Peter Heemeijer, Cars Hommes, Joep Sonnemans and Jan Tuinstra

Inflation Expectations and Stability in an Overlapping Generations Experiment with Money Creation

DNB Working Paper

De Nederlandsche Bank
Eurosysteem

Inflation Expectations and Stability in an Overlapping Generations Experiment with Money Creation

Peter Heemeijer, Cars Hommes, Joep Sonnemans and Jan Tuinstra *

Inflation Expectations and Stability
in an Overlapping Generations Experiment
with Money Creation

By Peter Heemeijer (DNB),

Cars Hommes (UvA),

Joep Sonnemans (UvA),

Jan Tuinstra (UvA).

We investigate how non-specialists form inflation expectations by running an experiment using a basic Overlapping Generations (OLG) model. The participants of the experiment are students of the University of Amsterdam, who predict inflation during 50 successive periods and are rewarded based on their accuracy. We include a central bank in the OLG model which increases the money supply at a constant rate. Participants are placed in separate OLG economies and are divided over two treatments: one with a "low" and one with a "high" money supply growth. We find that participants in the second treatment have substantially more difficulty in stabilizing inflation development by submitting accurate predictions than participants in the first treatment. However, when linear prediction rules are estimated on individual predictions, there is little difference between the two treatments. In both treatments, the most popular rules are Fundamentalist Expectations (predictions equal to the inflation sample mean) and Focal Expectations (predictions equal to a constant close to equilibrium). To verify whether participants adjust their prediction rules during the experiment, the estimated rules are checked for structural breaks. We find a surprisingly small number of structural breaks in both treatments.

1. Introduction

In economic systems with expectations feedback, in which agents' expectations about variables influence their realized values, the kinds of prediction rules that are applied have a major impact on the dynamics. This holds in particular for the dynamics close to the system's steady states, which often converges for certain rules and diverges for others, thus creating a set of stability conditions that can be helpful in judging the steady state's practical relevance. Determining stability conditions by testing, analytically or by simulation, a range of boundedly rational prediction rules, is especially important in expectations feedback systems with multiple equilibria. By determining the rules under which the equilibria are stable, the equilibria can in many cases be ordered from most to least relevant, given a good understanding of the kinds of prediction rules applied by individuals.

A widely used model that in the last two decades has been the subject of debate regarding the stability of its equilibria under bounded rationality is the Overlapping Generations (OLG) Model, originally developed by Samuelson (1958). Individuals in this model live for several, typically two or three, time periods after which they die and are replaced by a generation of new individuals. During each of the periods, individuals are active as consumers; in all but the last generation they may also be producers. What gives the OLG model its name and makes it suitable for studying many long-term macroeconomic phenomena is the co-existence of all generations, creating interaction between the old and young, which by market clearance determines prices through time. The main areas of interest in the model are the price development of the, typically single, goods market, and the welfare effects implied for the different generations. The price development is often, by simple

variable transformation, expressed as inflation or interest development.

In a much discussed paper on the interpretation of economic theory in general and economic equilibria in particular, Lucas (1986) argued that equilibria, even if they are derived under a rational expectations or perfect foresight assumption, should properly be understood as "steady states of some adaptive process" (*p.* 402). Viewing an individual as a set of preferences and decision rules constantly tested against experience, Lucas defines the term "adaptive" in the most general sense, namely as a "trial-and-error process through which our modes of behavior are determined." (*p.* 401) Applying this idea to a standard overlapping generations model, he shows that, while the model has two steady-state equilibria (an "autarkic" equilibrium in which no trade occurs and a "quantity-theoretic" equilibrium which is the social optimum), only one of them (the quantity-theoretic equilibrium) is stable when individuals use Adaptive Expectations with a low adjustment parameter to predict next period's price (for definition see Section 6.1). This result suggests, according to Lucas' interpretation of equilibria, that the stable steady state is the economically relevant one. However, realizing that he had only tested a single boundedly rational prediction rule while individuals can use many different rules not necessarily all converging to the same equilibrium, he warned against discriminating between the steady states based on this one example.

Lucas' warning turned out to be justified. Bullard (1994) shows in a similar over-lapping generations model that the model's non-autarkic or monetary steady state, at an inflation rate equal to the rate of money growth, becomes unstable for sufficiently high money growth, given that individuals assume that the inflation rate is constant and estimate it by applying OLS estimation to previous prices. In doing so he questioned the relevance of the socially desirable steady state that Lucas, in a similar setting with a constant money stock, had shown to be stable under adaptive expectations. At the same time, Bullard's paper demonstrated, together with later research by Schönhofer (1999), that his OLS prediction rule produces complicated, sometimes even chaotic, inflation fluctuations around the monetary steady state for high rates of money growth. Because these fluctuations are not the result of specifically constructed utility functions or other model structure as in Grandmont (1985), they are purely a by-product of the chosen prediction rule. Given their persistence in the long run, they are referred to as *learning equilibria*. Learning equilibria are a complication when testing rational expectations equilibria using boundedly rational prediction rules. These prediction rules may indicate which of the Rational Expectations (RE) equilibria are economically interesting, but may just as well add new equilibria to the model; in the end the most relevant equilibrium may very well be a learning equilibrium not present among the

initial set of RE equilibria.

Bullard's learning equilibria are the result of the OLS prediction rule he introduced but more specifically of the unusual way in which OLS estimation is applied by the young generations of his OLG model. Individuals first assume that inflation is constant, but then proceed by estimating it using realized prices. As noted by Tuinstra & Wagener (2007, *p.* 494), this estimation procedure is both unnatural from an individual's viewpoint and unsuitable from an econometric viewpoint since prices under a positive rate of money creation in a stationary economy increase exponentially and are therefore non-stationary. Tuinstra & Wagener (2007) show that if the estimation procedure is changed so that individuals predict the inflation rate by taking the average of all realized inflation rates of earlier periods, the learning equilibria found by Bullard disappear and the stability of the monetary steady state, analogous to the result in Lucas (1986), is restored.

The continuing discussion on the economic relevance of the steady-state equilibria of OLG models makes clear that it is hard to judge their relevance from their stability under specific boundedly rational prediction rules. This is the case first because there are many such rules and second because even subtle changes in a single rule can completely change the local dynamics around steady states. By itself, a stability analysis using boundedly rational expectation formation therefore has limited added value next to a traditional equilibrium analysis assuming perfect rationality. However, this added value would be greatly increased if it were known what sort of prediction rules individuals tend to use in economic environments resembling the overlapping generations model. An effective way of obtaining this knowledge, as Lucas (1986, *pp.* 420-1) recognizes, is by running experiments that place participants in a simulated OLG economy and observing their expectations of prices or inflations in the immediate future. In this paper the results of an experiment along these lines are analysed. Examples of similar OLG experiments are Marimon & Sunder (1993, 1994) and Marimon, Spear & Sunder (1993).

The paper is structured as follows. Section 2 formally introduces the OLG model that provides the framework of the experiment. In Section 3 the specific utility and savings functions used in the experiment are given, as well as the experimental parameters and treatment descriptions. Section 4 provides an overview of the experimental results, including the earnings of participants. Subjects' prediction errors are analysed in Section 5, followed by tests of the predictions' consistency with the Rational Expectations Hypothesis (prescribing minor fluctuations around the steady state of the experimental model). Individual prediction rules are studied in Section 6, starting with simulations of several simple rules, followed by

estimation using first a general linear specification and then a more restricted specification combining several rules of thumb. Afterwards the estimated individual rules are checked for structural breaks during the experiment. Section 7 summarizes the results and adds some concluding remarks. The details of the experimental procedure are found in the appendix, as well as a copy of the instructions for the participants, preceded by an English translation.

2. The Overlapping Generations Model with Money Creation

The experimental model consists of an overlapping generations economy with two generations, a single good and a central bank creating money at a constant rate. The generations are the young and old segments of the economy's population, with the typical retiring age dividing the segments. The "young" generation therefore represents the working part of the population, receiving units of the consumption good as a reward for their labor. The "old" generation represents the retired part, buying goods with the money saved when they were still working. Both segments of the population meet at the market for consumption goods, where the young generation sells part of their goods in order to save for retirement and the old generation buys them with its savings made when it was young. Clearing of the market determines the price of the consumption good given the amount of money circulating in the economy. The experimental model is dynamic with periods lasting a generation, meaning that at the end of each period the young generation turns old, the old generation dies and a new, young generation is born.[1]

The generations are represented by a single individual having a utility function $u(c_1, c_0)$ defined on the amounts of goods c_0 and c_1 consumed when "young" and "old" respectively. Let p_t be the price of the consumption good at a current period t and p_{t+1}^e the expectation of the individual of the price in the next period t+1. Suppose the individual is endowed with a total of w_0 goods when he is young and receives no goods at all when retired. Then a young individual at period t maximizes his utility $u(c_1, c_0)$ given the budget constraint

---

[1] The general structure of the OLG model, i.e. up to the choice of utility and savings functions, random disturbances and parameters like the rate of money creation, is identical to the model in Bullard (1994), Schönhofer (1999) and Tuinstra & Wagener (2007). The model in Lucas (1986) can be seen as a special case of this paper's model with zero money creation. The standard version of Grandmont's OLG model (1985) also has no money creation; it differs from the above mentioned models by allowing for relatively strong income effects in the intertemporal utility function which generate non-monotonous savings functions, resulting under certain conditions in rational expectations equilibria with a complicated cyclical character.

$p_t c_0 + p^e_{t+1} c_1 \leq p_t w_0$. Assuming that $u(c_1, c_0)$ is continuous, strictly increasing and strictly quasiconcave a unique consumption optimum $(c^*_0, c^*_1) \geq 0$ exists, maximizing utility under the assumption that the price in period $t+1$ is correctly anticipated.[2] The optimal quantities (c^*_0, c^*_1) are calculated by dividing the first-order conditions of utility maximization under the budget constraint on each other, and rearranging terms:

$$\frac{\partial u(c_0, c_1)}{\partial c_0} p^e_{t+1} = \frac{\partial u(c_0, c_1)}{\partial c_1} p_t. \tag{1}$$

Since the optimal consumption quantities are functions of p_t and p^e_{t+1} but are not dependent on the absolute price level, the savings function S can be defined as follows:

$$S\left(\frac{p^e_{t+1}}{p_t}\right) = w_0 - c^*_0\left(\frac{p^e_{t+1}}{p_t}\right). \tag{2}$$

Taking M_t as the amount of money in the economy at period t, which is demanded by the young generation to acquire savings for their old age, equality of demand and supply on the goods market requires the following restriction:

$$\frac{M_t}{p_t} = S\left(\frac{p^e_{t+1}}{p_t}\right). \tag{3}$$

The total money quantity M_t in this equation is independently defined by the money creation policy of the central bank:

$$M_t = \theta M_{t-1}, \quad \theta > 1, \tag{4}$$

with θ the rate of increase in the money quantity, or *seigniorage*, used by the govern-ment to make expenditures in the economy by diluting the real value of individual savings. Solving (3) for M_t and substituting in both sides of (4) gives:

$$S\left(\frac{p^e_{t+1}}{p_t}\right) p_t = \theta S\left(\frac{p^e_t}{p_{t-1}}\right) p_{t-1}. \tag{5}$$

---

[2] When the price expectation is not perfectly accurate the individual will either receive less than c^*_1 units of the good during old age (if $p_{t+1} > p^e_{t+1}$) or more than c^*_1 units (if $p_{t+1} < p^e_{t+1}$). In general these quantities do not in hindsight optimize his lifetime utility given the true prices p_t and p_{t+1}, except in the case that $(\partial u(c_0, c_1)/\partial c_0)(\partial u(c_0, c_1)/\partial c_1) = p_t / p_{t+1}$ happens to hold at (c_0, c_1) $= (c^*_0, (p^e_{t+1} / p_{t+1})c^*_1)$.

Defining the rate of inflation as $\pi_t = p_{t+1} / p_t$ and rearranging reveals the law of motion of the inflation rate:

$$\pi_t = \theta \frac{S(\pi_t^e)}{S(\pi_{t+1}^e)}, \tag{6}$$

which is dependent on two successive price expectations p_t^e and p_{t+1}^e of young individuals and can therefore have many different explicit forms. When the individuals correctly anticipate prices, i.e. have Rational Expectations with $\pi_t^e = \pi_t$ holding for all t, then it follows immediately from (6) that a unique constant equilibrium exists at $\pi_t = \theta$, $t \geq 0$. This steady-state equilibrium is referred to as the *monetary stationary state*, since in it inflation is both constant and fully determined by the monetary policy of the central bank. Because of the monetary steady state, θ is also referred to as the *natural rate of inflation*. The equilibrium is of central importance because of its simplicity, its uniqueness and the fact that rational expectations is one of the most common descriptions of individual expectation formation.[3] Note that the monetary stationary state is independent of the specific form of the savings function. In order to apply the inflation law of motion (6) in an experiment, it is necessary to fully specify its components. In the next section the savings function $S(\pi_t^e)$ used in the experiment is determined from the underlying utility function $u(c_0, c_1)$ and from the values chosen for the natural rate of inflation θ.

3. Experimental Treatments and Parameters

In the experiment, equation (6) is used to calculate realized inflation rates and the natural inflation rate θ is the only treatment parameter. The experiment has two treat-\ments, one with a relatively low and one with a relatively high value of θ. Both are single-agent treatments, i.e. each participant provides the inflation estimates necessary to determine the realized inflation in (6), generating a series of inflations that is not influenced by the actions of other participants.[4] The savings function $S(\pi_t^e)$ and utility function $u(c_0, c_1)$ remain the

---

[3] In principle another constant equilibrium, the *autarkic steady-state equilibrium*, exists for series of expectations $\pi_t^e = \pi^a$, $t \geq 0$, such that $S(\pi^a) = 0$. In this equilibrium there is no saving and therefore no trade on the goods market, leaving prices and actual inflation rates undetermined. The autarkic steady state does not exist however in the experiment, because the savings function S is positive for all expected inflation rates $\pi_t^e \geq 0$ (see Section 3).

[4] This experimental set-up is analogous to the use of a single representative agent in most literature on

same in both treatments; for practical reasons the savings function is actually a slight modification of the theoretical savings function generated by optimizing utility and applying equation (2).

Let the utility of the representative individual have a Constant Elasticity of Substitution (CES) form:

$$u\left(c_0,c_1\right)=\left(\left(\nu c_0\right)^{\rho}+c_1^{\rho}\right)^{1/\rho},\quad \nu>0,\quad 0\neq\rho<1, \tag{7}$$

with the coefficient ν indicating the individual's appreciation of consumption during working age relative to consumption during retirement. Calculating the lifetime demands for the consumption good $c_0^*(\pi_t^e)$ and $c_1^*(\pi_t^e)$ and applying (2) results in the following savings function:

$$S\left(\pi_t^e\right)=\frac{w_0}{1+\left(\nu\pi_t^e\right)^{\frac{\rho}{1-\rho}}}, \tag{8}$$

which is a strictly positive and, for positive ρ, monotonously decreasing function. This savings function has the practical disadvantage that whenever a subject expects a high enough inflation rate π_t^e, assuming a positive ρ, saving drops to nearly zero. If the previous expectation was more moderate, producing substantially larger savings, the law of motion (6) implies that the realized inflation rate at period t-1 will be very large, potentially reaching any positive number. In an experiment extremely high inflations would be confusing to participants as well as difficult to plot on a computer screen, making (8) unsuitable as a savings function. Preventing (8) from dropping to zero for large inflation predictions, as in the modification below, removes the disadvantage:

$$S\left(\pi_t^e\right)=\delta+\left(1-\delta\right)\frac{w_0}{1+\left(\nu\pi_t^e\right)^{\frac{\rho}{1-\rho}}},\quad \delta\in\left(0,1\right). \tag{9}$$

δ is a lower boundary on the amount of savings by the young generation, thereby preventing the inflation rate from going to infinity. Savings function (9) also provides an upper bound on the amount of savings, $S(\pi_t^e)\leq\delta+(1-\delta)w_0$, given that inflation estimates are always non-

expectation formation and learning in macro-economic models. See e.g. Bullard (1994), Schönhofer (1999), Tuinstra (2003).

negative. Taking (6) into account, this implies that the realized inflation rate lies in an open interval:

$$\frac{\delta}{\delta + (1-\delta)w_0} < \frac{\pi_t}{\theta} < \frac{\delta + (1-\delta)w_0}{\delta}. \tag{10}$$

The parameter choices in the experiment are $w_0 = 0.9$, $\nu = 0.92$, $\rho = 0.965$ and $\delta = 0.4$. This implies that $4/10 < S(\pi_t^e) < 94/100$ and $20/47 < \pi_t/\theta < 47/20$.[5] The savings function (9) under these parameter values is drawn in Figure 1(a). The steep drop in savings just above an expected inflation of 1 is caused by the closeness of ρ to 1, which gives the exponent in the denominator of (9) a high value, making the denominator increase quickly as π_t^e exceeds 1. In the figure the natural rates of inflation θ are accentuated at $\theta_1 = 1.01$ and $\theta_2 = 1.11$ respectively. Using these parameter values the first treatment, with $\theta_1 = 1.01$, is titled *Low Theta* (LTh), and the second, with $\theta_2 = 1.11$, *High Theta* (HTh). Both treatments are made up of 16 participants, each generating a series of realized inflations and associated predictions.

The sensitivity of the savings function (9) to changes in the expected inflation rate is measured in Figure 1(b) by the inflation elasticity $a(\pi_t^e) = -(\pi_t^e / S(\pi_t^e))(dS(\pi_t^e)/d\pi_t^e)$. Because the inflation predictions are non-negative and the savings rate positive and decreasing, this expression is everywhere non-negative, increasing with the reduction in the savings amount following an inflation rise.[6] The treatment values of the natural inflation rate θ are also indicated in Graph (b) of Figure 1, which shows a fraction of the domain of Graph (a). The elasticities of the savings function at the natural inflation rates $a(\theta_i)$, $i = 1, 2$, represent the sensitivity of the savings amounts around the equilibrium inflation levels of Treatments 1

---

[5] Note that the maximum amount of goods that can be sold by the young generation slightly exceeds the endowment w_0 available to it. This is the result of the modification of the theoretical savings function (8). The modified savings function (9) may be interpreted though as if a proportion δ of the young generation receives an endowment of 1 good and saves this completely, while the remaining proportion $1 - \delta$ receives $w_0 = 0.9$ goods and has the CES utility function (7) and the associated saving behavior (8).

[6] Subsituting (8) and its derivative in the definition of $a(\pi_t^e)$ gives the explicit form of the elasticity curve:

$$a(\pi_t^e) = \frac{\rho}{1-\rho} \frac{w_0 \left(\nu\pi_t^e\right)^{\rho/1-\rho}}{\left[\dfrac{\delta}{1-\delta} + \dfrac{w_0}{1 + \left(\nu\pi_t^e\right)^{\rho/1-\rho}}\right]\left(1 + \left(\nu\pi_t^e\right)^{\rho/1-\rho}\right)^2},$$

part of which is shown in Figure 1(b) after subsituting the experimental parameter values.
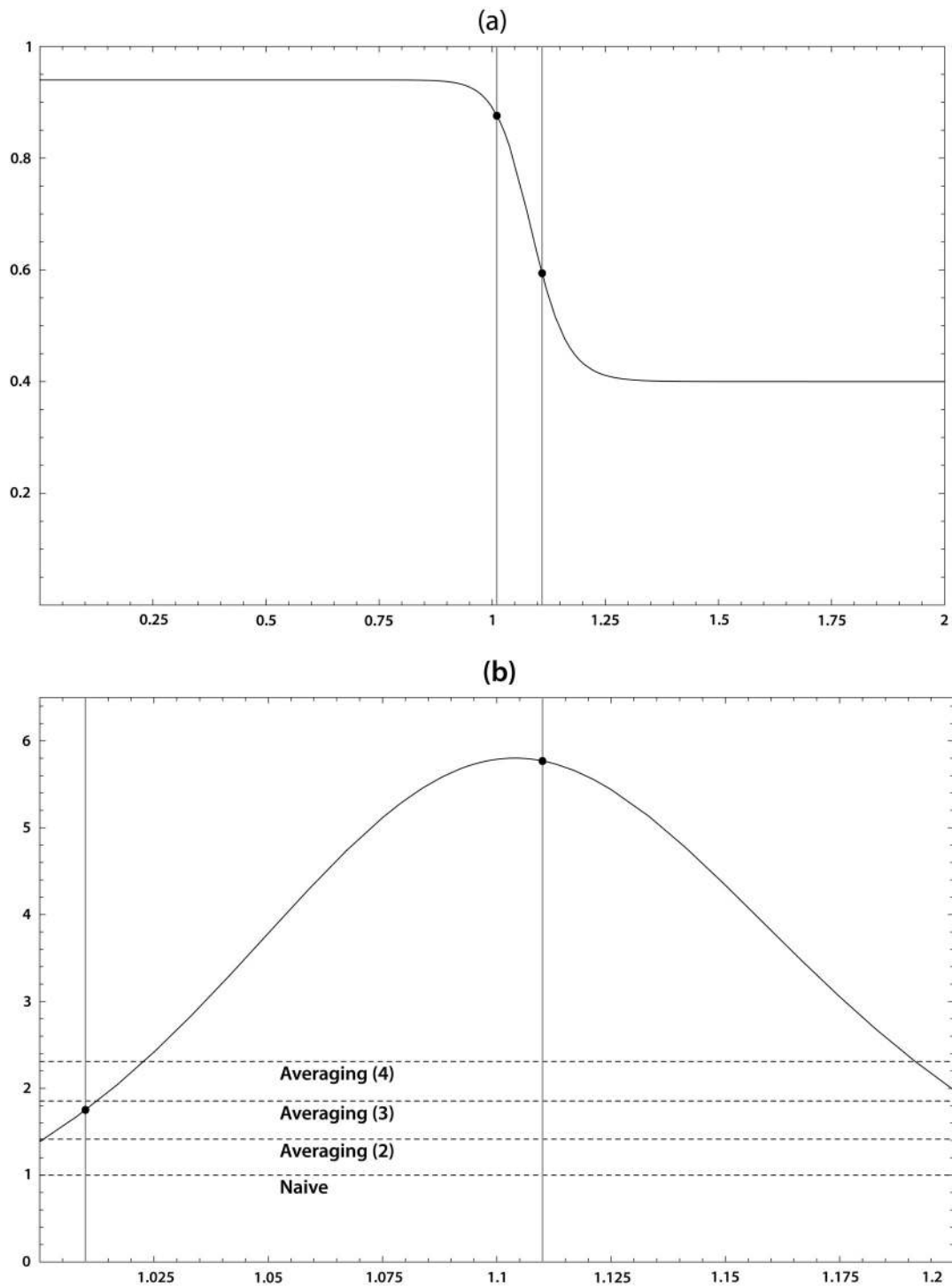
Figure 1: Savings rate as a function of expected rate of inflation (Graph (a)) and inflation elasticity of the savings rate (Graph (b)). Gray lines indicate the treatment parameters; horizontal dashed lines are boundaries above which averaging rules become unstable (from bottom to top: Naive Expectations; Averaging Expectations of last 2, 3 and 4 predictions resp.) Dots indicate the savings rates (Graph (a)) and the inflation elasticities of the savings rate (Graph (b)) in the steady-state equilibria of both treatments.

(LTh) and 2 (HTh). Looking at the law of motion (6) this sensitivity determines the degree of instability of the constant rational expectations equilibria in the two treatments. This means that the lower $a(\pi_t^e)$ is for $\pi_t^e = \theta$, with θ the natural inflation rate, the more prediction rules make (6) a difference equation with an asymptotically stable equilibrium at $\pi_t = \theta$. The horizontal lines in Figure 1(b) show that Naive Expectations, i.e. predictions equal to the most recently observed inflation rate, as well as an averaging rule over the last two realized rates, make the constant equilibrium unstable in Treatment 1 (LTh). Averaging over three or four of the most recent inflation lags suffices for stability.[7] In the case of the high natural rate $\theta_2 = 1.11$, none of the above mentioned rules generates convergence to the constant equilibrium, implying that subjects will have a harder time stabilizing inflation in the second treatment than in the first. A simulation of the prediction rules appearing in Figure 1(b) is provided in Section 6.1.

