Long-run expectations in a Learning-to-Forecast Experiment: A Simulation Approach

Annarita Colasante[1], Simone Alfarano[1], Eva Camacho Cuena[1], Mauro Gallegati[2]

[1] *University Jaume I (Spain),* [2] *Università Politecnica delle Marche(Italy)*

Abstract

In this paper, we elicit both short and long-run expectations about the evolution of the price of a financial asset by conducting a Learning-to-Forecast Experiment (LtFE) in which subjects, in each period, forecast the the asset price for each one of the remaining periods. The aim of this paper is twofold: on the one hand, we try to fill the gap in the experimental literature of LtFEs where great effort has been made in investigating short-run expectations, i.e. one step-ahead predictions, while there are no contributions that elicit long-run expectations. On the other hand, we propose an alternative computational approach with respect to the Heuristic Switching Model (HSM), to replicate the main experimental results. The alternative learning algorithm, called Exploration-Exploitation Algorithm (EEA), is based on the idea that agents anchor their expectations around the last market price, rather than on the fundamental value, with a range proportional to the recent past observed price volatility. Both algorithms perform well in describing the dynamics of short-run expectations and the market price. EEA, additionally, provides a fairly good description of long-run expectations.

JEL: D03 G12 C91
Keywords: Expectations, Experiment, Evolutionary Learning

1 Introduction

The economy can be formally thought as *an expectation feedback system*, i.e. a system where agents' expectations, formed based on available information and on past realizations of economic variables, influence future realizations of those variables Hommes [2001]. The comprehension of how individual agents form their expectations constitutes a crucial aspect, on the one hand, for understanding the evolution of the economic system itself and, on the other hand, in devising efficient economic policies to guide it towards social desirable outcomes. One of the main problems when dealing with expectations is that they are not directly observable, contrary to prices, volumes, interest rates and all economic variables recorded every day in world-wide markets. There are several cleverly designed methods to directly or indirectly estimate agents' expectations based on surveys (for a overview see Manski [2004]). However surveys do not typically provide incentives depending on the performance of the responders, so their validity turns out to be limited.

A valid alternative to surveys are controlled laboratory experiments, which have the advantage of a perfect monitoring of the information available to the subjects and, furthermore, allow for the elicitation of the expectations using appropriate performance-based incentives. Learning

to Forecasts Experiments, introduced by Marimon et al. [1993], are laboratory control experiments to elicit subjects' expectations in an expectation-feedback environment, where the feedback between the subjects' expectations and the aggregate quantities, typically prices, is designed by the experimenter. They turn out to be a powerful and flexible tool to study subjects' expectation formation under different feedback systems. Many LtFEs have been conducted to study how agents form their **short-run expectations** in financial markets (Hommes et al. [2005]), real estate markets (Bao and Ding [2015]), commodity markets (Bao et al. [2013]) and in simple macroeconomic frameworks (Assenzaa et al. [2011], Anufriev et al. [2013a], Cornand et al. [2013]). We conduct a Learning to Forecast Experiment (LtFE) in which, unlike the standard settings[1], subjects should submit a prediction for the asset price at different time horizons. In other words, we explicitly elicit subjects' **long-run expectations** at the beginning of every period, giving the possibility to revise their expectations as new information becomes available. The novelty of our experimental design is that it incorporates into the LtFEs the elicitation of long-run expectations, in order to study how expectations form and co-evolve with the market price. Our setting is extremely simple, since long-run expectations do not enter directly in the feedback mechanism, which is influenced just by one-step-ahead predictions.[2] Therefore, we can study how subjects form their long-run expectations based solely on price dynamics. We consider our setting a first step in the direction of a better understanding of the dynamics of expectations in more complex environments. With our paper, we want to fill the gap in the existing experimental literature in eliciting the whole spectrum of subjects' individual expectations. To the best of our knowledge, the only experimental work that elicit the long-run expectations in an asset market with bubbles is the work of Haruvy et al. [2007]. Moreover, Hanaki et al. [2016] investigate the impact of forecast elicitation on the miss-pricing in an experimental asset market. A kind of natural experiment has been conducted by Galati et al. [2011] , where they elicit short, medium and long-run inflation expectations using professional forecasters from central banks, academics and students. They provide, when possible, a reward based on the performances. Some effort has been devoted to elicit long-run expectations using data from surveys (Ashiya [2003], Fujiwara et al. [2013]), but they are not immune to the critical aspects related to the survey methodology.

Why are we interested in eliciting long-run expectations? Several empirical as well as theoretical contributions have stressed the importance of taking into account the whole time spectrum of agents' expectations when designing effective economic policies (Gurkaynak et al. [2005], Cœuré [2013]). Since early 2000s, Central Banks follow a "forward guidance" communication strategy (Woodford [2001]) to try to influence expectations by releasing public announcements on different macroeconomic indicators. Central Banks might also establish medium-term inflation target, to discipline the expectations of economic actors. In both cases, Central -Banks, when devising their monetary policy, take into consideration agents' expectations at different time horizons. These are just examples of why understanding the way agents form and revise their expectations at different time horizons is a relevant issue for policy design.

In the second part of the paper we introduce a learning algorithm able to reproduce the properties of the short as well as long-run expectations observed in our experiment. In the literature of LtFE we find several computational attempts to describe short-run expectations in different experimental settings and information sets using learning algorithms to simulate human behavior. For example, LtFEs with positive and negative feedback, with exogenous shocks, or in a macroeconomic environment as in Heemeijer et al. [2009], Assenzaa et al. [2011], Bao et al. [2013], Hommes and Lux [2013]. A unified framework has been proposed with the aim of reproducing

[1] For a comprehensive survey on the macroeconomic experiments on expectations see Assenza et al. [2014]

[2] We use the term "prediction" referring to the forecasts submitted by the subjects during the experiments. We assume that the subjects submit their predictions based on their expectations, which are not observable. Therefore, across the paper, we use the word "prediction" and "expectation" as (almost) interchangeable.

all experimental results od LtFEs, the so-called Heuristic Switching Model. This approach is based on the idea that each subject considers a limited number of simple extrapolative rules (heuristics), based on the seminal paper of Brock and Hommes [1998]. The artificial agents can switch heuristic depending on the forecasting performance in the recent past. The learning mechanism is based on a performance measure proportional to the quadratic forecasting error (see for example Anufriev and Hommes [2012]).

We propose an alternative approach to HSM to model individual behavior in a LtFE. We introduce an algorithm that we can loosely define as "non-parametric" since it does not impose any predetermined forecasting rule. The main idea arises from the analysis based on professional forecasters as in Campbell and Sharpe [2009] and Nakazono [2012]). According to these studies, professional forecasters, in order to reduce the uncertainty about the future, use the last observed price as an anchor. Looking at the experimental results in LtFEs, especially in the one step-ahead predictions, subjects predict the next price anchoring their predictions around the last realized price. A similar mechanism holds for the long-run predictions, as shown in Colasante et al. [2016]. The alternative algorithm proposed in this paper, called the Exploration-Exploitation Algorithm, is based on the empirical as well as experimental evidence that agents anchor their expectations around the last market price conditioning the range of variability of expectations on the past observed price volatility. The EEA is similar to those algorithms used to solve the multi-armed bandit problem (see Auer et al. [2002] and Koulouriotis and Xanthopoulos [2008]). In this kind of computational problems artificial agents face a trade-off between *exploitation* and *exploration*, i.e. taking a decision using "known and cheap" information up to period t or gathering "new and costly" information about the environment by exploring the "neighbourhood space".

All in all, the HSM and EEA are based on the well-known behavioral principle of *anchor-and-adjustment*, with the difference that the HSM has a predetermined set of few rules, while the EEA has a wider set of variability of the feasible actions available to the agents[3]. We show that EEA and HSM perform well in describing the dynamics of short-run expectations and the market price. EEA additionally provides a fairly good description of long-term expectations, contrary to benchmark version of current version HSM, which is structurally designed to describe short-term expectations.

2 The Learning to Forecast Experiment

2.1 Experimental Design

Our goal is to study expectations formation both, in the short and in the long-run. We implement a LtFE similar to Heemeijer et al. [2009], where the task of subjects is to predict the future price of an asset. In each of the 7 sessions implemented, 6 subjects play the role of professional forecasters for 20 periods (see the instructions in the complementary material). At the beginning of period t, subject i submits her short-run prediction for the asset price at the end of period t, denoted as ${}_ip^e_{t,t}$, as well as her set of long-run predictions for the price at the end of each one of the $20-t$ remaining periods. Long-run predictions are denoted as ${}_ip^e_{t,t+k}$with $1 \leq k \leq 20-t$. Subjects must submit a total of 190 predictions, since we elicit contemporaneously both short and long-run predictions. The choice of 20 periods, contrary to the approximately 50 one-step-ahead predictions typically used in the literature, results from a trade-off between having a time series sufficiently long to conduct a meaningful statistical analysis and, at the same time, avoid

[3]According to this principle, the strategy for estimating unknown quantities is to start with information one does know, i.e. an anchor, and then adjust until an acceptable value is reached, see Tversky and Kahneman [1974].

a too demanding task to the subjects.[4]

When submitting their predictions, subjects are informed about: (i) the constant interest rate (r) and average dividend (d), (ii) the asset prices until period $t-1$, (iii) all their own (short and long-run) past predictions and their corresponding profits. However, they are not informed about the predictions submitted by the other subjects and have just qualitative information on the price generating mechanism. In the instructions subjects are informed that there is a positive relationship between their one-step-ahead predictions and the next realized price. We provide in the appendix a screen shot of the experiment and the translated instructions in the supplementary material.

We follow the approach of Heemeijer et al. [2009] in deriving the pricing equation. The function connecting the prediction and the price is the following:

$$p_t = p_f + \frac{1}{1+r}(\bar{p}^e_{t,t} - p_f) + \epsilon_t \tag{1}$$

where r= 0.05 in all sessions and d is equal to 3.5 or 3.25 depending on the session. The fundamental price is computed as $p_f = \frac{d}{r}$. $\bar{p}^e_{t,t}$ is the average of the six one-step-ahead predictions submitted at the beginning of period t $\bar{p}^e_{t,t} = \frac{1}{6}\sum_{i=1}^{6} {}_ip^e_{t,t}$, and the term $\epsilon_t \sim N(0, 0.25)$ is an iid Normal shock. Looking at eq. (1), there is a strong positive feedback between expectations and market price, according to Hommes [2013].

Individual earnings at the end of each period depend on both, short and long-run prediction errors and are computed as ${}_i\pi_t = {}_i\pi^s_t + {}_i\pi^l_t$. We denote as ${}_i\pi^s_t$ the subject pay-off that depends on her short-run prediction error:

$${}_i\pi^s_t = \frac{250}{1+\beta} \qquad \text{with} \qquad \beta = \left(\frac{{}_ip^e_{t,t} - p_t}{2}\right)^2 \tag{2}$$

and as ${}_i\pi^l_t$ the subject pay-off that depends on long-run prediction error. We define ${}_i\pi^l_t = \sum_{j=1}^{t-1} {}_i\pi^l_{t-j,t}$, where ${}_i\pi^l_{t-j,t}$ represents the individual profit associated with the accuracy of the prediction submitted by subject i at the beginning of period $t-j$ about the asset price in period t, where $1 \leq j \leq t-1$. It is computed according to the following payment schedule:[5]

$${}_i\pi^l_{t-j,t} = \begin{cases} 25 & \text{if} \quad 0 \leq {}_i\delta_{t-j,t} \leq 5 \\ 12 & \text{if} \quad 5 < {}_i\delta_{t-j,t} \leq 10 \\ 5 & \text{if} \quad 10 < {}_i\delta_{t-j,t} \leq 15 \\ 0 & \text{otherwise} \end{cases}$$

where ${}_i\delta_{t-j,t} = |{}_ip^e_{t-j,t} - p_t|$. The final payment of each subject is the sum of pay-offs across all periods. We calibrated the parameters of the pay-off functions such that approximately $\max \sum_{t=1}^{20} {}_i\pi^s_t = \max \sum_{t=1}^{20} {}_i\pi^l_t$, in order to give to the subjects the same incentive to provide accurate predictions in the short as well as in the long-run. Note that subjects have an immediate feedback about the accuracy of their short-run predictions, while they experience a delay in evaluating the accuracy of their long-run predictions.

[4] It is important to underline that subjects attention do not decrease over time. To stress this aspect, we compute the average time taken by subjects to submit their predictions. Subjects take about 3 minutes to submit 20 predictions in the first period, and, at the end of the session, they take on average 40 seconds to submit just one predictions. We claim that each prediction, from the first to the last period, is a reasoned choice and it is not the result of an hasty choice.

[5] We used a pay-off mechanism similar to Haruvy et al. [2007].

The experiment involves 42 undergraduate students and it was conducted in the Laboratory of Experimental Economics at University Jaume I. Each session lasted approximatively 40 minutes and the average gain was 20 Euros.

We can compute the Rational Expectations Equilibrium (Lucas Jr [1978]) of our experiment. According to the law of motion of eq. (1), the REE implies that the realized price p_t should converge to the fundamental value with very small fluctuations due to the idiosyncratic shock ϵ_t. If we assume all subjects have rational expectations, their predictions[6] in each period t and **for each forecasting horizon** k should fluctuate around the constant fundamental value, i.e. ${}_ip^e_{t,t+k} \approx p_f$. This condition can be easily tested using our experimental data. It might be that the price as well as short-run expectations converge to the fundamental value, a case observed in the literature, however long-run expectations do not. This scenario is, in principle, not in line with REE in our setting with constant fundamental value. On the other hand, it might be that long-run expectations converge to the fundamental value, while the short-run ones together with the price dynamics do not. This hypothetical scenario would signal that subjects would consider the fundamental value as a kind of asymptotic equilibrium. Our experimental design allows to measure whether and how the price dynamics together with entire spectrum of expectations, short and long-run, converges to the fundamental value.

