REM WORKING PAPER SERIES

Time pressure reduces financial bubbles: Evidence from a forecasting experiment

Mikhail Anufriev, Frieder Neunhoeffer, Jan Tuinstra

REM Working Paper 0351-2024

October 2024

REM – Research in Economics and Mathematics
Rua Miguel Lúpi 20,
1249-078 Lisboa,
Portugal

ISSN 2184-108X

REM – Research in Economics and Mathematics

Rua Miguel Lupi, 20
1249-078 LISBOA
Portugal

Telephone: +351 - 213 925 912
E-mail: rem@iseg.ulisboa.pt

https://rem.rc.iseg.ulisboa.pt/

https://twitter.com/ResearchRem

https://www.linkedin.com/company/researchrem/

https://www.facebook.com/researchrem/

Time pressure reduces financial bubbles: Evidence from a forecasting experiment

Mikhail Anufriev* Frieder Neunhoeffer† Jan Tuinstra‡

March 3, 2024 (link to latest version)

Abstract

We investigate whether time pressure exacerbates or mitigates bubbles in laboratory experiments. We find that under high time pressure price volatility is lower and market prices are closer to their fundamental value. This is due to participants using simpler adaptive forecasting strategies, instead of the self-reinforcing extrapolative expectations that they use under low time pressure, and which are conducive to the emergence of bubbles. In addition, by substantially increasing the number of decision periods in our experiment we find that in the long run prices eventually tend to converge to their fundamental value, also in the absence of time pressure.

Keywords: expectation formation, learning-to-forecast, time pressure, long run dynamics, forecasting strategies.

JEL Classification: C91, G11, G41.

Acknowledgements: For many useful comments we thank Aleksandr Alekseev, Simone Alfarano, John Duffy, Myrna Hennequin, Cars Hommes, Johan de Jong, Anita Kopányi-Peuker, Michael McBride, Andreas Ortmann, Valentyn Panchenko, Rupert Sausgruber, Stefan Trautmann, Matthias Weber, and participants at the 2019 ESA meeting, Vancouver, the 2019 Asian Experimental Finance Meeting, Singapore, the 2021 Experimental Finance Online Conference, the 2022 WEHIA workshop, Catania, the 4th Behavioral Macroeconomics Workshop, Bamberg, the 2023 CEF conference, Nice, the 2023 CREST/CEFM Workshop on Experimental Economics, Paris, and seminars at the University of Amsterdam and the University of California Irvine.

*University of Technology Sydney, Department of Economics, Sydney, Australia, mikhail.anufriev@uts.edu.au, and Technical University of Ostrava, Faculty of Economics, Department of Finance, Ostrava, Czech Republic, ORCID: 0000-0002-1842-6814

†Bocconi University, frieder.neunhoeffer@unibocconi.it, and ISEG – School of Economics and Management, University of Lisbon; REM – Research in Economics and Mathematics, UECE – Research Unit on Complexity and Economics, ORCID: 0000-0002-3084-2570

‡Corresponding author. University of Amsterdam, Amsterdam School of Economics, Amsterdam, The Netherlands, j.tuinstra@uva.nl, ORCID: 0000-0001-6156-4940

1 Introduction

Expectations play a crucial role in economics due to their profound influence on decision-making. Individuals rely on their expectations about the future to make choices today, whether it is the consumption or production of goods and services, investments in the financial market, or the purchase of real estate. Aggregate choices shape economic outcomes, which will be one of the determinants of the decision-makers' expectations about the future. A thorough understanding of this feedback loop between expectations and market outcomes is important for evaluating the functioning of markets and predicting the effect of economic policies. As a shortcut through this feedback loop, one may assume that expectations are always formed rationally, consistent with the underlying economic model. However, this view has been challenged by experimental evidence (Palan, 2013; Duffy, 2016) as well as survey evidence (Case and Shiller, 2003; Coibion et al., 2018). Furthermore, recent theoretical work has shown that extrapolative expectations can become self-fulfilling, at least temporarily (Fuster et al., 2010; Anufriev and Hommes, 2012; Barberis et al., 2018). The impact of these extrapolative expectations is especially pronounced in demand-driven asset markets where they can lead to large and apparently persistent price deviations from the underlying fundamentals, which has been confirmed in a series of laboratory experiments (Hommes et al., 2005, 2008, 2021; Kopányi-Peuker and Weber, 2021).

This paper contributes to a deeper understanding of forecasting in self-referential financial markets. It is motivated by two observations. Firstly, the existing literature on expectation formation has largely overlooked the role of decision time available to decision-makers. However, in financial markets, the ability to respond quickly to perceived profit opportunities is crucial for traders' success. This is evident not only from the fast-paced nature of open outcry exchanges on trading floors, where market prices respond to news announcements within seconds (Busse and Green, 2002), but also from the proliferation of high-frequency trading which employs algorithms that enable rapid trades and operate with a short-term horizon. Traders in financial markets therefore often face substantial *time pressure*, which may impact their behavior and affect market dynamics.

Secondly, experimental research on forecasting typically focuses on behavior in stationary market environments for relatively short time spans, typically 50 decision periods or less. However, recent experimental work in other domains, e.g., Friedman et al. (2015), suggests that this may not be sufficient to capture all aspects of learning and market dynamics. It is therefore of interest to investigate whether the main findings regarding expectation formation hold true over a larger number of decision periods or merely correspond to transitory phenomena. By examining the long-run dynamics of expectation formation and market outcomes, we can gain valuable insights into the stability and sustainability of expectation-driven behaviors in financial markets.

In this paper, we present the results of a *Learning-to-Forecast* (LtF) experiment aimed at studying the impact of time pressure on the emergence of asset price bubbles and at investigating their persistence. To that end (i) we vary, between treatments, the time available to participants for making each decision; and (ii) we extend, within each treatment, the number of decision periods from 50 to about 150.

LtF experiments have recently gained significance in studying financial markets and macroeconomic models, as highlighted in comprehensive surveys by Hommes (2011) and Hommes (2021), respectively. These experiments focus solely on participants' ability to forecast future prices, with computerized agents executing corresponding trades. These automated agents rely on participants' price forecasts to make supply and demand decisions, while market clearance is governed by the experimental software. By separating forecasting from trading tasks, LtF experiments enable researchers to obtain clean information about participants' forecasts and to analyze their influence on market outcomes, considering the self-referential nature of financial markets.[1]

A robust finding from LtF asset pricing experiments is the endogenous emergence of bubbles and crashes (e.g., Hommes et al., 2008). This pattern has been replicated under various conditions, such as in large groups of up to 100 participants per market (Hommes et al., 2021) and when participants are experienced with the decision environment (Kopányi-Peuker and Weber, 2021).[2] Bubbles and crashes also emerge in the presence of stabilizing 'fundamentalist' robot traders (Hommes et al., 2005).

Our experimental data reveal that time pressure has a *mitigating effect* on the occurrence of bubbles and crashes, particularly in the first 50 periods of the experiment. To identify the underlying mechanism, we examine the prediction rules participants use. We find that under low time pressure, which provides a setting comparable to previous LtF experiments, participants tend to rely on rules that extrapolate trends in past asset prices. When prices increase, most participants expect a further price increase. The increased price expectation leads to more demand for the asset from the (computerized) traders, resulting in an instantaneous increase in the market clearing price. Participants' expectations of rising prices are thus confirmed, reinforcing their prediction strategy and ultimately contributing to a large bubble in asset prices. In contrast, under high time pressure, participants tend to adopt simpler prediction rules that are typically based only on the last observed price and inhibit the emergence of

[1] While our paper reports on an experiment in an asset pricing environment, it is worth noting that LtF experiments have been conducted in macroeconomic settings as well, with pioneering contributions by Marimon et al. (1993); Marimon and Sunder (1993, 1994), and recent experiments presented in Assenza et al. (2021); Mokhtarzadeh and Petersen (2021); Kryvtsov and Petersen (2021). Note that time pressure and long-term dynamics are clearly not exclusive to financial markets but have broader relevance.

[2] Kopányi-Peuker and Weber (2021) also show that bubbles and crashes emerge with experienced participants when these participants actively trade assets as in the experimental design of Smith et al. (1988). The similarity in aggregate market outcomes between LtF experiments and so-called *Learning-to-Optimize* experiments, where participants are trading (instead of only forecasting) is further supported by the experiments presented in Bao et al. (2017) and Arifovic et al. (2019). Moreover, Carlé et al. (2019) and Füllbrunn et al. (2022) find that forecasting behavior in experimental asset markets with trade is generally in line with participants' trading behavior.

bubbles.

We also find that in the long run the effect of time pressure is somewhat reduced. On the one hand, when participants have sufficient decision time, the incidence and size of bubbles and crashes diminish over time, providing evidence that the patterns found in earlier experiments, which lasted for about 50 periods, have a transitory character. However, convergence is fragile, and bubbles may sporadically re-emerge. On the other hand, in markets where participants have limited decision time, bubbles sometimes endogenously emerge later in the experiment. As participants adjust to higher time pressure, they may learn to use extrapolative forecasting rules.

Literature Review

By studying expectation channels that may facilitate price bubbles and subsequent crashes, we contribute to a large financial literature documenting repeated bubble-and-crash dynamics in real markets (see Phillips et al., 2011, 2015 for bubbles in stock markets, Cheah and Fry, 2015; Corbet et al., 2018, for bubbles in crypto-markets, and Case and Shiller, 2003; Case et al., 2012, for bubbles in housing markets). These boom-and-bust cycles in asset prices have been successfully replicated in laboratory experiments with paid human subjects under a wide variety of conditions (see Hommes, 2011 and Palan, 2013 for overviews). However, participants' decisions in these experiments are usually made in the absence of time pressure. In addition, the number of trading periods is typically low, making it difficult to evaluate whether experimental bubbles are persistent or are of a more transient nature so that they may disappear over time.

Despite its long-standing tradition in psychology, the use of laboratory experiments to study the effects of time pressure on decision-making is relatively new in economics (see Spiliopoulos and Ortmann, 2018, for an excellent review). The study closest to ours is Moritz et al. (2014), which presents an experiment where participants repeatedly predict the next realization of an exogenous time series. The authors show that forecasting performance decreases under time pressure. Ferri et al. (2021) investigate the effect of time pressure in the classic Smith et al. (1988) asset market experiment. Their findings indicate that market prices are on average below fundamental values. However, higher time pressure leads to positive deviations from the fundamental value, in particular at the beginning of the experiment, and also increases price volatility. Kocher and Sutter (2006) provide further evidence that time pressure may impede decision-making quality. An increase in time pressure slows down convergence to the Nash equilibrium in a beauty contest experiment. In contrast to these studies, we find that time pressure enhances forecasting performance.[3] Specifically, under higher time pressure participants switch to simpler forecasting strategies (e.g., conditioning the forecast on the last observed price instead of extrapolating a trend in previous prices) and these strategies tend to stabilize the dynamics of realized market prices, thereby facilitating their predictability.[4]

Our results are consistent with adaptive decision-making theory, see Payne et al. (1993). For example, Payne et al. (1988) and Rieskamp and Hoffrage (2008) demonstrate that in individual decision-making problems, participants adapt to high time pressure by using less complex decision rules. Spiliopoulos et al. (2018) extend these findings to decision problems involving strategic interaction. In our post-experiment questionnaire, participants reveal that they rely on simpler rules under high time pressure. This is confirmed by the prediction rules we estimate from the experimental data. Gigerenzer and Goldstein (1996) and Goldstein and Gigerenzer (2009) show that simple decision or prediction rules may outperform complex rules. This is also the case in our experiment.

[3]Lindner and Sutter (2013) also find that under higher time pressure participants' decisions are closer to the mixed strategy Nash equilibrium in the '11-20 money request game', but the authors suggest that this may be coincidental.

[4]The finding that people tend to use less information under time pressure is referred to as "filtration" in the psychological literature (e.g., Zur and Breznitz, 1981).

To an extent, our results are also consistent with dual-process theory, see, e.g., Stanovich and West (2000); Kahneman (2003). This theory assumes people make decisions through either a fast and automatic process relying on intuition (System 1) or a slower and more deliberate process (System 2). Some previous experimental evidence suggests that System 1 is dominant under higher time pressure (Moritz et al., 2014; Ferri et al., 2021). The scores of our participants on the Cognitive Reflection Test (Frederick, 2005) indeed suggest that the use of System 1 is more prevalent under high time pressure than under low time pressure.

Earlier LtF experiments run for maximally 50 consecutive periods except for those LtF experiments divided into separate blocks. Bao et al. (2012), for example, employ 65 periods with a structural change in the asset's fundamental value in periods 21 and 44. Kopányi-Peuker and Weber (2021) study the effect of experience in an LtF experiment by running three consecutive repetitions with a duration of 28, 32, and 26 periods, respectively. In both studies, bubbles occur also in the later stages of the experiment.

In these earlier experiments, the number of periods appears to have a limited impact on the viability of bubbles. However, in our experiment, which consists of approximately 150 periods per block and is considerably longer than previous LtF experiments, we see that bubbles eventually tend to disappear in the low time pressure treatment.[5] The number of decision periods has shown a similar effect in other decision environments. Berninghaus and Ehrhart (1998), for example, present an experiment on the minimal effort game and find that increasing the number of repetitions helps participants to coordinate on the Pareto dominant equilibrium. In contrast to earlier (but shorter) market entry experiments, Duffy and Hopkins (2005)

[5]In experiments where the asset is traded, in the tradition of Smith et al. (1988), trade typically takes place for around 15 periods with bubbles and crashes being a very robust feature. These bubbles persist for trade experiments with more periods. Lahav (2011) lets participants trade assets for 200 periods and finds multiple bubbles in prices until the end of the experiment. Smith et al. (2014) run an experiment with 50 trading periods and find a single bubble and crash pattern in most of their markets.

observe a tendency for participants to coordinate on one of the asymmetric pure strategy Nash equilibria in a market entry experiment repeated 100 times. However, it takes almost all 100 periods for convergence on such an equilibrium to occur. Finally, Friedman et al. (2015) consider Cournot games repeated for 1200 periods, with group rematching after each 400 periods. Consistent with earlier contributions, quantities approximate the competitive outcome (characterized by zero payoffs) in the initial 50 periods. However, this behavior turns out to be transient, as quantities fall below the Cournot-Nash equilibrium level or even approach collusive levels in the long run.

The remainder of this paper is organized as follows. In Section 2, we outline the experimental design and the hypotheses to be tested. We discuss our results in Section 3 and provide some concluding remarks in Section 4. Supplementary information, including the experimental instructions and additional data analyses, can be found in the appendices.

2 Experimental design

The experiment was conducted in the CREED laboratory of the University of Amsterdam.[6] In total 186 subjects participated in 12 sessions encompassing three different treatments. None of the subjects took part in more than one session. The sample of participants was fairly gender balanced (54% females), with an average age of 22 (see Table D1 in Internet Appendix D). Most participants (65%) were enrolled in a study program from the Economics and Business faculty of the University of Amsterdam. The experimental sessions lasted for about two hours, and average earnings (including a show-up fee) were €27.64, ranging from €13 to €40.

The experimental setup is based on the classical asset-pricing model with heterogeneous expectations (see Campbell et al., 1997 and Brock and Hommes, 1998) and is similar to the LtF experiments in Hommes et al. (2005, 2008). We follow their

[6]The experimental software, programmed in oTree (Chen et al., 2016), is available upon request.

design to focus solely on the implications of time pressure and long-run dynamics. In our experiment, several institutional traders invest their wealth in a financial market comprising a risk-free asset and a risky asset. Each human participant assumes the role of an "advisor to a pension fund" as explained in the instructions. They do not trade but only make price forecasts that the fund they advise uses to determine its demand for the risky asset. The price of the risky asset is determined by the forecasts of the human participants and some additional robot traders, and reported to the participants. Subsequently, the subjects make a point forecast for the next period, and this process continues for multiple periods. Each participant is paid based on the accuracy of their own forecasts, as explained below.

Before presenting the treatments in Section 2.3 and the hypotheses in Section 2.4, we discuss the price-generating mechanism of the model and the experimental procedures and software in the next two sections. Complete experimental instructions can be found in Internet Appendix C.

2.1 The price generating mechanism

As previously mentioned, there are two assets in the market. The risk-free asset pays a fixed interest rate $r > 0$ per period. The infinitely lived risky asset pays dividend y_t in period t. The dividend is independently and identically distributed (IID) with an expected value $\bar{y}$. The price of the risky asset p_t is endogenously determined by the aggregate demand and aggregate supply of the asset. In this market setting, all traders are assumed to be myopic mean-variance maximizers with full information about the dividend process. The model can then be solved for the market clearing price p_t, see Appendix A. There we show that traders' demand for the risky asset is an increasing linear function of the *expected return* $p^e_{i,t+1} + \bar{y} - (1+r)p_t$, where $p^e_{i,t+1}$ is trader i's expectation for the price in period $t + 1$. This expectation is formed at the beginning of period t, prior to the realization of the market clearing price p_t. In the experiment participants' task is to provide these price expectations for their trader (pension fund).

There are two types of traders in each experimental market: six large pension funds, each advised by one of the participants, and a fraction $n_t \in [0, 1)$ of computerized 'robot' investors. The underlying predictions for those 'robot' investors represent the fundamental price, given by the discounted expected value of all future dividends.[7] For our IID dividend process, the *fundamental price* of the risky asset is

$$p^f = \frac{\bar{y}}{r}\,. \tag{1}$$

Depending on the composition of the market, the market clearing price in period t is then given by

$$p_t = \frac{1}{1+r}\left[(1-n_t)\bar{p}^e_{t+1} + n_t p^f + \bar{y}\right]\,, \tag{2}$$

where $\bar{p}^e_{t+1}$ is the *average expected price* for period $t+1$, averaged over all human participants in the market that submitted their predictions in period t. Using Eq. (1), we can rewrite Eq. (2) as

$$p_t - p^f = \frac{1-n_t}{1+r}\left(\bar{p}^e_{t+1} - p^f\right)\,. \tag{3}$$

Therefore, the price deviation from the fundamental value in period t is positively related to the price deviation that the participants expect for period $t+1$. In other words, there is *positive expectations feedback* between predicted and realized asset prices. If traders expect the price to go up, they start buying the asset to reap the benefits from potential capital gains. This rise in demand for the risky asset instantaneously drives up its market clearing price. As the right side of Eq. (3) decreases

[7] Our computerized robots are therefore similar to so-called "value investors" in financial markets. The use of robot traders in experimental asset markets is fairly common (e.g., Bloomfield and O'Hara, 1999; Veiga and Vorsatz, 2010). In our experiment, as previously in Hommes et al. (2005), these robot traders play a stabilizing role.

with n_t, the expectation feedback is mitigated by robots, representing fundamentalists in this market. As in Hommes et al. (2005), we assume that these fundamentalists become more active as the price diverges further from its fundamental value and set

$$n_t = 1 - \exp\left(-\frac{1}{20}\left|\frac{p_{t-1} - p^f}{p^f}\right|\right) . \tag{4}$$

As a result, the impact of the robots grows with the extent of mispricing. The fundamentalist robots serve as an endogenous mechanism that inhibits bubbles from growing indefinitely.[8] However, specification (4) ensures that the mispricing must be substantial for the fundamentalist traders to become active. Even when the price is twice as high as the fundamental value, the robots' aggregate weight is below 5%. This is less than one-third of the weight of each human participant's prediction.[9]

2.2 Experimental procedure and software

Upon arrival in the lab, participants receive instructions, both on paper and on their computer screens. The instructions (Internet Appendix C) provide participants with background information about the market environment, their task, the structure of the experiment, and how their payoffs will be determined.

