Are Long-Horizon Expectations (De-)Stabilizing? Theory and Experiments*

George W. Evans[a,b,*], Cars Hommes[c], Bruce McGough[a], Isabelle Salle[c]

[a] *Department of Economics, University of Oregon*
[b] *School of Economics and Finance, University of St Andrews*
[c] *Amsterdam School of Economics, University of Amsterdam, Tinbergen Institute & Bank of Canada, Ottawa.*

Abstract

The impact of finite forecasting horizons on price dynamics is examined in a standard infinite-horizon asset-pricing model. Our theoretical results link forecasting horizon inversely to *expectational feedback*, and predict a positive relationship between expectational feedback and various measures of asset-price volatility. We design a laboratory experiment to test these predictions. Consistent with our theory, short-horizon markets are prone to substantial and prolonged deviations from rational expectations, whereas markets with even a modest share of long-horizon forecasters exhibit convergence. Longer-horizon forecasts display more heterogeneity but also prevent coordination on incorrect anchors – a pattern that leads to mispricing in short-horizon markets.
JEL classification codes: C92; D84; E70.

*Acknowledgments:** We thank the participants of the CREED seminar at the University of Amsterdam on June 6, 2016, the internal seminar at Utrecht University on February 9, 2017, the "Expectations in Dynamic Macroeconomic Models" workshop on August 28-30, 2017, at the Federal Reserve Bank of St Louis, the internal seminars at UNSW on March 6 and UTS on March 9, 2018, the workshop on Theoretical and Experimental Macroeconomics at the Bank of Canada on June 25-26, 2019 and the ESA meeting on July 7-11 at SFU, Vancouver for helpful discussions. This research has been partly financed by the EU FP7 project *MACFINROBODS*, grant agreement No. 612796. Isabelle Salle is grateful to the International Network on Expectational Coordination, funded by the Institute for New Economic Thinking, for financing her stay at the University of Oregon during spring 2016. None of the above are responsible for potential errors in this paper. The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of Canada. In compliance with the CREED lab rules, an institutional review board (IRB) approval is only required if the participants are from vulnerable groups, if the experiment may have lasting effects on the well-being of the participants or if in the CREED lunch seminar members of the audience have serious ethical objections. As none of those cases applied to our experiment, no IRB approval has been obtained. The experimental data and codes used in this paper are available at Evans et al. (2022).

*Corresponding author
Email address: gevans@uoregon.edu (George W. Evans)

Highlights:

- An asset-pricing model with heterogeneous finite-horizon planning is developed.
- Longer horizons are shown to reduce price volatility and mispricing.
- A lab experiment confirms the predictions from the model.
- Disagreement in forecasts at longer horizon prevents coordination on wrong anchors.

Keywords: Learning, Long-horizon expectations, Asset pricing, Experiments.

1. Introduction

Most macroeconomic and finance models involve long-lived agents making dynamic decisions in the presence of uncertainty. The benchmark modeling paradigm is the rational expectations (RE) hypothesis, which, in a stationary environment, can be captured by a one-step-ahead formulation of the model dynamics together with boundary conditions;[1] the impact of future plans at all horizons are fully summarized by one-step-ahead forecasts. Thus, under RE the issue of the decision horizon is hidden. When agents are more plausibly modeled as boundedly rational (BR), a stand must be taken on the decision and forecasting horizon employed. In this paper, using a simple asset-pricing model, we study the importance of the forecasting horizon length, both theoretically and in a lab experiment.

Forecast horizons are clearly relevant to many macroeconomic and financial issues, including, for example, forward guidance in monetary policy, the impact of fiscal policy, or trading strategies in asset markets. Under BR the forecast horizon of households and firms affects their economic and financial decisions and their reaction to policies.

Financial markets provide motivation for the specific focus of both our theoretical model and our experiment. If agents have long horizons, does this lead to greater or smaller price volatility than if agents use shorter horizons? The answer is not obvious. There is a long-standing view that short-horizon agents are likely to induce greater instability because of their tendency to chase short-term gains. This argument was forcefully stated by Keynes (1936, Chap. 12, Sect. V-VI) in well-known passages in which he discusses price fluctuations and instability resulting from a market emphasis on short-term speculation.[2] On the other hand, in a standard RBC model that is known to be very stable under short-horizon adaptive learning, Evans et al. (2019) find that long-horizon decision-making instead leads to greater instability.

Therefore, a question of considerable importance is how the behavior of asset prices depends on the decision horizon of agents and on how they form expectations over this horizon. In reality, agents' behavior needs not be invariant to the forecasting horizon or the

[1] These boundary conditions include initial conditions on the state, as well as no-Ponzi scheme and transversality conditions. Typically, a non-explosiveness condition ensures these latter two.

[2] Early findings in survey data report how short-horizon investors tend to use extrapolative investment strategies, whereas longer-horizon investors tend to use mean-reverting trading rules (Frankel and Froot, 1987). In the heterogeneous-agent literature, interactions between fundamental traders and chartists are key to generating short-run deviations from fundamentals but mean reversion in the long run; see, e.g. LeBaron (2006). In lab experiments, it has also been shown that short-run forecasters tend to coordinate on trend-following rules, which amplifies bubbles; see, e.g., Hommes (2021).

nature of the forecasting task; and agents need not operate on the same planning horizon. This variety of behaviors may have non-trivial implications for expectations and prices. Ultimately, whether these implications materialize is an empirical question.

The primary goal of this paper is to design an asset pricing model populated by boundedly rational agents with finite forecasting horizons that can be analyzed for different configurations of horizons, and implemented in the lab. By tuning the horizon of the expectations, our lab experiment allows us to test how forecasting horizons affect price dynamics. What is novel in our experiment, among other important features, is that we study the role of the forecasting horizon and use the experimental data to test different theories of learning and how these fit with short-horizon and long horizon forecasting.

Our contribution stands at the crossroad of two literatures: the learning literature, as implemented, e.g. in dynamic general equilibrium models (Evans and Honkapohja, 2001), and the experimental literature concerned with behavioral finance; see, e.g., Palan (2013); Noussair and Tucker (2013). While our focus lies in the former, we borrow from the latter the laboratory implementation that allows us to design a group experiment whose main features remain as close as possible to the theoretical learning setup (see Section 3).

We choose the framework of a consumption-based asset pricing model *à la* Lucas (1978). We replace the standard rational expectations and representative agent assumptions with heterogeneous expectations and BR decision-making based on an approach developed in Branch et al. (2012).[3] Heterogeneous expectations about future prices constitute a motive for trade between otherwise identical agents.

We show that our implementation of bounded rationality in the Lucas setting leads to a particularly simple connection between individual decisions and expectations about future asset prices: an individual agent's conditional asset demand schedule reduces to a linear function of their endowment, the market clearing price and the agent's expectation of the *average* asset price over the given horizon. This latter feature facilitates elicitation of forecasts from the human subjects in the lab. In this setting, expectations about future asset prices constitute a central element of the price determination and impart positive feedback into the price dynamics: higher price forecasts translate into higher prices.

[3] Under BR, the decision horizon in general equilibrium settings has been considered by a variety of authors. The widely used one-step-ahead "Euler equation" learning is extensively discussed in Evans and Honkapohja (2001). An infinite-horizon approach developed by Preston (2005) has been utilized in several settings, e.g. Eusepi and Preston (2011). The intermediate finite decision-horizon approach used in this paper also relates to Woodford (2018); Woodford and Xie (2019).

We find, in our theoretical setting, that expectational feedback depends negatively on forecast horizon length. This in turn implies that under a standard adaptive learning rule, the rate at which market price converges to the fundamental price is increasing in the planning horizon. These results, together with other findings from the adaptive learning literature (discussed in detail in Section 2.2) lead to several hypotheses which we then test experimentally. For example, our results suggest that longer forecast horizons lead to reduced price volatility and result in prices that are closer to their fundamental value.[4]

We design an experiment that belongs to the class of "learning-to-forecast" experiments (LtFEs),[5] which focuses on the study of expectation-driven dynamics. In these experiments, participants' beliefs are elicited and the implied boundedly optimal economic decisions, conditional on beliefs, are computerized. This specification is in line with how economic theory models market clearing, and it isolates the effects of interactions between planning horizons and expectation formation by eliminating other price determinants which arguably influence the real-world prices, e.g. interactions between price dynamics and speculation or price dynamics and liquidity.

As we will see, the model's strong expectational feedback permits expectation-driven fluctuations and (nearly) self-fulfilling price dynamics. Expectational feedback is paramount in modern macroeconomic models, and the strength of the feedback can be policy dependent.[6] Our findings suggest that the degree of expectational feedback in macro models, and the potential for self-fulfilling dynamics, will also depend on the agents' forecast horizons.[7]

The asset-pricing model underlying our lab experiment is easily summarized: there is a fixed quantity of a single durable asset, yielding a constant, perishable dividend that comprises the model's single consumption good. The initial allocation of assets is uniform across agents (referred to, in the experiment, as participants). Each period, each agent forms forecasts of future asset prices and, based on these forecasts and their current asset holdings, their asset demand schedules are determined. These schedules are coordinated by a competitive market-clearing mechanism, yielding equilibrium price and trades. If expec-

[4]The formal statement of the corresponding hypothesis is given in Section 3.4.

[5]See the earlier contribution of Marimon et al. (1993). More recent experimental studies within macro-finance models include Adam (2007); Assenza et al. (2021); Kryvtsov and Petersen (2021). This literature is surveyed in Duffy (2016) and Arifovic and Duffy (2018).

[6]This is evident in textbook new-Keynesian models, but also generically featured in DSGE models.

[7]Data collected in LtFEs are informative about broad classes of markets and behaviors: see, e.g., Kopányi-Peuker and Weber (2021) who compare price dynamics in LtFEs with experimental call markets, and Cornand and Hubert (2020) who compare forecasts in LtFEs and real-world forecasts from surveys.

tations of all agents were fully rational, they would make optimal decisions. Participants' payoffs reflect forecast accuracy and utility maximization. A random termination method emulates an infinite-horizon setting and yields a constant effective discount rate induced by the probability of termination. This economy has a unique perfect-foresight equilibrium price – the "fundamental price" – determined by the dividend and the discount factor.

We consider four experimental treatments, based on horizon length, T: short horizon ($T = 1$), long horizon ($T = 10$), and two treatments with mixtures of short and long horizons. We are interested in several questions: Does the horizon of expectations matter for the aggregate behavior of the market? If so, how do the horizon and heterogeneity of horizons affect this behavior? In particular, are long-horizon expectations (de)stabilizing?

In line with our theoretical results, we find that markets populated only by short-horizon forecasters are prone to significant and often prolonged deviations from the fundamental price. By contrast, if all traders are long-horizon forecasters, the price path is consistent with convergence to the fundamental price. Note that our specification does not predetermine the results. Our experimental findings need not have agreed with our theoretical predictions. In particular, if subjects had held fully rational expectations, the results across the four treatments would have been identical. Instead, the price behaviors across treatments differ greatly, which is reflected in distinct forecasting behaviors across horizons, including the treatments involving mixed horizons.

A detailed analysis of individual forecasts reveals that the failure of convergence in short-horizon markets reflects the coordination of participants' forecasts on patterns derived from price histories, e.g. "trend-chasing" behavior. In contrast, coordination of subjects' forecasts appears more challenging in longer horizon treatments: long-horizon forecasters display more disagreement. The resulting heterogeneity of long-horizon expectations impedes coordination on trend-chasing behavior and favors instead adaptive learning, leading to convergence towards the fundamental price. Given these two polar cases, a natural question arises: what share of long-horizon forecasters would be large enough to stabilize the market price? Our findings suggest that even a modest share of them is enough.

A substantial literature has investigated financial markets in a laboratory setting. Existing LtFEs involve environments where only one- or occasionally two-step-ahead expectations (as in, e.g., Rholes and Petersen (2021)) matter for the resulting price dynamics. An exception is Anufriev et al. (2020), who allow for forecast horizons of up to three periods. Like us, they report more market volatility associated with shorter horizons. In contrast to them, we provide a micro-founded model of BR decision making with heterogeneous forecast horizons, which allows us to study expectation formation over different horizons in the *same market environment*. Our theoretical model is closely connected to our lab implementation, and is based on a standard macro asset-pricing model rather than a mean-variance framework.