Note that in the eq. (1) we explicitly exclude the dependence of the price on long-run expectations. This allows us for a direct comparison of our results with the outcome of other LtFEs present in the literature. In principle, one might expect that, since long-run expectations do not enter in the price generating equation, they do not significantly impact the price dynamics. However, since subjects' pay-off does depend on the accuracy of their long-run expectations, we cannot exclude *a priori* that subjects use long-run expectations for some kind of inter-temporal hedging, influencing therefore the price dynamics. In this case, in fact, short-run predictions should be influenced by past long-run predictions. In this paper, we will show that such effect, if it exits, has a negligible impact on the price dynamics. Therefore, short-run expectations are independent of the elicitation of long-run expectations. In particular, we can study how past price dynamics is incorporated into the subjects' short-run expectations and whether and how short-run expectations and the past time series of prices influence long-run expectations. Do short- and long-run expectations follow the same pattern? If subjects learn to coordinate their short-run expectations, as shown in the learning to forecast literature, how their long-run expectations behave? Do they coordinate? Do they converge to the asset fundamental value?

2.2 Experimental results

Figure 1 shows individual short-run predictions and realized market prices for all groups. In the Appendix, Figures from 21 to 27 describe individual long-run predictions as well as the evolution of the market price for the 20 periods and for all 7 groups. As an example, Figure 2 shows the evolution over time of the market price together with individual long-run predictions of one of the groups.

From a visual inspection of those Figures, we observe two interesting regularities. The first regularity is an apparent pivotal role of the last realized price in the formation of expectations. This effect is particularly evident if we compare the whole set of expectations submitted in the first period, when subjects have no price available, to those submitted in period 2 (see panels (a) and (b) in Figures from 21 to 27 in the Appendix). In particular, the pivotal effect of the last realized price on the expectations' dynamics can be clearly identified if one considers the strong reduction in the heterogeneity of the entire spectrum of expectations submitted in the second period with respect to the first one; such reduction persists for several of the subsequent

[6] Across the

periods. The second interesting feature is that subjects' long-run expectations are persistently heterogeneous across periods and the heterogeneity appears to increase with the time horizon. An important question arises: Which is the origin of such heterogeneity? We are sure that it cannot be a difference in the subjects' information set, since the past price dynamics is common and common knowledge among subjects. We can conjecture that subjects have different *interpretations* of whether and how past prices influence future prices. In this respect, our experimental design allows to better measure the heterogeneity in the way individual subjects form their expectations using their available information as compared to other LtFEs in the literature, that are typically limited to one-period-ahead predictions, since we have a more comprehensive measures of their expectations.

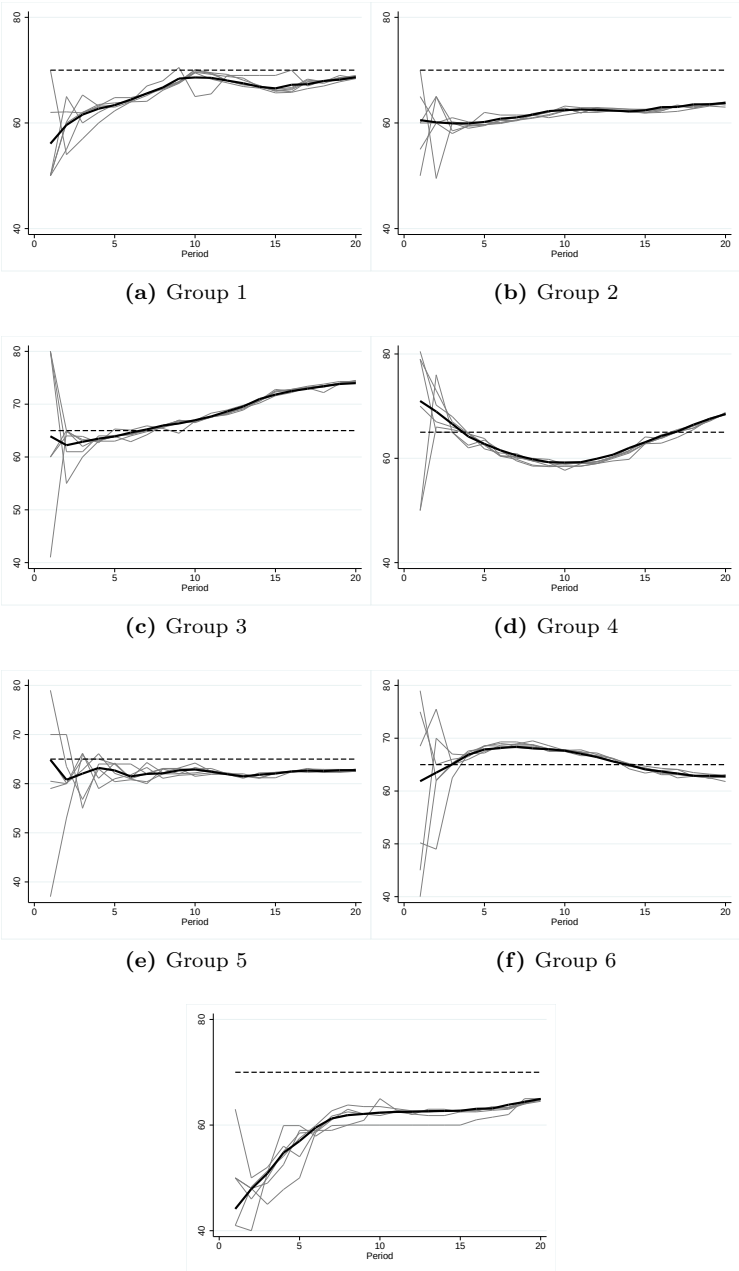
2.2.1 Coordination of short and long-run expectations

Given the strong positive feedback between the short-run predictions and market prices from eq. (1), each subject has to guestimate the expectations of the other subjects when submitting her short-run predictions. It exists, in fact, a strategic interaction among subjects: each subject has an incentive to coordinate her expectations around the others' expectations: subjects' expectations are strategic complements. We, therefore, expect a strong coordination of subjects' short-run predictions, in line with the literature (see for example Heemeijer et al. [2009]). What about long-run predictions? Whereas subjects have a direct incentive to coordinate their short-run expectations, their coordination motive for long-run expectations is more complex. When submitting their long-run predictions, the subject's task is to forecast at the beginning of period t the price at the end of period $t+k$, with $k > 0$. The price at the end of period $t+k$ depends on the subjects' short-run predictions submitted at the beginning of period $t+k$. Therefore they should guestimate, k-periods in advance, the short-run expectations of the other subjects at the beginning of period $t+k$. We should expect, then, a lower degree of coordination the longer is the forecasting horizon, given the increasingly degree of uncertainty in guestimating the future short-run behavior of the other subjects.

In order to measure the degree of coordination of subjects' expectations, we compute the standard deviation of their predictions submitted at a given period and for different time horizons. Figure 3 shows the average standard deviation of subjects' predictions submitted in period t for the price at the end of period $t+k$. In line with the LtFEs literature, we observe a fast coordination of subjects' short-run predictions. The heterogeneity of subjects' short-run predictions declines rapidly during the first 5 periods, to reach afterwards an almost stable value. Also for long-run predictions the degree of coordination increases over time. However, long-run expectations clearly need more time to reach the same coordination degree as short-run predictions. Moreover, we observe that the heterogeneity of subjects' expectations submitted in period t systematically increases with the time horizon, confirming our previous conjecture.

2.2.2 Pivotal role of the realized price

We have seen that subjects learn to coordinate their expectations: but, at which level do they coordinate? From the visual inspection of the data it seems that they coordinate using the last realized price as an anchor. The realized price, in fact, plays two important roles: on the one hand, it determines subjects' profit via their forecasting errors; on the other hand, the realized price is a public signal, carrying information on the others' expectations. Our conjecture is that, due to this double role, the realized price becomes an anchor for the coordination of expectations. The availability of prices provides to the subjects new information on the market mechanism, which immediately impacts, although in different degree, the entire spectrum of expectations.



(a) Group 1 **(b)** Group 2

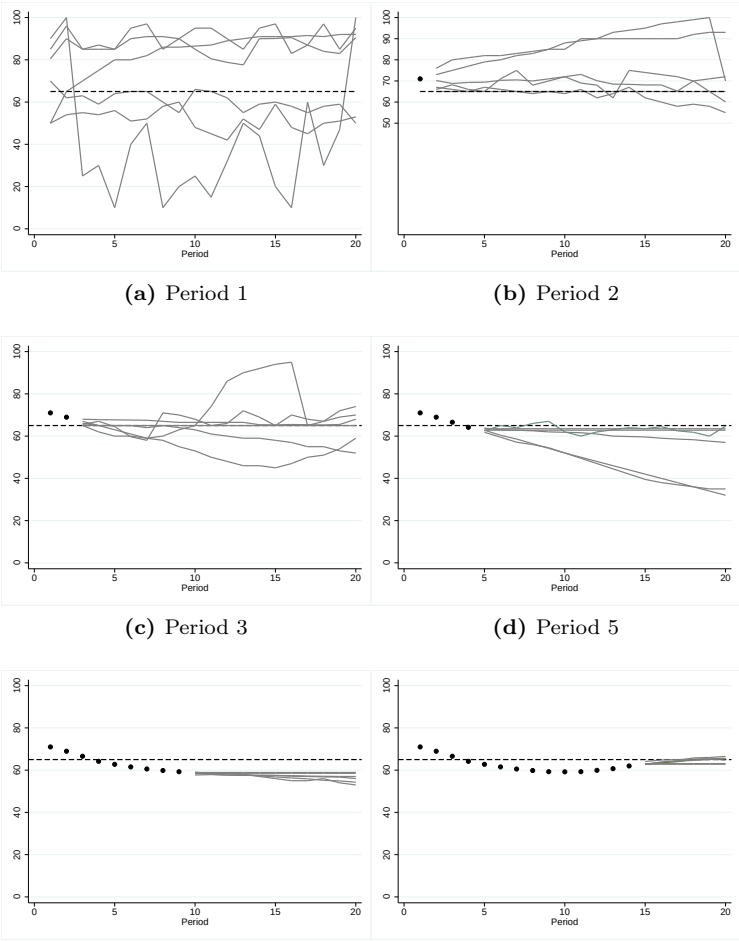
(c) Group 3 **(d)** Group 4

(e) Group 5 **(f)** Group 6

(g) Group 7

Figure 1: Realized price and individual short-run predictions of all groups. The black solid line is the market price, the grey lines are the individual one-step-ahead predictions and the dashed line represents the fundamental value.

Figure 3 shows the distribution of the correlation coefficients between the time series of previous realized prices $p_{t-1,t-1}$ and the individual predictions ${}_ip^e_{t,t+k}$ at different time horizons,



(a) Period 1 **(b)** Period 2

(c) Period 3 **(d)** Period 5

(e) Period 10 **(f)** Period 15

Figure 2: Realized prices (dots), the asset fundamental value (broken line) and all individual predictions (solid lines) submitted in periods 1, 2, 3, 5, 10 and 15 of Group 4 as a function of the time horizon.

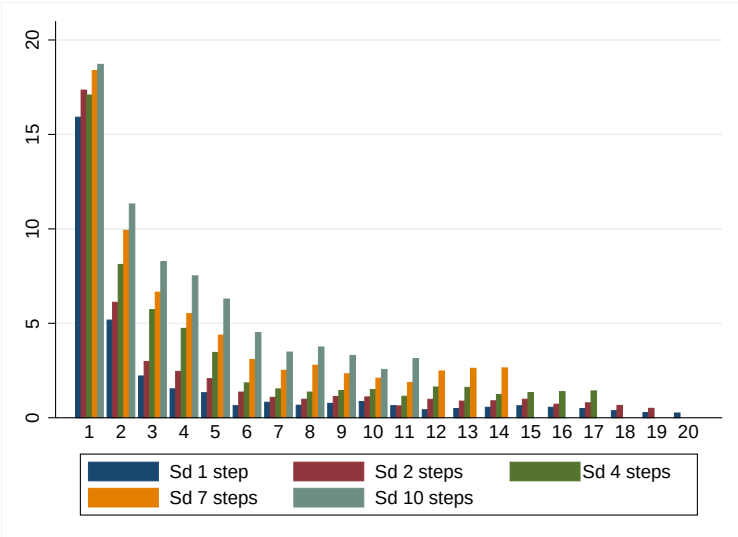


Figure 3: Coordination of expectations. For each period $t = 1, ..., 20$ is displayed the average standard deviation of the subjects' forecasts in period t for the price at the end of period $t + k$, where k is 0, 1, 3, 6 and 9.

represented as box-plots. We can see that the median of the correlation coefficient is decreasing with the time horizon starting from a value very close to 1. Even for a larger horizon ($k = 10$), the value of the correlation is significantly different from zero. It is clear that the last realized price becomes an anchor for the short-run predictions. Moreover, the realized price remains a stable anchor even for longer horizons, helping the subjects to reduce the uncertainty in guestimating the others' future short-run expectations.

2.2.3 Convergence of price and expectations to the REE

Subjects' learn to coordinate their expectations and the expectations are centred in the last realized price. Recall that, according to REE, the price and whole spectrum of expectations should converge to p_f independently of the time horizon. In line with the LtFE literature, Figure 5 shows that in our experimental markets there is no an immediate convergence of realized prices to the fundamental value. Apparently, in some cases prices exhibit a slow monotonically or oscillatory patter towards the fundamental value, whereas, in other cases the price seems to diverge[7]. So, the price, as an aggregate variable, does not converge to the REE.