Following the standard experimental protocol for LtF experiments (Hommes, 2011), each participant assumes the role of a financial advisor to a fund that has to decide how to divide its wealth between a risky and a risk-free asset over a number of consecutive periods. The fund's decision depends on the future price of the risky

[8] Many LtF experiments (Hommes et al., 2008, 2021) do not incorporate robot traders, and the price follows Eq. (2) with $n_t \equiv 0$ for all t. In all these studies, the price bubbles emerge early and grow until they hit an artificial upper bound, typically set at 1 000. Participants are informed about this upper bound only when they try to submit a prediction that is higher than this bound. In our experiment, the upper bound is set to 10 000. Due to the presence of robot traders in our experiment, it is never reached. The highest price in our experiment equals 2 504. See also footnote 13.

[9] The average weight of the robots in our experiment, over all periods and all groups, is about 2%. In only about 0.4% of the 9393 periods in our experiment, the weight n_t is larger than 25%.

asset, and the participants' objective is to predict that price as accurately as possible. Although participants are not explicitly given the equations (2) and (4), they do know that the realized price is determined through the equilibrium between the aggregate demand of all funds and a fixed supply of the risky asset.[10] In addition, participants are informed that higher price forecasts lead to increased investment in the risky asset by the fund they advise. They are also informed that some of the other funds in the market are advised by other human participants, and some use a fixed investment strategy.

The instructions explain that the experiment consists of two phases of price forecasting and concludes with a brief questionnaire. In both phases, participants are tasked with predicting the price for a large number of consecutive periods. They are informed that the number of periods is between 120 and 180 in each phase. The range is provided to prevent strategic effects that may arise toward the end of each phase. The instructions emphasize that the time to make each forecast will be limited, participants may also experience a "waiting" time between decisions (Section 2.3), and that the waiting and decision times, market parameters, and market composition remain constant during each phase but might be different between phases.

Earnings are based on participants' forecasting accuracy and are determined as follows. At the end of the experiment, the computer program randomly chooses 10 periods from the first phase and 10 periods from the second phase for payment. The aggregate points accumulated during these 20 periods determine the participant's payoff. If period t is selected, the number of points e_t for this period is based on the participant's absolute prediction error in this period as follows

$$e_t = \frac{200}{1 + |p_t - p_t^e|}\,, \tag{5}$$

[10] Sonnemans and Tuinstra (2010) show that exact knowledge of the price-generating mechanism in LtF experiments does not affect forecasting behavior and price dynamics.

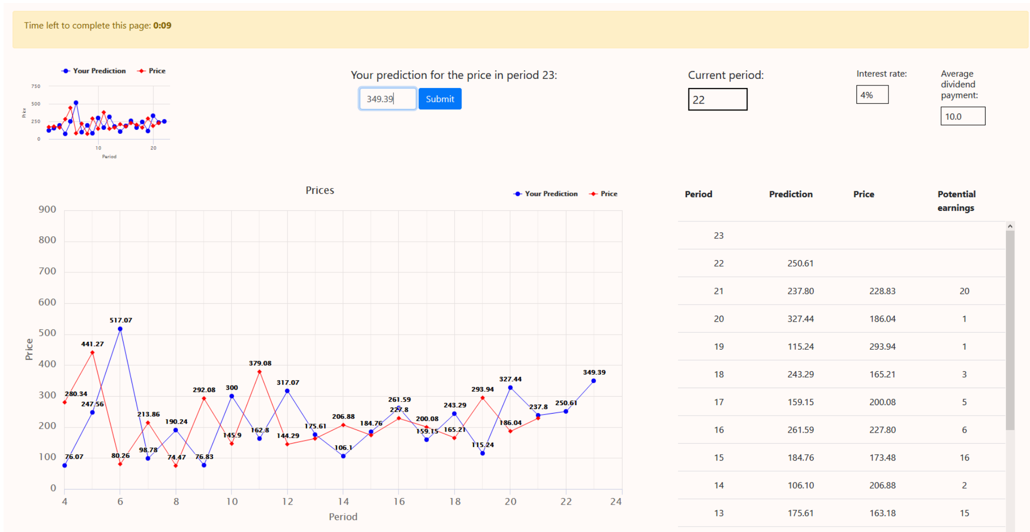


Figure 1: An example of the computer screen. The graph shows past predictions (blue) and past prices (red). The table provides the same information, and in addition, the "Potential earnings", that is, the points for each period if that period is selected for payment (using Eq. (5)). A prediction can either be typed into the box in the top middle of the screen or indicated by clicking the mouse button on the graph in the lower part. In the **H** and **L** treatments, the prediction is submitted automatically as the decision time runs out. To submit the prediction in the **HS** treatment, participants must either press the 'Enter' button on the keyboard or click the blue 'Submit' button on the top of the screen.

where p_t^e is the participant's prediction for price p_t. If in period t, $e_t = 0$. We choose the hyperbolic scoring rule (5), as it provides strong incentives to submit the exact forecast, by penalizing even small forecasting errors. A graphical representation of this scoring function can be found in Internet Appendix C, as included in the instructions provided to participants. Our payoff structure ensures that participants have a strong incentive to be accurate in each period, regardless of their performance in other periods. The total number of points earned is converted at a rate of one euro per 100 points at the end of the experiment. A participant can earn between 0 and 2 euros per period, summing up to €40 for the entire experiment. In addition, every participant receives a €10 show-up fee.

After reading the instructions and completing a short comprehension test, participants engage in a practice round to familiarize themselves with the software.[11]

[11] In this practice round, participants are presented with a pre-determined time series of prices and past predictions and have to choose predictions for several periods. Predictions do not impact

Fig. 1 provides an illustration of the interface.[12] Predictions can be submitted by either typing a number in the submission box at the top of the screen or by clicking with the computer mouse in the area of the main graph.[13] In the latter case, the value selected appears as the prediction both in the graph and in the submission box. The value can be changed by clicking on different points, and participants are free to change predictions as often as they like within the decision time. Once the decision time expires, the value in the submission box is automatically selected as the price prediction in treatments **L** and **H**. In the **HS** treatment, a blue 'Submit' button is added to the screen next to the prediction box. The computer program considers the prediction from the submission box only when this button is clicked or, alternatively, when the 'Enter' button on the keyboard is pressed.

Our sessions consist of either 6, 12, or 18 participants.[14] Once all participants have completed the practice round, the computer randomly forms groups of six participants, and the first phase starts. The starting screen announces the interest rate r, the mean dividend $\overline{y}$, and the decision and waiting times. This information remains visible in the upper part of the screen throughout the entire phase, accompanied by a countdown timer. The realized price for period $t = 1$ is determined by participants' predictions for period $t = 2$. Thus, participants start out with two predictions, for periods 1 and 2.[15] Then, at each period $t > 1$, every participant can submit their prediction for the price at $t+1$. The computer interface features a graph displaying the

prices in the practice round.

[12]This is an example of the screen shown in period 22 of treatment **HS**.

[13]Participants can choose any value in the interval $(0, 10\,000]$ up to two digits. Following previous LtF studies, our experiment does not mention the upper limit of 10 000 in the instructions and only informs subjects when their predictions exceed this upper limit. We find that out of the 186 participants only five chose one or more predictions that were at least once reasonably close (i.e., within 5%) to this upper bound.

[14]The total number of participants in one session depends on the turnout rate and has to be a multiple of six. There are eight sessions with 18 participants, three sessions with 12 participants, and one session with six participants.

[15]There are no fundamentalist robots at $t = 1$. To limit dispersion in the initial predictions, the instructions mention that the first two prices in the first phase are "likely to lie between 0 and 200" and that the first two prices in the second phase are "likely to lie between 0 and 100". These ranges include the fundamental price of 126.4 in the first phase and 71.2 in the second phase (Section 2.3).

participant's own predictions (blue) and the realized asset prices (red) for the most recent 20 periods (Fig. 1). A smaller graph in the upper left of the screen depicts the entire time series from the beginning of the phase. The vertical axis of both graphs automatically adjusts to accommodate increases in prices or predictions beyond the current scale. Additionally, a table on the right-hand side presents the participant's own predictions, the realized prices, and the number of points per period computed using Eq. (5). The latter are called 'Potential Earnings' on the screen as they are only earned if the corresponding period is chosen for the final payoff. At the end of the phase, the screen notifies participants that they are entering a new market. The second phase starts with new market parameters, including the decision and waiting times. Participants are randomly re-matched in groups of six.

Upon completing the two phases, the participants' cognitive ability is assessed by the standard three-question Cognitive Reflection Test (Frederick, 2005). Subsequently, participants fill out a questionnaire that elicits demographic information (gender, age, study program), the extent to which they experienced time pressure during the experiment, and the prediction strategies they used (Internet Appendix D). Finally, the computer informs each participant which of the 20 rounds were chosen for payment and what their payoff is. Participants are paid out privately.

2.3 Treatments

We solicit predictions from participants under two different conditions, which we refer to as the **low time pressure** (LTP) condition and the **high time pressure** (HTP) condition. These conditions vary in the amount of time allocated to participants for submitting a decision.

Each decision period is separated into a *waiting time* and a *decision time*. During the waiting time, participants have access to the computer interface, including the graphs and table displaying past prices and predictions. They can navigate the computer screen using their mouse and select potential predictions. However, only after the waiting time is over, and the decision time begins, the submission box will appear, showing the last value that was clicked, and predictions can be submitted.

In the low time pressure (LTP) condition, we set the waiting time at 10 seconds and the decision time at 15 seconds. In contrast, in the high time pressure (HTP) condition, the waiting time is set to 0 seconds and the decision time is set to 6 seconds. Therefore, the total time for arriving at a decision in the LTP condition equals 25 seconds per period, which is in line with the average time taken per prediction by participants in previous LtF experiments.[16] The total time for making a decision in the HTP condition is six seconds, which imposes substantial time pressure on participants (as corroborated by the questionnaire responses).

To study the long-run dynamics of individual predictions and market prices, we run the LTP condition for 146 periods and the HTP condition for 159 periods. Thus, the number of periods in each phase is approximately three times higher than in previous LtF experiments. Each participant engages in two prediction phases, one under the LTP condition, and one under the HTP condition. This ensures comparability in session length across treatments.

The treatments differ in the order in which participants experience these conditions (Table 1). We will refer to the treatment that starts with the LTP condition as treatment **L** and the treatment that starts with the HTP condition as treatment **H**. In all sessions, regardless of the time pressure condition imposed, the interest rate and mean dividend in the first phase are $r = 0.05$ and $\overline{y} = 6.32$, implying a fundamental value of $p^f = 126.4$, and $r = 0.05$ and $\overline{y} = 3.56$ for the second phase, corresponding to a fundamental value of $p^f = 71.2$.

We run 12 markets in treatment **L** and 10 markets in treatment **H**. After running the first sessions of our experiment, we observed a series of downward spikes in

[16] We include a waiting time in the LTP condition specifically to avoid participants rushing to their decision. See Moritz et al. (2014) for a discussion of how a waiting time can limit under-thinking.

Treatment	Time Pressure condition	Submit button	Number of markets	First Phase	Second Phase
L	LTP → HTP	No	12	146 pds, 25 sec/pd	159 pds, 6 sec/pd
H	HTP → LTP	No	10	159 pds, 6 sec/pd	146 pds, 25 sec/pd
HS	HTP → LTP	Yes	9	159 pds, 6 sec/pd	146 pds, 25 sec/pd
				$\bar{y}=6.32$, $p^f=126.4$	$\bar{y}=3.56$, $p^f=71.2$

Table 1: Overview of the treatments. The last two columns provide the values of the parameters in the two phases of the experiment. The interest rate is $r = 0.05$ in both phases. There are six human participants per group.

individual predictions under the HTP condition (e.g., panels 4, 8, and 10 in Figs. E4 and E5, Internet Appendix E). They appear to be due to participants being unable to complete their prediction in the submission box within the time limit.[17] These outliers, which do not occur under the LTP condition, may have a long-lasting effect on the price dynamics. To rule out these outliers as the main driver for the observed differences between time pressure conditions, we run the additional treatment **HS** as a robustness check.[18] This treatment is identical to treatment **H** in all respects, except that participants are required to either press the 'Enter'-key on the keyboard or click the 'Submit' button, added next to the submission box on the computer screen to register their predictions. Note that, for all treatments, the average prediction used in Eq. (2) is determined by averaging only over the submitted predictions. It can therefore be based upon fewer than six predictions if some participants fail to submit a prediction in time. We run nine markets in treatment **HS**.

2.4 Hypotheses

The first research question aims to determine whether the emergence of large bubbles and crashes in asset prices, as often observed in LtF experiments, is a transient

[17] Participants may be cut off when the decision time lapsed after entering the first digit only, despite their intention to enter two or three digits for the price forecast. This explanation is corroborated by some questionnaire comments.

[18] Note, however, that such outliers also occur in actual financial markets. So-called *fat-finger* errors through keyboard input or mouse click mistakes in human-placed orders are not uncommon in financial trading (Financial Times, 2019).

phenomenon. That is, will market prices eventually stabilize and converge to their fundamental value? This question is difficult to answer on the basis of existing LtF experiments, which do not provide systematic evidence that price volatility decreases or increases over the relatively short span of around 50 periods. It is plausible to speculate that, if participants are given enough opportunities to learn in the stationary environment provided by the experiment, they should be able to improve their predictions, eventually leading market prices to converge to their fundamental value. Also note that participants' earnings, which are based on absolute forecast errors, tend to be much lower in markets with significant volatility than in markets where prices closely stick to their fundamental value. This should provide a strong incentive for participants to adapt their prediction strategies when those strategies lead to pronounced volatility. This leads us to the following hypothesis.

Hypothesis 1. *Asset price volatility and mispricing decrease in the long run.*

Our second research question focuses on the effect of time pressure on price volatility and mispricing. The direction of this effect is not evident *a priori*. On the one hand, time pressure has been shown to reduce the quality of decision-making (Kocher and Sutter, 2006) and to increase market volatility in experimental asset markets (Ferri et al., 2021). Increased time pressure restricts participants' ability to engage in thorough deliberation, which may result in forecasting errors, ultimately leading to elevated price volatility and mispricing. This motivates the following hypothesis.

Hypothesis 2A. *Increased time pressure increases price volatility and mispricing.*

On the other hand, bubbles and crashes in earlier LtF experiments appear to be driven by participants coordinating on a common trend extrapolating prediction strategy, where the predicted price depends on the last two observed prices. When all participants in a market extrapolate the current trend in past prices, the positive feedback characteristic of the underlying price-generating mechanism perpetuates the price trend. This self-confirming nature of trend-extrapolating prediction strategies naturally gives rise to large bubbles in asset prices when participants coordinate on such a strategy. Under increased time pressure, coordination of participants on a common prediction strategy may become more difficult, and the emergence of large bubbles and crashes may therefore be inhibited. Instead of increasing price volatility, higher time pressure may therefore lead to more stable price dynamics. This motivates an alternative version of our second hypothesis.

Hypothesis 2B. *Increased time pressure reduces price volatility and mispricing.*

We will now discuss our experimental results in view of Hypotheses 1, 2A, and 2B.

3 Experimental results

In this section, we present the experimental data and use those to test the hypotheses formulated above. In Section 3.1, we analyze the effect of time pressure and the number of decision periods on market prices. In Section 3.2, we try to gain a better understanding of the results by investigating 'market expectations', which refer to the average price predictions across all participants in the same market. Finally, we explore the extent to which participants' performance can be predicted based on demographic variables and their CRT scores in Section 3.3.

The design of our experiment allows us to investigate the effect of time pressure by comparing the first phase of treatment **L** with those of treatments **H** and **HS**. In addition, we can study the long-run dynamics of prices by comparing prices at the beginning of each treatment with those at the end of that treatment.[19]

[19] We focus our analysis exclusively on the data from the first phase, because the data from the second phase are not independent: behavior and market dynamics may be affected by participants' experiences in the first phase. Moreover, after rematching, second-phase markets are composed of participants from either one, two, or three first-phase markets, which makes it difficult to control for different experiences effectively. For the interested reader, the market price data from the second phase are shown in Internet Appendix F. As our primary focus is on the first phase, we will refer to treatment **L** as the low time pressure treatment, and to treatments **H** and **HS** as the high time pressure treatments.

Since the number of periods in treatments **H** and **HS** is greater than in treatment **L**, our analysis considers the *common* periods 1–145. Moreover, to compare long-run versus short-run dynamics, we primarily examine prices and predictions in periods 11–50 versus those in periods 106–145. The selection of periods 11–50 is based on the convention in earlier LtF experiments, which typically involve around 50 periods and exclude the first ten periods from analysis to allow for some initial learning of the experimental task (Hommes et al., 2005, 2008). We follow that approach and also include the last 40 common periods, corresponding to periods 106-145, to have an equal number of periods for analysis.

3.1 Market prices

The three panels of Fig. 2 depict the market prices (thin gray lines) for each market in treatments **L**, **H**, and **HS**. The black line in each panel represents the median price across all markets in that treatment for each period.[20] In treatment **L**, there are 145 prices in each market, while treatments **H** and **HS** have 158 prices per market. To facilitate comparisons of price dynamics between treatments, the vertical axis ranges from 0 to 500 (i.e., approximately $4 \times p^f$) in all panels. Note that in some periods, particularly in treatment **L**, market prices exceed 500.[21]

We can make a few preliminary observations by comparing the median prices in Fig. 2, further supported by inspection of the market prices in Figs. B1 to B3 from Appendix B. First, median prices exhibit much higher overvaluation in the low time pressure treatment **L** than in the high time pressure treatments **H** and **HS**. In contrast, there does not seem to be a systematic difference between the two high time

[20] See Appendix B for the price dynamics in each market. Figs. B1 to B3 show markets **L**1-**L**12, **H**1-**H**10, and **HS**1-**HS**9, respectively. We will refer to some of these individual markets below to illustrate our results.

[21] The market price incidentally rises above 500 in nine out of 12 markets in treatment **L**. Overall, this happens during 199 periods (corresponding to about 11% of the total number of periods in that treatment). In contrast, the price does not move beyond 500 during any period in treatment **H** and does so in only five periods (0.35% of all periods) in treatment **HS**.

