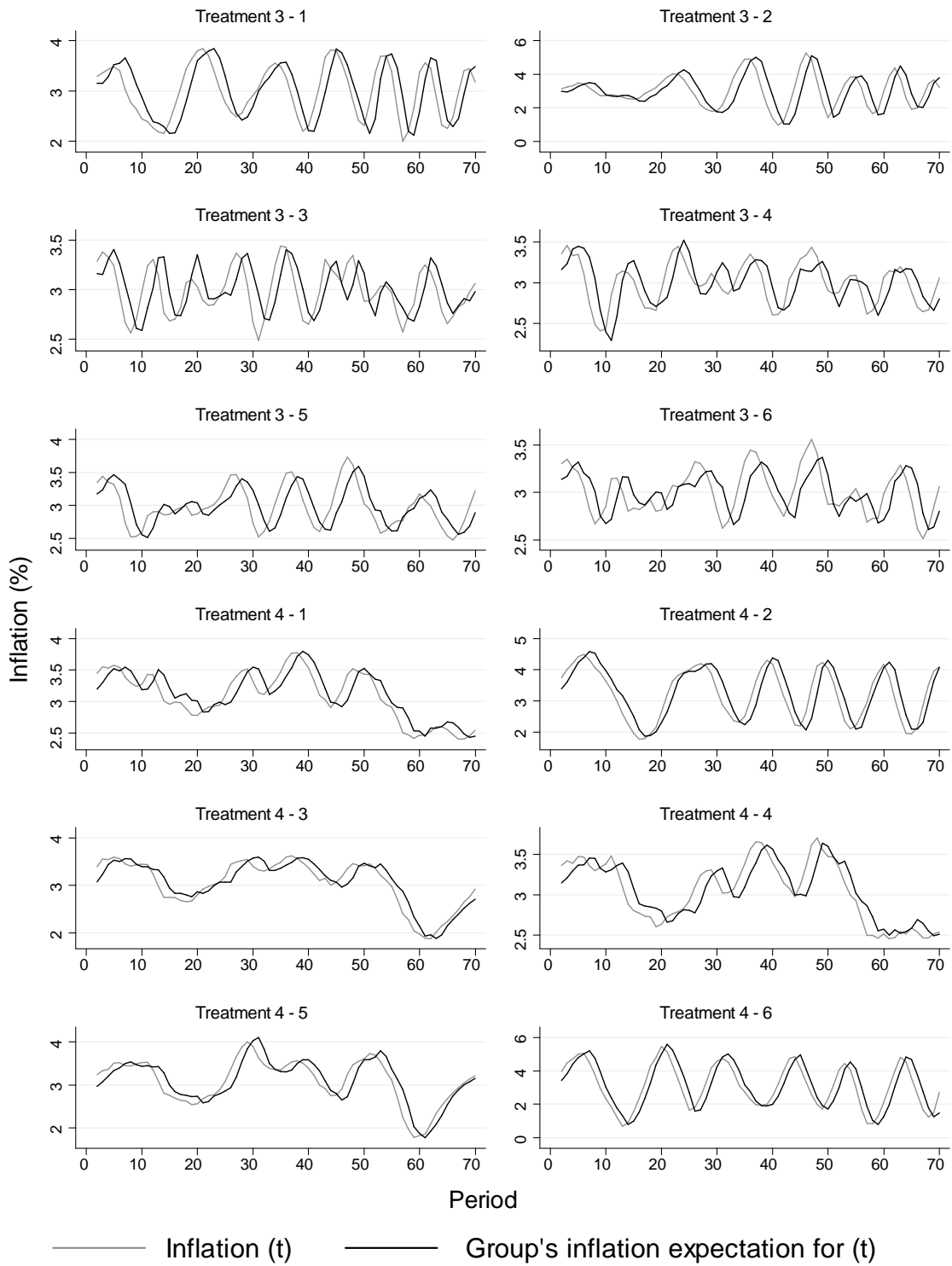


Figure A3: Inflation and inflation expectations per group, Part 2