The emerging patterns of the price dynamics and short-run expectations are very similar to those reported in other LtFEs eliciting solely short-run expectations Heemeijer et al. [2009]. This similarity indicates that the potential intertemporal hedging activity of the subjects, if it exists, has a negligible impact on both price dynamics and short-run expectations formation. Otherwise short-run expectations would be "closer" to long-run expectations. In the computational part of the paper we give a more detailed analysis on this issue.

If the price does no converge to the REE, do the individual expectations converge? In order to test whether individual expectations converge to the fundamental value, we compute the Relative Mean Square Error (RMSE) as the difference between the fundamenal value and

[7] A Wilcoxon test shows that the difference between observed prices and the fundamental value is statistical significant, except for group 5 with a p-value of 0.12.

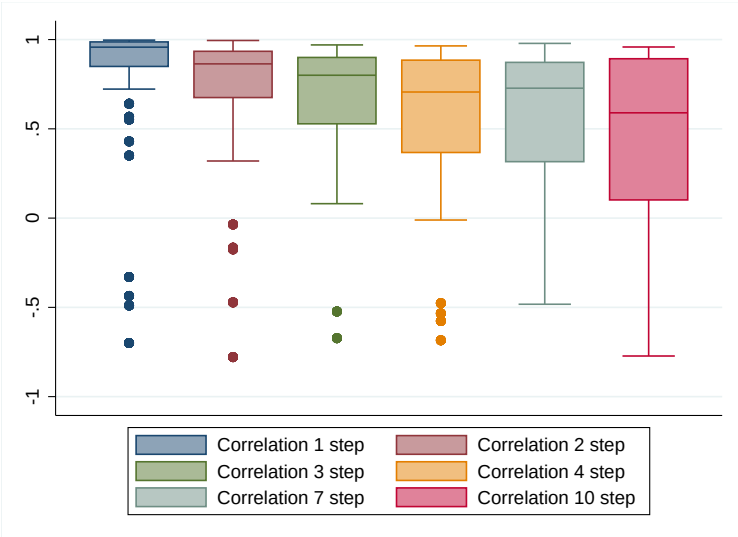
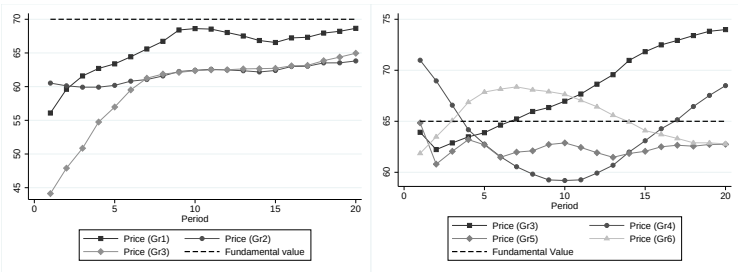


Figure 4: Box-plots of the correlation coefficients between the time series of prices and the individual subjects' expectations at different time horizons.



(a) Fundamental value equal to 70 **(b)** Fundamental value equal to 65

Figure 5: Convergence of Expectations: for each period $t = 1, ..., 20$ is displayed the RMSE, averaged across groups, of individual predictions from the fundamental value for different forecasting horizons $k = 1, 2, 4, 7, 10$.

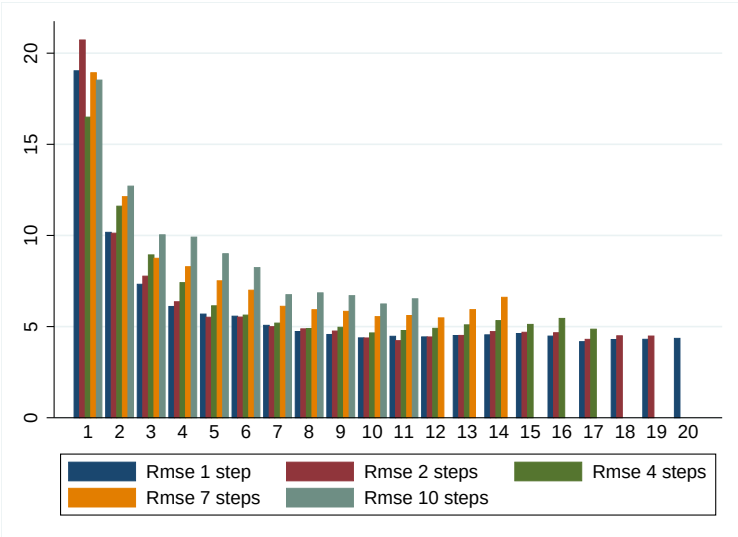


Figure 6: Average RMSE across groups as a function of the period and for different forecasting horizons.

individual predictions in period t for the price k-periods ahead:

$$RMSE_{t,k} = \sqrt{\frac{\sum_{i=1}^{6}(_i p^e_{t,t+k} - p_f)^2}{6}}$$

From Figure 6, it is evident that short-run expectations do not converge to the fundamental value, although the RMSE reduces over time. The same pattern is observed for long-run expectations. When comparing convergence for different time horizons from Figure 6, we observe that the RMSE increases marginally with the forecasting time horizon. As a first approximation, we can state that the degree of convergence of expectations is largely independent of the time horizon. It seems that the fundamental value is not the main determinant of the dynamics of short- and long-run expectations. Given our results, we can conclude then that the REE is not a good descriptor for subjects' expectations as well as price dynamics.

3 Computational Learning Approach

3.1 Learning Algorithms

In this section, we implement the EEA to test its ability in describing the stylized facts of the experimental data of our LtFEs with the target of reproducing the long-run expectations. We then compare the performances of the EEA as compared to the HSM. This exercise has two main goals: first, the goodness of fit of the EEA can be compared to the HSM, which is considered the benchmark. Second, by comparing the goodness of fit of HSM in describing the short-run expectations, we provide a quantitative criterion to support our conjecture that the elicitation of the long-term expectation does not affect the short-run expectations dynamics, because of the potential intertemporal hedging activity of the subjects trying to smooth their forecasting errors. This means that, if the dynamics of our experimental data is similarly replicated by the HSM as in the literature, we can conclude that the elicitation of long-run expectations has a

negligible effect on the price dynamics. So we can assume that long-run expectations dynamics reflect short-run expectation dynamics and the price evolution.

3.1.1 The Heuristic Switching Model

In the following, we list the four heuristics of the HSM as introduced in the original paper by Anufriev and Hommes [2012]. We briefly describe those rules adapting them to our notation. We label the four rules according to the index $h = 1, .., 4$, indicating the corresponding forecasting price as ${}_h p^e_{t,t}$. Note that the left sub-index h denotes now the heuristic instead of the subject. The heuristics are:

- Adaptive rule (ADA): it is a weighted average of the last prediction and the last realized price.

$$ {}_1 p^e_{t,t} = \alpha p_{t-1} + (1-\alpha) p^e_{t-1,t-1} \qquad \alpha = 0.65 \tag{3} $$

- Weak trend following rule (WTR): according to this rule, agents take into account the last realization of the market price and adjust their prediction extrapolating the market trend. The coefficient of proportionality is smaller than one, so the "weakness" in the name of the rule.

$$ {}_2 p^e_{t,t} = p_{t-1} + w(p_{t-1} - p_{t-2}) \qquad w = 0.4 $$

- Strong trend following rule (STR): it is structurally identical to the WTR. The difference is given by the weight assigned to the extrapolative parameter, in this case higher than 1.

$$ {}_3 p^e_{t,t} = p_{t-1} + s(p_{t-1} - p_{t-2}) \qquad s = 1.3 $$

- Learning and adjustment rule (LAA): the first part is the *time-dependent anchor* given by the average of the last observed price and the mean of the past prices. The second term of the equation represents the extrapolative term. Note here the unitary coefficient of the extrapolative term.

$$ {}_4 p^e_{t,t} = 0.5(p^{av}_{t-1} + p_{t-1}) + (p_{t-1} - p_{t-2}) $$

The learning mechanism is based on the relative profitability of each forecasting rule among the four fixed rules. Agents do not learn new rules and do not modify them neither. They rank the different rules and choose the one that better performed in the recent past. The switching mechanism is based on a performance measure $U_{h,t}$ that depends on the quadratic forecasting error. The performance measure $U_{h,t}$ is given by:

$$ U_{h,t-1} = -\left(p_{t-1} - {}_h p^e_{t-1,t-1}\right)^2 + \eta\, U_{h,t-2}, \qquad h = 1, ..., 4 $$

where the parameter $0 \leq \eta \leq 1$ represents the "memory" of agents, meaning the weight assigned to past errors. We set $\eta = 0.7$, following Hommes [2013].

To generate the market price, it is necessary to compute the proportion of agents that use each one of the heuristics. We compute the proportion of agents choosing each rule, i.e. $n_{h,t}$,

using the discrete choice model with asynchronous updating as in Diks and Van Der Weide [2005]. The updating equations are:

$$n_{h,t} = \delta \; n_{h,t-1} + (1-\delta) \; \frac{\exp(\beta \; U_{h,t-1})}{Z_{t-1}} \; , \tag{4}$$

$$Z_{t-1} = \sum_{h=1}^{4} \exp(\beta \; U_{h,t-1}) \; , \tag{5}$$

where $0 < \delta \leq 1$ denotes the share of agents that update their choice; the parameter $\beta \geq 0$ represents the intensity of choice and it determines the switching speed to the most successful rule. Z_{t-1} is a normalization factor. We consider $\delta = 0.9$ and $\beta = 0.4$ as in Hommes [2013]. We compute the expected price as a weighted average across the different expectations given by the four rules:

$$\bar{p}^e_{t,t} = \sum_{h=1}^{4} n_{h,t-1} \; {}_h p^e_{t,t} \; .$$

and we insert this value in eq. (1) to compute the resulting market price.

The literature of LtFE has shown that the HSM can fairly well reproduce the properties of short-run expectations in different experimental settings. However, the HSM, at least in its present form, cannot be a suitable model to reproduce the observed proprieties of long-run expectations. Note, in fact, that three out of four rules (heuristics 2, 3 and 4) depend exclusively on past prices. As a consequence, when using those rules, all artificial agents will share exactly the same prediction for future prices. It means that the agents' long-run predictions using those rules are constant values, independent of the forecasting horizon and the agent identity:

$${}_h p^e_{t,t+k} = {}_h p^e_{t,t} \qquad h = 2, 3, 4$$

The variable ${}_h p^e_{t,t}$ is a function of the past price, whose value depends on the particular heuristic considered. The only rule that can generate heterogeneous predictions across agents is the first one (ADA), since it depends on the past individual short-run prediction. If we iterate this rule, we obtain:

$${}_i p^e_{t,t+k} = {}_i p^e_{t-1,t-1} (1-\alpha)^k + p_{t-1} \; (1 - (1-\alpha)^{k+1}) \; . \tag{6}$$

where k denotes the forecasting horizon[8]. Note that we have replaced the index $h = 1$ with the identity of the agent i, since the k-steps-ahead prediction depends on the agent identity through the short-run expectation submitted in period $t-1$, i.e. ${}_i p^e_{t-1,t-1}$. Eq. (6) implies that, independently of the subjects' one-step-ahead prediction, long-run predictions exponentially converge to the last realized price p_{t-1}. Already for a forecasting horizon approximatively equal to 3, we obtain essentially ${}_i p^e_{t,t+k} \approx p_{t-1}$. Such behaviour is in contradiction with the observed data. According to eq. (6), in fact, the heterogeneity of subjects' long-run expectations vanishes, instead of increasing as in Figure 3. In its present form, therefore, the HSM cannot be used to describe in a meaningful way the subjects long-run predictions in our LtFE experiment, since it systematically underestimates the heterogeneity of subjects' long-run expectations.

[8] For $k = 0$ we obtain the one-step-ahead expression as in eq. (3).

3.1.2 The Exploration-Exploitation Algorithm

Given the drawbacks of the HSM in reproducing long-run expectation properties, we propose the Exploration-Exploitation Algorithm to describe both, short and long-run expectations and the corresponding price dynamics. Within the EEA we do not impose a finite set of decision rules. Instead, agents choose their actions from a distribution of feasible actions whose main determinants, mean and variance, evolve adaptively based on past price dynamics. In particular agents' learning is based on the ability to adapt the range of variation of their actions as they acquire more information when observing price dynamics.

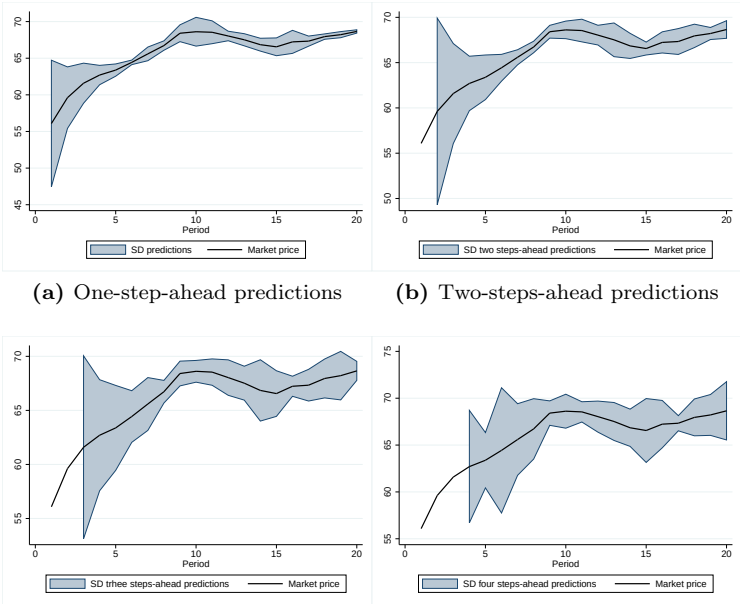
In order to adapt the EEA to our experimental setting, we consider the subjects' predictions for future prices at different time horizons to be the *actions* of artificial agents. The *exploration space* is the considered range of prices divided in a discrete grid of 100 steps, so that each step corresponds to a feasible action. Each action will have a different probability to be chosen, depending on its past performance. Moreover, the exploration space adapts over time, i.e. the range of actions changes according to the realized price in the last period and to its past fluctuations. In particular, the range is centred in the last observed price (p_{t-1}) and the amplitude of the range is proportional to the standard deviation of the past prices. In other words, the anchor is the last price and the adjustment is made according to the observed volatility of the time series of market prices. The mechanics of the algorithm is based on the experimental evidence suggesting that subjects coordinate their predictions anchoring them around the last observed price. The exploration space adapts to the new market conditions, typically reducing over time.

