Asset price volatility and investment horizons: An experimental investigation*

Mikhail Anufriev [a,†] Aleksei Chernulich [b,‡] Jan Tuinstra [c,§]

30 October 2021

[a] *University of Technology Sydney, Business School, Economics Discipline Group*
[b] *New York University Abu Dhabi, Center for Behavioral Institutional Design (C-BID)*
[c] *University of Amsterdam, Amsterdam School of Economics, CeNDEF*

Abstract

We study the effects of the investment horizon on asset price volatility using a Learning to Forecast laboratory experiment. We find that, for short investment horizons, participants coordinate on self-fulfilling trend-extrapolating predictions. Price deviations are then reinforced and amplified, possibly leading to large bubbles and crashes in asset prices. For longer investment horizons such bubbles do not emerge and price volatility tends to be lower. This is due to the fact that, for longer horizons, there is more dispersion in participants' forecasts, and participants extrapolate trends in past prices to a lesser extent. We also show that, independent of the investment horizon, if the initial history of asset prices is already relatively stable before participants start their prediction task, price volatility remains small, with prices close to their fundamental values for the duration of the experiment.

Keywords: Experimental Economics, Expectations, Asset Pricing, Investment Horizons, Behavioral Finance.

JEL Classification: C91, G12, G14, G41.

---

*We thank the editor, Daniel Houser, an anonymous reviewer, David Goldbaum, Anita Kopányi-Peuker, Joep Sonnemans, and participants at the 2018 BEEF workshop at the University of Sydney, the 2018 BEAM-ABEE workshop at the University of Amsterdam, the IMEBESS 2019 meeting in Utrecht, the 2019 Asia Meeting of the Econometric Society at the Xiamen University, and the WEHIA 2019 workshop for useful comments. The authors acknowledge financial support from the Australian Research Council's Discovery Projects funding scheme (projects DP140103566 and DP200101438) and from the ORA project "BEAM" (NWO 464-15-143) which is partly financed by the Netherlands Organization for Scientific Research (NWO). Aleksei Chernulich is grateful for financial support from Tamkeen under the NYUAD Research Institute award for Project CG005.

† *E-mail:* Mikhail.Anufriev@uts.edu.au.
‡ *E-mail:* Aleksei.Chernulich@nyu.edu.
§ *E-mail:* J.Tuinstra@uva.nl.

1 Introduction

The extent to which asset prices can be characterized by persistent deviations from fundamental values continues to spark the debate on the functioning of financial markets. This debate has spawned a sizable literature in experimental and behavioral finance. Mispricing as well as bubbles and crashes in asset prices are robust features of laboratory experiments where human subjects engage in trading a financial asset. These phenomena, which have been linked to bounded rationality both in speculative trading and in forecasting future market prices, occur under a vast number of different conditions, see Palan (2013) for a review.

The present paper belongs to the growing literature of so-called "*Learning to Forecast*" (LtF) laboratory experiments.[1] In LtF experiments participants act as professional forecasters for hypothetical institutional investors. Given the participants' forecasts, the investors' trading decisions are optimal, so that any deviation of prices from fundamental values is due to bounded rationality in forecasting. Participants forecast the future price of an infinitely lived risky asset for many consecutive periods, and are rewarded according to their forecasting accuracy.[2] The self-referential nature of financial markets, where expectations of future prices are an important determinant of market-clearing prices, is explicitly taken into account in such an experiment. In fact, there exists a *positive* feedback link between forecasts and prices in financial markets: an expected future increase in the price of a stock increases the current demand for that stock and, consequently, increases the market-clearing price as well. This feedback creates an opportunity for destabilizing trend-extrapolating forecasting strategies to outperform other strategies, thereby creating persistent mispricing, see for example the theoretical arguments in Anufriev and Hommes (2012b) and Anufriev et al. (2019).

In previous LtF experiments – which typically feature large bubbles and crashes in asset prices – the investment (and hence forecast) horizon is fixed to one period, which seems quite restrictive. The literature on behavioral finance shows that short investment horizons may be conducive to mispricing, see, e.g., De Long et al. (1990) and Dow and Gorton (1994) for theoretical papers, and Cella et al. (2013) and Cremers and Pareek (2015) for empirical support. The main goal of

[1] The Learning to Forecast experimental design was introduced in Marimon et al. (1993). Early contributions are reviewed in Hommes (2011). Recent examples include Bao et al. (2017), Kopányi et al. (2019), Kopányi-Peuker and Weber (2020), and Hommes et al. (2021).

[2] The advantage of the LtF experimental design is that it allows the experimenters to obtain clean data on participants' forecasts. A potential drawback is that participants' forecasts might not be consistent with their trading behavior. However, Bao et al. (2013, 2017) demonstrate that the results from LtF experiments are similar to those of so-called "Learning to Optimize" experiments, where participants are allowed to actively trade in the asset.

the current paper is to understand the role that the investment or forecast horizon plays in the emergence of excess volatility in general, and these bubbles and crashes in particular. To that end, we design and run a novel LtF experiment, where we vary the investment horizon of the institutional investors and, as a consequence, the forecast horizon of the participants. Depending on the treatment, participants have to predict, at the beginning of period t, the price for either period $t+1$, period $t+2$, or period $t+3$. The variation in the horizon is our first treatment variable. Our second treatment variable is the initial price history. Before participants start their prediction task, they observe the ten previous prices, where in some treatments this price history is more volatile than in others.[3] This treatment variation allows us to study the effect of the initial history on price volatility. Importantly, in all our treatments, we have the same rational expectation equilibrium characterized by a constant fundamental price augmented by small, unpredictable noise.

We find that price fluctuations tend to decrease with the investment horizon. One explanation for this may be that an increase in the investment horizon decreases the relevant discount factor and thereby weakens the strength of the expectations feedback. This would decrease price volatility, even if participants' behavior is unaffected by the investment horizon. To control for this effect, we add a treatment that compensates for the decrease in feedback strength for longer investment horizons. Even with this correction, price fluctuations are typically smaller for longer horizons, confirming that an increase in the horizon tends to stabilize asset price dynamics. In particular, large bubbles and crashes only emerge for the shortest possible investment horizon. In relation to the second treatment variable, we find that in the treatments with a relatively stable initial price history price volatility remains small until the end of the experiment, with prices staying in the vicinity of their fundamental value. Fluctuations tend to be much larger in the treatments with a more volatile initial price history. Taken together, these results suggest that policy measures that limit the possibilities for trading with a short investment horizon, e.g., by imposing a minimum holding period for the asset, may contribute to stabilizing asset markets and reducing excess price volatility, and that the effect of these measures will last beyond the period for which they are in place.

[3]In many LtF experiments – where participants typically start with a blank slate – there is substantial heterogeneity in the endogenous price dynamics between groups in the same treatment (see, e.g., Hommes et al., 2005 and Kopányi et al., 2019). Anufriev and Hommes (2012a,b) explain these outcomes by the Heuristic Switching Model (HSM). This model generates path-dependent price dynamics, suggesting that the between-group heterogeneity in LtF experiments can be a consequence of the differences in price dynamics that endogenously emerge in the first few periods of the experiments.

The large bubbles that emerge for a short investment horizon can be attributed to participants coordinating on a trend-extrapolating prediction strategy. An increase in prices then leads all participants to expect a further price increase which, due to the positive feedback nature of the asset pricing model, leads to higher prices, confirming participants' expectations. We provide two explanations for the decrease in price volatility for longer horizons. First, participants have a much stronger tendency to extrapolate trends over shorter investment horizons than over longer horizons. Second, participants fail to coordinate their expectations: dispersion of individual forecasts is typically higher when the investment horizon is longer. This inhibits the emergence of large self-confirming deviations from the fundamental value.[4] Finally, we do not find a significant difference in the forecasting rules that participants use when we compare treatments with small and large initial price volatility. This suggests that, in order for trend-extrapolating rules to be destabilizing, the initial price dynamics should be sufficiently volatile to begin with.

We contribute to the growing experimental literature on forecasting behavior along two dimensions, by studying: (i) the impact of "priming" participants by manipulating the initial price history, and (ii) the effect of the forecasting horizon. To the best of our knowledge, the only other LtF experiment that controls for the price history is Hennequin (2019). She finds that the incidence and size of bubbles increases in markets that contain more participants that have experienced bubbles before. This is consistent with our finding that an unstable initial price history induces volatile price dynamics, although in our experiment participants only *observe*, but do not *experience*, the price history.[5]

For the vast majority of LtF experiments, a short investment horizon is assumed, with participants predicting the price for either the current or the next period. One exception is Colasante et al. (2020) who consider an LtF experiment where in each period participants have to predict both the price for the current period, and the prices for all remaining periods, with realized prices only depending on the forecasts for the current period. They also find more dispersion in long-run expectations than in short-run expectations, although – in their case – this does not affect price volatility. Evans et al. (2019) develop a Lucas tree asset pricing model where trading decisions depend upon the *average* expected price

[4] Related to this, Patton and Timmermann (2010), who study survey data on macroeconomic forecasts of GDP growth and inflation by private sector forecasters, find that there is much more dispersion in long-horizon forecasts than in short-horizon forecasts.

[5] Malmendier and Nagel (2011, 2016) provide empirical evidence that expectations of economic agents are affected by the experiences they have during their lives, although that information is also available for those who did not experience it. This effect is sometimes called the *experience-description gap* (see Hertwig et al., 2004).

over a finite horizon, and run an LtF experiment where they vary that horizon. They find that for longer horizons dispersion of forecasts is higher and that the price is closer to the fundamental value. These results are in line with ours, even if the underlying model is different. The maximal investment horizon that Evans et al. (2019) use is substantially longer than in our experiment (10 instead of 3 periods), which implies a much stronger decrease in feedback strength. We find that already a minor increase in the investment horizon can have a substantial effect upon the ensuing price dynamics, even when correcting for the decrease in feedback strength.

Finally, the effect of the investment horizon has been studied in a number of experimental markets, where participants can trade in the asset. Hirota and Sunder (2007) and Hirota et al. (2020), for example, distinguish long-horizon traders, who can trade in the asset until it reaches maturity and pays out its dividend, and short-horizon traders, who leave the market before the asset matures and can therefore only profit from speculating on capital gains. These short-horizon traders contribute substantially to mispricing and bubble formation. However, Razen et al. (2017), who also consider trade in an asset with a single dividend that is paid out when the asset reaches maturity, find that an *increase* in the date of maturity (i.e., the investment horizon) may lead to bubbles, which they attribute to the decreased salience of the dividend payment.

The remainder of the paper is organized as follows. The next section presents the experimental design and the hypotheses. Results are presented in Section 3. Section 4 concludes. Experimental instructions, statistical tests and additional data analysis can be found in the appendices.

2 Experimental Design

The experiment took place at the University of Technology Sydney's Behavioural Lab in May, August, and September 2018. Using the Online Recruitment System for Economic Experiments (Greiner, 2015), we recruited 210 students who participated in seven experimental treatments. No student participated in more than one session. At each session, participants were randomly divided into groups of six that operated in the same market during the whole experiment. The experimenter read the instructions aloud (participants had printed copies of the instructions as well), after which participants did a short quiz that tested their familiarity with the task. Once all participants completed the quiz, the computerized experiment, programmed in z-Tree (Fischbacher, 2007), took place. After the experiment, participants filled in a questionnaire and received payment under the double-blind protocol. Sessions lasted approximately 80 minutes, and participants earned on average 27 Australian dollars, including a show-up fee.

In the remainder of this section, we introduce the price generating mechanism used in our LtF experiment (Section 2.1), discuss our experimental treatments (Section 2.2), and outline the main hypotheses (Section 2.3).

2.1 Price forecasts and market prices: the price generating mechanism

Our experiment is based on the standard Learning-to-Forecast (LtF) experimental setup, see Hommes (2011) for an overview. LtF experiments use the present value model of asset pricing, see, e.g., Campbell et al. (1997), micro-founded and extended to heterogeneous beliefs in Brock and Hommes (1998). In the model, mean-variance investors divide their wealth between a risk-free and a risky asset. The risk-free asset has gross return $R = 1 + r > 1$. The risky asset pays an IID dividend with mean $\bar{y}$ each period; the dividend process is common knowledge. The price of the risky asset, p_t, is determined from the market clearing condition. In the original model the price evolves as

$$p_t = p^f + \frac{1}{R}\left(\bar{p}^e_{t+1} - p^f\right) , \tag{1}$$

where $p^f = \bar{y}/r$ denotes the so-called *fundamental value* of the risky asset, i.e., the discounted expected value of future dividends, and $\bar{p}^e_{t+1}$ are average investor expectations made in the beginning of time t. The original model assumes that traders are myopic and the investment horizon of each trader is one period. Consequently, the expectations at time t are about the next period price, p_{t+1}.

In Appendix A, we generalize this model to the case where traders with arbitrary investment horizons are present in the market, in fixed proportions. The interpretation that we use in our experiment is that institutional investors, such as pension funds, trade every period, but each time operate on behalf of retirees with a specific and fixed investment horizon.[6]

In our experiment we focus on the situation where *all* investors have the same investment horizon H. We leave the case where different investors have different investment horizons for future research. To determine optimal demand in

[6]For example, a pension fund may pursue a strategy with a horizon of 10 years, investing the funds every month on behalf of a new cohort of clients.

period t, investor i forms expectations about the dividend stream, $\mathrm{E}_{t,i}[y_{t+s}]$ for $s = 1, \ldots, H$, and about the market clearing price in period $t+H$, $\mathrm{E}_{t,i}[p_{t+H}]$. Only these expectations are relevant for the cohort of clients of i investing at time t. Under our assumption about the dividend process, the stream of discounted expected dividend payments over the next H periods is the same for all investors and given by

$$\sum_{s=1}^{H} R^{H-s}\mathrm{E}_{t,i}[y_{t+s}] = \bar{y}\,\frac{R^H - 1}{R-1} = p^f\left(R^H - 1\right).$$

Instead, the price forecasts are heterogeneous, and the demand of mean-variance investor i, given his expectations, can be shown to be

$$z_{t,i} = \frac{1}{a\sigma^2}\left(\mathrm{E}_{t,i}[p_{t+H}] + p^f(R^H - 1) - R^h p_t\right) \,,$$

where a is the risk aversion coefficient, and σ^2 is the perceived variance of the risky asset return during H periods, assumed to be the same for all investors. The market-clearing price is then given by

$$p_t = p^f + \frac{1}{R^H}\left(\bar{p}^e_{t+H} - p^f - \varepsilon_t\right), \tag{2}$$

where $\bar{p}^e_{t+H}$ is the average prediction for the price in period $t+H$ and ε_t is a small random outside supply of the asset from noise traders, with mean zero. Equation (2) governs the relation between individual forecasts and market-clearing prices that we use in our LtF laboratory experiment.[7] Note that in period t, subjects have to predict the price for time $t+H$ on the basis of information up to time $t-1$, since p_t is not in their information set yet.[8]

The *rational expectation equilibrium* is given by the fundamental value p^f. Indeed, if all traders predict the fundamental price, the realized price will be equal to p^f in expectation, and forecasts will be correct, on average. Importantly, the fundamental value is independent of the investment horizon.

The price generating mechanism (2) defines the current market price p_t as a weighted average of the fundamental value and average expectations for the price in period $t+H$. This creates *positive expectations feedback*: an increase in average expectations of the future price increases the demand for the asset and, through

[7] As all investors have the same investment horizon, there is no *structural* dependence between subsequent periods for $H > 1$. For example, when $H = 2$, price expectations for the odd periods determine prices in the odd periods, and price expectations for even periods determine prices in even periods. However, there could be a *behavioral* dependence: expectations could be informed by past prices from both odd and even periods.

[8] For this reason, experiments with $H = 1$ are referred as "two-periods ahead" LtF experiments in the literature. To avoid confusion we will not use this terminology.

market clearing, leads to an instantaneous increase in the current market price. The strength of this feedback is given by the discount factor $1/R^H$ in (2) which diminishes with the investment horizon H, with R^H reflecting the opportunity cost of investing in the risky asset. As a consequence, an expected price increase will have a smaller effect on the market-clearing price, if it lies in the more distant future.[9]

A potential issue for LtF experiments based on the pricing equation (2) is that extreme "outlier" forecasts have a substantial effect on the realized price dynamics. One way around this problem would be to let the realized price be a function of the median, instead of the average, forecast, as in Arifovic and Petersen (2017). Another possibility would be to exclude forecasts that deviate too much from the last price when determining the average forecast, as in Kopányi-Peuker and Weber (2020). We use the standard LtF setup in our experiment, as this is the most straightforward implementation of the underlying asset pricing model.

2.2 Treatments

In our experiment, participants have to forecast asset prices for about 50 consecutive periods. The instructions (see Appendix B) specify the participant's role as a "financial forecaster" for a large pension fund. It is explained that the participant's task is to give point forecasts of the future asset price, and that based on those forecasts, the pension fund the participant advises will make trading decisions. Following the standard practice of LtF experiments, the instructions do not specify the exact pricing equation. However, they highlight qualitative features of the market, e.g., that a higher price forecast leads to a larger demand for the asset by the fund, that there are several funds affecting the total demand (the number of funds is unknown to the participants), and that the realized price follows from the equilibrium between aggregate demand and fixed supply. The numerical values of the interest rate r and the mean dividend $\bar{y}$ are provided to the participants which, in principle, allows them to calculate the fundamental value $p^f = \bar{y}/r$. We choose values of r and $\bar{y}$ in such a way that this fundamental value is 60 in all treatments. The instructions carefully explain the timing of the forecasting task, i.e., for how many periods ahead the price forecast is made, and how forecasting accuracy is determined.