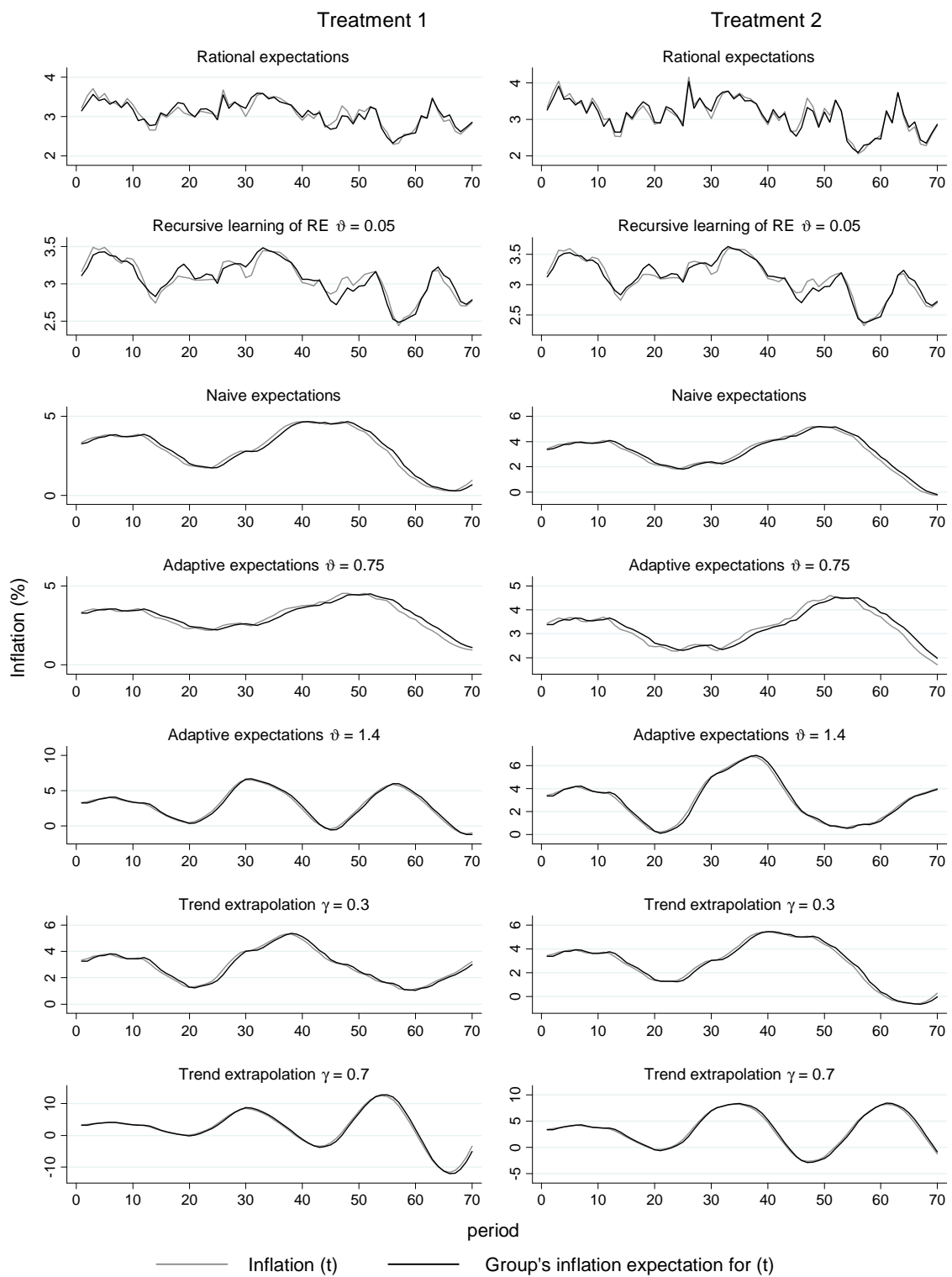


Figure A4: Alternative expectation formation rules (treatments 1 and 2).