As we have shown when describing the experimental results, the heterogeneity among individual predictions reduces rapidly in the first few periods. Within the framework of the EEA, we can interpret this dynamics as the fact that subjects tend to explore a large space in the first periods, when they possess few pieces of information to characterize the price dynamics, while they prefer to adopt the exploitation strategy afterwards, when the "cost of exploration" is relatively high.[9] In order to better illustrate our interpretation of the EEA, we show in Figure 7 an example of how the exploration space reduces over time. We compute the standard deviation of the subjects' individual predictions from the experimental data each period for the price one, two, three and four steps-ahead. We observe that, in early periods, subjects count with a few pieces of information to have a precise idea about the price evolution and therefore the range of their short and long-run predictions is wider. After few periods of learning, subjects tend to play the exploitation strategy, i.e. to submit forecasts narrowly centred in p_{t-1}. This is particularly true for the one-step-ahead predictions for which subjects receive an immediate feedback Colasante et al. [2016].

Let us formalize the mathematical description of the EEA algorithm. All agents have a set of $n = 100$ feasible actions denoted as $\mathsf{A}_t = \{a_{1t}, a_{2t}, ..., a_{nt}\}$ where $a_{1t} < a_{2t} < ... < a_{nt}$, that is common to all agents, where a_{jt} denotes a single element in A_t, being $j \in \{1, ..., 100\}$. The set A_t changes every period depending on the last market price and the standard deviation of past prices. The range of the set is given by $(p_{t-1} - 5\sigma_p,\ p_{t-1} + 5\sigma_p)$ where σ_p is the standard deviation of the last three market prices. Note that $a_{1t} = p_{t-1} - 5\ \sigma_p$ and $a_{100t} = p_{t-1} + 5\ \sigma_p$.

At the beginning of period t, agent i selects an action ${}_i\tilde{a}_t = {}_ip^e_{t,t}$ from A_t, that corresponds to the agent i's expected price at the end of period t, i.e. the one-step-ahead prediction. Besides the one-step-ahead predictions, agent i chooses three more actions ${}_i\tilde{a}^k_t = {}_ip^e_{t,t+k}$ from the corresponding set of actions $\mathsf{A}^k_t = \{a^k_{1t}, a^k_{2t}, ..., a^k_{nt}\}$, where $k \in \{1, 2, 3\}$. Each one of these actions represents the agent i's long-run expectations up to four-steps-ahead. Also the elements of those sets change every period. In particular, the range of A^k_t is centred on p_{t-1}, as in the case of

[9] If subjects' short-run expectations are coordinated in a narrow interval, as we observed after period 5 in our experimental data, the opportunity cost of deviating becomes high.



(a) One-step-ahead predictions **(b)** Two-steps-ahead predictions

(c) Three-steps-ahead predictions **(d)** Four-steps-ahead predictions

Figure 7: The continuous line represents the market price and the shadow area refers to one standard deviation of individual predictions. The data refers to Group 1.

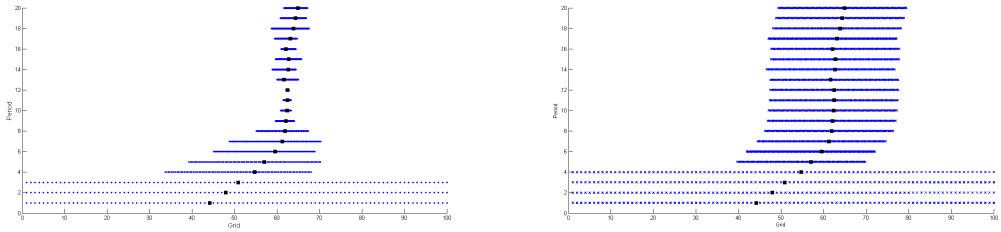
short-run predictions, but, differently from A_t, its amplitude is constant over time and does not depend on the price evolution. Since the long-run profit function used in the experimental design is a step function, we assume that the maximum range for the agents long-run predictions is 15.[10] The range for the actions that belong to A_t^k is therefore $(p_{t-1}-15, p_{t-1}+15)$.[11] As an illustrative example, Figure 8 shows in panel (a) how the range of set A_t (short-run predictions) reduces over time, and in panel (b), how the range of set A_t^k (long-run predictions) remains constant over time.

Once all agents choose their actions the market price is computed according to eq. (1). Agents then evaluate the performance of all feasible actions using a fitness function. We introduce two different measures: the first one to evaluate the actions in the set A_t, that refers to the agents' short-run predictions, denoted as V_t; the second one to evaluate the actions in the sets A_t^k that refer to the agents' long-run predictions, denoted as V_t^k. The main difference is that we consider the quadratic distance between the agent's action and the realized price to evaluate short-run predictions, while we consider the absolute distance to evaluate the subject's long-run predictions. This difference is introduced in order to replicate the profit functions we use in the experiment. The value of the fitness measures are:

$$ {}_i\mathsf{V}_{j,t} = -(p_t - {}_ia_{jt})^2 + \phi^s \; {}_i\mathsf{V}_{j,t-1} \; , \tag{7} $$

[10]Note that, following the payment schedule used to reward the subjects long-run expectations in the experiment, if the absolute difference between the price and the long-run prediction is higher than 15, the profit is equal to zero.

[11]The dynamics of the range of the exploration space can be generalize in many directions to include an individual dependence, a higher number of past prices to compute the standard deviation etc...



(a) Short-run predictions (A_t) **(b)** Long-run predictions (A_t^4)

Figure 8: Evolution of the range of the agents' decision sets (one-step-ahead and four-steps-ahead predictions) for different periods.

$$ {}_i\mathsf{V}_{j,t}^k = -|p_t - {}_ia_{jt}^k| + \phi^k \; {}_i\mathsf{V}_{j,t-1}^k \tag{8} $$

The parameters ϕ^s and ϕ^k represent the weight assigned to the past forecasting errors. We set $\phi^s = 0.3$ and $\phi^k = 0.5$.[12]

We then introduce a probability distribution associated to the set A_t for each agent i. Essentially, all agents have the same set of actions, however the fitness measures and the associated probability distributions differ among agents depending on the individual past performance. Note the dependence in eqs. (7) and (8) of the measures on the individual agent's identity. For the choice of short-run predictions, let ${}_ip_{jt}$ be the probability that agent i selects action ${}_ia_{jt}$ from the set A_t, such that $0 \leq {}_ip_{jt} \leq 1$ and $\sum_{j=1}^{100} {}_ip_{jt} = 1$. The probability to select action ${}_ia_{jt}$ is given by:

$$ {}_ip_{jt} = \frac{\exp(\gamma \cdot {}_i\mathsf{V}_{j,t})}{\sum_{j=1}^{100} \exp(\gamma \cdot {}_i\mathsf{V}_{j,t})} \; , \tag{9} $$

where $\gamma \in [0, \infty)$ represents the intensity of choice as in the HSM. We set $\gamma = 0.2$. For the choice concerning the long-run predictions we introduce analogue functions to compute the probabilities attached to each one of the actions referring to the long-run expectations:

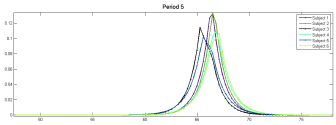
$$ {}_ip_{jt}^k = \frac{\exp(\gamma \cdot {}_i\mathsf{V}_{j,t}^k)}{\sum_{j=1}^{100} \exp(\gamma \cdot {}_i\mathsf{V}_{j,t}^k)} \tag{10} $$

According to the probability distributions in eqs. (9) and (10) in each period t each agent randomly chooses four actions ${}_i\tilde{a}_t$ and ${}_i\tilde{a}_t^k$, where $k \in \{1, 2, 3\}$ (one short-run prediction and three long-run predictions). Figures 9 and 10 illustrate an example of the probability distributions associated to each action for the different agents and time horizon, specifically one and four-steps-ahead actions at different time periods of a simulation. We can clearly observe a certain degree of heterogeneity among agents and time horizons, which depends on the past history of the agents' choices.

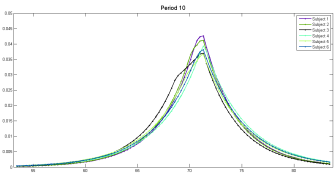
3.2 Simulation results

After a comprehensive description of the two algorithms, in this section we compare their performance in replicating the experimental results.

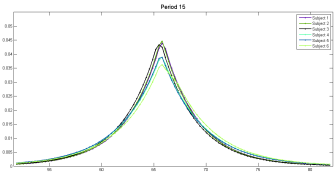
[12] We use two different values for these parameters following the results of Colasante et al. [2016]. Estimating a panel regression model, they show that subjects assign a higher weight to their previous forecasts when forming their the long-run expectations compared to the weight assigned to past forecast when forming their short-run expectations.



(a) Period 5

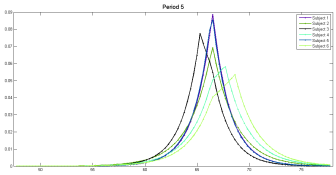


(b) Period 10

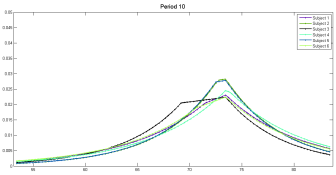


(c) Period 15

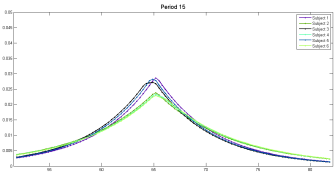
Figure 9: Examples of the probability distributions of one-step-ahead prediction for each agent for three different periods.



(a) Period 5

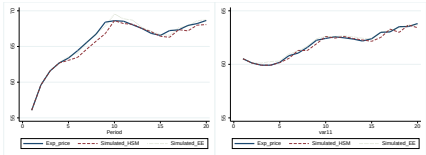


(b) Period 10



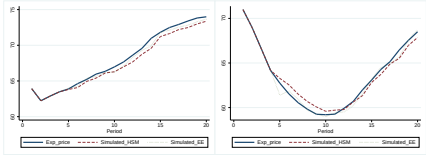
(c) Period 15

Figure 10: Examples of the probability distributions of four-step-ahead prediction for each agent for three different periods.



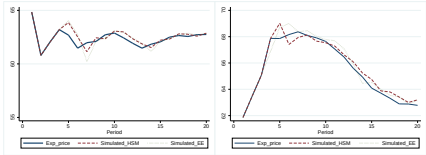
(a) Group 1

(b) Group 2



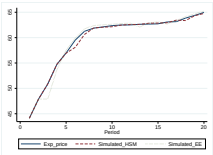
(c) Group 3

(d) Group 4



(e) Group 5

(f) Group 6



(g) Group 7

Figure 11: Simulation results based on the HSM and the EEA. The continuous black line is the experimental market price, the blue line is the market price generated using the HSM and the dashed grey line is the market price generated using the EEA.

3.2.1 Comparing the HSM and EEA to describe short-run expectations

We calibrate the HSM using the experimental market price in the first three periods. We need, in fact, three prices to compute the expectations in some of the heuristics. At the beginning of the simulation, we assign the same weight to each rule, i.e. $n_{h,1} = 0.25, \forall h$. Starting from period three we compute the fitness measure and weight $n_{h,3}$ associated to each heuristic. For the subsequent periods we iterate the algorithm detailed in the previous section.

We calibrate then the EEA using the experimental individual predictions and the first three realized prices, since the range of the action sets depend on the three past realized prices. To compute the market price using the EEA, in each period we make six (the number of subjects in the group) independent draws from the six different distributions of the short-run expectations. Once we have the individual predictions we use eq. (1) to compute the market price and iterate then the algorithm.

As we can see from Figure 11, the two algorithms are able to replicate the experimental price dynamics. In order to have a more quantitative comparison, Table 1 presents the Mean Squared Error (MSE) for the two algorithms. We observe that the errors with respect to the experimental

data are of the same order of magnitude. In particular, the EEA shows a better performance in half of the groups. So we can safely conclude that the EEA achieves similar results as the HSM in replicating the price dynamics.

Table 1: Mean Squared Error of the HSM and EEA in describing the time series of experimental market prices.

	HSM	EEA
Group 1	0.37	**0.29**
Group 2	**0.06**	**0.06**
Group 3	0.44	**0.13**
Group 4	0.33	**0.13**
Group 5	**0.22**	0.38
Group 6	0.19	**0.16**
Group 7	**0.15**	0.77

The next step is to check the performance of the two algorithms in replicating coordination and convergence of individual expectations. To analyse the coordination of expectations, we compute the standard deviation among the individual one-step-ahead predictions in each group of the EEA. In the case of HSM, the procedure is not straightforward since the HSM does not replicate the individual predictions, but instead the frequencies in the use of the heuristics across a large population of agents. So, we compute the standard deviation of the expectations considering the frequencies in the use of heuristics as given by eq. (4). Figure 12 shows the comparison of the two algorithms with respect to the experimental data. Note that the first three periods coincide with the experimental data because of the calibration procedure of the two algorithms. The qualitative behavior of the experimental data is well captured by the simulated results of the tow algorithms, without any systematic difference. Figure 13 shows a good agreement between simulated and experimental data.