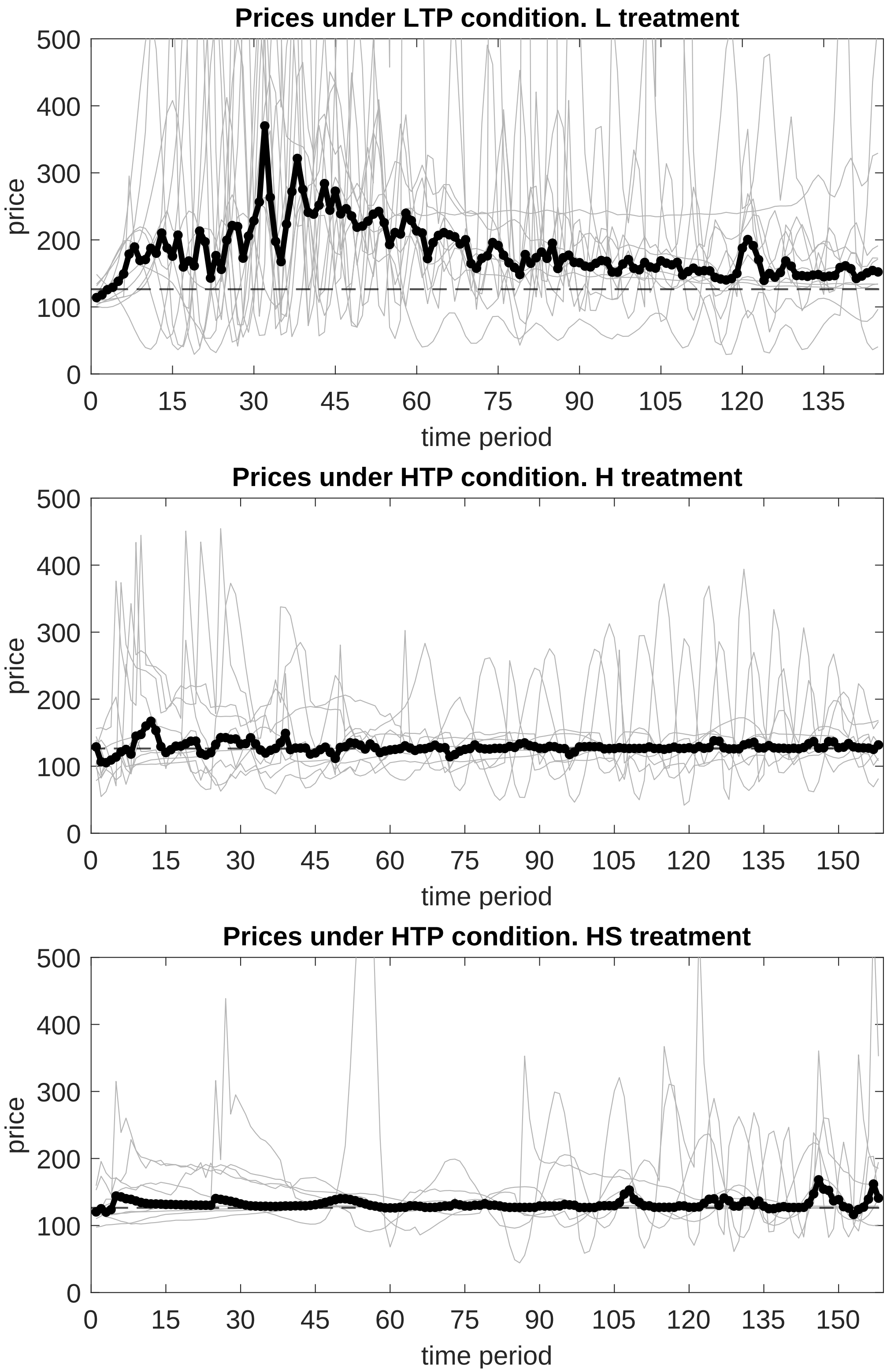


Figure 2: Median prices (thick, black) and prices in individual markets (gray) in the experimental treatments. The fundamental price $p^f = 126.4$ is indicated by the dashed horizontal line.

pressure treatments, suggesting that differences between treatments **L** and **H** cannot be attributed to participants submitting incomplete predictions in treatment **H**.

Second, price volatility is notably high in treatment **L**, with prices oscillating wildly in all markets, see also footnote 21. In contrast, price volatility in both HTP treatments is considerably lower, as most markets experience relatively small oscillations and quick convergence to the fundamental value.

Third, in the LTP treatment, oscillations start almost from the beginning of the experiment, whereas in the HTP treatments, they seem to be more prevalent in the second half of the experiment, following an initial phase of relatively minor volatility. Moreover, fluctuations in the LTP treatment seem to diminish over time. Note, however, that this tendency is rather fragile: in some markets, oscillations re-emerge after prices have seemingly converged (e.g., markets **L**2 and **L**9). In some cases, market prices converge to levels significantly higher than the fundamental value, after which they become unstable again (e.g., markets **L**1, **L**2 and **L**8). Conversely, in the HTP treatments, there is no apparent structural decrease in price volatility over time.

To corroborate these observations we consider two quantitative measures: the *interquartile range* (IQR) to evaluate price volatility and the median of the *relative absolute deviations* (RAD) from the fundamental value to evaluate mispricing.[22] Fig. 3 shows the IQR and the median RAD for all 31 experimental markets in the left and right panel, respectively. Each panel is divided into three areas that correspond to the **L**, **H**, and **HS** treatments. Each area, in turn, shows the statistics for each market computed over three different time periods: 1–145 (black dots in the middle), 11–50 (blue dots to the left), and 106–145 (red dots to the right). The disks indicate the median value for the corresponding treatment/time periods. Note the logarithmic

[22] The IQR is the length of the interval containing the middle half of the (ordered) market prices. The RAD from the fundamental value is defined as $\left|p_t - p^f\right| / p^f$ and was introduced as a measure for mispricing in Stöckl et al. (2010). Whereas they consider the mean RAD, we focus on the median of the RAD because the latter is more robust to outliers. We have considered other measures as well, e.g., the standard deviation of prices (for volatility) and the number of periods that the price is within 5% of the fundamental value p^f (for mispricing), but they tell essentially the same story.

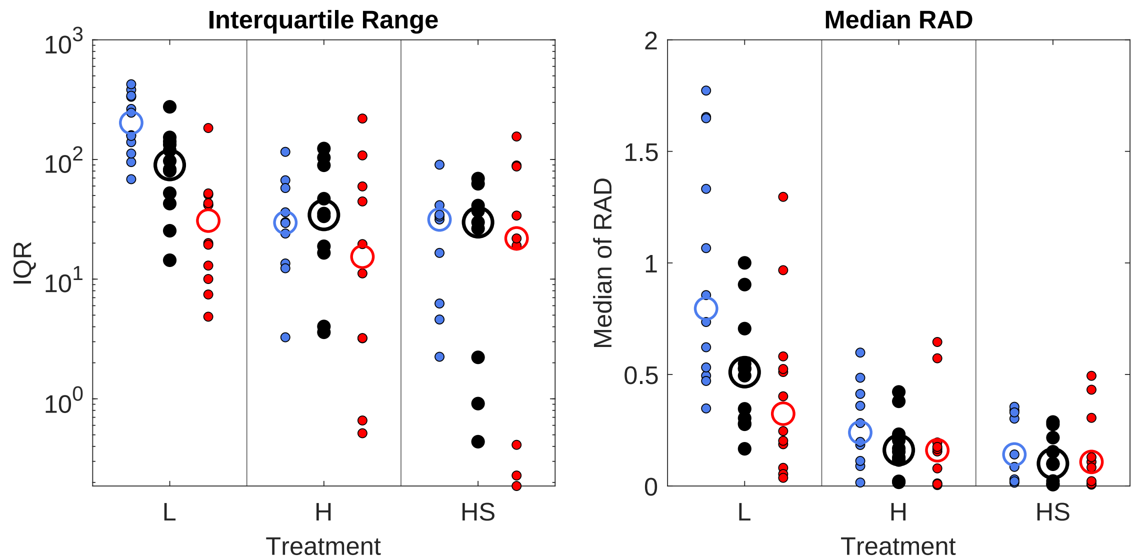


Figure 3: Measures for price volatility (IQR) and mispricing (Median of RAD).

scale of the vertical axis for the IQR. The corresponding numerical values for both measures can be found in Table B1 in Appendix B.

Fig. 3 illustrates that, for the LTP condition, both measures tend to decrease over time. Table B1 further shows that the IQR decreases for *all* markets except one (**L**12) and the median RAD decreases for all markets except two (**L**8 and **L**12).[23] This trend is statistically confirmed by a Wilcoxon signed-rank test, comparing the statistics between periods 11–50 and periods 106–145. See Table 2 for the p-values of the tests discussed here and in the rest of this section.

The p-values of the one-sided tests are 0.0005 for the IQR and 0.0105 for the median RAD. Thus, in treatment **L**, the IQR and the median RAD are significantly lower in the later periods of the experiment compared to the initial periods, at the 1% and 5% level, respectively. In contrast, for the HTP treatments **H** and **HS**, there is no statistically significant decrease in either price volatility or mispricing between periods 11–50 and periods 106–145. Therefore, we can reject Hypothesis 1 for treatments **H**

[23]The volatility increase in market **L**12 is due to one participant submitting extreme predictions (1 or 10 000) in 30 of the 34 periods after period 54. The price in **L**8 converges to a level of around 240 after period 60 and stays in this neighborhood for about 60 periods. Therefore, price volatility is low during that phase, whereas mispricing is high.

Data to compare		Statistics		Test
Data I	**Data II**	**IQR**	**Median RAD**	
Short-run vs Long-run				
L periods 11–50	**L** periods 106–145	0.0005***	0.0105**	one-sided
H periods 11–50	**H** periods 106–145	0.4229	0.1875	Wilcoxon
HS periods 11–50	**HS** periods 106–145	0.5449	0.4551	signed-rank
Time pressure				
L periods 11–50	**H** periods 11–50	0.0001***	0.0003***	one-sided
L periods 11–50	**HS** periods 11–50	0.0001***	0.0001***	MWW
L periods 106–145	**H** periods 106–145	0.2876	0.0463**	
L periods 106–145	**HS** periods 106–145	0.3746	0.0408**	
L periods 1–145	**H** periods 1–145	0.0259**	0.0011***	
L periods 1–145	**HS** periods 1–145	0.0038***	0.0003***	
Submission Button				
H periods 11–50	**HS** periods 11–50	0.6607	0.2507	two-sided
H periods 106–145	**HS** periods 106–145	0.8421	0.8421	MWW
H periods 1–145	**HS** periods 1–145	0.4002	0.4002	

Table 2: The p-values of the corresponding test (see the last column) for various comparisons. *, **, and *** indicate p values that are below 10%, 5% and 1%, respectively.

and **HS**, but not for treatment **L**.

Result 1. *Price volatility and mispricing decrease in the long run for the low time pressure treatment* **L**, *but not for the high time pressure treatments* **H** *and* **HS**.

Recall that in earlier LtF experiments, prices do not converge to their fundamental value within the standard duration of around 50 periods. This is consistent with the behavior we observe in our treatment **L**. Until period 50, all 12 markets exhibit large fluctuations in prices, with substantial reductions in only three of those markets (**L**4, **L**6, **L**10). However, as Result 1 indicates, price fluctuations tend to decrease after these first 50 periods. Long-run dynamics are significantly different compared to the beginning – although sometimes boom-and-bust cycles are reignited at a later stage. We conclude that, even if bubbles and crashes appear to be a persistent phenomenon, they are of a *transient* nature (at least in this stationary environment) and tend to diminish over time.

Turning now to the comparison of the two time pressure conditions reveals striking differences in price patterns during the first 50 periods. In particular, the HTP treatments demonstrate smaller fluctuations compared to the LTP treatment and earlier LtF experiments. For example, only two of the 19 HTP markets (**H**6 and **HS**5) exhibit larger fluctuations in periods 11–50, measured by the IQR, than the *most stable* market **L**8 in the LTP treatment (see Table B1). Moreover, when price fluctuations occur in the HTP treatments, they are often initiated by extreme predictions from individual participants.[24] In contrast, during the first 50 periods of the LTP treatment such 'outliers' are very rare, and price fluctuations can thus not be attributed to the idiosyncratic behavior of individual participants. Hence, increased time pressure seems to inhibit price volatility and the emergence of bubbles and crashes during the initial 50 periods. This is confirmed by comparing the IQR and the median RAD for periods 11–50 across treatments, by means of a one-sided Mann-Whitney-Wilcoxon (MWW) test. We find a statistically significant difference at the 1% level between treatments **L** and **H**, and between treatments **L** and **HS**, see Table 2.

Figs. 3 and B1 to B3 also suggest that the effect of time pressure on price behavior dissipates towards the end of the experiment. In particular, the mean value of the interquartile range for periods 106–145 is lower in treatment **L** than in treatments **H** and **HS**, while six of the seven markets with the highest IQR for these periods belong to one of the HTP treatments. While the differences in IQR between treatment **L** and treatments **H** and **HS** are not statistically significant for periods 106-145, the differences in median RAD are significant at the 5% level. This brings us to our second result, which confirms Hypothesis 2B.

Result 2. *Price volatility and mispricing are lower in treatments* **H** *and* **HS** *than in treatment* **L** *for periods* 11–50*. For periods* 106–145*, there is no statistically significant*

[24]See Figs. E6 and E7 in Internet Appendix E. For example, in markets **HS**2, **HS**5, and **HS**8 prices are quite stable until there is a sudden spike due to one participant predicting a price that is about 10 times the last observed market price.

difference in price volatility.

Lower price volatility and mispricing in HTP markets are most pronounced during the first 50 periods of the experiment. However, even when considering all periods (i.e., periods 1–145), price volatility and mispricing remain significantly lower than in the LTP markets.

Finally, regardless of which periods are examined, we find no significant difference in either IQR or median RAD between HTP treatments **H** and **HS**, indicating that incomplete predictions have a minor effect on price dynamics.

Non-responses. In all three treatments, there are instances where participants fail to submit their predictions in time. Over all groups and periods such 'non-responses' occur 236 times in treatment **L** (i.e., 2.25% of all $6 \times 12 \times 146$ predictions in that treatment), 317 times in treatment **H** (3.32%), and 941 times in treatment **HS** (10.96%). The higher number of non-responses in treatments **H** and **HS** may be attributed to the elevated time pressure imposed, whereas the large difference between these two treatments may be explained by the requirement in treatment **HS** that participants need to explicitly submit a prediction. Whereas in the HTP treatments the fraction of non-responses is highest in the first 10 periods and decreases thereafter, there is only one non-response (out of 720 predictions) in the first 10 periods of the LTP treatment, but the number of non-responses increases over time in that treatment.[25] This suggests that a majority of the non-responses in treatment **L** are due to inattention or boredom, whereas the non-responses in the early periods in treatments **H** and **HS** may be more likely caused by time pressure.

These non-responses may have an impact on the price dynamics through a *selection effect*. For example, participants with more non-responses may differ in their

[25]In particular, for treatment **L** the fraction of non-responses increases from 0.14% in the first 10 periods to 0.59% in periods 11–50 and 3.78% in periods 106–145. In treatment **H** (**HS**) the fraction of non-responses is 13.50% (17.22%) in the first 10 periods and decreases to 3.71% (10.28%) in periods 11–50 and 2.42% (11.90%) in periods 106–145.

average prediction accuracy from other participants (for a discussion and analysis of the selection bias induced by time pressure in an individual decision-making experiment, see Kocher et al., 2019). However, despite the higher number of non-responses in treatment **HS**, we do not find significant differences in interquartile range (IQR) or median RAD compared to treatment **H**, see Table 2. This suggests that the number of non-responses is not a significant factor in our experiment. Furthermore, we assess the correlation between the number of non-responses in each group during periods 1-10 and the IQR and median RAD in that group for periods 11-50 for treatments **H** and **HS**. None of these correlations is statistically significant,[26] indicating that the effect of higher time pressure on price dynamics cannot be explained by an increase in the number of non-responses in the high time pressure treatments.

3.2 Market expectations

Bubbles and crashes are often attributed to participants' tendency to coordinate on trend-extrapolating prediction strategies (Hommes et al., 2005; Anufriev and Hommes, 2012). The lower incidence of bubbles and crashes under the HTP condition may be due to participants' failure to coordinate on a common prediction strategy. To evaluate this possible channel, we compute the standard deviation of individual predictions for each period. When this measure is low, it indicates a higher level of coordination or, equivalently, a lower level of mis-coordination. To illustrate the initial evolution of this measure, Fig. 4 plots its median across all markets within a given treatment for each of the first 20 periods. The high standard deviation in the **H** treatment is consistent with the possibility of incomplete predictions submitted in this treatment. However, no noticeable distinction in the level of coordination

[26]The correlation between the number of non-responses and the IQR is -0.2677 for treatment **H** and 0.0655 for treatment **HS**, with p-values 0.4546 and 0.8670, respectively. The correlation between the number of non-responses and the median RAD is 0.0158 for **H** and 0.5059 for **HS**, with p-values 0.9655 and 0.1647, respectively. Note that the number of non-responses in the first 10 periods varies between 4 and 13 (out of 60 predictions) per market in treatment **H** and between 1 and 19 per market in treatment **HS**.

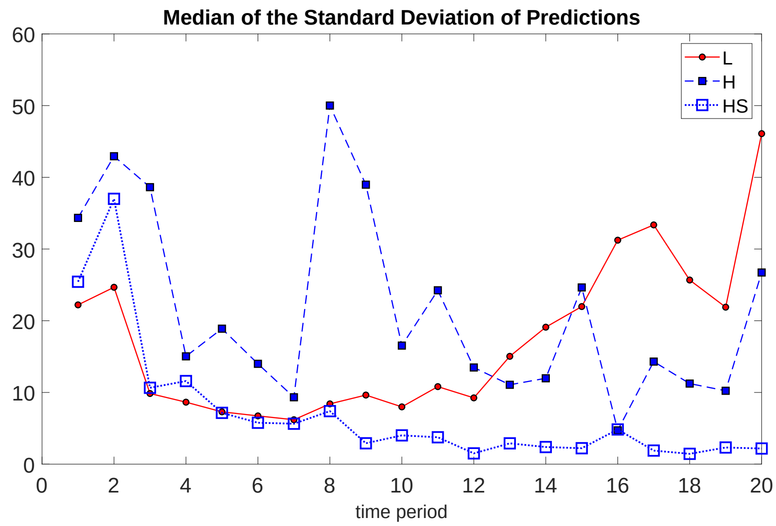


Figure 4: Time evolution of the median (over markets) standard deviation of the individual predictions in the three treatments.

between the **L** and **HS** treatments is apparent within the initial 10 periods. This observation is supported by a one-sided Mann-Whitney-Wilcoxon test.[27] Furthermore, after these periods, the coordination *decreases* in the **L** but not in **H** and **HS** treatments. We conclude that we cannot attribute the failure for bubbles to occur in the HTP treatment to a breakdown of coordination.

As another possibility to explain the differences between the treatments, we investigate the prediction strategies used in our experiment and study whether they change over time or are affected by the level of time pressure. We consider *market expectations*, $\bar{p}^e_{t+1}$, defined, for each market and for each period, as the average prediction taken over all participants who submitted a prediction for that period in that market. We then estimate the following prediction rule

$$\bar{p}^e_{t+1} = a + b_1 p_{t-1} + b_2 p_{t-2} + \nu_t \,, \tag{6}$$

[27] The hypothesis asserting that the standard deviations of predictions in the **L** treatment are not lower than in the **HS** treatment cannot be rejected at the 5% level when tested for each period between 1 and 20, except for period 2. It can also not be rejected when the data are aggregated over multiple periods, e.g., 1-10. Note that the forecasts for the first two periods are made without prior price information, while subsequently forecast variability could be influenced by price dynamics.

Treatment	Market	Periods 11-50			Periods 106-145		
		Const	Past prices		Const	Past prices	
		a	b_1	b_2	a	b_1	b_2
L	**Average**	105.17	1.61	-0.90	75.11	1.14	-0.57
	Median	94.06	1.63	-0.95	67.52	1.17	-0.63
H	**Average**	52.27	0.84	-0.19	55.75	1.24	-0.61
	Median	43.26	0.84	-0.05	60.52	1.44	-0.78
HS	**Average**	-0.90	1.19	-0.17	29.77	1.31	-0.48
	Median	-1.82	1.16	-0.05	26.48	1.43	-0.39

Table 3: The average and median of the estimated coefficients for prediction rule (6).

where a, b_1, and b_2 are fixed coefficients and ν_t is an error term. Prediction rules of type (6) have been shown to fit expectations quite well in previous LtF experiments (Hommes et al., 2005). Setting $g_0 = b_1 + b_2$ and $g_1 = -b_2$, Eq. (6) can be rewritten as

$$\bar{p}^e_{t+1} = a + g_0 p_{t-1} + g_1(p_{t-1} - p_{t-2}) + \nu_t \, .$$

Thus, expectations extrapolate trends in past prices, whenever g_1 is positive, or equivalently, b_2 in Eq. (6) is negative.