In the experiment, participants submit inflation predictions that, applying the law of motion (6) with the modified savings function (9), result in actual inflation rates. As soon as these inflation rates are realized, they are made available to participants for possible use in the formation of further predictions. The cycle then repeats itself, with the set of previous inflation rates constantly expanding, until a predetermined number of predictions has been submitted. The law of motion used in the experiment consists of equation (6), with random disturbances added to represent fluctuations in the expansion of the money supply by the central bank. The disturbances at the same time prevent trivial developments in the realized inflation rates. The law of motion has the following form in the experiment:

$$\pi_t = \theta \, \varepsilon_t \, \frac{S\left(\pi_t^e\right)}{S\left(\pi_{t+1}^e\right)}, \tag{11}$$

with ε_t uniformly distributed on the interval $[0.975, 1.025]$. This implies that in the constant rational expectations equilibrium $\pi_t^e = \theta$ should hold for all t, while the actual inflation rate

---

[7] Substituting these prediction rules in the law of motion (6) and linearizing the resulting difference systems around the constant equilibrium $\pi_t = \theta$ allows one to compute (exactly for very simple rules, otherwise approximately) the elasticity values above which specific rules lead to instability of the equilibrium (cf. Tuinstra 1999, *pp.* 91-2 for the case of naive expectations). For the prediction rules shown in Figure 1(b), from bottom to top, these elasticity values are $a(\theta) = 1$ (naive expectations), $a(\theta) = \sqrt{2}$ (averaging over 2 inflations), $a(\theta) \approx 1.854$ (averaging over 3 inflations), $a(\theta) \approx 2.309$ (averaging over 4 inflations).

fluctuates around the natural rate with an amplitude of at most 2.5 percent of that rate.

The law of motion (11) illustrates the order in which variables are determined during the experiment. Since both inflation predictions π_t^e and π_{t+1}^e are required to determine the actual inflation rate π_t, a participant predicts the inflation rate two periods ahead. In an overlapping generations economy this makes sense, because the young generation in period $t+1$ needs π_{t+1}^e to choose how much it will save, and the resulting supply of goods together with the demand from the old generation (based on π_t^e from the previous period t) are required to determine p_{t+1} up to the existing price level, i.e. $p_{t+1}/p_t = \pi_t$. The experiment lasts a total of 50 periods, meaning that each parti-cipant submitted 51 predictions producing a series of 50 realized inflation rates, with the last prediction only serving to determine the 50th inflation. In an arbitrary period t, during which a participant i would be asked to submit an inflation prediction π_{t+1}^e, the information set available to him has the following form:

$$I_{i,t} = \left\{ \left\{ \pi_{i,1}^e, \pi_{i,2}^e, \ldots, \pi_{i,t}^e \right\}, \left\{ \pi_1, \pi_2, \ldots, \pi_{t-1} \right\} \right\}, \quad t \leq 50. \tag{12}$$

The final element of the experimental design concerns the way in which sub-jects were rewarded for their predictions. They received points or experimental credits for each prediction, positively related to its accuracy. The accuracy of predictions was measured through the absolute prediction error, so a Linear Scoring Rule (LSR) was used. The LSR awards points to each inflation prediction inversely proportional to the absolute prediction error, up to a certain magnitude, beyond which no points are awarded. Define $P(\pi_t^e)$ as the points awarded for a prediction $\pi_{i,t}^e$ of the inflation π_t by an individual i. Then $P(\pi_t^e)$ is the following LSR:

$$P\left(\pi_{i,t}^e\right) = Max\left\{100 - 400\left|\pi_{i,t}^e - \pi_t\right|, 0\right\}. \tag{13}$$

Participants therefore earn a reward for each prediction with an error of no more than 0.25, or, as it would be expressed to them, 25%.[8] For each participant all rewards per period were rounded and then added at the end of the experiment to be converted into euros at the rate of 200 points for 1 euro. Subjects could earn a maximum of 25 euros for the whole experiment, though it was practically impossible to achieve earnings very close to the maximum because of

---

[8] In the experiment, inflation rates were expressed as percentual changes in the price level in order to make them easier to interpret for the participants. Any inflation rate $\pi \geq 0$ was expressed as $(100\pi - 100)\% \geq -100\%$.

the random disturbances in (11). The reason for using a LSR instead of a Quadratic Scoring Rule (QSR) in this paper is that an LSR has a derivative that does not go to zero when approaching the perfectly accurate prediction $\pi_t^e = \pi_t$, while the derivative of a QSR does go to zero. This motivates participants to be precise in their. Assuming that the subjective inflation distributions of participants through time are not broader than 0.5 (i.e., the difference between the maximum and minimum inflations to which they assign positive probability is no larger than 50%) and that subjects are expected utility optimizers, a standard result says that the LSR extracts the median of these distributions, while the QSR extracts the mean. Both can be considered useful representations of the underlying beliefs.

4. An Overview of the Experimental Results

The difference between the two experimental treatments lies in the natural rate of inflation θ, which next to fixing the equilibrium level of inflation influences the magnitude of the random disturbances in each period (see (11)). It can be expected that changing the natural rate will have a noticeable effect on the inflation dynamics, since out-of-equilibrium realizations of inflation are proportional to θ. One research question is whether differences in dynamics between the treatments are merely differences in the scale of inflation values or involve other, qualitative changes. This question will be formally investigated in the next two sections, but before doing so the experimental results are presented as a whole, allowing for some preliminary hypotheses to be formulated.

Since the experiment has only single-agent treatments, results for each of the 32 participants, consisting of 16 replications of the experiment for each of the two treatments, can be studied separately. In this section the inflation prediction and realization series associated with the participants will be introduced, organized by treatment. After that the total earnings of participants will be looked at.

4.1   Low Equilibrium Inflation Rate Treatment

The experimental results of each participant are characterized by the series of inflation predictions and realized values, of length 51 and 50 respectively.[9] These series are shown for

---

[9] The difference in length is caused by the fact that inflation predictions are submitted two periods ahead in the OLG model with two generations (cf. (11)), requiring subjects to end the experiment with a prediction for period 51 from which the realized inflation in period 50 follows. Note that the

all 16 participants of Treatment 1 (LTh) in separate graphs of Figure 2. Immediate questions are whether the inflation rate tends to converge to the natural rate of 1% and whether the prediction behavior of the participants has easily discernible properties affecting the inflation rate's degree of convergence. The results in Figure 2 are helpful in answering both questions. There is a clear separation between participants who manage to stabilize inflation around the natural rate and those who do not; also, at first sight many participants seem to have used simple prediction rules, with their choice of rules strongly affecting their success in forecasting the inflation rate.

Regarding the convergence of realized inflation to the natural rate of 1%,[10] the experimental results of most subjects fall into one of three qualitative categories that can be named "stable," "unstable" and "first unstable then stable." The first category consists of subjects that succeed from the beginning of the experiment in keeping the inflation rate close to the natural rate, without explicitly knowing this since they were not informed of the exact value of θ. In Figure 2, Graphs (b), (h), (i), (k), (l), (n), (o) and (p) depict the results of participants that can be put in this category, which is exactly half the number of subjects in the treatment. The second category contains subjects unable to stop inflation from fluctuating wildly. In the first treatment only Graphs (d) and (m) display this sort of movement, making permanent instability an uncommon but not practically irrelevant phenomenon. The third and last category has participants who manage to stabilize inflations around the natural rate, but only after going through a phase of substantial instability. Graphs (a), (c), (g), and to some degree (e) and (f) fit this description, representing almost a third of the first treatment and thus making this category the second most important.[11] In the end only Graph (j) does

---

law of motion (11) requires subjects to submit their expectations for the first two periods when no realized inflation values are yet available.

[10] The term "convergence" is used in this section in the sense of approaching relatively closely the natural rate of inflation, taking into account the exogenous disturbance term and the fact that inflation fluctuations can easily have an amplitude of 40 (see e.g. Graphs (d), (j) and (m) of Figure 2; also cf. the theo-retical bounds in (10)). In Section 5.2 a statistical test of convergence is performed on the infla-tion series of both treatments to determine unambiguously whether convergence is achieved or not.

[11] Classifying the participants of Graphs (e) and (f) as "from unstable to stable" requires some flexibility. Graph (e) shows a considerable reduction in inflation volatility from about period 30 but never a clear convergence (which is confirmed by the fluctuating prediction series), while in Graph (f) there is near-convergence starting in period 21 which breaks down though in the last periods of the experiment.
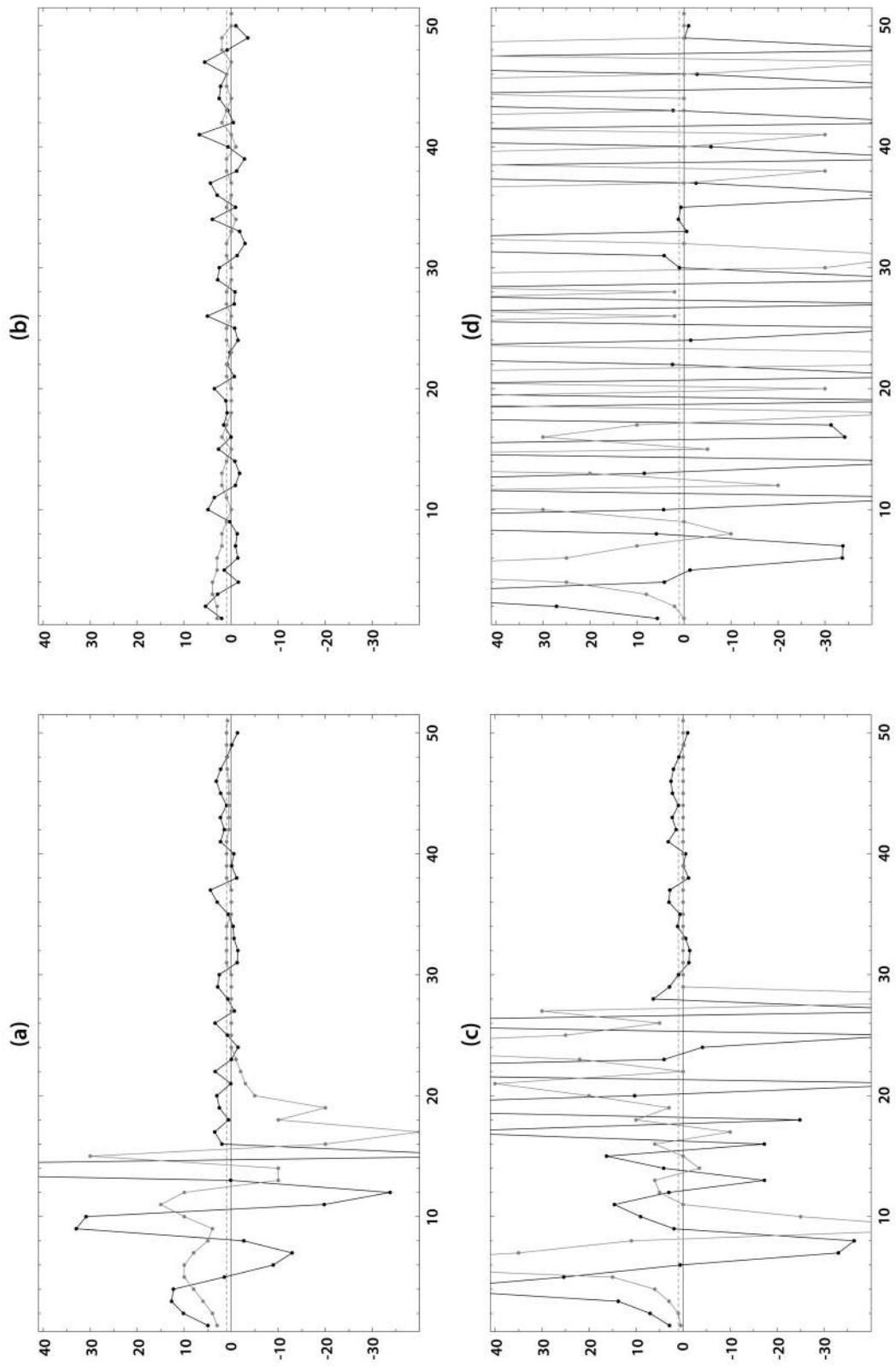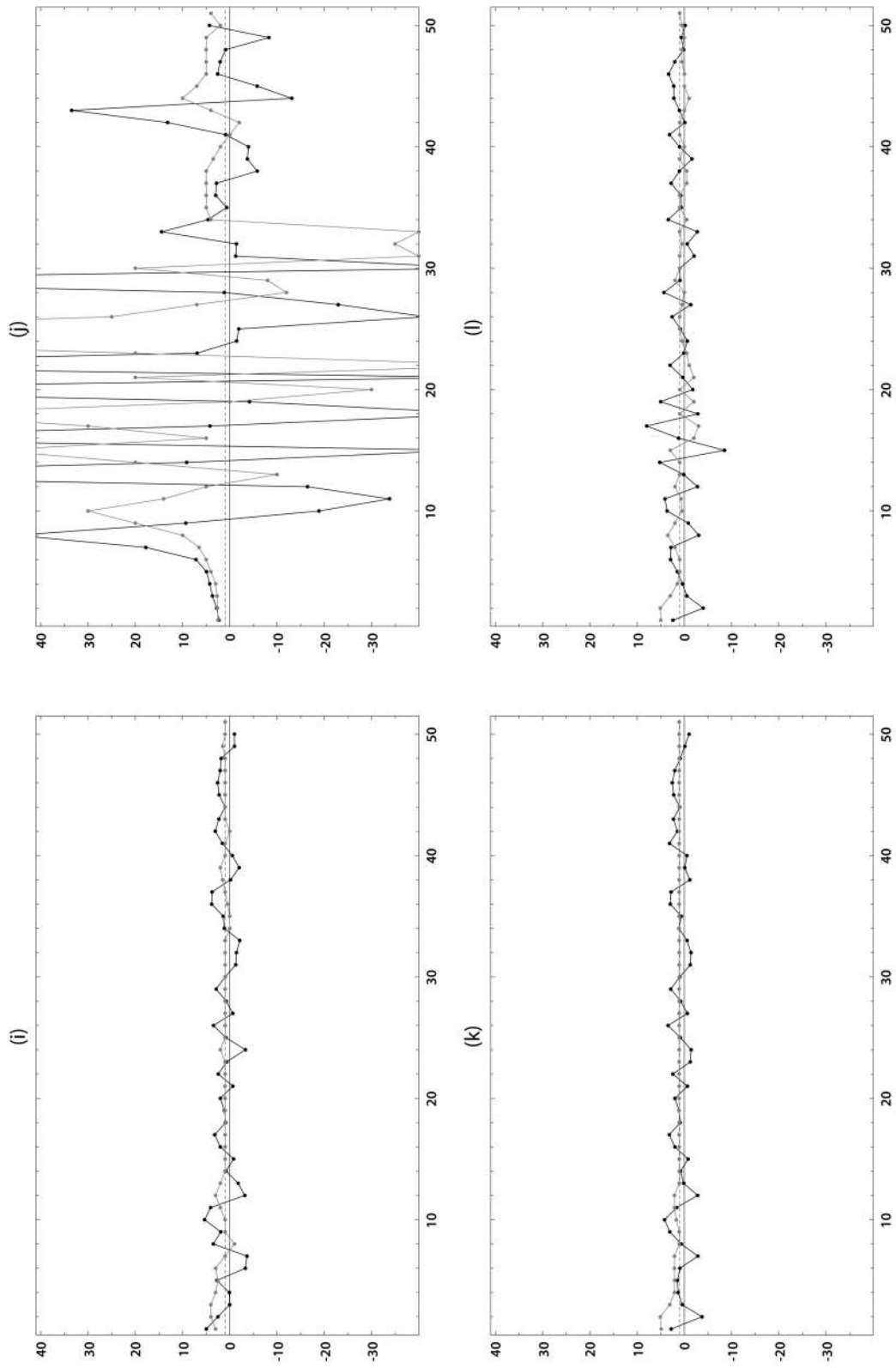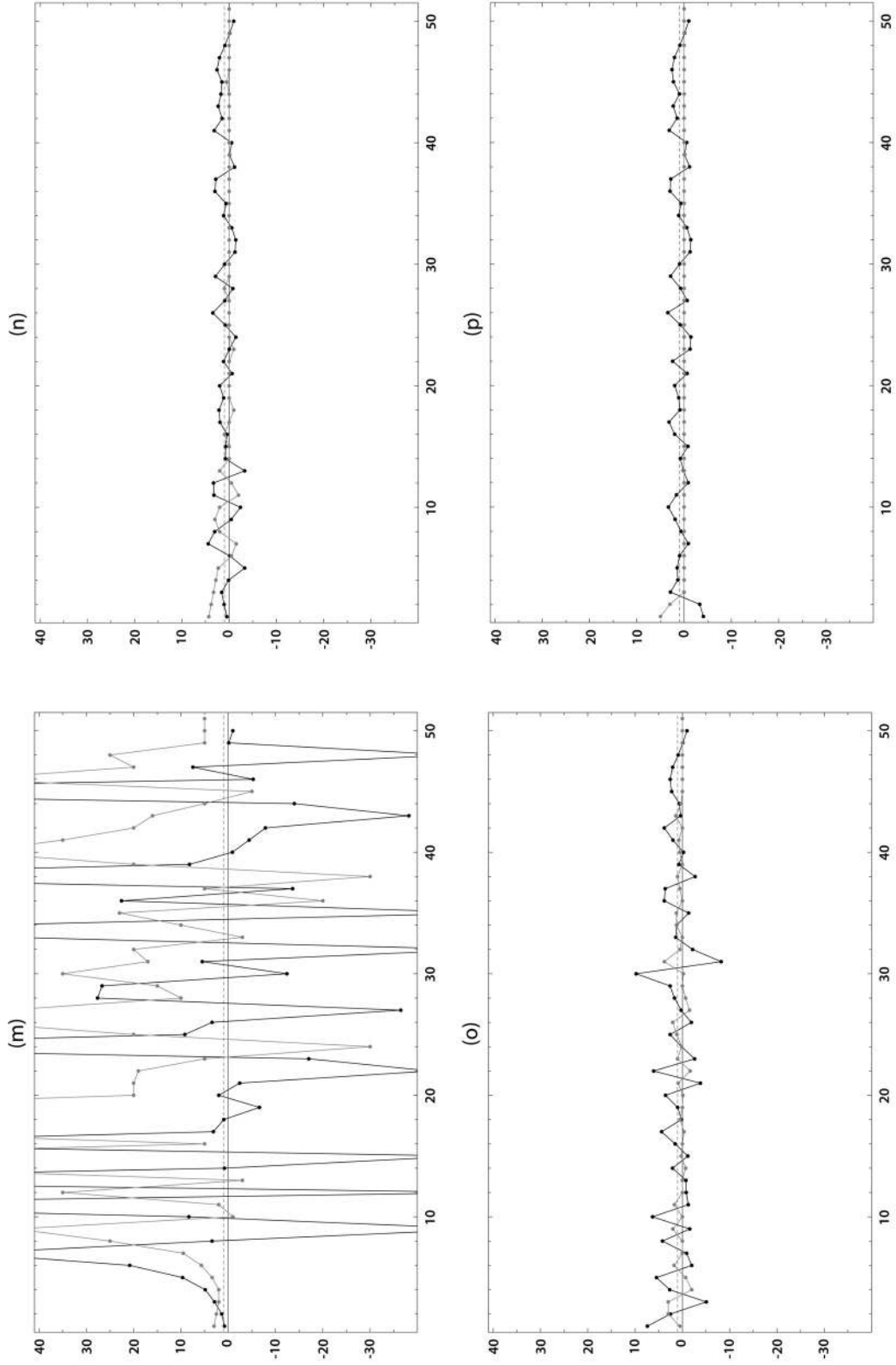
Figure 2 (part 1 of 4): Inflation rates, predictions and the natural rate of inflation for subjects 1 through 4 of Tr. 1 (LTh). Vertical axis is in percentages. Black lines are realized rates of inflation and gray lines inflation predictions. The dashed horizontal lines indicate the natural rate of inflation. *Continued on next page.*
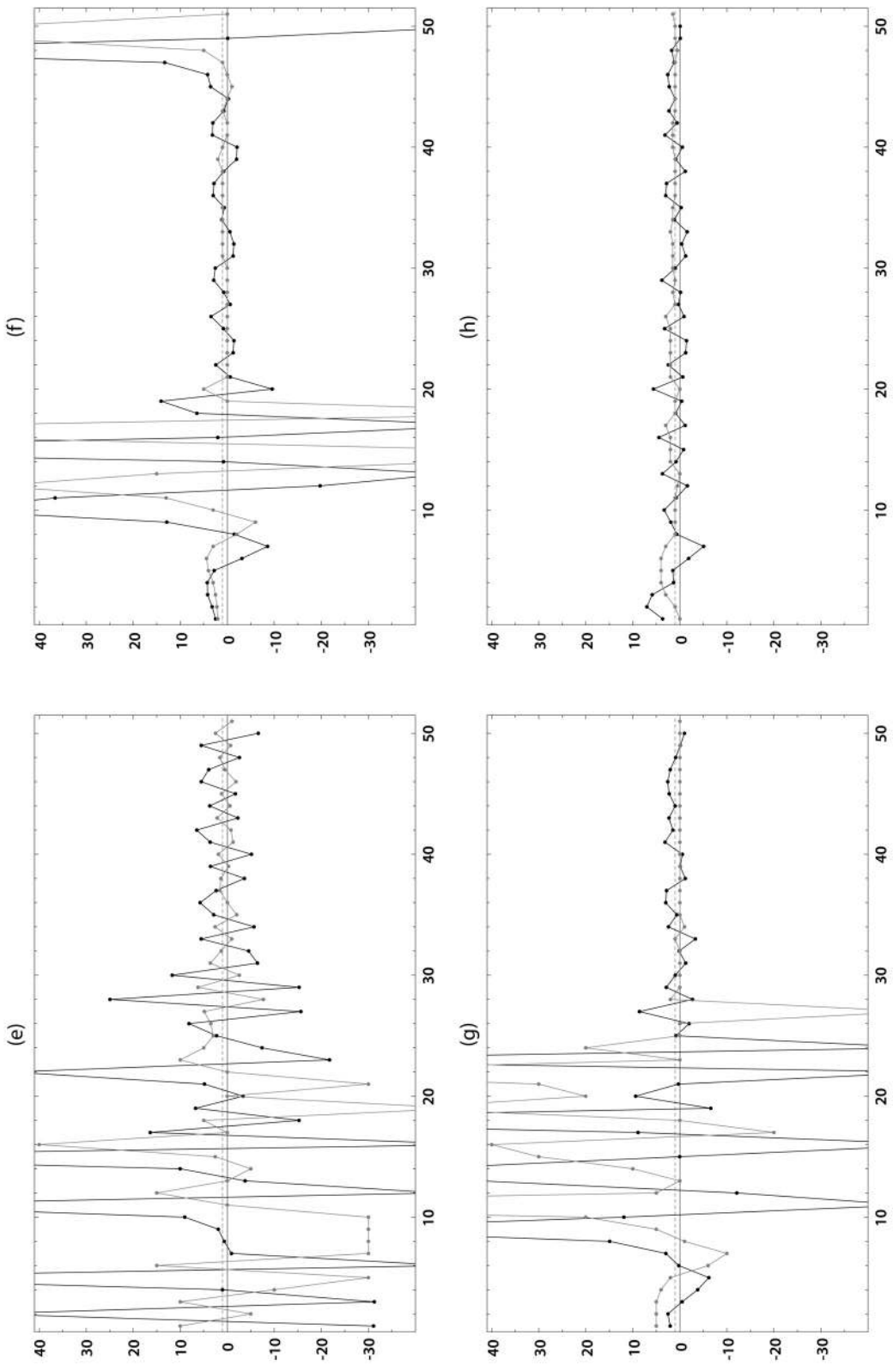
Figure 2 (part 2 of 4): Inflation rates, predictions and the natural rate of inflation for subjects 5 through 8 of Treatment 1 (LTh). Vertical axis is in percentages. Black lines are realized rates of inflation and gray lines inflation predictions. The dashed horizontal lines indicate the natural rate of inflation.