[9]Mean-variance optimization underlying our theoretical model results in a constant discount factor *per period*, inversely related to the gross return of the risk-free asset. The discount factor gets smaller with a longer horizon, reflecting that investors betting on growing prices of the risky asset require a larger expected price increase to hold the asset for a longer time.

Figure 1: An example of the experimental screen.

During the experiment, the information at the disposal of a participant consists of all past realized prices and own past predictions shown in the computer screen, see Fig. 1. This information is presented in both graph and table format. The last period earnings and cumulative earnings are also shown on the screen. In each decision period, participants should type their forecast in the box in the lower part of the screen and submit it by pressing the "OK" button.[10] The timer in the upper part of the screen suggests participants to submit their forecast within 30 seconds, but this is not a binding restriction.

Participants are rewarded for their forecasting accuracy under the following scoring rule widely used in LtF experiments:

$$e_{i,t} = \max \left\{ 1300 - \frac{1300}{49} \left(p_t - p^e_{i,t} \right)^2, 0 \right\}, \tag{3}$$

where $e_{i,t}$ are the points earned by participant i in period t and $p^e_{i,t}$ is the participant's point prediction for price p_t. Rule (3) is decreasing in the quadratic forecasting error, providing incentives to make accurate predictions, and is truncated at zero to avoid that participants suffer losses. Points accumulated during 50 periods (as specified below) are converted to Australian dollars (AUD), with the exchange rate of 0.5 AUD for 1300 points, and are paid to the participants on top

[10] No restrictions are given to the participants about the prices they can submit. However, to be consistent with earlier LtF experiments, only numbers between 0 and 1000 with up to two decimal places are accepted. A participant will only learn about these restrictions after trying to submit a prediction that is not acceptable with a prompt indicating the bounds and asking to submit another prediction. This rarely happens in our experiment.

Treatment	Initial price	Horizon H	Interest r	Dividend $\bar{y}$	Feedback $1/(1+r)^H$	Paid periods	Number of groups
ConH1		1	5%	3	0.95	12 – 61	3
ConH2	Converging	2	5%	3	0.91	13 – 62	4
ConH3		3	5%	3	0.86	14 – 63	3
OscH1		1	5%	3	0.95	12 – 61	8
OscH2	Oscillating	2	5%	3	0.91	13 – 62	5
OscH3		3	5%	3	0.86	14 – 63	6
OscH3SF		3	1.6%	0.96	0.95	14 – 63	6

Table 1: Information on experimental design.

of a show-up fee.[11] Note that expected payoffs are maximized if all participants predict the fundamental value in each period.

Our experiment features seven treatments. They are summarized in Table 1. Fundamental price, $p^f = \bar{y}/r$, is the same in each treatment and equal to 60. There are $N = 6$ participants in each group and the same realization of shocks $\varepsilon_t \sim N(0, 0.25)$ is used in every group of each treatment. The treatments differ along three dimensions. First, we use two different initial price histories as initial conditions for the experiment, to study the effect of previous price developments on the ensuing price dynamics in a positive expectations feedback environment. For the two initial price histories, we take the first 10 prices in two groups from the standard LtF experiment in Hommes et al. (2005). The left panels of Fig. 3 below start with the first 10 prices observed in group 6 of the Hommes et al. (2005), and the right panels start with the first 10 prices observed in group 7 of the same experiment. Both price histories are hill-shaped and deviate from the fundamental price, which is also 60 in the Hommes et al. (2005) experiment. However, the deviations of the first price history (all 10 prices are between 56.0 and 60.0) are much smaller than those of the second price history, where all prices are between 44.8 and 67.1. From now on we will refer to the first price history as "Converging" and to the second price history as "Oscillating".[12] Note that, given the initial price history of 10 periods, the first price that is determined endogenously in the experiment is the price in period $t = 11$.

[11] We set a show-up fee of 10 AUD and communicated this in advance to the participants. Because payoffs in some treatments were quite low, we increased show-up fees for those treatments ex post. In the end show-up fees ranged between 10 and 20 AUD.

[12] The two price histories are qualitatively relatively similar, with prices first increasing up to, or larger than the fundamental value, and then decreasing again. Alternatively, we could have chosen price histories that are farther apart (e.g., one where prices slowly approach the fundamental value from below versus one where prices are monotonically increasing and go far beyond the fundamental value in the first 10 periods). In order to focus on the effect of small variations in initial conditions, we did not want to make the differences between the two price histories too large.

The second, and most important, variation in treatments is the investment horizon H that affects pricing equation (2). All participants in the same treatment have the same investment horizon, which is either $H = 1$, $H = 2$, or $H = 3$. For treatments with $H = 1$, participants start the experiment by predicting the price for period 12. These forecasts determine, through equation (2), price p_{11}. After learning this price, participants predict the price for period 13, which determines p_{12}, and so on. The final forecast is for period 62, and determines price p_{61}. The forecasting accuracy of the participants' predictions can only be evaluated for periods 12 up to 61. Thus, participants are rewarded for their forecasts for these periods only (see column 'Paid periods' in Table 1). Similarly, for treatments with $H = 2$, the first prediction, which determines p_{11}, is for the price in period 13, and the final prediction, determining p_{62}, is for period 64. In these treatments, participants are rewarded for the predictions for periods 13 to 62. Finally, for treatments with $H = 3$, the first prediction, determining p_{11}, is for period 14, and the final prediction, determining p_{63}, is for period 66. In the treatments with $H = 3$, the participants are rewarded for predictions for periods 14 to 63. The corresponding dynamic structure of the experiment, relating the current period, periods for which information is available, and the period for which the forecast is made, is carefully explained in the instructions. Participants are explicitly informed about the periods for which their forecasting accuracy affects earnings.[13]

By having two price histories and three investment horizons, we have a 2×3 between-subjects design. We run one more treatment with the oscillatory initial price history and investment horizon $H = 3$, but with a lower interest rate, $r = 0.016$ and a mean dividend of $\bar{y} = 0.96$. This combination of lower interest rate and lower mean dividend gives rise to the same fundamental value, $p^f = 60$, as in the other treatments. However, the feedback strength, $1/R^H$, for this new treatment (with $H = 3$) is now the same as that for the original treatment with $H = 1$. We refer to this treatment as the **OscH3SF** treatment, where **SF** stands for "strong feedback".

2.3 Conjectures

For $H = 1$, the pricing equation (2) corresponds to the price generating mechanism used in earlier LtF experiments (see, e.g., Hommes et al., 2005, 2008, 2021). The following results from these experiments are remarkably robust, although there are

[13]Note that in all treatments participants were incentivized for fifty periods. Forecasts for some periods (e.g., for periods 64, 65 and 66, in treatments with $H = 3$) were not incentivized.

minor differences in the experimental design.[14] Price dynamics often do not converge to the fundamental value within 50 periods, and instead are characterized by the emergence of large bubbles and crashes, with prices regularly approaching the upper bound of 1000, exceeding the fundamental value by a factor of 16. Moreover, although participants cannot observe each others' forecasts, individual forecasts of participants in the same group are highly coordinated. These results have been attributed to the positive expectations feedback of the price generating mechanism. For instance, for a standard choice of $R = 1.05$, the market price in (2) will be close to the average expected price, even when this average deviates from the fundamental value. Participants that submit forecasts that are close to the average forecast in their group will perform relatively well. This feature promotes coordination of expectations and, as a consequence, forecast errors are typically strongly correlated across participants. Moreover, as soon as price deviates from the fundamental value, participants tend to coordinate on a so-called *trend-extrapolating heuristic*

$$p^e_{i,t+1} = p_{t-1} + \theta \left(p_{t-1} - p_{t-2} \right), \tag{4}$$

with positive θ. This forecasting rule extrapolates the most recent trend in prices into the future. For large enough θ, the use of such a rule is consistent with the endogenous emergence of bubbles.[15] Indeed, such rules have been frequently observed in LtF experiments, for example, in Hommes et al. (2005).

Our experiment aims at understanding the effect of providing a price history and increasing the investment horizon on the emergence of asset price bubbles. To the best of our knowledge, in all LtF experiments, except Hennequin (2019), participants start the experiment without any prior information about prices.[16] Given the path dependency observed in these earlier experiments, i.e., the finding that the price dynamics in the initial periods can help explain the price dynamics in the remaining periods of the experiment (see Anufriev and Hommes, 2012a), we hypothesize that the difference in price histories will affect price dynamics in our experiment, as formulated in our first conjecture.

Conjecture 1. *Price volatility will be higher in groups with an oscillating initial price history than in groups with a converging initial price history.*

[14]For example, in Hommes et al. (2005) stabilizing "robot" traders are added to most groups, and the number of participants per group varies across experiments from $N = 6$ in Hommes et al. (2005, 2008), to around 100 in some of the groups in Hommes et al. (2021).

[15]See the discussion of Fig. 2 below.

[16]In most LtF experiments, participants are only told that the first two prices are "likely" to lie between 0 and 100.

To test for differences in volatility, we will compare price volatility between treatments with the same horizons (and interest rate) but different price histories. That is, we will compare treatment **ConH1** with treatment **OscH1**, treatment **ConH2** with treatment **OscH2**, and treatment **ConH3** with treatment **OscH3**.

Our remaining conjectures are concerned with the effect of the investment horizon on aggregate price behavior and individual predictions. When comparing the pricing equation (2) for different values of H, two effects stand out. First, as explained above, the feedback strength $(1/R^H)$ decreases with H. This effect is stabilizing for the price dynamics, because with longer horizons a given deviation of the average prediction from the fundamental value will result in a smaller price deviation (see Sonnemans and Tuinstra, 2010, for the effect of feedback strength on price dynamics). Second, participants have to predict further into the future. In particular, if they believe that prices will increase between two subsequent periods, they have to extrapolate this price increase accordingly over their investment horizon. To be specific, assume that all participants believe that prices between two subsequent periods evolve as in (4) with $\theta \geq 0$. By iterating this equation forward, we obtain

$$p^e_{i,t+H} = p_{t-1} + \theta \left(1 + \theta + \cdots + \theta^H\right) \left(p_{t-1} - p_{t-2}\right), \tag{5}$$

for a participant with investment horizon H. As one would expect, an increase in H will lead to an increase in the expected price change. Hence, the trend-extrapolating heuristic will amplify price trends stronger, when the horizon is longer. Thus, this effect is destabilizing for the price dynamics. On the other hand, it assumes a consistent use of the trend-extrapolating heuristic over time, as in (5), which is computationally more demanding for longer horizons.

It depends on the values of θ, R and H, how the destabilizing effect of a further extrapolation compares to the stabilizing effect of a lower feedback strength. To illustrate this dependence, assume that all participants use trend-extrapolating heuristic (5) and that there are no supply shocks in the market-clearing equation (2). Then price dynamics in deviations from the fundamental level $x_t = p_t - p^f$ are given by

$$x_t = \frac{1}{R^H}\Big(x_{t-1} + \theta \left(1 + \theta + \cdots + \theta^H\right) \left(x_{t-1} - x_{t-2}\right)\Big). \tag{6}$$

This is a second-order linear difference equation whose steady state, $x = 0$, corresponds to the fundamental value. The speed of convergence to this steady state (or divergence from it) depends on the largest eigenvalue in modulus of this system. Fig. 2 depicts this largest eigenvalue as a function of the extrapolation coefficient

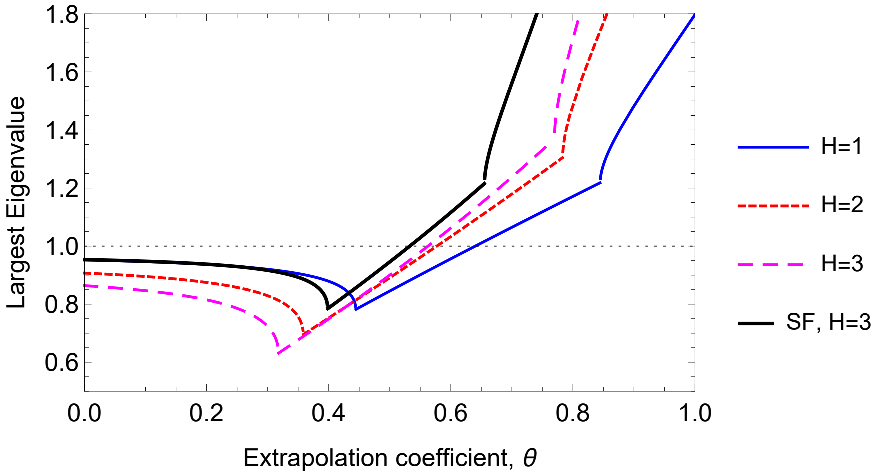

Figure 2: The largest eigenvalue for trend-extrapolating behavior in system (6).

θ for all combinations of H and R considered in this paper. The horizontal black line corresponds to the threshold value 1. Price dynamics are stable if the largest eigenvalue is below this threshold and unstable otherwise. Generally, the higher the largest eigenvalue is, the faster the price diverges (or the slower it converges).

The three curves labeled "H=1", "H=2" and "H=3" show the largest eigenvalue for the corresponding horizon, when $R = 1.05$. We observe that for small θ an increase in H stabilizes the system, and for large θ such an increase destabilizes it.[17] The effect of an increase in the investment horizon is, therefore, *a priori* ambiguous, even if participants form expectations consistently across different investment horizons. Estimations of the forecasting rules reported in Hommes et al. (2005) and supported by later LtF experiments with $H = 1$ and $R = 1.05$ show that the value of θ in the trend-extrapolating heuristic (4) is typically above 0.4 and can be as large as 1. Although it does not resolve all ambiguity, this provides the basis for the following conjecture.[18]

Conjecture 2. *An increase in the investment horizon H leads to an increase in price volatility.*

To verify if this conjecture is correct, we will compare price dynamics across treat-

[17]All three curves intersect for $\theta = \theta^c \approx 0.44$, when the horizon effect reverses. The fundamental value is stable for $\theta < 0.64$ if $H = 1$, for $\theta < 0.58$ if $H = 2$, and for $\theta < 0.56$ if $H = 3$. The fundamental value is stable for $\theta < 0.53$ if $H = 3$ and the feedback is strong as in the **OscH3SF** treatment. All threshold values are computed numerically.

[18]Fig. 2 is obtained assuming that each individual uses the same trend-extrapolating rule (5). However, simulations also support Hypothesis 2 for the case where the extrapolation coefficient for each individual is uniformly distributed around the mean value of $\bar{\theta} > 0.5$. Also note that, applying Jensen's inequality, for a given horizon, to the trend-extrapolating rule (5) we find that the amplification of price trends will be stronger in a population with heterogeneous θ than in the homogeneous case.

ments **ConH1**, **ConH2** and **ConH3**, and across treatments **OscH1**, **OscH2**, and **OscH3**.

As discussed, the ambiguity of the horizon effect arises because the change of the horizon affects the feedback strength. Treatment **OscH3SF** is designed to remedy this effect: it features the same feedback strength as treatment **OscH1**, but the investment horizon is longer. The black solid line in Fig. 2 shows the largest eigenvalue for treatment **OscH3SF**, when $R = 1.016$ and $H = 3$. For small θ, the stability is very similar to treatment **OscH1**, but for large θ the system is more unstable than in the other three cases. Therefore, provided that trend-extrapolating participants take the investment horizon into account when forecasting future prices, we expect to see less stable price dynamics in treatment **OscH3SF** than in treatment **OscH1**. In addition, price dynamics in treatment **OscH3SF** should be less stable than in treatment **OscH3** as well, because feedback strength is higher in **OscH3SF** than in **OscH3**, but the investment horizon is the same.

Conjecture 3. *Prices in treatment **OscH3SF** are more volatile than prices in treatments **OscH1** and **OscH3**.*

If we do not find a significant difference between treatments **OscH3SF** and **OscH1**, this may be because participants do not take into account the longer investment horizon when extrapolating in the former treatment. In fact, both Conjecture 2 and Conjecture 3 are based upon the behavioral assumption that participants extrapolate past prices stronger in their predictions when they face a longer investment horizon. We use this observation to formulate the following conjecture.

Conjecture 4. *Forecasts of participants in treatments with longer investment horizons are characterized by a stronger degree of trend extrapolation.*

Earlier LtF asset pricing experiments are characterized by strong coordination of expectations on a common trend-extrapolating prediction strategy. Strong coordination further contributes to the endogenous emergence of bubbles. Because in our experiment participants have to predict further ahead in treatments with $H > 1$, forming predictions there might be cognitively or computationally more demanding. This may lead to a larger heterogeneity across individual forecasts, impeding coordination and the emergence of bubbles. This observation gives rise to our final conjecture.