For each of the 31 markets in our experiment, we estimate the prediction rule (6) separately for periods 11–50 and periods 106–145. Table 3 reports the average and median of the estimated coefficients in each treatment. We observe that prediction rules structurally differ between the beginning and the end of the experiment as well as between the low and high time pressure treatments. In treatment **L**, for periods 11–50 the median value of the estimates of b_2 is -0.95, which suggests participants tend to extrapolate trends in their predictions. Specifically, if the price increases by one unit over the two most recent periods, participants' average price prediction increases by almost one additional unit. However, this trend extrapolation tendency declines towards the end of the experiment, and for periods 106–145 the median value of the estimates changes to -0.63. In contrast, there is virtually no trend extrapolation in the high time pressure treatments in the initial periods as the median estimates for periods 11–50 are -0.05 in both **H** and **HS** treatments.

Table B2 in Appendix B presents the estimated coefficients (and their statistical significance) for each market separately.[28] These coefficients are visually summarized in Fig. 5, with the left and right panels corresponding to periods 11–50 and 106–145, respectively. The scatter plots in each panel display the estimated coefficients of Eq. (6) for the 31 individual markets, in coordinates (b_1, b_2). The estimates for treatments **L**, **H**, and **HS** are depicted by the dots, filled squares, and non-filled squares, respectively. The triangle and parabola divide the space (b_1, b_2) to the regions with qualitatively different dynamics that rule (6) generates in the market.[29] Prices converge to the steady state p^f if the pair (b_1, b_2) lies inside the triangle and diverge if it is outside, and dynamics are monotone if (b_1, b_2) is above the parabola and oscillate if below. Particularly, when b_2 is slightly above the lower horizontal edge of the triangle, asset prices will oscillate and converge to the steady state value. The convergence will be slower when b_2 is closer to the edge. Moreover, when b_2 is below the edge, prices will fluctuate perpetually without converging to a steady state.

The location of the pairs of estimated coefficients in Fig. 5 illustrates what we discussed above. For periods 11–50 most of the estimated combinations of b_1 and b_2 for the LTP treatment are in the lower right corner of the triangle, often in close proximity to the boundary. This pattern indicates oscillating and slowly converging price dynamics. In contrast, the estimated combinations of b_1 and b_2 for the two HTP

[28] Three observations supporting the discussion above can be inferred from the table. First, for all 12 markets in treatment **L**, the estimated coefficient b_2 for periods 11–50 is significantly different from zero. Moreover, these coefficients are all negative. Second, for 11 of the 12 markets in treatment **L**, the estimate of b_2 is lower (in absolute value) for periods 106–145 than for periods 11–50. Third, taking the two high time pressure treatments together, the estimate of b_2 is only significantly different from zero in seven (out of the 19) markets for periods 11–50 and for 12 markets for periods 106–145. In fact, market expectations in a number of the markets in the HTP treatments closely resemble naïve expectations (i.e., $\bar{p}^e_{t+1} = p_{t-1}$), e.g., markets **H**2, **HS**3, and **HS**7.

[29] Eqs. (2) and (4) are complemented by rule (6) with $\nu_t = 0$ and $a = p^f(1 - b_1 - b_2)$. The latter restriction makes the rule 'consistent', i.e., predicting p^f in the fundamental steady state. The edges of the triangle are given by $b_2 - b_1 = 1 + r$ (left edge), $b_1 + b_2 = 1 + r$ (right edge), and $b_2 = -1 - r$ (bottom edge). The parabola is given by $b_2 = -b_1^2/(4(1 + r))$. See Proposition 2, Appendix B in Anufriev and Hommes (2012) for a formal derivation.

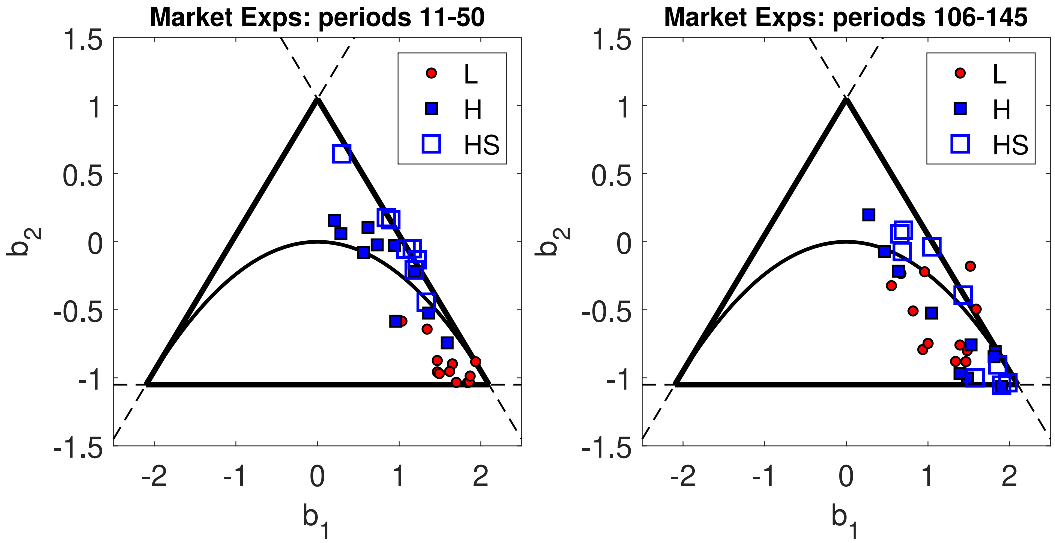


Figure 5: Estimated prediction rules across treatments.

treatments tend to cluster much closer to the center of the triangle, often with an estimated value of b_2 close to 0, implying that prices converge more rapidly to the fundamental value.

As for periods 106–145 (right panel of Fig. 5), in most of the markets in the two HTP treatments, the estimated market expectation rules drift towards the lower right corner of the triangle. A few exceptions exist where the estimated value of b_2 equals zero. Conversely, the rules for the LTP treatment move in the opposite direction.

The evolution of prediction rules demonstrates distinct patterns between treatments. In the initial stages of the experiment, the degree of trend extrapolation is much higher in the LTP treatment than in the HTP treatments. However, by the end of the experiment, this has been more or less reversed. We summarize this as

Result 3. *Trend-extrapolating prediction rules are used frequently under low time pressure, but are less common under high time pressure, especially for periods* 11–50.

3.3 Time pressure and the evolution of forecasting

How can the contrasting evolution of market expectation rules between low and high time pressure be explained?

One interpretation of Result 3 is that, as participants have more time to form predictions under LTP, they form predictions in a more sophisticated way than participants in the HTP condition. While in the latter, predictions tend to rely only on the last observed price, participants in the LTP condition have time to look for *patterns* in the time series of prices. This might allow them to identify trends in past prices and extrapolate those when forming predictions.

This explanation finds support in the responses obtained from the questionnaire, where participants described the prediction strategies they followed under each time pressure condition. Some examples related to the LTP condition include:

- "*I tried to mimic the price line and extrapolate based on its volatility*"
- "*I had enough time to look back and make forecasting decisions based on past trends*"
- "*I was mainly following previous patterns of the graph, but then I tried to stay inside a stable price*".

Among the descriptions for the strategies used in the HTP condition, we find:

- "*I had no time to think, so I almost always filled in the previous price*"
- "*I did what first came to mind and didn't have time for second thoughts which ended up being a good thing*".

These examples indeed suggest a higher tendency to use trend extrapolation strategies under low time pressure.[30]

[30]This is confirmed by a semantic analysis of the questionnaire answers. For example, the terms

Moreover, high time pressure may promote intuitive thinking. Besides collecting information on gender and age, we administered the standard three-question CRT test (Frederick, 2005) to all participants (further details about the questionnaire and the CRT test can be found in Internet Appendix D). To investigate whether gender, age, or the CRT score affect forecasting performance, we construct a variable E_i for each of the 186 participants, representing the median of that participant's absolute forecast errors in the first phase of the experiment. There is no significant effect of gender and age on forecasting accuracy. However, while the CRT score does not significantly affect forecasting performance in treatment **L**, it does influence forecasting accuracy in treatments **H** and **HS**. Specifically, one additional correct answer in the CRT test reduces E_i by 5% to 10% (Table B3). This suggests that intuitive thinking may indeed have been more prevalent in treatments **H** and **HS** than in treatment **L**.

As discussed above, more time for deliberation under low time pressure results in a tendency among participants in treatment **L** to adopt and coordinate on trend-extrapolating prediction rules. This can lead to significant fluctuations in market clearing prices. Higher price volatility often produces larger forecast errors for participants, consequently generating lower earnings. This can prompt participants to adapt their prediction rule and avoid strong extrapolation. This change in prediction behavior eventually leads to more stable prices.

For participants in the high time pressure treatments, the situation is different. Initially, due to the time constraints imposed, they tend to rely primarily on the most recent observed price. This reliance on the last price leads to relatively stable time series of market clearing prices. However, as time goes by and the participants gain more experience with the decision environment, they may get used to the time

"trend", "follow", and "pattern" appear significantly more often in descriptions of prediction strategies used in the LTP condition. In contrast, terms like "same" and "time" are more frequent in descriptions of strategies used in the HTP condition. Similarly, whereas "price" and "period" appear more often in the HTP condition, their plural versions, "prices" and "periods", crop up more frequently when discussing the LTP condition.

pressure and learn to quickly identify and extrapolate price trends.

4 Conclusion

In this paper, we investigated the effect of time pressure on market dynamics and price volatility in a Learning-to-Forecast experiment. Increased time pressure leads participants to switch away from complex prediction strategies (see e.g., Payne et al., 1988; Rieskamp and Hoffrage, 2008; Spiliopoulos et al., 2018, for similar findings).

Under low time pressure, participants tend to use prediction strategies that extrapolate trends in past prices. The positive feedback nature of asset markets reinforces these price expectations and leads to further realized and expected price increases. This process eventually produces large price bubbles. However, when decision time is limited participants tend to rely on less information when forming predictions. Typically, they only consider the most recent market price, which discourages the use of destabilizing trend-extrapolating strategies. These simpler prediction strategies often give rise to stable price dynamics, with prices quickly converging to their fundamental values and remaining relatively constant thereafter. As a consequence forecasting errors tend to be smaller, and these simpler prediction strategies outperform the more complex strategies used under low time pressure (cf. Gigerenzer and Goldstein, 1996; Goldstein and Gigerenzer, 2009). Overall, our experiment shows that time pressure, which is prevalent in financial markets, substantially impacts market price dynamics.

In addition, we show that the bubbles and crashes in asset prices that emerge under low time pressure conditions are often *transitory* in nature. Over time, participants learn to improve their price predictions, leading to a convergence of prices to the vicinity of the fundamental value. As a result, price volatility decreases, although this convergence remains relatively fragile and can be easily disrupted. We note that convergence typically sets in after a substantial number of periods, which is why it has not been identified in earlier LtF experiments that run for around 50 periods. Our results, therefore, suggest that while bubbles may initially emerge in a stationary market environment, they eventually dissipate with experience, but it may take a substantial amount of time.

There are, however, two important qualifications to this conclusion that future experimental work can address. First, as seen in this experiment, the market environment here is inherently prone to bubbles and crashes even after convergence. Relatedly, Anufriev et al. (2019) study the long-run ($2,000$ periods) behavior of artificial agents using genetic algorithms to learn their forecasting strategies in a similar market. They find that long phases of stability can suddenly be interrupted by phases of high price volatility. Second, excess price volatility can also emerge if the market environment is not stationary. For example, Xiong and Yu (2011) suggest that the lack of investor learning in the Chinese warrants bubble (2005–2008) is due to the inflow of new and naïve investors into the market.

References

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

Anufriev, M., C. Hommes, and T. Makarewicz (2019). Simple forecasting heuristics that make us smart: Evidence from different market experiments. *Journal of the European Economic Association 17*(5), 1538–1584.

Arifovic, J., C. Hommes, and I. Salle (2019). Learning to believe in simple equilibria in a complex OLG economy - evidence from the lab. *Journal of Economic Theory 183*, 106–182.

Assenza, T., P. Heemeijer, C. H. Hommes, and D. Massaro (2021). Managing self-organization of expectations through monetary policy: A macro experiment. *Journal of Monetary Economics 117*, 170–186.

Bao, T., C. Hommes, and T. Makarewicz (2017). Bubble formation and (in) efficient markets in learning-to-forecast and optimise experiments. *Economic Journal 127*(605), F581–F609.

Bao, T., C. Hommes, J. Sonnemans, and J. Tuinstra (2012). Individual expectations, limited rationality and aggregate outcomes. *Journal of Economic Dynamics and Control 36*(8), 1101–1120.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2018). Extrapolation and bubbles. *Journal of Financial Economics 129*(2), 203–227.

Berninghaus, S. K. and K.-M. Ehrhart (1998). Time horizon and equilibrium selection in tacit coordination games: Experimental results. *Journal of Economic Behavior & Organization 37*(2), 231–248.

Bloomfield, R. and M. O'Hara (1999). Market transparency: Who wins and who loses? *Review of Financial Studies 12*(1), 5–35.

Brock, W. and C. Hommes (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics & Control 22*(8-9), 1235–1274.

Busse, J. A. and T. C. Green (2002). Market efficiency in real time. *Journal of Financial Economics 65*(3), 415–437.

Campbell, J. Y., A. W. Lo, and A. C. MacKinlay (1997). *The Econometrics of Financial Markets*. Princeton University Press.

Carlé, T. A., Y. Lahav, T. Neugebauer, and C. N. Noussair (2019). Heterogeneity of beliefs and trade in experimental asset markets. *Journal of Financial and Quantitative Analysis 54*(1), 215–245.

Case, K. E. and R. J. Shiller (2003). Is there a bubble in the housing market? *Brookings Papers on Economic Activity 2003*(2), 299–362.

Case, K. E., R. J. Shiller, and A. Thompson (2012). What have they been thinking? Home buyer behavior in hot and cold markets. Technical report, National Bureau of Economic Research.

Cheah, E.-T. and J. Fry (2015). Speculative bubbles in bitcoin markets? An empirical investigation into the fundamental value of bitcoin. *Economics Letters 130*, 32–36.

Chen, D. L., M. Schonger, and C. Wickens (2016). oTree – An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance 9*, 88–97.

Coibion, O., Y. Gorodnichenko, and R. Kamdar (2018). The formation of expectations, inflation, and the Phillips curve. *Journal of Economic Literature 56*(4), 1447–1491.

Corbet, S., B. Lucey, and L. Yarovaya (2018). Datestamping the bitcoin and ethereum bubbles. *Finance Research Letters 26*, 81–88.

Duffy, J. (2016). Macroeconomics: a survey of laboratory research. In J. H. Kagel and A. E. Roth (Eds.), *Handbook of Experimental Economics Volume 2*, pp. 1–90. Princeton: Princeton University Press.

Duffy, J. and E. Hopkins (2005). Learning, information, and sorting in market entry games: Theory and evidence. *Games and Economic Behavior 51*(1), 31–62.

Ferri, G., M. Ploner, and M. Rizzolli (2021). Trading fast and slow: The role of deliberation in experimental financial markets. *Journal of Behavioral and Experimental Finance 32*, 100593.

Financial Times (2019). 'Fat finger' error sends airline stock on wild ride. https://www.ft.com/content/56188540-a2cf-11e9-974c-ad1c6ab5efd1, accessed December 8, 2020.

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives 19*(4), 25–42.

Friedman, D., S. Huck, R. Oprea, and S. Weidenholzer (2015). From imitation to collusion: Long-run learning in a low-information environment. *Journal of Economic Theory 155*, 185–205.

Füllbrunn, S., C. Huber, C. Eckel, and U. Weitzel (2022). Heterogeneity of beliefs and trading behavior: A reexamination. *Journal of Financial and Quantitative Analysis First View*, 1–25.

Fuster, A., D. Laibson, and B. Mendel (2010). Natural expectations and macroeconomic fluctuations. *Journal of Economic Perspectives 24*(4), 67–84.

Gigerenzer, G. and D. G. Goldstein (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review 103*(4), 650.

Goldstein, D. G. and G. Gigerenzer (2009). Fast and frugal forecasting. *International Journal of Forecasting 25*(4), 760–772.

Hennequin, M. (2018). Experiences and expectations in asset markets: an experimental study. Technical report, Working Paper, University of Amsterdam.

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

Hommes, C. (2021). Behavioral and experimental macroeconomics and policy analysis: A complex systems approach. *Journal of Economic Literature 59*(1), 149–219.

Hommes, C., A. Kopányi-Peuker, and J. Sonnemans (2021). Bubbles, crashes and information contagion in large-group asset market experiments. *Experimental Economics 67*(3), 1–20.

Hommes, C., J. Sonnemans, J. Tuinstra, and H. van de Velden (2005). Coordination of expectations in asset pricing experiments. *Review of Financial Studies 18*(3), 955–980.

Hommes, C., J. Sonnemans, J. Tuinstra, and H. Van de Velden (2008). Expectations and bubbles in asset pricing experiments. *Journal of Economic Behavior & Organization 67*(1), 116–133.

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review 93*(5), 1449–1475.

Kocher, M. G., D. Schindler, S. T. Trautmann, and Y. Xu (2019). Risk, time pressure, and selection effects. *Experimental Economics 22*(1), 216–246.

Kocher, M. G. and M. Sutter (2006). Time is money – Time pressure, incentives, and the quality of decision-making. *Journal of Economic Behavior & Organization 61*(3), 375–392.

Kopányi-Peuker, A. and M. Weber (2021). Experience does not eliminate bubbles: Experimental evidence. *Review of Financial Studies 34*(9), 4450–4485.

Kryvtsov, O. and L. Petersen (2021). Central bank communication that works: Lessons from lab experiments. *Journal of Monetary Economics 117*, 760–780.

Lahav, Y. (2011). Price patterns in experimental asset markets with long horizon. *Journal of Behavioral Finance 12*, 20–28.

Lindner, F. and M. Sutter (2013). Level-k reasoning and time pressure in the 11–20 money request game. *Economics Letters 120*(3), 542–545.

Marimon, R., S. E. Spear, and S. Sunder (1993). Expectationally driven market volatility: An experimental study. *Journal of Economic Theory 61*(1), 74–103.

Marimon, R. and S. Sunder (1993). Indeterminacy of equilibria in a hyperinflationary world: Experimental evidence. *Econometrica 61*(5), 1073–1107.

Marimon, R. and S. Sunder (1994). Expectations and learning under alternative monetary regimes: An experimental approach. *Economic Theory 4*(1), 131–162.

Mokhtarzadeh, F. and L. Petersen (2021). Coordinating expectations through central bank projections. *Experimental Economics 24*(3), 883–918.

Moritz, B., E. Siemsen, and M. Kremer (2014). Judgmental forecasting: Cognitive reflection and decision speed. *Production and Operations Management 23*(7), 1146–1160.

Palan, S. (2013). A review of bubbles and crashes in experimental asset markets. *Journal of Economic Surveys 27*(3), 570–588.

Payne, J. W., J. R. Bettman, and E. J. Johnson (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition 14*(3), 534.

Payne, J. W., J. R. Bettman, and E. J. Johnson (1993). *The Adaptive Decision Maker*. Cambridge University Press.

Phillips, P. C., S. Shi, and J. Yu (2015). Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. *International Economic Review 56*(4), 1043–1078.

Phillips, P. C., Y. Wu, and J. Yu (2011). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values? *International Economic Review 52*(1), 201–226.

