

## Evolutionary Selection of Individual Expectations and Aggregate Outcomes in Asset Pricing Experiments<sup>†</sup>

By MIKHAIL ANUFRIEV AND CARS HOMMES\*

*In recent “learning to forecast” experiments (Hommes et al. 2005), three different patterns in aggregate price behavior have been observed: slow monotonic convergence, permanent oscillations, and dampened fluctuations. We show that a simple model of individual learning can explain these different aggregate outcomes within the same experimental setting. The key idea is evolutionary selection among heterogeneous expectation rules, driven by their relative performance. The out-of-sample predictive power of our switching model is higher compared to the rational or other homogeneous expectations benchmarks. Our results show that heterogeneity in expectations is crucial to describe individual forecasting and aggregate price behavior. (JEL C53, C91, D83, D84, G12)*

In the economy, today’s individual decisions crucially depend upon expectations or beliefs about future developments. For example, in speculative asset markets such as the stock market, an investor buys (sells) stocks today when she expects stock prices to rise (fall) in the future. Expectations affect individual trading decisions and the realized market outcome (e.g., prices and traded quantities) is an aggregation of individual behavior. A market is therefore an *expectations feedback system*: market history shapes individual expectations which, in turn, determine current aggregate market behavior and so on. How exactly do individuals form market expectations, and what is the aggregate outcome of the interaction of individual market forecasts? To answer these questions in this paper we reinvestigate individual learning behavior observed in the laboratory experiments of Hommes et al. (2005) and Hommes et al. (2008) (HSTV05 and HSTV08, henceforth), specifically designed to study expectations feedback. On the basis of the experimental evidence we propose a

\*Anufriev: Economics Discipline Group, University of Technology Sydney, PO Box 123 Broadway, NSW 2007, Australia (e-mail: Mikhail.Anufriev@uts.edu.au); Hommes: CeNDEF, School of Economics, University of Amsterdam, Roetersstraat 11, NL-1018 WB Amsterdam, Netherlands (e-mail: C.H.Hommes@uva.nl). We thank the participants of the “Complexity in Economics and Finance” workshop, Leiden, October 2007, the “Computation in Economics and Finance” conference, Paris, June 2008, the “Learning and Macroeconomic Policy” workshop, Cambridge, September 2008, and seminar participants at the University of Amsterdam, Sant’Anna School of Advanced Studies (Pisa), University of New South Wales, University of Technology Sydney, Catholic University of Milan, University of Warwick, Paris School of Economics, and the European University Institute (Florence) for stimulating discussions. We benefitted from helpful comments by Buz Brock, Cees Diks, John Duffy, Doyne Farmer, Frank Heinemann, and Valentyn Panchenko and from detailed comments of three anonymous referees on earlier versions of this paper. This work is supported by the EU FP7 collaborative projects POLHIA, grant no. 225408, and CRISIS (Complexity Research Initiative for Systemic Instabilities), grant no. 288501, and by a Complexity grant of the Netherlands Organisation for Scientific Research (NWO).

<sup>†</sup>To comment on this article in the online discussion forum, or to view additional materials, visit the article page at <http://dx.doi.org/10.1257/mic.4.4.35>.

*behavioral model of heterogeneous expectations*, which fits individual forecasting behavior as well as aggregate market outcomes in the experiments quite well.

After the seminal work by Muth (1961) and Lucas (1972), it has become common practice in economic theory to assume that all individuals have rational expectations. In a rational world individual expectations coincide, on average, with market realizations, and markets are efficient with prices fully reflecting economic fundamentals (Samuelson 1965; Fama 1970). In the traditional view, there is no room for market psychology and “irrational” herding behavior. An important underpinning of the rational expectations approach comes from an early evolutionary argument made by Friedman (1953), that “irrational” traders will not survive competition and will be driven out of the market by rational traders, who will trade against them and earn higher profits.