Several experimental studies have been concerned with belief elicitation at longer horizons: see, e.g., Haruvy et al. (2007) and Colasante et al. (2020). However, in these studies, players' forecasts do not affect price dynamics. Hirota and Sunder (2007) and Hirota et al. (2015) studied the influence of trading horizons on prices in a setting that differs greatly from ours, and found that longer forecast horizons lead to convergence of prices to fundamentals. Noussair and Tucker (2006) show how futures markets can prime subjects toward thinking about more distant prices, which contributes to stabilizing current prices; see also De Jong et al. (2022). Duffy et al. (2019), among others, study prices in an experimental market with an indefinitely lived asset, for example due to bankruptcy. They find that "horizon uncertainty" does not significantly affect traded prices. Their framework also differs greatly from ours.

The paper is organized as follows. Section 2 gives the theoretical framework. Section 3 details the experimental design and our hypotheses based on predictions from the learning model. Section 4 provides the results of the experiment and Section 5 concludes.

2. Theoretical framework: an asset-pricing model

The underlying framework of our experiment is a consumption-based asset-pricing model *à la* Lucas (1978). This model can be interpreted as a pure exchange economy with a single type of productive asset; at time t, each unit of the asset costlessly produces y_t units of consumption. The textbook model refers to this asset as a "tree" that produces "fruit." In the experiment, we use the framing of a "chicken" producing "eggs." This terminology reduces the likelihood that participants with a background in economics or finance would recognize the textbook asset-pricing model, and it also facilitates the implementation of an infinite-horizon environment in the lab by suggesting an asset with a finite life.

2.1. The infinite-horizon model

There are many identical agents, each initially endowed with $q > 0$ chickens, where each chicken lays $y > 0$ non-storable eggs per period. In each period, there is a market for chickens. Each agent collects the eggs from her chickens, consumes some, and sells the balance for additional chickens. Alternatively, the agent can sell chickens to increase current egg consumption. This decision depends on both the current price of chickens, and forecasts of future chicken prices.

To formalize the model, we consider the representative agent's problem:

$$\max E \sum_{t\geq 0} \beta^t u(c_t), \text{ s.t. } c_t + p_t q_t = (p_t + y) q_{t-1}, \text{ with } q_{-1} = q \text{ given,} \tag{1}$$

where $u' > 0$ and $u'' < 0$, q_{t-1} is the quantity of chickens held at the beginning of period t, c_t is the quantity of eggs consumed, and p_t is the goods-price of a chicken. Finally, E denotes the subjective expectation of the agent.

Under RE, which, in our non-stochastic setting reduces to perfect foresight (PF), the Euler equation is $u'(c_t) = p_t^{-1}(p_{t+1} + y) u'(c_{t+1})$. There is no trade in equilibrium, i.e. $c_t = q_t y$. Thus the perfect foresight steady state is given by $c = qy$ and $p = (1-\beta)^{-1}\beta y$. We refer to $p = (1-\beta)^{-1}\beta y$ as the fundamental price (value) of the asset, and often refer to the PF equilibrium as the RE equilibrium, or REE. Note that in REE, the representative agent holds wealth constant and consumes her dividend each period; this same behavior obtains even if agents are endowed with different initial wealth levels.

2.2. The model with finite-horizon agents

We relax the assumption of perfect foresight over an infinite horizon and consider the behavior of a BR agent with a finite planning horizon $T \geq 1$. This relaxation introduces the need to specify a terminal condition for the agent's decision problem, in the form of an expected wealth target q^e_{t+T}, i.e. the number of the chickens the agent expects to hold at the end of the planning period. We *assume* $q^e_{t+T} = q_{t-1}$: the agent views his current wealth as a good estimate for his terminal wealth. This assumption is based on the following principle: if, at a given time t, current price and expected future prices coincide with the PF steady state, then the agent's decision rule should reproduce fully optimal behavior.[8] It follows that if the forecasts of all agents align with the PF steady state then REE obtains.

The BR agent's problem may now be presented as follows: in each period t, taking as given wealth q_{t-1}, prices p_t and price expectations p^e_{t+k} for $k = 1,\ldots,T$, the agent chooses current and future planned consumption and savings, c_{t+k} for $k = 0,\ldots,T$ and q_{t+k} for $k = 0,\ldots,T-1$, to maximize $\sum_{k=0}^{T} \beta^k u(c_{t+k})$ subject to the budget constraints

[8] See On-line Supplementary Materials: Appendix A.1 for discussion. This is a bounded optimality extension of the principle, introduced by Grandmont and Laroque (1986), which in particular requires that forecast rules reproduce steady states.

$c_t + p_t q_t = (p_t + y)q_{t-1}$, $c_{t+k} + p^e_{t+k} q_{t+k} = (p^e_{t+k} + y)q_{t+k-1}$ for $1 \leq k < T$, and $c_{t+T} + p^e_{t+T} q_{t-1} = (p^e_{t+T} + y)q_{t-1}$. In this last equation, the period $t+T$ expected terminal wealth q^e_{t+T} has been replaced with q_{t-1}, as per our assumption. On-line Supplementary Materials: Appendix A.2 derives the individual demand curves for assets, which depend negatively on prices and positively on price forecasts.

We now consider equilibrium price dynamics in the BR market. We allow for heterogeneous forecasts and planning horizons, and it is convenient to work with the linearized model, and to thin notation we reinterpret variables as deviations from the non-stochastic steady state. Formally, we distinguish agents by type $i \in \{1, \ldots, I\}$, where agents of type i have planning horizon T_i and price forecasts $p^e_{i,t+k}$. Let α_i be the proportion of agents of type i. Finally, let $\bar{p}^e_{it}(T_i) = T_i^{-1} \sum_{k=1}^{T_i} p^e_{i,t+k}$ be agent i's forecast of the average price over his planning horizon. The following result characterizes equilibrium price dynamics:

Proposition 2.1 *There exist type-specific expectation feedback parameters* $\xi_i > 0$ *such that* $\xi \equiv \sum_i \xi_i < 1$ *and* $p_t = \sum_i \xi_i \cdot \bar{p}^e_{it}(T_i)$.

All proofs are in the On-line Supplementary Materials: Appendix A. We note that the each of the feedback parameters ξ_i depends on the weights $\{\alpha_j\}_{j=1}^I$ as well as the corresponding planning horizons $\{T_j\}_{j=1}^I$. From this result, we see that the time t price only depends on the agents' forecasts of the *average* price of chickens over their planning horizon, i.e. $\{\bar{p}^e_{it}(T_i)\}_{i=1}^I$. The asset-pricing model with heterogeneous agents is therefore an *expectational feedback* system, in which the perfect foresight steady-state price is exactly self-fulfilling and is unique.

If expectations are homogeneous across planning horizons, i.e. $\bar{p}^e_{it}(T_i) = p^e_t$, $\forall i$, then the model's dynamics become $p_t = \xi p^e_t$, where, by Proposition 2.1, $\xi \in (0,1)$. More can be said about this expectational feedback parameter in the homogeneous case.

Proposition 2.2 *Let* $I \geq 1$, $\alpha_i \geq 0$, $\sum \alpha_i = 1$, $T_i \geq 1$, *and assume* $\bar{p}^e_{it}(T_i) = p^e_t$, $\forall i$. *Then:*

1. *If planning horizons are homogeneous then* $1 \leq T < T' \implies \xi > \xi'$.
2. *For the case of two planning horizons, if* $T_1 < T_2$ *then* $\frac{\partial}{\partial \alpha_1}\xi > 0$.

Proposition 2.2 says that the expectational feedback in this system is always positive but less than one. When there is a single planning horizon, increasing its length reduces the feedback. The strongest feedback occurs when $T = 1$, where $\xi = \beta$. Finally, for two agent types, increasing the proportion of agents using the shorter horizon increases the feedback.

Next we consider whether agents using simple learning rules would eventually coordinate their forecasts on the REE. Put differently, is the REE stable under adaptive learning? In Section 4.4, where we analyze subject-level forecasts from the experiment, we consider several types of forecast rules; here, for theoretical considerations, we focus on one prominent class of adaptive learning rules which has each of the N agents updating beliefs via

$$\bar{p}^e_{it}(T_i) = \bar{p}^e_{it-1}(T_i) + \gamma_t(p_{t-1} - \bar{p}^e_{it-1}(T_i)). \tag{2}$$

Here, $0 < \gamma_t \leq 1$ is called the "gain" sequence, which is assumed to satisfy $\sum_t \gamma_t = \infty$. There are two prominent cases in the literature: "decreasing gain" with $\gamma_t = t^{-1}$, which provides equal weight to all data; and "constant gain" with $\gamma_t = \gamma \leq 1$, which discounts past data.

Corollary 1 *Under decreasing and constant gain, $\bar{p}^e_{it}(T_i)$ and p_t converge to the REE price as $t \to \infty$. Furthermore, asymptotically, agents make fully optimal savings decisions.*

Corollary 1 shows that under adaptive learning of the form (2), the price dynamics converge to the fundamentals price. This result is independent of the number of agent-types and the lengths of their horizons, and can be extended to include heterogeneous gains.

The empirical macro literature employing adaptive learning is almost exclusively based on constant gain algorithms, and the analysis of our experimental results will be similarly focused. Under constant gain learning, the rate of convergence, i.e. $1-\zeta$ where $\zeta = p_t/p_{t-1}$, is time invariant: see On-line Supplementary Materials: Appendix A. In the homogeneous horizon case $1-\zeta = \gamma(1-\xi)$, which emphasizes that the rate of convergence is inversely related to the magnitude of ξ. The following result identifies the dependence of $1-\zeta$ on the planning horizon.

Corollary 2 *Under constant gain learning, the rate at which market price converges to its fundamental value is increasing in individual planning horizons T_i.*

Numerical investigations indicate that this result can be extended to allow for heterogeneous (constant) gains that are held fixed as planning horizons are varied.

Stochastic versions of model like $p_t = \xi p^e_t$ have been studied under constant gain learning. It is known that the extent and speed of convergence depend on the expectational feedback parameter ξ.[9] In short-horizon settings a number of authors have noted the possibility

[9]See, e.g. Evans and Honkapohja (2001, Ch. 3.2, 3.3 and 7.5).

that when the expectational feedback parameter is near one, near-random-walk behavior of asset prices is almost self-fulfilling, in that the associated forecast errors can be small, while also leading to significant departures from REE and excess volatility.[10] In our model this phenomenon arises most forcefully when $T = 1$ and β is near one so that ξ is near one.

Values of ξ near one also have implications for forecast accuracy. In particular, for some simple salient forecast rules, including those based on possibly-weighted sample averages (γ small) or near random walks (γ large), as well as higher-order trend-chasing models, expectations are nearly self-fulfilling. Thus in this case, even if the price level is far from the REE, the agents' forecast errors can be small. We will come back to this point later when interpreting our experimental results.

The results and discussion above point to the following implications for this model under learning, which we would expect to be reflected experimentally:

Implication 1: Prices and individual forecasts converge over time towards the REE.

Implication 2: The extent and speed of convergence toward the REE will be greater the smaller is the expectational feedback parameter ξ.

Implication 3: Deviations of forecasts from REE will be smaller for smaller ξ.

Implication 4: The level of price volatility will be lower the smaller is ξ.

These implications are reflected in the hypotheses we develop and test in the experiment.

3. The experimental design

The experiment is couched in terms of a metaphorical asset market in which assets are chickens (and thus finite-lived), and dividends are eggs (and thus perishable), comprising the experiment's unique consumption good. Participants are traders who make saving decisions based on forecasts of future chicken prices. In the experiment, participants submit price forecasts that are then coupled with the decision rules derived in Section 2 to determine their demand-for-saving schedules. Equilibrium prices and saving decisions are determined each period via market clearing.

3.1. Environment and procedures

Each group in the experiment is composed of $J = 10$ participants. At the opening of a market, each forecaster/trader is endowed with a given number of chickens. This

[10] See, e.g., Blanchard and Watson (1982), Branch and Evans (2011) and Adam et al. (2016)

number is the same across all forecasters/traders, but participants can only observe their own endowment and do not know the total number of chickens in the market.