Conjecture 5. *Coordination of expectations in treatments with longer investment horizons will be lower than in the treatments with shorter investment horizons.*

We will the validity of this last conjecture by comparing the dispersion of forecasts across treatments **ConH1**, **ConH2** and **ConH3**, and across treatments **OscH1**, **OscH2**, and **OscH3**.

3 Price dynamics, investment horizons and price histories: Experimental results

We start the discussion of our experimental results with analyzing the impact of the initial price history and the length of the investment horizon on market price volatility in Section 3.1. To understand the findings presented there, we subsequently turn our attention to forecasting behavior in Section 3.2, and discuss the effect of the investment horizon on the coordination of expectations and on participants' tendency to extrapolate trends in prices. For our analysis, we restrict attention to the incentivized periods, see Table 1. Thus we have observations on prices and predictions for fifty periods in each group.[19]

3.1 Aggregate market dynamics

We have 35 groups in our experiment in total, divided over seven treatments, see Table 1. Fig. 3 shows the evolution of prices for each of these groups, organized by treatment.[20] The left panels of Fig. 3 present market prices for the groups in the **ConH1**, **ConH2** and **ConH3** treatments, and the right panels present market prices for the groups in the **OscH1**, **OscH2**, **OscH3** and **OscH3SF** treatments, respectively. The first ten periods in each panel show the initial price history (in black). This initial history is the same for all groups in the same treatment but differs between the **Con** and **Osc** treatments, as explained in Section 2.2. To

[19]As discussed in Section 2.1, the LtF experiments are vulnerable to forecasts that are seemingly unrelated to the price dynamics. Such outliers were also present in our experiment, which forced us to exclude some data from the analysis. First, nine of the last ten forecasts of participant 5 in group 1 of treatment **OscH3SF** are equal to either 0, 999 or 1000. These extreme predictions have a large impact on the realized market price in periods $54-63$ in that group. We excluded those periods in that group from the analysis. Second, we excluded two groups from the analysis altogether. In a group from treatment **ConH1**, all predictions of one participant are much lower (5.0 on average and never higher than 15.4) than both the initial prices (all ten of which are at least 56.0) or the realized prices (which are 23.8 on average). In a group from treatment **OscH2**, the predictions of one participant start with 60.59 for period 13, then monotonically decrease to 2.15 for period 32, after which they monotonically increase up to 102.85 for period 64. In both cases, these peculiar forecasting rules have an important effect on the entire price dynamics, which is why we excluded these two groups. Note, however, that our results are qualitatively robust with respect to including these outliers (see the Online Appendix).

[20]See Appendix D for the evolution of prices and predictions for each group separately.

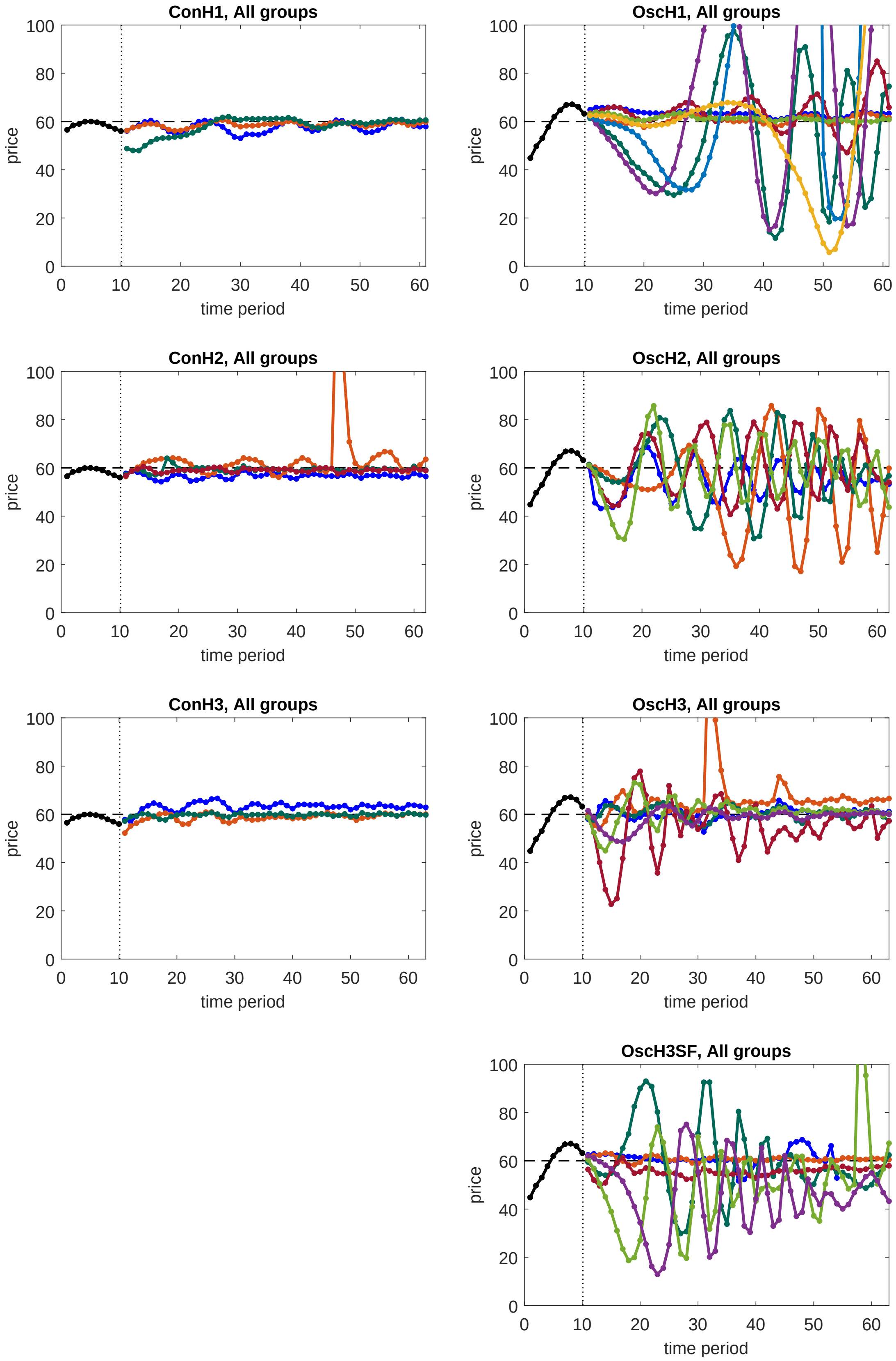

Figure 3: Price dynamics in all experimental treatments. The initial price history for periods 1-10 is shown in black. Experimental price dynamics in different groups is shown with different colours. The constant fundamental price of 60 is shown by the dashed line.

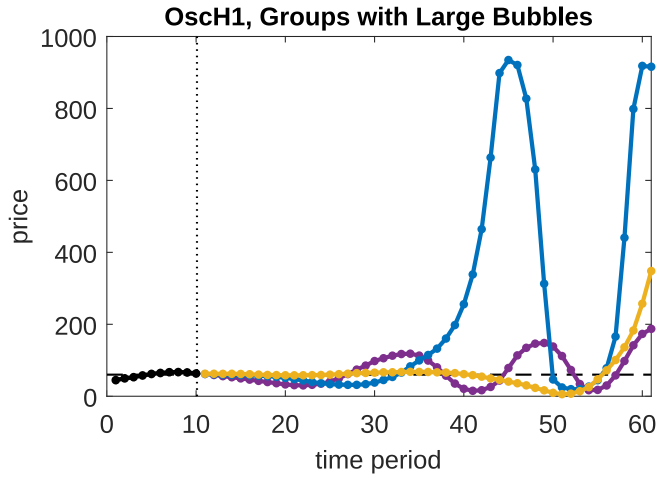

Figure 4: Price dynamics in groups 6 (purple), 7 (blue) and 8 (yellow) of the **OscH1** treatment.

facilitate the comparison, all panels in Fig. 3 have the same vertical scaling, from 0 to 100. However, in three out of the 35 groups prices systematically go above 100. These three groups are all from treatment **OscH1**, and prices in these groups (groups 6, 7 and 8) are shown separately in Fig. 4 with a larger vertical scale.[21]

We make several observations based on Figs. 3 and 4. First, there is an apparent effect of the initial price history on market price dynamics. Price oscillations in all 10 groups that start with the converging initial price history are limited, whereas more than half of the groups that start with the oscillating initial price history lead to substantial and persistent oscillations. In addition, the treatments with $H = 1$ can be compared to earlier LtF experiments. As discussed above, these experiments – where no initial price history was given to participants – often lead to large bubbles in asset prices with prices rising to more than 10 times the fundamental value.[22] In our experiment, such a large bubble only emerges in one out of the eight groups (group 7) in treatment **OscH1** and in none of the three groups in **ConH1**. This observation suggests that, compared to giving no price history at all, the oscillating initial price history also tends to reduce the incidence of large bubbles substantially. That may happen because in the last

[21]There are two other instances where the price is above 100 for only one period. These are period 32 in group 2 of treatment **OscH3** and period 58 in group 5 of treatment **OscH3SF**, when the price suddenly jumps to 145 and 134, respectively, and drops below 100 again in the subsequent period. In both cases, this is due to a sudden extreme prediction by one of the participants. The reason might be an accidental typo by that participant. For example, in the first case a participant predicts 647 instead of 64.7, where the latter would be very much in line with the last observed price and the participant's preceding forecasts.

[22]For example, this is the case for five of the six groups in Hommes et al. (2008) whose design, apart from the initial price history, is essentially equivalent to our treatments with $H = 1$.

couple of periods of this oscillating initial price history prices decrease again. This effect alone is sufficient to inhibit the emergence of large bubbles, although the oscillating initial prices history does tend to increase price volatility.

The effect of the investment horizon on price volatility is more difficult to distill from Fig. 3. Price dynamics in all groups from the **Con** treatments are relatively stable with no visible effect of the investment horizon on price volatility (which is low for all groups in these treatments). For the **Osc** treatments, the general picture is somewhat blurred by the heterogeneity between outcomes in the same treatment. For example, three of the groups in treatment **OscH1** are characterized by relatively stable price dynamics. In contrast, the other five groups either show persistent oscillations from the start or feature such oscillations in the second half of the experiment. Similarly, one of the six groups in treatment **OscH3** and three of the six groups in treatment **OscH3SF** show persistent oscillations, with price dynamics in the other groups in those treatments relatively stable. Only in treatment **OscH2**, dynamics across the different groups are relatively homogeneous with persistent price fluctuations in all five groups of that treatment.

Notwithstanding this heterogeneity, the **Osc** treatments do not seem to provide much evidence in support of a positive effect of the investment horizon on price volatility. Although the fraction of groups that are characterized by persistent fluctuations goes up from 5 out of 8 to 5 out of 5, when going from treatment **OscH1** to treatment **OscH2**, it falls to 1 out of 6 in treatment **OscH3**. Also, the maximal amplitude of the price oscillations is different between treatments. In three groups in treatment **OscH1**, illustrated in Fig. 4, prices eventually grow much larger than 100, whereas prices in treatments **OscH2**, **OscH3** and **OscH3SF** consistently stay below 100, even in groups that exhibit persistent oscillations. All this points towards the conclusion that increasing the time horizon impedes the emergence of large bubbles. Finally, when comparing treatment **OscH3SF** – which compensates for the decrease in feedback strength – with treatment **OscH1**, the former does not appear to lead to higher price volatility, with only 3 out of 6 groups showing persistent fluctuations; these fluctuations are much smaller than the largest fluctuations in treatment **OscH1**. Therefore, an increase in the investment horizon seems more likely to *decrease* price volatility, rather than to increase it.[23]

[23]The following observation further supports this conclusion. Bao and Hommes (2019) vary feedback strengths in an LtF experiment with $H = 1$. Their treatment L has a feedback strength 0.86, coinciding with the feedback strength in our treatment **OscH3**. However, where we only find persistent oscillations in 1 out of 6 groups in treatment **OscH3**, Bao and Hommes (2019) find persistent oscillations in all 5 groups in their treatment L. Note that this difference may also partly be attributed to the absence of an initial price history in Bao and Hommes (2019).

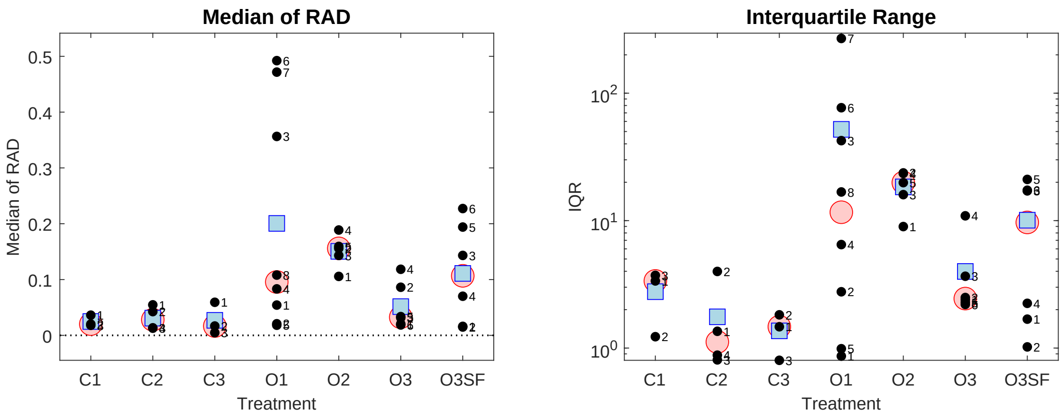

Figure 5: *Left:* Level of mispricing measured as the median of the relative absolute deviation. *Right:* Level of volatility measured as the Interquartile Range (log scale). For both panels, the data are organized by treatments, where the statistics for individual groups (black dots), the median for the treatment (red circle), and the average for the treatment (blue square) are shown.

To corroborate these impressions, we consider two measures that quantify price volatility. The first measure is the *median RAD*, i.e., the median (over all incentivized periods in a group) of the relative absolute deviation of price from the fundamental price, $|p_t - p^f|/p^f$. The RAD is one of the measures proposed by Stöckl et al. (2010) to study bubbles in asset market experiments.[24] Our second measure is the *interquartile range* (IQR), i.e., the length of the interval that contains the middle 50% of the prices from the given group. Fig. 5 depicts both measures for all 35 groups in the experiment, as well as the mean (blue squares) and median (red circles) of these measures for every treatment. The numeric values of these and other descriptive statistics can be found in Table 2 in Appendix C.1.

The quantitative measures depicted in Fig. 5 are consistent with our earlier observations. All groups from the **Con** treatments have low values of both measures. In contrast, a substantial number of groups from the **Osc** treatments have much higher values of these measures (note the log-scale of the right panel for the IQR). At the same time, there are some groups from these four treatments for which the measures are comparable to the values from the **Con** treatments. We perform pairwise tests of similarity of both measures between the treatments. The p-values of three one-sided non-parametric tests (the Kolmogorov-Smirnov, Fischer-Pitman and Mann-Whitney-Wilcoxon tests) are reported in Appendices C.2 and C.3 for the median RAD and IQR, respectively. All tests show that there

[24] Instead of looking at the mean value of the relative absolute deviations, as in Kirchler et al. (2012), for example, we look at the median value. This limits the effect of outliers caused by extreme predictions.

is a significant difference, at the 1% level, for both measures, between treatment **OscH2** and treatment **ConH2**. The other relevant comparisons (i.e., keeping the horizon the same) are not as strong but point in the same direction. For the comparison of treatments **ConH1** and **OscH1**, the difference in median RAD is significant for two tests at the 5% level (and at the 10% level for the third test), but the difference in IQR is not statistically significant. For the comparison of treatments **ConH3** and **OscH3**, the difference in median RAD is significant for one test only (at the 10% level), while for all three tests the difference in IQR is significant, at least at the 5% level. The heterogeneity in outcomes in treatment **OscH1** and, to a lesser extent, in treatment **OscH3**, which can be seen in Fig. 5, may explain the lack of strongly significant differences in outcomes. This heterogeneity is small in treatment **OscH2**, which is precisely the treatment for which the most convincing results are obtained.

Based upon our discussion and statistical tests, we conclude that we find evidence that supports Conjecture 1. In particular, we find support for the following result.

Result 1. *The initial price history plays an important role in subsequent price dynamics. Price volatility is higher for treatments with the oscillating initial price history than for treatments with the converging initial price history.*

Fig. 5 confirms another observation. There seems to be only a small effect of an increase in the investment horizon in the **Con** treatments, and there is no clear support for an increase in volatility with the investment horizon in the **Osc** treatments. In fact, for treatments **OscH1**, **OscH2** and **OscH3** the average values of both the mean RAD and the IQR are *lower* for longer horizons, and the same holds when comparing treatments **OscH1** and **OscH3SF**. From the statistical tests presented in Appendices C.2 and C.3, we find that the only statistically significant difference is that volatility is *lower* in **OscH3** than in **OscH2** (for both measures and all three tests, at the 1% level). Also the differences between **OscH3SF** and treatments **OscH1** and **OscH3** are not statistically significant with only one exception for each measure, with significance at the 10% level. We therefore conclude that both Conjectures 2 and 3 do not have empirical support, which brings us to our second result.