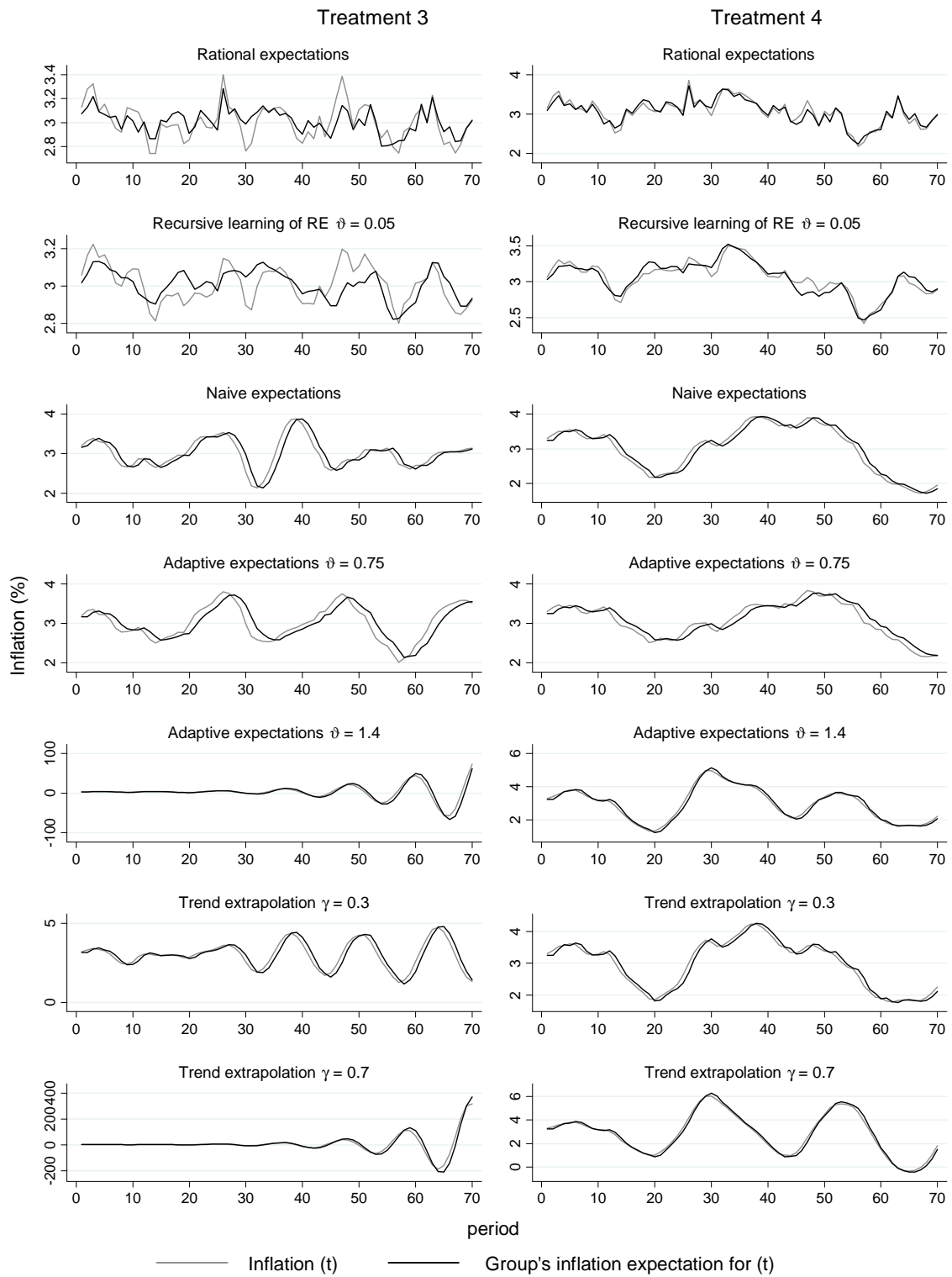


Figure A5: Alternative expectation formation rules (treatments 3 and 4).

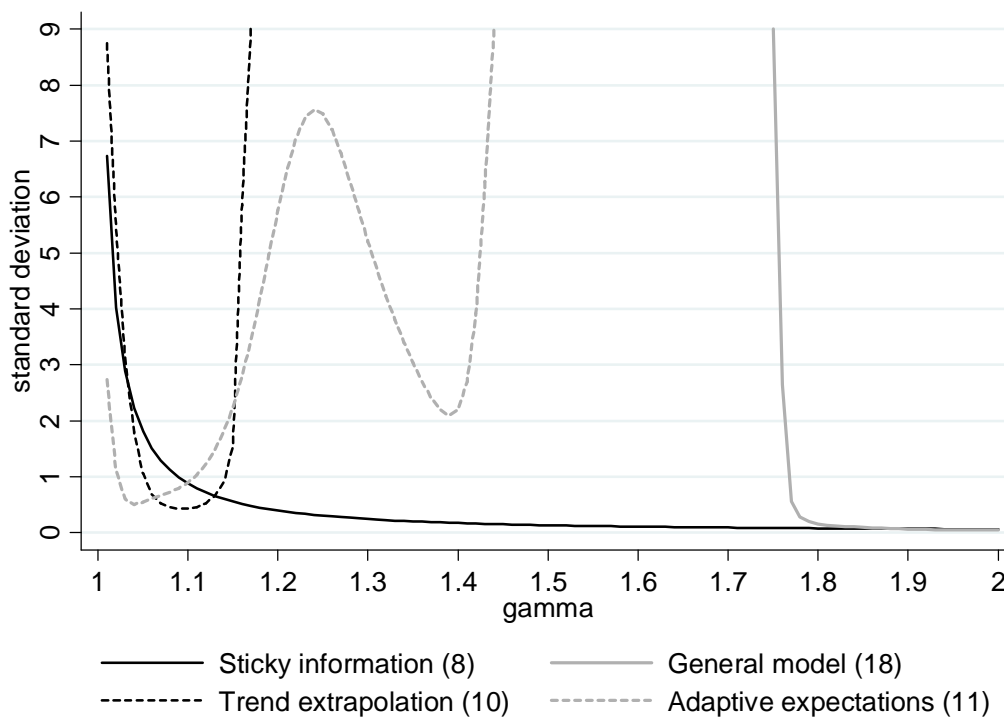


Figure A6: Variability of inflation and alternative expectation formation rules (inflation forecast targeting).

model (eq.)\ Treatments	1	2	3	4	All
Rational expectations (8)	35.2	48.1	5.6	25.9	28.7
AR(1) process (M1)	0.0	0.0	0.0	1.9	0.5
Sticky information type (M2)	3.7	1.9	16.7	3.7	6.5
Adaptive expectations (M3)	9.3	3.7	7.4	9.3	7.4
Trend extrapolation (M5)	35.2	25.9	25.9	33.3	30.1
Recursive - lagged inflation (M7)	3.7	5.6	24.1	13.0	11.6
Recursive - REE (M8)	0.0	1.9	9.3	0.0	2.8
Recursive - trend extrapolation (M9)	0.0	0.0	0.0	1.9	0.5
Recursive - AR(1) process (M10)	13.0	13.0	11.1	11.1	12.0