Upon entering the lab, each participant is assigned the *single* task of forecasting the average market price of a chicken in terms of eggs over a given horizon, and this horizon remains the same throughout the experiment. Trading and the resulting egg consumption levels are computerized on behalf of the subjects. Each period, elicited forecasts are inserted into individual asset demand schedules, which are then aggregated, yielding the market clearing price. This price determines the market's trade volume, and is used to update individual asset holdings, egg consumption and utility level. Thus, conditional on forecasts, the outcomes in the lab are determined exactly as in our theoretical framework. Individual and aggregate asset demand schedules are given in the On-line Supplementary Materials: Appendix A, by (A.11) and (A.12), respectively, and the timing of events is given in Figure 1.

The dividend is common knowledge, and participants operate under no-short-selling and no-debt constraints. Each period, they must consume at least one egg. Eggs are both the consumption good and the medium of exchange, but only chickens are transferable between periods (see Crockett et al. 2019 for a similar setup).

Transposing this type of model to a laboratory environment requires resolving a number of issues, as discussed for instance in Asparouhova et al. (2016). Two major concerns are the emulation of stationarity and infinitely lived agents. Stationarity is an essential feature as it rules out rational motives to deviate from fundamentals, hence allowing us to get cleaner data on potential behavioral biases. An infinite-lifetime setting, together with exponential discounting and the dividend process, determines the fundamental value of the asset. This may play an important role in the belief formation process of the participants.

We use the standard random termination method originally proposed by Roth and Murnighan (1978) to deal with infinite lifetime in the laboratory. If each experimental market has a constant and common-knowledge probability of ending in each period, the probability of continuation is known to theoretically coincide with the discount factor. In the instructions of our experiment, the metaphor of the chickens allows us to tell the participants the story of an avian flu outbreak that may occur with a 5% probability in each period (corresponding to a discount factor $\beta = 0.95$). If this is the case, the market terminates: all chickens die and become worthless.

As for the stationarity issue, we choose a constant dividend process. The fundamental value associated with this dividend value and discount factor was not given to the participants. However, we think it likely that the experimental environment, including in particular the constant dividend process, is concrete enough to induce the idea of a fundamental value for a chicken in terms of eggs to the participants.

As discussed in Asparouhova et al. (2016), a major difficulty lies in the constant termination probability (discount factor). Participants should perceive the probability of a market to end to be the same at the beginning of the experimental session as towards the end of the time span for which they have been recruited. We therefore use the "repetition" design of Asparouhova et al. (2016): we recruited the participants for a given time and ran as many markets as possible within this time frame. Furthermore, we recruited them for 2 hours and 30 minutes but completed most of the sessions within 2 hours so as to keep the participants' perception of the session's end in the distant future throughout the experiment (see also Charness and Genicot (2009) for such an implementation). We did so by starting a new market only if not more than 1 hour and 50 minutes had elapsed since the participants entered the lab. If market was still running after this time constraint, the experimenter would announce that the current 20-period block (see below) was the last one.

[Figure 1 about here.]

Finally, our framework involves two additional difficulties. Most importantly, participants have to form forecasts over a given horizon, say over the next 10 periods, but the market may terminate before period 10. In this case, the average price corresponding to their elicited predictions is not realized, and participants' tasks cannot be evaluated (see below how the payoffs are determined). In order to circumvent this issue, we use the "block" design proposed by Fréchette and Yuksel (2017): each market is repeated in blocks of a given number of periods, and the termination or continuation of the market is observed only at the end of each block. This design allows the experiment to continue at least for the number of periods specified in the block, without altering the emulation of the stationary and infinite living environment from a theoretical viewpoint.

In our experiment, the length of a block is taken to be 20 periods, which corresponds to the expected lifetime of a chicken with a 5% probability of termination. The random draws in each period are "silent," and participants observe only every 20 periods whether the chickens have died during the previous 20 periods. If this occurred, the market terminates and they enter a new market from period 1 on. If this did not occur, the market continues for another 20-period block. In period 40, participants observe whether a termination draw has occurred between periods 20 and 40. If this is the case, the market terminates and a new one starts; if not, participants play another 20-period block till period 60, etc. Only periods during which the chickens have been alive count towards the earnings of the participants.

To prevent knowledge of the fundamental being carried over across markets we vary the dividend y, and thus the equilibrium price, between markets. We also vary the initial endowment of chickens to match the symmetric equilibrium distribution and keep liquidity and utility levels constant across markets: see Table 1.[11] On entering each new market, participants receive the corresponding values through a pop-up message, and those values remain on the screen throughout the market (see On-line Supplementary Materials: Appendix B, Figure 1). To avoid perfect predictions, we add a small noise term υ to the price, with $\upsilon \sim \mathcal{N}(0, 0.25)$.

[Table 1 about here.]

3.2. Payoffs

We elicit price forecasts from participants, but those forecasts translate into trade decisions, and the predictions of our theoretical model partly rely on the properties of the utility function and the incentive to smooth consumption over time. For this reason, the payoff of the participants consists of two parts: at the end of each market, all participants receive experimental points based *either* on forecast accuracy *or* on their resulting egg consumption with equal probability. This design avoids "hedging" and maintain equal incentives towards the two objectives (forecasting and consuming) throughout each market. Payoff tables are reported in the On-line Supplementary Materials: Appendix D.

The consumption payoff is $u(c) = 120 \cdot \ln(c)$ ($c \geq 1$). Specifying a concave utility function provides tight control on subjects' preferences and induces the consumption smoothing behavior that underlies the predictions from the theoretical model (see also Crockett et al. (2019)). Participants are paid only for periods during which chickens are alive. The payoff based on utility is simply the sum of their utility realized in each of those periods.[12]

To limit the cognitive load of the experiment and impart fairness between the payments for the forecasting and the utility maximization tasks, predictions are rewarded using a quadratic scoring rule, which ensures a decreasing and concave relationship between the

[11] We remark that only integer values of chickens and eggs are allowed to be traded/consumed. The large number of chickens renders this imposition inconsequential.

[12] These widely used cumulative payments align with discounted utility maximization with random termination under risk neutrality. Sherstyuk et al. (2013) find that the potential bias if agents are risk averse is of little empirical importance. Moreover, it would not impact our treatment differences.

payoffs and the forecast errors: max $\left(1100 - {}^{1100}\!/\!_{49}(\text{error})^2, 0\right)$. If the error is higher than 7, the payoff is zero. We must take account of the fact that there are necessarily periods before the death of the chickens for which forecast errors are not available. Consequently, the number of realized average prices over T periods, and the associated forecasting payments, is lower than the number of utility payments that take place in every period. To circumvent this discrepancy, the last rewarded forecast is paid $T+1$ times to the participants. This also incentivizes them to submit accurate forecasts for every period, as they are uncertain about which one will be the last and, hence, the most rewarded. If the chickens die in the first block before $T+1$ periods, participants were paid on utility. At the end of all the markets, the total number of points earned by each participant was converted into euros at a pre-announced exchange rate, and paid privately.

3.3. Instructions and information

Participants were given instructions that they could read privately at their own pace (see On-line Supplementary Materials: Appendix D). The instructions contain a general description of the markets for chickens, explanations about the forecasting task and how it translates into computerized trading decisions, information about the payoffs, and payoff tables, as well as an example. The instructions convey a qualitative statement of the expectations feedback mechanism that characterizes the underlying asset pricing model. This information set implies that subjects know the form of, and the sign restrictions on, the price law of motion, but do not know the exact coefficient value, which is consistent with the theoretical model. Qualitative knowledge of the fundamentals is also in line with the functioning of real-world markets, while keeping the cognitive load of the instructions reasonable.

At the end of the instructions, participants had to answer a quiz on paper. Two experimenters were in charge of checking the accuracy of their answers, discussing their potential mistakes and answering privately any question. The first market opens only after all participants had answered accurately all questions of the quiz. This procedure allows us to be confident that all participants start with a reasonable understanding of the experimental environment and their task. Of the participants, 90% (218) reported that the instructions were understandable, clear or very clear.

3.4. Hypotheses and experimental treatments

The testable implications discussed in Section 2.2 relate the feedback parameter ξ to the price dynamics. In the experiment, we adopt the setup considered in Item 2 of Prop. 2.2:

two types of agents, distinguished by forecast horizon. This setup implies that ξ depends on the horizon lengths and the share of each agent-type. We design four treatments, labeled `L`, `M50`, `M70` and `S`, and summarized in Table 2.

First, we consider homogeneous planning horizons. Item 1 of Proposition 2.2 establishes that the feedback ξ is inversely related to horizon length. In treatment Tr. `S` (for 'short'), all subjects forecast price over the planning horizon $T = 1$, and ξ reaches its upper bound $\beta < 1$. In Treatment `L` (for 'long') all subjects forecast average price over the next $T = 10$ periods, giving the lowest value of ξ that we explore. Ten is chosen as a compromise between the feasibility in the lab and reduction in ξ: see Figure 2b for the comparison of the expectational feedback across our different treatments.

Second, we allow for two planning horizons. Item 2 of Prop. 2.2 shows that the feedback parameter $\xi \in (0,1)$ increases with the share of short-horizon forecasters α. Figure 2 illustrates the effect of α on ξ for calibration of the model implemented in the laboratory. As is clear from Figure 2a, the impact on ξ is nonlinear, magnifying the stabilization power of even a small share of long-horizon agents. We add two intermediate treatments where the fraction $\alpha \in (0,1)$ of short-horizon planners takes the values 70% and 50% (Tr. `M70` and Tr. `M50` respectively, for 'mixed'), and the rest of the subjects are long-horizon forecasters. With this set up, the law of motion of the price, based on Eq. (12), is

$$p_t = p + \xi_s \left(\frac{\sum_s (p^e_{s,t} - p)}{\alpha J} \right) + \xi_l \left(\frac{\sum_l (p^e_{l,t} - p)}{(1-\alpha) J} \right), \tag{3}$$

where p is the fundamental price, and

$$\xi_s = \frac{\alpha h(1)}{\alpha g(1) + (1-\alpha) g(10)}, \; \xi_l = \frac{(1-\alpha) h(10)}{\alpha g(1) + (1-\alpha) g(10)},$$
$$g(T) = \left(1 - \beta^{T+1}\right)^{-1} \left(1 - \beta^T\right) \text{ and } h(T) = \left(1 - \beta^{T+1}\right)^{-1} (1-\beta) T \beta^T.$$

The sums are over the short (s) and long (l) horizon participants, respectively, and $p^e_{i,t}$ is the expectation of average price over agent i's forecast horizon (short=1 and long=10). Finally, ξ_s and ξ_l measure the expectational feedback induced by the short- and long-horizon forecasters, respectively. SeeTable 2 for specific values used in the experiment.

Proposition 2.2 and the implications established in Section 2.2, provide the first three main hypotheses to be tested through the experimental treatments. Corollary 1, suggests convergence in all treatments since the feedback parameter is always less than one. However, the implications at the end of Section 2.2 suggest that convergence to the REE can be tenuous if ξ is near one, as in Tr. S. These considerations suggest the following hypotheses:

Hypothesis 1a (Price convergence) *Under each treatment, participants' average forecasts and the price level converge towards the REE.*

Hypothesis 1b (Price deviation) *The higher the share of short-horizon forecasters, the more likely average forecasts and the price level will fail to converge towards the REE.*

Hypothesis 2 (Price volatility) *Increasing the share of short-horizon participants increases the level of price volatility.*

Our theoretical results suggest coordination of agents' expectations will increase over time as agents learn the REE. Since heterogeneous expectations provide a motive for trade in our experiment, we test the following in all treatments:

Hypothesis 3 (Eventual coordination) *Price predictions of participants become more homogeneous over time. As a consequence, trade decreases over time.*

[Table 2 about here.]

Besides providing an empirical test of the theoretical implications of the model, one further advantage of learning-to-forecast experiments is that they make it possible to collect "clean" data on individual expectations because the information, underlying fundamentals, and incentives are under the full control of the experimenter. Knowledge of fundamentals renders the measurement of mispricing patterns trivial; specification of the information received by the participants makes it possible to filter out which information really affected agents' expectations, which are the only degree of freedom in the experiment. We can then use this rich dataset to test additional hypotheses regarding participants' forecasting behavior. In the current context, it is of interest to compare the forecasts of short-horizon and long-horizon participants. A variety of factors suggest that long-horizon forecasting is more challenging than short-horizon forecasting. Long-horizon forecasting involves accounting for a sequence of endogenous outcomes, whereas short-horizon forecasting involves contemplation of only a single data point, and hence a lighter cognitive load.