Result 2. *Price volatility does not increase for longer investment horizons, even after correcting for differences in feedback strength. In fact, it tends to decrease.*

We take one more look at prices by characterizing price dynamics in the following way. The simplest linear model that might give rise to oscillations in prices is the following AR(2) specification

$$p_t = \beta_0 + \beta_1 p_{t-1} + \beta_2 p_{t-2} + \nu_t \,. \tag{7}$$

Depending on coefficients β_1 and β_2, the dynamics in such a model can be converging or diverging. This dependence is captured by the triangles[25] in (β_1, β_2) coordinates, plotted in Fig. 6. In particular, the dynamics are converging for the coefficient combinations inside the triangles, diverging for the combinations outside the triangles, and oscillatory for combinations (β_1, β_2) below the parabola.

We fitted the AR(2) model (7) on the prices for each experimental group. The estimation results can be found in Table 9 in Appendix C.4. The price dynamics in 26 of the 35 markets can be described well by (7), in the sense that the estimated rules pass two specification tests (on autocorrelation and heteroskedasticity) for residuals. We superimposed the estimated values $(\hat{\beta}_1, \hat{\beta}_2)$ for each experimental group on the triangles of Fig. 6. Note that almost all estimated pairs lie in the region where the dynamics are characterized by converging oscillations. The only two exceptions are groups 4 and 6 in treatment **OscH1**, where oscillations are diverging. The regularities we identified before are nicely represented by the plots. Estimations in most groups from the **Con** treatments are relatively far away from the instability boundary $\beta_2 = -1$. In the **Osc** treatments, the estimated combinations tend to become more stable (i.e., they lie further away from that boundary) when the investment horizon increases.

3.2 Trend extrapolation and coordination of expectations

The results from the previous section suggest that there is a tendency for price volatility to decrease when the investment horizon increases. In fact, large bubbles with prices eventually growing to a level that is a multiple of the fundamental value only appear in our experiment for an investment horizon of $H = 1$. As discussed in Section 2.3 the prevalence and size of bubbles in earlier LtF experiments have been attributed to participants in the same group coordinating their forecasts on a trend-extrapolating strategy. The absence of large bubbles and the tendency to have fewer groups with persistent asset price fluctuations for longer investment horizons in our experiment might therefore be caused by either a failure of participants to coordinate their expectations, or a failure to coordinate on trend-extrapolating strategies, or a combination of both. In this section we will study participants' individual forecasts to better understand how an increase in

[25] The edges of the triangles are given by $\beta_2 = 1 - \beta_1$, $\beta_2 = 1 + \beta_1$ and $\beta_2 = -1$.

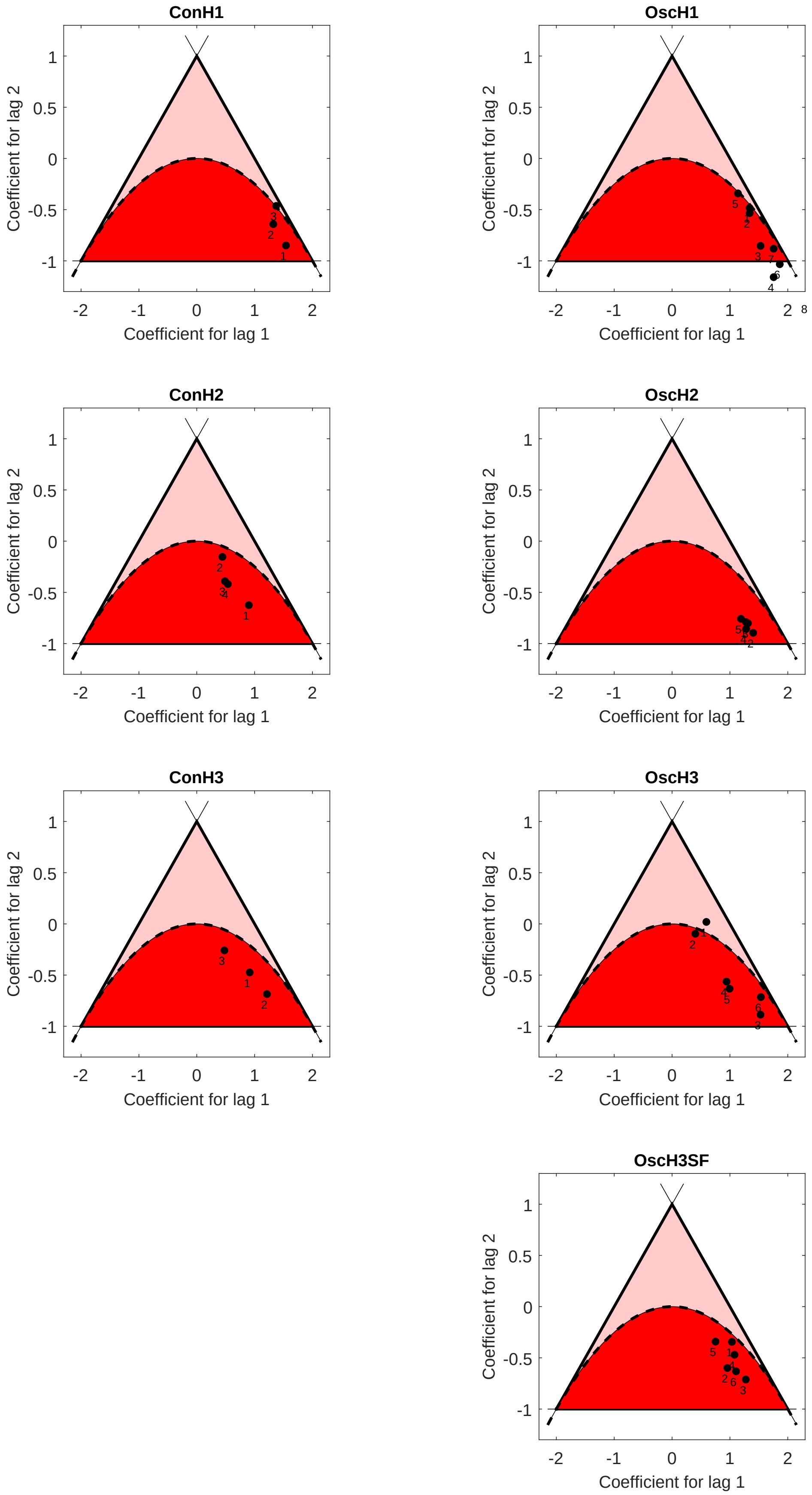

Figure 6: Stability triangles for the AR(2) model and estimated AR(2) rules.

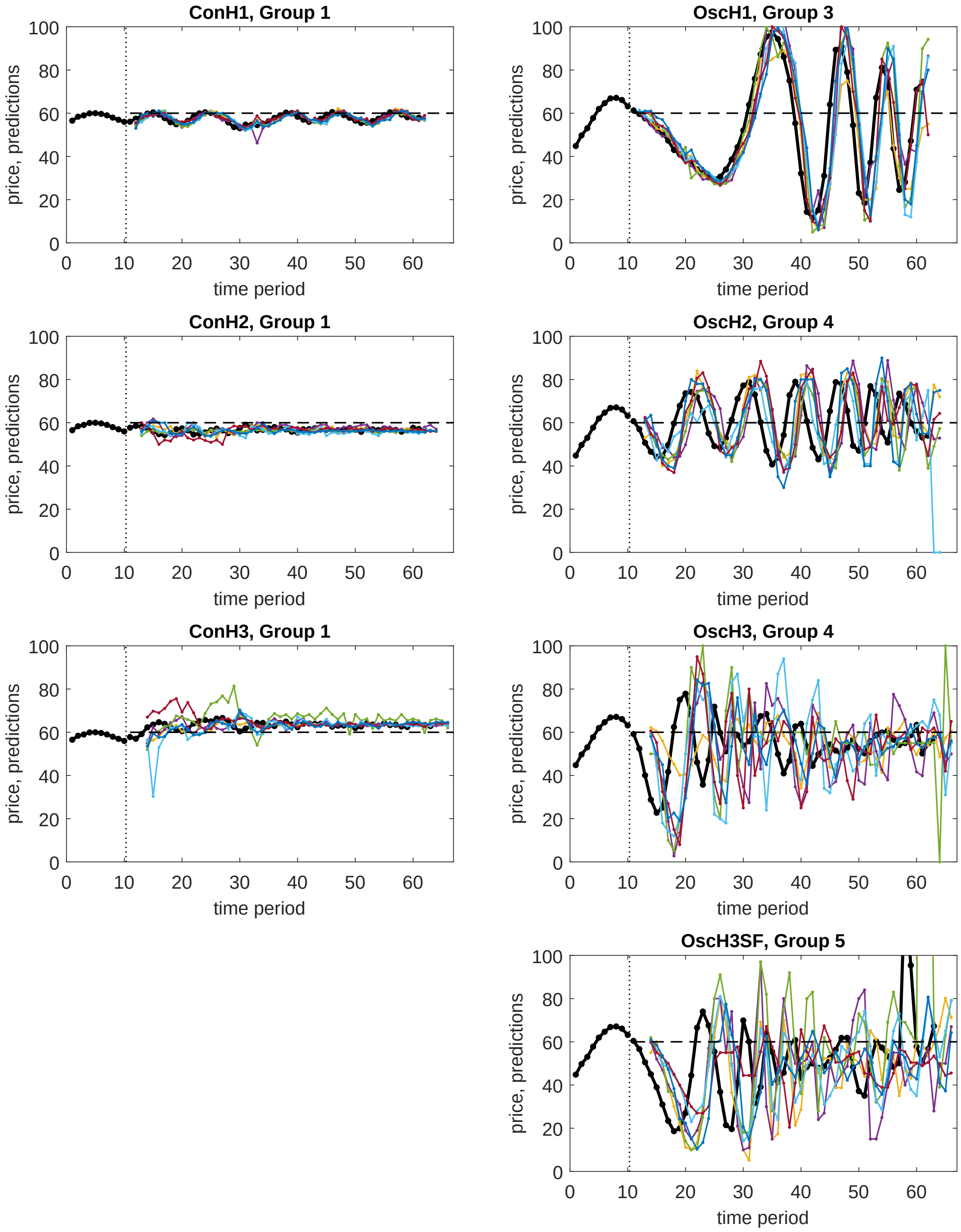

Figure 7: Examples of individual forecast dynamics. Price is shown by the thick black line. The constant fundamental price of 60 is shown by the dashed line.

the investment horizon affects their tendency to coordinate on trend-extrapolating strategies. For an illustration of the individual predictions see Fig. 7, which shows, for one group from each treatment, both the market price (in black) and participants' individual predictions (in color). See Appendix D for plots for all groups.

Recall from Section 2.3 that, if participants believe prices evolve according to

$p_t = p_{t-1} + \theta(p_{t-1} - p_{t-2})$, their predictions should be given by Eq. (5), implying

$$p^e_{i,t+H} - p_{t-1} = \theta\left(1 + \theta + \cdots + \theta^H\right)(p_{t-1} - p_{t-2}) \,, \tag{8}$$

for investment horizon H. Therefore, the effect of the last observed price change, $p_{t-1} - p_{t-2}$, on the expected price increase, $p^e_{i,t+H} - p_{t-1}$, should be stronger for longer investment horizons H.

To investigate this, Fig. 8 shows scatter plots where, for each of the seven treatments, the averages of participants' expected price changes, $\bar{p}^e_{t+H} - p_{t-1}$, are plotted against the last observed price change, $p_{t-1} - p_{t-2}$. The red line in each panel represents a linear regression line, the slope of which is shown in the title of the panel. According to (8), the slope should increase with the investment horizon, but it actually *decreases* substantially.[26] Therefore, participants not only insufficiently account for the investment horizon when extrapolating trends, but their tendency to extrapolate these trends diminishes when the investment horizon increases. We conclude the following:

Result 3. *An increase in the investment horizon leads to a substantial decrease in the extent to which trends in past prices are extrapolated by participants.*

Result 3, together with the effect of horizons on the feedback strength in the pricing equation, provides an explanation for why price volatility tends to go down for longer investment horizons. A related, but separate, question is whether participants' ability to coordinate on a common prediction strategy is affected by the investment horizon as well.

From Fig. 7, it appears that predictions within groups are well-coordinated, although there are some differences between treatments. For example, coordination seems to be stronger in group 3 of treatment **OscH1** than in group 4 of treatment **OscH2** and in group 4 of treatment **OscH3**. To quantify coordination of expectations, we compute, for each group in our experiment, the standard deviation of the six predictions in each period and take the median of those 50 standard deviations. We will refer to this quantity as the *discoordination* measure. It is depicted for all groups, and organized by treatment, in Fig. 9. The numeric values can be found in Table 2 from Appendix C.1.

[26] For example, the slope changes from 0.81 in treatment **OscH1** to 0.64 in treatment **OscH2** to 0.01 in treatment **OscH3**. The slope for treatment **OscH1** is significantly higher (at the 5% level) than the slopes in treatments **OscH2**, **OscH3** and **OscH3SF**, see Appendix C.5. The slope of 0.81 for $H = 1$ is consistent with a value $\tilde{\theta} \approx 0.53$ for the extrapolation coefficient. This would imply slopes $\tilde{\theta}(1 + \tilde{\theta} + \tilde{\theta}^2) \approx 0.97$ and $\tilde{\theta}(1 + \tilde{\theta} + \tilde{\theta}^2 + \tilde{\theta}^3) \approx 1.05$ for treatments **OscH2** and **OscH3**, respectively, for forecasting behavior to be consistent between treatments.

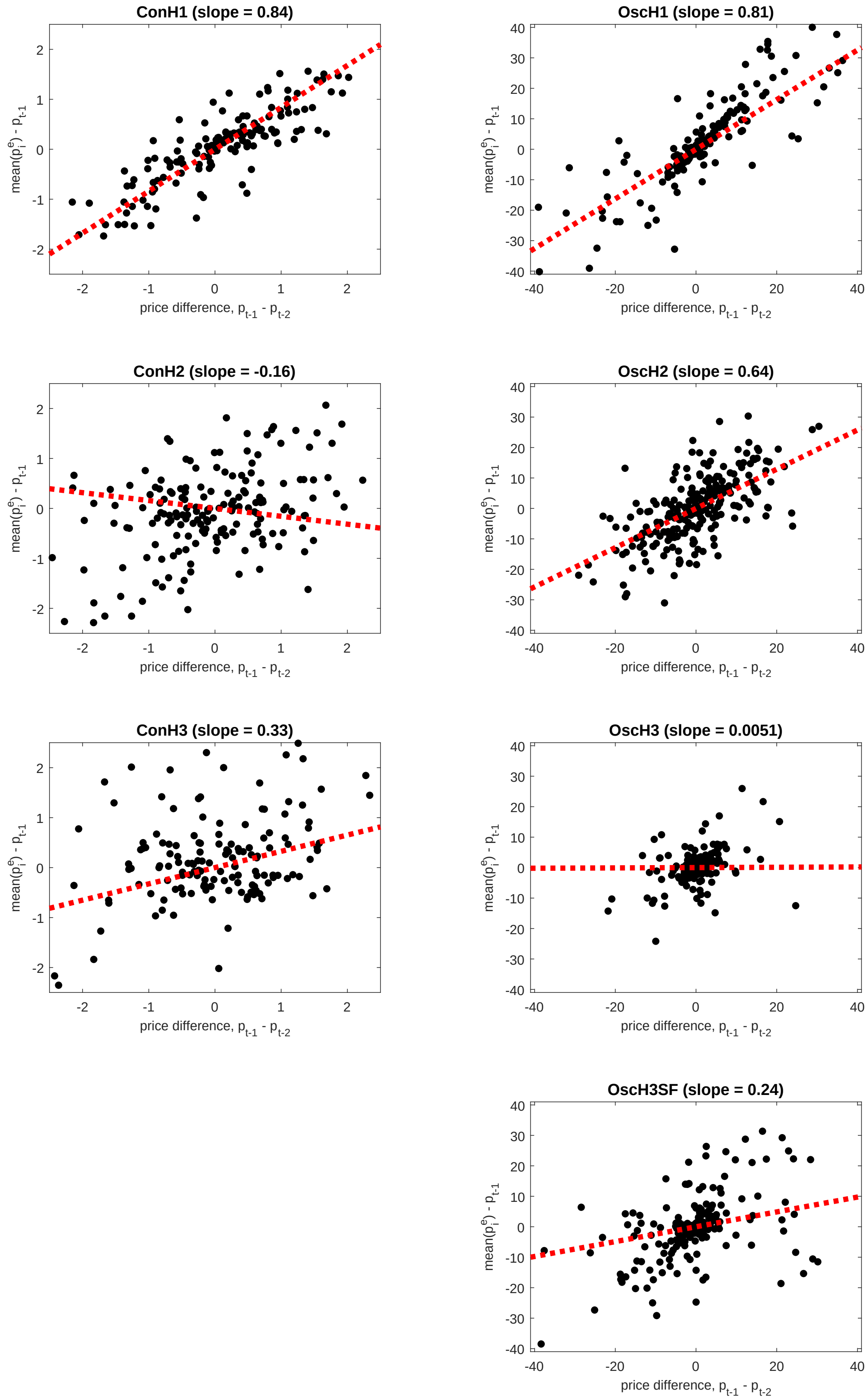

Figure 8: Scatter plots of $\bar{p}^e_{i,t+h} - p_{t-1}$ and the most recent price change, $p_{t-1} - p_{t-2}$.