Table A1: Inflation expectation formation across treatments (percent of subjects, Comparison 1)

	Probit RE	Probit PA	Logit RE	Logit PA	Logit FE
Cons.	-0.2502 (0.3537)	-0.221 (0.1884)	-0.4139 (0.3149)	-0.3552* (0.1817)	
$ \pi_{t-1} - \pi_{t-2} $	0.0422 (0.0376)	0.0402 (0.0317)	0.0661 (0.0600)	0.0639 (0.0559)	0.0545 (0.0632)
$\pi_{t-1}$	-0.0568 (0.3717)	-0.0533 (0.1945)	-0.0919 (0.3099)	-0.0857 (0.1611)	-0.076 (0.2747)
$y_{t-1}$	-0.1702*** (0.0518)	-0.1596*** (0.0471)	-0.2747*** (0.0808)	-0.2577*** (0.0769)	-0.2540*** (0.0925)
$i_{t-1}$	0.044 (0.2493)	0.0415 (0.1308)	0.0715 (0.2075)	0.067 (0.1072)	0.0575 (0.1843)
$\left(\pi_{t-1} - \pi_{t-1 t-2}^k\right)^2$	0.0061 (0.0432)	0.006 (0.0367)	0.011 (0.0688)	0.0099 (0.0602)	0.0089 (0.0518)
$\ln(\sigma^2)$ (panel)	-1.5874*** (0.2882)		-0.5814*** (0.3014)		
$\sigma$ (panel)	0.4522*** (0.0652)		0.7478*** (0.1127)		
$\rho$ (panel)	0.1697*** (0.0406)		0.1453*** (0.0374)		
N	14040	14040	14040	14040	13975
Groups	216	216	216	216	215
Obs per Group	65	65	65	65	65
Wald $\chi^2(9)$	26.2	38.5	25.1	25.1	27.1

Table A2: Determinants of swithing behavior. Notes: RE stands for random effects, PA population averages, while FE is for fixed effects model. Standard errors in parentheses. \*/\*\*/\*\* denotes significance at 10/5/1 percent level. Standard errors are calculated using bootstrap procedures (1000 replications) that take into account potential presence of clusters in treatments.

	reg1	reg2	reg3	reg4
$sd_{s,t}$	1.0036*** (0.0081)	1.0026*** (0.0085)	1.0026*** (0.0085)	1.0007*** (0.0070)
Gen. mod. (M6), $\zeta = 0$	0.0015** (0.0007)	0.0009 (0.0008)	0.0015 (0.0010)	
Sticky info. (M2)	-0.0027** (0.0013)	-0.0034** (0.0014)	-0.0028** (0.0014)	-0.0037*** (0.0011)
ADE DGL (M4)	-0.0023** (0.0011)	-0.0028** (0.0014)	-0.0023* (0.0012)	-0.0024** (0.0011)
Trend Ext. (M5)	0.0036** (0.0014)	0.0031** (0.0014)	0.0037** (0.0017)	0.0034** (0.0013)
ADE CGL (M3)		-0.0017 (0.0010)	-0.0011 (0.0008)	
Recursive V1 (M7)		-0.0005 (0.0010)		-0.0008 (0.0009)
Recursive V4 (M9)			0.0005 (0.0010)	
T2	0.0446 (0.0347)	0.0426 (0.0377)	0.0426 (0.0377)	0.0476 (0.0384)
T3	-0.0873 (0.0553)	-0.0918 (0.0567)	-0.0918 (0.0567)	-0.0829 (0.0549)
T4	-0.1246* (0.0701)	-0.1296* (0.0719)	-0.1296* (0.0719)	-0.1242* (0.0725)
cons	0.0221 (0.0612)	0.0847 (0.0909)	0.0323 (0.0827)	0.0777* (0.0404)
N	1560	1560	1560	1560
$\chi^2$	74609.6	101746.4	101746.4	68326.8

Table A3: Decision model's influence on standard deviation of inflation. Notes: Estimations are conducted using system GMM estimator of Blundell and Bond (1998) for dynamic panels. Robust standard errors in parentheses. \*/\*\*/\*\* denotes significance at 10/5/1 percent level.

## B (Not for Publication) Results on Individual Rules

### B.1 Tests of Rational Expectations

Several econometric tests are designed to check the rationality of forecasts. In this subsection we apply some standard tests commonly employed in the survey data literature.<sup>36</sup> We assess different degrees of forecast efficiency and check whether forecasts yield predictable errors. The simplest test of efficiency is a test of bias:

$$\pi_{t+1} - \pi_{t+1|t}^k = \alpha, \quad (15)$$

<sup>36</sup>See Pesaran (1987), Mankiw, Reis, and Wolfers (2004) and Bakhshi and Yates (1998) for a review of these methods.

where  $\pi_{t+1}$  is inflation at time  $t + 1$  and  $\pi_{t+1|t}^k$  is  $k^{th}$  subject's inflation expectations for time  $t + 1$  made at time  $t$  (with information set  $t - 1$ ). By regressing expectational errors on a constant we check whether inflation expectations are centred around the right value. Majority of agents produce unbiased estimates of inflation. Overall, only 7.9% of them produce biased estimates at a 5% significance level and only 4.6% at a 1% threshold. Most of them are from treatments 2 and 4.