Comparing the results of Table 1 to the performance of the HSM reported in other papers (Anufriev and Hommes [2012], Anufriev et al. [2013b] and Bao et al. [2012], we can see that the MSEs are of the same order of magnitude. Given that, filtering our experimental prices with the HSM, we obtain essentially similar aggregate results, we can state that eliciting long-run expectations does not affect the price dynamics with respect the baseline LtFE, i.e. when eliciting just one-step-ahead expectations. This conclusion allows us to discard the existence of a significant impact of inter-temporal hedging strategies followed by the subjects. Similar behavior has been reported by Hanaki et al. [2016] in an experimental setting with trading studied by Haruvy et al. [2007].

3.2.2 Long-run expectations in EEA

To the best of our knowledge, this is the first attempt in the literature on LtFEs to reproduce individual long-run expectations using a learning algorithm. Despite the fact that in the experiment we elicit the expectations for the whole time horizon, we replicate the individual expectations up to four steps-ahead. Our choice represents a good compromise between considering the whole time-span and having a sufficient statistics to analyse the properties of the experimental data as a function of the time horizon.

We analyse the performance of the EEA in replicating the main properties of the experiment: (i) the time series of the individual expectations, (ii) the role of the realized price in forming expectations, (iii) the coordination of expectations as a function of the time horizon and, finally,

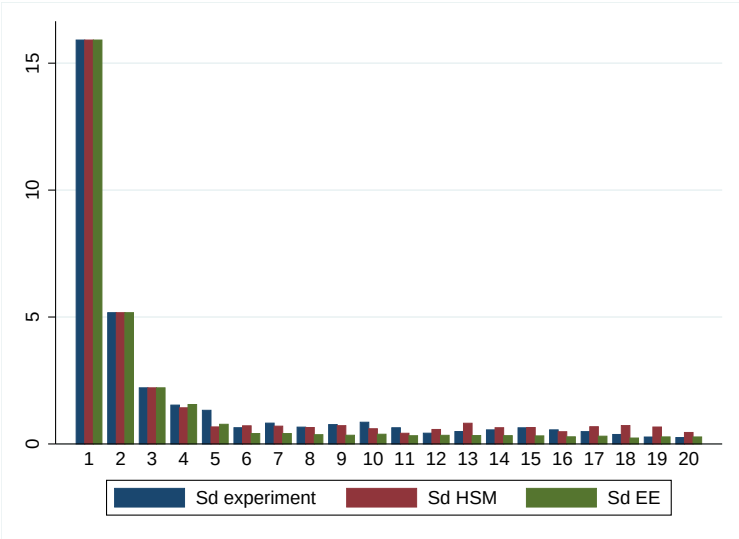


Figure 12: Coordination: average across groups of the standard deviation of one-step-ahead predictions per period. Blue bars refer to experimental data, red bars refer to HSM and green bars refer to EEA. The histograms of the simulated data are an average over 100 Monte Carlo iterations of the EEA and HSM.

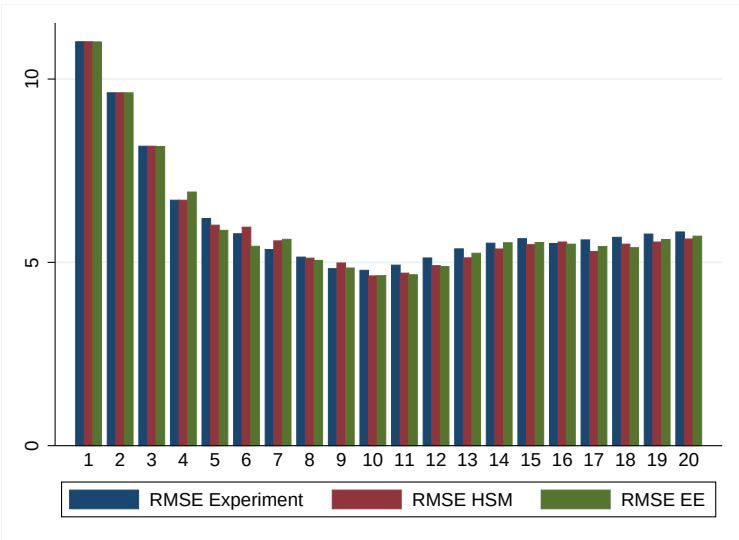


Figure 13: Convergence to the fundamental value: RMSE of the distance between one-step-ahead predictions and the fundamental value per period. Blue bars refer to experimental data, red bars refer to HSM and green bars refer to EEA. The histograms of the simulated data are an average over 100 Monte Carlo iterations of the EEA and HSM.

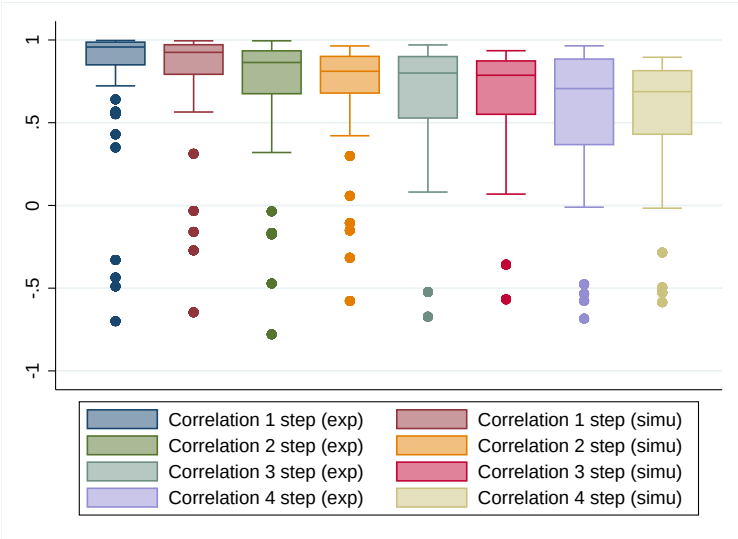


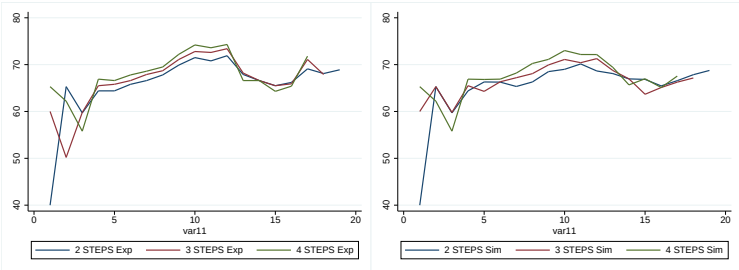
Figure 14: Box-plots of the correlation coefficients between the time series of prices and the individual subjects' expectations at different time horizons compared to the corresponding values for the simulated data.

(iv) the convergence of long-run expectations to the fundamental value. Note that here we do not have an aggregate variable like the market price, but just individual long-run predictions. Furthermore, we cannot compare the performance of our algorithm to an alternative learning specification, since,as previously illustrated, the HSM, in its present form can not be implemented to model individual long-run expectations.

Figures 15, 16, 17 show the simulated individual long-run expectations confronted to the experimental data of three representative subjects belonging to groups 1, 4 and 5. Each line represents the predictions submitted by subject i in period t for the price at the end of the following 2, 3 and 4 periods ahead. In other words, each series represents respectively: ${}_ip^e_{t,t+1}$ (blue line), ${}_ip^e_{t,t+2}$ (red line) and ${}_ip^e_{t,t+3}$ (green line) $\forall t$. The EEA describes fairly well the individual long-run predictions. However, the simulated time series of expectations exhibit a "rougher" path compared to the smooth time series observed in the experimental data, We conjecture that the experimental expectations possess a higher degree of time-correlations among the predictions at different time horizons than what results from the EEA. Consider that in the EEA, the three predictions (${}_i\tilde{a}^1_t$, ${}_i\tilde{a}^2_t$ and ${}_i\tilde{a}^3_t$) are independent draws from the corresponding probability distributions of actions ${}_ip^k_{jt}$, that evolve independently without any explicit conditional dependence. Our results speak in favour of the existence of a conditional dependence among the long-run expectations, that can be implemented in future modifications of the EEA or alternatively in a modified version of the HSM.

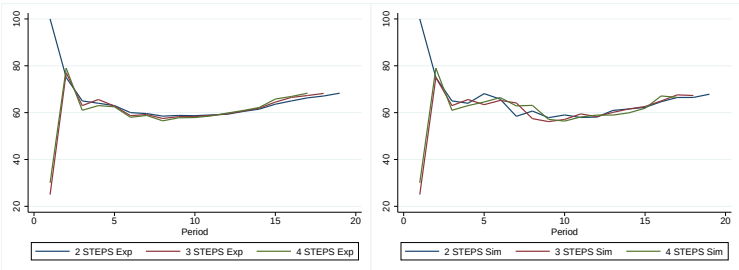
The EEA assumes that agents actions are anchored around the last realized price. The good performance of the algorithm leads us to infer that the last price turns out to be a strong anchor for long-run expectations (at least up to four period ahead) even if the feedback given by the profit comes with a delay. Figure 14 shows a stable correlation between the market price and the expectations at different time horizons. The EEA is structurally built on using the last price as an anchor. However, it is noticeable how the EEA captures extremely well the subjects' heterogeneity in incorporating the last realized price in their expectations.

Figure 18 displays the standard deviation of individual predictions for the price 2, 3 and 4



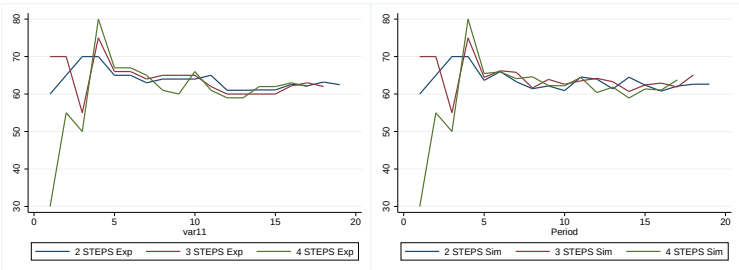
(a) Experimental data **(b)** Simulated data

Figure 15: Illustrative example of a time series of long-run predictions of an individual subject in group 1 for different time horizons (2, 3 and 4 steps ahead) as compare to the simulated data.



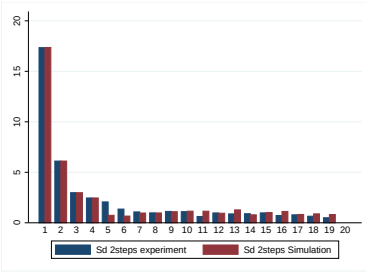
(a) Experimental data **(b)** Simulated data

Figure 16: Illustrative example of a time series of long-run predictions of an individual subject in group 4 for different time horizons (2, 3 and 4 steps ahead) as compared to the simulated data.

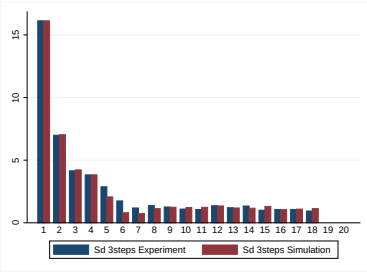


(a) Experimental data **(b)** Simulated data

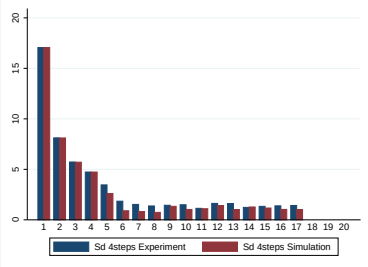
Figure 17: Illustrative example of a time series of long-run predictions of an individual subject in group 5 for different time horizons (2, 3 and 4 steps ahead) as compared to the simulated data.



(a) Two-steps-ahead



(b) Three-steps-ahead



(c) Four-steps-ahead

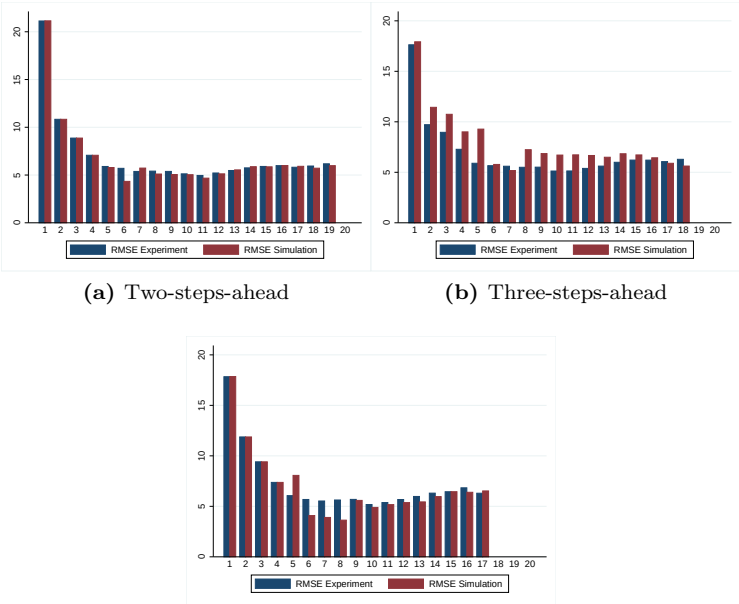
Figure 18: Coordination of individual long-run expectations comparing simulated and experimental data. The histograms of the simulated data are an average over 100 Monte Carlo iterations of the EEA.

periods ahead. It shows that the degree of coordination resulting from the simulated data is fairly close to the degree of coordination of experimental data. Additionally, the EEA is able to reproduce the more persistent heterogeneity observed for the long-run predictions as compared to the degree of coordination of the one-step-ahead predictions.

In order to measure the performance of the EEA to reproduce the convergence of individual expectations to the fundamental value, in Figure 19, we compare the RMSE of simulated and experimental data. We are able to conclude that the EEA is able to replicate also the lack of convergence.