Figure 2 (part 3 of 4): Inflation rates, predictions and the natural rate of inflation for subjects 9 through 12 of Treatment 1 (LTh). Vertical axis is in percentages. Black lines are realized rates of inflation and gray lines inflation predictions. The dashed horizontal lines indicate the natural rate of inflation.

Figure 2 (part 4 of 4): Inflation rates, predictions and the natural rate of inflation for subjects 13 through 16 of Treatment 1 (LTh). Vertical axis is in percentages. Black lines are realized rates of inflation and gray lines inflation predictions. The dashed horizontal lines indicate the natural rate of inflation.

not easily fit any of the three categories. An initial phase of high volatility is followed up around period 30 by moderately successful attempts to control the inflation. These do not result in convergence though within the length of the experiment, making it impossible to classify the last 20 periods as either converging or unstable.

Looking at the inflation predictions of the participants grouped by the three categories introduced above, certain regularities stand out within each of them. Participants in the first, "stable" category tend to use adaptive predictions with a small adjustment parameter in the beginning of the experiment, while often fixing their predictions to specific numerical values or "focal points" after some time.[12] Graphs (i) and (k) clearly fit this description, with predictions slowly moving in the direction of recent realized inflations in the initial periods and sticking to specific values somewhere between periods 10 and 20. These values are very close but not identical: Graph (i)'s participant fixes his predictions exactly at the natural rate of 1%, but deviates from it several times during the experiment; Graph (k)'s participant on the other hand fixes at 1.11% from period 16 onwards.[13]

The prediction rules of the two participants falling in the "unstable" category, corresponding to Graphs (d) and (m) in Figure 2, are harder to analyse. Both start out the experiment with expectations that are approximately naive, with especially Graph (m) offering a clear demonstration of naivety in the first 8 periods. Under naive expectations the inflation rate diverges from the natural rate (as will be illustrated in Section 6.1; also cf. Figure 1), leading in both cases to a strong upward inflationary movement in the beginning of the experiment. What happens afterward is difficult to see in the graphs, because of the large amplitude of the fluctuations. Given that prediction rules will be investigated statistically in Section 6, it suffices to say that in neither case the subject discovers the key to stabilizing inflation, namely rigidity in the development of predictions. Until the end that is, since both

---

[12] Adaptive predictions are here predictions similar to Adaptive Expectations. See Section 6.1 for the definition of Adaptive Expectations in an OLG context and an illustration of its stability properties under different adjustment speeds.

[13] Within the stable category of subjects the most frequently submitted predictions are at or around the natural rate. The two most frequently submitted values are 0% and 1%, making up resp. 31.4% and 22.1% of all predictions within the category (which consists of 8 subjects). The maximum at 0% instead of the natural rate seems odd, but is caused mainly by the fact that Graph (p)'s participant chooses this value in 49 periods. The percentages fall to 27.5% and 14.0% when all subjects in the treatment are included.

participants keep their predictions constant in the last three periods, catching just a glimpse of its stabilizing effect on inflation.

Permanent high volatility in the inflation rate is exceptional in Treatment 1 (LTh), since most participants experiencing violent fluctuations manage to stabilize inflation after a while. Predictions of subjects in the "from unstable to stable" category have a combination of characteristics already observed in the "unstable" and "stable" categories. The initial phase of high volatility resembles the permanently high volatility of Graphs (d) and (m), in that it is started by expectations which are approximately naive (as in Graphs (c) and (f)) or adaptive with high adjustment (Graph (g)). The subsequent phase of relative stability is, like the prediction series in the stable category, dominated by fixation on focal values (e.g. 0 in Graphs (c) and (g), 0 and 1 in Graphs (a) and (f)), in some cases arrived at through an adaptive rule with slow adjustment (as in Graph (a), note that such a rule is also clearly used in Graph (j) after period 30). Of the five "first unstable then stable" series only the one in Graph (e) seems unaffected by these regularities, showing no signs of a naive or adaptive rule and no fixation at a certain prediction value. The identification of the rule that is used, if indeed a single prediction rule is used in this case, will therefore be left to the regression analysis of Section 6. There it will turn out that participant 5 in the second half of the experiment applies a linear rule slightly more complicated than the ones discussed above (see Table 2, Tr. 1).

The substantial number of participants experiencing high inflation volatility but managing to eliminate it during the experiment suggests that their prediction rules may have undergone structural change. In other words, participants may have changed prediction rules during the experiment to achieve better results. However, the fact that a majority of subjects was in an environment of uniformly high or low inflation volatility (10 out of 16 subjects fall into the "stable" or "unstable" category) is likely to limit the number of structural breaks. Also, while in many graphs in the "first unstable then stable" category there is a dramatic change in inflation development at the beginning of the convergence to the equilibrium rate, this does not necessarily mean that a similarly distinct change happened in the prediction rule. As is explained in Section 6.1, small changes in parameters of prediction rules may cause bifurcation points to be crossed and the stability of the monetary steady state to be reversed. The dynamics in Graph (a) are an example of this, being apparently driven by adaptive expectations with a fairly high adjustment rate in the beginning periods and a lower adjustment from around period 20 onwards. This means that a structural break should occur just before the middle of the experiment, though it also means that it may be small in terms of

the prediction rule parameters and therefore difficult to confirm through statistical testing. Structural breaks are further analysed statistically in Section 6.4.

4.2   High Equilibrium Inflation Rate Treatment

Increasing θ from 1% to 11% has an effect on inflation development that goes far beyond the shift in the equilibrium steady state and the increase in the amplitude of random disturbances (see the law of motion (11)). Looking at the 16 time series of Figure 3, corresponding with each of the subjects in Treatment 2 (HTh), it is clear that with the natural rate the volatility of inflation on the whole increases dramatically (average standard deviation in Tr. 1 is 20.28, in Tr. 2 44.68; a t test with H_0: $\sigma_\pi^2 = \sigma_\pi^1$ at 30 df. gives t value 3.620, $p < 0.01$). Compared to the first treatment, this increase in volatility clearly reflects a diminished ability of participants to stabilize the inflation dynamics. The dynamics of the second treatment are therefore more than just a scale increase due to the higher natural rate of inflation θ.

Classifying subjects' predictions in the same categories as in the previous subsection ("stable," "unstable" and "first unstable then stable"), a very different result is obtained. The number of subjects that manages to keep the inflation rate at or around the natural rate drops to 0 in the second treatment, from 8 in the first treatment, clearly indicating that the higher equilibrium rate of inflation made it harder for subjects to stabilize the system. The absence of early convergence in the second treatment is caused to a considerable degree by the fact that in the first two period, when no previously realized inflation rates were available, all participants submitted predictions lower than the equilibrium rate. In the directly following periods many adopted a prediction rule close to naive or high-adjustment adaptive expectations (clear examples are depicted in Graphs (a), (j) and (m) of Figure 3). Given the relatively low initial predictions, which in general underestimate actual inflation rates, these expectations lead first to an increased prediction, followed by a magnification of this increase in the realized inflation rate moving it far away from equilibrium (as implied by the law of motion (11)).

A substantial number of participants managed to move the inflation rate close to equilibrium in the course of the experiment, though several did not do so in the rational expectations sense, i.e. with their inflation predictions similarly close to the equilibrium rate. Six of the 16 subjects, corresponding to Graphs (b), (d), (e), (f), (j) and (k), can be assigned to the "first unstable then stable" category. Four more, shown in Graphs (a), (g), (i) and (o), are stabilizing later in the experiment but have not completed the process at period 50 (the subjects of Graphs (g) and (i) temporarily use focal prediction rules close to 0 after period 40;
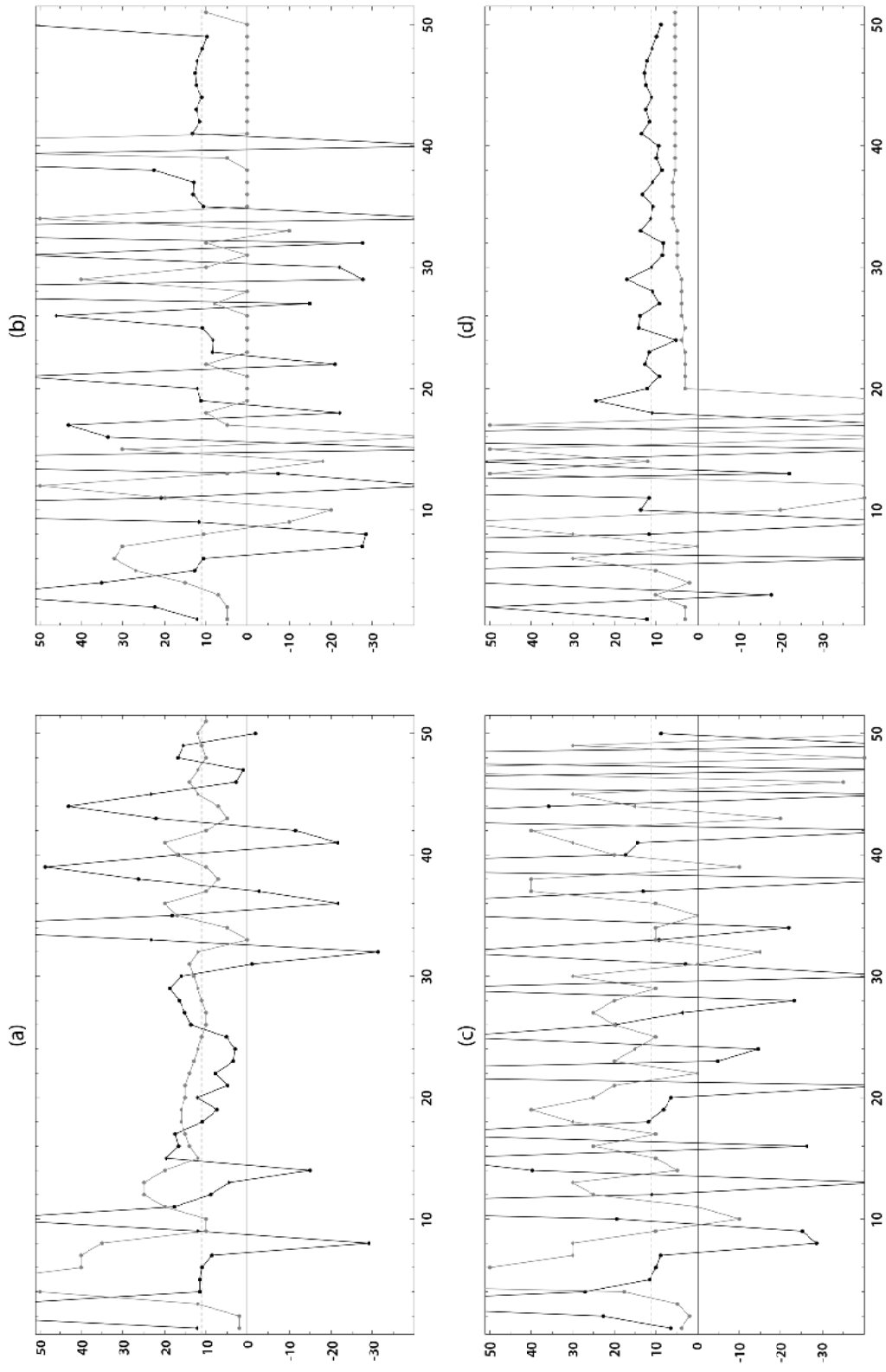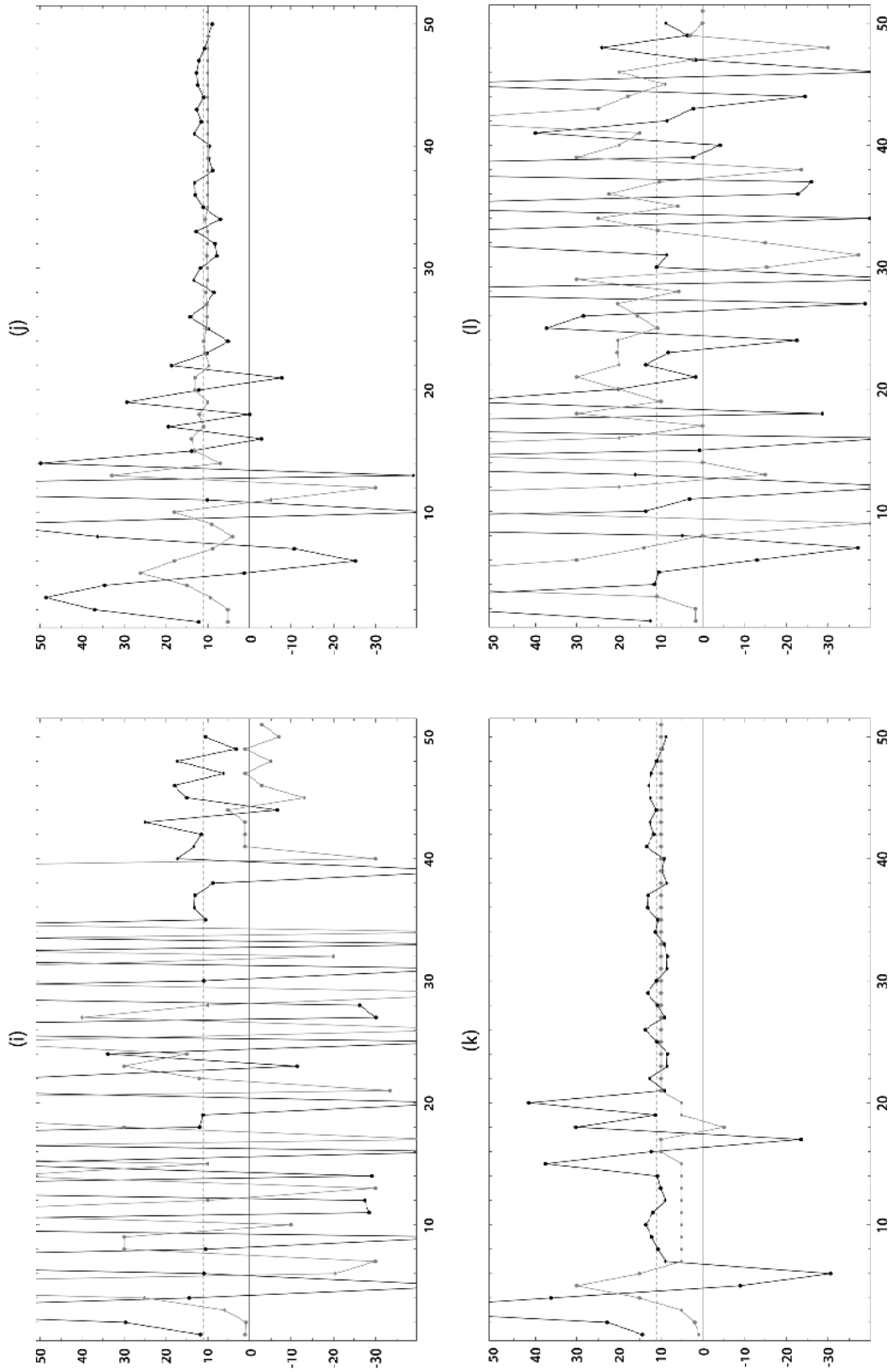
Figure 3 (part 1 of 4): Inflation rates, predictions and the natural rate of inflation for subjects 1 through 4 of Treatment 2 (HTh). Vertical axis is in percentages. Black lines are realized rates of inflation and gray lines inflation predictions. The dashed horizontal lines indicate the natural rate of inflation.
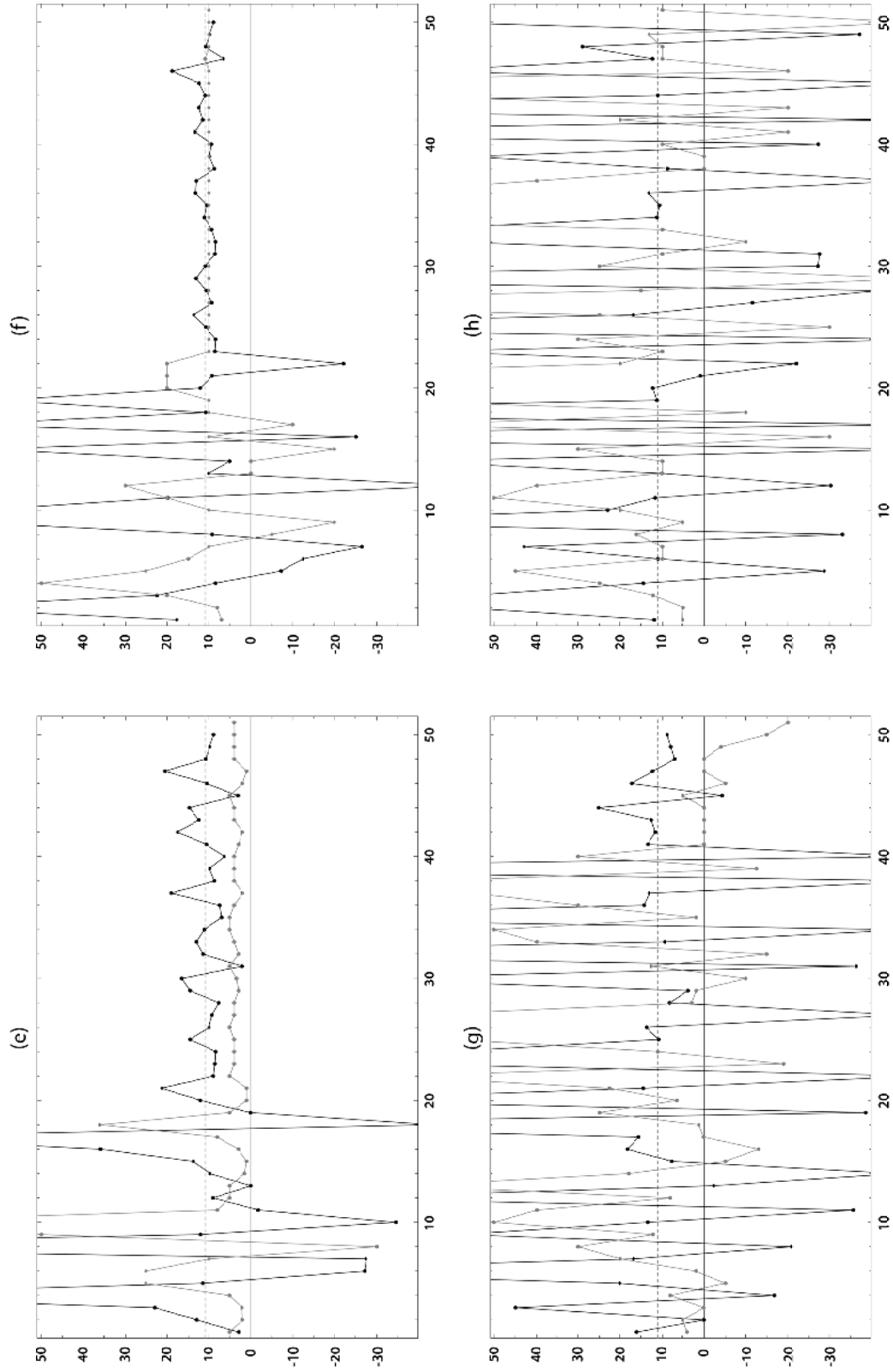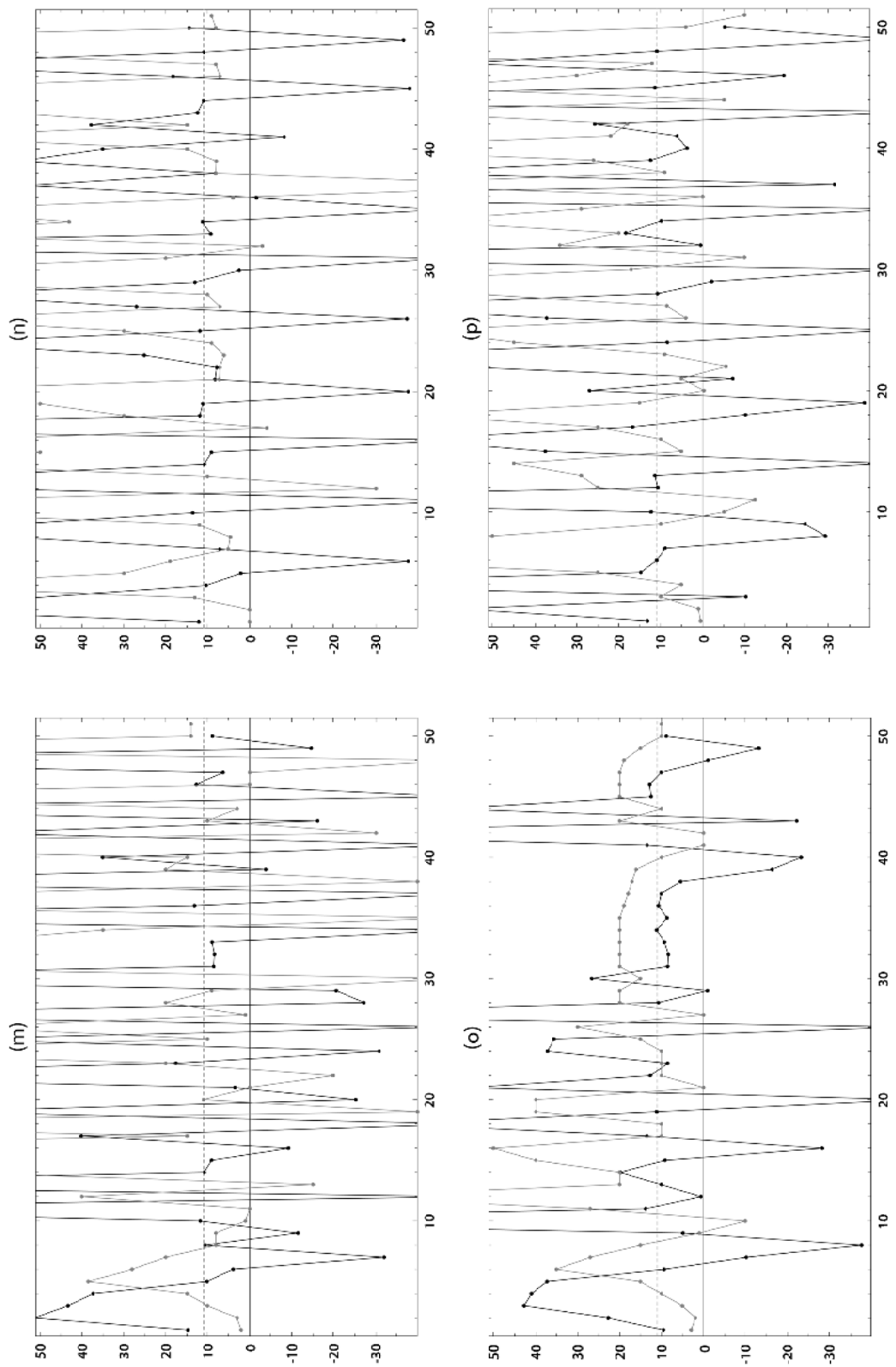
Figure 3 (part 2 of 4): Inflation rates, predictions and the natural rate of inflation for subjects 5 through 8 of Treatment 2 (HTh). Vertical axis is in percentages. Black lines are realized rates of inflation and gray lines inflation predictions. The dashed horizontal lines indicate the natural rate of inflation.

Figure 3 (part 3 of 4): Inflation rates, predictions and the natural rate of inflation for subjects 9 through 12 of Treatment 2 (HTh). Vertical axis is in percentages. Black lines are realized rates of inflation and gray lines inflation predictions. The dashed horizontal lines indicate the natural rate of inflation.