The next regression represents a further test for rationality:

$$\pi_{t+1} = a + b\pi_{t+1|t}^k. \quad (16)$$

As in Mankiw, Reis, and Wolfers (2004) the last expression can be simply augmented to test whether information in forecasts are fully exploited:

$$\pi_{t+1} - \pi_{t+1|t}^k = a + (b - 1)\pi_{t+1|t}^k, \quad (17)$$

where rationality implies jointly that  $a = 0$  and  $b = 1$ . As in the test for bias, under the null of rationality these regressions are meant to have no predictive power. The latter model is a more strict test of rationality and is seldomly fulfilled in the survey data literature. On the contrary, our results suggest that 28.7% of agents exploit all the available information at a 5% significance level and 42.1% of them when we decrease the threshold to 1%. Treatment 2 is associated with the highest proportion of rational agents (48% and 57%, accordingly). Compared to other experimental studies, these tests suggest that a significant proportion of subjects behave rationally, although in asset pricing experiments Heemeijer, Hommes, Sonnemans, and Tuinstra (2009) find a significant proportion of fundamental traders. These can be associated with rational expectations. Also Roos and Luhan (2008) show that about 23% of subjects do not have biased price expectations.<sup>37</sup>

## B.2 Sticky Information Type Regression

In this section we estimate a simple weighted average regression similar in formulation to sticky information model by Carroll (2003a) and adaptive expectations. In our framework we have forecasts derived under the assumption of rational expectations while Carroll (2003a) implements professional forecasters predictions. We estimate the following equation:

$$\pi_{t+1|t}^k = \lambda_1 \pi_{t+1|t}^{RE} + (1 - \lambda_1) \pi_{t|t-1}^k; \quad (18)$$

$$\pi_{t+1|t}^k = \lambda_1 \eta_0 + \lambda_1 \eta_1 y_{t-1} + (1 - \lambda_1) \pi_{t|t-1}^k, \quad (19)$$

---

<sup>37</sup>In field experiments by Berlemann and Nelson (2005) similar rationality tests were conducted suggesting that most participants exploit all available information.

where  $\pi_{t+1|t}^{RE}$  is a rational expectations prediction of inflation for period  $t + 1$  at period  $t$ . This type of models are important for forecasting, especially in our framework where some agents are backward-looking and also rational agents have to incorporate this into their forecasts. Thus we estimate the model (19) that is stated in terms of observable variables with the restrictions on all coefficients, where  $\eta_0$  and  $\eta_1$  are REE coefficients. Our formulation is inherently different than the one by Carroll (2003a, 2003b) as epidemiological framework that he proposes is no longer valid in our setup where subjects in principle observe all relevant information.<sup>38</sup> About 97% of agents display a significantly positive  $\lambda_1$ , while the average  $\lambda_1$  is 0.20. Groups in treatment 3 had the highest average  $\lambda_1$  (0.37), while subjects in treatment 2 had the lowest (0.11). It is not straightforward to define rationality in our framework and thus the results can be challenged on these grounds. The definition used in this subsection corresponds to REE if all agents in the group form expectations rationally.<sup>39</sup> Similar weighted average regressions are estimated in Heemeijer, Hommes, Sonnemans, and Tuinstra (2009), where they replace RE prediction with the equilibrium price.

### B.3 Trend Extrapolation Rule

We also evaluate simple trend extrapolation rules. These are pointed out as particularly important rules for expectation formation process in Hommes, Sonnemans, Tuinstra, and van de Velden (2005). We specify the following process:

$$\pi_{t+1|t}^k - \pi_{t-1} = \tau_0 + \tau_1 (\pi_{t-1} - \pi_{t-2}), \quad (20)$$

where we estimate  $\tau_0$  and  $\tau_1$ . We find that constant is significant at 5% level in 28.7% of cases while the  $\tau_1$  is significant in 78.2% of cases at the same level. Most of the times  $\tau_1$  is between 0 and 1, but there are a few cases when  $\tau_1$  is significantly lower than 0 (6.9%) and for 15.3% of subjects it is significantly higher than 1. We refer to the latter rules as strong trend extrapolation. Hommes, Sonnemans, Tuinstra, and van de Velden (2005) find that about 50% of subjects in their experiment behave consistently with the trend extrapolation rule.

### B.4 Estimating Simple Learning Rules

In order to test for adaptive behavior, we apply different learning rules to experimental data. For a discussion on learning rules and convergence to rational expectations see Evans and Honkapohja (2001). We first test learning on a model with constant gain

---

<sup>38</sup>He argues that news about inflation spreads slowly across agents and reaches only a fraction  $\lambda_1$  of population in each period.