Following Simon (1957), many economists have argued that rationality imposes unrealistically strong informational and computational requirements upon individual behavior and that it is more reasonable to model individuals as *boundedly rational*, using simple rules of thumb in decision making. Laboratory experiments indeed have shown that individual decisions under uncertainty are at odds with perfectly rational behavior, and can be much better described by simple *heuristics*, which sometimes may lead to persistent biases (Tversky and Kahneman 1974; Kahneman 2003; Camerer and Fehr 2006). Models of bounded rationality have also been applied to forecasting behavior, and several adaptive learning algorithms have been proposed to describe market expectations. For example, Sargent (1993) and Evans and Honkapohja (2001) advocate the use of *adaptive learning* in modeling expectations, where agents learn unknown parameters of the model using econometric techniques on past observations. In some models (Bray and Savin 1986) adaptive learning enforces convergence to rational expectations, while in others (Bullard 1994) learning may not converge at all but instead lead to excess volatility and persistent deviations from rational expectations equilibrium similar to field settings (Shiller 1981; DeBondt and Thaler 1989). While most of the initial adaptive learning models have been used in the macroeconomics literature, several of them recently have been applied to explain various phenomena of financial markets. The results are mixed. Adam, Marcet, and Nicolini (2011) show that when agents predict price returns with least-squares learning, the moments of time series of returns and price-dividend ratios are significantly closer to the actual values than under rational expectations. On the other hand, Carceles-Poveda and Giannitsarou (2008) study similar least-squares learning in a general equilibrium framework and find that it is not sufficient to bring moments toward plausible values. Branch and Evans (2011) show that a model with least-squares learning of returns and the conditional variance of returns generates repeated bubbles and crashes.

Recently, models with *heterogeneous expectations* and evolutionary selection among forecasting rules have been proposed, in Brock and Hommes (1997) and Branch and Evans (2006), see Hommes (2006) for an extensive overview. By taking a larger deviation from the rational expectation framework, these models were able to match moments of different financial variables and generate several “stylized facts” of financial markets, such as excess volatility, volatility clustering, and fat tails of the return distribution. See Gaunersdorfer, Hommes, and Wagener (2008);

Anufriev and Panchenko (2009); Franke and Westerhoff (2011), among others. At the same time Boswijk, Hommes and Manzan (2007) and de Jong, Verschoor, and Zwinkels (2009) estimated heterogeneous expectation models on financial market data and found evidence of heterogeneity. Fitting such models to “clean” experimental data, as we do in this paper, provides an important additional test for heterogeneous expectations models. Moreover, we will use the experimental data to provide an empirical microfoundation for the model, since the forecasting rules which we will employ were among the individual rules estimated in the laboratory.

Laboratory experiments with human subjects are well suited to study individual expectations and how their interaction shapes aggregate market behavior (Marimon, Spear, and Sunder 1993; Peterson 1993). But the results from laboratory experiments are mixed. Early experiments by Smith (1962) show convergence to equilibrium, while more recent asset pricing experiments exhibit deviations from equilibrium with persistent bubbles and crashes, see, e.g., Smith, Suchanek, and Williams (1988) and Lei, Noussair, and Plott (2001). It is important to recognize that in many earlier experiments expectations only play a secondary role, being intertwined with other aspects, such as market architecture and trading behavior of participants. In order to provide clean data on expectations and control all other underlying model assumptions, HSTV05 designed a so-called “learning to forecast” experiment.<sup>1</sup> In a typical session six human subjects had to predict the price of an asset for 50 periods, having knowledge of the fundamental parameters (mean dividend and interest rate) and previous price realizations. Trading is computerized, using an optimal demand schedule derived from maximization of myopic constant absolute risk aversion (CARA) mean-variance utility, given the subject’s individual forecast. Hence, subject’s only task in every period is to give a two period ahead point prediction for the price of the risky asset, and their earnings are inversely related to their prediction errors. Learning to forecast experimental data can be used as a test bed for various expectations hypotheses, such as rational expectations or adaptive learning models, in any benchmark dynamic economic model with all other assumptions controlled by the experimenter; see Duffy (2008).

HSTV05 ran 14 sessions of the learning to forecast experiment in three different treatments. In seven sessions of one of the treatments, the outcomes were quite different, despite identical settings. While in some sessions price convergence did occur, in other sessions prices persistently fluctuated and temporary bubbles emerged (see Figure 2, panel C and the right panels of Figure 3). HSTV08 ran six sessions in another treatment and show that the typical outcome of the experiment is the emergence of long-lasting bubbles followed by crashes (see Figure 2, panel B). Another striking and robust finding of the learning to forecast experiments is that in all sessions individuals were able to coordinate on a common prediction rule,

<sup>1</sup> See Hommes (2011) for an overview of learning to forecast experiments, where also earlier experiments on expectation formation are discussed, e.g., Williams (1987) and Hey (1994) among others. The experiments in HSTV05 and HSTV08, which we discuss in this paper, study expectations in financial markets. Experiments in Heemeijer et al. (2009); and Bao et al. (forthcoming) consider markets of agricultural goods in a cobweb framework, while Adam (2007); Assenza et al. (2011); and Pfajfar and Zaketj (2010) ran experiments in a macroeconomic environment. Anufriev, Hommes, and Philipse (forthcoming) fit a version of the heuristic switching model to the Heemeijer et al. (2009) experiment.

without any knowledge about the forecasts of others (see Figure 3, in the Predictions portion of the panels).