Figure 3 (part 4 of 4): Inflation rates, predictions and the natural rate of inflation for subjects 13 through 16 of Treatment 2 (HTh). The vertical axis is in percentages. Black lines are realized rates of inflation and gray lines inflation predictions. The dashed horizontal lines indicate the natural rate of inflation.

the subjects of Graphs (a) and (o) use rules close to slow-adjustment adaptive expectations in a substantial part of the experiment but are not careful enough in applying them). These two groups taken together form a majority of participants in Treatment 2 (HTh), indicating that the increased natural rate of inflation, makes it harder, but not impossible, for participants to converge to the rational expectations equilibrium. Looking at the prediction rules used by participants in the "first unstable then stable" category, similar patterns emerge as in the same category of Treatment 1 (LTh). Approximately naive or high-adjustment adaptive rules are often applied at the beginning of the experiment (see e.g. Graphs (f), Graph (j)) and low-adjustment adaptive or focal rules in the course of it (see e.g. Graphs (d), Graph (k)).[14]

The number of participants having a uniformly high inflation volatility throughout the experiment increases from 2 to 6 when moving from the first to the second treatment (the relevant graphs in Figure 2 are (d) and (m); in Figure 3 (c), (h), (l), (m), (n) and (p)). As in the first treatment (see Section 4.1), little more can be said about the unstable category than that its subjects tend to initially use adaptive prediction rules with high adjustment towards the most recent realized inflation rate (e.g. subject 12 of Graph (l) uses an approx. naive rule in roughly the first 13 periods of the experiment). Combined with the underestimation relative to the equilibrium inflation rate in the first two periods of the experiment, which holds for all participants in the second treatment, destabilization of the inflation development follows within a few periods. Since the high-adjustment adaptive expectations of the beginning are typically only used for a limited amount of periods, which is understandable given their extremely poor performance, subjects try to find a prediction rule better suited to their environment. What rules result from this process, and whether the changes in applied rules are large enough to show a structural break in the prediction series, are questions that cannot be answered by casual inspection of the data and are therefore left to the statistical analysis of Section 6.

4.3   Participants' Earnings

The total money earnings of participants, for the two treatments separately, are shown in

---

[14] Of course, the applied focal rules are often different from the ones in the first treatment because of the different equilibrium inflation rate. The prediction values most frequently chosen in the "first unstable then stable" category including the 4 boundary cases of Graphs (a), (g), (i) and (o) in Figure 3, are 10 and 0, making up 11.2% and 7.3% resp. of all predictions in the category (percentages drop to 10.4% and 6.3% for the whole treatment). Note that none of the subjects in Tr. 2 (HTh) repeatedly chooses the equilibrium value of 11.

Figure 4. There are substantial differences in earnings both within and between the treatments. Within treatments major differences exist, with the top-ranked participants earning more than five times as much as the bottom-ranked one. Unsurprisingly, the top-ranked subjects of Treatments 1 (LTh) and 2 (HTh) fall into the "stable" and "from unstable to stable" categories respectively as defined in Section 4.1, while the bottom-ranked subjects were labelled "unstable" and "from unstable to stable" respectively (the latter subject, nr. 9 in Tr. 2, having an extraordinary long and volatile "unstable" phase). This correspondence between the stability classification of the previous subsections and participants' earnings generalizes to all of Figure 4. In Graph (a) for example, the 8 earnings above 20 euros constitute the "stable" category, while the 2 earnings that are by far the lowest (nr. 4 and nr. 13) constitute the "unstable" category; in Graph (b) 5 out of 7 above-average earnings (excluding the unclassified nr. 1 and nr. 15) are from subjects in the "from unstable to stable" category, and all earnings associated with the "unstable" category are below average.

Two further remarks, looking at Figure 4, are first that the average earnings in both treatments are distinctly below the rational expectations average, and second that average earnings in the second treatment are distinctly below those of the first. Both statements are verified with one-sided t tests, giving resp. $H_0 : \mu_1 = \mu_1^{RE}$, $t \approx -3.91$ (15 df.), $p < 0.001$; $H_0 : \mu_2 = \mu_2^{RE}$, $t \approx -11.2$ (15 df.), $p < 0.001$; $H_0 : \mu_1 = \mu_2$, $t \approx 4.159$ (30 df.), $p < 0.001$. These tests confirm that subjects in neither treatment come close on average to the RE value, even though half of the subjects in the first treatment, corresponding to the "stable" category, do approach it very closely individually. Additionally, the greater difficulty subjects had in stabilizing inflation development under the high natural rate of inflation θ is reflected in the significantly lower average earnings of the second treatment.

5. Testing the Rational Expectations Hypothesis

In the description of the experimental model in Section 2 it was pointed out that in the unique Rational Expectations (RE) equilibrium an individual predicts the natural rate of inflation, as determined by the degree of seigniorage θ maintained by the central bank. The inflation rate then fluctuates around the natural rate with low amplitude such that the individual only makes prediction errors due to the random shocks in the economy, thus maximizing his expected earnings. As could be expected, none of the subjects in the experiment acted in accordance with this equilibrium; first because the instructions did not give them sufficient information to derive it theoretically and second because no market
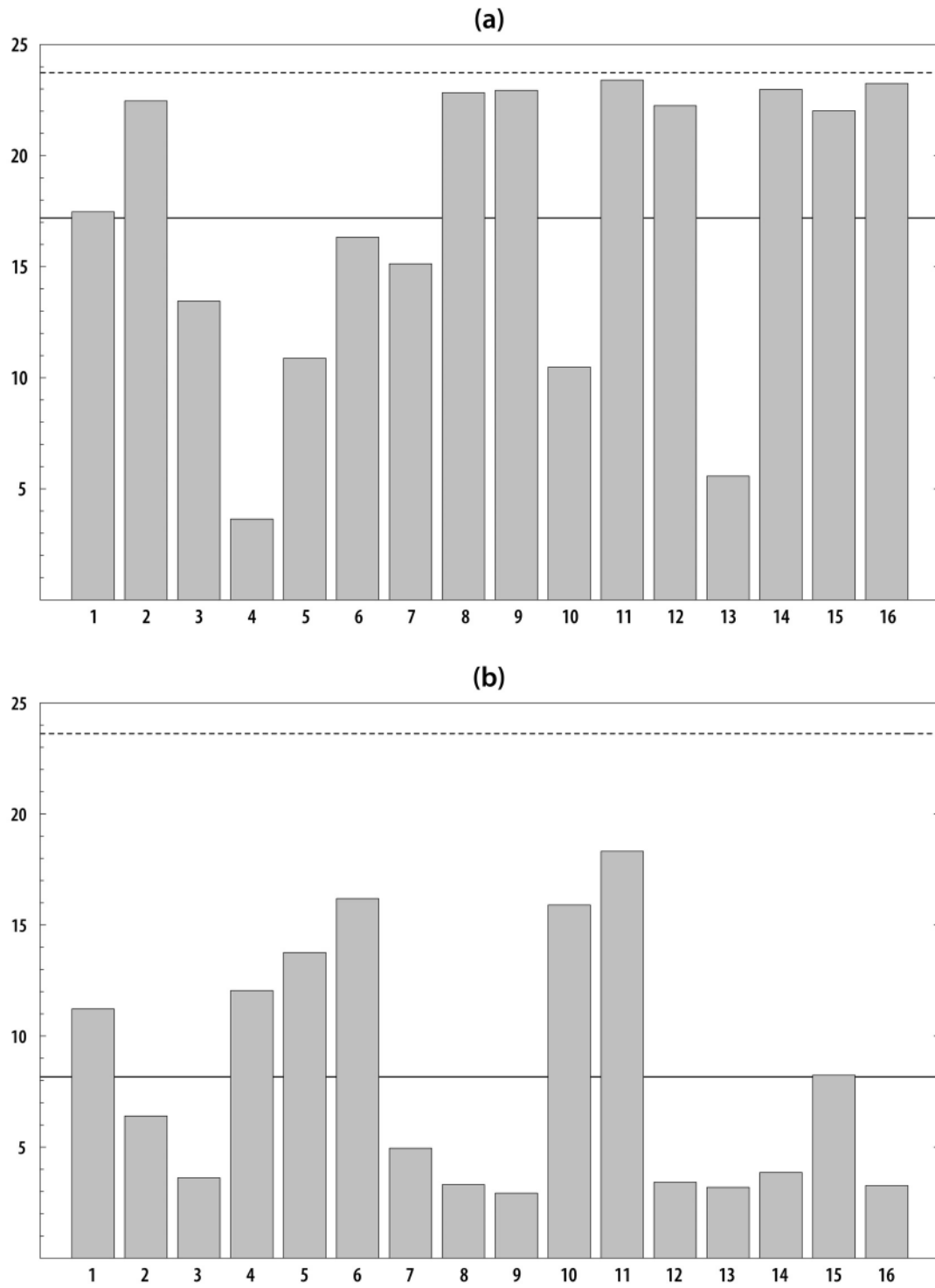
Figure 4: Earnings of the participants in Treatments 1 (LTh) and 2 (HTh) resp., in euros. Solid horizontal lines indicate the average earnings per treatment. Dashed horizontal lines indicate average earnings under the Rational Expectations equilibrium.

information was available during the first two periods, making it virtually impossible to guess the right inflation rate from the start. Nevertheless, a considerable number of participants was able to learn the constant RE equilibrium or approach it closely in the course of the experiment, making it potentially a useful long-run description of the situation in overlapping generation economies like the one employed in the experiment. The RE equilibrium is attractive as a long-run description of the economy's state because of its simplicity and because it optimizes the expected utility of all generations, given that the prediction errors are random and unpredictable. In order to judge the relevance of the RE equilibrium, first the absolute prediction errors of participants will be studied. These should be minimal in the case of perfect rationality. Then the compatibility with the Rational Expectations Hypothesis of the mean and standard deviation of realized inflation values is subjected to testing, while allowing for a suitable "learning phase" at the start of the experiment.

5.1 Absolute Prediction Errors and a Prediction Error Classification

The amplitude of the random disturbances in both treatments is 2.5% of the inflation rate (cf. (11)) and therefore varies with the level of inflation, increasing when inflation goes up and vice versa. However, in the unique RE steady state the distribution of disturbances remains constant, and by implication the distribution of absolute disturbances. The disturbance distribution under the RE equilibrium is uniform on the domain $[0.975\theta, 1.025\theta]$, with θ the natural rate of inflation. Assuming that an individual acts rationally and consistently chooses θ (or $(\theta-1)*100\%$ in the terminology of the experiment), his prediction errors are distributed uniformly on $[-0.025\theta, 0.025\theta]$, and his absolute errors uniformly on $[0, 0.025\theta]$. The latter distribution easily translates into a test of the RE equilibrium in any single period, since, at a significance level of 5%, the rational individual would make an absolute prediction error of no more than $19/20*0.025\theta \approx 0.0238\theta$ in 95% of all periods. This means that the one-sided critical values at 5% for the absolute prediction error under the RE equilibrium are 2.4% and 2.6% approximately for Treatments 1 (LTh) and 2 (HTh) respectively.

Figures 5 and 6 display absolute errors for all periods and both treatments. In Graphs (a) all errors less than or equal to 50 percent are plotted, with columns of dots linked to periods of a single treatment. In addition, the median errors, i.e. the averages between the eighth and ninth ranked errors, are connected and the critical values under the RE equilibrium, as calculated above, are indicated. The dashed lines at 40, following from the Linear Scoring Rule (13), draw the threshold above which predictions errors are so large that incremental earnings are zero. Graphs (b) are derived from Graphs (a), counting respectively

the number of participants with an absolute prediction error below the RE critical value, between the critical value and the zero incremental earnings point, and above this point. Consequently, the bars at the bottom of Graphs (b) indicate the number of subjects satisfying the REH in any single period, the middle bars count those too inaccurate for the RE benchmark but still accurate enough to be rewarded for their predictions, and the bars at the top represent the predictions so inaccurate as to be worthless to the participants.

In Treatment 1 (LTh) the accuracy of predictions both tends to increase during the experiment, as is shown by the decreasing trend in the median errors of Figure 5, and reaches a level at the end of it with a clear majority of all participants satisfying the REH (in the last four periods at least 10 subjects do so; in the last 10 periods the average number of REH-compatible errors is 9.7). Also, at the end of the experiment virtually all participants manage to keep their prediction error below 40, implying that they had achieved at least a basic understanding of their influence on the inflation dynamics. A test at 5% significance level of the RE equilibrium on all 16 subjects yields poor results however. Given that the critical value drawn in Graph (a) should not be exceeded in 95 percent of the cases, at least 13 out of 16 predictions should fall into the bottom category of Graph (b) under the REH, for the treatment as a whole. Graph (b) shows that this is only true for periods 43 and 50, so it cannot be said that the steady-state equilibrium is a good description of individual behavior for the entire first treatment. Whether there are subjects with realized inflation series within REH limits will be tested in the next subsection.

It is apparent from Figure 6 that absolute errors in Treatment 2 (HTh) are on the whole much higher than in Treatment 1 (LTh). In fact, for a considerable number of periods up to nearly period 20, the median error surpasses the zero earnings increment boundary of 40, meaning that at least half of the subjects earned nothing during those periods. At the same time the median error series clearly shows that the accuracy of predictions tends to rise during the experiment. At the end of the experiment however the median absolute error still is more than double the 95% critical value under the RE equilibrium, showing that the number of REH-compatible predictions, at least within the length of the experiment, never exceeds half of the participants. This statement is confirmed in Graph (b) of Figure 6, which makes clear more specifically that the number of prediction errors staying absolutely below the REH critical value in any period never exceeds 6. Obviously this means that the RE equilibrium cannot be assumed to hold for the second treatment as a whole, or for any subset of periods within the experiment. At the same time, Graph (b) shows that the number of prediction errors exceeding 40 tends to decrease, especially towards the end of the experiment, leaving
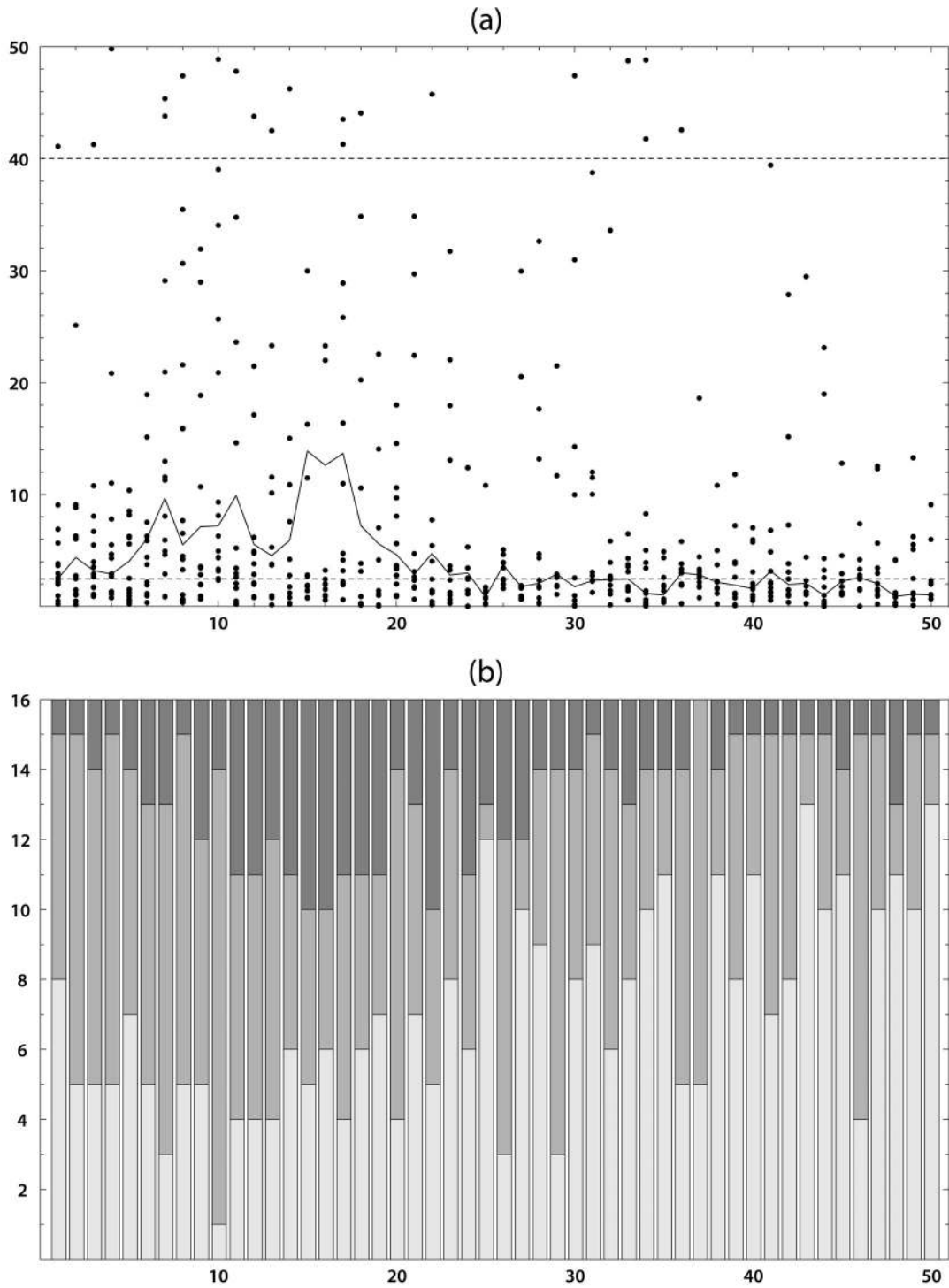
Figure 5: Absolute prediction errors for all 16 subjects of Tr. 1 (LTh), and classification of prediction errors based on the REH and Linear Scoring Rule (13). Dots in Graph (a) indicate absolute prediction errors for each period (some lie beyond the graph's range); the solid line connects the median errors; the lower dashed line shows the critical error value at 5% level under the REH; the upper dashed line shows the error value beyond which a subject's earnings in a period are zero. In Graph (b) the numbers of prediction errors are counted that satisfy the REH (5% level), yield positive earnings and yield no earnings resp., starting at the bottom of the bars.
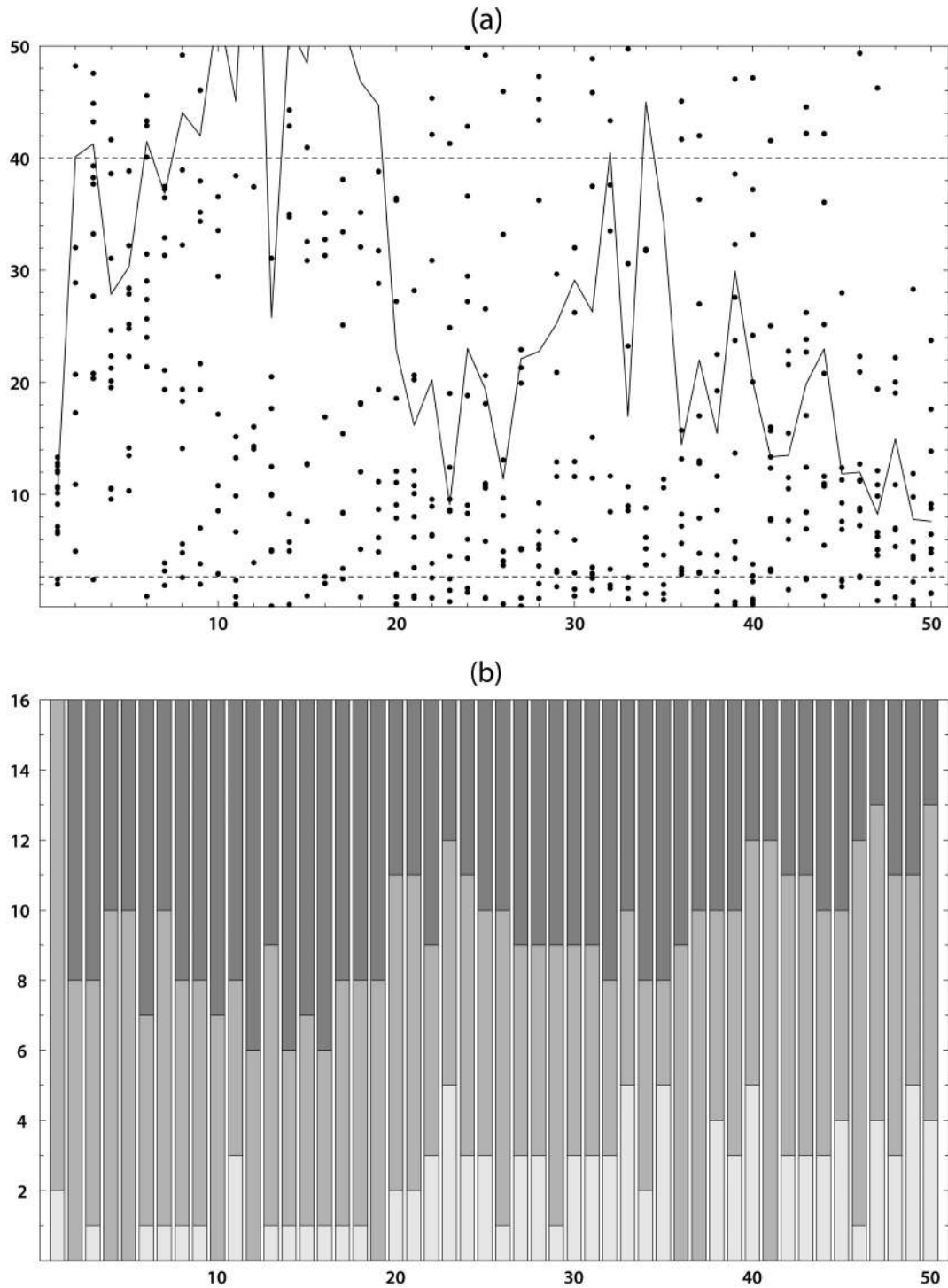
Figure 6: Absolute prediction errors for all 16 subjects of Tr. 2 (HTh), and classification of prediction errors based on the REH and Linear Scoring Rule (13). Dots in Graph (a) indicate absolute prediction errors for each period (some lie beyond the graph's range); the solid line connects the median errors; the lower dashed line shows the critical error value at 5% level under the REH; the upper dashed line shows the error value beyond which the subject's earnings in a period are zero. In Graph (b) the numbers of prediction errors are counted that satisfy the REH (5% level), yield positive earnings and yield no earnings resp., starting at the bottom of the bars.

only a small minority with worthless predictions in the last periods (in periods 41 through 50 the average number of participants in the zero incremental-earnings category of Graph (b) is 4.6).

We have found that prediction errors in both treatments tend to decrease during the experiment and that, especially in Treatment 2 (HTh), errors in the beginning of the experiment are at a level motivating many subjects to change their prediction rules. It is therefore necessary for the further analysis of the experimental results to exclude an initial phase in which participants are often forming their expectations and have not yet chosen a definite prediction rule. The exclusion of such a "learning phase" makes the assumption of stationarity in the prediction series more credible, which is important in the statistical derivation of prediction rules in the next section. An objective criterium could be defined to determine the length of the learning phase at the treatment- or participant-level, but because of the complexity involved a more heuristic approach is chosen here. Since the median absolute prediction error (see Graphs (b) of Figures 5 and 6) in both treatments makes a substantial drop in the periods leading to period 20 and has local minima in period 21, the first 20 periods of the experiment are designated "learning phase," for all subjects of both treatments. This means there are 30 periods for the analysis of individual prediction rules, leaving a sufficient amount of degrees of freedom to make the estimation results meaningful.

5.2   Testing Inflation Means and Standard Deviations against the REH

In order to determine whether the inflation series generated by participants satisfy the Rational Expectations Hypothesis, the inflation means and standard deviations are compared to their approximate equilibrium distributions. Approximations of the equilibrium distributions of the first two inflation moments under the REH are used because the exact distributions, given the uniform distribution of the disturbances, are very complicated. The Central Limit Theorem (CLT) gives an alternative due to the 30 available observations per subject. In order to reduce the chance of making type I errors for subjects who learn to act rationally only later in the experiment, the sample sizes are limited to 30 by excluding the learning phase determined in Section 5.1. Table 1 lists inflation averages and standard deviations for all subjects of both treatments and compares these with the critical values under the RE equilibrium.