<sup>39</sup>Note if we would use naive expectations this model would correspond to adaptive expectations in equation (21).

updating (CGL), where subjects learn from their past observed errors. The model below is equivalent to the adaptive expectations formula:

$$\pi_{t+1|t}^k = \pi_{t-1|t-2}^k + \vartheta (\pi_{t-1} - \pi_{t-1|t-2}^k), \quad (21)$$

where  $\vartheta$  is the constant gain parameter. Under this learning rule agents revise their expectations according to the last observed error. In the experiment subjects are asked to forecast inflation in the next period (hence they make their forecast for period  $t + 1$  at time  $t$ ), therefore the revision regards their previous period's forecast ( $t - 1$ ), which is made at time  $t - 2$ . Note that this rule corresponds to the second order adaptive scheme in Marimon, Spear, and Sunder (1993). All participants have  $\vartheta$  positive and significant at a 5 percent level. 13.4% of participants have a constant gain parameter significantly lower than 1, while 53.7% of them update their forecasts with an error correction term significantly greater than 1. This means that the latter agents possibly overreact to their past errors. Their prevalence might imply problems with dynamic stability in certain treatments.

Below we present a learning mechanism with decreasing gain parameter (DGL):

$$\pi_{t+1|t}^k = \pi_{t-1|t-2}^k + \frac{\iota}{t} (\pi_{t-1} - \pi_{t-1|t-2}^k). \quad (22)$$

If the estimated parameter ( $\iota$  in this version) is significantly different from 0, we conclude that agents actually learn from their past mistakes with a decreasing gain over time. Our tests do not support the hypotheses that the coefficient decreases over time as the  $R^2$  is always greater (for all subjects) for a constant gain model.

Several versions of these models are estimated in Arifovic and Sargent (2003), Hommes, Sonnemans, Tuinstra, and van de Velden (2005), Marimon and Sunder (1995) and Bernasconi and Kirchkamp (2000). Hommes, Sonnemans, Tuinstra, and van de Velden (2005) argue that some subjects (about 5%) behave consistently with this rule, while Marimon and Sunder (1995) and Bernasconi and Kirchkamp (2000) put forward that most subjects in their OLG experiments use either first or second order adaptive expectations.

#### B.4.1 Recursive Representation of Simple Learning Rules

The above specification mainly aims at testing whether data support the existence of adaptive behavior. As in the adaptive learning literature in this subsection we assume that subjects behave like econometricians, using all available information at the time of the forecast. In the following specifications, we test whether agents update their coefficients with respect to the last observed error. We assume four different perceived



laws of motion (PLM):

$$\pi_{t+1|t}^k = \phi_{0,t-1} + \phi_{1,t-1}\pi_{t-1} + \varepsilon_t. \quad (23)$$

$$\pi_{t+1|t}^k = \phi_{0,t-1} + \phi_{1,t-1}y_{t-1} + \varepsilon_t. \quad (24)$$

$$\pi_{t+1|t}^k = \phi_{0,t-1} + \phi_{1,t-1}\pi_{t|t-1}^k + \varepsilon_t. \quad (25)$$

$$\pi_{t+1|t}^k - \pi_{t-1} = \phi_{0,t-1} + \phi_{1,t-1}(\pi_{t-1} - \pi_{t-2}) + \varepsilon_t. \quad (26)$$

Note that equation (24) represents a PLM of the REE form and equation (26) a version of the trend extrapolation rule. When agents estimate their PLMs they exploit all available information up to period  $t - 1$ . As new data become available they update their estimates according to a stochastic gradient learning (see Evans, Honkapohja, and Williams, 2010) with a constant gain. Let  $X_t$  and  $\widehat{\phi}_t$  be the following vectors:  $X_t = \begin{pmatrix} 1 & \pi_t \end{pmatrix}$  and  $\widehat{\phi}_t = \begin{pmatrix} \phi_{0,t} & \phi_{1,t} \end{pmatrix}'$ . In this version of constant gain learning (CGL) agents update coefficients according to the following rule:

$$\widehat{\phi}_t = \widehat{\phi}_{t-2} + \vartheta X_{t-2}' \left( \pi_t - X_{t-2} \widehat{\phi}_{t-2} \right). \quad (27)$$

The empirical approach consists in searching the parameter  $\vartheta$  that minimizes the sum of squared errors (SSE), i.e.  $\left( \pi_{t+1|t}^s - \pi_{t+1|t}^k \right)^2$  (see Pfajfar and Santoro, 2010 for details). The implicit problem in this approach is that we have to assume the initial values for  $\widehat{\phi}_t$  for 2 periods. Setting up the initial values is one of the main problems when we recursively estimate learning. This issue is extensively discussed in Carceles-Poveda and Giannitsarou (2007). Strictly speaking, this problem should not occur in our case since we simply try to replicate our time-series data as closely as possible. Thus, in the following recursive learning estimations, we design an exercise in order to search for the best combinations of the gain parameter and initial values to match each subjects' expectations as closely as possible. This strategy can also be considered as a testing procedure for the detection of learning dynamics for each individual. If the gain is positive under this method of initialization, then the series would exhibit learning for all other initialization methods with higher (or equal) gain.