Anufriev and Hommes (2012) showed by three 50-periods-ahead simulations that a simple learning model of heterogeneous expectations can generate dynamics qualitatively similar to the three outcomes observed in some of the HSTV05 experiments. In this paper we fit the model to all 20 different market sessions in the HSTV05 and HSTV08 experiments and perform one period ahead and out-of-sample forecasting. Our model is an extension of the *evolutionary selection* mechanism introduced in Brock and Hommes (1997), applied to simple prediction rules which individuals often used in the experiment. From the point of view of the participants several plausible behavioral rules can be used for price prediction at each time step. Participants can choose any of these rules, but at the aggregate level the rules which generated relatively good predictions in the recent past will be chosen more often. Participants also exhibit some inertia in their decisions, by temporarily sticking to their rule. Combining these ideas we arrive at a *behavioral model of learning* with only three free parameters. We show that this model fits the experimental data surprisingly well for a large range of parameters. The model outperforms several models with homogeneous expectations and a simple non-structural AR (2) model both in and out of sample. Our model of individual learning can explain different observed aggregate patterns within the same experimental environment (Figure 6) and is robust with respect to changes in the experimental environments (Figure 8).

The fact that the model with heterogeneous expectations can explain different aggregate outcomes observed within the same environment has a clear intuitive explanation. The key feature driving the result is the path-dependent property of the nonlinear switching model. If participants start to coordinate on an adaptive rule, the resulting (stable) price dynamics is such that the adaptive rule performs better than other rules, reinforcing the coordination and explaining convergence and stable price behavior. If, on the other hand, a majority of participants coordinates on a trend-following rule, price oscillations and temporary bubbles arise. In that case trend-following rules predict better than the adaptive rule, thus reinforcing and amplifying price trends and temporary bubbles.

The paper is organized as follows. In Section I, we review the findings of the laboratory experiments of HSTV05 and HSTV08. Section II discusses individual behavior of participants in the experiments, and studies the price dynamics under *homogeneous* forecasting rules. A learning model based on evolutionary selection between simple forecasting heuristics is presented in Section III. In Section IV we discuss how our model fits the experimental data. Finally, Section V concludes.

## I. Learning-to-Forecast Experiments

A number of computerized *learning-to-forecast experiments* (LtFEs) have been performed in the CREED laboratory at the University of Amsterdam, see Hommes (2011) for an overview. This paper proposes a theoretical explanation of the results obtained in 20 different sessions of the LtFEs based on the asset pricing model, see HSTV05 and HSTV08. In each session of the experiment, 6 participants were advisers to large pension funds and had to submit point forecasts for the price of a risky

asset for 50 consecutive periods. There were two assets in the market, a risk-free asset paying a fixed interest rate  $r$  every period, and a risky asset paying stochastic IID dividends with mean  $\bar{y}$ . Subjects knew that the price of the risky asset is determined every period by market clearing, as an aggregation of individual forecasts of all participants. They were also informed that the higher the forecasts are, the larger the demand for the risky asset is. Stated differently, they knew that there was positive feedback from individual price forecasts to the realized market price. Trading had been computerized with the price  $p_t$  determined in accordance with a standard mean-variance asset pricing model with heterogeneous beliefs (Campbell, Lo, and MacKinlay 1997; Brock and Hommes 1998):

$$(1) \quad p_t = \frac{1}{1+r} (\bar{p}_{t+1}^e + \bar{y} + \varepsilon_t), \quad t = 0, \dots, 50,$$

where  $\bar{p}_{t+1}^e = \frac{1}{6} \sum_{i=1}^6 p_{i,t+1}^e$  is the average of 6 individual forecasts, and a small stochastic term  $\varepsilon_t$  represents demand/supply shocks. The same realizations of the shocks, drawn independently from a normal distribution with mean 0 and standard deviation 0.5, has been used in all sessions of HSTV05. An individual forecast,  $p_{i,t+1}^e$ , could be any number (with two decimals) in the range  $[0, 100]$  to be submitted at the beginning of period  $t$ , when the last observed price is from period  $t - 1$ . Subjects, therefore, had to make *two-periods-ahead* forecasts. After all forecasts were submitted, every participant was informed about the realized price  $p_t$ . The earnings per period were determined by a quadratic scoring rule

$$(2) \quad e_{i,t} = \begin{cases} 1 - \left(\frac{p_t - p_{i,t}^e}{7}\right)^2 & \text{if } |p_t - p_{i,t}^e| < 7, \\ 0 & \text{otherwise,} \end{cases}$$