The left upper part of the table deals with the inflation means of Treatment 1 (LTh). Remarkably, all of these are positive, while the natural rate in the first treatment is only

Tr.	Subj.	Mean π_t	Cr.val.(5%)	Cum.Pr.	H_0	St.Dev. π_t	Cr.val.(5%)	Cum.Pr.	H_0
1	1	0.962	{0.478; 1.522}	44.4	Not Rej.	1.721	{1.084; 1.830}	92.2	Not Rej.
1	2	0.789	id.	21.4	Not Rej.	2.633	id.	100	Rej.
1	3	2.867	id.	100	Rej.	33.49	id.	100	Rej.
1	4	17.83	id.	100	Rej.	72.12	id.	100	Rej.
1	5	1.601	id.	98.8	Rej.	12.18	id.	100	Rej.
1	6	2.711	id.	100	Rej.	21.17	id.	100	Rej.
1	7	0.774	id.	19.8	Not Rej.	23.78	id.	100	Rej.
1	8	0.787	id.	21.2	Not Rej.	1.618	id.	81.9	Not Rej.
1	9	0.826	id.	25.7	Not Rej.	1.892	id.	98.8	Rej.
1	10	3.879	id.	100	Rej.	36.87	id.	100	Rej.
1	11	0.820	id.	25.0	Not Rej.	1.585	id.	77.0	Not Rej.
1	12	0.949	id.	42.5	Not Rej.	1.731	id.	93.0	Not Rej.
1	13	7.465	id.	100	Rej.	46.97	id.	100	Rej.
1	14	0.820	id.	24.9	Not Rej.	1.514	id.	64.7	Not Rej.
1	15	0.812	id.	24.0	Not Rej.	3.233	id.	100	Rej.
1	16	0.820	id.	25.0	Not Rej.	1.585	id.	77.0	Not Rej.
2	1	11.99	{10.43; 11.57}	100	Rej.	22.64	{1.192; 2.012}	100	Rej.
2	2	20.30	id.	100	Rej.	48.57	id.	100	Rej.
2	3	24.63	id.	100	Rej.	66.40	id.	100	Rej.
2	4	11.11	id.	64.7	Not Rej.	2.259	id.	99.9	Rej.
2	5	11.14	id.	68.9	Not Rej.	4.606	id.	100	Rej.
2	6	9.66	id.	0.0	Rej.	6.450	id.	100	Rej.
2	7	19.21	id.	100	Rej.	55.34	id.	100	Rej.
2	8	24.26	id.	100	Rej.	65.13	id.	100	Rej.
2	9	24.24	id.	100	Rej.	62.90	id.	100	Rej.
2	10	10.31	id.	0.9	Rej.	4.275	id.	100	Rej.
2	11	10.81	id.	24.9	Not Rej.	1.744	id.	77.4	Not Rej.
2	12	16.40	id.	100	Rej.	47.66	id.	100	Rej.
2	13	30.40	id.	100	Rej.	71.63	id.	100	Rej.
2	14	18.27	id.	100	Rej.	42.56	id.	100	Rej.
2	15	16.84	id.	100	Rej.	36.84	id.	100	Rej.
2	16	22.31	id.	100	Rej.	58.19	id.	100	Rej.

Table 1: Testing the Rational Expectations Hypothesis (REH) through the mean and standard deviation of the inflation rate. The first two columns show treatment and subject number of the tested inflation series. In the 3rd through 6th columns inflation means are compared to their REH distribution; in the 7th through 10th columns inflation standard deviations are compared to their REH distributions. The 4th and 5th column resp. show critical values at 5% level (normal approximation) and the cumulative probability of the observed inflation mean under the null hypothesis; the 8th and 9th columns show the same relative to the inflation standard deviations.

slightly above zero (when expressed in percentages). This makes it easy for the realized inflation to drop below zero, as is confirmed by the graphs of Figure 2. Comparing the inflation means with the critical values under the REH gives mixed results. For 10 out of 16 subjects the null hypothesis of rational behavior is not rejected, while the remaining subjects all have inflation means significantly higher than the equilibrium rate. When restricting attention to the inflation mean, the majority of subjects is therefore in accordance with the REH, though this majority is smaller than the number of 14 required under the null hypothesis that the REH holds for the treatment as a whole.[15] Looking next at the right upper part of Table 1 it is seen that the REH fares slightly worse when testing the standard deviation in the first treatment. Six of the subjects, all of whom have an inflation mean consistent with the REH, have an inflation volatility low enough not to reject the possibility of rational behavior, while the others exceed the upper critical value. Similar to the mean, the REH is unfit as a descripton of inflation volatility in Treatment 1 (LTh) as a whole, but for 6 out of 16 subjects the REH as a post-learning phase description of realized inflation rates cannot be rejected.

In the second treatment the REH fails almost completely as a description of inflation development, as could have been expected from the much higher degree of instability and lower accuracy of predictions compared to the first treatment, observed in Sections 4.2 and 5.1 respectively. Only in three cases does the inflation mean fall outside the critical region, while the standard deviation with a single non-rejection does even worse. Subject 11, who uses a focal rule at 10% starting exactly after the end of the learning phase, is the only subject having both an inflation mean and standard deviation in accordance with the REH. Subjects 6 and 10, who are the only ones with inflation means significantly below the equilibrium value, experienced very low inflation values directly following the learning phase, after which they both committed to near-focal strategies. The rejection of both tests for these individuals is mainly due to the fact that their learning phases were a few periods longer than the standard 20 periods.[16] Summarizing, it is reasonable to say that the REH is a good description of

---

[15] Assuming the REH is true, the number of non-rejections is binomially distributed with success rate 95% and 16 trials. The critical value at 5% significance level for a one-sided test is 14.

[16] Extending the learning phase for both individuals by two periods to 22 in total, the averages are increased to 10.81 and 10.65 for subjects 6 and 10 resp. Adjusting for the change in the number of remaining observations, cumulative probabilities become 26.4% and 12.5%, cancelling the rejections

almost a third of the inflation series when looking at their means. At the same time, the REH cannot explain the standard deviations, with 15 out of 16 inflation series exhibiting significant excess volatility. Obviously, this means that the REH is not suitable as a description of inflation development for Treatment 2 (HTh) as a whole.

6. Estimation of Individual Linear Prediction Rules

It has been observed in the experimental results (see Section 4) that small changes in the prediction rules used by participants can have major consequences on the stability of the inflation development. Typical rules that were found to lead to unstable inflation series were close to adaptive expectations with a high adjustment towards the latest realized inflation rate, while many subjects that succeeded in stabilizing inflation seemed to apply adaptive expectations with a low adjustment parameter. The fact that minor changes in the parameters of prediction rules can have a major impact on qualitative dynamics emphasizes the need to accurately determine these parameters for each participant separately.

A difficulty that is likely to arise when estimating individual prediction rules is that the high volatility of realized and predicted inflation rates observed with many subjects makes it hard to detect the underlying structure in the prediction series, if indeed any consistent structure exists. Another difficulty is the possibility of structural breaks occurring in the prediction series of participants. Many of them, especially in the second treatment, learned to stabilize inflation development only later in the experiment. This suggests that they switched between different rules at least once, which compromises the representativeness of a single estimation on their prediction series. And finally, the specification of the applied prediction rules is not a priori clear. Subjects may have applied more complicated rules than the linear ones fitted on their predictions in this section, though given the prevalence of adaptive- and focal-like rules observed in Section 4 it seems unlikely that non-linearities have been used extensively.

In this section several well-known prediction rules will be simulated for both treatments of the experiment to see what kind of inflation dynamics they produce. Then, a general linear prediction rule specification is estimated on the expectations of all participants in order to detect any linear structure in them. Next, the linear specification is restricted to the so-called first-order heuristics form, with the aim of obtaining a simple interpretation of the

---

for both subjects. Standard deviations both decrease to 2.433 and 2.233, but still substantially exceed the upper critical value, giving cumulative probabilities of 100% and 99.9% resp.

subjects' expectation formation. The section ends with a verification of the hypothesis that multiple prediction rules were used in succession by individual participants.

6.1 Simulation of Several Benchmark Prediction Rules

Before proceeding with the estimation of prediction rules used by participants, several simple rules are simulated in order to study their effects on inflation development. Comparing these effects with those observed in the experiment gives a better idea of the kind of rules that have been applied by various participants. Several variations of rules that were repeatedly observed in the experimental results in Section 4 are applied in the model for a large number of periods and both treatment values of the equilibrium inflation rate θ. The first set of simulated rules includes naive expectations and averaging expectations over the last three and all available realized inflations respectively. These rules are part of one family of rules since naivety can be seen as "averaging" over only the most recently available inflation rate. The second set of simulated rules consists of adaptive expectations, with parameter values representing "high," "medium" and "low" adjustment successively. All of these prediction rules are simulated for 100 periods, i.e. twice the length of the experiment, to show what their effects are beyond the experimental time limitation while still being able to plot the results such that all details are easily discernable.

Figure 7 shows the simulation results for naive and averaging expectations. Given the information (12) available to subjects during the experiment, these rules are defined as follows:

$$\text{(I)} \ \pi_t^e = \pi_{t-2};$$

$$\text{(II)} \ \pi_t^e = \frac{1}{3}\left(\pi_{t-2} + \pi_{t-3} + \pi_{t-4}\right); \tag{14}$$

$$\text{(III)} \ \pi_t^e = \frac{1}{t-2}\sum_{k=2}^{t-1}\pi_{t-k} \quad \text{if} \quad t > 2; \quad \pi_t^e = \pi_{t-2} \quad \text{otherwise.}$$

Note that these prediction rules are not defined for the first two periods and that Rule II is also not defined for the third and fourth period. In the simulation this problem was solved by adding initial random inflation values to make all rules well-defined for all $t \geq 1$.[17] Rule I, naive expectations, is simulated in Figure 7(a) under Treatment 1 (LTh) conditions. The prediction series copies the most recent realized inflation which, as has been observed

---

[17] These random inflations have a uniform distribution ranging from -10% to 10% plus the equilibrium rate of inflation. In the simulations the random disturbances were kept constant between runs, but the initial inflation rates were redrawn.
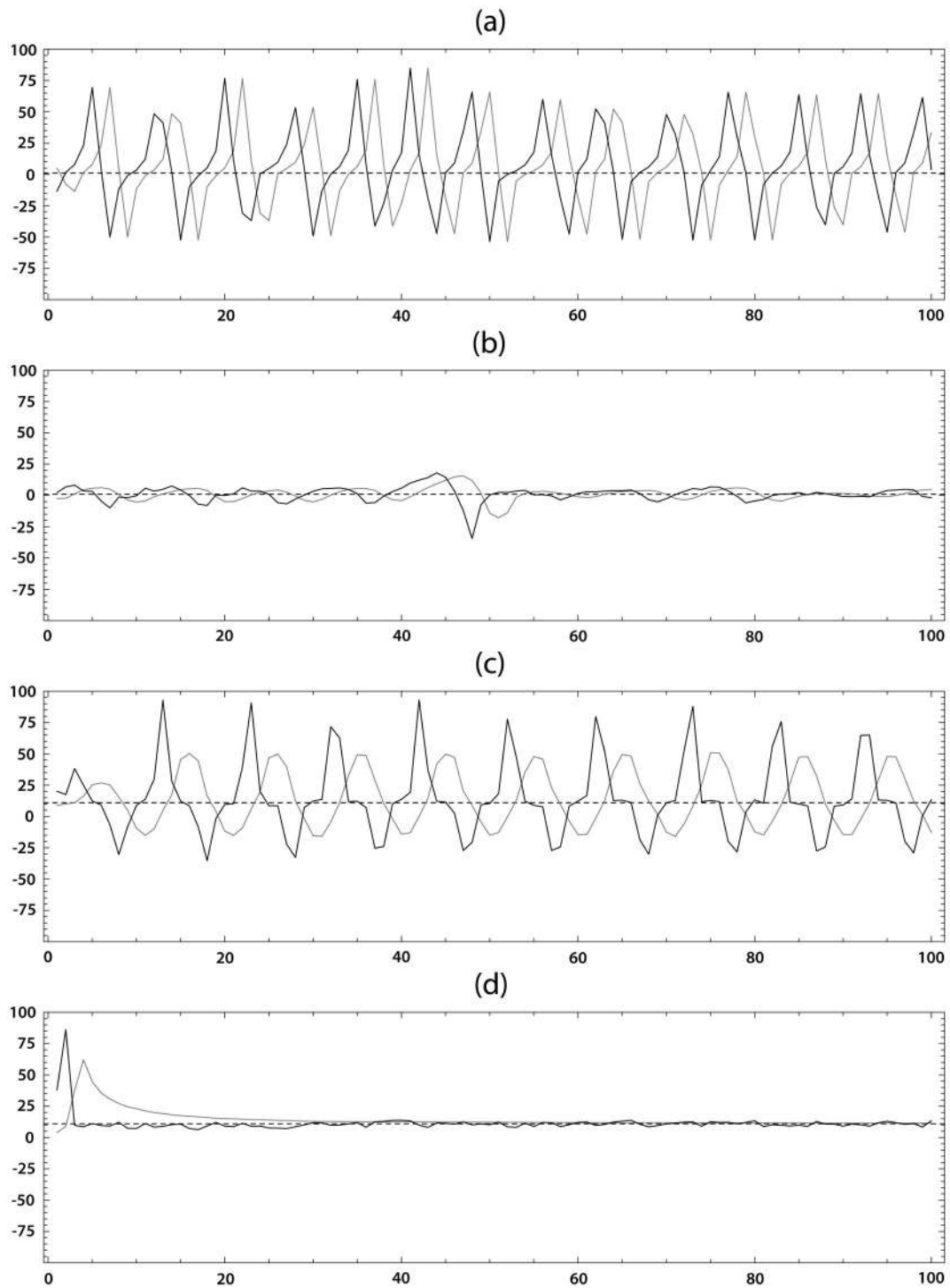
Figure 7: Simulations of inflation rate development under Averaging Expectations, for both treatments. Black lines connect inflation rates; gray lines inflation predictions; dashed lines indicate the fundamental inflation rate. Graph (a) displays Naive Expectations under Tr. 1 (LTh); Graph (b) averaging over the last 3 inflations under Tr. 1; Graph (c) averaging over the last 3 inflations under Tr. 2 (HTh); Graph (d) averaging over all previous inflations under Tr. 2. Disturbances are identical across graphs (but different from those in the experiment).

repeatedly in the experimental results, leads to wild and persistent fluctuations in the inflation development. There is a near-repeating pattern in the series of realized inflation rates, with a period of roughly 5 to 10 periods, which gradually changes because of the random disturbances. Viewing Rule I as an averaging rule, it can be expanded into Rule II by adding two inflation lags. Graph (b) of Figure 7 demonstrates that the three-term averaging rule immediately stabilizes inflation development in Treatment 1 (LTh). The one diverging movement in realized inflation, between periods 40 and 50, is caused by exceptional values of the disturbances around period 44 and illustrates that the stability of the dynamics under Rule II is not produced by favorable initial values.[18] Why this must be so can be seen in Figure 1(b): moving along the left vertical line (the first treatment) from the bottom dashed line to the third one from the bottom (corresponding to Rules I and II resp.), the elasticity curve is crossed, indicating that the second rule produces stable dynamics.

Graph (c) of Figure 7 also shows the effects of Rule II, but under Treatment 2 (HTh) conditions. Looking back again at Figure 1(b) it is easy to understand why the inflation series loses its stability when going from Graph (b) to Graph (c) in Figure 7, since this amounts to moving to the right between the vertical lines along the "Average (3)" dashed line. One then crosses the elasticity curve again and ends up far below it, resulting in unstable dynamics. The fact that Prediction Rule II suffices to keep inflation stable under the low equilibrium value of θ but not under the high one is an example of the more stringent stability requirements for the second treatment, which is reflected in the greater difficulty experienced by subjects in Treatment 2 (HTh) to keep the inflation rate from diverging. As in Figure 7(a), in Figure 7(c) a regular pattern emerges in both predicted and realized inflations that does not disappear in the long run. The amplitude of the predictions is distinctly lower than that of the inflations since Rule II averages over inflation lags and therefore dampens any fluctuation. The pattern is again slowly changing because of the random disturbances.

As Figure 1(b) shows, adding a fourth inflation lag to Rule II will not stabilize inflation development under the second treatment. Stabilization is achieved however when the number of inflation lags is set to maximum, i.e. the mean of all available realized inflation rates is taken as the prediction. Rule III is this averaging rule, with the addition, in order to make it well-defined for all $t > 0$, that it reverts to naive expectations when less than two realized inflations are available. Figure 7(d) shows that Rule III causes the inflation rate to

---

[18] The distribution of the disturbances is uniform with a maximum amplitude of 0.025 θ, meaning 2.5% in Tr. 1 and 2.8% in Tr. 1. See also the introduction of the law of motion (11) in Section 3.

converge to the natural rate of the second treatment, though it takes many periods for the predictions to also approach the natural rate closely due to an exceptionally high realized rate in the beginning of the simulation. This outlier influences the predictions for such a long time because it is used, with hyperbolically declining weight in the determination of all predictions following its realization.

Averaging expectations clearly increase the stability of inflation development as the number of included terms increases, with naive expectations producing strong fluctuations in Treatment 1 (LTh) and averaging expectations over all available realized rates resulting in fast convergence in Treatment 2 (HTh). At least three terms are necessary for convergence in the first treatment, and more than four in the second. An alternative way to achieve the same qualitative results is by applying adaptive instead of averaging expectations, thereby using information about previous inflation rates contained in the most recent prediction to simplify the prediction rule. Because realized inflations are always one period behind predictions in the experiment, adaptive expectations takes the following form:

$$(\text{IV}) \; \pi_t^e = \pi_{t-1}^e + \alpha \left(\pi_{t-2} - \pi_{t-1}^e \right), \tag{15}$$

with α the parameter of adjustment from the previous prediction π_{t-1}^e in the direction of the last realized inflation rate π_{t-2}. The effects of applying the above rule for several values of α and for both treatments are displayed in Figure 8.

Figure 8(a) starts out with a relatively high adjustment parameter $\alpha = 3/4$ in Treatment 1 (LTh). This evidently does not stabilize the inflation development. The persistent fluctuations are unsurprising since adaptive expectations with $\alpha = 1$ is equivalent to naive expectations (Rule I in (14)), which resulted in strong and persistent fluctuations in Figure 7(a), so any adaptive rule with high enough adjustment will also lead to instability. The adjustment parameter in Rule IV is subsequently lowered to 1/2. This parameter shift suffices, as Graph (b) shows, to stabilize the dynamics, though some minor oscillations remain because of the influence of the random disturbances through the most recent inflation. A necessary condition for stabilizing inflation in the experimental model is to change predictions very little if at all, especially in an area between roughly 0 and 20 percent where the savings curve of the young population segment decreases steeply (see Figure 1(a)). Further lowering of α will therefore produce even closer convergence to the natural rate of inflation, though it will simultaneously slow the convergence of predictions towards the natural rate.

In Graph (c) of Figure 8, $\alpha = 1/2$ is maintained but θ is increased to Treatment 2 (HTh) level. The resulting destabilization makes clear that this adjustment value is still too
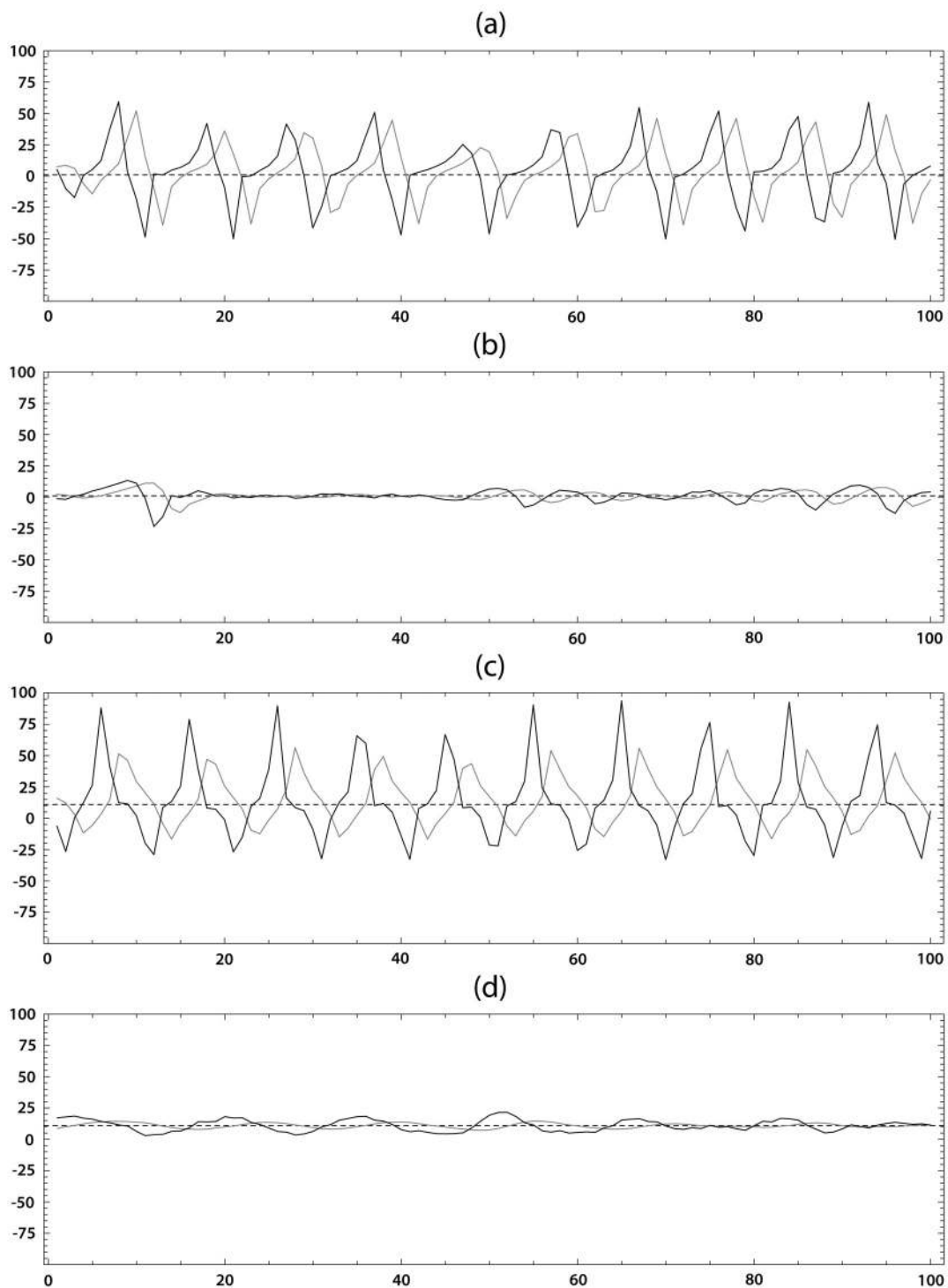
Figure 8: Simulations of inflation rate development under Adaptive Expectations (AE), for both treatments. Black lines connect inflation rates; gray lines predictions of those rates; dashed lines indicate the fundamental inflation rate. Graph (a) displays AE with adjustment parameter $\alpha = 3/4$ under Tr. 1 (LTh); Graph (b) AE with $\alpha = 1/2$ under Tr. 1; Graph (c) AE with $\alpha = 1/2$ under Tr. 2 (HTh); Graph (d) AE with $\alpha = 1/6$ under Tr. 2. Disturbances are identical across graphs (but different from those in the experiment).

high for convergence of the inflation rate under the second treatment. The repeating patterns in both inflations and predictions in Graph (c) look similar to those in Figure 7(c), with predictions following the wild inflation fluctuations in a dampened form. This is understandable because both Rule II and Rule IV with $\alpha = 1/2$ take averages of recent experimental values (the most recent realized inflation being in common) and therefore follow their movements but never exceed them. A further lowering of the adjustment parameter α to 1/6 does stabilize inflation under Treatment 2 (HTh) conditions, though even at this value minor but persistent oscillations in the inflation rate remain. Comparing Graphs (d) of Figures 7 and 8, it is seen that the inflation rate converges more closely under Rule III in Figure 7. The lower average distance to the equilibrium rate comes at a computational cost however, since Rule III after several periods consists of many and increasingly more terms than Rule IV. Also, the stability of the inflation development produced by the adaptive expectations rule can be enhanced by further lowering the adjustment parameter.