We find that 56.5% of participants learn according to the first setup with lagged inflation as in model (23). The gain parameter  $\vartheta$  is in the range between 0.0001 and 0.1000, with a mean value of 0.02900 and the median is 0.01125. We also estimate adaptive learning with the PLMs consistent of the REE form and AR(1) form, however these models rarely outperform other models studied here. In the learning version of the trend extrapolation model (26) 31.5% of subjects have positive gains. The optimal gains are on average slightly higher than before as they range between 0.0003 and 0.7900 with a mean value of 0.0654 (the median is 0.0310).

This version of the PLM (26) often performs better than previous versions of learning

in terms of SSE. Below we compare different models and find that this version of constant gain learning indeed best represents the behavior of a significant proportion of our subjects. For a comparison with other studies, we exclude from our sample all subjects for which learning does not represent the best model.<sup>40</sup> In this case, we find that the average gain of these subjects is 0.0447 with a standard deviation of 0.0537 (median is 0.0260). The standard deviation is quite high as there are a few very high values, but most of the gains fall in the range between 0.01 and 0.07.

There are only a few estimates of the gain coefficient in the literature. Orphanides and Williams (2005a) suggest a gain between 0.01 – 0.04 and Milani (2007) estimates it at 0.0183, while Pfajfar and Santoro (2010) find smaller gains (around 0.00021 for a similar version of learning). Results in this paper suggest slightly higher gains than most of the above papers, but our data might be more volatile than the actual US inflation.

## B.5 "General" Models of Expectation Formation

Simple learning rules do not capture all macroeconomic factors that can affect inflation forecasts. In this subsection we estimate some general models of expectation formation. We specify the following regression:<sup>41</sup>

$$\pi_{t+1|t}^k = \alpha + \gamma\pi_{t-1} + \beta y_{t-1} + \mu i_{t-1} + \zeta \pi_{t|t-1}^k + \varepsilon_t. \quad (28)$$

We find that 81.9% of agents take into account inflation when making their predictions. About 56.0% of the subjects take interest rate into account, while 66.7% also regard their own forecast from the previous period. Under some restrictions this equation could represent the form of the RE solution of the model ( $\zeta = 0$ ).<sup>42</sup>

For a comparison we also estimate a simple AR(1) model:

$$\pi_{t+1|t}^k = \phi_0 + \phi_1 \pi_{t|t-1}^k + \varepsilon_t. \quad (29)$$

Similar model was already estimated with recursive learning. Model with constant coefficients, in general, is not often used by subjects for forecasting inflation.

---

<sup>40</sup>We will consider Comparison 1 in the Table 6 and exclude model (24) as it is generally associated with extremely high values of gain parameter.

<sup>41</sup>Models in groups 19-24 do not have interest rate as dependent variable as this would imply multicollinearity due to the design of monetary policy in our framework.

<sup>42</sup>We also investigate more in depth the nature of the forecast error. We estimate the model where we regress the forecast error on past observed forecast error and changes of other macroeconomic variables. Subjects often do not exploit the informational content of the output gap and most importantly subjects overreact to last observed change of inflation. As the coefficient in front of the change in inflation is in most cases higher than 1, we can say many subject are pessimistic about future developments of inflation. This feature is repeatedly found in the survey data literature.

## C (Not for Publication) Experimental Instructions<sup>43</sup>

Thank you for participating in this experiment, a project of economic investigation. Your earnings depend on your decisions and the decisions of the other participants. There is a show up fee of the 4 Euros assured. From now on until the end of the experiment you are not allowed to communicate with each other. If you have some question raise your hand and one of the instructors will answer the question in private. Please do not ask aloud.

### The Experiment

All participants receive exactly the same instructions. You and 8 other subjects all participate as agents in the *same* fictitious economy. You will have to predict future values of given economic variables. The experiment consists of 70 periods. The rules are the same in all the periods. You will interact with the same 8 subjects during the whole experiment.

Imagine that you work in a firm where you have to predict inflation for the next period. Your profit depends on the accuracy of your inflation expectation.

### Information in Each Period

The economy will be described with 3 variables in this experiment: the *inflation rate*, the *output gap*, and the *interest rate*.

- **Inflation** measures general rise in prices in the economy. Each period it depends on the inflation expectations of the agents in economy (you and other 8 participants in this experiment), output gap and small random shocks.
- The **output gap** measures for how much (in %) the actual Gross Domestic Product differs from the potential one. If the output gap is greater than 0, it means that the economy is producing more than the potential level, if negative, less than potential level. It depends each period on inflation expectations of the agents in economy, past output gap, interest rate and small random shocks.
- The **interest rate** is (in this experiment) the price of borrowing the money (in %) for one period. The interest rate is set by the monetary authority. Their decision mostly depends on inflation (expectations) of the agents in economy.

All given variables might be relevant for inflation forecast, but it is up to you to work out their relation and possible benefit of knowing them. The evolution of variables will

---

<sup>43</sup>Instructions used for experiments at Universitat Pompeu Fabra are in Spanish language. In experimental sessions, they were accompanied with the screenshots of the experimental interface and the profit table with earnings for various combinations of *estimation error* and *confidence interval*.

partly depend on the inputs of you and other subjects and also different random shocks influencing the economy.