This discussion suggests that there may be more variation of price forecasts for long-horizon forecasters than for short-horizon forecasters. To measure this heterogeneity we use *cross-sectional dispersion*, defined in terms of the relative standard deviation of subjects' forecasts within each period. We have the following two hypotheses:

Hypothesis 4 (Coordination and forecast horizons) *Long-horizon forecasters exhibit more heterogeneity of forecasts, than short-horizon forecasters.*

Hypothesis 5 (Trade volume and forecast horizons) *Higher shares of long-horizon forecasters result in greater heterogeneity of forecasts and, hence, higher trade volumes.*

[Figure 2 about here.]

3.5. Implementation

The experiment was programmed using the Java-based PET software.[13] Experimental sessions were run in the CREED lab at the University of Amsterdam between October 14 and December 16, 2016. Most subjects (124 out of 240) had participated in experiments on economic decision making in the past, but no person participated more than once in this experiment. Each of the four treatments involved six groups of ten participants, for a total of 240 subjects, who participated in a total of 63 markets, ranging from 20 to 60 periods. The average earnings per participant amount to €22.9 (ranging from €10.8 to €36.6).

4. The experimental results

In Section 4.1, we provide a graphical overview of the price data from the experimental markets. In Section 4.2 we examine our hypotheses using cross-treatment statistical comparisons. Section 4.3 conducts an empirical assessment of convergence to REE using price data. Finally, Section 4.4 connects the cross-treatment differences in terms of aggregate behavior to distinct forecasting behaviors across horizons by analyzing individual data.[14]

4.1. A first look at the data

Figure 3 displays an overview of the realized prices in the experimental markets for each of the four treatments. Each line represents a market, with the reported levels corresponding to the deviations from the market's fundamental value, expressed in percentage

[13] The PET software was developed by AITIA, Budapest under the FP7 EU project CRISIS, Grant Agreement No. 288501.

[14] We adopt a 5% confidence threshold to assess statistical significance. When carrying out econometric analysis, we use OLS estimates, autocorrelation in error terms is detected by Breusch-Godfrey tests, and heteroskedasticity using Breusch-Pagan tests. When needed, we use the consistent estimators described in Newey and West (1994). Significant differences between distributions are established using K-S tests and Wilcoxon rank sum tests to address non-normality issues.

points.[15] Plots with individual forecast data for each single market are given in On-line Supplementary Materials: Appendix B: see Figures 2 to 4. In those figures, blue corresponds to long-horizon forecasts, red to short-horizon forecasts, dots to rewarded forecasts and crosses to non-rewarded forecasts. Finally, the solid line is the realized price and the dashed horizontal line is the fundamental price.

A first visual inspection of the market price data in Figure 3 leads us to identify three different emerging patterns: (i) *convergence* to the fundamental price (see, for instance, in Figure 3d, Tr. `L`, Gp. 2 in purple or Gp. 6 in orange); (ii) *mispricing*, that we characterize by mild or dampening oscillations around a price value that is different from the fundamental value; either above the fundamental price, i.e. *overpricing*, or below the fundamental price, i.e. *underpricing* (see, for an example of each type of mispricing, the two markets played by Gp. 1 in Tr. `M70` on Figure 3b, red lines); and (iii) *bubbles and crashes*, described by large and amplifying oscillations (where the top of the "bubble" reached several times the fundamental value); see, e.g., the markets of the first group in Tr. `S` (Figure 3a, red lines).

This first glance at the data already leads us to question Hypothesis 1a, as it is clear that not every market exhibits price convergence towards the fundamental value. On the other hand, we see patterns in the data that are in line with Hypothesis 1b: while large deviations from fundamentals are observed in the short-horizon treatments (Tr. `S` and Tr. `M70`), they are absent from the long-horizon treatments (Tr. `M50` and Tr. `L`). Moreover, the problem of mispricing seems particularly acute in the short-horizon markets.

[Figure 3 about here.]

Interestingly, though, the observed bubbles break endogenously, which is *not* usual in LtFEs.[16] Several features of our setting may be behind this phenomenon: (i) the framing in terms of chickens and eggs, or (ii) incentives related to the payoff-relevant utility: in the end-of-experiment questionnaire some participants reported attempting to lower the price because they experienced low payoff along a bubble.[17]

[15]The apparent asymmetry around zero in the proportional deviations from fundamental values reflects that the price cannot be negative, while there is no upper bound except for the artificial one of 1000 that is unknown to the subjects until they hit it.

[16]The only exception is Market 2 of Group 2, in Tr. `S`, where one participant hits the upper-bound of 1000 and receives the message that his predictions have to be lower than this number. Note that this bound has been implemented for technical reasons, and none of the participants were aware of this bound, unless they reach it. This bound was reached 25 times out of the 18,170 forecasts elicited across all markets and subjects (which is about 0.1% of all forecasts).

[17]We also note that a high price provides incentives to sell – and therefore to submit a lower prediction

In the rest of this section, we explore the differences between treatments and confront these with our theoretical implications and experimental hypotheses. We now formulate five main results in the context of our five hypotheses.

4.2. Cross-treatment comparison

Table 3 reports cross-treatment comparisons of aggregate data. The first rows show significant cross-treatment differences regarding the price deviation (from fundamental), price volatility and, to a lesser extent, forecast dispersion: see Table 3 for definitions of these terms. These differences confirm the visual impression that the horizon of the forecasters matters for price dynamics and convergence towards the REE. The discrepancy between the realized price and the fundamental is strikingly lower in Tr. `L` than in Tr. `S`. Moreover, while the discrepancy from the REE is not statistically different between Tr. `L` and Tr. `M50`, prices are significantly closer to the fundamental price in those two treatments than in Tr. `M70`. These difference lead us to reject Hypothesis 1a in favor of Hypothesis 1b:

Finding 1 (Price convergence) *Increasing the share of long-horizon forecasters from 0% to 30% and also from 30% to 50% significantly reduces price deviation from the REE.*

Turning to Hypothesis 2, we find long-horizon forecasters have a stabilizing influence on prices. The price in Tr. `S` is significantly more volatile than in all other treatments, while price volatility is not significantly different between Tr. `M50` and Tr. `L`. Those observations yield the following finding, consistent with Hypothesis 2:

Finding 2 (Price volatility) *Increasing the share of long-horizon forecasters from zero percent to 30% and also from 30% to 50% significantly reduces price volatility.*

Our results suggest a *threshold effect* in the share of short-horizon forecasters on price convergence and volatility. A large share of short-horizon forecasters (more than half of the market) is necessary to hinder stabilization and convergence.

[Table 3 about here.]

Regarding Hypothesis 3, we consider the issue of coordination between participants. The trade volume significantly decreases in all treatments except Tr. `S`. Similar dynamics are observed for the within-participants forecast dispersion over time. In Tr. `S`, neither

---

than the average of the group – a strategy that was also reported a few times.

the forecast heterogeneity nor the trade volume shrinks over time.[18] Therefore, in partial support of Hypothesis 3, we obtain the following result:

Finding 3 (Eventual coordination) *In all treatments except in Tr.* `S`*, participants' forecasts become more homogeneous over time and, hence, the trade volume decreases over time.*

Our last two hypotheses relate to the differences across treatments of participants' degree of coordination. Table 3 gives some evidence that the presence of more short-horizon forecasters leads to more homogeneous forecasts: forecast dispersion is higher in Trs. `L` and `M70` than in Tr. `S`. In mixed treatments, coordination among agents with common forecast horizons can be assessed. For example, in Tr. `M50`, looking at the first market of Gp. 4, or at all markets in Gp. 5 and 6, it is clear that short-horizon forecasts are closer to each other than the long-horizon ones (see Figure 3 in the On-line Supplementary Materials: Appendix B). This is confirmed by statistical analysis: in this treatment, the average dispersion between short-horizon forecasters is 0.057, versus 0.163 among the long-horizon forecasters, and the difference is significant (p-value $< 2.2e-16$). Using also the trade-volume and forecast-dispersion rows in Table 3, and in line with Hypotheses 4 and 5, we find the following:

Finding 4 (Coordination and forecast horizons) *Long-horizon forecasters exhibit greater cross-sectional forecast dispersion than do short-horizon forecasters.*

Finding 5 (Trade volume and forecast horizons) *The higher the share of long-horizon forecasters in a market, the greater the cross-sectional dispersion of price forecasts and the higher the trade volume.*

These findings align with the survey-data analysis of Bundick and Hakkio (2015) and the experimental work of Haruvy et al. (2007) (done in non-self-referential environments).

There are two additional considerations of interest that are less directly connected to our hypotheses: first, possible learning effects resulting from repetition; second, the implications of performance metrics based on received utility versus forecast accuracy.

[18] A regression of the trade volume on the period leads to the coefficients -0.433, -0.348, -0.699 and 0.021 for, respectively, Tr. `L`, `M50`, `M70` and `S`, with the associated p-values $< 2e-13$ except for Tr. `S` with 0.493. Similarly, with the forecast dispersion as a dependent variable, the same estimated coefficients are -0.004, -0.004, -0.005 and 6.185e-05 with the associated p-values of 0.020, 5.4e-06, 0.002 and 0.935.

The repetition design of our experiment allows us to look at *learning effects* in sequential markets with the same group of subjects. Replications of the seminal Smith et al. (1988) bubble experiment find that large deviations from fundamentals disappear if the market is repeated several times with the same participants (Dufwenberg et al., 2005).

Results from our experiment convey the impression that price fluctuations do not decrease with participants' experience: see figures in the On-line Supplementary Materials: Appendix B. On the contrary, a bubble can take several markets to arise, and price deviations from fundamental tend to amplify with market repetitions. This is especially the case in Groups 1, 2 and 4 of Tr. S. Deviations from fundamental tend also to increase with market repetition in Gp. 5 of Tr. L.[19] Not only are learning effects absent, in fact our results suggest that volatility in the form of bubbles and crashes persists across markets.

Turning to the role of performance metrics, we return to Table 3 and consider the earnings of participants in different treatments. While not directly connected to our hypotheses, incentives are an essential ingredient of theory testing using laboratory experiments. The data from the last two rows of Table 3 reveal that there is no noticeable difference in participants' earnings across treatments, whether based on utility or forecasting.

4.3. Assessing convergence to the REE

Since Hypotheses 1a-1b are the primary focus of the experiment, this subsection and the next complement Finding 1. Here we formally test whether convergence to the fundamental value occurs in the experimental markets. We follow the method presented in Noussair et al. (1995), which consists in estimating the value to which the price would converge asymptotically if a market were extrapolated indefinitely.[20] As the lengths of our markets differ and most are short due to the stochastic termination rule, this approach appears well suited to our experiment.

We estimate the following equation for each of the four treatments separately:

$$\frac{p_{g,m,t} - p_{g,m}}{p_{g,m}} = \frac{1}{t}\sum_{g=1}^{6}\sum_{m\in\Omega_{M_g}} D_{g,m} b_{1,g,m} + \frac{t-1}{t}\sum_{g=1}^{6}\sum_{m\in\Omega_{M_g}} D_{g,m} b_{2,g,m}, \tag{4}$$

[19] Linear regressions of the absolute deviations of prices and forecasts from the REE on the order of the market confirms the absence of convergence along sequential markets. By design, repeated markets had different fundamental prices, which makes it difficult to carry over knowledge from one market to the next.

[20] Duffy (2016) identifies circumstances in which Noussair et al. (1995)'s method has shortcomings, and suggests an alternative regression to address them. In our case, these circumstances only arise in one out of the 63 markets (in the first market of Tr. S, Group 2).

with $p_{g,m,t}$ the realized market price in period t in Group $g \in \{1,...,6\}$ and market m; Ω_{M_g} the number of markets played by Group g; $D_{g,m}$ a dummy taking the value one if the price comes from Group g and market m and zero otherwise; and $p_{g,m}$ is the fundamental value of the price in Group g and market m.

The estimated coefficients of these regressions provide the fitted initial ($\hat{b}_{1,g,m}$) and asymptotic ($\hat{b}_{2,g,m}$) prices. If $\hat{b}_{2,g,m}$ is not significantly different from zero, we cannot reject the hypothesis of *strong convergence* towards the fundamental, i.e. $b_{2,g,m} = 0$. If $\left|\hat{b}_{1,g,m}\right| > \left|\hat{b}_{2,g,m}\right|$ holds significantly, the evidence supports *weak convergence* towards the fundamental. The results are collected in Figure 4. Details of the estimations are in On-line Supplementary Materials: Appendix C.