6.2   *Estimation and Complexity of General Linear Prediction Rules*

To uncover as much information on the participants' prediction rules as possible, using a standard estimation procedure, a general linear specification is fitted to all separate prediction series excluding their learning phases. We name this specification the Generalized Adaptive (GA) prediction rule. The parameters of the rule has been adapted to the information offered to participants during the experiment. It can be described as GA(4,3), meaning that is consists of 4 prediction lags and 3 inflation lags, resulting in the following regression model:

$$\pi_{i,t}^{e} = c + \sum_{k=1}^{4} s_k \pi_{i,t-k}^{e} + \sum_{l=2}^{4} o_l \pi_{t-l} + \varepsilon_t. \tag{16}$$

The four "subjective" and three "objective" coefficients, weighting the lagged predicti-ons and inflations respectively, are estimated in such a way that the rule as a whole best approximates the prediction series of a subject i. Since the learning phase generally needed by subjects to adjust themselves to the experimental environment was set in Section 5.1 to 20 periods, a total of 31 predictions and 30 realized inflation rates is available per subject to estimate (16). To avoid using experimental results from the learning phases for the lagged regressors, the effective number of predictions used for the dependent variable is 27 (exceptions were made in two cases; see footnote 19).

Table 2 lists the estimation results of an OLS regression.[19] Before looking at the coefficient values it is important to observe that in the great majority of cases the GA(4,3) specification yields estimated rules free of autocorrelation, meaning that by far the most coefficient estimates pass an important first test of being both efficient and unbiased. In the second treatment the 3 subjects whose predictions cannot be described with a GA(4,3) rule due to autocorrelation is substantially higher than that in the autocorrelation-free first treatment, but on the whole more than 90 percent of subjects can be represented by a linear prediction rule in a statistically valid sense. The fact that the R^2 statistic is often zero and, when positive, often not very high (means of positive R^2 values are 0.504 in Tr. 1 and 0.408 in Tr. 2), does not diminish the representativeness of the estimated rules. Many rules namely describe a focal or near-focal strategy with nothing but a constant, resulting in an R^2 of zero (examples are subject 7 of Tr. 1 and subject 6 of Tr. 2).

The first thing that can be said about the estimated coefficient values is that on average not many of them are significantly different from zero. In Table 3 the number of significant subjective and objective coefficients of the rules listed in Table 2 are counted for each treatment separately. The complexity of the estimated prediction rules, which is defined as the total number of significant explanatory variables, lies in both treatments on average not far above 1, with Treatment 1 (LTh)'s average exceeding that of Treatment 2 (HTh) by 0.3. Unsurprisingly, given these numbers, the amount of rules having multiple subjective or objective coefficients is negligible in both treatments. The most frequent rule by complexity, making up nearly half of the successfully estimated rules in both treatments, is the constant rule with zero complexity (recall that a learning phase of 20 periods was subtracted from the experimental results before estimating the prediction rules, deleting for many subjects the most volatile part of the inflation and prediction series). An interesting observation regarding the composition of prediction rules is furthermore that participants in the first treatment on

---

[19] Special circumstances regarding coefficient estimation are the following. Subjects 3 and 7 of Tr. 1 clearly settled to specific prediction rules only a few periods after the learning phase of 20 periods. OLS regression was therefore started on $t = 29$ for subject 3 and on $t = 28$ for subject 7. Furthermore, subjects 3, 11 and 16 of Tr. 1 have a constant prediction series in the estimation sample, so their coefficients are undetermined and the R^2 and Ljung-Box autocorrelation statistics non-existent; constant rules were chosen for these subjects. Subject 11 of Tr. 2 similarly had a near-constant prediction series making it impossible to computationally determine the coefficients maximizing the fit; again a constant rules was chosen.

Subj.	C	π^e_{-1}	π^e_{-2}	π^e_{-3}	π^e_{-4}	π_{-1}	π_{-2}	π_{-3}	R^2	AC	Break
1	0.056*	0.587	0	0	0	0.111	0	0.077	0.643	No	No
2	0.112*	0	0	0.487	0	0.191	0	0	0.491	No	No
3‡	0.000	0	0	0	0	0	0	0	—	—	Yes
4	80.87	0	-0.649	0	0	-0.530	0	0	0.207	No	No
5	1.470	0	-0.629	0	0	-0.370	0	0	0.430	No	No
6	4.471	0	0	0	-5.956	0.502	0	0	0.683	No	No
7†,‡	0.083*	0	0	0	0	0	0	0	0.000	No	Yes
8	1.005	0.342	0	0	0	-0.181	0	0	0.527	No	No
9	0.755	0	0	0	0	0.086	0.112	0	0.524	No	No
10	3.502*	0.900	-0.266	0	-0.169	-0.483	-0.152	-0.378	0.889	No	Yes
11‡	1.110	0	0	0	0	0	0	0	—	—	No**
12	0.481	0	0	0	0	0	0	0	0.000	No	No
13	17.26	0	0	0	0	0	0	0	0.000	No	No
14	0.006*	0	0	0	0	0.053	0	0	0.146	No	No
15	0.447	0	0	0	0	0	0	0	0.000	No	No
16‡	0.000	0	0	0	0	0	0	0	—	—	No**
1	11.19	0	0	0	0	0	0	0	0.000	Yes	Yes
2	9.741*	0	0	0	0	0	0	0	0.000	No	No
3	9.630*	0	0	0	0	0	0	0	0.000	No	No
4	1.447	0	0.742	0	0	0	0	0	0.754	No	No
5	8.014	0	0	-0.683	0	0	-0.158	0	0.262	No	No
6	10.04	0	0	0	0	0	0	0	0.000	No	No
7	15.30	0	0	0	0	0	0	0	0.000	Yes	No
8	23.63	0	0	0	0	0	0	0	0.000	No	No
9	7.123	0	0.428	0	0	0	0	0.520	0.353	No	No
10	5.386	0	0	0.296	0.167	0	0	0	0.433	No	No
11‡	10.00	0	0	0	0	0	0	0	0.000	No	No
12	0.574	0.429	0	0	0	0	0	0.258	0.411	No	No
13	90.15	-0.564	-0.450	-0.563	0	0	0	0	0.421	No	Yes
14	30.11	0	0	0	0	0.481	0	0	0.222	No	No
15	15.70	0	0	0	0	0	0	0	0.000	No	Yes
16	37.80	0	0	0	0	0	0	0	0.000	Yes	No

Table 2: Estimated GA(4,3) prediction rules (OLS) for subjects of both treatments. The 1st column shows subjects' numbers, clustered according to treatment. The 2nd through 9th columns show the estimated GA coefficients, starting with the constant. The 10th and 11th column contain the R^2 statistic and autocorrelation report up to 8th order (Q statistics, 5%). The last column indicates the presence of a structural break in the GA rule at period 25 (Chow Breakpoint Tests, 5% level). Insignificant explanatory variables have been eliminated successively, largest p value first (5% level). An asterisk denotes an insignificant constant; a dagger elimination of apparently significant variables because of autocorrelation; a double dagger special circumstances due to outliers etc. (see footnote 19); a double asterisk non-standard determination of the structural break due to multiplicity in the prediction rule parameters (see footnote 29).

	(a)					
Obj. \ Subj.	*0*	*1*	*2*	*3*	*4*	*Total*
0	7	0	0	0	0	**7**
1	1	5	0	0	0	**6**
2	1	1	0	0	0	**2**
3	0	0	0	1	0	**1**
Total	**9**	**6**	**0**	**1**	**0**	**16**

Averages: 0.81 \ 0.56 — 1.38 total

	(b)					
Obj. \ Subj.	*0*	*1*	*2*	*3*	*4*	*Total*
0	6	1	0	1	0	**8**
1	1	3	1	0	0	**5**
2	0	0	0	0	0	**0**
3	0	0	0	0	0	**0**
Total	**7**	**4**	**1**	**1**	**0**	**13**

Averages: 0.38 \ 0.69 — 1.08 total

Table 3: Complexities of GA prediction rules, for both treatments separately (cf. Table 2). *Complexity* is defined as the number of significant terms, divided in *objective terms* (lagged realized inflations) and *subjective terms* (lagged own inflation predictions). The average complexities per treatment are listed at the bottom of the table.

Part.	*Orig.*	*p val.*	α_1	α_2	$1-\alpha_1-\alpha_2$	β	*Classification*
1	1	0.231	0.111	0.801	0.088	0.023	Adaptive, Weakly Obstinate
2	2	0.108	0.082	-0.102	1.020	0.088	None
3	4	0.211	-0.775	0.202	1.573	0.393	Fundamentalist
4	5	0.085	-0.272	-0.076	1.348	0.073	None
5	6	0.211	1.988	-1.564	0.576	-0.777	None
6	9	0.303	0.184	0.142	0.674	-0.095	Weakly Fundamentalist, Tr. Reversing
7†	11	—	0.000	1.000	0.000	0.000	Obstinate
8	12	0.513	-0.167	0.450	0.717	0.046	Fundamentalist
9	13	0.438	0.220	0.102	0.678	-0.074	Fundamentalist
10	14	0.096	-0.005	0.932	0.073	0.063	Obstinate
11	15	0.099	-0.116	0.138	0.978	0.085	Fundamentalist
12†	16	—	0.000	1.000	0.000	0.000	Obstinate
13	2	0.707	-0.149	0.178	0.971	0.023	Fundamentalist
14	3	0.111	-0.416	0.165	1.251	0.162	Fundamentalist
15	4	0.106	-0.130	0.986	0.144	0.048	Weakly Obstinate
16	6	0.064	-0.020	0.877	0.143	0.000	Obstinate
17	8	0.694	0.265	0.034	0.701	-0.012	Fundamentalist
18	10	0.212	-0.049	0.593	0.465	0.033	Trend Following
19†	11	1.000	0.000	1.000	0.000	0.000	Obstinate
20	12	0.053	-0.398	0.553	0.845	0.217	Weakly Obstinate, Tr. Following
21	13	0.078	0.409	-0.502	1.093	-0.134	Fundamentalist
22	14	0.216	0.388	0.166	0.446	0.028	Adaptive
23	15	0.494	0.041	0.040	0.919	-0.016	Fundamentalist

Table 4: Estimated coefficients of subjects whose predictions are well-described by an FOH rule (17). Estimates are clustered by treatment. *p* values in the 3rd column give the likelihood of the FOH description (Wald tests, 5% level). The classification in the last column uses mostly well-known rules, see footnotes 24, 25 for definitions (categorization based on series of Wald tests, 5% level).

average relied more on lagged inflation rates (i.e. objective variables) than they did on their own past predictions (subjective variables), while in the second treatment the reverse is true. Part of the explanation of this difference is the inflation rate's much lower degree of stability in the second treatment compared to the first, making it more difficult for subjects to use lagged inflations as determinants of future realizations. The same holds for the lagged predictions, but less strongly because their volatility tends to be much smaller than that of the associated inflation rates. A shift from objective to subjective variables in the second treatment is therefore generally sensible given the accuracy-based reward of the experiment.

Moving back to Table 2 the above statement on the relative amounts of objective and subjective variables in each of the treatments can be verified at a glance. The difference between the treatments consists mainly of the frequent use of the most recent lagged inflation by the subjects of the first treatment, as opposed to its virtual absence in the second treatment. Though π_{t-1} is often used there is no obvious pattern in the estimated values of o_1, except for the fact that they all have an absolute value smaller than 0.6. A pattern does exist in the estimated coefficients s_1 through s_4. Aside from having with one exception absolute values smaller than one, the coefficients alternate in sign starting with positive signs for the coefficients with the first lag of predictions. This alternation could be the result of some participants choosing a prediction series with negative serial correlation (e.g. because they perceive an alternating pattern in realized inflation rates), so that a prediction depending positively on the previous prediction would also depend negatively on the second prediction lag, etc. It is not easy to find proof for this conjecture in Figure 2, though in particular subjects 2 and 4 do seem to provide examples of alternating prediction series.[20]

The coefficients of the second treatment are too few and too irregular to easily derive interpretations of the estimated prediction rules from them. Properties that do stand out though are first that they are all absolutely smaller than one, like all but one of the coefficients in the first treatment, and second that none of the subjects have coefficients of differing signs entering in their prediction rules, as opposed to several subjects in the first treatment. The lack of evident meaning in the coefficient estimates is compensated by the often sensible regression constant. As was pointed out above, nearly half of all estimated rules are constant

---

[20] Subject 4's predictions often lie beyond the range of Figure 2(d), but they fluctuate with fairly constant amplitude and period. Both series have negative first-order autocorrelation statistics on the estimation sample, but neither is close to being significant. Statistics are -0.012 for subject 2 ($p > 0.9$, based on Ljung-Box Q statistic) and -0.212 for subject 4 ($p > 0.2$, idem).

in both treatments, and many of them describe focal or near-focal predictions that mostly are close to the equilibrium infla-tion rate (examples are subjects 11 and 15 of Tr. 1 and subjects 6 and 11 of Tr. 2; exceptions with high expectations volatility and a constant rule also exist though, e.g. subject 3 of Tr. 2). This leads to the question whether the general prediction rule (16) can be simplified in such a way that most of the estimated expectations structure is retained but the meaning of the non-constant rules is simplified, allowing the estimated rules to be more easily interpreted.

In the next subsection the First-Order Heuristics (FOH) rule specification will be applied to those prediction series that are successfully described (i.e. without creating residual autocorrelation) by a GA(4,3) rule, and not significantly worse described when the necessary restrictions are put on the coefficients of the GA specification. Several characteristics of the estimates in Table 2 suggest that the FOH rule might do well.

First, the frequent usage of the most recently available inflation rate in Treatment 1 (LTh) fits well within the FOH framework, since the first-order heuristic assigns a central role to the first lag of realized inflation. Whether the near-absence of this variable in the estimated rules of the second treatment makes the FOH rule less suitable in that case remains to be seen. Second, the closeness of a substantial part of the regression constants in both treatments to the natural rate of inflation indicates that this rate, or subjects' estimation of it based on realized inflation rates and previous predictions, might play an independent role in expectation formation. This is consistent with the use of an equilibrium inflation estimate as a separate term in the FOH specification (see next subsection). And third, the fact that all but one of the estimated coefficients in Table 2 lie in absolute value within the unit interval suggests that sub-jects use weighted averages between the available variables, possibly including a con-stant, in their prediction rules. The FOH rule specifically assumes that they average over the previous prediction, the most recent inflation rate and a constant equilibrium estimate (which will be defined below); whether this specification is general enough depends on the correlations between the variables in the estimated GA rule (16). Finally, the Chow test on structural breaks in the individual prediction rules, reported in the last column of Table 2, is the subject of Section 6.4.

6.3   Estimation and Interpretation of First-Order Heuristics

The Generalized Adaptive (GA) prediction rule (16) has been defined to detect as much regularities as possible in subjects' predictions within a linear framework, taking into account the limited amount of available observations. While the GA rule specification has been

successfully applied to subjects' prediction series of both treatments, the results in many cases lack a clear behavioral interpretation due to the broadness of the specification. This makes it hard in particular to see whether the estimated rules conform to well-known elementary rules (e.g. naive, adaptive, trend-following and trend-reversing expectations). This problem can be solved by using a restricted form of the GA specification that has just several coefficients with clear behavioral interpretations and at the same time is not a significantly worse description of actual predic-tions. The First-Order Heuristics (FOH) rule provides such a restricted form; in the context of this paper it is described for an individual i as follows:

$$\pi_{i,t}^{e} = \alpha_1 \pi_{t-2} + \alpha_2 \pi_{i,t-1}^{e} + \left(1 - \alpha_1 - \alpha_2\right)\frac{1}{30}\sum_{t=21}^{50}\pi_t + \beta\left(\pi_{t-2} - \pi_{t-3}\right) + \varepsilon_t, \tag{17}$$

with the parameters α_1, α_2 and β replacing the eight parameters (including the constant) of the GA(4,3) specification estimated in Section 6.2. Note that the third term in (17) is constant, because it is an average of realized inflation rates that does not depend on t.

The FOH rule gets its name from the decomposition of its non-stochastic terms in a zeroth-order inflation estimate (the first three terms of (17) with α coefficients, counting the sum over π_t as a single term) and a first-order estimate (the fourth term of (17) with coefficient β). The zeroth-order estimate is a weighted average of the most recent prediction, the most recent realized inflation and the post-learning phase sample mean; this average can be interpreted as a "reference point" constructed in each period to serve as a basis in the formation of the prediction.[21] The first-order addition to the zeroth-order estimate simply consists of the slope between the two most recently realized inflation rates $\pi_{t-2} - \pi_{t-3}$, scaled with a factor β. This difference can be interpreted as a short-run trend estimate of inflation

---

[21] The estimation sample mean of inflations $\frac{1}{30}\sum_{t=21}^{50}\pi_t$ per definition is not available during the experiment, but it is still included here for three reasons. First, the mean of all previously realized inflation rates can be calculated by subjects and it generally converges fast to the sample mean, making the latter a suitable proxy for the former. Second, if the FOH rule (17) is understood as resulting from a set of parameter restrictions on the GA(4,3) rule (16), then the term representing an estimate of the inflation equilibrium in (17) must be either a constant or an average of the last 3 realized inflations; the latter is generally not an accurate estimate because of the small number of terms. Third, including the true natural rate of inflation in (17) instead of the sample average substantially reduces the descriptive value of the FOH rule, since many subjects evidently did not learn the true value and therefore could not have incorporated it in their prediction rules.

development. Either a trend-following or trend-reversing compensation is therefore added to the reference point, depending on the sign of β.

Before applying the FOH specification (17) to the individual prediction series, Wald tests on joint parameter restrictions are done to check that the explanatory value of the estimated rules is not significantly reduced when simplifying the GA(4,3) into the FOH specification.[22] Table 4 gives the p values for the subjects that passed this test, i.e. for whom the null hypothesis of no reduction in explanatory value was not rejected. 12 out of 16 and 11 out of 16 subjects in Treatments 1 (LTh) and 2 (HTh) respectively are be equally well described by a FOH rule as by a more general GA(4,3) rule.[23] The prediction rules of these subjects were re-estimated based on samples identical to the ones in the GA(4,3) estimation (starting at $t = 25$; see footnote 19 for exceptions) and using the FOH specification in a deviation-from-average form:

$$\pi_{i,t}^e - \sum_{t=21}^{50}\frac{\pi_t}{30} = (\alpha_1 + \beta)\left(\pi_{t-2} - \sum_{t=21}^{50}\frac{\pi_t}{30}\right) + \alpha_2\left(\pi_{i,t-1}^e - \sum_{t=21}^{50}\frac{\pi_t}{30}\right)$$
$$-\beta\left(\pi_{t-3} - \sum_{t=21}^{50}\frac{\pi_t}{30}\right) + \varepsilon_t. \tag{18}$$

The estimated coefficients as well as the implied weight $1 - \alpha_1 - \alpha_2$ of the inflation ave-rage are listed in the table. The coefficients are also plotted as vectors $(\alpha_1, \alpha_2, \beta)$ in the so-called Prism of First-Order Heuristics shown in Figure 9. The reason that the parameter space of the FOH specification (17) has a prismatic shape is that the coefficients of its first three terms

---

[22] The parameter restrictions tested simultaneously are, using the notation of specification (3.16):

$$c + \frac{1}{30}\sum_{t=21}^{50}\pi_t\,(s_1 + o_1 + o_2 - 1) = 0, \ \ s_2 = 0, \ \ s_3 = 0, \ \ s_4 = 0, \ \ o_3 = 0.$$

The first restriction follows after multiplying and dividing the constant in (16) by the inflation sample mean and requiring that this mean and the most recent prediction and realized inflation form a weighted average, while allowing for a slope correction. The resulting Wald statistic has an F distribution with 5 and 19 df. The tests used the original GA estimation results, i.e. before the removal of insignificant variables; subjects without a valid GA rule due to autocorrelation were excluded from FOH testing.

[23] The daggers in Table 4 with the 7th and 12th estimated FOH rules (corresponding to subjects 11 and 16 in Tr. 1) indicate that F tests could not be done due to perfect multicollinearity in the regressors caused by constant prediction series (cf. footnote 19). In these cases α_2 was set to 1 and both other parameters to 0, resulting in a perfect fit of the predictions.
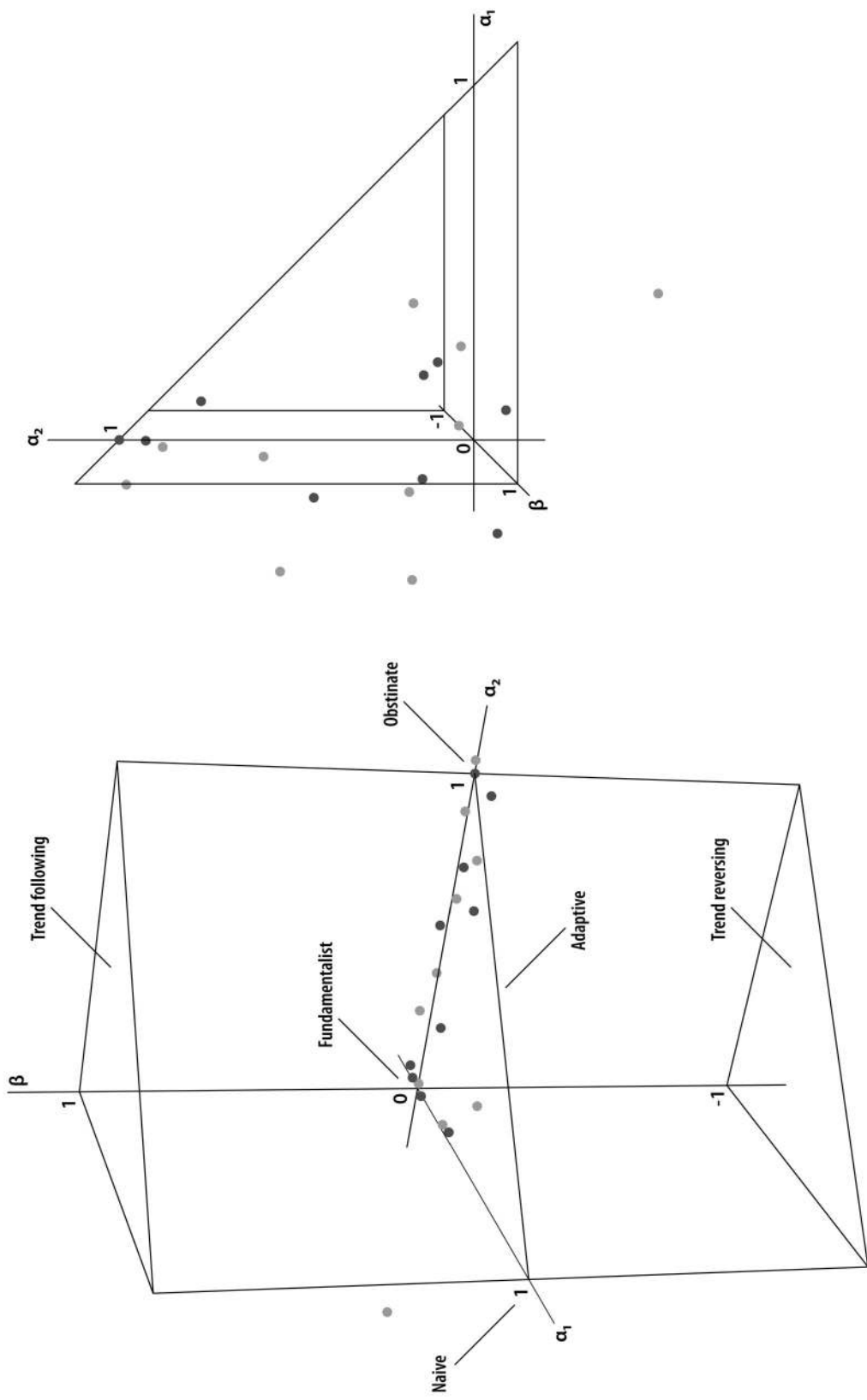
Figure 9: Prism of First-Order Heuristics containing parameter vectors of estimated FOH predicton rules (cf. Table 4). The smaller graph is a top-down view of the Prism. Black and gray dots represent subjects in Tr. 1 (LTh) and Tr. 2 (HTh) resp. Several dots fall beyond the graphs' ranges. The labels in the prism point to parameter subsets corresponding to well-known prediction rules (see footnote 24).