so that forecasting errors exceeding 7 would result in no reward at a given period. At the end of the session, the accumulated earnings of every participant were converted to euros (1 point computed as in (2) corresponded to 50 cents). Subjects of the experiments neither knew the exact functional form of the market equilibrium equation (1) nor the number and identity of other participants. They were informed about the scoring rule (2) and the values of the fundamental parameters,  $r$  and  $\bar{y}$ , at the beginning of the experiment. Participants could, therefore, compute the *rational fundamental price* of the risky asset, given by the discounted sum of expected dividends, which reduces to  $p^f = \bar{y}/r$ . See, for example, Campbell, Lo, and MacKinlay (1997, 256) and Brock and Hommes (1998, 1239). The information set of participant  $i$  at period  $t$  consisted of past prices up to  $p_{t-1}$ , past own predictions up to  $p_{i,t}^e$ , the fundamental parameters  $r$  and  $\bar{y}$ , and past own earnings.

On the basis of variations in the implementation of the experiment, four different experimental settings (treatments) can be identified. The setup described above has been used in sessions 11 – 14 of HSTV05, with  $\bar{y} = 3$  and  $r = 0.05$ , resulting in a fundamental price  $p^f = 60$ . For reasons which will become clear below, we use the acronym **NoRo** (no robots) for these four sessions. Figure 1 shows the simulation of prices and prediction errors, which would occur when all individuals use the fundamental forecasting rule,  $p_{i,t+1}^e = 60$ , for all  $i$  and  $t$ . Under rational expectations, the

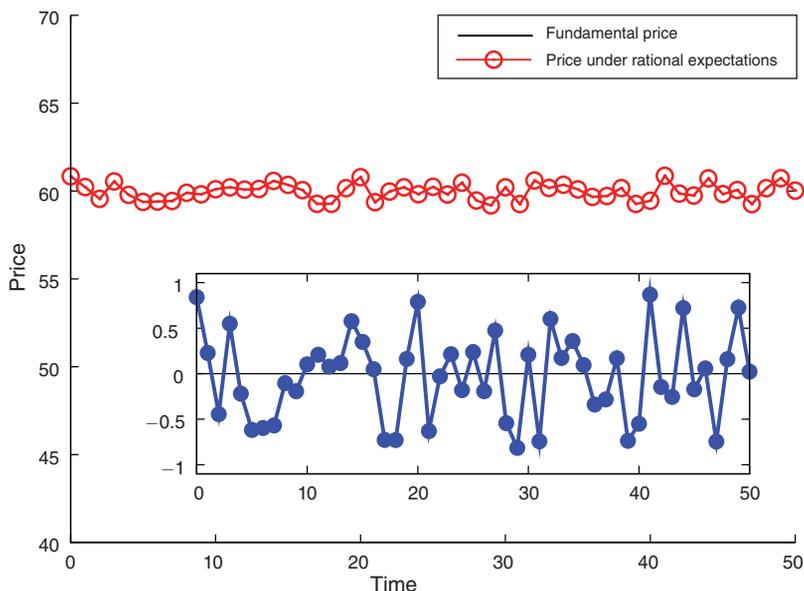


FIGURE 1. PRICE EVOLUTION AND PREDICTION ERRORS UNDER RE FUNDAMENTAL EXPECTATIONS

Notes: Simulation of equation (1) with  $\bar{y} = 3$  and stochastic shocks  $\varepsilon_t$  from the HSTV05 experiment. Prediction errors are shown in the inner frame.

realized price  $p_t = 60 + \varepsilon_t / (1 + r)$  randomly fluctuates around the fundamental level with small amplitude. In the experiment, one cannot expect rational outcome at the outset, but aggregate prices might converge to their fundamental value through individual learning. The actual price dynamics in the four sessions of the experiment are shown in Figure 2, panel A. Interestingly, only in one session the price converges to the fundamental value, while in three other sessions constant or dampened price oscillations with high amplitude are observed.

HSTV08 ran six sessions of a different treatment of the LtFE, where the price is determined according to (1) without shocks  $\varepsilon_t$ , and where participants could submit their predictions in a much larger range  $[0, 1000]$ . The experimental results in all six sessions are shown in Figure 2, panel B under the acronym **LFR** (large forecasting range). An increase of the allowable prediction range results in long-lasting bubbles.