[Figure 4 about here.]

The distributions of the estimated coefficients in Figure 4 reveal a net decrease in the estimated distances of the price to fundamental in Tr. `M70`, `M50` and `L` (compare the paired box plots per treatment).[21] However, a decrease is not observed in Tr. `S`. The estimated final distances are particularly concentrated around zero in Tr. `L`, and even more strikingly in Tr. `M50`. Econometric analysis shows that weak convergence obtains in all but one market in Tr. `L`, and most markets in Tr. `M50`. By contrast, fewer than two-thirds of the markets in Tr. `M70` exhibit weak convergence, and fewer than one-half of the markets in Tr. `S`. Results on strong convergence show a similar pattern.

As a complement to Finding 1, we draw from this exercise the following insight:

Finding 6 (Statistical convergence) *Convergence to the REE is more frequently observed when the share of long-horizon forecasters is increased.*

This finding conforms with Hypothesis 1b and Figure 4 rejects Hypothesis 1a.

We now examine factors that contribute to the convergence failures observed in Tr. `M70` and Tr. `S`. Initial conditions in a given market may be correlated with terminal conditions in the previous market: see figures in On-line Supplementary Materials: Appendix B. Price patterns, such as systematic mispricing and oscillatory behaviors, sometimes appear to

[21] A box plot illustrates a distribution by reporting the four quartiles, with the thick line being the median, and the two whiskers being respectively Q1 and Q4 within the lower limit of $Q1 - 1.5(Q3 - Q1)$ and the upper limit of $Q3 + 1.5(Q3 - Q1)$. Outside that range, data points, if any, are outliers and represented by the dots. In the figure, each pair of box plots represents a treatment. The first box plot of each pair gives the distribution of the estimated initial values $\hat{b}_{1,g,m}$, the second one the estimated asymptotic values $\hat{b}_{2,g,m}$ in (4). The zero line represents convergence to fundamental.

carry over from one market to another even though the information from previous markets is not displayed to participants.

We compute the correlation between the estimated initial price values $\{\hat{b}_{1,g,m}\}$ and the price levels prevailing in the preceding market. This correlation is 0.6644 (p-value 0.0000) when the previous prevailing prices is measured as the average price over the last 10 periods of the previous market, and is 0.3444 (p-value: 0.0057) when measured as simply the last observed price in the preceding market.[22]

Equation (4) can also be used to assess the role of price histories in convergence failures, by conducting an analysis of the variance of the estimated asymptotic coefficients $\{\hat{b}_{2,g,m}\}$ in terms of three factors: the fundamental value; the price in period one; and the last price in the previous market.[23] Results, reported in Figure 5, reveal a striking pattern: asymptotic price values are almost entirely driven by fundamental values in Tr. `L` and `M50`, while initial price levels and price histories explain a considerable amount of the asymptotic price values in Tr. `M70`, and an even larger amount in Tr. `S`. This analysis confirms the dynamics reported in Figure 4, and sheds further light on Hypotheses 1a and 1b: coordination of subjects' forecasts on an incorrect anchor, namely past observed prices, is responsible for the lack of convergence observed in Tr. `M70` and Tr. `S` and, hence, the rejection of Hypothesis 1a.

Finding 7 (Fundamental and non-fundamental factors)

(i) When the share of long-horizon forecasters is large enough, the asymptotic market price is driven by fundamentals only.

(ii) If short-horizon forecasters dominate, the asymptotic market price is partly driven by non-fundamental factors, in particular past observed price levels.

[Figure 5 about here.]

To shed some light on the causal mechanisms behind those results, we now seek to understand how the participants formed their price forecasts and how those individual behaviors connect to the observed market prices in the experiment.

[22] For first markets, we took 50 as the previous value because it corresponds to the middle point of the empty price plot that the participants observe before entering their first forecast; see the screen shots, On-line Supplementary Materials: Appendix B, Figure 1. Removing first markets results in fewer data points, but the correlation pattern persists.

[23] The variance decomposition was done using the Fourier amplitude sensitivity test.

4.4. Participants' forecasts and aggregate outcomes

At the end of the experiment, participants were asked to describe in a few words their strategies. Analysis of the answers makes clear that the vast majority of participants, aside from strategic deviations for trading purposes, made use of past prices. The observation that expectations about future market prices depend on past trends has also found wide support in the experimental literature – see the early evidence reported in Smith et al. (1988) and Andreassen and Kraus (1990), and more recent evidence found in Haruvy et al. (2007); see also the empirical literature, starting from early contributions such as Shiller (1990).

To estimate the dependence of participants' forecasts on past data, we begin with the following class of simple, yet flexible, agent-specific forecasting models:

$$p^e_{j,t} = \beta_0 + \beta_1 p_{t-1} + \beta_2 p_{t-2} + \delta_1 p^e_{j,t-1}. \tag{5}$$

This class extends the constant gain implementation of equation (2) to include models conditioning on p_{t-2}. Clearly, participants could have paid attention to even more lags of the observable variables – a few reported to have done so – but most referred to at most the last two of prices in their strategy. Of course, including lagged expectations is an indirect way of accounting for the influence of additional lags of prices.[24]

We focus on the following three special cases of the forecasting model (5):

Naive expectations:	$\beta_0 = \beta_2 = \delta_1 = 0$ and $\beta_1 = 1$
Adaptive expectations:	$\beta_0 = \beta_2 = 0$, $\beta_1 \in (0,1)$, and $\beta_1 + \delta_1 = 1$
Trend-chasing expectations:	$\beta_0 = \delta_1 = 0$, $\beta_1 > 1$, and $\beta_1 + \beta_2 = 1$

Under naive expectations, $p^e_{j,t} = p_{t-1}$. Although we label this "naive," these are the optimal forecasts if the price process follows a random walk, and naive expectations are therefore "nearly rational" when prices follow a near-unit root process. We note that naive expectations corresponds to constant-gain adaptive learning with $\gamma = 1$: see Section 2.2. Under adaptive expectations, agents forecast as $p^e_{j,t} = p^e_{j,t-1} + \beta_1(p_{t-1} - p^e_{j,t-1})$. This rule, which corresponds to the constant-gain adaptive learning rule of Section 2.2 with $0 < \gamma < 1$, is known to be optimal if the price process is the sum of a random walk component and white noise, i.e. a mix of permanent and transitory shocks: see Muth (1961).

[24] In principle, this forecasting model could generate negative price forecasts, in which case it would be natural for agents to impose a non-negativity condition.

Under trend-chasing expectations, agents forecast as $p^e_{j,t} = p_{t-1} + \phi\,(p_{t-1} - p_{t-2})$ where $\phi = \beta_1 - 1 > 0$. This rule performs well in bubble-like environments in which price changes are persistent. In fact, this forecasting rule is optimal if the first difference in prices follows a stationary AR(1) process. Intuitively, agents are forecasting based on the assumption that the proportion ϕ of last period's price change will continue into the future. Finally, we note that trend-chasing expectations can lead to stable cyclical price dynamics.

We focus on the class of simple rules (5) for parsimony and because they nest salient special cases. However, adaptive learning is much more general, both in terms of included regressors and in allowing parameters to evolve over time as new data become available.

[Figure 6 about here.]

Figure 6 illustrates the potential for these simple forecasting rules to explain the price data in five different experimental markets: see graphs (a) to (e). The dashed horizontal line is the fundamental price and the dotted line is the realized price in the experimental market. Dots correspond to simulated price forecasts and the solid line gives the implied, simulated market prices. To construct the simulated price forecasts, a parametric specification of a particular forecasting model is chosen, and, for each agent, is initialized using their forecasts in the first two periods of the experiment. In each subsequent period, agents' forecasts are determined using the forecasting model, previously determined simulated prices, and a small, idiosyncratic white noise shock. Note that the simulated and experimental price time series are close to each other. Figure 6 also highlights the systematic differences between treatments and horizons in belief formation and links them to the observed price patterns.

Graph (a) provides an example of trend-chasing behavior that emerged from treatment `S`. The simulated data are based on setting $\phi = \beta_1 - 1 = 0.3$, strikingly illustrate the possibility of a bubble and crash being generated by trend-chasing forecast rules. Graph (b) gives an example of adaptive expectations associated to treatment `L`, with parameterization $\beta_1 = 0.7$ and $\delta_1 = 0.3$, showing apparent convergence to the fundamental price.

Graphs (c) and (d) correspond to treatment `M50`, in which short-horizon forecasters are naive and trend-following, respectively, and long-horizon forecasters form expectations adaptively. The simulated price paths depend on the individuals' initial forecasts in each market, a significant factor in the observed dynamics. Graph (c) exhibits persistent departures from fundamentals, while in graph (d) the short-horizon trend-chasers generate cyclic dynamics as well as apparent convergence. Finally, graph (e) corresponds to `M70` with short-horizon trend-chasing forecasters and long-horizon forecasters forming expectations adaptively. Here the cyclicality arising from the trend-followers is even more pronounced. The presence of only 30% long-horizon types appears insufficient to impart convergence.

Using step-by-step elimination, we examined individual participant-level forecast data, pooled across markets, and looked for simplifications of the model (5) in an attempt to determine if, and to what extent, participants used one of the three simple rules listed above, and whether there exist systematic differences in forecasting behaviors across horizons. We found, considering all 240 participant forecast series,[25] that more than half the short-horizon participants had forecasts consistent with trend-chasing rules, and more than a third of the long-horizon participants had forecasts consistent with adaptive expectations.[26]

The estimated coefficients $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ and $\hat{\delta}_1$ from (5) for each participant are illustrated in Figure 7: smaller, solid triangles identify long-horizon forecasters and larger triangles identify short-horizon forecasters.[27] Panel 7a shows a scatterplot of the components $\hat{\beta}_1$ and $\hat{\beta}_2$ for each participant. Under the restrictions $\hat{\beta}_0 = \hat{\delta}_1 = 0$ and $\hat{\beta}_1 > 1$, the trend-chasing model aligns with the constellation of points on the part of the downward-sloping dashed line that lies within the shaded region. Clearly, there are striking differences in the behaviors of participants tasked with short-horizon versus long-horizon forecasting.

A substantial number of the short-horizon points in Panel 7a lie on, or close to, the trend-chasing constellation. The trend-chasing restrictions cannot be rejected for 56% of the short-horizon forecasters. Panel 7b shows the corresponding scatterplot of the components $\hat{\beta}_1$ and $\hat{\delta}_1$. Under the assumptions that $\hat{\beta}_0 = \hat{\beta}_2 = 0$ and $0 < \hat{\beta}_1 < 1$, the adaptive-expectations model aligns with the constellation of points on the part of the downward-sloping dashed line that lies within the shaded region in panel 7b.[28] In contrast with the behavior exhibited by short-horizon forecasters, a substantial number of the long-horizon points in panel 7b lie on, or close to, the adaptive-expectations constellation. The adaptive-expectations restrictions cannot be rejected for more than one-third of the participants in long-horizon treatments. We summarize these findings as follows:

[25] The experiments included 18 treatment `S`, 14 treatment `L`, 18 treatment `M70`, and 13 treatment `M50` markets, with 10 participants in each market, giving 630 market-participant forecast series.

[26] For 212 of 240 participants, the step-by-step elimination process leads to a forecasting model in which at least one variable other than the intercept is significant. Also, the average R^2 is high for each treatment (ranging from an average of 0.884 in Tr. `L` to 0.962 in Tr. `M50`), which confirms the ability of the simple class of rules (5) to capture the main features of participants' behavior.

[27] A few of the participants' estimated coefficients lie outside the ranges chosen for Figure 7.

[28] Naive expectations corresponds to limiting cases (i.e. $\hat{\beta}_1 \to 1$) of both trend-chasing and adaptive-expectations forecasting models.

Finding 8 (Individual forecast behaviors) *Short-horizon and long-horizon forecasters display different forecasting behaviors: (i) More than one-half of the short-horizon forecasters form forecasts consistent with trend-chasing behavior. (ii) More than one-third of the long-horizon forecasters form forecasts consistent with adaptive expectations.*

These results align with Hypothesis 1b: distinct forecasting behaviors across horizons imply differences in price patterns. Trend-chasing behavior tends to preclude, and adaptive expectations tend to impart convergence to REE. Finding 8 also suggests greater forecast-model heterogeneity in long-horizon treatments, providing some support to Hypothesis 4.