(the zeroth-order part of the prediction) form a weighted average. Given non-negative weights, the subvector (α_1, α_2) lies in the two-dimensional simplex, which becomes a prism by adding the slope parameter β on the vertical axis.

The parameter vectors shown in Figure 9 facilitate the interpretation of the estimated coefficients in Table 4. While there is a considerable spread in the sets of estimated vectors of both treatments, the prism makes clear that the β and, to a lesser extent, α_1 components are relatively unimportant in the prediction rules. In so far as any pattern can be observed in the parameter vectors, it is that they are roughly clustered around the α_2 axis of the simplex at $\beta = 0$, ranging from the origin (corresponding to Fundamentalist Expectations, i.e. expectations always equal to the sample mean of inflations) to the vertex at $(0, 1, 0)$ (Obstinate Expectations, i.e. expectations always equal to the previous prediction).[24] The fact that parameter vectors of both treatments are often located in this area is consistent with the observations in Section 4 that many subjects commit to focal or near-focal prediction rules somewhere during the experiment. Focal rules are special cases of obstinate expectations, and simultaneously of fundamentalist expectations if the focal inflation rate is sufficiently close to the sample mean of realized inflation rates.

On the other hand, naive expectations or adaptive expectations with a high adjustment parameter lead to violent instability in the inflation development (cf. the simulations of Section 6.1), so it is logical that they are mostly dropped well before the end of the experiment. The low absolute value of the slope parameter β is similarly unsurprising, since neither trend-following nor trend-reversing expectations are any help in stabilizing the inflation rate in the experimental model. Finally, no within-treatment clustering is visible among the parameter vectors, though those belonging to the second treatment are more

---

[24] The parameter regions of the other labels in Figure 9 and the prediction rule definitions they follow from are as follows. $(1, 0, 0)$ equals Naive Expectations, meaning that predictions are equal to the most recent realized inflation rate; the line between $(0, 1, 0)$ and $(1, 0, 0)$ represents Adaptive Expectations, with predictions equal to a weighted average of the most recent prediction and realized inflation; the simplex at $\beta = \pm 1$ equals Trend-Following or Trend-Reversing Expectations, meaning any combination of the above prediction rules plus the positive or negative slope between the last two realized inflations. Note that Fundamentalist Expectations does not mean here that predictions are always equal to the equilibrium inflation rate. The estimation sample mean approximates the equilibrium rate in many cases, but is sufficiently different from it in general to improve the fit of the FOH rule by using the sample mean instead of the equilibrium rate (cf. footnote 21).

spread out than those of the first, reflecting the greater difficulty subjects from Treatment 2 (HTh) had in finding a rule suitable for taming inflation.

To arrive at an accurate interpretation of the estimated FOH rules, it is tested whether they are statistically equal to any of the simple prediction rules that are special cases of the FOH specification (17), indicated as parameter subsets in the prism in Figure 9. Each estimated rule is subjected to a series of Wald tests with coefficient restrictions derived from the six simple rules indicated in Figure 9.[25] The rightmost column of Table 4 shows the test results in the form of the basic prediction rules that the estimated FOH rules are equivalent to. This classification of rules confirms the observations made above, namely that fundamentalist and obstinate expectations are the most frequently occurring and that there is no clear clustering of rules within the treatments. In both treatments rules are classified most frequently as fundamentalist expectations (also counting its "weak" form, see for definition footnote 25), directly followed by obstinate expectations. Furthermore, there is remarkable similarity between the classifications of the two treatments (again ignoring weakness of rules), with fundamentalism occurring five times, obstinacy four times and adaptive expectations once in both treatments. The differences between the treatments on the other hand are limited to the first treatment having three subjects with FOH parameters significantly different from any of the defined subrules, while the second treatment has none, and the first treatment having a single trend-reversing rule component, while the second has a single trend-following one.

6.4   Structural Changes in Prediction Rules during the Experiment

When overviewing the experimental results in Section 4, the inflation dynamics of a substantial number of participants (5 in Tr. 1, 6 in Tr. 2) were categorized as "from unstable to stable." This means that somewhere in the course of the experiment these participants changed such that the inflation development settled down to fluctuations around the equilibrium rate with a relatively low amplitude. Assuming that the classification of the

---

[25] The restrictions in terms of the parameters α_1, α_2 and β are as follows. Naivety: $\alpha_1 = 1$, $\alpha_2 = 0$, $\beta = 0$; Obstinacy: $\alpha_1 = 0$, $\alpha_2 = 1$, $\beta = 0$; Fundamentalism: $\alpha_1 = 0$, $\alpha_2 = 0$, $\beta = 0$; Adaptive expectations: $\alpha_1 + \alpha_2 = 1$, $\beta = 0$. A rule is qualified as "weak" if the restrictions are rejected as listed above, but not rejected if $\beta = 0$ is dropped (i.e. if the rule is allowed a trend component). If a rule is classified as obstinate or naive then it is technically also adaptive, which is left implicit in Table 4. Trend-following or trend-reversing expectations are tested through the respective one-sided t tests of the hypothesis $\beta = 0$. All tests have been done at 5% significance level.

inflation dynamics is correct and that subjects' expectations are sufficiently consistent to be described by prediction rules, this implies that a structural change in the applied rules must have occurred at least once during the experiment. However, as was remarked in Section 4 and confirmed in the simulations of Section 6.1, the difference between a stable and an unstable inflation development may lie in a small change of a single prediction rule parameter. Examples are a shift in the adjustment parameter in the case of adaptive expectations or the addition or removal of an inflation lag in the case of averaging expectations. It is therefore uncertain whether the structural breaks in the prediction series of participants in the "from unstable to stable" category are large enough to be confirmed by statistical testing. Conversely, the remaining subjects in the "stable" or "unstable" categories, especially those in the second category, may have changed their prediction rules during the experiment without this change affecting the stability of the inflation series.

Before testing directly whether structural breaks have occurred in the prediction rules of participants during the experiment, possible evidence for them is studied through the development of the participants' earnings. Figure 10 shows average earnings of participants in terms of experimental points per period (see the linear scoring rule (13)), with the experiment divided in two equal halves of 25 periods. When a structural change in a subject's prediction rule changes the inflation dynamics from unstable to stable, the subject will, largely independent of the exact period in which the change is implemented, earn substantially more in the second half of the experiment than in the first. Such a subject should show up in Figure 10 in the left upper part, well above the diagonal. In those cases in which the inflation dynamics with a subject are characterized by uniform instability or stability, earnings should be low or high respectively in both halves of the experiment, translating into positions around the far ends of the diagonal in Figure 10.

All of the above cases are indeed represented in Figure 10. Close to the diago-nal at the right upper corner lies a cluster of black dots associated with subjects from Treatment 1 (LTh), some of which approach the average earnings under the rational expectations equilibrium in the second half of the experiment (as indicated by the black horizontal line). The uniformly high earnings of these individuals suggest that their inflation dynamics fall in the "stable" category, which is correct for the whole cluster.[26] At the other side of the diagonal lies a more spread out cluster, made up almost entirely of subjects from the second treatment.

---

[26] The cluster consists of 8 dots in total, corresponding to subjects 2, 8, 9, 11, 12, 14, 15 and 16 of Tr. 1. This set of subjects is identical to the "stable" category of this treatment (see Section 4.1).
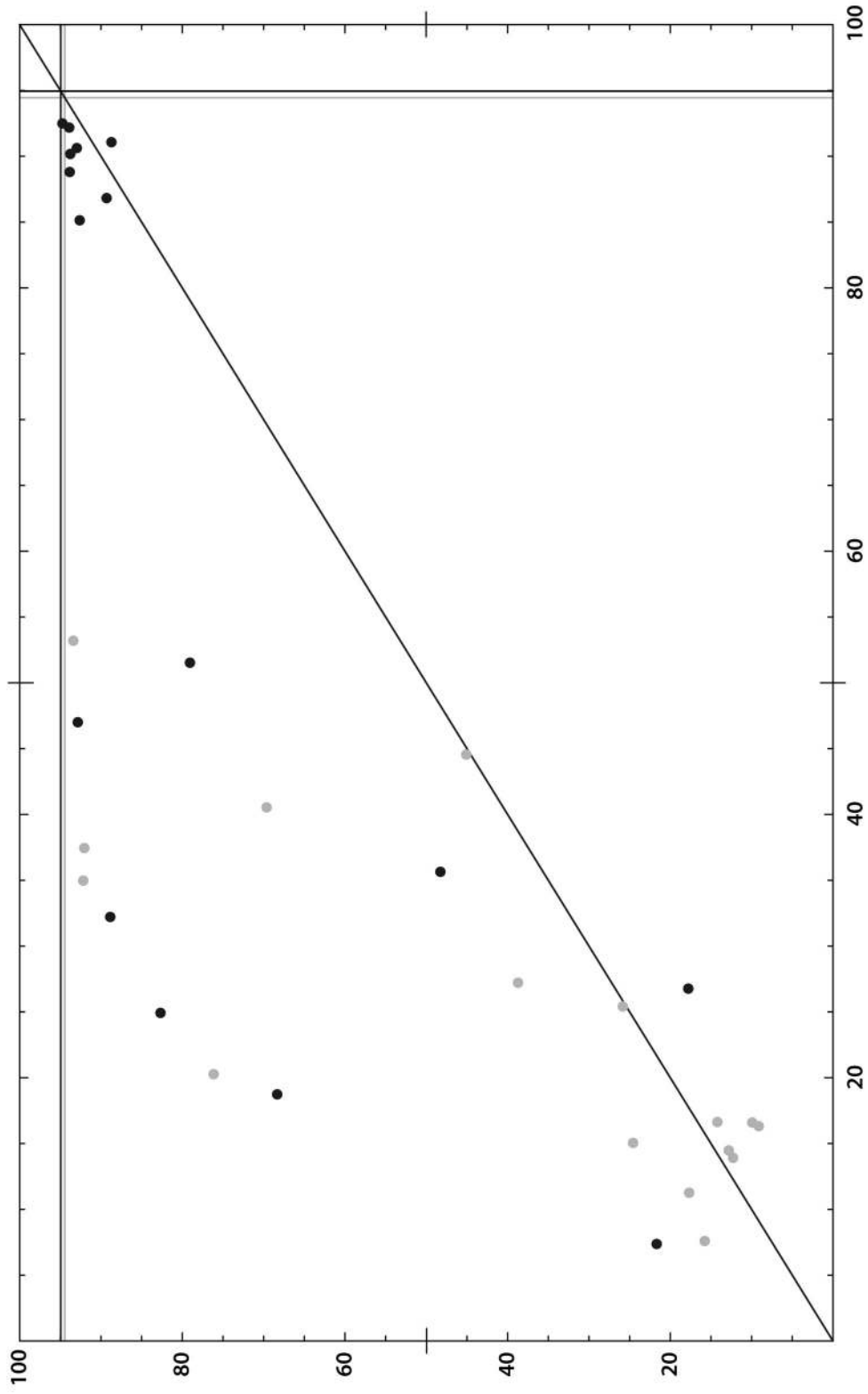
Figure 10: Average earnings of subjects in the first half (horizontal axis) and second half (vertical axis) of the experiment. Earnings are expressed in mean Prediction Score, i.e. do not have a money value (see Appendix). Black dots belong to subjects from Tr. 1 (LTh); gray dots to participants from Tr. 2 (HTh). The pair of crossing lines in the upper right of the figure indicates the mean Prediction Score under the REH; black lines for the first treatment and gray lines for the second.

The densest part of this cluster, in the rectangle $[0,20] \times [0,30]$, consists of nine subjects who all have inflation dynamics classified as "unstable."[27] As expected, the subjects with the poorest earnings in both halves of the experiment are those who were unable to stabilize inflation development. The composition of the clusters reflects the facts that in the first treatment few participants were unable to stabilize inflation development within the duration of the experiment, while in the second none succeeded in achieving stability in the beginning and maintaining it to the end.

A third group of dots is located well above the diagonal and towards the left of Figure 10. These dots represent participants whose earnings were poor in the first half of the experiment, but much improved in the second half. These beneficial changes strongly suggest that the participants succeeded in changing their prediction rules during the experiment in such a way that it ended initial phases of instability. Defining the group of dots as those falling short of 50% of the hypothetical maximum reward in the first half of the experiment and exceeding 50% in the second half, two sets of four dots from both treatments remain. All of these dots represent participants that in Section 4 were classified in the "from unstable to stable" category.[28] This is consistent with the hypothesis that a substantial number of participants from both treatments changed their prediction rules during the experiment and as a result strongly improved inflationary stability and their own predictive accuracy.

To verify structural breaks in prediction rules directly, the Chow Breakpoint Test is applied to the prediction series of all participants. To include all participants and test as broadly as possible for structural changes within a linear framework, the GA(4,3) specification (16) estimated in Section 6.2 is used to model the prediction rules. The main objective is to verify the existence of structural breaks linking phases of unstable inflation development to phases of stable inflation development. Moreover, the number of observations for each subject is limited relative to the number of coefficients to be estimated. Chow's test is therefore set to

---

[27] Out of the 9 subjects lying in the rectangle $[0,20] \times [0,30]$, 1 is from Tr. 1 (subject 4) and 8 are from Tr. 2 (subjects 3, 7, 8, 9, 12, 13, 14 and 16). The inflation series of all these subjects were classified as "unstable" in Sections 4.1 and 4.2; the only subject with an "unstable" inflation series not in the rectangle $[0,20] \times [0,30]$ is subject 13 of Tr. 1 (this subject falls just to the right of the rectangle, with earnings coordinates (26.8, 17.8)).

[28] The 4 black dots correspond to subjects 1, 3, 5 and 7 from Tr. 1; the 4 gray dots correspond to subjects 4, 5, 6 and 10 of Tr. 2. Both sets of subjects are subsets of the "from unstable to stable" categories (see Sections 4.1 and 4.2 resp.)

detect a single breakpoint in each prediction series, i.e. each estimation sample is split in two subsamples. The size of the subsamples is held constant for all participants, for the reason that the most likely period in general for a structural prediction change to take place is at the end of the learning phase, which was given a fixed length in Section 5.1. Determining the length of the second subsample in such a way that the lags of the GA(4,3) specification only use information from after the learning phase (which ends at $t = 20$) results in a split of the full prediction sample in parts $\{\pi^e_{i,1}, \ldots, \pi^e_{i,24}\}$ and $\{\pi^e_{i,25}, \ldots, \pi^e_{i,51}\}$ for any subject i. Note that the second subsample is identical to the estimation sample used in Section 6.2 to arrive at the GA(4,3) estimates in Table 2.

The last column of Table 2 lists the results from the Chow Breakpoint Test at 5% significance level. The test estimates the GA rule for both subsamples and then compares the total sum of squared residuals with the sum resulting from estimating the GA rule on the full sample. This means that it remains undetermined for subjects 11 and 16 of Treatment 1 (LTh), for whom the GA coefficients could not be estimated in all three cases. However, both subjects have a focal prediction rule for nearly the whole experiment (see Graphs (k) and (p) of Figure 2), so they are listed in Table 2 as having no structural break.[29] Looking at the whole column of test results, two remarkable things are the lower than expected number of rejections and the equal amount of rejections between the treatments. The null hypothesis of structural consis-tency is rejected three times in both treatments, which is strictly less than the size of the group of two times four subjects identified in Figure 10 as displaying signs of a structural break. It is also less than the five and six subjects respectively making up the "from unstable to stable" category in the first and second treatment, who were classified as such because of an apparent structural change stabilizing their initially highly volatile inflation development. Moreover, the prediction series in which a break was detected do not all belong to the "from unstable to stable" category (subject 13 of Tr. 2 was classified as "unstable" and subjects 10 of Tr. 1 and 15 of Tr. 2 could not be classified), nor do they all fall in the left upper part of Figure 10 which was identified above as the most likely region for subjects with structurally unstable prediction series (only subjects 3 and 7 are in this region).

---

[29] Subject 11 has a perfectly focal prediction rule in the second subsample, creating multiplicity in the prediction lags and the constant of the GA(4,3) specification (cf. footnote 19). If a constant rule is assumed to hold though, the Chow statistic can be calculated and confirms non-rejection of the absence of a structural break. Subject 16 follows a focal rule from $t = 3$, making it impossible to determine the GA coefficients in both subsamples and the full sample.

Apparently many of the changes participants made to their prediction rules during the experiment were either, given the limited amount of available observations, too small to be detected through standard statistical testing, or did not occur sufficiently close to $t = 25$ to be picked up by the Chow statistic. An example of a participant that would have demonstrated a structural break if subsamples would have been tailored to separate prediction series is participant 11 of Treatment 2 (HTh) (see Figure 3(k)). Setting the end of the first subsample at $t = 20$ instead of 24, to fully capture the near-focal rule at $\pi = 10$ in the second subsample, gives a Chow Breakpoint statistic (F distribution with 8 and 31 df.) of 2.717, p value 0.022.[30] It should finally be noted that even though the number of structural breaks indicated by Chow's Break-point Test is low in both treatments, it is still high enough to reject the hypotheses that structural changes do not happen in the experimental treatments as a whole. Evaluating these two hypotheses at 5% significance level through the number of rejections following the Chow tests, they are rejected in one-sided tests by the smallest possible margin. A maximum of two apparent structural breaks per treatment is allowed in the case of structural consistency (under the null hypothesis, the number of rejections of the Chow tests is binomially distributed with 16 trials and rejection rate 5%).

7.   Concluding Remarks

In this paper the way in which individuals who are not professional analysts tend to predict the long-run development of inflation is studied. The overlapping generations environment with money creation gives participants the task of supplying a hypothetical population with future inflation estimates. These estimates determine the population's saving behavior and thus supply and demand on the goods market, resulting in market-clearing prices that translate into actual inflation rates. This paper has two main topics: first, the ability of subjects to control the translation of predictions into inflation rates, in the sense of submitting successive predictions that turn out to be accurate when the actual inflation rates materialize; second, the prediction rules through which subjects attempt to achieve this control. The ability to successfully predict and thereby control inflation has been shown to vary widely, both between subjects of the same treatment and between the treatments as a whole.

---

[30] Redefining subsamples does not always work though, such as in the case of subject 6 of the second treatment (see Figure 3(f)). Ending the first subsample at $t = 22$ to exactly restrict the second subsample to the near-focal rule, again at $\pi = 10$, gives a Chow statistic (F distribution with 8 and 31 df.) of 1.348, p value 0.258.

Regarding the difference between the treatments, it is clear that the inflation elasticity of the savings rate around the equilibrium rate of inflation (cf. Figure 1(b)) is related inversely to the number of subjects succeeding in stabilizing inflation development within the 50 periods of the experiment. Concerning the formation of inflation expectations, it has been shown that subjects' strategies to maximize their forecasting accuracy, estimated as linear rules dependent on previously realized inflations and predictions, are essentially few in number and remarkably similar between treatments.

The sensitivity of the goods market in the OLG economy to expected changes in the inflation rate, which is deliberately higher in the second treatment (with the high equilibrium inflation rate), is such that small changes in prediction rules can lead to major changes in the dynamics of the inflation rate. This property of the experimental model is clearly reflected in the stability characteristics of the inflation series produced by participants. These differ strongly between participants but often also change dramatically during the experiment, typically moving from highly volatile instability to small fluctuations around the equilibrium rate of inflation. In the first treatment, with the low equilibrium rate making it relatively easy for subjects to prevent inflation from wildly fluctuating, almost all subjects (14 out of 16) managed to stabilize inflation to some extent, either at the start of the experiment or somewhere in the course of it. In the second treatment less subjects succeeded in doing so (10 out of 16) and none succeeded in keeping the inflation rate stable during the entire experiment.

The Rational Expectations Hypothesis, prescribing predictions always equal to the equilibrium inflation rate, on the whole poorly explains the experimental results. It does much better though in the first treatment than in the second and in both treatments the REH describes the inflation means much better than the inflation volatility. Also, many subjects stabilize inflation long after the beginning of the experiment. It is therefore reasonable to suppose that others, who do not achieve stabilization at all, would manage to do so in an experiment that lasts longer than 50 periods. In an economic environment analogous to the one in this paper that exists for a longer time than the duration of the experiment, the REH might be a reasonable explanation of the average inflation rate. To what extent excess volatility decreases under an extended experimental duration remains uncertain however, since Table 1 demonstrates that excess volatility may remain even after apparent stabilization of the inflation development.

Fitting linear prediction rules composed of previous predictions and inflation rates to the prediction series of individual participants has proven successful, in the sense that nearly

all participants' predictions can be represented by a linear rule free of residual autocorrelation. The estimated rules do often consist of few terms and in many cases are simply constant. The generally low complexity of the rules has made it possible to reduce the Generalized Adaptive (GA) specification in a large majority of cases to the First-Order Heuristics (FOH) specification, a composition of simple and well-known linear prediction rules. A re-estimation and classification of rules based on the FOH specification has clearly established that in both treatments the two expectation rules most used are fundamentalist and obstinate expectations, i.e. predictions given by the sample mean of realized inflations and predictions equal to their first lag respectively. The fact that the classification of prediction rules following from the FOH estimation is very similar in both treatments shows that subjects in the second treatment are, as much as those in the first, able to learn the simple prediction rules that stabilize the inflation dynamics. However, they apparently do not apply them strictly enough in general to achieve the same degree of inflation rate stabilization as participants in Treatment 1 (LTh).

A future experiment expanding the one discussed in this paper could use groups instead of separate individuals to form inflation predictions in an overlapping generations economy, e.g. by taking in each period the group mean of predictions and treating it in the same way as an individual inflation prediction in this paper's experiment. Whether multi-agent treatments on the whole produce more stable inflation series is a priori unclear. Averaging over a group of inflation expectations will on the one hand have a smoothing effect on the aggregate prediction series, by limiting the effect of individual outliers and cancelling out fluctuations in individual predictions, but on the other hand will make it less likely for the aggregate prediction series to remain approximately constant. It will depend on the elasticity of the savings rate at the chosen equilibrium rate of inflation whether a series of average predictions has a sufficiently low volatility to keep the realized inflation rate close to equilibrium. A second research question would be whether functioning in a group leads participants to use different prediction rules than in a single-agent environment. A preliminary hypothesis regarding that question is that averaging over individual expectations reduces the necessity for participants to accurately learn the rules that by themselves lead to convergence to the equilibrium inflation rate. Instead they may rely, because of the averaging over expectations within groups, on the equally imperfect predictions of other group members to compensate for their lack of insight into the inflation-generating law of motion.

Finally, the analysis of structural breaks deserves to be extended in future research. In this paper, Chow's Breakpoint Test has been applied to two prediction subsamples divided by

a fixed period for all participants of both treatments. The resulting number of rejections of the null hypothesis of coefficient stability is, though significant in each treatment, smaller than expected based on either the number of inflation series classified as "first unstable then stable" (see Sections 4.1 and 4.2), or the earnings development of subjects (see Figure 10). The reason that the subsamples were kept constant across subjects in Section 6.4 is that to make them variable requires an objective criterium, in terms of changes in predictions or realized inflations, to identify periods in which a structural break might occur. An alternative would be to test for breaks in a broad range of periods and to determine the most likely breakpoints for each subject based on the results. Both would complicate the analysis on this topic and therefore fall beyond the scope of this paper. In future research, the period at which the prediction series are tested for a break should be determined at the individual level, and if the series seem to undergo more than one structural change multiple tests should be done if the size of the subsamples allows it. A more complete analysis of structural breaks would also benefit from an increase in the number of available observations beyond the 51 predictions per subject of this paper. A larger number of observations would increase the power of Chow's Breakpoint Test, reducing the number of failures to detect structural breaks, which seems to be substantial with the breakpoint tests of this paper.

References

Bullard, J.B. (1994) Learning Equilibria, *Journal of Economic Theory* 64, *pp.* 468-485.

Grandmont, J. (1985) On Endogenous Competitive Business Cycles, *Econometrica* 53-5, *pp.* 995-1045.