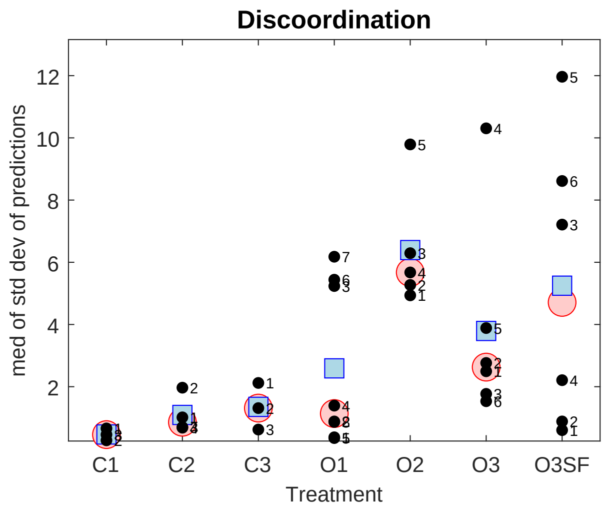

Figure 9: Level of discoordination of predictions in all groups measured as the median of the standard deviation of predictions. For each treatments, the statistics for individual groups (black dots), the median for the treatment (red circle), and the average for the treatment (blue square) are shown.

We see substantial heterogeneity in this discoordination measure between groups in the same treatment. Given the heterogeneity in price volatility within treatments, this is hardly surprising: if prices converge to the fundamental value, then typically all forecasts also converge to that value, which reduces the dispersion of the forecasts considerably. However, when restricting attention to the groups with a higher price volatility, there seems to be a tendency for an increase in the investment horizon to *adversely* affect the level of coordination. Although one might expect that the dispersion of forecasts is highest in large bubbles, the five groups with the highest dispersion in forecasts are all from treatments with $H = 2$ or $H = 3$, whereas the largest bubbles emerge in treatment **OscH1**. The left panel of Fig. 10 explores the relationship between discoordination and price volatility further. It plots, for each group, the discoordination measure against the median RAD (which measures the price volatility as discussed in Section 3.1). The plot suggests that coordination is easier if the investment horizon is shorter. To see this, note that for groups from different treatments with a similar price volatility, the dispersion of predictions is higher if the investment horizon is longer. Remarkably, the groups with the highest price volatility (groups 3, 6 and 7 from treatment **OscH1**) are characterized by a smaller dispersion of forecasts than groups with lower price volatility but a longer investment horizon.

Figs. 9 and 10 suggest that the ability of participants to coordinate their forecast on a common prediction strategy decreases when the investment horizon in-

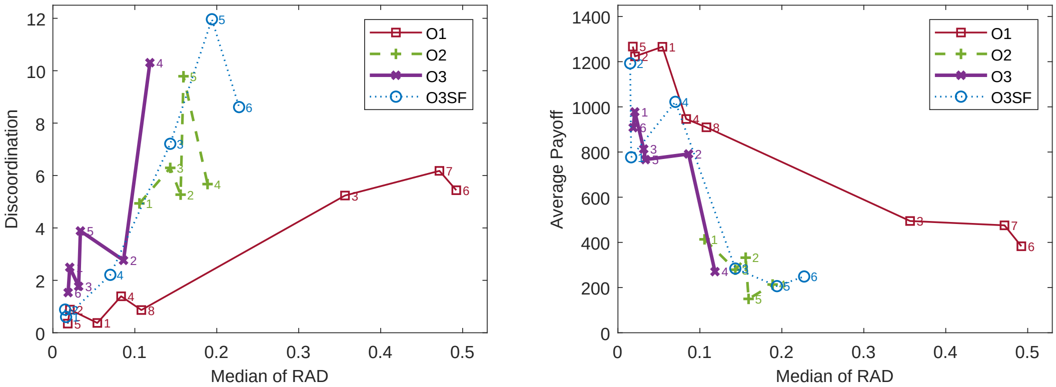

Figure 10: *Left:* Median of the standard deviation of predictions plotted against median of RAD. *Right:* Average payoff plotted against median of RAD. The data from the **OscH1**, **OscH2**, **OscH3** and **OscH3SF** treatments are used.

creases. Due to the heterogeneity of outcomes within treatments, the differences between treatments are not always statistically significant: the three pairwise differences between **OscH1** and **OscH2** are statistically significant, but most other comparisons between **Osc** treatments are not (see the p-values of the three tests in Appendix C.6). We conclude the following:

Result 4. *The extent to which participants coordinate their expectations decreases with the investment horizon.*

The fact that coordination of expectations becomes more difficult may be responsible for the fact that large bubbles do not emerge for longer investment horizons. Note that along such a large bubble, participants are still able to do relatively well if they coordinate their forecasts. As long as a participant's forecasts are close to the average forecast and thereby relatively close to the market price, forecast errors remain reasonably small (although they tend to be smaller when prices are closer to the fundamental value). Participants are, therefore, more willing to "ride the bubble" if they can coordinate their expectations. If dispersion of the price forecasts is large, at least some participants will make large forecast errors, and low payoffs, along such a bubble. Those participants are then likely to change their forecasting behavior, and this makes large bubbles unsustainable. The right panel of Fig. 10 supports this explanation for the absence of large bubbles for longer investment horizons. The panel plots the average payoff (the number of earned points, averaged over periods and participants in a group) versus price volatility in that group (represented by the median RAD again). It shows that, for groups that have approximately the same level of price volatility but are from different treatments, the average number of points earned by participants is indeed lower if the investment horizon is longer than one period. Participants in the groups with large bubbles in treatment **OscH1** still do better than participants in groups from treatments **OscH2**, **OscH3** and **OscH3SF** that experience persistent, but much smaller, fluctuations in asset prices.

Large bubbles can develop because of the self-confirming nature of financial markets. If all traders believe the price of an asset will go up, many traders will buy the asset, and the price will indeed go up, because of the increase in demand. For such an expectation-induced asset price bubble to emerge and be sustainable it is, however, necessary that there is a sufficient degree of agreement between traders about the price development of the asset. From the analysis of the individual forecasts in our experiment, we learn that this agreement is difficult to maintain when the investment horizon increases.

4 Conclusion

In this paper we discussed a Learning to Forecast laboratory experiment to investigate the effect of the investment horizon on asset price dynamics. In previous LtF asset pricing experiments, characterized by short investment horizons and the absence of an initial price history, large bubbles and crashes often emerge, with prices regularly exceeding their fundamental values by one order of magnitude, see, e.g., Hommes et al. (2008, 2021). We only observe such bubbles for an investment horizon of one period and even then they are unlikely to occur if the market has a history of relatively stable prices, which suggests that a short investment horizon is a necessary (but not sufficient) condition for the emergence of bubbles. An increase in the investment horizon tends to lead to smaller fluctuations and lower price volatility.

The large bubbles that emerge for a short investment horizon seem to be driven by the tendency of participants to coordinate on trend-extrapolating prediction strategies that reinforce and amplify initial price fluctuations. If the investment horizon increases, both the extent to which participants extrapolate price trends, and participants' success in coordinating their predictions, decrease and as a consequence price volatility goes down.[27]

As mentioned in the Introduction, in the existing behavioral finance literature, deviations from fundamental values have often been associated explicitly

[27]Remarkably, participants' expectations, therefore, are consistent with the empirical finding that asset returns show momentum in the short run, but revert to fundamental values in the longer run, see, e.g., Hong and Stein (1999) for a discussion.

with short-horizon speculative trading. Note, however, that our results do not require limits to arbitrage for participants with shorter investment horizons, as in De Long et al. (1990) or Dow and Gorton (1994). In particular, the information asymmetries between traders assumed in these papers and responsible for creating arbitrage opportunities, are absent in our experimental framework. Combined with the observation that many decisions of institutional investors are driven by short-term gains, possibly at the expense of earnings in the long-run,[28] several measures have been suggested to curb this so-called “short-termism”, see, e.g., Bolton and Samama (2013), and Crouzet et al. (2020). Our results suggest that policy measures that inhibit the possibility of short-horizon trading, e.g., by imposing a minimum holding period for the asset, may indeed help in reducing price volatility. However, such measures might also decrease the speed with which private information about the asset is reflected in its price (note that, because information asymmetries are absent, this does not play a role in our experiment). Moreover, from our findings on the effect of initial price histories it follows that the impact of such measures may be mitigated by the historical performance of the market.

There are several avenues for extending our research. First, in our stylized experiment we only consider investment horizons of up to three periods, and we assume that all investors active in the same market have the same investment horizon. An obvious extension would be to consider longer horizons (and possibly compensate for the implied reduction in feedback strength, as we did in treatment **OscH3SF**) or to consider markets composed of participants with different investment horizons. Because both extensions are likely to increase the dispersion of expectations even further (and the first extension is likely to reduce the level of trend extrapolation) we expect a similar (but stronger) effect of the investment horizon. A second, potentially more interesting, extension is to consider mixed investment horizons, and make the impact that participants with different investment horizons have on the market-clearing price depend on their relative performance. If participants with a shorter time horizon are more accurate forecasters, and therefore attract more investors, their impact on the market-clearing price increases. This means that – even in an environment with mixed horizons – large bubbles may still emerge, making the use of short investment horizons viable.[29]

[28]See Bushee (2001), for empirical evidence, and Bolton et al. (2006), for a theoretical explanation.

[29]The impact of relative performance can, for example, be formalized by letting the weights on individual expectations in pricing equation (2) depend upon the forecast accuracy of the participants (as in Kopányi et al., 2019). In addition, note that performance rankings play an important role in risk-taking behavior of financial professionals (see e.g., Kirchler et al., 2020), which might contribute further to price volatility.

Another relevant extension concerns the initial price history. In treatment **OscH1**, a relatively small fraction of groups experiences large bubbles, when compared to the results in Hommes et al. (2008, 2021), which use an experimental design that is, apart from the initial price history, virtually the same as that of treatment **OscH1**. So, even though the initial price history in treatment **OscH1** features oscillations, it still seems to induce more stability than not providing an initial history of prices at all. The question, therefore, is whether a treatment without any initial price history, or with an initial price history that features monotonically increasing prices that already go (far) beyond the fundamental value might be able to give rise to large bubbles and crashes, even for longer investment horizons.

References

Anufriev, M., Hommes, C., 2012a. Evolution of market heuristics. Knowledge Engineering Review 27, 255–271.

Anufriev, M., Hommes, C., 2012b. Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. American Economic Journal: Microeconomics 4, 35–64.

Anufriev, M., Hommes, C., Makarewicz, T., 2019. Simple forecasting heuristics that make us smart: Evidence from different market experiments. Journal of the European Economic Association 17, 1538–1584.

Arifovic, J., Petersen, L., 2017. Stabilizing expectations at the zero lower bound: Experimental evidence. Journal of Economic Dynamics & Control 82, 21–43.

Bao, T., Duffy, J., Hommes, C., 2013. Learning, forecasting and optimizing: An experimental study. European Economic Review 61, 186–204.

Bao, T., Hommes, C., 2019. When speculators meet suppliers: Positive versus negative feedback in experimental housing markets. Journal of Economic Dynamics & Control 107, 103730.

Bao, T., Hommes, C., Makarewicz, T., 2017. Bubble formation and (in) efficient markets in learning-to-forecast and optimise experiments. The Economic Journal 127, F581–F609.

Bolton, P., Samama, F., 2013. Loyalty-shares: Rewarding long-term investors. Journal of Applied Corporate Finance 25, 86–97.

Bolton, P., Scheinkman, J., Xiong, W., 2006. Executive compensation and short-termist behaviour in speculative markets. The Review of Economic Studies 73, 577–610.

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

Bushee, B.J., 2001. Do institutional investors prefer near-term earnings over long-run value? Contemporary Accounting Research 18, 207–246.

Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. The econometrics of financial markets. Princeton University press.

Cella, C., Ellul, A., Giannetti, M., 2013. Investors' horizons and the amplification of market shocks. The Review of Financial Studies 26, 1607–1648.

Colasante, A., Alfarano, S., Camacho-Cuena, E., Gallegati, M., 2020. Long-run expectations in a learning-to-forecast experiment: a simulation approach. Journal of Evolutionary Economics 30, 75–116.

Cremers, M., Pareek, A., 2015. Short-term trading and stock return anomalies: Momentum, reversal, and share issuance. Review of Finance 19, 1649–1701.

Crouzet, N., Dew-Becker, I., Nathanson, C.G., 2020. On the effects of restricting short-term investment. The Review of Financial Studies 33, 1–43.

De Long, J.B., Shleifer, A., Summers, L.H., Waldmann, R.J., 1990. Noise trader risk in financial markets. Journal of Political Economy 98, 703–738.

Dow, J., Gorton, G., 1994. Arbitrage chains. The Journal of Finance 49, 819–849.

Evans, G.W., Hommes, C.H., McGough, B., Salle, I., 2019. Are long-horizon expectations (de-) stabilizing? Theory and experiments. Technical Report. Bank of Canada Staff Working Paper.

Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics 10, 171–178.

Greiner, B., 2015. Subject pool recruitment procedures: Organizing experiments with ORSEE. Journal of the Economic Science Association 1, 114–125.

Hennequin, M., 2019. Experiences and expectations in asset markets: an experimental study. Working Paper, University of Amsterdam.

Hertwig, R., Barron, G., Weber, E.U., Erev, I., 2004. Decisions from experience and the effect of rare events in risky choice. Psychological Science 15, 534–539.

Hirota, S., Huber, J., Stock, T., Sunder, S., 2020. Speculation, money supply and price indeterminacy in financial markets: An experimental study. Journal of Economic Behavior and Organization, doi:10.1016/j.jebo.2020.06.010. forthcoming.

Hirota, S., Sunder, S., 2007. Price bubbles sans dividend anchors: Evidence from laboratory stock markets. Journal of Economic Dynamics & Control 31, 1875–1909.

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

Hommes, C., Kopányi-Peuker, A., Sonnemans, J., 2021. Bubbles, crashes and information contagion in large-group asset market experiments. Experimental Economics 24, 414–433.

Hommes, C., Sonnemans, J., Tuinstra, J., Van de Velden, H., 2005. Coordination of expectations in asset pricing experiments. Review of Financial Studies 18, 955–980.

Hommes, C., Sonnemans, J., Tuinstra, J., Van de Velden, H., 2008. Expectations and bubbles in asset pricing experiments. Journal of Economic Behavior & Organization 67, 116–133.

Hong, H., Stein, J.C., 1999. A unified theory of underreaction, momentum trading, and overreaction in asset markets. The Journal of Finance 54, 2143–2184.

Kirchler, M., Huber, J., Stöckl, T., 2012. Thar she bursts: Reducing confusion reduces bubbles. American Economic Review 102, 865–83.

Kirchler, M., Lindner, F., Weitzel, U., 2020. Delegated investment decisions and rankings. Journal of Banking & Finance 120, 105952.

Kopányi, D., Rabanal, J.P., Rud, O.A., Tuinstra, J., 2019. Can competition between forecasters stabilize asset prices in learning to forecast experiments? Journal of Economic Dynamics & Control 109, 103770.

Kopányi-Peuker, A., Weber, M., 2020. Experience does not eliminate bubbles: Experimental evidence. Review of Financial Studies, DOI: https://doi.org/10.1093/rfs/hhaa121, forthcoming.

Malmendier, U., Nagel, S., 2011. Depression babies: Do macroeconomic experiences affect risk taking? Quarterly Journal of Economics 126, 373–416.

Malmendier, U., Nagel, S., 2016. Learning from inflation experiences. Quarterly Journal of Economics 131, 53–87.

Marimon, R., Spear, S.E., Sunder, S., 1993. Expectationally driven market volatility: An experimental study. Journal of Economic Theory 61, 74–103.

Palan, S., 2013. A review of bubbles and crashes in experimental asset markets. Journal of Economic Surveys 27, 570–588.

Patton, A.J., Timmermann, A., 2010. Why do forecasters disagree? Lessons from the term structure of cross-sectional dispersion. Journal of Monetary Economics 57, 803–820.

Razen, M., Huber, J., Kirchler, M., 2017. Cash inflow and trading horizon in asset markets. European Economic Review 92, 359–384.

Sonnemans, J., Tuinstra, J., 2010. Positive expectations feedback experiments and number guessing games as models of financial markets. Journal of Economic Psychology 31, 964–984.

Stöckl, T., Huber, J., Kirchler, M., 2010. Bubble measures in experimental asset markets. Experimental Economics 13, 284–298.

APPENDIX

A An asset pricing model with different investment horizons

A.1 A generalization of the standard asset pricing model

Following Brock and Hommes (1998), we consider a market populated by a number of investors, e.g., institutional investors, such as pension funds. These investors maximize their wealth by investing in two assets, a risk-free asset and a risky asset. The supply of the risk-free asset is infinitely elastic, and this asset gives a fixed net return of r per period. We denote gross return of this risk-free asset by $R = 1 + r$. The risky asset pays stochastic dividend y_t in period t and its price p_t, which may vary over time, follows from temporary equilibrium between demand and supply in period t.