[Figure 7 about here.]

The plots in Figure 7 include estimates that do not appear, even after accounting for statistical significance, to align with any of the special cases identified above. There are several possible explanations. First, it is possible that some subjects use less parsimonious forecasting rules than are captured by the class (5). Second, given that most subjects participated in multiple markets, it is quite possible that some of these participants used different rules in different markets. Our pooling estimation strategy does not account for this. Third, in general, under adaptive learning , in addition to the intercept, the other coefficients in the subjects' forecasting rules may evolve over time to reflect recent patterns of the data. Finally, we note that if ξ is near one then *any* collective forecast of the deviation of price from fundamentals is nearly self-fulfilling; this point is particularly germane for Tr. `S`.

Finding 8 sheds further light on the observed treatment differences. Admittedly, it is difficult, using the whole dataset, to distinguish between the effects on prices of changes in ξ and differences in how expectations are formed over different horizons. It is more revealing to look into Trs. `M50` and `M70` only, where all subjects, whether long- or short-horizon forecasters, operate in the same market environment – only the nature of their forecasting task differs. In these treatments, Finding 8 still holds: subjects systematically used distinct rules to forecast over short and long horizons; see Figures 7c-7d. It follows that prices display different patterns across Trs. `M50` and `M70` in part because the respective participants' forecasting tasks differ, and not only because the expectational feedback differs.

Visually, we cannot identify a different pattern between the top panels and the bottom panels. We verify this visual impression by using chi-squared tests for equality of proportions. The proportions of trend-chasers in the four treatments pooled together, in the two heterogeneous-horizon treatments pooled together, in Tr. M50 only, and in Tr. M70 only, are not significantly different from each other $\left(\chi^2(3) = 1.38,\ \text{p-value } = 0.71\right)$. Similarly, the proportions of adaptive learners in the four treatments pooled together, in the two heterogeneous-horizon treatments pooled together, in Tr. M50 only, and in Tr. M70 only, are not significantly different from each other $\left(\chi^2(3) = 1.85,\ \text{p-value } = 0.60\right)$.

In summary, longer forecast horizons induce lower expectation feedback and long-horizon treatments are empirically associated with adaptive expectations; both of these features induce price stability and more frequent convergence to the fundamental price. By contrast, shorter forecast horizons result in higher expectation feedback and short-horizon treatments are empirically associated with trend-chasing behavior; both of these features lead to persistent departures from the fundamental price.

5. Conclusions

We have investigated the impact of forecast horizons on price dynamics in a self-referential asset market. We developed a model with BR agents and heterogeneous planning horizons, and derived theoretical predictions for the effects of the planning horizon on the dynamic and asymptotic behavior of market price. We then tested our predictions by implementing our asset market in a lab experiment, eliciting price forecasts at different horizons from human subjects and trading accordingly.

The central finding of this paper is that key features of price dynamics are governed by the forecast horizons of agents. This was demonstrated analytically in a simple asset-pricing model, and then tested in a laboratory experiment. Our experimental design, which holds everything fixed except for the proportions of long-horizon and short-horizon subjects, finds dramatically different pricing patterns in the different treatments.

Prices in markets populated by only short-horizon forecasters fail to converge to the REE, with large and prolonged deviations from fundamentals. By contrast, in line with our theoretical predictions, we find that even a relatively modest share of long-horizon forecasters is sufficient to induce convergence toward the REE.

In our design, payoffs are determined in part by discounted consumption utility, as reflected in our forecast-based trading mechanism. This eliminates incentives to obtain capital gains arising from speculation about future crowd behavior, which is the focus of models like (De Long et al., 1990). Because dividends are known to be constant, we rule out the possibility that heterogeneous beliefs about future dividends cause price deviations from fundamentals. Nor do fluctuations arise from confusion about how the market works, as the vast majority of participants reported to understand their experimental task. We can exclude the role of liquidity in mispricing, as this is kept constant across all treatments.

Our finding that even a modest proportion of long-horizon subjects tends to guide the economy to the REE can be related both to the magnitude of the model's expectational feedback and to the systematically different forecasting behaviors identified for short and long horizons. Trend-chasing behavior is widely observed among short-horizon forecasters while adaptive expectations better describes long-run predictions. Hence, long-horizon forecasts induce stability around the REE, whereas coordination of forecasts on trend-following beliefs, and anchoring of individual expectations on non-fundamental factors, are largely responsible for mispricing in short-horizon markets. Instability of this type has been noted in the adaptive learning literature. Our experiment shows that this theoretical outcome constitutes an empirical concern as well.

Our study employs a framing that does not use the vocabulary of speculative asset markets; we emulate a stationary and infinite environment that induces discounting with a stochastic ending; and our payoff scheme incentivizes participants to smooth consumption. Despite these features, we obtain systematic mispricing when only short-horizon subjects are present, which implies an expectational feedback parameter close to one. We also identify systematic variations in the behaviors of short-horizon and long-horizon forecasters that are consistent with the distinct price patterns across horizons.

Long-horizon forecasting is more challenging than short-horizon forecasting: participants must average over a number of future periods; further, the observability of the forecast errors and the resulting feedback from the experimental environment is delayed to the end of the forecast horizon, when the average price is realized. Long-horizon forecasters also display more disagreements. Despite these obstacles, their presence stabilizes the market.

An interesting insight from our findings is that heterogeneity in behavior need not be detrimental to market stabilization. In our setup, when short-horizon agents are present, introducing long-horizon agents contributes to breaking the coordination of participants' beliefs on non-fundamental factors. We also find that the type of forecast rule used by a given subject depends on the exogenously imposed planning horizon. This suggests that BR agents are not characterized by invariant behavioral types.

Our study has implications for macro-finance models with heterogeneous, BR agents. Our findings that agents' forecast horizons play a central role in the determination of asset prices clearly suggest that the forecasting horizon of agents must be taken into account when assessing economic models and designing policy. For example, in new-Keynesian models a key issue is how to design the interest rate policy rule. Currently there is discussion about the possibility of targeting the average inflation rate over a stated interval of time. Over how many periods remains an open question, and our findings suggest that forecast horizon should be taken into consideration when designing such a policy.

We have assumed a stationary setup, but policy in macro models often is concerned with announced temporary changes. Examples include forward guidance in monetary policy and fiscal stimulus with announced durations. Clearly the efficacy of these policies depends on the expectations of agents, and thus on their forecast horizons. There are well-known puzzles related to announced policy under rational expectations, which can be ameliorated when RE is replaced by adaptive learning. A fruitful area for research would be to extend the approach in this paper to study how the forecast horizon affects theoretical and experimental results in the context of announced policy changes.

References

Adam, K., April 2007. Experimental Evidence on the Persistence of Output and Inflation. Economic Journal 117 (520), 603–636.

Adam, K., Marcet, A., Nicolini, J. P., 02 2016. Stock Market Volatility and Learning. Journal of Finance 71 (1), 33–82.

Andreassen, P. B., Kraus, S. J., 1990. Judgmental extrapolation and the salience of change. Journal of Forecasting 9 (4), 347–372.

Anufriev, M., Chernulich, A., Tuinstra, J., Aug. 2020. Asset Price Volatility and Investment Horizons: An Experimental Investigation. Working Papers 20200053, New York University Abu Dhabi, Department of Social Science.

Arifovic, J., Duffy, J., 2018. Chapter 9 - Heterogeneous Agent Modeling: Experimental Evidence. In: Hommes, C., LeBaron, B. (Eds.), Handbook of Computational Economics. Vol. 4 of Handbook of Computational Economics. Elsevier, pp. 491 – 540.

Asparouhova, E., Bossaerts, P., Roy, N., Zame, W., 2016. 'Lucas' In The Laboratory. Journal of Finance 71 (6), 2727–2780.

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

Blanchard, O. J., Watson, M. W., 1982. Bubbles, Rational Expectations and Financial Markets. In: Wachtel, P. (Ed.), *Crisis in the Economic and Financial Structure*. Lexington, MA.

Branch, W., Evans, G. W., 2011. Learning about Risk and Return: A Simple Model of Bubbles and Crashes. American Economic Journal: Macroeconomics 3 (3), 159–191.

Branch, W., Evans, G. W., McGough, B., 2012. Finite Horizon Learning . In: Sargent, T., Vilmunen, J. (Eds.), *Macroeconomics at the Service of Public Policy*. Oxford University Press.

Bundick, B., Hakkio, C. S., Mar 9 2015. Are longer-term inflation expectations stable? Macro Bulletin, 1–3Federal Reserve Bank of Kansas City.

Charness, G., Genicot, G., 2009. Informal Risk Sharing in an Infinite-Horizon Experiment. Economic Journal 119 (537), 796–825.

Colasante, A., Alfarano, S., Camacho-Cuena, E., Gallegati, M., 2020. Long-run expectations in a learning-to-forecast experiment: a simulation approach. Journal of Evolutionary Economics 30 (1), 75–116.

Cornand, C., Hubert, P., 2020. On the external validity of experimental inflation forecasts: A comparison with five categories of field expectations. Journal of Economic Dynamics and Control 110, 1037–46.

Crockett, S., Duffy, J., Yehuda, I., 2019. An Experimental Test of the Lucas Asset Pricing Model. Review of Economic Studies 86, 627–667.

De Jong, J., Sonnemans, J., Tuinstra, J., 2022. The effect of futures markets on the stability of commodity prices. Journal of Economic Behavior & Organization 198, 176–211.

De Long, J. B., Shleifer, A., Summers, L. H., Waldmann, R. J., 1990. Positive Feedback Investment Strategies and Destabilizing Rational Speculation. The Journal of Finance 45 (2), 379–395.

Duffy, J., 2016. Macroeconomics: A Survey of Laboratory Research. In: Kagel, J. H., Roth, A. E. (Eds.), *The Handbook of Experimental Economics, Volume 2*. Introductory Chapters. Princeton University Press.

Duffy, J., Jiang, J. H., Xie, H., Jul. 2019. Experimental Asset Markets with an Indefinite Horizon. Cahiers de recherche 08-2019, Centre interuniversitaire de recherche en Economie quantitative, CIREQ.

Dufwenberg, M., Lindqvist, T., Moore, E., 2005. Bubbles and Experience: An Experiment. American Economic Review 95 (5), 1731–1737.

Eusepi, S., Preston, B., 2011. Expectations, Learning and Business Cycle Fluctuations. American Economic Review 101 (6), 2844–2872.

Evans, G., Hommes, C., McGough, B., Salle, I., 2022. Data and codes for 'Are Long-Horizon Expectations (De-)Stabilizing? Theory and Experiments'. Mendeley Data, V1, doi: 10.17632/y2sjc22nyy.1.

Evans, G. W., Guesnerie, R., McGough, B., 2019. Eductive stability in real business cycle models. Economic Journal 129, 821–852.

Evans, G. W., Honkapohja, S., 2001. Learning and Expectations in Macroeconomics. Princeton University Press.

Frankel, J., Froot, K., 1987. Short-term and long-term expectations of the yen/dollar exchange rate: evidence from survey data. Journal of the Japanese and International Economies 1, 249–274.

Fréchette, G., Yuksel, S., 2017. Infinitely Repeated Games in the Laboratory: Four Perspectives on Discounting and Random Termination. Experimental Economics 20 (2), 279–308.

Grandmont, J. M., Laroque, G., 1986. Stability of cycles and expectations. Journal of Economic Theory 40, 138–151.

Haruvy, E., Lahav, Y., Noussair, C., 2007. Traders' expectations in asset markets: Experimental evidence. American Economic Review 97 (5), 1901–1920.

Hirota, S., Huber, J., Stock, T., Sunder, S., Jun. 2015. Investment Horizons and Price Indeterminacy in Financial Markets. Cowles Foundation Discussion Papers 2001, Cowles Foundation for Research in Economics, Yale University.

Hirota, S., Sunder, S., 2007. Price bubbles sans dividend anchors: Evidence from laboratory stock markets. Journal of Economic Dynamics and Control 31 (6), 1875–1909.

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

Keynes, J. M., 1936. The General Theory of Employment, Interest and Money. Macmillan.

Kopányi-Peuker, A., Weber, M., 2021. Experience Does Not Eliminate Bubbles: Experimental Evidence. The Review of Financial Studies 34, 4450–4485.