4 Conclusion

In this paper we present the results of a LtFE where we elicit subjects' short and long-run expectations about the future prices of a financial asset, generalizing the usual focus of the existing literature on LtFEs. Our experimental results generalize previous findings showing that subjects' expectations are not consistent with the REE, neither in the short not in the long-run. The market price dynamics exhibits in some cases an oscillatory or monotonically tendency to converge to the asset fundamental value. We have shown that the elicitation of subjects' long-run expectations does not to have a significant impact on the price dynamics. On the contrary, we observe a clear influence of price dynamics on the formation and evolution of subjects' long-run expectations. Subjects learn to coordinate the entire spectrum os expectations. In particular, subjects anchor their short and long-run expectations around the last realized price. We have observed that short-run expectations coordinate around the market price much faster than the



(a) Two-steps-ahead **(b)** Three-steps-ahead

(c) Four-steps-ahead

Figure 19: Convergence of individual long-run predictions to the fundamental value for different time horizons comparing simulated and experimental data. The histograms of the simulated data are an average over 100 Monte Carlo iterations of the EEA.

long-run expectations, due to a higher degree of uncertainty faced by a subject when predicting the future behavior of the others.

In the second part of the paper, we introduce an adaptive learning algorithm in order to reproduce individual short and long-run expectations: the Exploration-Exploitation Algorithm. Such algorithm incorporates the bounded rational behavior of subjects by assuming that their expectations are centred on the last observed price and their range varies according to the most recent price fluctuations. We can surely cast our algorithm into the well-known anchor and adjustment behavioral framework. In order to evaluate the goodness of fit of our algorithm, we compare it to the well-established Heuristic Switching Model proposed by Anufriev and Hommes [2012] to model short-run expectations and market price dynamics. The computational part of the paper shows that the two learning algorithms perform equally well in describing the short-run dynamics of the experimental data. Although structurally different, the two algorithms share the same basic behavioral principle: anchor and adjustment. The fact that both satisfactorily describe the experimental data signals that subjects follow a similar general heuristic principle. Additionally, the good performance of the EEA to reproduce the long-run expectations dynamics generalizes and reinforces such conclusion.

Based on our experimental and computational analysis, it seems that the subjects condition their long-run expectations on shorter-term expectations. Such feature can possibly be cast in a more general version of HSM and/or EEA, which should incorporate a conditional dependence of expectations on past prices and past individual expectations. Future research will focus on the incorporation of more complex dependence of current expectations on past expectations in order to take into account the empirical evidence of our experimental setting. Additionally, we can study the formation of short and long-run expectations under more general settings including

other information sources, such as aggregate information on subjects' long-run expectations, public announcements of policy measures (monetary policies with or without targeted level of inflations) or future changes of the fundamentals.

Acknowledgement

The authors are grateful for funding the Universitat Jaume I under the project P11B2015-63 and the Spanish Ministry Science and Technology under the project ECO2015-68469-R.

Compliance with Ethical Standards

Funding: This study is funded by the Universitat Jaume I under the project P11B2015-63 and the Spanish Ministry Science and Technology under the project ECO2015-68469-R. The authors declare that they have no conflict of interest.

References

Mikhail Anufriev and Cars Hommes. Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. *American Economic Journal: Microeconomics*, 4(4):35–64, 2012.

Mikhail Anufriev, Cars Assenza, Tiziana and, and Domenico Massaro. Interest rate rules and macroeconomic stability under heterogeneous expectations. *Macroeconomic Dynamics*, 17(08): 1574–1604, 2013a.

Mikhail Anufriev, Cars H Hommes, and Raoul HS Philipse. Evolutionary selection of expectations in positive and negative feedback markets. *Journal of Evolutionary Economics*, 23(3):663–688, 2013b.

Masahiro Ashiya. Testing the rationality of japanese gdp forecasts: the sign of forecast revision matters. *Journal of economic behavior & organization*, 50(2):263–269, 2003.

Tiziana Assenza, Te Bao, Cars Hommes, and Domenico Massaro. Experiments on expectations in macroeconomics and finance. In *Experiments in macroeconomics*, pages 11–70. Emerald Group Publishing Limited, 2014.

Tiziana Assenzaa, Peter Heemeijerc, Cars Hommesb, and Domenico Massarob. Individual expectations and aggregate macro behavior. Technical report, 2011.

Peter Auer, Nicolo Cesa-Bianchi, and Paul Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2-3):235–256, 2002.

Te Bao and Li Ding. –nonrecourse mortgage and housing price boom, bust, and rebound. *Real Estate Economics*, 2015.

Te Bao, Cars Hommes, Joep Sonnemans, and Jan Tuinstra. Individual expectations, limited rationality and aggregate outcomes. *Journal of Economic Dynamics and Control*, 36(8):1101–1120, 2012.

Te Bao, John Duffy, and Cars Hommes. Learning, forecasting and optimizing: An experimental study. *European Economic Review*, 61:186–204, 2013.

William A Brock and Cars H Hommes. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic dynamics and Control*, 22(8-9):1235–1274, 1998.

Sean D Campbell and Steven A Sharpe. Anchoring bias in consensus forecasts and its effect on market prices. *Journal of Financial and Quantitative Analysis*, 44(02):369–390, 2009.

Benoit Cœuré. Monetary policy in the crisis – confronting short-run challenges while anchoring long-run expectations. Technical report, Speech by Benoit Cœuré, Member of the Executive Board of the ECB at the Journées de l' AFSE 2013, 2013.

Annarita Colasante, Simone Alfarano, Eva Camacho-Cuena, Mauro Gallegati, et al. Long-run expectations in a learning-to-forecast experiment. Technical report, 2016.

Camille Cornand et al. Does inflation targeting matter? an experimental investigation. *An Experimental Investigation (October 17, 2013)*, 2013.

Cees Diks and Roy Van Der Weide. Herding, a-synchronous updating and heterogeneity in memory in a cbs. *Journal of Economic dynamics and control*, 29(4):741–763, 2005.

Ippei Fujiwara, Hibiki Ichiue, Yoshiyuki Nakazono, and Yosuke Shigemi. Financial markets forecasts revisited: Are they rational, stubborn or jumpy? *Economics Letters*, 118(3):526–530, 2013.

Gabriele Galati, Peter Heemeijer, and Richhild Moessner. How do inflation expectations form? new insights from a high-frequency survey. 2011.

Refet S Gurkaynak, Brian Sack, and Eric Swanson. The sensitivity of long-term interest rates to economic news: evidence and implications for macroeconomic models. *The American Economic Review*, 95(1):425–436, 2005.

Nobuyuki Hanaki, Eizo Akiyama, and Ryuichiro Ishikawa. A methodological note on eliciting price forecasts in asset market experiments. 2016.

Ernan Haruvy, Yaron Lahav, and Charles Noussair. Traders' expectations in asset markets: experimental evidence. *The American Economic Review*, 97(5):1901–1920, 2007.

Peter Heemeijer, Cars Hommes, Joep Sonnemans, and Jan Tuinstra. Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation. *Journal of Economic dynamics and control*, 33(5):1052–1072, 2009.

Cars Hommes. *Behavioral rationality and heterogeneous expectations in complex economic systems*. Cambridge University Press, 2013.

Cars Hommes and Thomas Lux. Individual expectations and aggregate behavior in learning-to-forecast experiments. *Macroeconomic Dynamics*, 17(02):373–401, 2013.

Cars Hommes, Joep Sonnemans, Jan Tuinstra, and Henk van de Velden. A strategy experiment in dynamic asset pricing. *Journal of Economic Dynamics and Control*, 29(4):823–843, 2005.

Carsien Harm Hommes. Financial markets as nonlinear adaptive evolutionary systems. 2001.

Dimitris E Koulouriotis and A Xanthopoulos. Reinforcement learning and evolutionary algorithms for non-stationary multi-armed bandit problems. *Applied Mathematics and Computation*, 196(2):913–922, 2008.

Robert E Lucas Jr. Asset prices in an exchange economy. *Econometrica: Journal of the Econometric Society*, pages 1429–1445, 1978.

Charles F Manski. Measuring expectations. *Econometrica*, 72(5):1329–1376, 2004.

Ramon Marimon, Stephen E Spear, and Shyam Sunder. Expectationally driven market volatility: an experimental study. *Journal of Economic Theory*, 61(1):74–103, 1993.

Yoshiyuki Nakazono. Heterogeneity and anchoring in financial markets. *Applied Financial Economics*, 22(21):1821–1826, 2012.

Amos Tversky and Daniel Kahneman. Judgment under uncertainty: Heuristics and biases. *science*, 185(4157):1124–1131, 1974.

Michael Woodford. Monetary policy in the information economy. Technical report, National Bureau of Economic Research, 2001.

A Screenshot and Instructions

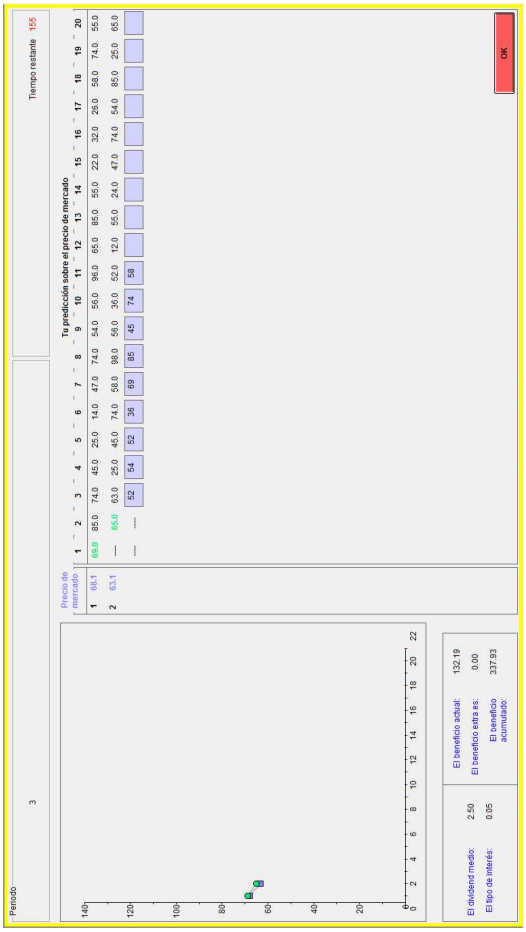
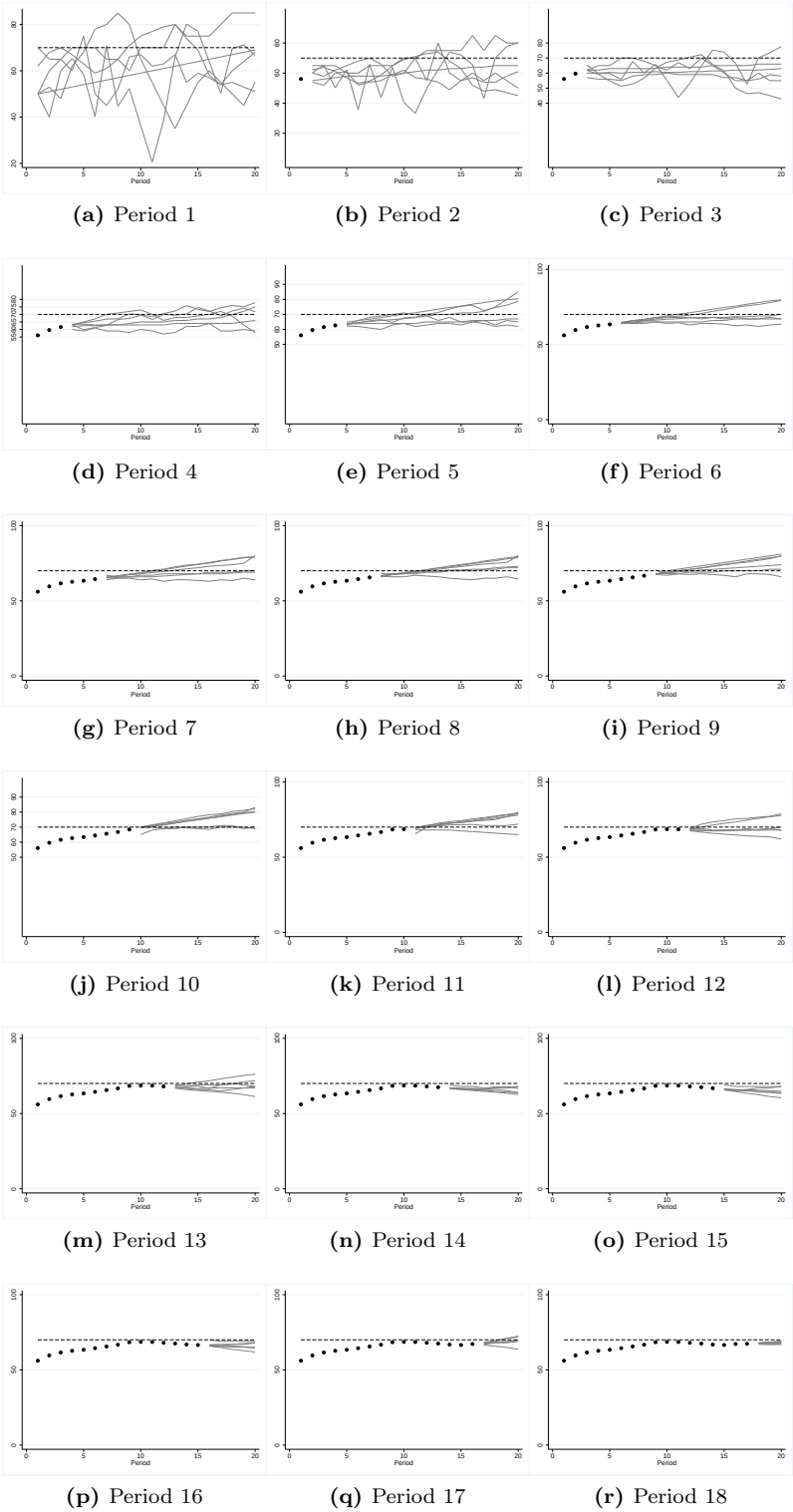