In order to make long-lasting bubbles less likely, HSTV05 used a slightly different pricing rule:

$$(3) \quad p_t = \frac{1}{1+r} \left( (1 - n_t) \bar{p}_{t+1}^e + n_t p^f + \bar{y} + \varepsilon_t \right),$$

where  $n_t$  represents the weight of the robot trader, whose forecast is always  $p^f$ ; the average forecast of the six participants has weight  $1 - n_t$ . In field settings, robot traders correspond to fundamentalists, who have a better understanding of the

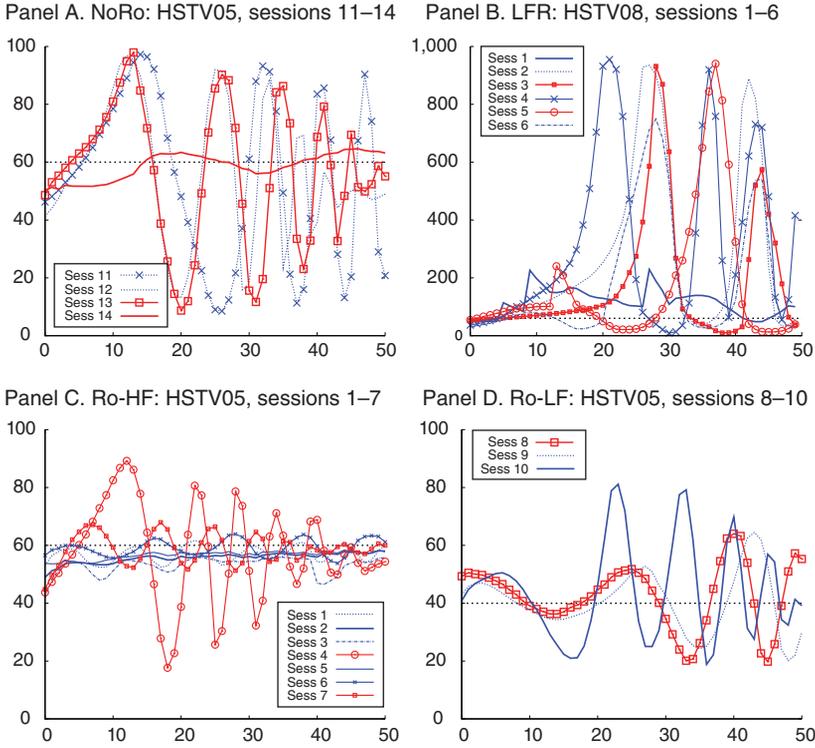


FIGURE 2. AGGREGATE MARKET PRICES IN FOUR DIFFERENT VARIATIONS OF THE LtFE

market environment than other traders. The weight of the robot traders increases in response to the deviations of the asset price from its fundamental level:

$$(4) \quad n_t = 1 - \exp\left(-\frac{1}{200} |p_{t-1} - p^f|\right).$$

This mechanism reflects the feature that in financial markets there is more agreement about over- or undervaluation of an asset when the price deviation from the fundamental level is large.<sup>2</sup> Seven experimental sessions, 1–7, had the fundamental parameters  $r = 0.05$  and  $\bar{y} = 3$  with fundamental price  $p^f = \bar{y}/r = 60$ . For this specification, an acronym **Ro-HF** (robots, high fundamental) is used. As Figure 2, panel C shows, in the presence of a stabilizing robot traders the amplitude of the oscillations decreases. Finally, in the remaining sessions 8–10, the market had a smaller dividend  $\bar{y} = 2$ , resulting in a smaller fundamental price  $p^f = \bar{y}/r = 40$ , see Figure 2, panel D. The oscillations are quite large in this **Ro-LF** (robots and low fundamental) case.

A closer look at six different sessions of the LtFEs is given in Figure 3. The left panels show time series of prices (upper parts of panels), individual predictions

<sup>2</sup>In the experiments, the fraction of robot trader never became larger than 0.25.