Rieskamp, J. and U. Hoffrage (2008). Inferences under time pressure: How opportunity costs affect strategy selection. *Acta Psychologica 127*(2), 258–276.

Smith, A., T. Lohrenz, J. King, P. R. Montague, and C. F. Camerer (2014). Irrational exuberance and neural crash warning signals during endogenous experimental market bubbles. *Proceedings of the National Academy of Sciences 111*(29), 10503–10508.

Smith, V. L., G. L. Suchanek, and A. W. Williams (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica 56*(5), 1119–1151.

Sonnemans, J. and J. Tuinstra (2010). Positive expectations feedback experiments and number guessing games as models of financial markets. *Journal of Economic Psychology 31*(6), 964–984.

Spiliopoulos, L. and A. Ortmann (2018). The BCD of response time analysis in experimental economics. *Experimental Economics 21*(2), 383–433.

Spiliopoulos, L., A. Ortmann, and L. Zhang (2018). Complexity, attention, and choice in games under time constraints: A process analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition 44*(10), 1609.

Stanovich, K. E. and R. F. West (2000). Individual differences in reasoning: Implications for the rationality debate? *Behavioral and Brain Sciences 23*(5), 645–665.

Stöckl, T., J. Huber, and M. Kirchler (2010). Bubble measures in experimental asset markets. *Experimental Economics 13*(3), 284–298.

Veiga, H. and M. Vorsatz (2010). Information aggregation in experimental asset markets in the presence of a manipulator. *Experimental Economics 13*(4), 379–398.

Xiong, W. and J. Yu (2011). The Chinese warrants bubble. *American Economic Review 101*(6), 2723–53.

Zur, H. B. and S. J. Breznitz (1981). The effect of time pressure on risky choice behavior. *Acta Psychologica 47*(2), 89–104.

APPENDIX

A Asset-pricing Model

Our experiment is built around the standard asset pricing model of a long-lived risky asset. As in Brock and Hommes (1998), we make a number of simplifying assumptions to focus solely on the impact of price expectations.

Consider a market with a large number of investors, and two assets, a risk-free bond and risky equity. The time is discrete and indexed by t. Let r be the return of the risk-free bond, and y_t and p_t be the dividend and price per share of the risky asset, respectively. Let $W_{t,i}$ be the wealth of investor i, and $z_{t,i}$ be the investor's holdings of the risky asset bought at time t. Then, the investor i's wealth evolves as

$$W_{t+1,i} = W_{t,i}(1+r) + \left(p_{t+1} + y_{t+1} - (1+r)p_t\right) z_{t,i}\,, \tag{A.1}$$

where the expression in the parentheses of the second term on the right-hand side is the excess return of the risky asset. Investors are mean-variance maximizers[31] solving the problem

$$\max \left\{ \mathrm{E}_{t,i}[W_{t+1,i}] - \frac{a_i}{2}\, \mathrm{V}_{t,i}[W_{t+1,i}] \right\}\,, \tag{A.2}$$

where a_i is the risk aversion of investor i, and $\mathrm{E}_{t,i}[\,\cdot\,]$ and $\mathrm{V}_{t,i}[\,\cdot\,]$ are the beliefs about expected wealth and variance of wealth, respectively. All investors have the same risk aversion, $a_i \equiv a$, and have the same belief about the variance of price p_{t+1}, denoted as σ^2. Problem (A.2) subject to (A.1) gives the investor i's demand for the risky asset

$$z_{t,i} = \frac{\mathrm{E}_{t,i}\left(p_{t+1} + y_{t+1} - (1+r)p_t\right)}{a\sigma^2}\,. \tag{A.3}$$

[31]Equivalently, investors assume that their next period wealth is normally distributed and maximize the expected utility with the utility function given by $\exp(W_{t+1,i})$.

The price of the risky asset in each period is determined from equilibrium between the aggregate demand for the risky asset and its exogenous supply, set to zero. The temporary market equilibrium reads $\sum_i z_{t,i} = 0$. Solving it, we find the price

$$p_t = \frac{1}{1+r} \sum_i \left(\mathrm{E}_{t,i}[p_{t+1}] + \mathrm{E}_{t,i}[y_{t+1}] \right) = \frac{1}{1+r} \left(\sum_i \mathrm{E}_{t,i}[p_{t+1}] + \bar{y} \right) ,$$

where for the last equality we assume, as in the experiment, that the dividends are IID with mean $\bar{y}$.

Equation (2), which generates prices in the experiment, coincides with the last equation. Indeed, in the experiment, we have the fraction $n_t \in [0, 1)$ of 'robot' traders who expect that the price will be at the fundamental value $p^f = \bar{y}/r$, as given by (1). That is, $\mathrm{E}_{t,i}[p_{t+1}] = p^f$ for all these investors. The remaining weight of $1 - n_t$ is given to the human subjects who advise the large pension funds. The average of their forecasts (taken with equal weights) is denoted as $\bar{p}^e_{t+1}$ in Eq. (2).

B Data and Data Analysis

Figs. B1 to B3 display the market prices (blue lines) for each market in treatments **L**, **H**, and **HS**, respectively. The fundamental value $p^f = 126.4$, constant across all markets, is indicated by the dashed line.

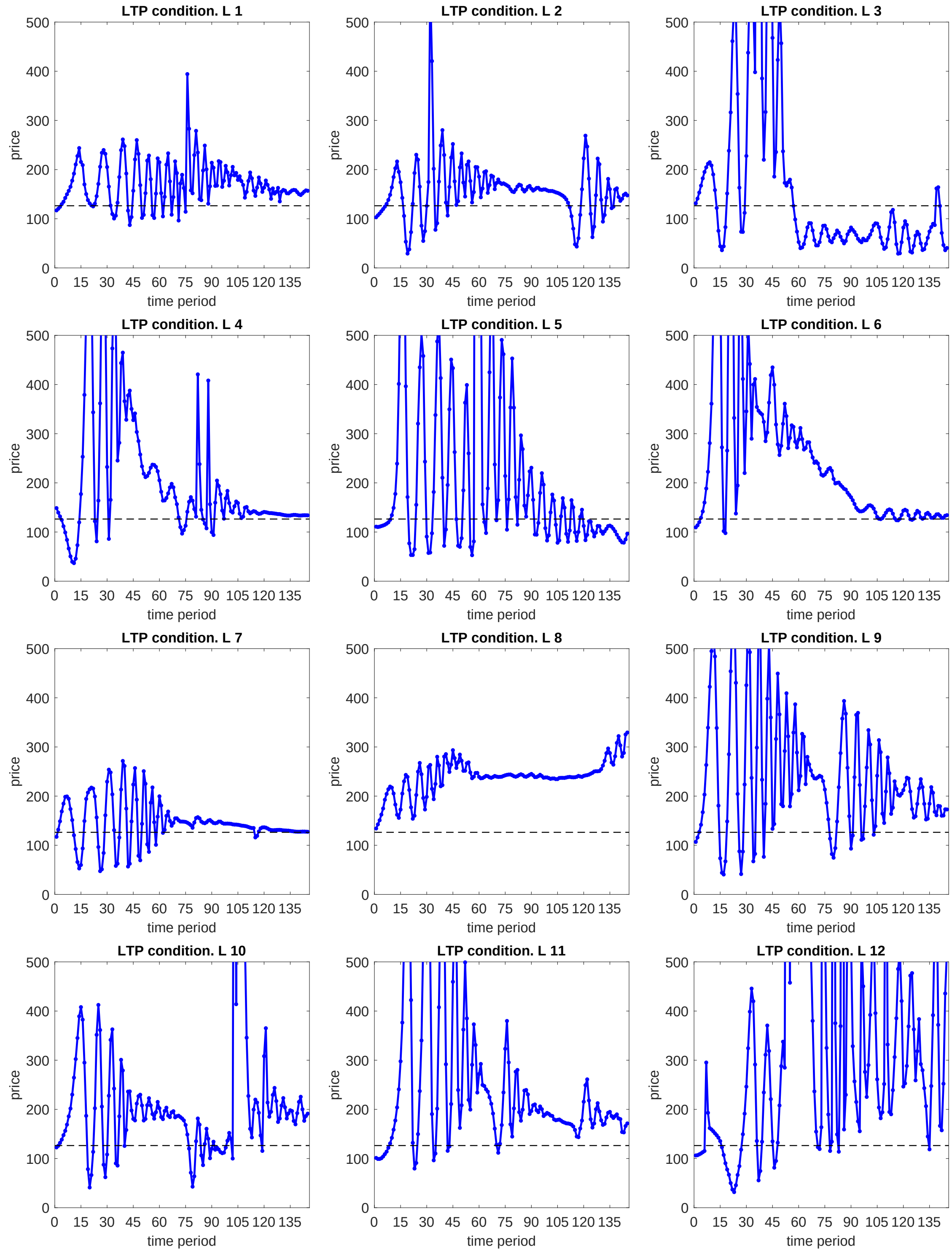


Figure B1: Prices in the **L** treatment (blue thick line). The dashed horizontal line indicates the fundamental price, $p^f = 126.4$.

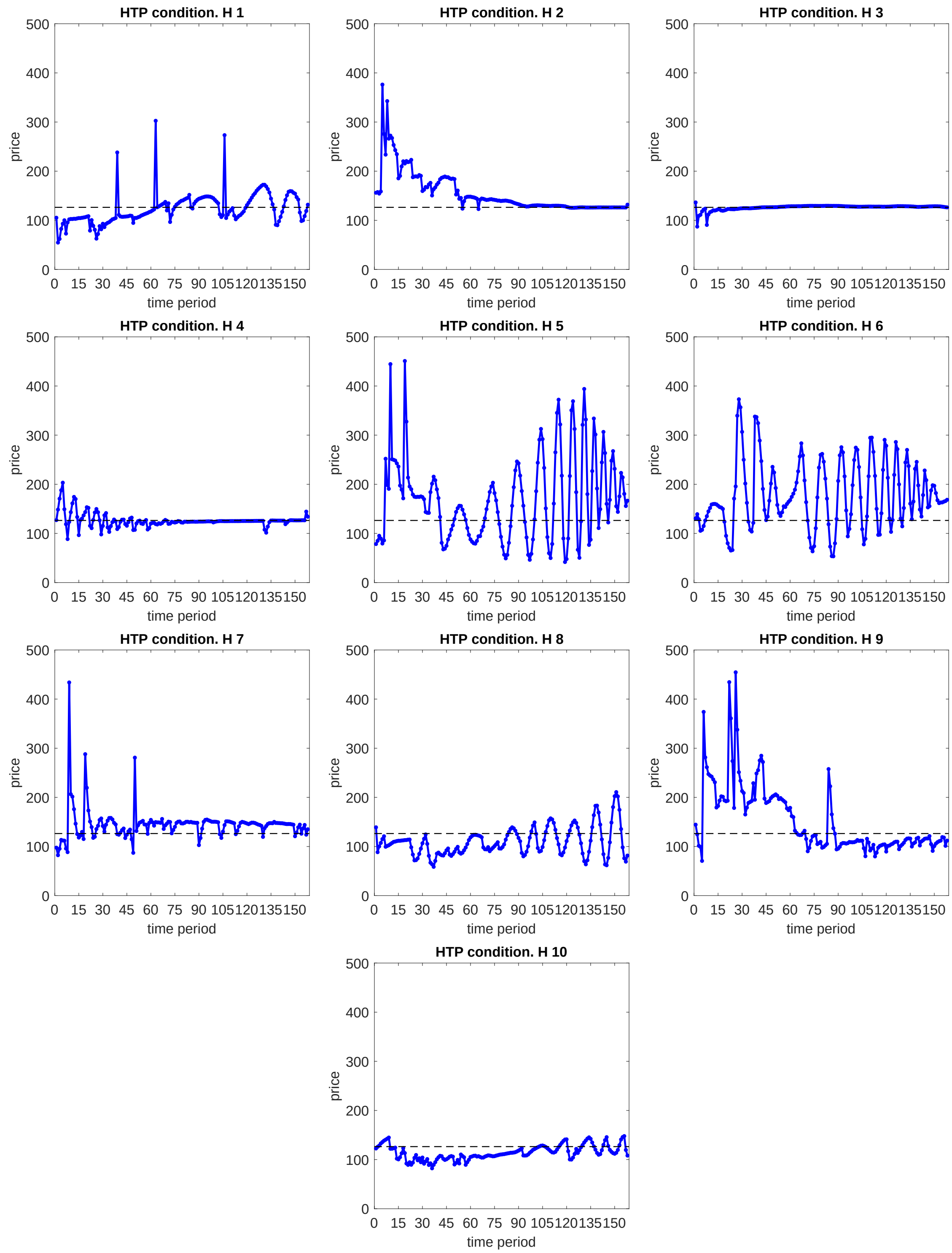


Figure B2: Prices in the **H** treatment (blue thick line). The dashed horizontal line indicates the fundamental price, $p^f = 126.4$.

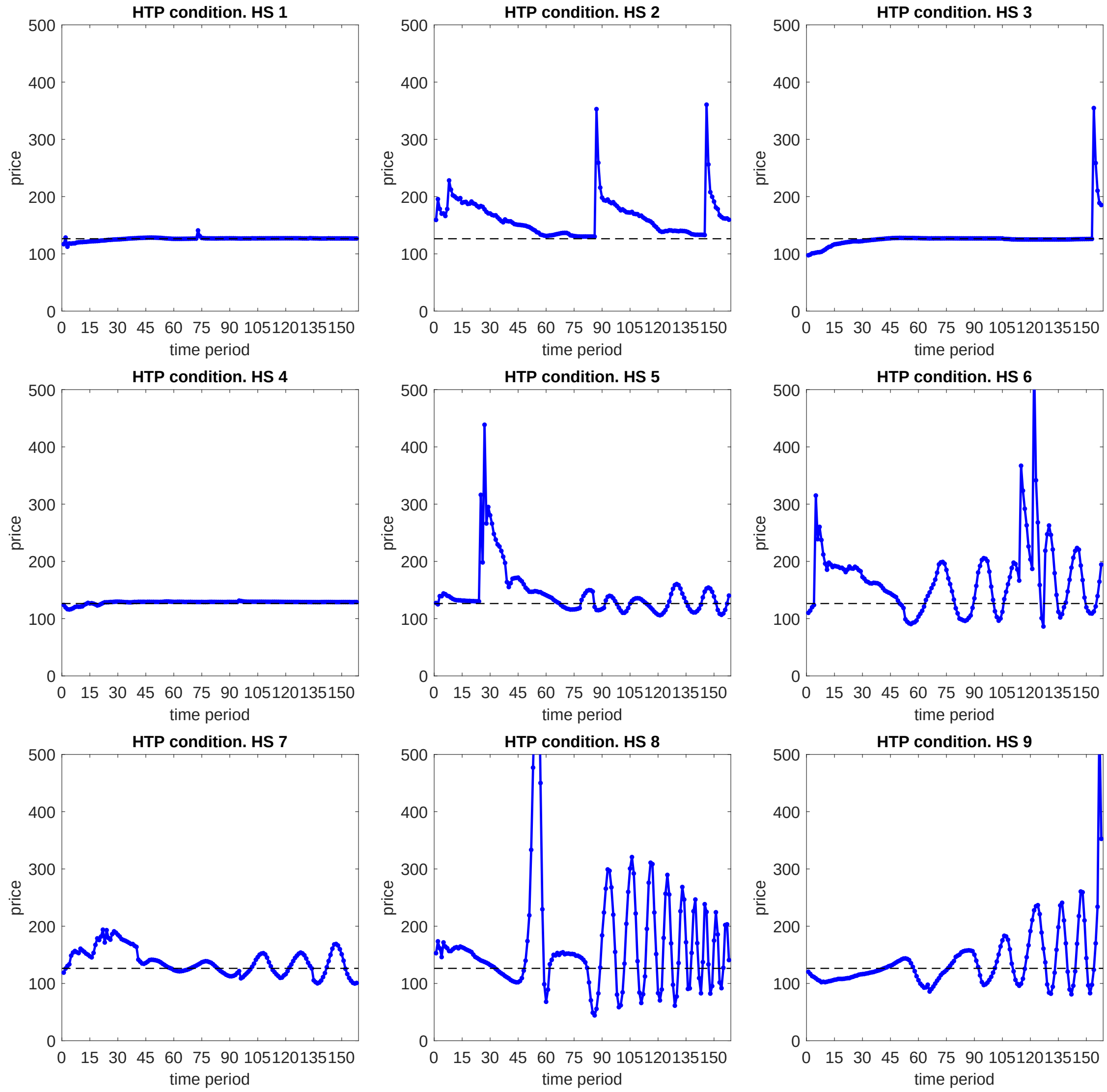


Figure B3: Prices in the **HS** treatment (blue thick line). The dashed horizontal line indicates the fundamental price, $p^f = 126.4$.

Table B1 contains the descriptive statistics of the experimental data for each market separately and as average and median by treatment. The market statistics are illustrated in Fig. 3 of the paper.

Treatment		**Interquartile range**			**Median of RAD**		
	Market	11-50	1-145	106-145	11-50	1-145	106-145
L	**L** 1	95.13	52.33	12.98	0.35	0.30	0.25
	L 2	112.10	42.80	50.82	0.49	0.28	0.19
	L 3	382.51	119.05	41.75	1.65	0.53	0.51
	L 4	264.41	97.25	4.85	1.65	0.28	0.08
	L 5	334.23	141.54	20.01	0.62	0.35	0.20
	L 6	245.96	152.97	10.03	1.77	0.55	0.05
	L 7	139.54	25.37	7.44	0.53	0.17	0.04
	L 8	68.42	14.39	41.64	0.86	0.90	0.97
	L 9	340.97	132.94	52.05	1.07	0.71	0.52
	L 10	159.67	82.60	43.14	0.74	0.49	0.58
	L 11	425.36	80.51	19.40	1.33	0.53	0.40
	L 12	157.32	275.52	183.22	0.47	1.00	1.30
	Average	227.14	101.44	40.61	0.96	0.51	0.42
	Median	202.82	89.92	30.83	0.80	0.51	0.32
H	**H** 1	13.52	35.33	44.60	0.18	0.15	0.15
	H 2	36.18	47.01	3.20	0.49	0.12	0.00
	H 3	3.27	3.60	0.66	0.02	0.02	0.01
	H 4	24.05	4.02	0.52	0.09	0.02	0.01
	H 5	67.02	123.48	219.85	0.41	0.42	0.65
	H 6	115.82	103.65	108.17	0.36	0.38	0.57
	H 7	29.91	16.57	3.22	0.11	0.17	0.17
	H 8	29.44	33.49	59.55	0.28	0.20	0.20
	H 9	57.73	89.44	11.17	0.60	0.23	0.18
	H 10	12.33	18.82	19.63	0.20	0.13	0.08
	Average	38.93	47.54	47.06	0.27	0.18	0.20
	Median	29.67	34.41	15.40	0.24	0.16	0.16
HS	**HS** 1	4.60	0.91	0.23	0.01	0.01	0.01
	HS 2	31.50	41.21	19.06	0.34	0.22	0.11
	HS 3	6.26	2.22	0.19	0.03	0.01	0.01
	HS 4	2.24	0.44	0.41	0.02	0.02	0.02
	HS 5	90.44	26.48	21.89	0.30	0.10	0.08
	HS 6	33.36	69.26	89.12	0.36	0.29	0.49
	HS 7	34.53	29.88	33.98	0.33	0.10	0.13
	HS 8	41.50	62.66	155.62	0.14	0.28	0.43
	HS 9	16.56	37.12	87.39	0.09	0.15	0.31
	Average	29.00	30.02	45.32	0.18	0.13	0.18
	Median	31.50	29.88	21.89	0.14	0.10	0.11

Table B1: The interquartile range (IQR) and median of the relative absolute deviation (RAD) from the fundamental value per market and selected periods.