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

LeBaron, B., 2006. Agent-based computational finance. In: Tesfatsion, L., Judd, K. (Eds.), Handbook of Computational Economics, Volume 2: Agent-Based Computational Economics. Elsevier Science B.V., pp. 1109–1186.

Lucas, R. E., 1978. Asset Prices in an Exchange Economy. Econometrica 46, 1429–1445.

Marimon, R., Spear, S. E., Sunder, S., 1993. Expectationally Driven Market Volatility: An Experimental Study. Journal of Economic Theory 61 (1), 74–103.

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

Newey, W. K., West, K. D., 1994. Automatic Lag Selection in Covariance Matrix Estimation. Review of Economic Studies 61, 631– 653.

Noussair, C., Plott, C., Riezman, R., 1995. An Experimental Investigation of the Patterns of International Trade. American Economic Review 85 (3), 462–491.

Noussair, C., Tucker, S., 2006. Futures Markets And Bubble Formation In Experimental Asset Markets. Pacific Economic Review 11 (2), 167–184.

Noussair, C. N., Tucker, S., 2013. Experimental research on asset pricing. Journal of Economic Surveys 27 (3), 554–569.

Palan, S., 2013. A Review Of Bubbles And Crashes In Experimental Asset Markets. Journal of Economic Surveys 27 (3), 570–588.

Preston, B., 2005. Learning about Monetary Policy Rules when Long-Horizon Forecasts Matter. International Journal of Central Banking 1 (2), 81–126.

Rholes, R., Petersen, L., 2021. Should central banks communicate uncertainty in their projections? Journal of Economic Behavior & Organization 183, 320–341.

Roth, A. E., Murnighan, J. K., 1978. Equilibrium behavior and repeated play of the prisoner's dilemma. Journal of Mathematical Psychology 17, 189–198.

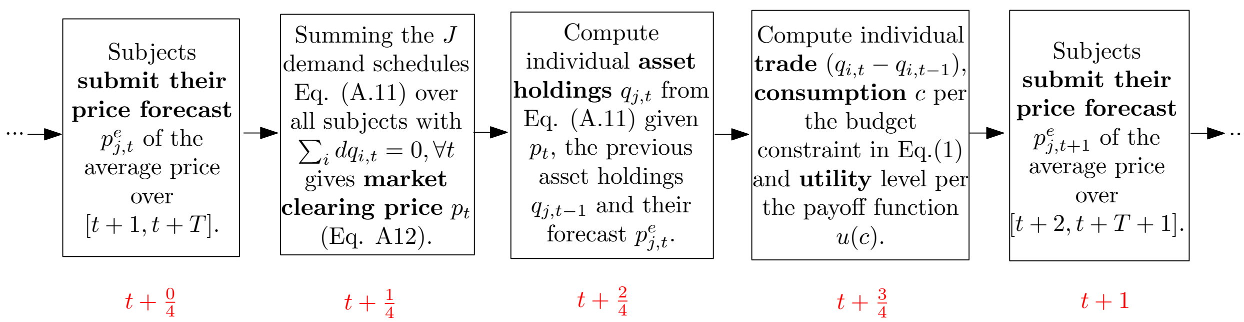
Sherstyuk, K., Tarui, N., Saijo, T., 2013. Payment schemes in infinite-horizon experimental games. Experimental Economics 16, 125–153.

Shiller, R. J., 1990. Speculative Prices and Popular Models. Journal of Economic Perspectives 4 (2), 55–65.

Smith, V. L., Suchanek, G. L., Williams, A. W., 1988. Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets. Econometrica 56 (5), 1119–1151.

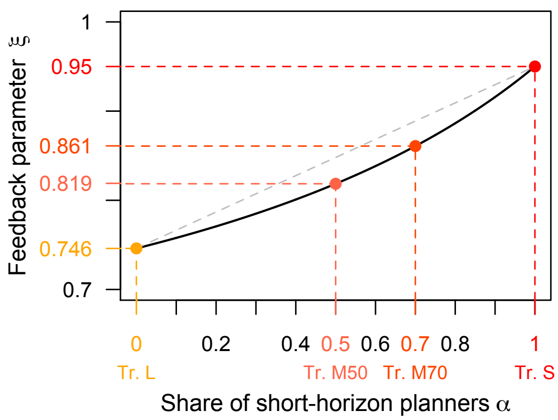
Woodford, M., 2018. Monetary policy analysis when planning horizons are finite. NBER Macroeconomics Annual 33, 1 – 50.

Woodford, M., Xie, Y., 2019. Policy Options at the Zero Lower Bound When Foresight is Limited. AEA Papers and Proceedings 109, 433–37.

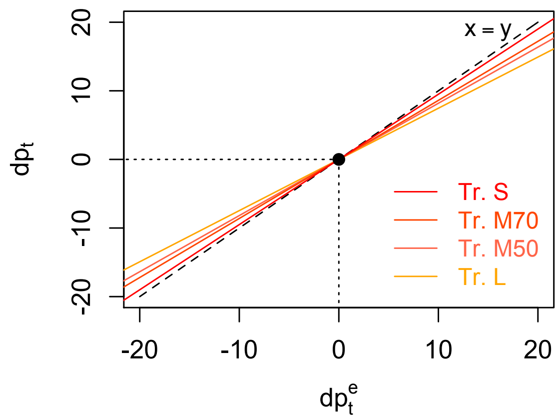


Note: in the experiment, we use a two-type version of the model with $T_i = \{1, 10\}$, $i = 1, 2$ and $J = 10$ subjects. The share α of short-horizon forecasters is a treatment variable; see Table 2. The steady state values of the price p, the chicken endowment q and the egg dividend y vary in each market; see Table 1.

Figure 1: Timing of events within one period of an experimental market

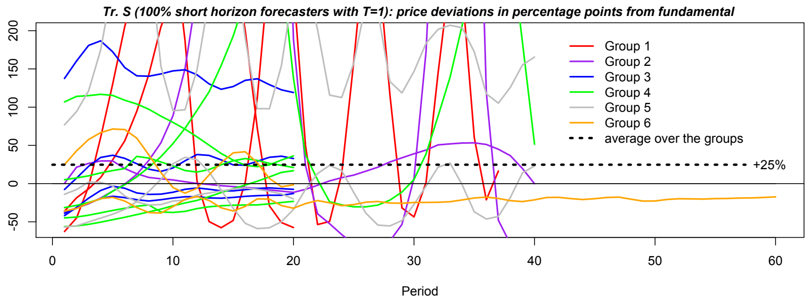


(a) Feedback ξ as a function of α

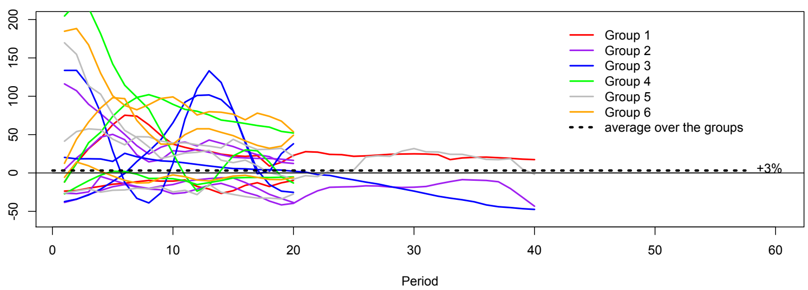


(b) Relationship between forecasts and prices

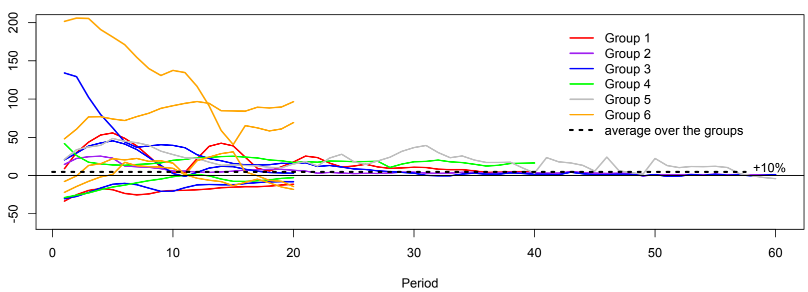
Figure 2: Price equation in the four experimental treatments assuming homogeneous expectations



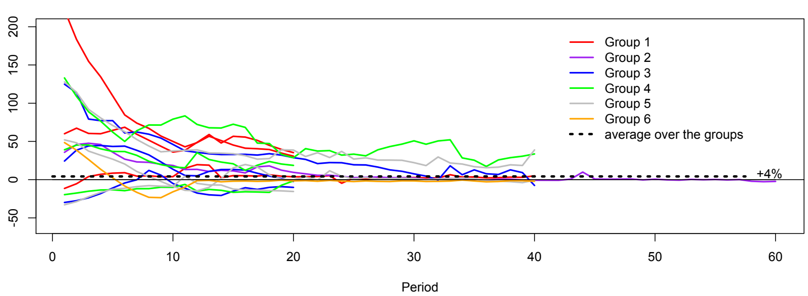
(a) Treatment S: 100% of short-horizon forecasters



(b) Treatment M70: 70% of short-horizon forecasters, 30% of long-horizon forecasters



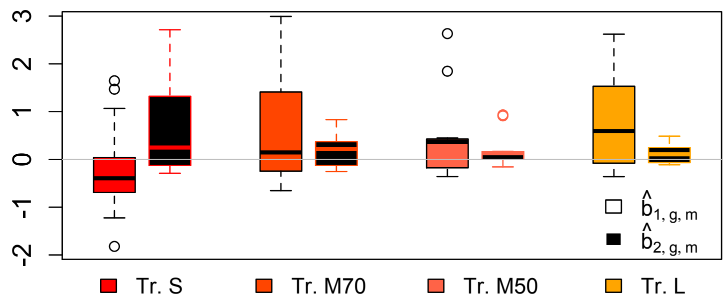
(c) Treatment M50: 50% of short-horizon forecasters, 50% of long-horizon forecasters



(d) Treatment L: 100% of long-horizon forecasters

Note: the plots report deviations in percentage points from the fundamental value.

Figure 3: Overview of the realized price levels in all experimental markets



	Market level	
	weak conv: $\mid \hat{b}_{1,g,m} \mid > \mid \hat{b}_{2,g,m} \mid$	strong conv: $\mid \hat{b}_{2,g,m} \mid = 0$
Tr. `S`	$7/18 \simeq 39\%$	$3/18 \simeq 17\%$
Tr `M70`	$11/18 \simeq 61\%$	$2/18 \simeq 11\%$
Tr. `M50`	$10/13 \simeq 77\%$	$3/13 \simeq 23\%$
Tr. `L`	$13/14 \simeq 93\%$	$4/14 \simeq 29\%$

Note: upper panel: distribution of estimated initial ($\hat{b}_{1,g,m}$) and final ($\hat{b}_{2,g,m}$) price values in relative deviation from fundamental per treatment. Lower panel: number of markets exhibiting weak and strong convergence, as defined in the main text, over the total number of markets in each treatment, and corresponding fractions of converging markets.