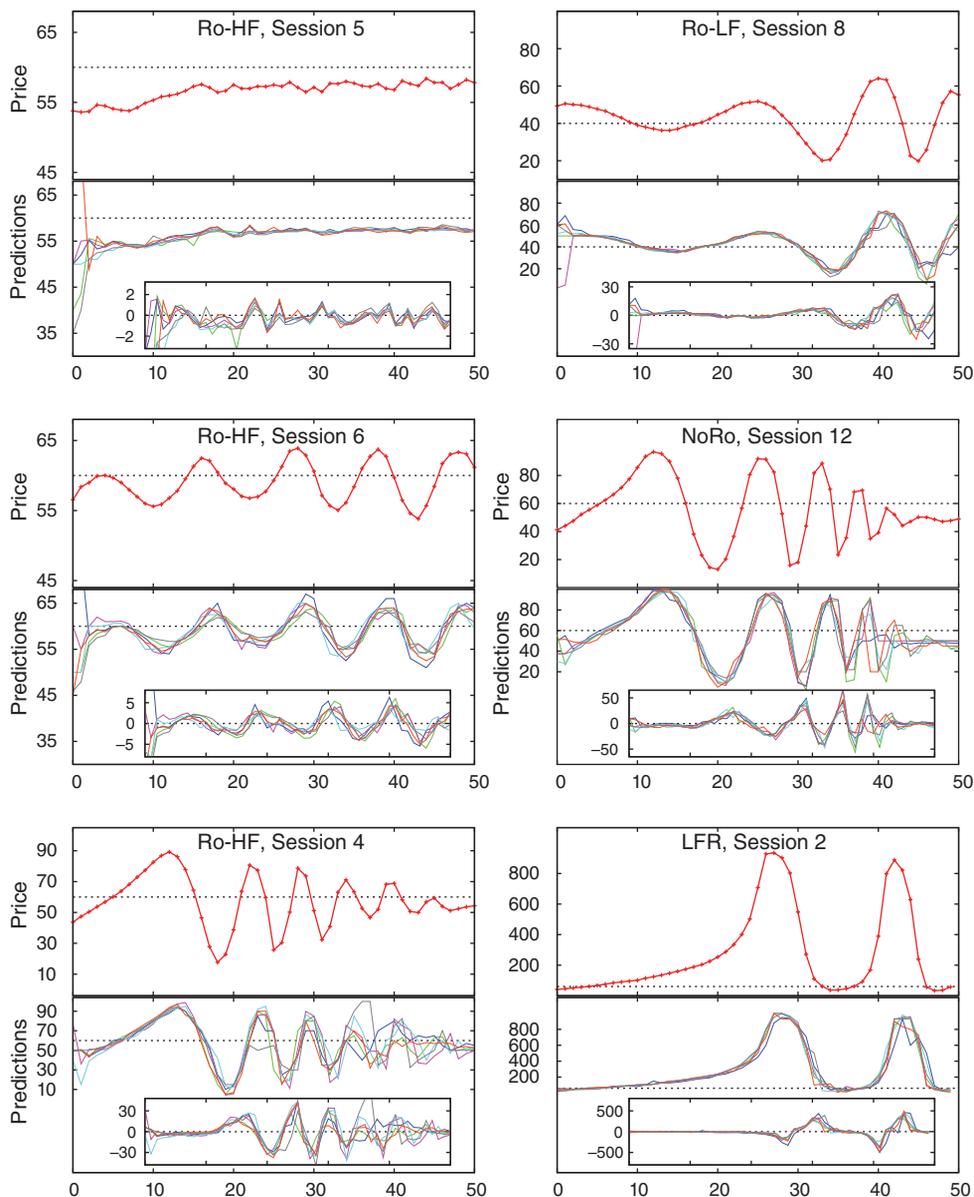


FIGURE 3. PRICES, PREDICTIONS, AND FORECASTING ERRORS IN SIX SESSIONS OF THE LTFES

*Notes:* The left panels illustrate three qualitatively different outcomes for the same **Ro-HF** treatment. The right panels illustrate typical examples of the **Ro-LF**, **NoRo**, and **LFR** treatments. Every panel shows time series of price (upper part), six individual predictions (lower part), and six forecasting errors (inner frame).

(lower parts of panels) and forecasting errors (inner frames) for three out of seven sessions in the **Ro-HF** treatment. A striking feature of this experiment is that *three different* qualitative patterns for aggregate price behavior emerge within the same environment. The prices in session 5 (and 2; not shown) converge slowly and almost monotonically to the fundamental price level. In session 6 (and 1; not shown), persistent oscillations are observed during the entire experiment. In session 4

(and 7; not shown), prices are also fluctuating but their amplitude is decreasing. The right panels of Figure 3 show the price fluctuations in the typical sessions of the other LtFEs: session 8 of **Ro-LF**, session 12 of **NoRo**, and session 2 of **LF**.<sup>3</sup>

Individual predictions in Figure 3 show another striking result of the LtFEs. In all experimental sessions, participants were able to *coordinate* their forecasting activity. The forecasts are dispersed in the first periods but then become very close to each other. The coordination of individual forecasts has been achieved in the absence of any communication between subjects and knowledge of past and present predictions of *other* participants.

LtFEs represent a tailored laboratory study to test different theories of *expectation* formation. A suitable model should be able to reproduce the following findings of the HSTV05 and HSTV08 learning-to-forecast experiments:

- Participants have not learned the RE fundamental forecasting rule; only in some cases individual predictions slowly moved in the direction of the fundamental price toward the end of the experiment.
- Three different price patterns were observed in the same **Ro-HF** experiment: (i) slow, (almost) monotonic convergence, (ii) persistent oscillations with almost constant amplitude, and (iii) large initial oscillations dampening toward the end of the experiment.
- LtFEs tend to produce long lasting bubbles followed by crashes in the **LF** treatment.