Table B2 shows the estimated coefficients of the AR(2) model (6) for the market expectations. For each market, we estimate the model separately for periods 11–50 and for periods 106–145. The average and median of the estimates by treatment are reported in the corresponding rows. The LB and H columns show the p-values for the Ljung–Box Q test for zero autocorrelation and for the Engel test for residual heteroscedasticity. Fig. 5 plots the estimates of coefficients b_1 and b_2 from this table.

Treatment		**Periods 11-50**					**Periods 106-145**				
		Const	**Past prices**		**LB**	**H**	**Const**	**Past prices**		**LB**	**H**
	Market	a	b_1	b_2			a	b_1	b_2		
L	**L** 1	61.54^{***}	1.70^{***}	-1.03^{***}	**0.58**	**0.27**	92.75^{***}	0.67^{***}	-0.23	**0.15**	**0.98**
	L 2	100.32^{***}	1.03^{***}	-0.58^{***}	**1.00**	**0.90**	59.51^{***}	1.46^{***}	-0.88^{***}	**0.14**	**0.06**
	L 3	87.81	1.94^{***}	-0.88^{***}	**0.90**	**0.49**	21.00^{***}	1.39^{***}	-0.76^{***}	**1.00**	**0.89**
	L 4	128.37^{***}	1.65^{***}	-0.90^{***}	**0.32**	**0.08**	95.55^{***}	0.82^{***}	-0.51^{***}	0.00	**0.77**
	L 5	84.34^{***}	1.84^{***}	-1.03^{***}	**0.23**	**0.10**	91.18^{***}	0.94^{***}	-0.79^{***}	0.01	**0.29**
	L 6	215.77^{***}	1.62^{***}	-0.95^{***}	**0.18**	**0.15**	72.96^{***}	1.34^{***}	-0.88^{***}	**0.74**	**0.35**
	L 7	77.11^{***}	1.47^{***}	-0.96^{***}	**0.42**	0.01	34.47^{**}	0.96^{***}	-0.22	**1.00**	**0.97**
	L 8	78.78^{***}	1.34^{***}	-0.64^{***}	0.00	**0.93**	-8.66	1.59^{***}	-0.50^{***}	0.00	**0.14**
	L 9	164.87^{***}	1.49^{***}	-0.97^{***}	0.00	**0.92**	153.47^{***}	1.00^{***}	-0.75^{***}	**0.10**	0.00
	L 10	102.49^{***}	1.47^{***}	-0.87^{***}	**0.75**	0.00	-69.24^{*}	1.52^{***}	-0.18	**0.99**	0.01
	L 11	131.74^{***}	1.86^{***}	-1.03^{***}	**0.22**	**0.10**	62.08^{***}	1.48^{***}	-0.80^{***}	**0.55**	**0.23**
	L 12	28.89^{***}	1.87^{***}	-0.99^{***}	**0.05**	**0.34**	296.22^{***}	0.55^{***}	-0.32^{*}	**0.93**	**0.35**
	Average	105.17	1.61	-0.90			75.11	1.14	-0.57		
	Median	94.06	1.63	-0.95			67.52	1.17	-0.63		
H	**H** 1	63.61^{***}	0.21	0.16	**0.99**	**0.92**	72.70^{**}	0.28	0.20	**0.99**	0.00
	H 2	20.52	0.94^{***}	-0.03	**0.15**	**0.56**	-2.21	1.82^{***}	-0.81^{***}	**0.73**	**0.85**
	H 3	4.53	1.18^{***}	-0.22^{*}	**0.99**	0.05	4.73	1.81^{***}	-0.85^{***}	**0.06**	**0.17**
	H 4	80.18^{***}	0.96^{***}	-0.58^{***}	**0.77**	**0.61**	59.49^{***}	1.04^{***}	-0.52^{***}	**0.98**	**0.66**
	H 5	48.71^{*}	0.62^{***}	0.11	**1.00**	**0.83**	112.25^{***}	1.48^{***}	-1.00^{***}	**0.09**	**0.79**
	H 6	37.81^{*}	1.36^{***}	-0.52^{***}	**0.13**	**0.87**	115.31^{***}	1.39^{***}	-0.97^{***}	**0.15**	**0.82**
	H 7	98.28^{***}	0.29	0.06	**1.00**	**0.54**	85.24^{***}	0.64^{***}	-0.22	**0.41**	**0.96**
	H 8	12.40^{**}	1.59^{***}	-0.74^{***}	**0.37**	**0.67**	19.67^{***}	1.90^{***}	-1.06^{***}	**0.73**	0.01
	H 9	128.96^{***}	0.57^{***}	-0.08	**0.63**	**0.91**	61.54^{***}	0.47^{**}	-0.07	**0.50**	**0.41**
	H 10	27.72^{**}	0.73^{***}	-0.02	**0.68**	**0.41**	28.74^{***}	1.53^{***}	-0.76^{***}	**0.87**	**0.97**
	Average	52.27	0.84	-0.19			55.75	1.24	-0.61		
	Median	43.26	0.84	-0.05			60.52	1.44	-0.78		
HS	**HS** 1	-3.50^{***}	1.08^{***}	-0.05	**0.98**	**0.28**	27.32	0.70^{***}	0.08	**0.79**	**0.97**
	HS 2	-8.60^{**}	0.90^{***}	0.16	**0.59**	**0.83**	-5.03^{*}	1.43^{***}	-0.39^{**}	**0.16**	0.01
	HS 3	2.93	1.18^{***}	-0.21	**0.75**	0.05	35.84^{***}	0.66^{***}	0.06	**0.88**	0.00
	HS 4	14.60^{**}	1.33^{***}	-0.44^{***}	**0.80**	0.03	-1.24	1.05^{***}	-0.04	**1.00**	**0.63**
	HS 5	17.05	0.30^{**}	0.65^{***}	**0.90**	**0.71**	7.42^{***}	1.98^{***}	-1.04^{***}	**0.66**	**0.81**
	HS 6	-17.34^{***}	1.16^{***}	-0.05	**0.30**	**0.09**	88.53^{**}	0.69^{***}	-0.07	**0.92**	**0.85**
	HS 7	-1.31	0.84^{***}	0.18	**0.84**	0.03	7.09	1.85^{***}	-0.90^{***}	**0.75**	0.03
	HS 8	-1.82	2.66^{***}	-1.63^{***}	**0.81**	0.00	81.53^{***}	1.58^{***}	-1.00^{***}	**0.63**	**0.88**
	HS 9	-10.13^{***}	1.22^{***}	-0.13	**0.85**	**0.94**	26.48^{***}	1.90^{***}	-1.05^{***}	**0.54**	**0.59**
	Average	-0.90	1.19	-0.17			29.77	1.31	-0.48		
	Median	-1.82	1.16	-0.05			26.48	1.43	-0.39		

Table B2: Market expectations. For parameters' estimates, * denotes the significance at the 10% level, ** at the 5% level, and *** at the 1% level. For the Ljung-Box and Engel specification tests whose p-values are in LB and H columns, the bold font indicates rejection of residual structure (autocorrelations or heteroskedasticity) at the 5% level.

Table B3 reports the results of regressions of variable E_i, the median of the absolute forecasting error of subject i over all prices in the first block of the experiment on age, gender, and the CRT score. See Section 3.2 for discussion.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
treatment	pooled		L		H		HS	
dependent variable	E_i							
CRT	−0.465	−0.971**	−0.122	−1.295	−0.685**	−0.731***	−0.732**	−0.667**
	(0.568)	(0.476)	(1.416)	(1.189)	(0.269)	(0.270)	(0.278)	(0.278)
H	−19.177***	−20.135***						
	(6.553)	(7.037)						
HS	−24.628***	−25.180***						
	(6.749)	(7.251)						
age		−0.269**		−0.376		−0.097		−0.109
		(0.127)		(0.247)		(0.120)		(0.081)
female		−0.350		−0.367		−0.887		1.218*
		(1.205)		(3.080)		(0.661)		(0.709)
Constant	30.436***	38.159***	29.897***	40.138***	11.599***	14.188***	6.258***	7.804***
	(4.507)	(5.576)	(2.755)	(5.905)	(0.518)	(2.684)	(0.554)	(2.020)
Random effects	Yes	Yes	No	No	No	No	No	No
Group fixed effects	No	No	Yes	Yes	Yes	Yes	Yes	Yes
Observations	186	185	72	71	60	60	54	54
R^2			0.00	0.06	0.12	0.16	0.14	0.22

Table B3: Linear regression of the median absolute forecasting error E_i of subject i on age, gender, and CRT score. Standard errors are in parentheses; *, **, and *** indicate that the corresponding p values are below 10%, 5% and 1%, respectively.

C Experimental Instructions – Internet Appendix

Thank you for participating in today's experiment. The experiment is anonymous, your choices will not be linked to your name. You will be paid privately at the end of the experiment, after all participants have finished. The main part of the experiment consists of two phases. In each phase, you have to make a large number of consecutive stock market predictions. After the main part of the experiment, but before the payments are made, you will have to do one short additional task and fill out a short questionnaire.

Please read these instructions carefully. You are not allowed to use your mobile phone or communicate with other participants. If at any moment in time you have a question, please raise your hand, and one of the experimenters will come to your seat to assist you.

General information.

Your role is that of a financial advisor to a pension fund that wants to optimally invest a large amount of money. The pension fund has two investment options: a risk-free investment and a risky investment. The risk-free investment is putting money into a bank account, paying a fixed interest rate. The alternative, risky investment is an investment in the stock market. In each time period, the pension fund has to decide which fraction of its money to put in the bank account and which fraction of the money to spend on buying stocks. To make an optimal investment decision, the pension fund needs an accurate prediction of the price of the stocks. As its financial advisor, you have to predict the stock market price (in euro) during a number of consecutive time periods.

Information about the stock market.

The stock market price is determined by equilibrium between the demand and supply of stocks. The supply of stocks is fixed during the experiment. The demand for stocks is determined by the aggregate demand of a number of large pension funds active in the stock market. Some of these pension funds are advised by a participant to the experiment, others use a fixed investment strategy. The price of the stocks is determined by market equilibrium, that is, the stock market price in period t will be the price for which aggregate demand equals supply.

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 bank account of the risk-free investment pays a fixed interest rate per time period. The holder of the stocks receives a dividend payment in each time period. These dividend payments are uncertain and may vary over time. Economic experts of the pension funds have computed the average dividend payments per time period. The return on the stock per time period for a pension fund is given by the dividend payment for that period and profits or losses from possible price changes of the stock. As the financial advisor of a pension fund, you are not asked to forecast dividends, but you are only asked to forecast the price of the stock in each time period. Based on your stock price forecast, the pension fund that you advise will make an optimal investment decision. The higher your price forecast for the next period is, the larger will be the fraction of money invested by this pension fund in the stock market in the current period, so the larger will be its demand for stocks.

Forecasting task of the financial advisor.

The only task of the financial advisors in this experiment is to forecast the stock market price in each time period as accurately as possible. For each period, there will be only a limited amount of time to make a prediction. After this *limited amount of time*, the next period starts. This **time limit** will be the same for each period in the same phase, but it may be different between the two phases. Additionally, there may be a **waiting time** in either one of the phases, which again will be the same for each period in the same phase. Submitting a prediction in a period is only possible after the waiting time for that period is over. The relevant time limit and, if applicable, waiting time will be announced immediately prior to the start of each phase.

In the first period of each phase of the experiment, you have to predict the price for the first period, and in the second period, you will have to predict the price for that second period. *Only after the second period has finished, the realized price for the first period will be announced.* After that, you have to give your prediction for the price in the third period. When the third period is finished, the realized stock price for the second period will be revealed, and so on. This process continues until the end of the phase.

To forecast the stock price p_t for period t, the available information provides:

- past prices up to period $t-2$
- your own past predictions up to period $t-1$

Your past predictions and prices are represented both in a table and graphically, see the computer interface below for illustration. The main graph and the table show prices (red) and your predictions (blue) in the last 20 periods. You can scroll back/up to see prices and predictions from earlier periods. In the upper left corner of the screen, a smaller graph shows all your predictions and prices over the entire horizon up to the current period, and in the upper right corner, you find the relevant interest rate and the average dividend payment. The vertical axis of the graphs rescales when prices or predictions increase beyond the current scale of the graph.

You can make a prediction by clicking with your **mouse** in the graph. The value you click will then be shown as your current prediction in the graph and the box on the top of the screen, and that value will automatically change if you use your mouse to click on another value. You can change that value as often as you want to, within the given time limit, by moving and clicking the mouse. During the *waiting time*, you can already click with the mouse, but only after the waiting time is over does the box appear and show a clicked value.

At the top of the screen, you see a timer with the remaining time for the current period. To submit your prediction, you can **either click the "Submit" button with the mouse or press "Enter" on the keyboard before the time for the period is up**. If the time is up and you have not pushed the "Submit" button or pressed "Enter" yet, you made no prediction for this period. Although using the mouse is the fastest way to submit a prediction, you can also directly *type* your forecast in the box.

The price of the stock must be positive. The vertical scale of the graphs does not represent an upper bound for prices or predictions: the vertical axis of the graphs will rescale when prices or predictions change. However, the *first two prices in the first phase* are likely to lie between 0 and 200, and the *first two prices in the second phase* are likely to lie between 0 and 100. Upon completion of the instructions, you will first enter two practice rounds to get familiar with the computer program.

Differences between the two phases

The experiment consists of two phases. Prices in the second phase will be independent of prices and predictions from the first phase. Your task, predicting the price of the risky asset, will be the same in both phases, but there are also some important differences between the two phases. In particular, the **time limit** for making a decision and the **waiting time** in each period may be different for the two phases, and the **interest rate** and **average dividend payment** might be different between the two phases as well. Just prior to the start of each phase, the time limit and waiting time for each period in that phase and the relevant interest rate and average dividend payment will be announced. Also, the likely price range for the first two prices (which are between 0 and 200 for the first phase and between 0 and 100 for the second phase) will be repeated then.

There are two other differences between the two phases. First, the set of pension funds (and therefore, the group of financial advisors/participants) active in your market may not be the same in the two phases. Second, the number of periods might be different between the two phases. In particular, the number of periods is unknown for each phase, but it will be between 120 and 180 periods for each phase.

Earnings.

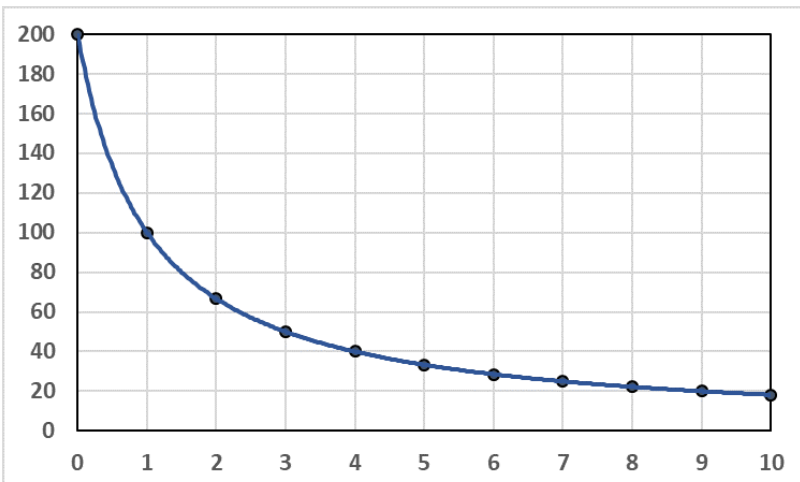
Earnings will be fully determined by forecasting accuracy. At the end of the experiment, 10 periods from the first and 10 periods from the second phase are randomly chosen, all with the same probability. You will be paid for your forecasting accuracy in these 20 periods. The better you predict the market price in these periods, the higher your earnings will be. Note, if a period is selected in which you did not make a prediction, your earnings for that period will be zero. The number of points you get for a period is given by:

$$\textbf{Points for period } t \ = \frac{200}{1 + \textbf{Prediction Error}},$$

where **Prediction Error** is your prediction error in that period (the difference between your forecast for period t and the realized market price p_t for period t). The relationship between the error and the number of points for some particular cases is given below:

Prediction Error	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Points	200	100	66.67	50	40	33.33	28.57	25	22.22	20	18.18

The figure below shows the relation between the number of points you score (vertical axis) and the prediction error. Notice that the table presents only some possibilities for your point earnings (the table is not exhaustive) and that the number of points you earn decreases more slowly as your prediction error increases.



The number of points earned in a period that is randomly selected to be paid will be transferred to euros at a rate of 100 points for one euro. For example, suppose for period 16 you predict that the price will be equal to 45, but the actual market price is 48, then your error is $48 - 45 = 3$, and you will, therefore, earn 50 points. If period 16 is randomly selected as one of the periods for which you are paid, you will earn €0.50 for that period.

As a second example: suppose you predict a price of 90 for period 60, and the price in that period turns out to be 88.3. Now your prediction error is $90 - 88.3 = 1.7$

and you would receive $\frac{200}{1+1.7} = 74.04$ points. This will give you €0.74 if that period is selected for payment. Note that if you are paid for a period in which you predicted the price correctly, you will obtain 200 points, or €2.00, for that period. On the computer screen that you see during the experiment, you can find (under "potential earnings"), for each period, the number of points that you would get if that period is selected for payment.

Your total earnings for the experiment will be the earnings from the 20 periods that are randomly selected, plus a participation fee of €10.00. Click the "Next" button on your screen to proceed to the practice rounds.

D CRT and Questionnaire – Internet Appendix

At the end of the experiment, participants had to answer 3 questions of the standard CRT test (Frederick, 2005), where we modified the first question. The questions are:

Question 1. A shirt and a jacket cost €220 in total. The jacket costs €200 more than the shirt. How many Euros does the shirt cost? (in Euros)

Question 2. If it takes 5 machines 5 minutes to make 5 widgets, how many minutes would it take 100 machines to make 100 widgets? (in minutes)

Question 3. In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how many days would it take for the patch to cover half of the lake? (in days)

Each of these questions have both a correct answer (10, 5, and 47, respectively) and an intuitive but wrong answer (20, 100, and 24, respectively).

Next, they completed the individual questionnaire asking a number of questions about their demographic characteristics, study programs, experience with the financial market and experiments, and their perception of time pressure. The mean age was 21.96, ranging between 18 and 59 years, with a standard deviation of 4.58. Table D1 provides the statistics of the responses to the other questions.

At the end of this questionnaire, we also asked participants to describe their strategies by answering three questions:

- How would you describe your strategy in the phase where you had a time limit of 6 seconds?
- How would you describe your strategy in the phase where you had a time limit of 25 seconds?
- Was there a difference between how you formed decisions in phase 1 and in phase 2, and if so, can you indicate how this differed?

The analysis of these answers is discussed in Section 3.3.

E Individual Predictions – Internet Appendix

Figs. E1 to E3 show the evolution of prices and individual predictions in all 12 markets of treatment **L**, under the LTP condition. Figs. E4 and E5 show the evolution of prices and individual predictions in all 10 markets of treatment **H**, under the HTP condition. Finally, Figs. E6 and E7 show the evolution of prices and individual predictions in all 9 markets of treatment **HS**, also under the HTP condition.

In all plots, the prices are shown by the black, thick line and the individual predictions (that are often close to the price) are shown by the thin, colored lines.