We extend the model in Brock and Hommes (1998) by allowing investors to differ in their investment horizons. Let us denote the wealth of investor i in period t by $W_{t,i}$. Suppose this investor has investment horizon h, i.e., a portfolio should maximize wealth in period $t + h$. Investor i's wealth in period $t + h$ is given by

$$\begin{aligned} W_{t+h,i} &= (1 - x_{t,i})W_{t,i}R^h + x_{t,i}\frac{W_{t,i}}{p_t}\left(p_{t+h} + \sum_{s=1}^{h} R^{h-s}y_{t+s}\right) \\ &= W_{t,i}R^h + \left(p_{t+h} + \sum_{s=1}^{h} R^{h-s}y_{t+s} - R^h p_t\right) z_{t,i}\,, \end{aligned} \tag{A.1}$$

where $x_{t,i}$ is the share of the investor's wealth invested in the risky asset, so that $z_{t,i} = x_{t,i}W_{t,i}/p_t$ are the investor's *holdings* of the risky asset bought at time t. In expression (A.1), it is assumed that all dividends are automatically reinvested in the risk-free asset every period.

Following Brock and Hommes (1998), we assume that all investors with investment horizon h are mean-variance maximizers with risk aversion parameter a and belief about the conditional variance of excess return (over h periods) given by σ_h^2. The optimal amount of the risky asset to be purchased by an investor with horizon h, given return expectations, is then

$$z_{t,i,h} = \frac{\mathrm{E}_{t,i,h}\left(p_{t+h} + \sum_{s=1}^{h} R^{h-s}y_{t+s} - R^h p_t\right)}{a\sigma_h^2}\,,$$

where $\mathrm{E}_{t,i,h}[\,\cdot\,]$ stands for the expectations that investor i holds, in period t, about future prices and dividends. The notation stresses that the expectations may be heterogeneous even for investors with the same horizon, and that expectations are formed at time t.

The asset price at period t is found from the equilibrium between the total demand, i.e., the sum of individual demands by all institutional investors (with different investment horizons) in the market, and supply, which is assumed to fluctuate randomly, due to noise traders, around the zero mean.[30] The temporary market equilibrium equation reads

$$\sum_{i,h} z_{t,i,h} = \sum_h \frac{1}{a\sigma_h^2} \left(\sum_i \mathrm{E}_{t,i,h}[p_{t+h}] + \sum_i \sum_{s=1}^{h} R^{h-s} \mathrm{E}_{t,i,h}[y_{t+s}] - N_h R^h p_t \right) = \epsilon_t,$$

where N_h represents the number of investors with investment horizon h. Dividing this equation by the total number of investors, $N = \sum_h N_h$, we obtain

$$\sum_h \frac{f_h}{N_h} \sum_i \left(\mathrm{E}_{t,i,h}[p_{t+h}] + \sum_{s=1}^{h} R^{h-s} \mathrm{E}_{t,i,h}[y_{t+s}] \right) - \sum_h f_h R^h p_t = \varepsilon_t,$$

where the "adjusted" fraction of investors with horizon h is $f_h = N_h/(a\sigma_h^2 N)$, and the normalized supply from the noise traders is $\varepsilon_t = \epsilon_t/N$. From this, it immediately follows that the market-clearing price is given as:

$$p_t = \frac{1}{\sum_h f_h R^h} \left(\sum_h \frac{f_h}{N_h} \sum_i \left(\mathrm{E}_{t,i,h}[p_{t+h}] + \sum_{s=1}^{h} R^{h-s} \mathrm{E}_{t,i,h}[y_{t+s}] \right) - \varepsilon_t \right). \quad \text{(A.2)}$$

A.2 Asset prices in the laboratory experiment

To use (A.2) as the price generating mechanism for the laboratory experiment, we make two further assumptions. First, we assume that the dividend process is IID with mean value $\bar{y}$, which is known to all investors. This implies that the stream of expected dividend payments can be computed as:

$$\sum_{s=1}^{h} R^{h-s} \mathrm{E}_{t,i,h}[y_{t+s}] = \bar{y}\, \frac{R^h - 1}{R - 1} = \frac{\bar{y}}{r} \left(R^h - 1 \right).$$

We define the *fundamental price* as the constant price solution under the assumption that all investors have rational expectations, i.e., $\mathrm{E}_{t,i,h}[p_{t+h}] = p_{t+h}$, for all t, i and h, and the total supply of noise traders is zero. Under this assumption, the pricing equation (A.2) becomes

$$p_t = \frac{1}{\sum_h f_h R^h} \sum_h f_h \left(p_{t+h} + \frac{\bar{y}}{r} (R^h - 1) \right).$$

Then it is straight-forward to check that the fundamental price is $p^f = \bar{y}/r$, i.e., it is equal to the discounted sum of all future expected dividends, and is independent of the distribution of traders over the different investment horizons.

[30] The assumption that, on average, there is no outside supply of the risky asset, is standard in the literature, see, e.g., Brock and Hommes (1998).

Second, in this paper we focus on the case where all N investors have the same horizon H. Pricing equation (A.2) then simplifies to

$$p_t = \frac{1}{R^H}\left(\frac{1}{N}\sum_i \mathrm{E}_{t,i}[p_{t+H}] + \bar{y}\left(\frac{R^H - 1}{R-1}\right) - \varepsilon_t\right) ,$$

which, using fundamental price $p^f = \bar{y}/r$, can be further rewritten as

$$p_t = p^f + \frac{1}{R^H}\left(\frac{1}{N}\sum_{i=1}^{N} \mathrm{E}_{t,i}[p_{t+H}] - p^f - \varepsilon_t\right) . \tag{A.3}$$

This equation defines the market-clearing price in period t that we use in the experiment, see Eq. (2) in the main text. In the experiment we implement Eq. (A.3) with $N = 6$ (the size of the group), $H = 1$, 2 and 3 (values of the investment horizon in different treatments), $p^f = 60$ and $R = 1.05$ or, in treatment **OscH3SF**, $R = 1.016$, see Table 1. Note that the fundamental price, $p^f = \bar{y}/r$, is the rational expectation solution of (A.3) regardless of the values of the parameters.

We note that, in the absence of noise traders (i.e., when $\varepsilon_t = 0$ for all t) other rational expectations equilibria, so-called *rational bubbles*, exist in this model. Such a rational bubble is an equilibrium where traders expect the deviation of the asset price from the fundamental value to grow with a factor R each period. That is, if $p_0 > p^f$ and expectations of all traders at time t are given by

$$\mathrm{E}_{t,i}[p_{t+H}] = p^f + R^{H+1}(p_{t-1} - p^f) ,$$

then substituting these expectations in (A.3), we derive that the price in period t is $p_t = p^f + R(p_{t-1} - p^f)$, confirming the expectation scheme. Hommes et al. (2008) study the experimental bubbles and reject that they correspond to the rational bubbles.

B Experimental Instructions

Instructions below were distributed to all subjects participating in the **OscH3SF** treatment. The wording of the instructions in all other treatments is the same. The instructions are only adjusted to the corresponding investment horizon and the specific values of parameters r and $\bar{y}$ used in a treatment, see Table 1.

INSTRUCTIONS

General information

Today you will participate in an experiment which will require you to predict the future price of a risky asset. During the experiment you will be able to earn a number of points. The better your predictions are, the more points will you earn. These points will be converted into Australian dollars after the experiment.

Information about your task

You are a **financial forecaster** working for a pension fund that wants to optimally invest a large amount of money for 3 periods. The pension fund has two investment options: a risk-free investment and a risky investment. The risk-free investment is putting money in **a savings account**, which pays a fixed and constant interest rate over 3 periods. The alternative for the pension fund is to invest its money in **a risky asset**, where risk comes from the uncertain future price of that asset.

In each period the pension fund has to decide which fraction of its money to put in the savings account and which fraction of its money to invest in the risky asset. To make the optimal investment decision, the pension fund needs an accurate prediction of the future price of the asset. The pension fund is only interested in the price of the risky asset after 3 periods.

As the **financial forecaster** of the fund, you have to predict the price for the risky asset 3 periods ahead during 53 subsequent periods. Your earnings during the experiment depend upon the accuracy of your predictions. **The smaller your errors in each period are, the higher your total earnings will be.**

Information about the asset market

The market price of the risky asset in each period is determined by demand and supply. The total supply of assets is fixed during the experiment. The demand for assets is mainly determined by the aggregate demand of several large pension funds active in the asset market. There is also some uncertain, small demand for assets by private investors but the effect of private investors upon the asset price is small.

Information about the investment strategies of the pension funds

The precise investment strategy of the pension fund that you are advising and the investment strategies of the other pension funds are unknown. The savings account, that provides the risk-free investment, pays a fixed interest rate of 1.6% per period. The owner of the risky asset receives an uncertain payment in each period, but economic experts have computed that this payment is 0.96 dollars per period on average. The return of the asset market per period depends upon these payments as well as upon price changes of the asset.

As the **financial forecaster** of a pension fund you are only asked to predict, in each period, the 3 periods ahead price of the asset. Based upon your future price predictions, your pension fund will make an optimal investment decision and hold the asset for 3 periods. The higher your predicted future price is, the larger will be the fraction of money invested by your pension fund in the asset market in the current period, so the larger will be its demand for assets.

Information during the experiment

At the beginning of the experiment, you have the history of the asset price in the first 10 periods, and you start in period 11 by giving your prediction of the price in period 14. After all participants have given their predictions, the realized asset price for period 11 will be revealed. Then you (as all other participants) will need to make a new prediction, now for the price in period 15, so that the asset price for period 12 can be defined. And so on. This process continues until period 63, where the last prediction, for the price in period 66, will be given.

To predict the asset price for period $t + 3$ in period t, the available information consists of

- past prices up to period $t - 1$,
- your previous predictions up to period $t + 2$,
- your past earnings up to period $t - 1$.

Starting from period 14, your earnings in each period will be based upon your prediction error, that is, the difference between the price you predicted for that period and the realized price in that period. The last period for which you will be paid is period 63.

The better you predict the asset price in each period, the higher your aggregate earnings will be. Earnings for each period in points will be automatically computed according to the following earnings table, where "error" denotes the absolute value of the difference between your prediction and price in that period. Information on your earnings in the current period and cumulative earnings will be reported to you during the experiment.

After the experiment your earned points will be converted into Australian dollars, with 1300 points equal to 50 cents. You will be paid the sum of show-up fee and all your earnings in AUD.

Earnings table									
Error	**Points**	**Error**	**Points**	**Error**	**Points**	**Error**	**Points**	**Error**	**Points**
0.1	1300	**1.5**	1240	**2.9**	1077	**4.3**	809	**5.7**	438
0.2	1299	**1.6**	1232	**3**	1061	**4.4**	786	**5.8**	408
0.3	1298	**1.7**	1223	**3.1**	1045	**4.5**	763	**5.9**	376
0.4	1296	**1.8**	1214	**3.2**	1028	**4.6**	739	**6**	345
0.5	1293	**1.9**	1204	**3.3**	1011	**4.7**	714	**6.1**	313
0.6	1290	**2**	1194	**3.4**	993	**4.8**	689	**6.2**	280
0.7	1287	**2.1**	1183	**3.5**	975	**4.9**	663	**6.3**	247
0.8	1283	**2.2**	1172	**3.6**	956	**5**	637	**6.4**	213
0.9	1279	**2.3**	1160	**3.7**	937	**5.1**	610	**6.5**	179
1	1273	**2.4**	1147	**3.8**	917	**5.2**	583	**6.6**	144
1.1	1268	**2.5**	1134	**3.9**	896	**5.3**	555	**6.7**	109
1.2	1262	**2.6**	1121	**4**	876	**5.4**	526	**6.8**	73
1.3	1255	**2.7**	1107	**4.1**	854	**5.5**	497	**6.9**	37
1.4	1248	**2.8**	1092	**4.2**	832	**5.6**	468	**7 or more**	0

Additional information

- By the end of the experiment, you will be paid privately. Before the payment you will be asked to answer a questionnaire. Inserted data will be processed in nameless form only. Please fill in the correct information.
- During the experiment any communication with other participants, whether verbal or written, is forbidden. The use of phones, tablets or any other gadgets is not allowed. Violation of the rules can result in exclusion from the experiment without any remuneration.
- Please follow the instructions carefully at all the stages of the experiment. If you have any questions or encounter any problems during the experiment, please raise your hand and the experimenter will come to help you.

Please ask any questions you have now!

C Descriptive statistics, tests, estimations

C.1 Descriptive statistics

Treatment	Group	Mean price	Std dev	Median RAD	IQR	Discoord	Average payoffs
ConH1	Group 1	57.67	2.07	0.04	3.36	0.66	1221.2
	Group 2	58.69	0.99	0.02	1.23	0.28	1272.4
	Group 3	58.23	3.65	0.02	3.72	0.46	1226.5
	Average	58.20	2.24	0.02	2.77	0.46	930.0
	Median	58.23	2.07	0.02	3.36	0.46	1223.8
ConH2	Group 1	56.63	0.99	0.05	1.35	1.01	1198.9
	Group 2	64.33	14.34	0.04	3.99	1.97	857.6
	Group 3	59.37	1.08	0.01	0.81	0.70	1215.5
	Group 4	59.08	0.70	0.01	0.88	0.68	1241.1
	Average	59.85	4.28	0.03	1.76	1.09	1128.3
	Median	59.22	1.04	0.03	1.12	0.86	1207.2
ConH3	Group 1	63.46	1.28	0.06	1.47	2.12	1016.7
	Group 2	58.96	1.24	0.02	1.82	1.31	1121.6
	Group 3	59.76	0.62	0.00	0.80	0.62	1252.2
	Average	60.73	1.05	0.03	1.36	1.35	1130.2
	Median	59.76	1.24	0.02	1.47	1.31	1121.6
OscH1	Group 1	63.28	1.26	0.05	0.86	0.37	1266.0
	Group 2	60.64	1.64	0.02	2.76	0.88	1223.5
	Group 3	52.25	23.91	0.36	42.42	5.24	494.7
	Group 4	63.85	6.96	0.08	6.48	1.39	946.3
	Group 5	61.30	1.07	0.02	0.99	0.35	1266.9
	Group 6	72.25	45.16	0.49	76.90	5.44	382.9
	Group 7	231.63	305.74	0.47	268.87	6.18	475.5
	Group 8	67.64	57.04	0.11	16.78	0.87	909.5
	Average	84.10	55.35	0.20	52.01	2.59	870.7
	Median	63.57	15.43	0.10	11.63	1.14	927.9
OscH2	Group 1	54.84	6.76	0.11	8.98	4.94	413.5
	Group 2	51.93	18.56	0.16	23.67	5.27	332.3
	Group 3	58.42	14.08	0.14	15.99	6.29	279.0
	Group 4	60.32	11.83	0.19	23.46	5.67	213.5
	Group 5	57.84	13.09	0.16	19.86	9.79	149.9
	Average	56.67	12.86	0.15	18.39	6.39	231.4
	Median	57.84	13.09	0.16	19.86	5.67	246.3
OscH3	Group 1	60.47	2.18	0.02	2.40	2.50	977.6
	Group 2	67.57	12.73	0.09	2.49	2.76	791.4
	Group 3	60.56	2.57	0.03	3.66	1.77	813.4
	Group 4	54.37	11.12	0.12	10.90	10.31	270.6
	Group 5	61.33	4.60	0.03	2.28	3.89	766.6
	Group 6	58.57	3.61	0.02	2.20	1.53	907.9
	Average	60.48	6.14	0.05	3.99	3.79	754.6
	Median	60.52	4.11	0.03	2.44	2.63	802.4
OscH3SF	Group 1	61.02	3.34	0.02	1.69	0.60	776.8
	Group 2	60.72	0.97	0.01	1.02	0.88	1192.8
	Group 3	59.76	15.61	0.14	17.11	7.21	284.2
	Group 4	55.73	1.82	0.07	2.24	2.21	1022.2
	Group 5	50.60	19.54	0.19	21.05	11.96	206.4
	Group 6	45.10	14.72	0.23	17.31	8.61	249.0
	Average	55.49	9.33	0.11	10.07	5.25	621.9
	Median	57.74	9.03	0.11	9.68	4.71	530.5

Table 2: Descriptive statistics in the experiment.

C.2 Nonparametric test for median of RAD comparison

Median RAD	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.318	0.633	0.043**	0.007***	0.273	0.099*
ConH2	0.318	X	0.601	0.075*	0.003***	0.220	0.068*
ConH3	0.161	0.601	X	0.083*	0.007***	0.099*	0.099*
OscH1	1.000	1.000	1.000	X	0.153	0.879	0.879
OscH2	1.000	1.000	1.000	0.348	X	1.000	0.472
OscH3	1.000	0.845	1.000	0.239	0.009***	X	0.160
OscH3SF	0.561	1.000	1.000	0.313	0.185	0.442	X

Table 3: p-values of the one-sided Kolmogorov-Smirnov test comparing the median of RAD statistics, pairwise for different treatments, where the alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively.