The purpose of this paper is to explain these “stylized facts” simultaneously by a simple behavioral model of individual learning.

## II. Individual Forecasting Behavior

In this section, we will look closer at the individual forecasting behavior in the LtFEs in order to develop some behavioral foundations of the heterogeneous expectations model.

Which forecasting rules did individuals use in the learning to forecast experiment? Comparison of the RE benchmark in Figure 1, with the lab experiments in Figures 2 and 3, suggests that rational expectations is not a good explanation of individual forecasting and aggregate behavior. The REs cannot be reconciled with the dynamics of the HSTV05 experiment, especially not in the oscillating sessions. Using individual experimental data HSTV05 estimated the forecasting rules for each subject based on the last 40 periods (to allow for some learning phase). Simple linear rules of the form

$$(5) \quad p_{i,t+1}^e = \alpha_i + \beta_i p_{t-1} + \gamma_i (p_{t-1} - p_{t-2}) + \delta_i p_{i,t}^e$$

<sup>3</sup>See online Appendix A for similar plots in all remaining sessions. Sessions 2, 1, and 7 of **Ro-HF** experiment have been analyzed in Anufriev and Hommes (2012). The dynamics in session 3 of **Ro-HF** were peculiar. Similar to session 6 it started with moderate oscillations, then stabilized at a level below the fundamental, suddenly falling significantly in period  $t = 40$ , probably due to a typing error of one of the participants.

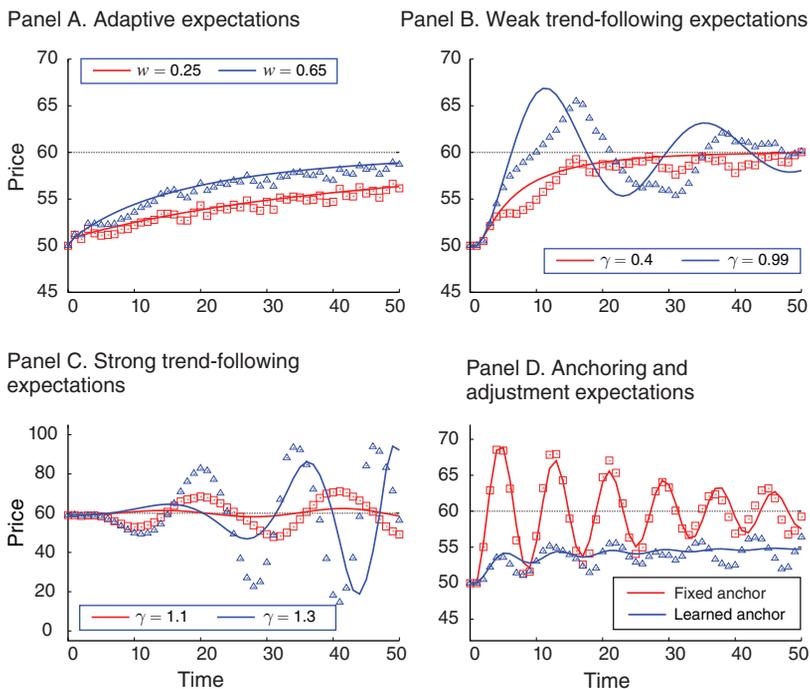


FIGURE 4. PRICE DYNAMICS WITH HOMOGENEOUS EXPECTATIONS

*Note:* The trajectories of the deterministic skeleton (the curves) and stochastic simulations with the same realization of shocks as in the experiment (triangles and squares) are simulated for different forecasting heuristics.

were estimated, where the index  $i$  stands for human subject. Notice that the information on the right-hand side is consistent with the experimental timing: at the moment when the forecast for the price  $p_{t+1}$  is submitted, the price  $p_t$  is still unknown, but his own last forecast  $p_{i,t}^e$  is known. Equation (5) is called *first-order heuristics*, because it only uses the last own forecast, the last observed price, and the last observed price change to predict the future price. Remarkably, individual forecasts are well explained by first-order heuristics: for 63 out of  $14 \times 6 = 84$  participants (i.e., for 75 percent) an estimated linear rule falls into this simple class with an  $R^2$  typically higher than 0.80.