CRT		**Share**
Question 1:	Correct	58.9%
	Intuitive	36.8%
	Other	4.3%
Question 2:	Correct	46.5%
	Intuitive	41.1%
	Other	12.4%
Question 3:	Correct	55.1%
	Intuitive	25.4%
	Other	19.5%
Correct Answers	3	32.4%
	2	21.6%
	1	20.0%
	0	25.9%
Intuitive Answers	3	11.9%
	2	19.5%
	1	28.6%
	0	40.0%
Questionnaire		**Share**
Gender		
	Female	54.1%
	Male	45.9%
Study program		
	Faculty of Econ and Business	65.4%
	Faculty of Law	7.0%
	Faculty of Science, Math, and Computer Science	2.7%
	Faculty of Humanities	4.3%
	Faculty of Medicine/Dentistry	1.6%
	Faculty of Behavioral Science – other than Psychology	13.0%
	Other	3.2%
Previous participation in economic experiments		
	No	13.0%
	Yes, only once	11.9%
	Yes, more than once	75.1%
Previous participation in the price forecasting experiments		
	No	32.4%
	Yes, only once	37.8%
	Yes, more than once	29.7%
Trading experience in financial markets		
	No	79.5%
	Yes, a little bit	17.8%
	Yes, I frequently trade	2.7%
Enough time to make a decision		
	No, I felt time pressure in some of the periods in HTP, but not in LTP	30.3%
	No, I felt time pressure in most of the periods in HTP, but not in LTP	40.5%
	No, I felt time pressure in almost all of the periods in HTP, and in some periods in LTP	8.6%
	No, I felt time pressure in most of the periods in both blocks	2.7%
	Yes, in all periods	17.8%

Table D1: CRT and Questionnaire information from 185 participants.

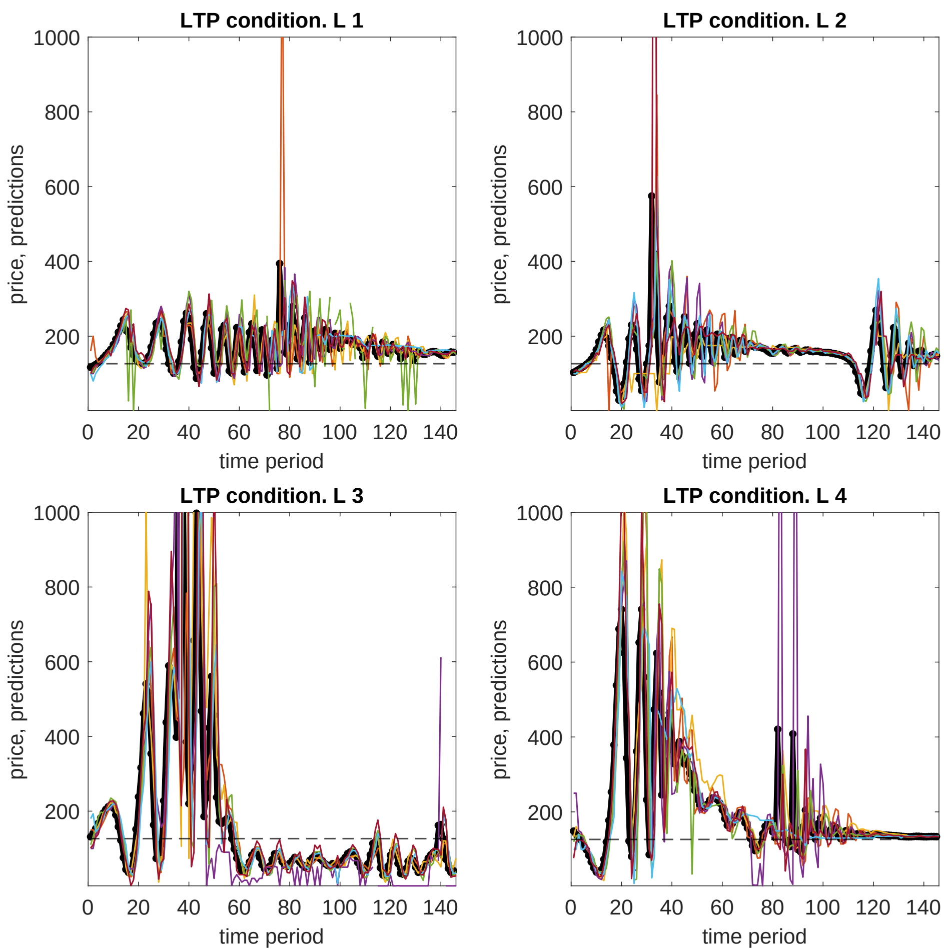


Figure E1: Prices (black) and predictions (colored) in markets 1-4 of treatment **L**.

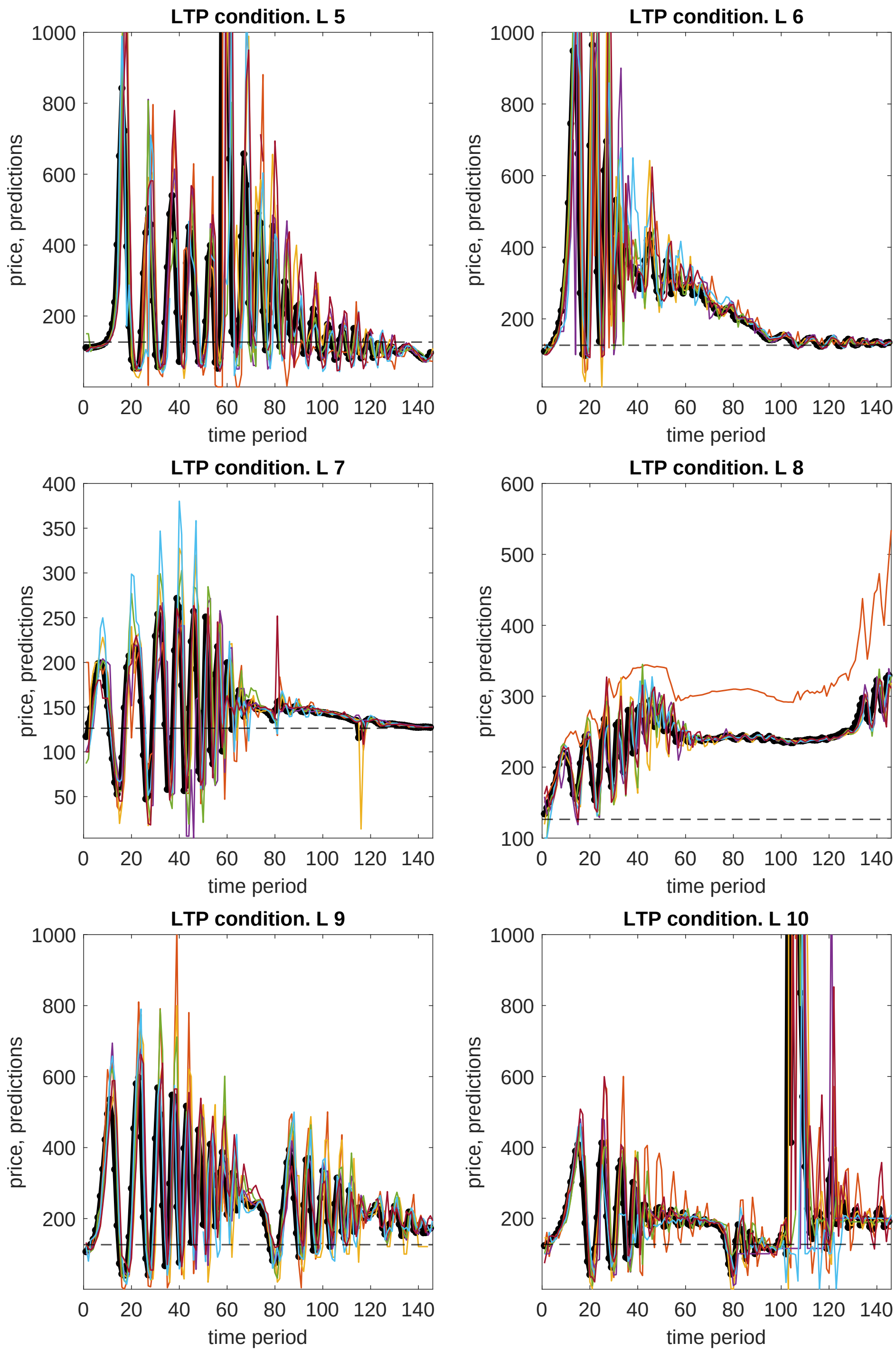


Figure E2: Prices (black) and predictions (colored) in markets 5-10 of treatment **L**.

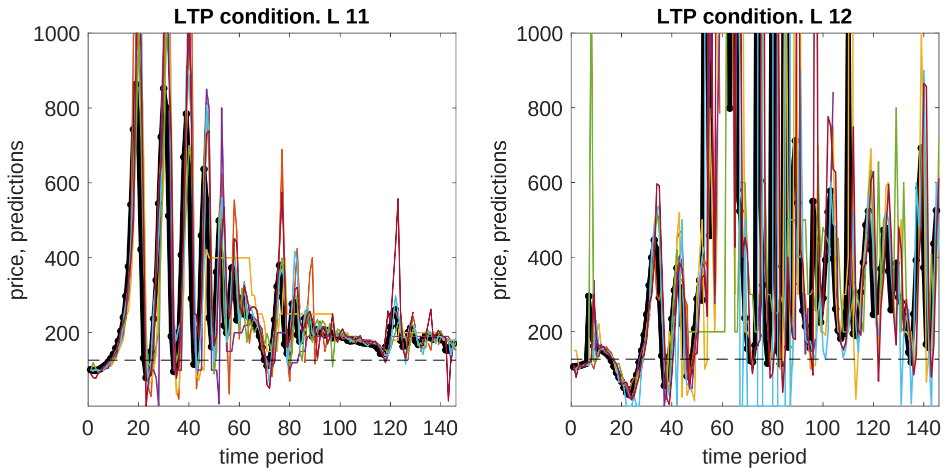


Figure E3: Prices (black) and predictions (colored) in markets 11-12 of treatment **L**.

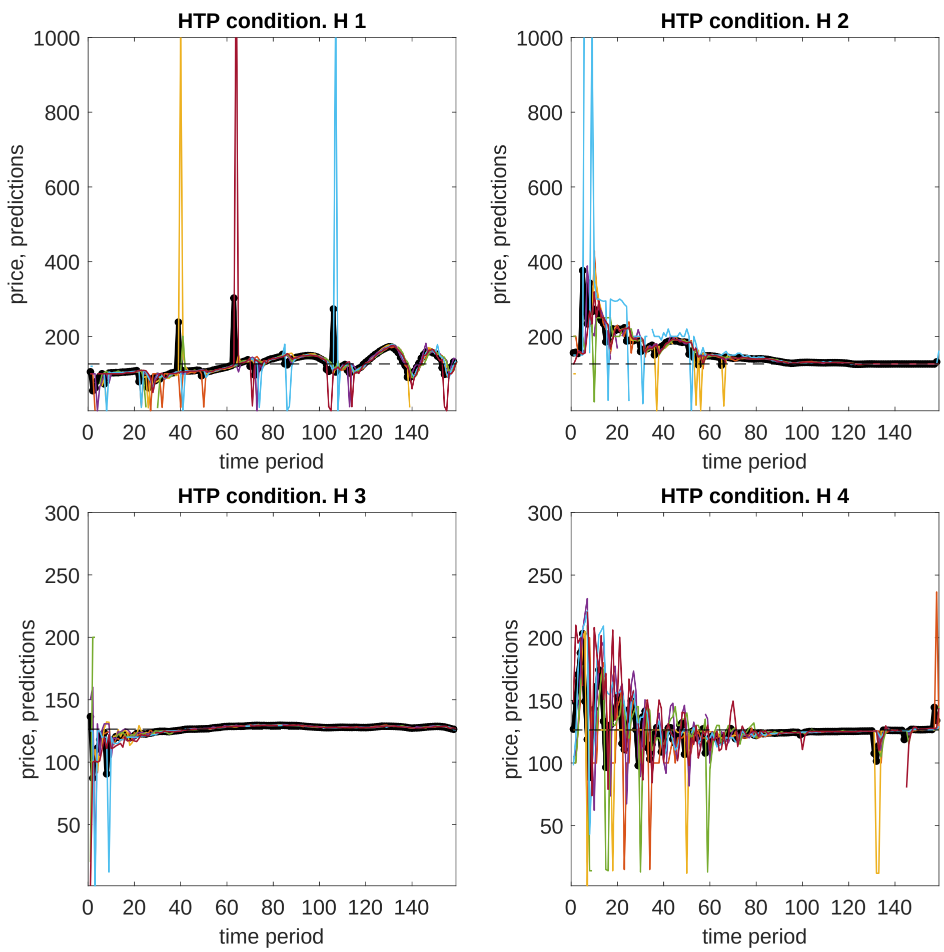


Figure E4: Prices (black) and predictions (colored) in markets 1-4 of treatment **H**.

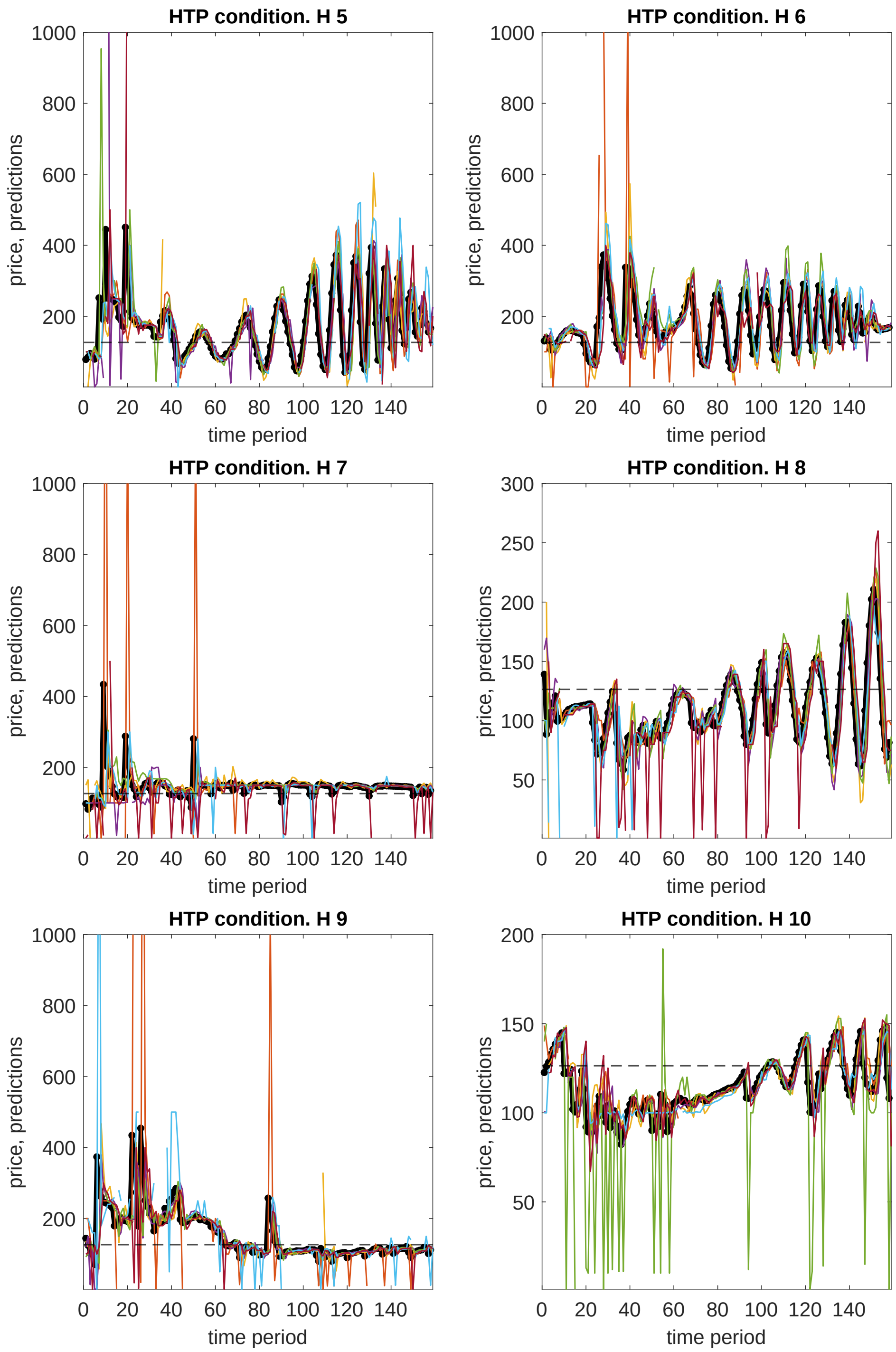


Figure E5: Prices (black) and predictions (colored) in markets 5-10 of treatment **H**.

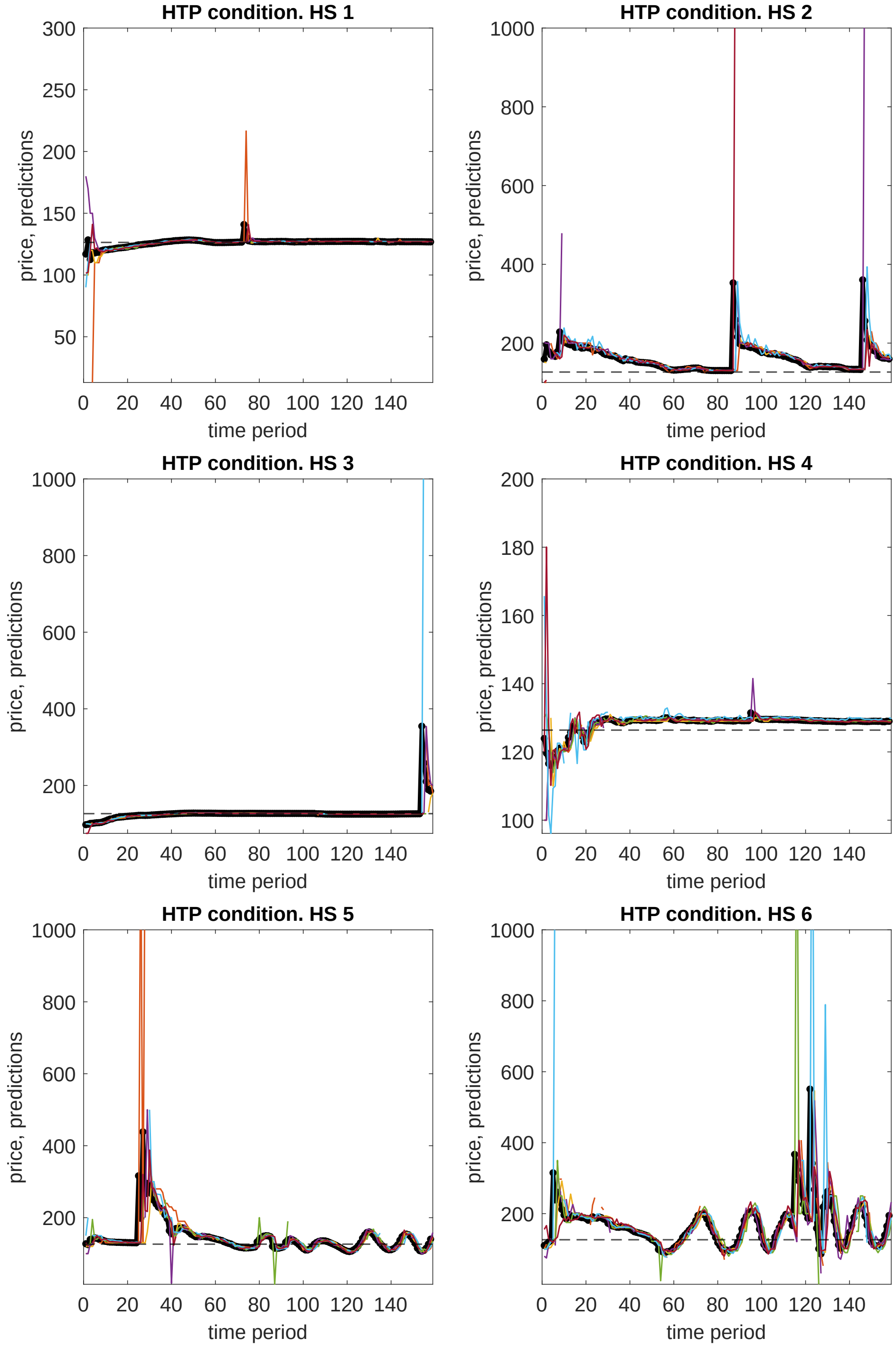


Figure E6: Prices (black) and predictions (colored) in markets 1-6 of treatment **HS**.