Median RAD	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.338	0.444	0.063*	0.005***	0.232	0.095*
ConH2	0.662	X	0.556	0.030**	0.005***	0.227	0.066*
ConH3	0.556	0.444	X	0.048**	0.024**	0.212	0.082*
OscH1	0.938	0.970	0.952	X	0.712	0.936	0.834
OscH2	0.991	0.995	0.976	0.288	X	0.996	0.815
OscH3	0.768	0.773	0.788	0.064*	0.004***	X	0.090*
OscH3SF	0.905	0.934	0.918	0.166	0.185	0.910	X

Table 4: p-values of the one-sided Fischer-Pitman permutation test comparing the median of RAD statistics, pairwise for different treatments, where the alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively.

Median RAD	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.571	0.800	0.042**	0.018**	0.190	0.274
ConH2	0.571	X	0.571	0.036**	0.008***	0.238	0.057*
ConH3	0.350	0.571	X	0.042**	0.018**	0.131	0.131
OscH1	0.976	0.976	0.976	X	0.311	0.886	0.793
OscH2	1.000	1.000	1.000	0.738	X	0.998	0.732
OscH3	0.869	0.824	0.917	0.141	0.004***	X	0.294
OscH3SF	0.810	0.967	0.917	0.245	0.331	0.758	X

Table 5: p-values of the one-sided Mann-Whitney-Wilcoxon test comparing the median of RAD statistics, pairwise for different treatments, where the alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively.

C.3 Nonparametric test for IQR comparison

IQR	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.751	1.000	0.113	0.007***	0.561	0.273
ConH2	0.318	X	0.451	0.075*	0.003***	0.033**	0.127
ConH3	0.161	0.601	X	0.043**	0.007***	0.006***	0.099*
OscH1	0.705	1.000	1.000	X	0.153	0.597	0.597
OscH2	1.000	1.000	1.000	0.348	X	1.000	1.000
OscH3	0.561	0.959	1.000	0.127	0.009***	X	0.160
OscH3SF	0.866	1.000	1.000	0.313	0.185	0.442	X

Table 6: p-values of the one-sided Kolmogorov-Smirnov test comparing the interquartile range statistics, pairwise for different treatments, where the alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively.

IQR	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.817	0.889	0.173	0.009***	0.393	0.156
ConH2	0.183	X	0.528	0.042**	0.005***	0.104	0.081*
ConH3	0.111	0.472	X	0.048**	0.012**	0.012**	0.071*
OscH1	0.827	0.958	0.952	X	0.710	0.953	0.870
OscH2	0.986	0.995	0.988	0.290	X	0.997	0.938
OscH3	0.607	0.896	0.988	0.047**	0.003***	X	0.090*
OscH3SF	0.844	0.919	0.929	0.130	0.062*	0.910	X

Table 7: p-values of the one-sided Fischer-Pitman permutation test comparing the interquartile range statistics, pairwise for different treatments, where the alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively.

IQR	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.800	0.900	0.248	0.018**	0.548	0.357
ConH2	0.314	X	0.571	0.055*	0.008***	0.086*	0.057*
ConH3	0.200	0.571	X	0.067*	0.018**	0.012**	0.083*
OscH1	0.812	0.964	0.958	X	0.362	0.886	0.669
OscH2	1.000	1.000	1.000	0.689	X	0.998	0.937
OscH3	0.548	0.943	1.000	0.141	0.004***	X	0.469
OscH3SF	0.726	0.967	0.952	0.377	0.089*	0.591	X

Table 8: p-values of the one-sided Mann-Whitney-Wilcoxon test comparing the interquartile range statistics, pairwise for different treatments, where the alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively.

C.4 Estimated AR(2) Rules

Treatment	Group	Const	Past Prices		Specification Tests		AdjRSq	NObs
			lag 1	lag 2	autoc	hskd		
	Group 1	17.72***	1.54***	−0.85***	0.71	0.35	0.92	48
ConH1		(2.55)	(0.08)	(0.08)				
	Group 2	18.66***	1.32***	−0.64***	0.68	0.70	0.78	48
		(4.24)	(0.11)	(0.11)				
	Group 3	5.41***	1.37***	−0.46***	0.34	0.24	0.97	48
		(1.39)	(0.12)	(0.11)				
	Group 1	40.99***	0.90***	−0.62***	0.13	0.25	0.57	48
ConH2		(6.05)	(0.12)	(0.11)				
	Group 2	45.76***	0.44***	−0.15	1.00	0.88	0.13	48
		(10.70)	(0.15)	(0.15)				
	Group 3	53.60***	0.49***	−0.39***	0.99	0.54	0.23	48
		(9.24)	(0.14)	(0.14)				
	Group 4	52.03***	0.54***	−0.42***	0.74	0.43	0.28	48
		(8.32)	(0.13)	(0.13)				
	Group 1	35.45***	0.91***	−0.47***	0.51	0.57	0.51	48
ConH3		(7.24)	(0.13)	(0.13)				
	Group 2	27.81***	1.21***	−0.69***	0.69	0.54	0.73	48
		(4.80)	(0.11)	(0.11)				
	Group 3	46.64***	0.48***	−0.26*	0.32	0.07	0.17	48
		(9.50)	(0.14)	(0.14)				
	Group 1	9.05***	1.34***	−0.48***	0.42	0.88	0.87	48
OscH1		(3.13)	(0.13)	(0.12)				
	Group 2	11.66***	1.34***	−0.53***	0.81	0.39	0.84	48
		(3.71)	(0.12)	(0.12)				
	Group 3	16.87***	1.53***	−0.85***	0.01	0.31	0.91	48
		(2.68)	(0.08)	(0.08)				
	Group 4	25.34***	1.76***	−1.16***	0.06	0.85	0.98	48
		(1.29)	(0.04)	(0.04)				
	Group 5	12.36***	1.14***	−0.34**	0.66	0.52	0.76	48
		(4.19)	(0.14)	(0.13)				
	Group 6	12.29***	1.86***	−1.03***	0.00	0.44	0.99	48
		(1.48)	(0.04)	(0.04)				
	Group 7	29.91**	1.76***	−0.88***	0.08	0.06	0.96	48
		(11.32)	(0.08)	(0.09)				
	Group 8	2.56**	2.33***	−1.37***	0.92	0.55	1.00	48
		(1.13)	(0.04)	(0.05)				
	Group 1	28.23***	1.28***	−0.79***	0.74	0.23	0.83	48
OscH2		(3.43)	(0.08)	(0.08)				
	Group 2	25.82***	1.40***	−0.89***	0.03	0.55	0.89	48
		(2.85)	(0.07)	(0.07)				
	Group 3	28.71***	1.31***	−0.80***	0.01	0.00	0.82	48
		(3.95)	(0.09)	(0.09)				
	Group 4	34.97***	1.28***	−0.86***	0.09	0.02	0.86	48
		(3.56)	(0.07)	(0.07)				
	Group 5	33.02***	1.19***	−0.76***	0.03	0.84	0.76	48
		(4.65)	(0.10)	(0.10)				
	Group 1	23.33***	0.59***	0.02	0.97	0.67	0.37	48
OscH3		(7.12)	(0.15)	(0.14)				
	Group 2	47.17***	0.40***	−0.10	1.00	0.88	0.10	48
		(11.37)	(0.15)	(0.15)				
	Group 3	21.56***	1.53***	−0.89***	0.12	0.41	0.92	48
		(2.64)	(0.07)	(0.07)				
	Group 4	34.48***	0.94***	−0.56***	0.16	0.19	0.60	48
		(4.90)	(0.11)	(0.10)				
	Group 5	39.42***	0.99***	−0.63***	0.00	0.33	0.69	48
		(4.63)	(0.10)	(0.08)				
	Group 6	10.59***	1.54***	−0.71***	0.01	0.10	0.93	48
		(2.09)	(0.09)	(0.09)				
	Group 1	18.98**	1.03***	−0.34*	0.27	0.66	0.57	38
OscH3SF		(7.16)	(0.18)	(0.18)				
	Group 2	38.99***	0.96***	−0.60***	0.21	0.93	0.61	48
		(5.56)	(0.11)	(0.10)				
	Group 3	26.25***	1.27***	−0.71***	0.46	0.13	0.77	48
		(4.61)	(0.10)	(0.10)				
	Group 4	21.73***	1.08***	−0.47***	0.13	0.60	0.66	48
		(5.12)	(0.12)	(0.11)				
	Group 5	30.25***	0.75***	−0.34**	1.00	0.98	0.36	48
		(7.03)	(0.14)	(0.14)				
	Group 6	23.44***	1.11***	−0.63***	0.26	0.05	0.66	48
		(4.29)	(0.11)	(0.11)				

Table 9: Estimated AR(2) rules, Eq. (7), for each experimental group. *, **, and *** for the estimations show significance at the 10%, 5% or 1% significance level, respectively, with the standard errors shown below in parentheses. Columns 'autoc' and 'hskd' show p-values of the specification tests for the residual autocorrelation (Ljung-Box test) and heteroscedasticity (Engle's ARCH tests), respectively. The group label is in bold when the estimated rule passes both tests at the 5% significance level.

C.5 t-test for correlations comparison

Slope	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.944	0.947	0.993	0.991	0.957	0.981
ConH2	0.962	X	0.963	0.988	0.973	0.970	0.975
ConH3	0.002***	0.866	X	0.964	0.882	0.868	0.901
OscH1	0.036**	0.033**	0.036**	X	0.028**	0.024**	0.026**
OscH2	0.861	0.895	0.862	0.956	X	0.887	0.901
OscH3	0.169	0.216	0.169	0.926	0.781	X	0.768
OscH3SF	0.935	0.946	0.935	0.970	0.941	0.942	X

Table 10: p-values of the t-test of the differences in correlations between $\bar{p}^e_{i,t+h} - p_{t-1}$ and the price change, $p_{t-1} - p_{t-2}$ (see Fig. 8), pairwise for different treatments, where the alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively.

C.6 Nonparametric test for discoordination comparison

Discoord	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.010**	0.161	0.043**	0.007***	0.006***	0.027**
ConH2	1.000	X	0.451	0.394	0.003***	0.033**	0.068*
ConH3	1.000	0.601	X	0.456	0.007***	0.099*	0.099*
OscH1	1.000	0.661	0.705	X	0.053*	0.040**	0.127
OscH2	1.000	1.000	1.000	1.000	X	0.829	0.544
OscH3	1.000	1.000	1.000	0.699	0.009***	X	0.442
OscH3SF	1.000	0.845	0.866	1.000	0.185	0.442	X

Table 11: p-values of the one-sided Kolmogorov-Smirnov test comparing the discoordination (median of the standard deviations of predictions) statistics, pairwise for different treatments, where the alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively.

D Dynamics of individual forecasts and price

Figures 11-17 show the evolution of individual forecasts and prices (in black with markers) for each of the 35 groups of our experiment. Figures are organized by treatment, see captions for the details. The first ten periods in each panel show the initial price history in the group. The point forecasts are displayed against those time periods that they are made for (and not *when* they are made). For instance, in treatment **ConH3**, in period $t = 15$, we show price p_{15} and individual forecasts, $p^e_{i,15}$ that were made in period 12.

Discoord	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.056*	0.083*	0.087*	0.005***	0.014**	0.062*
ConH2	0.944	X	0.278	0.200	0.005***	0.014**	0.062*
ConH3	0.917	0.722	X	0.301	0.012**	0.047**	0.141
OscH1	0.913	0.800	0.699	X	0.013**	0.236	0.101
OscH2	0.991	0.995	0.988	0.987	X	0.915	0.690
OscH3	0.986	0.986	0.953	0.764	0.085*	X	0.270
OscH3SF	0.938	0.938	0.859	0.899	0.310	0.730	X

Table 12: p-values of the one-sided Fischer-Pitman permutation test comparing the discoordination (median of the standard deviations of predictions) statistics, pairwise for different treatments, where the alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively.

Discoord	ConH1	ConH2	ConH3	OscH1	OscH2	OscH3	OscH3SF
ConH1	X	0.029**	0.100	0.067*	0.018**	0.012**	0.024**
ConH2	1.000	X	0.429	0.341	0.008***	0.019**	0.129
ConH3	0.950	0.686	X	0.461	0.018**	0.048**	0.190
OscH1	0.958	0.715	0.612	X	0.023**	0.141	0.091*
OscH2	1.000	1.000	1.000	0.985	X	0.974	0.669
OscH3	1.000	0.990	0.976	0.886	0.041**	X	0.531
OscH3SF	0.988	0.914	0.869	0.929	0.396	0.531	X

Table 13: p-values of the one-sided Mann-Whitney-Wilcoxon test comparing the discoordination (median of the standard deviations of predictions) statistics, pairwise for different treatments, where the alternative hypothesis is that the row treatment has a value not lower than the column treatment. *, **, and *** denote treatment comparisons when the null hypothesis of equality is rejected at the 10%, 5% or 1% significance level, respectively.

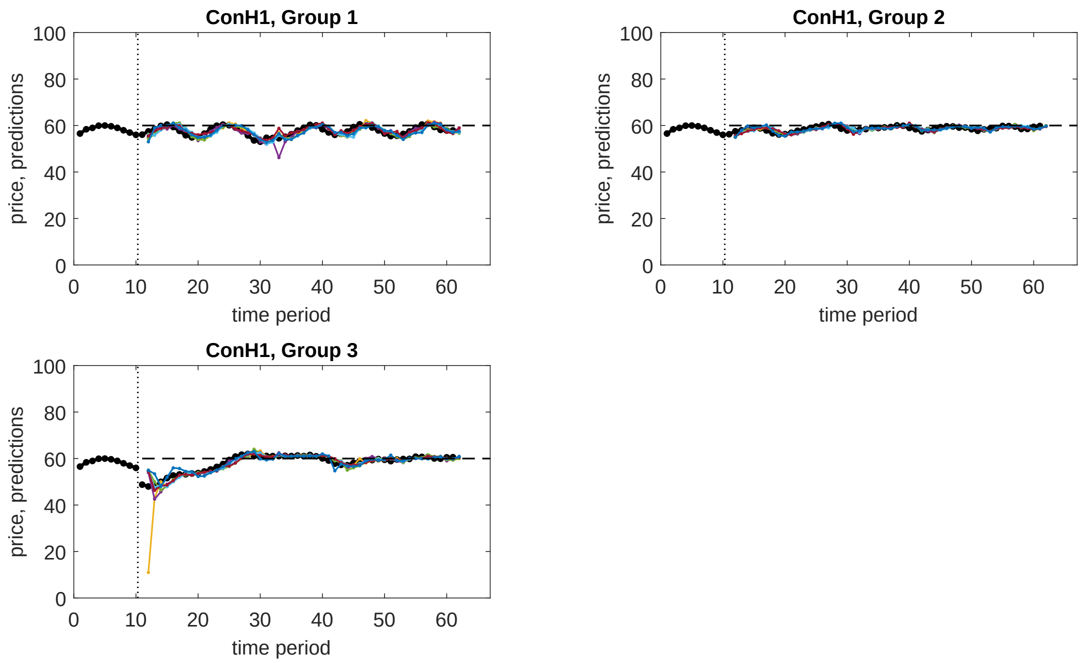

Figure 11: Forecasts and price dynamics (black) in the three groups of the **ConH1** treatment.

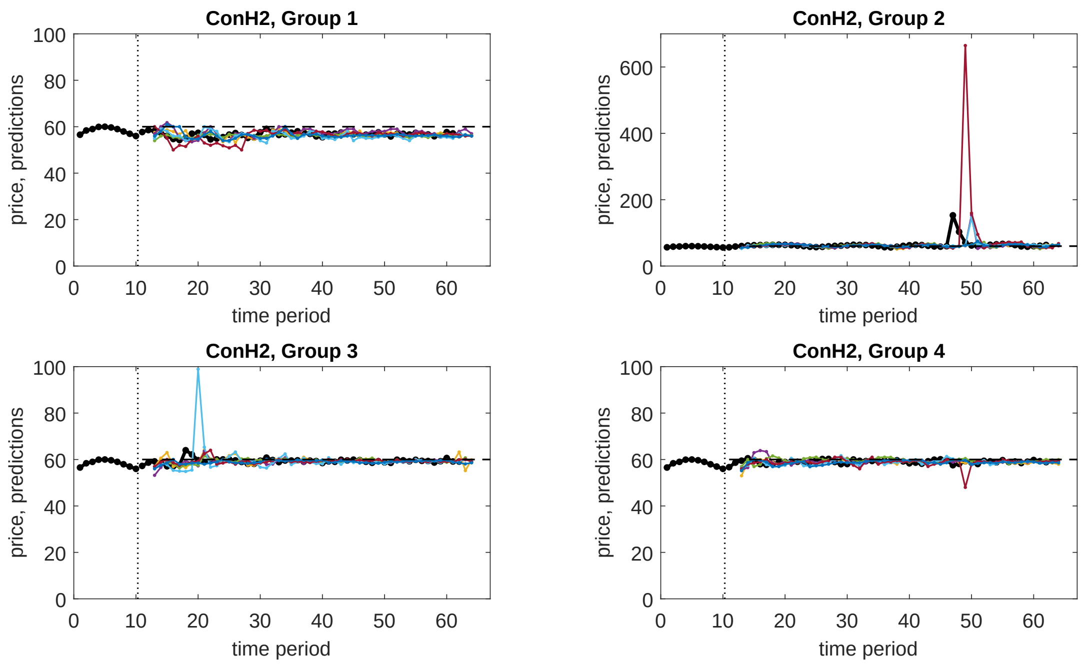

Figure 12: Forecasts and price dynamics in the four groups of the **ConH2** treatment.