Lucas, R.E. (1986) Adaptive Behavior and Economic Theory, *Journal of Business* 59-4-2, *pp.* S401-S426.

Marimon, R. & Spear, S.E. & Sunder, S. (1993) Expectationally Driven Market Volatility: An Experimental Study, *Journal of Economic Theory* 61, *pp.* 74-103.

Marimon, R. & Sunder, S. (1993) Indeterminacy of Equilibria in a Hyperinflationary World: Experimental Evidence, *Econometrica* 61-5, *pp.* 1073-107.

Marimon, R. & Sunder, S. (1994) Expectations and Learning under Alternative Monetary Regimes: An Experimental Approach, *Economic Theory* 4-1, *pp.* 131-162.

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

Samuelson, P.A. (1958) An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money, *Journal of Political Economy* , *pp.* 467-482.

Schönhofer, M. (1999) Chaotic Learning Equilibria, *Journal of Economic Theory* 89, *pp.* 1-20.

Tuinstra, J. (1999) *Price Dynamics in Equilibrium Models*, Tinbergen Institute Research Series No. 214, University of Amsterdam.

Tuinstra, J. (2003) Beliefs Equilibria in an Overlapping Generations Model, *Journal of Economic Behavior & Organization* 50, *pp.* 145-164.

Tuinstra, J. & Wagener, F.O.O. (2007) On Learning Equilibria, *Economic Theory* 30, *pp.* 493-513.

Appendix: Experimental Protocol

The experiment was performed in two sessions, both taking place on April 25th 2006. In both sessions an experimental treatment was completed. The experiment was conducted in the CREED laboratory at the economics faculty of the University of Amsterdam and organized by the research group CeNDEF.[31] The experimenters were J. Tuinstra (PhD) and P. Heemeijer (MSc), as part of a team that also includes prof. C. Hommes and prof. J. Sonnemans. The laboratory software was written in *Mathematica* bij P. Heemeijer.

During both sessions, when an excess of participants had left, a short welcoming message was read aloud from paper, after which participants were randomly assigned a place in the laboratory. Subjects were received in a room separate from the laboratory. The required amount of subjects was in both cases 16, making up one experimental treatment. In the case of an excess of subjects, volunteers were first asked to accept an immediate reward of 5 euros not to participate in the experiment; if the number of volunteers was not sufficient to remove the excess, people were randomly assigned to take 5 euros and leave until the required number was reached. The laboratory consists of several rows of cubicles equipped with computers, at the time of the experiment supplemented with pencil, paper and pocket calculator. The experiment was fully computerized, so no instructions were handed out in the laboratory.

When all participants were seated they were asked to begin reading the instruc-tions on their computer screens. Instructions were identical for both treatments. It had been made clear to the participants that they could at any time call one of the experimenters if they had a question. After everyone had finished reading the instructions, the experiment automatically started. When the experiment was finished, the participants were called to the reception room one by one to receive their earnings in cash. They left the laboratory immediately afterwards.

Below a translation follows of the instructions presented to the participants before the experiment started. A set of screenshots of the instructions in the original form is attached.

Translation of Experimental Instructions

Welcome to the economic laboratory

---

[31] CREED stands for Center for Research in Experimental Economics and political Decision-making; CeNDEF for Center for Non-linear Dynamics in Economics and Finance. The Faculty of Economics and Business (FEB), University of Amsterdam, is located at Roetersstraat 11, the CREED laboratory in room B515.

Experimental Team
Peter Heemeijer, Cars Hommes, Joep Sonnemans & Jan Tuinstra.
— Computer program written in *Mathematica* by Peter Heemeijer —

Structure of the experiment

You are part of an experiment about economic decision making. You will be rewarded based on the decisions you make during the experiment. The experiment will be preceded by a number of pages with instructions, which will explain to you how exactly it works. After the experiment you will be asked to answer a number of questions regarding its results.[32]

- The entire experiment, including the instructions and the questionnaire, is run on the computer. Therefore, you are not required to submit the sheet of paper on your desk, but you can use it to make notes.

- On your desk is a calculator. If necessary, you can use it during the experiment.

- If at any time you have a question, raise your hand, then someone will approach you for assistance.

General information about the experiment

You are a *statistical research bureau* that earns its income by making predictions of the *price level of consumption goods* in the economy. In particular you regularly make predictions concerning the change in the price level of consumption goods, i.e. the *inflation*. The experiment consists of a total of *50 periods*. In each period you are asked to predict the inflation in the price of consumption goods; your *reward at the end of the experiment* will be based on the accuracy of your predictions.

In the following instructions you will get more information about the *economy* in which you operate, about the *market for consumption goods* that your predictions are relevant to, and about the exact way in which *making predictions* works during the experiment. Also, the *computer program* used during the experiment will be explained.

Information about the economy

The economy you are part of splits into a *young population segment*, consisting of individuals of employable age (roughly 18 through 65 years old), and an *old population segment*, consisting of individuals who no longer work because of their age. Individuals from the young population segment receive an *income* consisting of a *fixed number of consumption goods*; individuals from the old segment, who no longer work, receive no income. In the economy the possibility exists

---

[32] In reality however subjects did not answer any questions at the end of the experiment. Originally, the intention was to present participants with a questionnaire, but due to a technical problem this was cancelled.

for young people to *save* part of their income in order to obtain consumption goods when they have reached old age.

Your predictions of the inflation are used by individuals from the young popu-lation segment to determine *which part* of their present income they will save for the time when they themselves will belong to the old segment. The *money* which individuals use to keep their savings, is brought into circulation by a *central bank*.

Information about the market for consumption goods

In every period you make a prediction about the inflation in the price of consumption goods. Based on this young people determine which part of their income they will save by *selling* goods; the remaining goods they consume before retiring. Old people are not sensitive to your prediction, since they will always use their accumulated savings to *buy* consumption goods. The *true price level* on the market for consumption goods is in each period determined in such a way that with the savings of old people the available goods, offered by young people, can exactly be bought. Also, the price level is in each period slightly influenced in an unpredictable way by circumstances in the rest of the economy.

As stated earlier, a *central bank* brings money into circulation that individuals use to save. It is known that this bank has the tendency to let the *total money supply* increase slowly.

Information about making predictions

As stated earlier, the experiment consists of a total of *50 periods*. In each period you make a prediction about the inflation in the price of consumption goods. Because the true price level in every period is partly determined by the expected price level in the following period, which after all makes young people decide how much of their consumption goods to sell, your predictions of the inflation will be *one period ahead*.

Suppose for example that the experiment has progressed to period 12. You will then predict the inflation in the next period 13, i.e. the change in the price level between periods 12 and 13 [this is an error; given the inflation definition $\pi_t = p_{t+1} / p_t$, it should be "between periods 13 and 14"]. When making your predic-tion you can use the following *information* (which will be shown on your computer screen): the *inflations* up to and including the previous period 11, and your *predictions of the inflation* up to and including the present period 12. Notice that in period 12 the inflation prediction for that period (i.e. the change in price level between periods 11 and 12) [should be "between periods 12 and 13"] is already available to you, because you also predicted one period ahead in the previous period 11.

Information about your reward (part 1 of 2)

Your *reward after the experiment has ended* increases with the accuracy of your predictions. In the experiment this accuracy is measured by the *absolute error* between your predictions of the

inflation and its true values. For each period this absolute error is calculated as soon as the true value of the inflation is known; you will then receive a Prediction Score that increases the smaller the absolute error becomes. The *table below* gives the relation between the absolute prediction error and the Prediction Score. If for a certain period you predict for example an inflation of 2%, and the true inflation turns out to be 7%, then you make an absolute error of 7% – 2% = 5%. You will therefore receive a Prediction Score of 80.

— *Paragraph continues on the next page* —

[Title of table's first row] *Absolute prediction error*

[Title of table's second row] *Prediction Score*

Remark: The table serves as an illustration and contains only part of all possible prediction errors.

Information about your reward (part 2 of 2)

If you predict an inflation of 2%, and it turns out to be –3%, you also make a prediction error of 2% – (–3%) = 5%. You will therefore receive the same Prediction Score of 80. For a perfect prediction, with a prediction error of zero, you will receive a Prediction Score of 100. When you make a prediction error of 25% or more, your Prediction Score for the relevant period will be zero. Generally speaking, your Prediction Score decreases with four points if your prediction error increases with one percent. Notice that the inflation and your predictions of it in the experiment are expressed in *percentages*; naturally the inflation can become negative, because the price level might decrease.

Your *Final Score* at the end of the experiment consists simply of a summation of all Prediction Scores you have earned during the experiment. During the experiment your scores are shown on your computer screen; at the end of the experiment you will be shown an overview of your Prediction Scores, followed by the resulting Final Score. Your monetary reward consists of *1/2 eurocent* for each point in your Final Score (so 200 points equals 1 euro).

Information about the computer program (part 1 of 3)

Directly below you see an example of the *left upper part* of the computer screen during the experiment.[33] It is a *graphical depiction* of the inflations in the price of consumption goods (red series) and your predictions of them (yellow series).[34] On the horizontal axis are the *time periods*; the vertical axis has percentage as unit. In the *imaginary situation* depicted in the graph, the experiment is in period 28 and you are predicting the inflation in period 29 (the

---

[33] The figures referred to are included in the original instructions directly following these translations.

[34] In the experiment the illustrations were shown in color. In the attached original instructions the yellow prediction series is the one showing an observation at $t = 28$.

experiment lasts for 50 periods). Notice that the graph only shows results of the last 25 periods and that the next period is always at the far-right side.

Information about the computer program (part 2 of 3)

To the side you see an example of the *right upper part* of the computer screen during the experiment. It consists of a *table* with information on the history of the experiment up to a maximum of 17 periods. This information is supplementary to the graph in the left upper part of the screen. The *first column* of the table shows the time period (the next period, in the example 29, is always on top). The *second and third columns* respectively show the inflations in the price of consumption goods and your predictions of them. Finally, the *fourth column* lists your Prediction Score for each period separately. Notice that you can use the sheet of paper on your desk to save data longer than 17 periods.

Information about the computer program (part 3 of 3)

Below you see an example of the *bottom part* of the computer screen during the experiment. In each period you will be asked to submit your prediction of the inflation in the next period (directly beneath *Submit your prediction*). In the *first two periods of the experiment* no information is yet known about the market (the true inflation in period 1 becomes known as soon as you have submitted your prediction for period 2); submit in these periods inflation predictions that seem reasonable at first glance. When entering your prediction, use the *decimal point* if necessary. Notice that your predictions and the true inflation in the experiment are rounded to two decimals. In addition Prediction Scores are rounded to whole numbers.

> — *You are now ready to start the experiment* —

> Please wait until all participants are ready to begin the experiment

> *Remark*: When all participants are ready, the experiment will start immediately. There are consequently no practice rounds. When the experiment begins, it may take a moment for the experimental program to load.
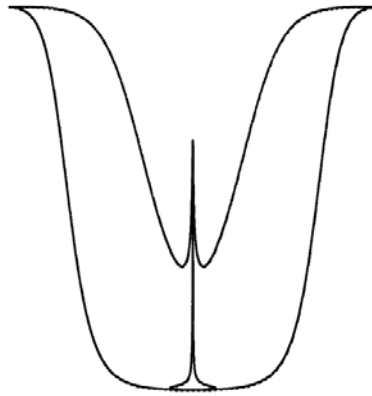
CeNDEF
Center for Non-linear Dynamics in Economics and Finance

WELKOM IN HET ECONOMISCH LABORATORIUM

<u>Experimenteel Team</u>
Peter Heemeijer, Cars Hommes, Joep Sonnemans & Jan Tuinstra.
— Computerprogramma geschreven in *Mathematica* door Peter Heemeijer —

Opzet van het experiment

U doet mee aan een experiment naar economische besluitvorming. Op basis van de beslissingen die u tijdens het experiment neemt, zult u worden beloond. Het experiment zal voorafgegaan worden door een aantal pagina's aan instructies, waarin wordt uitgelegd hoe het in zijn werk gaat. Na afloop van het experiment zult u worden gevraagd om een aantal vragen over het verloop ervan te beant-woorden.

- Het gehele experiment, inclusief de instructies en de vragenlijst, verloopt via de computer. U hoeft het papier dat op uw bureau ligt dus niet in te leveren, maar u kunt het gebruiken om aantekeningen te maken.

- Op uw bureau ligt een rekenmachine. Die kunt u, zo nodig, tijdens het experiment gebruiken.

- Heeft u op enig moment een vraag, steek dan uw hand op, dan komt iemand u helpen.

Algemene informatie over het experiment

U bent een *statistisch onderzoeksbureau* dat zijn inkomsten verkrijgt door het doen van voorspellingen over het *prijsniveau van consumptiegoederen* in de economie. In het bijzonder doet u regelmatig voorspellingen over de verandering in het prijsniveau van consumptiegoederen, ofwel de *inflatie*. Dit experiment bestaat in totaal uit *50 perioden*. In iedere periode wordt u gevraagd om een voorspelling te doen van de inflatie in de prijs van consumptiegoederen; uw *beloning na afloop van het experiment* is gebaseerd op de nauwkeurigheid van uw voorspellingen.

In de hierna volgende instructies krijgt u meer informatie over de *economie* waarin u zich bevindt, over de *markt voor consumptiegoederen* waarop uw voorspellingen betrekking hebben, en over de manier waarop *het doen van voorspellingen* tijdens het experiment in zijn werk gaat. Daarnaast zal het *computerprogramma* worden toegelicht dat tijdens het experiment wordt gebruikt.

Informatie over de economie

De economie waarin u zich bevindt, valt uiteen in een *jonge bevolkingsgroep*, bestaande uit personen van werkzame leeftijd (ruwweg 18 tot en met 65 jaar), en een *oude bevolkingsgroep*, bestaande uit personen die vanwege hun leeftijd niet langer werkzaam zijn. Personen uit de jonge bevolkingsgroep ontvangen een *inkomen* bestaande uit een *vast aantal consumptiegoederen*; personen uit de oude groep, die niet meer werken, ontvangen geen inkomen. In de economie bestaat de mogelijkheid voor jonge personen om een deel van hun inkomen te *sparen* om ook op hun oude dag over consumptiegoederen te kunnen beschikken.

Uw voorspellingen van de inflatie worden gebruikt door personen uit de jonge bevolkingsgroep om te bepalen *welk deel* van hun huidige inkomen ze zullen sparen voor de tijd waarin ze zelf tot de oude bevolkingsgroep behoren. Het *geld* waarin personen hun besparingen aanhouden, wordt in omloop gebracht door een *centrale bank*.

Informatie over de markt voor consumptiegoederen

In iedere periode van het experiment doet u een voorspelling over de inflatie in de prijs van consumptiegoederen. Op basis hiervan bepalen jonge personen welk deel van hun inkomen ze zullen sparen door goederen te *verkopen*; de rest maken ze op voordat ze met pensioen gaan. Oude mensen zijn niet gevoelig voor uw voorspelling, omdat ze altijd hun eerder opgebouwde besparingen gebruiken om consumptiegoederen te *kopen*. Het *werkelijke prijsniveau* op de markt voor consumptiegoederen krijgt steeds een hoogte zodanig dat oude personen met hun spaargeld precies de beschikbare goederen, aangeboden door jonge personen, kunnen aanschaffen. Daarnaast wordt het prijsniveau in iedere periode op onvoorspelbare wijze lichtelijk beïnvloed door omstandigheden in de rest van de economie.

Zoals gezegd brengt een *centrale bank* het geld in omloop dat mensen gebruiken om te sparen. Het is bekend dat deze de gewoonte heeft om de *totale geldhoeveelheid* langzamerhand te laten toenemen.

Informatie over het doen van voorspellingen

Zoals gezegd bestaat het experiment in totaal uit *50 perioden*. In iedere periode doet u een voorspelling over de inflatie in de prijs van consumptiegoederen. Omdat het werkelijke prijsniveau in iedere periode mede bepaald wordt door het verwachte prijsniveau in de periode erna, dat jonge mensen immers doet beslissen hoeveel van hun consumptiegoederen ze verkopen, voorspelt u de inflatie steeds *één periode vooruit*.

Stel bijvoorbeeld dat het experiment zich bevindt in periode 12. U voorspelt dan de inflatie in de volgende periode 13, ofwel de verandering in het prijsniveau tussen periode 12 en 13. Bij het doen van uw voorspelling kunt u gebruik maken van de volgende *informatie* (die op uw computerscherm zal worden afgebeeld): de *inflaties* tot en met de vorige periode 11, en uw *voorspellingen van de inflatie* tot en met de huidige periode 12. Merk op dat u in periode 12 reeds beschikt over uw inflatievoorspelling voor die periode (dat wil zeggen de verandering in prijsniveau van periode 11 naar 12), omdat u ook in de vorige periode 11 één periode vooruit voorspelde.

Informatie over uw beloning (deel 1 van 2)

Uw *beloning na afloop van het experiment* neemt toe met de nauwkeurigheid van uw voorspellingen. In het experiment wordt deze nauwkeurigheid gemeten door de *absolute fout* tussen uw voorspellingen van de inflatie en de werkelijke waarden daarvan. Voor iedere periode wordt deze absolute fout berekend zodra de werkelijke waarde van de inflatie bekend is; u ontvangt vervolgens een *Voorspelscore* die toeneemt naar mate de absolute fout kleiner is. De *onderstaande tabel* geeft het verband weer tussen de absolute voorspelfout en de Voorspelscore. Als u voor een zekere periode bijvoorbeeld een inflatie van 2% voorspelt, en de werkelijke inflatie blijkt 7% te zijn, dan maakt u een absolute fout van 7% – 2% = 5%. U ontvangt dan een Voorspelscore van 80.

— Alinea gaat verder op de volgende pagina —

— Alinea gaat verder op de volgende pagina —

Absolute voorspelfout	0	5	10	15	20	≥ 25
Voorspelscore	100	80	60	40	20	0

Opmerking: De tabel dient ter illustratie en bevat slechts een aantal van de mogelijke voorspelfouten.

Informatie over uw beloning (deel 2 van 2)

Als u een inflatie van 2% voorspelt, en deze blijkt –3% te zijn, maakt u eveneens een voorspelfout van 2% – (–3%) = 5%. U ontvangt daarom dezelfde Voorspelscore van 80. Voor een perfecte voorspelling, met een voorspelfout van nul, ontvangt u een Voorspelscore van 100; als u een voorspelfout maakt van 25% of meer, is uw Voorspelscore in de betreffende periode nul. In het algemeen daalt uw Voorspelscore met vier punten als uw voorspelfout met een procentpunt toeneemt. Merk op dat de inflatie en uw voorspellingen ervan in het experiment worden uitgedrukt in *procenten*; de inflatie kan vanzelfsprekend negatief worden, omdat het prijsniveau van consumptiegoederen kan dalen.

Uw *Totaalscore* aan het einde van het experiment bestaat simpelweg uit een opsomming van alle Voorspelscores die u tijdens het experiment heeft behaald. Tijdens het experiment worden uw scores afgebeeld op uw computerscherm; na afloop van het experiment krijgt u een overzicht te zien van uw Voorspelscores, gevolgd door de resulterende Totaalscore. Uw uiteindelijke beloning bestaat uit *1/2 eurocent* voor iedere punt in uw Totaalscore (200 punten staat dus gelijk aan 1 euro).

Informatie over het computerprogramma (deel 1 van 3)

Hieronder ziet u een voorbeeld van het *linkerbovendeel* van het computerscherm tijdens het experiment. Het bestaat uit een *grafische voorstelling* van de inflaties in de prijs van consumptiegoederen (rode reeks) en uw voorspellingen daarvan (gele reeks). Op de horizontale as staan de *tijdsperioden* vermeld; de verticale as is in procenten. In de *denkbeeldige situatie* die de grafiek toont, is het experiment in periode 28 en voorspelt u de inflatie in periode 29 (het experiment duurt 50 perioden). Merk op dat de grafiek slechts resultaten van ten hoogste de laatste 25 perioden toont en dat de eerstvolgende periode altijd helemaal rechts staat.
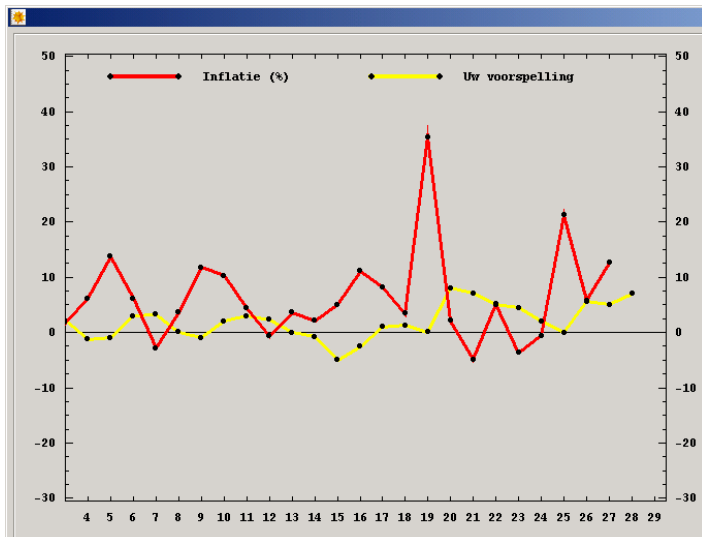
Informatie over het computerprogramma (deel 2 van 3)

Hiernaast ziet u een voorbeeld van het *rechterbovendeel* van het computerscherm tijdens het experiment. Het bestaat uit een *tabel* met informatie over het verloop van het experiment in ten hoogste de laatste 17 perioden. Deze informatie is een aanvulling op de grafiek in het linkerbovendeel van het scherm. De *eerste kolom* van de tabel laat de tijdsperiode zien (de eerstvolgende periode, in het voorbeeld 29, staat altijd bovenaan). De *tweede en derde kolom* respectievelijk tonen de inflaties in de prijs van consumptiegoederen en uw voorspellingen daarvan. De *vierde kolom* ten slotte geeft uw

Tijds-periode	Inflatie (%)	Uw Voor-spelling	Voorspel-score
29		???	
28		7.00	
27	12.67	5.00	69
26	5.83	5.56	99
25	21.28	0.00	15
24	−0.56	2.00	90
23	−3.64	4.50	67
22	5.16	5.00	99
21	−4.93	7.12	52
20	2.17	8.00	77
19	35.42	0.11	0
18	3.45	1.25	91
17	8.27	1.00	71
16	11.13	−2.56	45
15	5.03	−5.00	60
14	2.12	−0.75	89
13	3.61	0.00	86
12	−0.66	2.30	88
11	4.36	3.00	95

Voorspelscore bij iedere afzonderlijke periode. Merk op dat u het kladvel op uw bureau kunt gebruiken om gegevens langer dan 17 perioden te bewaren.

Informatie over het computerprogramma (deel 3 van 3)

Hieronder ziet u een voorbeeld van het *onderste deel* van het computerscherm tijdens het experiment. In iedere periode wordt u gevraagd om uw voorspelling van de inflatie in de volgende periode te geven (onder *Vul uw voorspelling in*). In de *eerste twee perioden van het experiment* is er nog geen informatie bekend over de markt (de werkelijke inflatie in periode 1 wordt bekend zodra u uw voorspelling voor periode 2 heeft opgegeven); doe in deze perioden inflatievoorspellingen die u op het eerste gezicht redelijk lijken. Maak bij het invoeren van uw voorspelling zo nodig gebruik van de *decimale punt*. Merk op dat uw voorspellingen en de werkelijke inflatie in het experiment worden afgerond op twee decimalen. Daarnaast worden Voorspelscores afgerond op gehele punten.

Het is nu periode **28**. Het experiment zal **50** perioden duren.
GELIEVE UW VOORSPELLING VAN DE INFLATIE IN PERIODE **29** IN TE VOEREN

Vul uw voorspelling in:

Doorgaan

— *U bent nu gereed om aan het experiment te beginnen* —

_____

GELIEVE TE WACHTEN TOT ALLE PARTICIPANTEN

GEREED ZIJN OM MET HET EXPERIMENT TE BEGINNEN

_____

Opmerking: Als alle participanten gereed zijn, begint het experiment meteen.
Er zijn dus geen oefenrondes. Het kan straks even duren
voordat het experimentele programma geladen is.