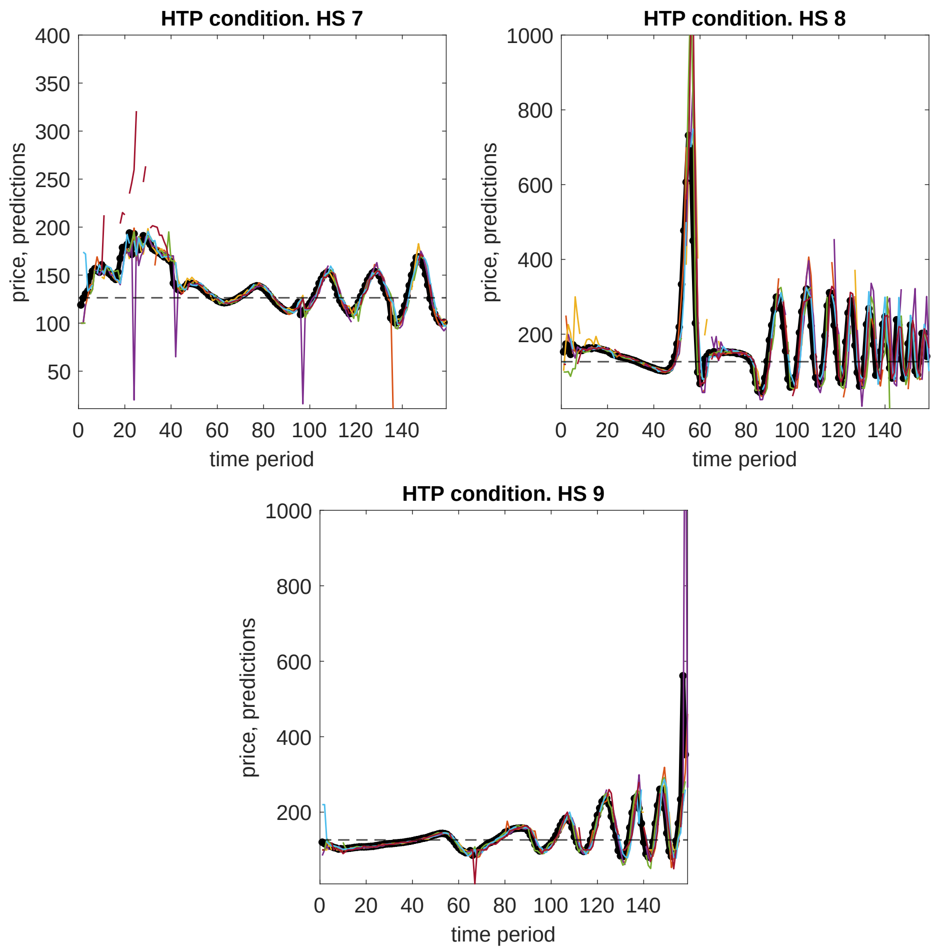


Figure E7: Prices (black) and predictions (colored) in markets 7-9 of treatment **HS**.

F Second Phase Data – Internet Appendix

In this appendix, we briefly discuss the data from the second phase of our experiment. To achieve equitable average payoffs across different sessions and run them for compatible durations, each session consisted of two phases. There was a change in the time pressure condition between these phases, see Table 1 of the paper for more details. In the paper, we refer to the treatments based on the time pressure conditions in the first phase. Consequently, the second phase data reflect the opposite time pressure conditions compared to the treatment names.[32]

In the paper, we analyze data from the first phase only. There are several reasons why data from the second phase are not suitable for conducting the analysis in our experiment. First and foremost, the time pressure experienced in the first phase may have a much stronger impact on the behavior of subjects in the second phase than any pre-experimental experiences they may have had. This makes it cleaner to compare the first-phase data between treatments, where participants are randomly assigned to the treatments. Moreover, our experiment involves groups, and during each phase, subjects interact within the same group. To reduce the effect of the group on their individual behavior, we rematched the subjects between phases. However, the group size (6) and the size of the lab (approximately 25 workstations) meant that there was some overlap in group compositions. Due to different turnout rates, this overlap varied in different cases and was beyond our control. In most cases, we had 2 or 3 groups of participants per session, but in one case, we had only 6 participants, leading to the same group after rematching. Since the experiment had some heterogeneity in group dynamics, especially in treatment **L**, different experiences in the first block, coupled with non-random rematching, may have had an uncontrolled impact on the

[32]Specifically, in treatment **L**, we formed 12 new groups (by re-matching the participants after the first phase) and placed them under the High Time Pressure (HTP) condition. In treatment **H**, we formed 10 new groups and placed them under the Low Time Pressure (LTP) condition. Finally, in treatment **HS**, we formed 9 new groups and placed them under the LTP condition, requesting them to press the 'Submit' button to register their forecasts.

dynamics after the rematching. For all these reasons, we did not intend to study the second phase data of this experiment.[33]

For the sake of completeness, this appendix presents the data from the second phase. We also report the results of the same statistical tests as in the main paper. To avoid possible confusion, and motivated by the discussion above, we will treat all the data of the second phases discussed in this Appendix as separate "treatments". The treatments will be referred to as $\mathbf{H^2}$, $\mathbf{L^2}$, and $\mathbf{LS^2}$ to reflect the actual time pressure conditions in this experimental phase and indicate that we are referring to the second phase of the data (via the superscript).[34]

The three panels in Fig. F1 depict the market prices (thin gray lines) for each market in these three treatments. This figure can be compared with Fig. 2 in the main paper. To facilitate such a comparison, we adjusted the price range based on the fact that the fundamental price is smaller in the second phase. (The upper level of the vertical range is about four times larger than the fundamental price in both Fig. F1 and Fig. 2). Visually, the dynamics under the HTP condition (lower panel) display lower mispricing and lower volatility than in the other two panels.

Table D2 provides descriptive statistics for all markets in the treatments $\mathbf{L^2}$, $\mathbf{LS^2}$, and $\mathbf{H^2}$. Similar to the first phase in the main paper, we report statistics over the common periods, which are periods 1 to 145 of this second phase. Additionally, we approximate the 'beginning' of the phase by considering periods 11 to 50, and the 'end' of the phase with periods 106 to 145.

Finally, Table D3 reports the p-values of various tests as described in the last column. In particular, its upper part compares statistics for the same time pressure conditions of the second-phase data with those in the first-phase data. We observe

[33] Although interesting, the question about the effect of experience is beyond the scope of this paper. It is addressed, for the learning to forecast experiment, in Kopányi-Peuker and Weber (2021) and Hennequin (2018).

[34] Thus, participants who were in treatment **L** moved to treatment $\mathbf{H^2}$, participants who were in treatment **H** moved to treatment $\mathbf{L^2}$, and, finally, participants who were in treatment **HS** moved to treatment $\mathbf{LS^2}$.

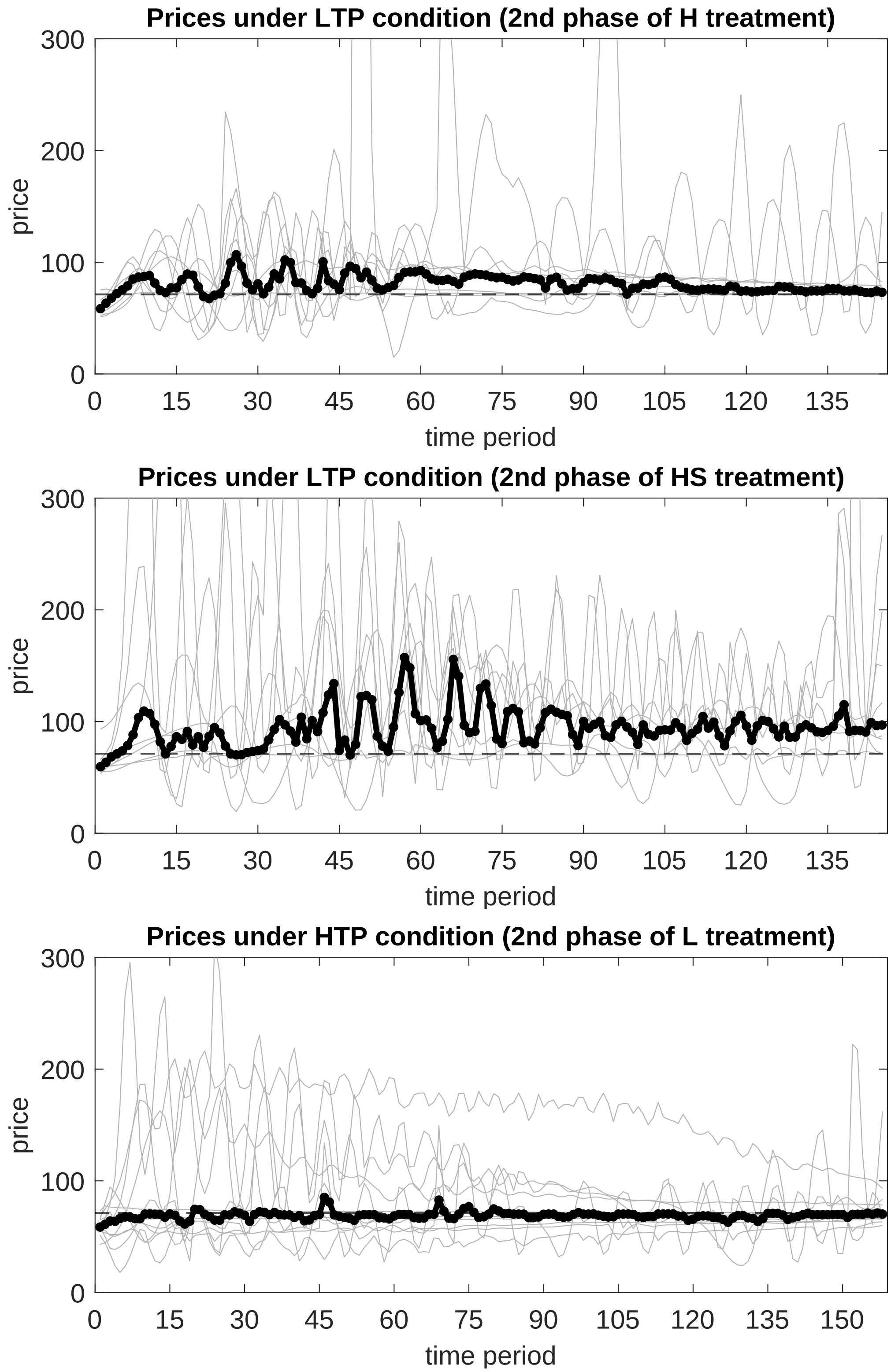


Figure F1: Median prices (thick, black) and prices in individual markets (gray) in the experimental treatments. The fundamental price $p^f = 71.2$ is indicated by the dashed horizontal line.

Treatment	Market	Interquartile range			Median of RAD		
		11-50	1-145	106-145	11-50	1-145	106-145
$\mathbf{L}^2$	$\mathbf{L}^2 1$	47.29	0.79	0.00	0.31	0.01	0.00
	$\mathbf{L}^2 2$	34.04	41.75	0.67	0.32	0.28	0.05
	$\mathbf{L}^2 3$	13.43	4.90	2.20	0.10	0.04	0.02
	$\mathbf{L}^2 4$	0.13	0.50	0.19	0.02	0.01	0.01
	$\mathbf{L}^2 5$	56.67	10.57	5.11	0.39	0.22	0.17
	$\mathbf{L}^2 6$	45.29	17.53	3.42	0.58	0.28	0.15
	$\mathbf{L}^2 7$	30.55	9.20	6.10	0.16	0.17	0.14
	$\mathbf{L}^2 8$	61.60	20.30	1.37	0.43	0.20	0.04
	$\mathbf{L}^2 9$	20.19	35.86	63.23	0.19	0.24	0.33
	$\mathbf{L}^2 10$	73.27	70.04	131.93	0.45	0.33	0.65
	Average	38.25	21.14	21.42	0.30	0.18	0.16
	Median	39.66	14.05	2.81	0.31	0.21	0.10
$\mathbf{LS}^2$	$\mathbf{LS}^2 1$	26.02	63.07	34.73	0.20	0.44	0.52
	$\mathbf{LS}^2 2$	1.09	0.16	0.18	0.02	0.02	0.01
	$\mathbf{LS}^2 3$	262.73	99.00	64.56	1.70	1.06	0.98
	$\mathbf{LS}^2 4$	40.22	39.67	15.66	0.38	0.51	0.48
	$\mathbf{LS}^2 5$	4.34	12.36	63.82	0.04	0.08	0.46
	$\mathbf{LS}^2 6$	83.79	30.86	12.47	0.49	0.43	0.31
	$\mathbf{LS}^2 7$	69.07	27.76	5.55	0.54	0.15	0.04
	$\mathbf{LS}^2 8$	116.98	38.54	8.49	0.78	0.43	0.32
	$\mathbf{LS}^2 9$	33.07	46.91	56.69	0.48	0.46	0.29
	Average	70.81	39.81	29.13	0.51	0.40	0.38
	Median	40.22	38.54	15.66	0.48	0.43	0.32
$\mathbf{H}^2$	$\mathbf{H}^2 1$	31.18	34.57	34.54	0.27	0.28	0.24
	$\mathbf{H}^2 2$	0.64	2.06	1.26	0.07	0.05	0.03
	$\mathbf{H}^2 3$	86.08	51.04	40.69	0.98	0.46	0.20
	$\mathbf{H}^2 4$	2.62	7.27	0.33	0.23	0.16	0.11
	$\mathbf{H}^2 5$	11.64	6.27	2.02	0.09	0.09	0.08
	$\mathbf{H}^2 6$	75.34	33.63	20.05	0.74	0.35	0.12
	$\mathbf{H}^2 7$	15.57	29.39	33.03	1.61	1.38	0.94
	$\mathbf{H}^2 8$	41.35	32.39	0.84	0.90	0.24	0.14
	$\mathbf{H}^2 9$	1.46	1.01	0.28	0.03	0.01	0.01
	$\mathbf{H}^2 10$	14.94	11.26	2.80	0.38	0.32	0.23
	$\mathbf{H}^2 11$	4.15	6.08	1.89	0.25	0.20	0.13
	$\mathbf{H}^2 12$	3.00	3.19	1.09	0.11	0.07	0.04
	Average	24.00	18.18	11.57	0.47	0.30	0.19
	Median	13.29	9.26	1.95	0.26	0.22	0.13

Table D2: Descriptive statistics for the data in the second phases.

that despite rematching, experience does have some effect, especially for the LTP condition. Note that the subjects in the $\mathbf{L^2}$ and $\mathbf{LS^2}$ treatments were initially in the HTP condition and thus experienced relatively stable dynamics. The data in the LTP condition in the second phase ($\mathbf{L^2}$ and $\mathbf{LS^2}$) are indeed more stable and less volatile than the data from the sessions where subjects experience the LTP initially in the experiment (**L**). On the other hand, the past experience of the LTP condition that subjects in the $\mathbf{H^2}$ treatment brought to the second phase does not seem to have an effect on dynamics when compared to subjects who had the HTP condition in the first phase. This is somewhat consistent with our finding that the HTP condition helps subjects avoid overusing past data and consequently achieve more stable dynamics.

The remaining part of Table D3 is analogous to Table 2 in the main text. It shows that the second-phase data exhibit a similar directional effect of time pressure as the first-phase data, even though this effect is less significant.

Data to compare		Statistics		Test
Data I	**Data II**	**IQR**	**Median RAD**	
Experience				
$\mathbf{L^2}$ periods 11–50	**L** periods 11–50	0.0001***	0.0005***	two-sided MWW
$\mathbf{L^2}$ periods 106–145	**L** periods 106–145	0.0192**	0.0321**	
$\mathbf{L^2}$ periods 1–145	**L** periods 1–145	0.0011***	0.0027***	
$\mathbf{LS^2}$ periods 11–50	**L** periods 11–50	0.0032***	0.0507*	
$\mathbf{LS^2}$ periods 106–145	**L** periods 106–145	0.7491	0.9151	
$\mathbf{LS^2}$ periods 1–145	**L** periods 1–145	0.0252**	0.2707	
$\mathbf{H^2}$ periods 11–50	**H** periods 11–50	0.1765	0.6209	
$\mathbf{H^2}$ periods 106–145	**H** periods 106–145	0.1213	0.9737	
$\mathbf{H^2}$ periods 1–145	**H** periods 1–145	0.0806*	0.5752	
$\mathbf{H^2}$ periods 11–50	**HS** periods 11–50	0.3028	0.1886	
$\mathbf{H^2}$ periods 106–145	**HS** periods 106–145	0.4555	0.6959	
$\mathbf{H^2}$ periods 1–145	**HS** periods 1–145	0.5458	0.1886	
Short-run vs Long-run				
$\mathbf{L^2}$ periods 11–50	$\mathbf{L^2}$ periods 106–145	0.1377	0.0322**	one-sided Wilcoxon signed-rank
$\mathbf{LS^2}$ periods 11–50	$\mathbf{LS^2}$ periods 106–145	0.0820*	0.1504	
$\mathbf{H^2}$ periods 11–50	$\mathbf{H^2}$ periods 106–145	0.0386**	0.0002***	
Time Pressure				
$\mathbf{L^2}$ periods 11–50	$\mathbf{H^2}$ periods 11–50	0.0882*	0.5913	one-sided MWW
$\mathbf{LS^2}$ periods 11–50	$\mathbf{H^2}$ periods 11–50	0.0549*	0.4295	
$\mathbf{L^2}$ periods 106–145	$\mathbf{H^2}$ periods 106–145	0.4606	0.6895	
$\mathbf{LS^2}$periods 106–145	$\mathbf{H^2}$ periods 106–145	0.0408**	0.0298**	
$\mathbf{L^2}$ periods 1–145	$\mathbf{H^2}$ periods 1–145	0.4345	0.7556	
$\mathbf{LS^2}$periods 1–145	$\mathbf{H^2}$ periods 1–145	0.0408**	0.1353	
Submission Button				
$\mathbf{L^2}$ periods 11–50	$\mathbf{LS^2}$ periods 11–50	0.7517	0.8612	two-sided MWW
$\mathbf{L^2}$ periods 106–145	$\mathbf{LS^2}$ periods 106–145	0.9606	0.9606	
$\mathbf{L^2}$ periods 1–145	$\mathbf{LS^2}$ periods 1–145	0.9333	0.9733	

Table D3: The p-values of the corresponding test (see the last column) for various comparisons. *, **, and *** indicate p values that are below 10%, 5% and 1%, respectively.