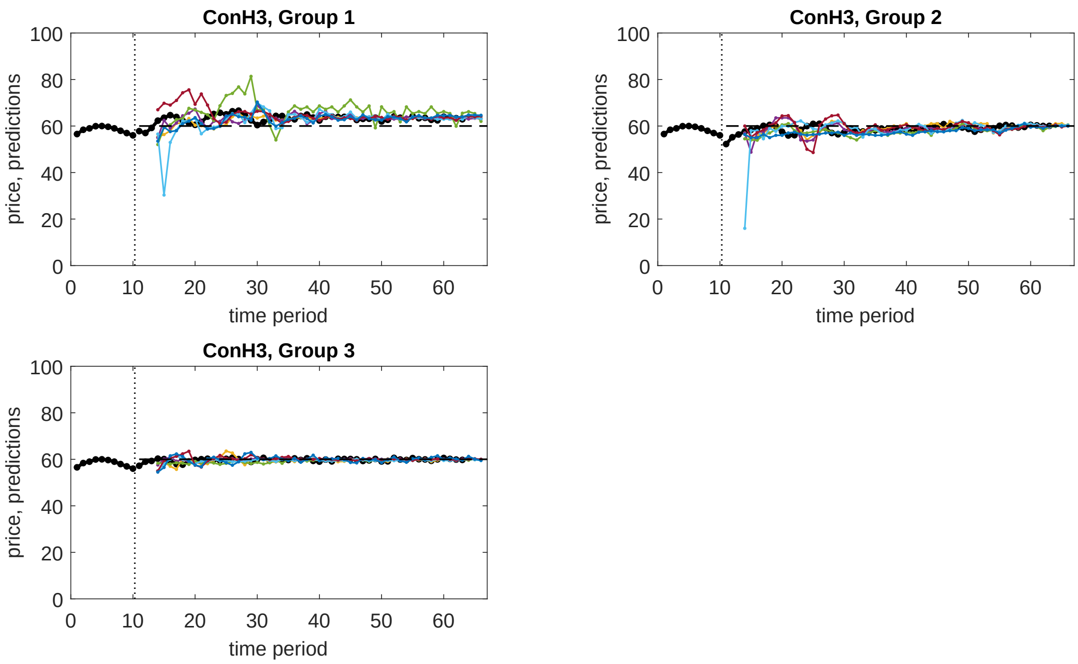

Figure 13: Forecasts and price dynamics in the three groups of the **ConH3** treatment.

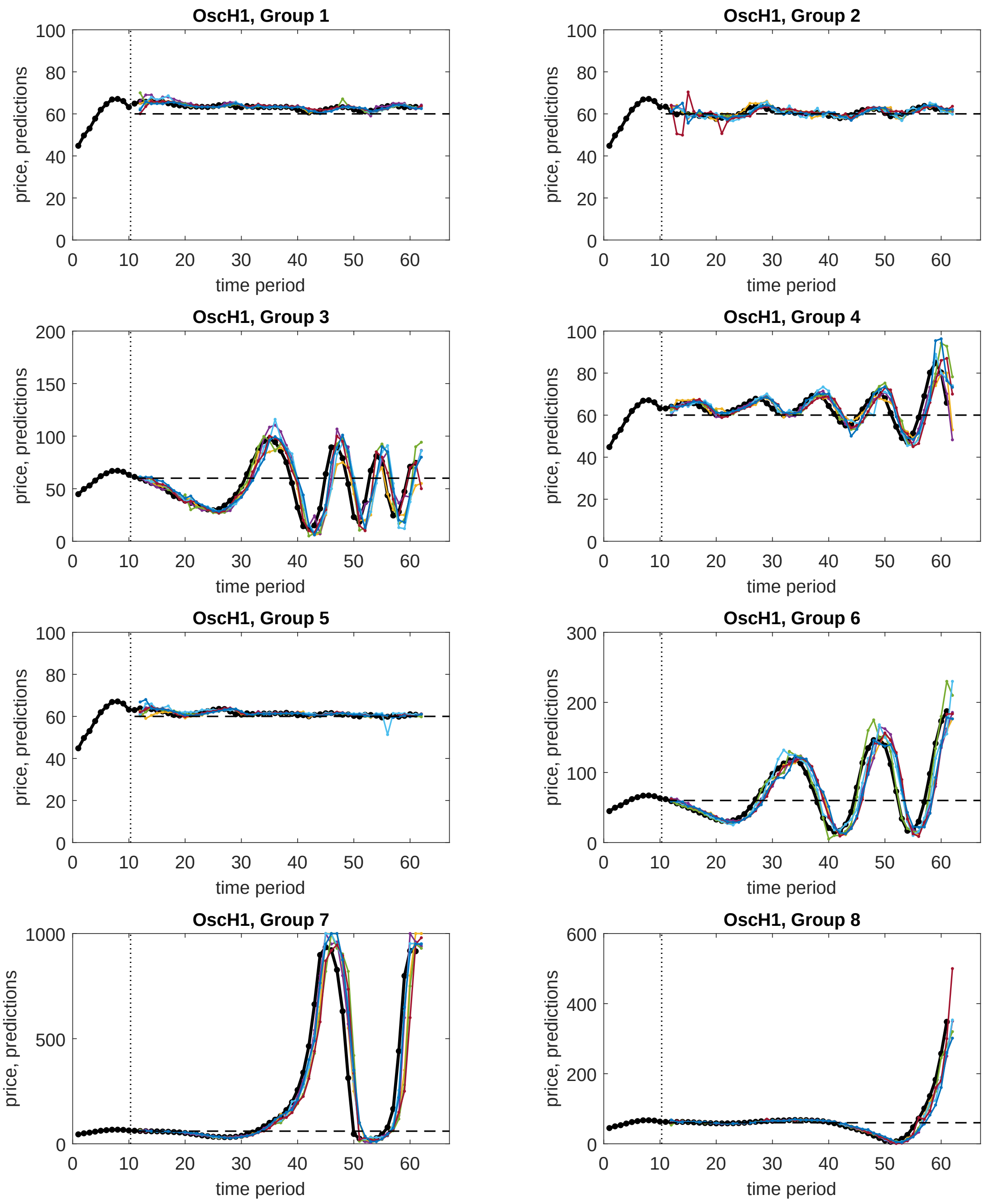

Figure 14: Forecasts and price dynamics in the eight groups of the **OscH1** treatment.

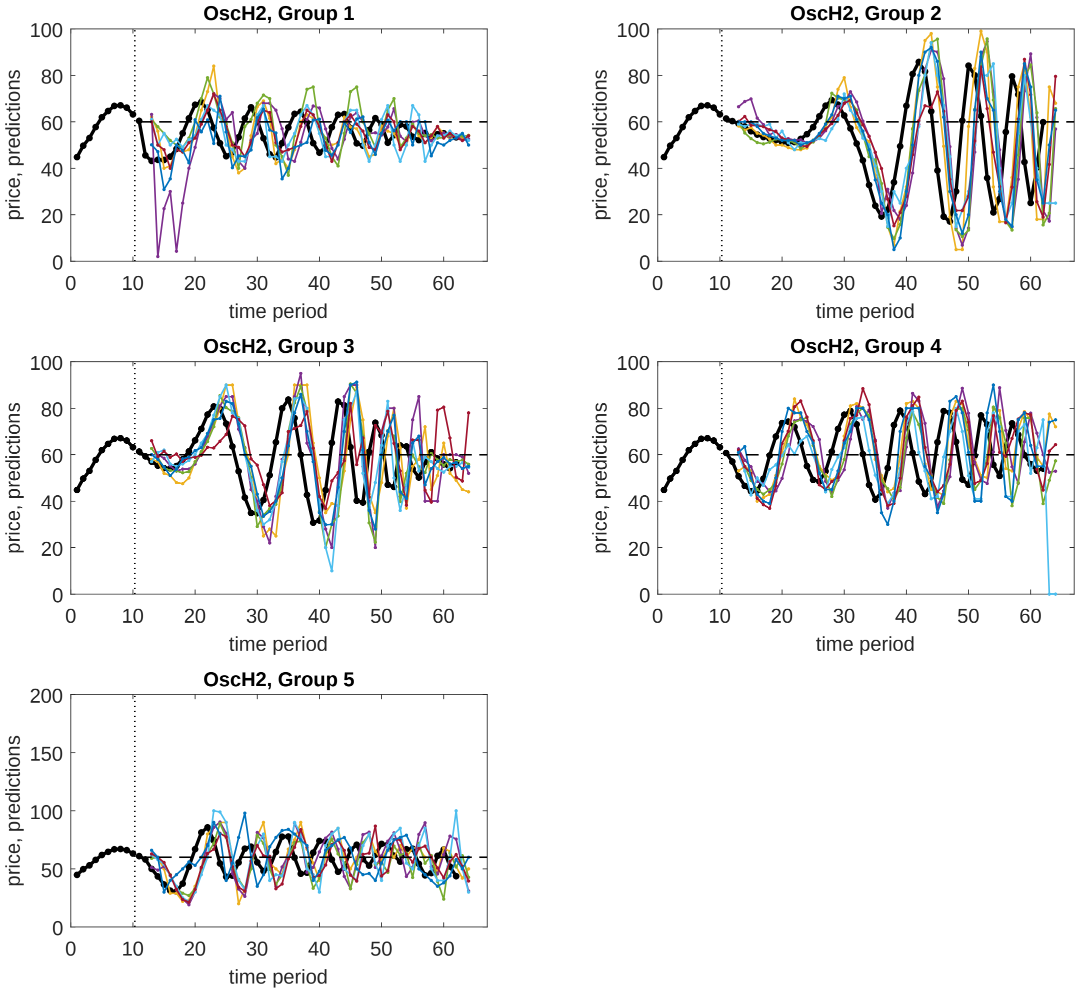

Figure 15: Forecasts and price dynamics in the five groups of the **OscH2** treatment.

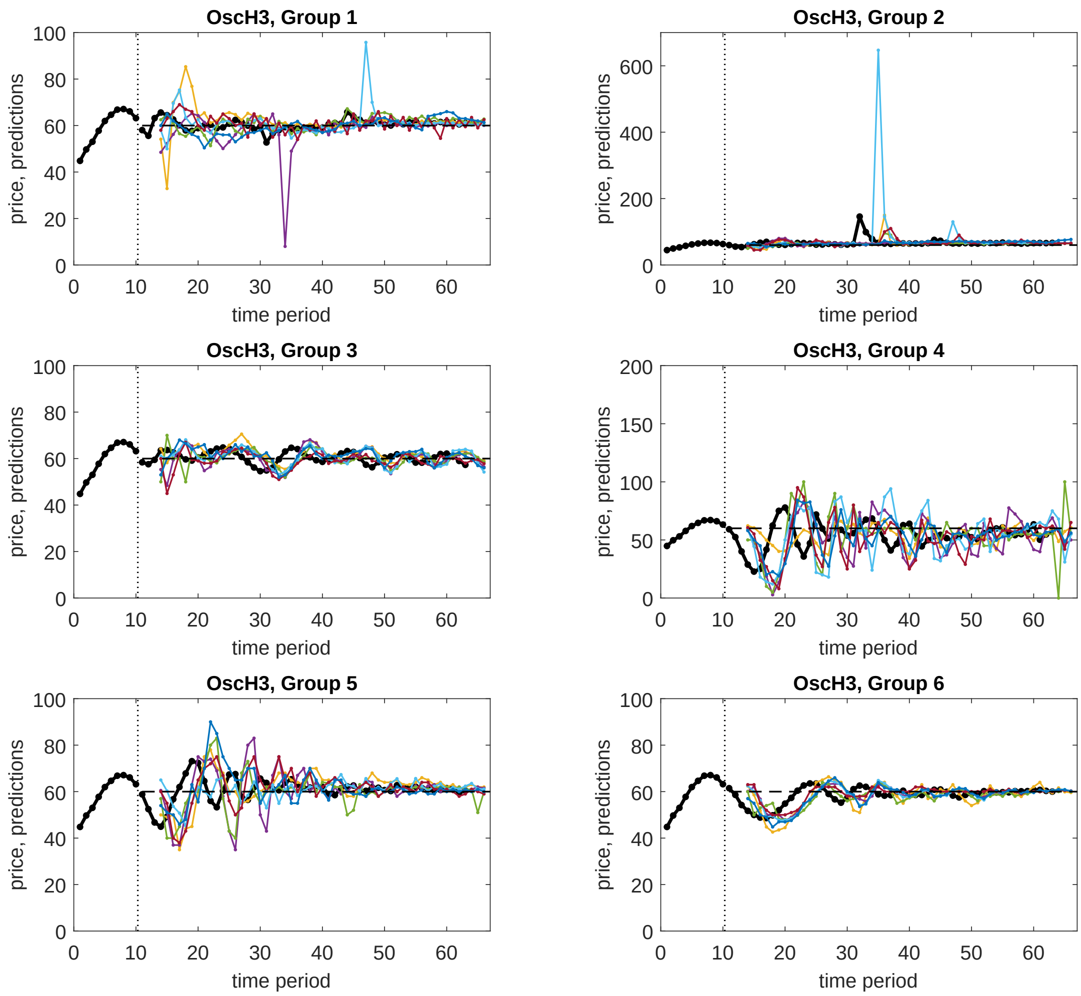

Figure 16: Forecasts and price dynamics in the six groups of the **OscH3** treatment.

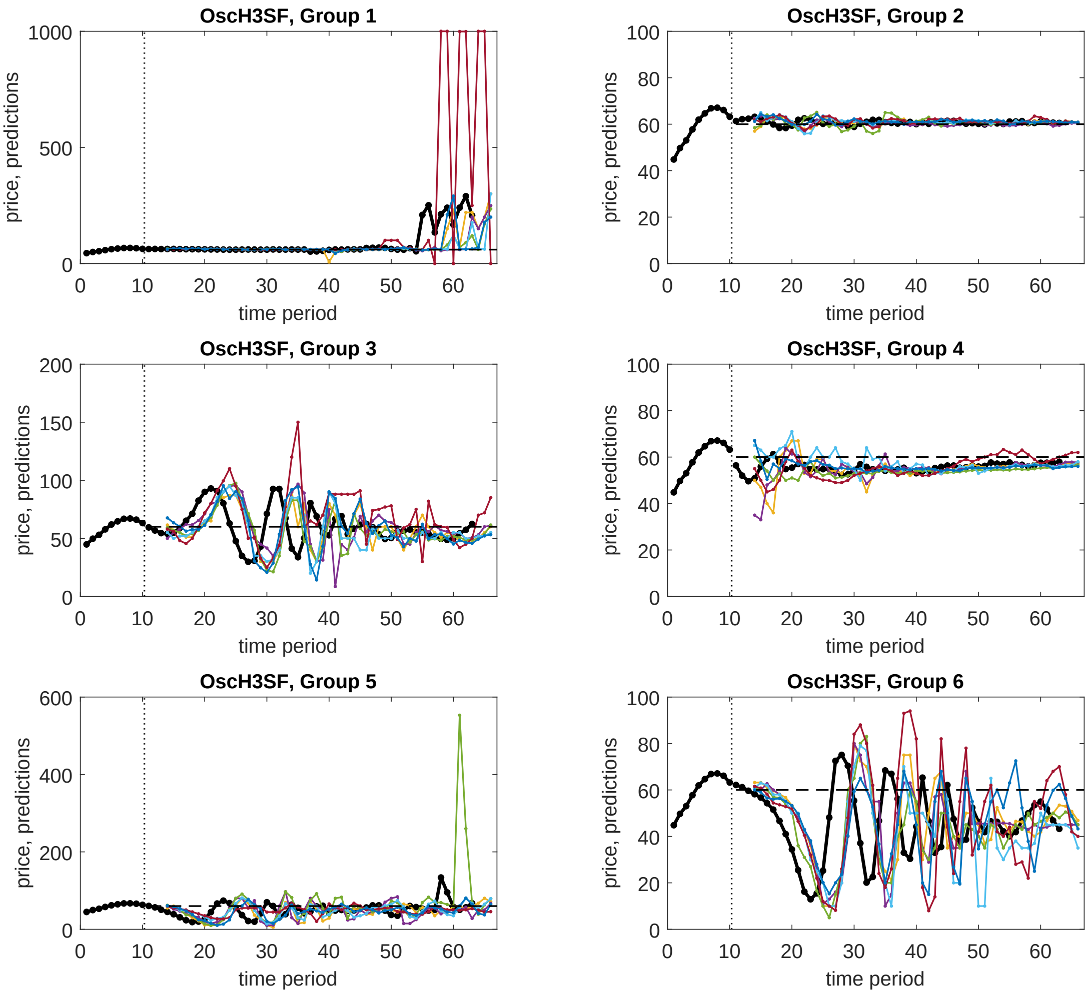

Figure 17: Forecasts and price dynamics in the six groups of the **OscH3SF** treatment. Note the abnormal behaviour of one of the participants in the last 10 periods in group 1, which significantly affected the price dynamics. We excluded the last 10 periods on this group from our analysis.