- You enter the economy in period 1. In this period you will be given computer generated past values of inflation, output gap and interest rate for 10 periods back (Called: -9, -8, ... -1, 0)
- In period 2 you will be given all past values as seen in period 1 plus the value from period 1 (Periods: -9, -8, ... 0, 1).
- In period 3 you will see all past values as in period 2 (Periods: -9, -8, ... 1, 2) plus YOUR prediction about inflation in period 2 that you made in period 1.
- In period  $t$  you will see all past values of actual inflation up to period (Periods: -9, -8, ... , ) and your predictions up to period (Periods: 2, 3, ... , ).

### What Do You Have to Decide?

Your payoff will depend on the accuracy of your prediction of the inflation in the future period. In each period your prediction will consist of two parts:

1. *Expected inflation*, (in %) that you expect to be in the NEXT period (*Exp.Inf.*)
2. The *Confidence Interval* (*Conf.Int.*) around your prediction for which you think there is 95% probability that the actual inflation will fall into. The interval is determined as the number of percentage points for which the actual inflation can be higher or lower.

**Example 1** *Let's say you think that inflation in the next period will be 3.7%. And you also think there is most likely (95% probability) that the actual inflation will not differ from that value for more than 0.7 percentage points. Therefore, you expect that there is 95% probability that actual inflation in the next period will be between 3.0% and 4.4% (3.7% ± 0.7%). Your inputs in the experiment will be 3.7 under 1) and 0.7 under 2).*

Your goal is to maximize your payoff, given with the equation:

$$W = \max \left\{ \frac{100}{1 + |\text{Inflation} - \text{Exp.Inf.}|} - 20, 0 \right\} + \max \left\{ \frac{100x}{1 + \text{Conf.Int.}} - 20, 0 \right\}$$

where *Exp.Inf.* is your expectation about the inflation in the NEXT period, *Conf.Int.* is the confidence interval you have chosen, *Inflation* is the actual inflation in the next period, and  $x$  is a variable with value 1 if

$$\text{Exp.Inf.} - \text{Conf.Int.} \leq \text{Inflation} \leq \text{Exp.Inf.} + \text{Conf.Int.}$$

and 0 otherwise.

This expression tells you, that  $x$  will be 1, if actual inflation falls between  $Exp.Inf. - Conf.Int.$  (3.0% in our example) and  $Exp.Inf. + Conf.Int.$  (4.4% in our example).

The *first part* of the payoff function states that you will receive some payoff if the actual value in the next period will differ from your prediction in this period for less than 4 percentage points. The smaller this difference will be, the higher the payoff you receive. With a zero forecast error ( $|Inflation - Exp.Inf.| = 0$ ), you would receive 80 units. However, if your forecast is 1 percentage point higher or lower than the actual inflation rate, you will get only 30 units ( $100/2 - 20$ ). If your forecast error is 4 percentage points or more, you will receive 0 units ( $100/5 - 20$ ).

The *second part* of the payoff function simply states that you will get some extra payoff if the actual inflation is within your expected interval and if that interval is not be larger than  $\pm 4$  percentage point. The more certain of the actual value you are, the smaller interval you give, and the higher will be your payoff if the actual inflation indeed is in the given interval but there will also be higher chances that actual value will fall outside your interval. In our example this interval is  $\pm 0.7$  percentage points. If the actual inflation falls in this interval you would receive 38.8 units ( $100/(1 + 0.7) - 20$ ) in addition to the payoff from the first part of the payoff function. If the actual values is outside your interval, your receive 0.

In the attached sheet you can find table which shows various combinations of *forecast error* and *confidence interval* needed to earn a given number of points. See also figure on the next page.

### Information After Each Period

Your payoff depends on your predictions for the next periods and actual realization in next period. Because the actual inflation will be only known in the next period, you will also be informed about you current period ( $t$ ) prediction and earnings after the end of NEXT period ( $t + 1$ ). Therefore:

- After Period 1 you will not receive any earnings, since you did not make any prediction for the period 1.
- In any other period, you will receive the information about the actual inflation rate in this period and your inflation and confidence interval prediction from previous period. You will also be informed if the actual inflation value is in your expected interval and what are your earnings for this period.

The units in the experiment are fictitious. Your actual payoff will be the sum of profits from all the periods converted to euros in 1/500 conversion.

If you have any questions please ask them now!

## Questionnaire<sup>44</sup>

1. If you believe that inflation in the next period will be   4.2%  , and you are quite sure that it will be higher than   3.5%   and lower than   4.9%  , you will type:  
Under (1)                               for inflation, and  
Under (2)                               for confidence interval.
2. If you are now in period                              , you have information about past inflation, output gap and interest rate up to period                               and you have to predict the inflation for period                              .

---

<sup>44</sup>Options (1) and (2) are pointing to the different fields on the screenshot of the experimental interface.