Figure 20: Screenshot of the game: all the feasible information like the previous predictions, the past realized price, the past profit, the interest rate and the mean dividend. On the right hand side there are the table in which players insert he prediction for the future periods

B Individual long term predictions



(a) Period 1 **(b)** Period 2 **(c)** Period 3

(d) Period 4 **(e)** Period 5 **(f)** Period 6

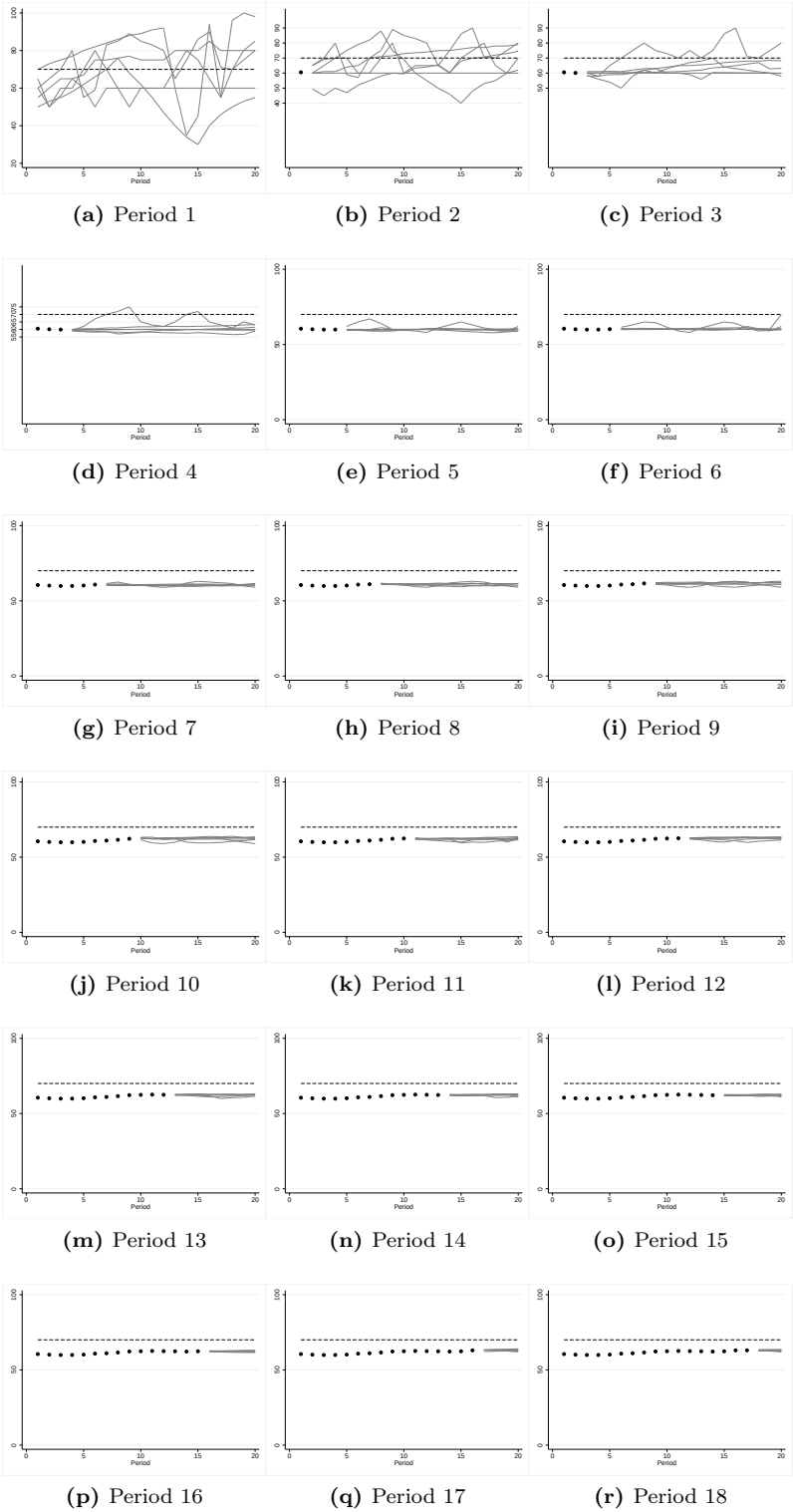
(g) Period 7 **(h)** Period 8 **(i)** Period 9

(j) Period 10 **(k)** Period 11 **(l)** Period 12

(m) Period 13 **(n)** Period 14 **(o)** Period 15

(p) Period 16 **(q)** Period 17 **(r)** Period 18

Figure 21: Individual long-run predictions of Group 1. The black dots are the market price, the grey lines are the individual forecasts and the dashed line represents the fundamental value.



(a) Period 1 **(b)** Period 2 **(c)** Period 3

(d) Period 4 **(e)** Period 5 **(f)** Period 6

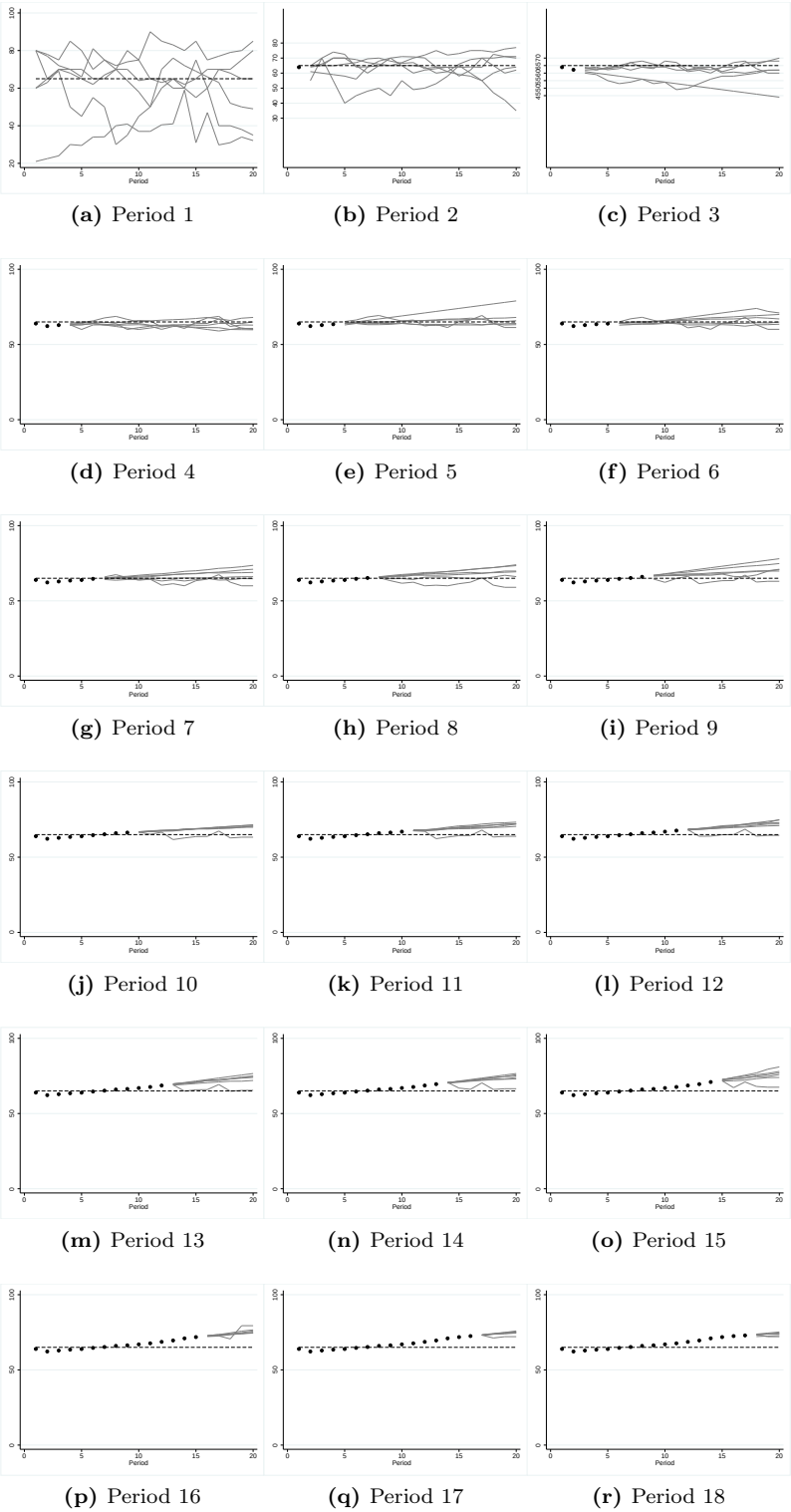
(g) Period 7 **(h)** Period 8 **(i)** Period 9

(j) Period 10 **(k)** Period 11 **(l)** Period 12

(m) Period 13 **(n)** Period 14 **(o)** Period 15

(p) Period 16 **(q)** Period 17 **(r)** Period 18

Figure 22: Individual long-run predictions of Group 2. The black dots are the market price, the grey lines are the individual forecasts and the dashed line represents the fundamental value.



(a) Period 1 **(b)** Period 2 **(c)** Period 3

(d) Period 4 **(e)** Period 5 **(f)** Period 6

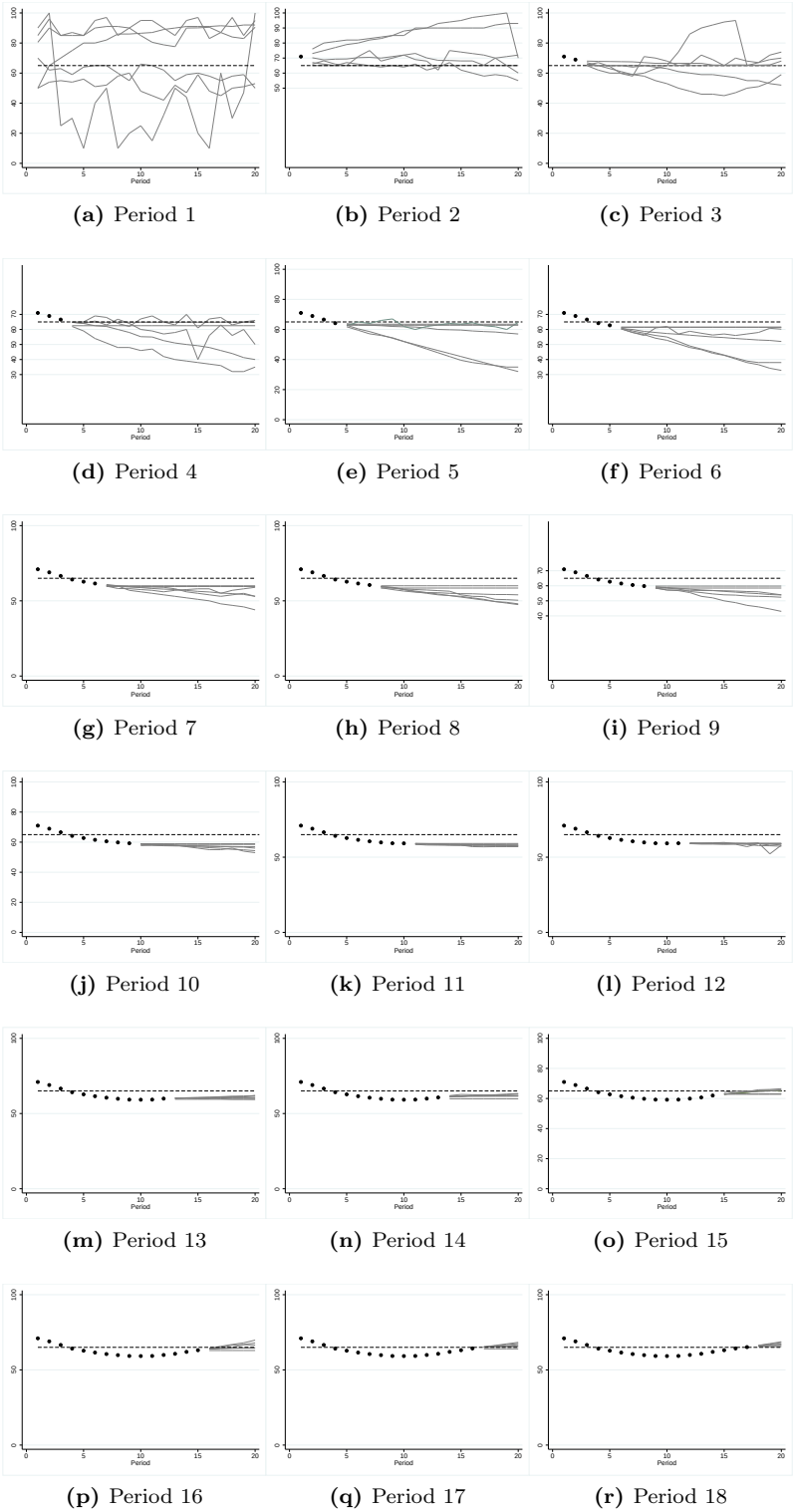
(g) Period 7 **(h)** Period 8 **(i)** Period 9

(j) Period 10 **(k)** Period 11 **(l)** Period 12

(m) Period 13 **(n)** Period 14 **(o)** Period 15

(p) Period 16 **(q)** Period 17 **(r)** Period 18

Figure 23: Individual long-run predictions of Group 3. The black dots are the market price, the grey lines are the individual forecasts and the dashed line represents the fundamental value.