Figure 4: Results of the convergence assessment

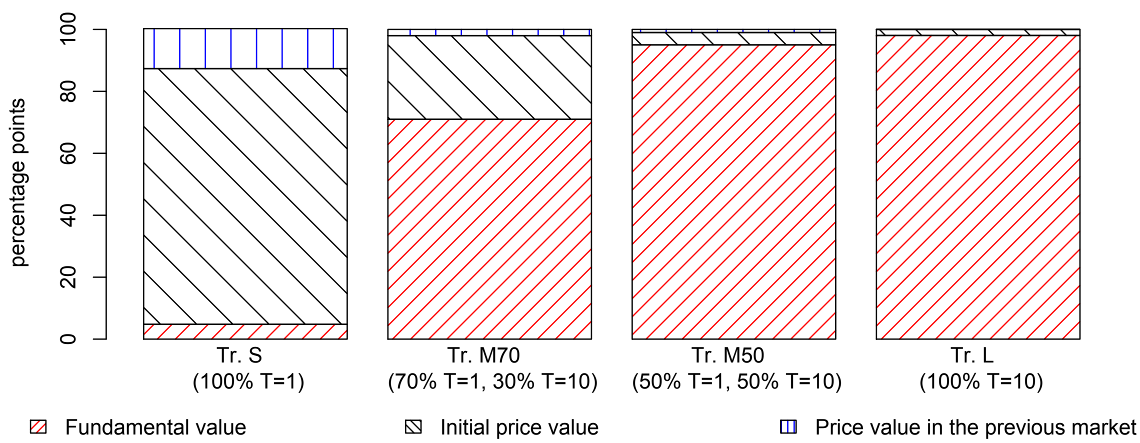
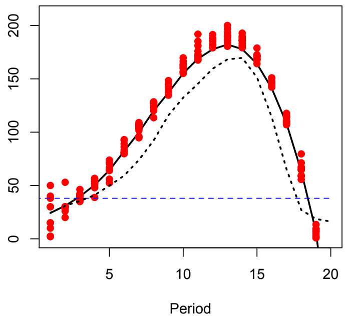
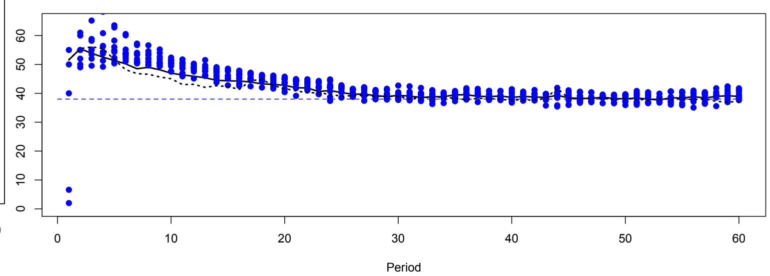


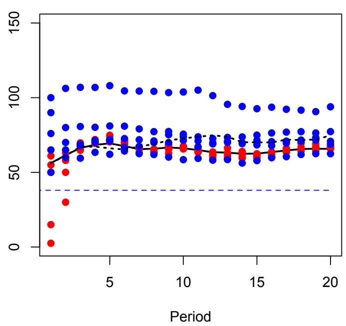
Figure 5: Contribution to the variance of the estimated final values $\hat{b}_{2,g,m}$



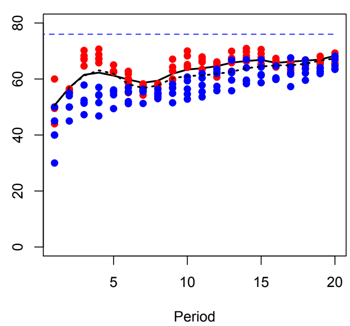
(a) Bubble and crash with trend-chasing forecasting



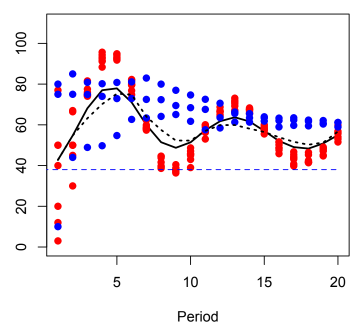
(b) Convergence with adaptive learners



(c) Overpricing with myopic and adaptive learners



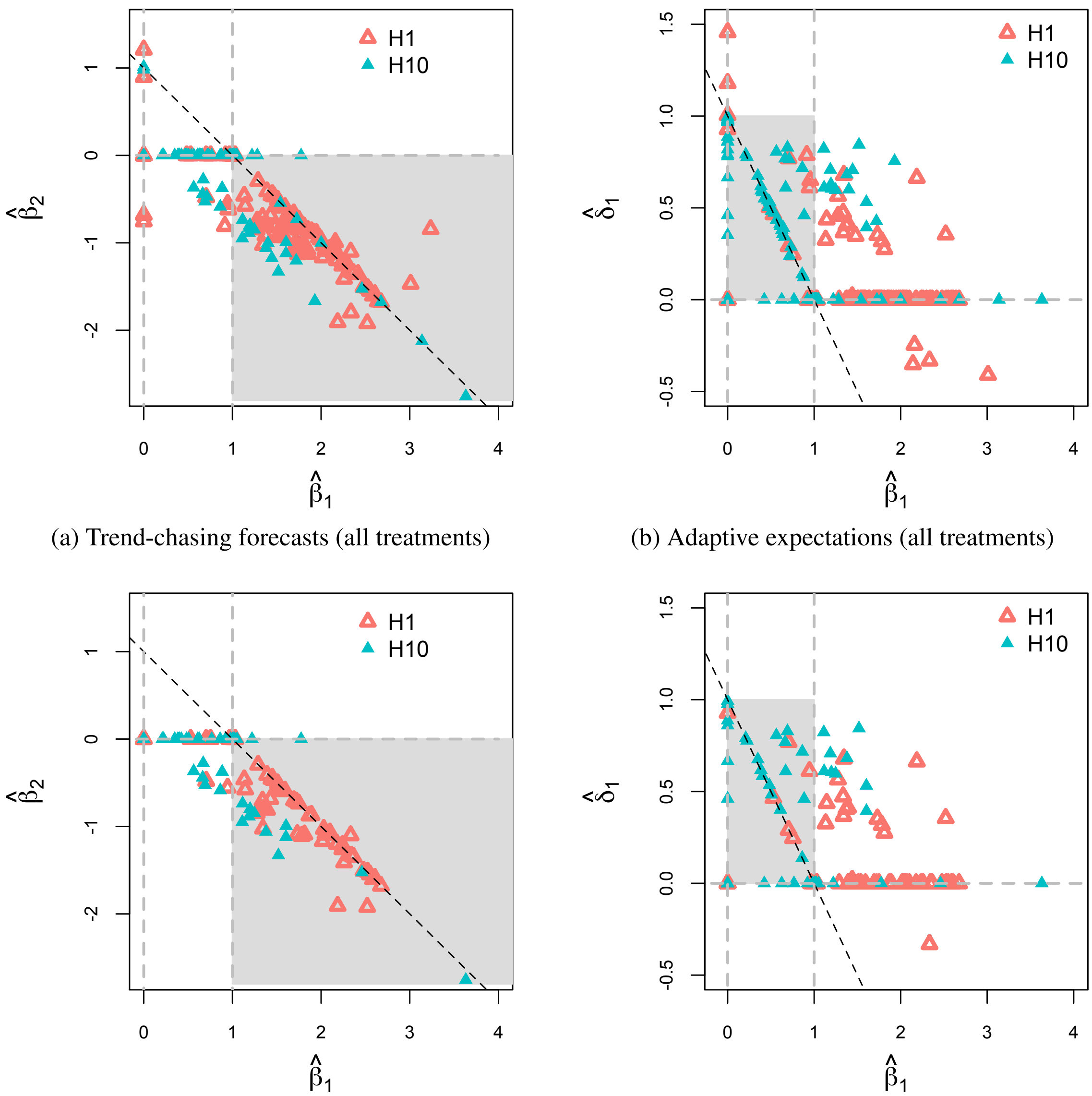
(d) Underpricing with trend-chasing and adaptive learners



(e) Overpricing with oscillations with trend-chasing and adaptive learners

Note: The blue dashed line is the fundamental price, the dotted lines represent the prices in the experimental markets, the dots and the solid lines are the simulated forecasts and prices. The forecasts in the first two periods are taken from the experiment. An idiosyncratic shock distributed as $\mathcal{N}(0,2)$ is added then in each subsequent period to the forecasts. Fig. (a): Tr. `S`, Gp. 1, Market 1, trend-chasing forecasting model with $\beta_1 = 1.3$ (see Eq. (5) below); Fig. (b): Tr. `L`, Gp. 2, Market 1, convergence with adaptive learning, $\delta_1 = 0.3$; Fig. (c): Tr. `M50`, Gp. 6, Market 1, overpricing with static short-horizon forecasters ($\beta_1 = 1$) and adaptive long-horizon forecasters ($\delta_1 = 0.1$); Fig. (d): Tr. `M50`, Gp. 1, Market 2, trend-chasing short-horizon forecasters ($\beta_1 = 1.3$) and adaptive long-horizon forecasters ($\delta_1 = 0.1$); Fig. (e): Tr. `M70`, Gp. 6, Market 1, trend-chasing short-horizon forecasters ($\beta_1 = 1.75$), adaptive long-horizon forecasters ($\delta_1 = 0.1$).

Figure 6: Simulated versus experimental time series for selected price patterns



(a) Trend-chasing forecasts (all treatments)

(b) Adaptive expectations (all treatments)

(c) Trend-chasing forecasts (Trs. `M50` and `M70` only)

(d) Adaptive expectations (Trs. `M50` and `M70` only)

Figure 7: Distribution of the estimated coefficients of Eq. (5)

	Markets				
	Market 1	Market 2	Market 3	Market 4	Market 5
Dividend y	2	4	1	5	3
Fundamental price p	38	76	19	95	57
Endowment q	4100	2100	8200	1700	2700

Table 1: Calibration of the markets, all groups, all treatments

	Treatments			
	Tr. `L`	Tr. `M50`	Tr. `M70`	Tr. `S`
Share α with horizon $T = 1$ (and number of forecasters)	0 (0 subject)	0.5 (5 subjects)	0.7 (7 subjects)	1 (10 subjects)
Share $1-\alpha$ with horizon $T = 10$ (and number of forecasters)	1 (10 subjects)	0.5 (5 subjects)	0.3 (3 subjects)	0 (0 subject)
Values of $\{\xi_s, \xi_l\}$	$\{0, 0.746\}$	$\{0.338, 0.481\}$	$\{0.534, 0.326\}$	$\{0.95, 0\}$
Number of independent observations (number of participants)	6 (60)	6 (60)	6 (60)	6 (60)

Notes: $\{\xi_s, \xi_l\}$ refer to the feedback parameters in equation (3) associated with the average forecast among, respectively, the short-horizon and the long-horizon subjects.

Table 2: Summary of the four experimental treatments

	Diff-diff treatments					
	L-S	*L-M70*	*L-M50*	*M70-S*	*M50-S*	*M50-M70*
Price deviation[a]	**-0.564**	**-0.111**	0.012	**-0.453**	**-0.576**	**-0.123**
(p-value)	(0.000)	(0.000)	(0.205)	(0.000)	(0.000)	(0.000)
Price volatility[b]	**-2.123**	**-0.111**	-0.029	**-2.013**	**-2.094**	**-0.082**
(p-value)	(0.000)	(0.000)	(0.315)	(0.000)	(0.000)	(0.000)
Trade volume[c]	**0.088**	**0.061**	**0.140**	**0.027**	**-0.052**	**-0.079**
(p-value)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Forecast dispersion[d]	**0.161**	0.080	**0.115**	**0.081**	0.046	**-0.035**
(p-value)	(0.030)	(0.970)	(0.094)	(0.010)	(0.968)	(0.049)
EER (forecasts)[e]	-0.071	-0.026	-0.083	-0.045	0.012	0.057
(p-value)	(0.231)	(0.924)	(0.452)	(0.304)	(0.990)	(0.622)
EER (utility)[e]	0.010	-0.003	0.002	0.013	0.008	-0.005
(p-value)	(0.965)	(0.984)	(0.614)	(0.663)	(0.754)	(0.414)

Note: The table reports the differences between treatments, and the associated p-values of the two-sided Wilcoxon rank-sum tests (except to compare the cross-treatment price volatility where we use a Levene test). In bold are the significant differences between treatments. K-S tests give the same predictions, except between treatments `M50` and `L` regarding the price deviation, in which case the pair-difference becomes insignificant.

[a] Average of the absolute price deviation from its fundamental value p_m, over all periods $t \geq 1$ of each market m, computed as $(p_m)^{-1} \mid p_{m,t} - p_m \mid$.

[b] Variance of the price normalized by the fundamental value computed as $Var\left(\frac{p_{m,t}}{p_m}\right)$.

[c] Sum over all periods t and all markets m of exchanged assets among subjects in proportion of the steady-state endowment q_m, i.e. $\sum_{j=1}^{10} \mid \frac{q_{j,t} - q_{j,t-1}}{q_m} \mid$.

[d] Relative standard deviation between subjects' forecasts $\frac{\sqrt{Var(p^e_{j,t})_{j \in J}}}{mean(p^e_{j,t})_{j \in J}}$, $t \geq 1$, averaged over all periods of each market.

[e] Earnings Efficiency Ratio (EER) computed over all periods of each market, averaged over the 10 subjects as follows: (i) for the forecasting task, it is the average number of forecasting points earned in each market over the total amount of points possible in the market (1100 per period in case of perfect prediction); (ii) for the consumption task, it is the average number of utility points earned in each market over the total utility points earned at equilibrium (1081 per period).

Table 3: Cross-treatment statistical comparisons