(a) Period 1 **(b)** Period 2 **(c)** Period 3

(d) Period 4 **(e)** Period 5 **(f)** Period 6

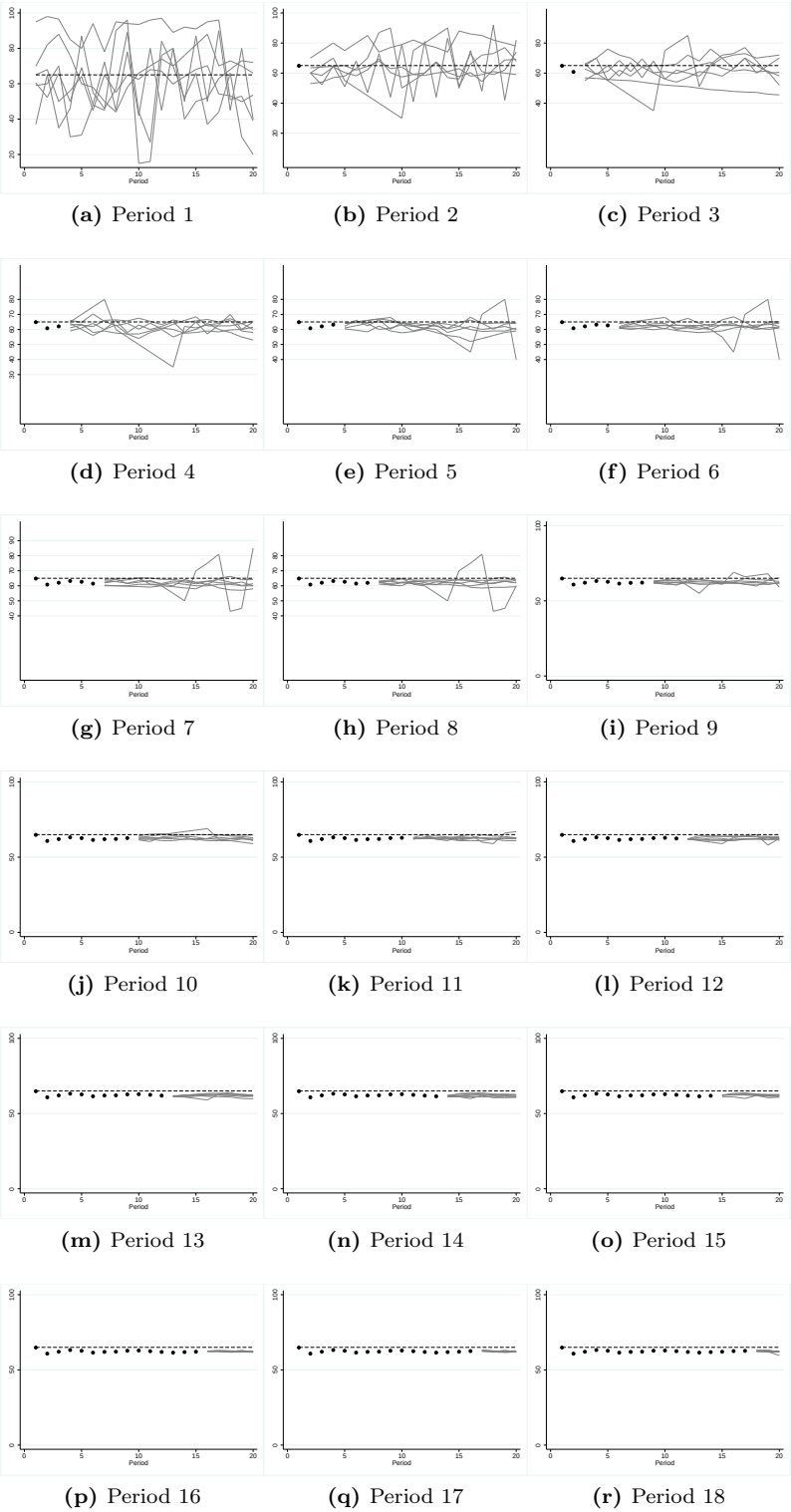
(g) Period 7 **(h)** Period 8 **(i)** Period 9

(j) Period 10 **(k)** Period 11 **(l)** Period 12

(m) Period 13 **(n)** Period 14 **(o)** Period 15

(p) Period 16 **(q)** Period 17 **(r)** Period 18

Figure 24: Individual long-run predictions of Group 4. The black dots are the market price, the grey lines are the individual forecasts and the dashed line represents the fundamental value.



(a) Period 1 **(b)** Period 2 **(c)** Period 3

(d) Period 4 **(e)** Period 5 **(f)** Period 6

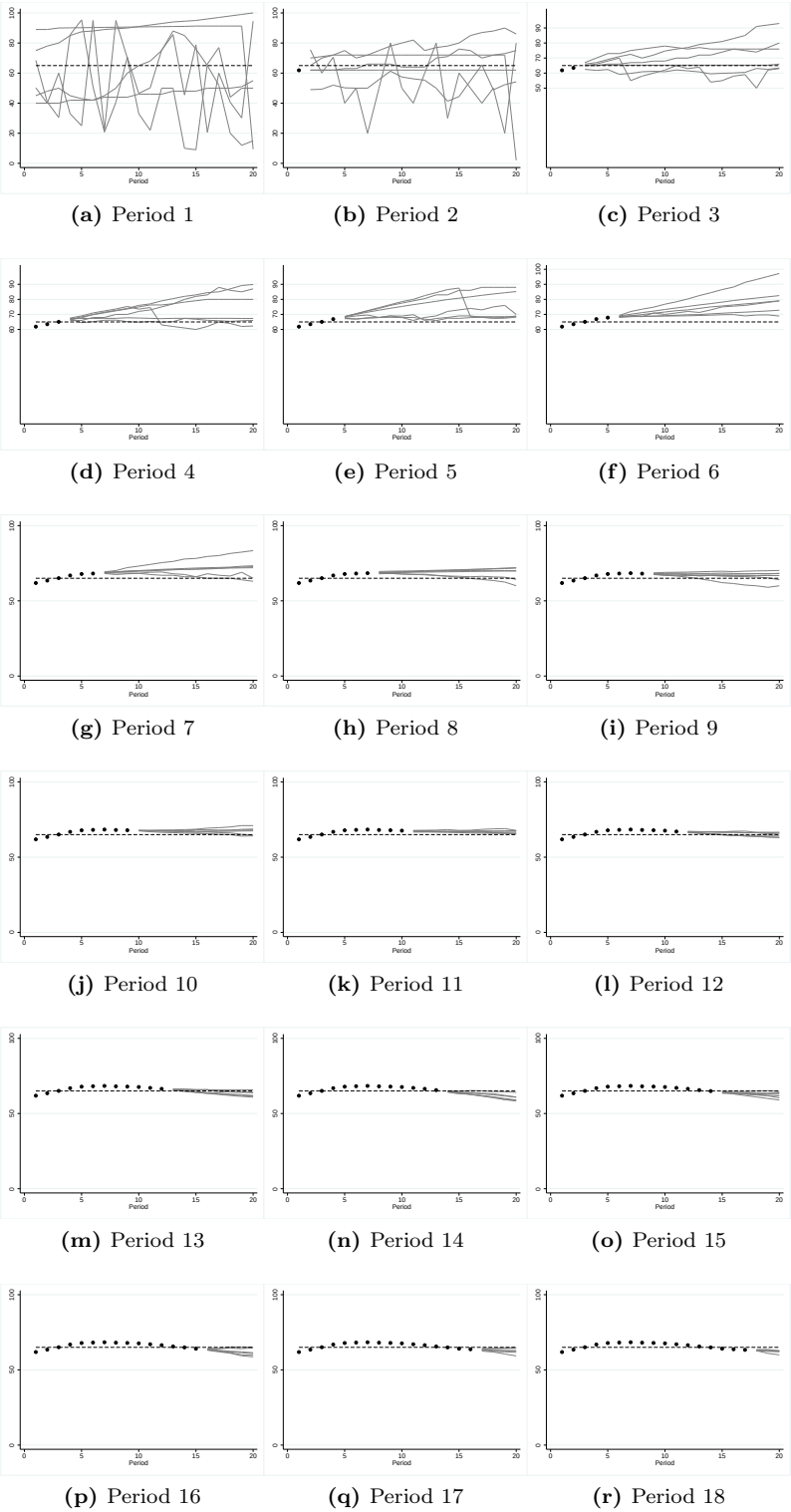
(g) Period 7 **(h)** Period 8 **(i)** Period 9

(j) Period 10 **(k)** Period 11 **(l)** Period 12

(m) Period 13 **(n)** Period 14 **(o)** Period 15

(p) Period 16 **(q)** Period 17 **(r)** Period 18

Figure 25: Individual long-run predictions of Group 5. The black dots are the market price, the grey lines are the individual forecasts and the dashed line represents the fundamental value.



(a) Period 1 **(b)** Period 2 **(c)** Period 3

(d) Period 4 **(e)** Period 5 **(f)** Period 6

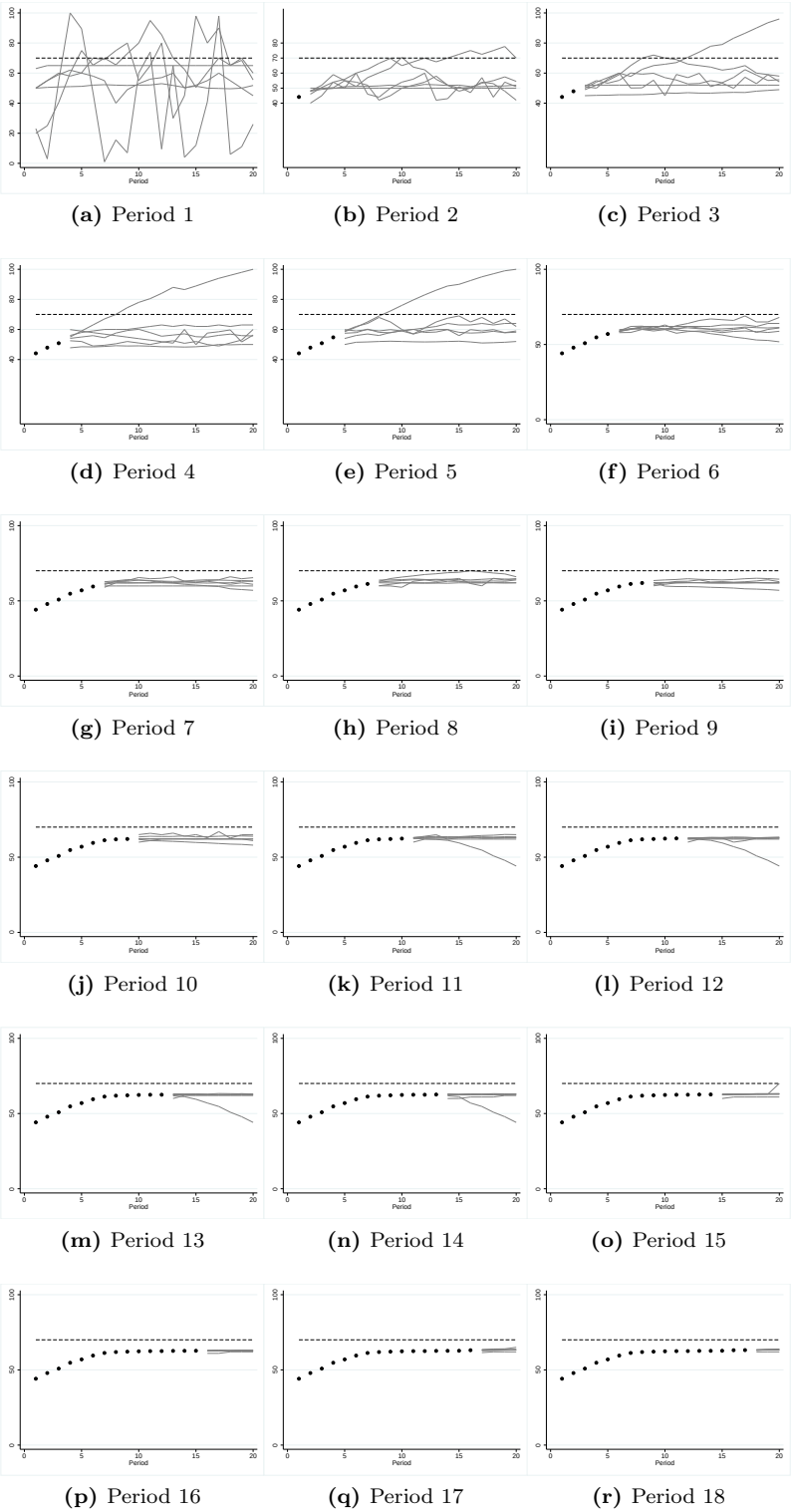
(g) Period 7 **(h)** Period 8 **(i)** Period 9

(j) Period 10 **(k)** Period 11 **(l)** Period 12

(m) Period 13 **(n)** Period 14 **(o)** Period 15

(p) Period 16 **(q)** Period 17 **(r)** Period 18

Figure 26: Individual long-run predictions of Group 6. The black dots are the market price, the grey lines are the individual forecasts and the dashed line represents the fundamental value.



(a) Period 1 **(b)** Period 2 **(c)** Period 3

(d) Period 4 **(e)** Period 5 **(f)** Period 6

(g) Period 7 **(h)** Period 8 **(i)** Period 9

(j) Period 10 **(k)** Period 11 **(l)** Period 12

(m) Period 13 **(n)** Period 14 **(o)** Period 15

(p) Period 16 **(q)** Period 17 **(r)** Period 18

Figure 27: Individual long-run predictions of Group 7. The black dots are the market price, the grey lines are the individual forecasts and the dashed line represents the fundamental value.