The experimental evidence suggests strong coordination on a common prediction rule. One can therefore suspect that this common rule (which, for whatever reason, turned out to be different in different sessions) generates the resulting pattern. It is therefore useful to first investigate price dynamics under *homogeneous expectations* when all participants use the same forecasting rule. The model with homogeneous expectations is obtained when the average  $\bar{p}_{t+1}^e$  is generated by the rule of type (5) and then plugged into the corresponding pricing equation: equation (1) for the experiment without robots or equation (3) for the experiments with robots. Figure 4 illustrates deterministic (setting the noise term  $\varepsilon_t \equiv 0$ ) as well as stochastic simulations of the model with the same realization of the shocks as in the experiment.

*Adaptive Heuristic.*—Participants from the converging sessions often used an *adaptive expectations* rule of the form

$$(6) \quad p_{t+1}^e = w p_{t-1} + (1 - w) p_t^e = p_t^e + w(p_{t-1} - p_t^e),$$

with weight  $0 \leq w \leq 1$ . Notice that for  $w = 1$ , we obtain the special case of *naive expectations*.<sup>4</sup>

Assume that all participants use the same adaptive heuristic  $p_{t+1}^e = w p_{t-1} + (1 - w) p_t^e$  in their forecasting activity. It is easy to show (see online Appendix B) that in the absence of stochastic shock  $\varepsilon_t$  the dynamics will monotonically converge to the fundamental price, independently of initial conditions. This is illustrated in Figure 4, panel A for two different values of the weight  $w$  assigned to the past price. Shocks slightly perturb the system, but the stochastic price series (shown by triangles and squares) still exhibit almost monotonic convergence. Coordination on adaptive expectations thus seems a good explanation of the aggregate price pattern observed in the experimental sessions with convergence.

*Trend-Following Heuristic.*—Especially for the subjects in the permanent and the dampened oscillating sessions of HSTV05, estimation revealed a *trend-following* forecasting rule of the form

$$(7) \quad p_{t+1}^e = p_{t-1} + \gamma(p_{t-1} - p_{t-2}),$$

where  $\gamma > 0$ . This rule has a simple behavioral interpretation: the forecast uses the last price observation and adjusts in the direction of the last price change. The extrapolation coefficient  $\gamma$  measures the strength of the adjustment. The estimates of this coefficient ranged from relatively small extrapolation values,  $\gamma = 0.4$ , to quite strong extrapolation values,  $\gamma = 1.3$ .

The price in the homogeneous model may either converge or diverge under the trend-following rule (7), depending upon the parameter  $\gamma$ . In the former case the trend-following rule is called weak (see Figure 4, panel B), whereas in the latter case it is referred to as strong (see Figure 4, panel C). For weak trend extrapolation, when the extrapolating coefficient is small (e.g.,  $\gamma = 0.4$ ), convergence is monotonic; for larger  $\gamma$ -values (e.g.,  $\gamma = 0.99$ ) convergence becomes oscillatory. For strong trend extrapolation the price dynamics diverges from the fundamental steady state through oscillations of increasing amplitude. The speed of divergence and amplitude of the long run fluctuations increase with  $\gamma$ , as shown by comparison of the cases  $\gamma = 1.1$  and  $\gamma = 1.3$ .

<sup>4</sup>The adaptive rule (6) was estimated for 5 out of 12 participants in sessions 2 and 5 of the HSTV05 experiment. Three participants had  $w$  insignificantly different from 1. Four other participants used an AR(1) rule,  $p_{t+1}^e = a + b p_{t-1}$ , conditioning only on the past price with a coefficient  $b < 1$ .

*Anchoring and Adjustment Heuristic.*—A number of participants, especially in the permanently oscillating sessions,<sup>5</sup> used slightly more sophisticated rules of the form

$$(8) \quad p_{t+1}^e = 0.5(p^f + p_{t-1}) + (p_{t-1} - p_{t-2}).$$

This is an example of an *anchoring and adjustment* (AA) rule (Tversky and Kahneman 1974), since it extrapolates the last price change from the reference point or *anchor*  $(p^f + p_{t-1})/2$  describing the “long-run” price level. One could argue that the anchor for this rule, defined as an equally weighted average between the last observed price and the fundamental price  $p^f$ , was unknown in the experiment, since subjects were not provided explicitly with the fundamental price. Therefore, in our evolutionary selection model in Section III, one of the rules will be a modification of (8) with the fundamental price  $p^f$  replaced by a proxy given by the (observable) sample average of past prices  $p_{t-1}^{av} = (1/t) \sum_{j=0}^{t-1} p_j$ , to obtain

$$(9) \quad p_{t+1}^e = 0.5(p_{t-1}^{av} + p_{t-1}) + (p_{t-1} - p_{t-2}).$$

We will refer to this forecasting rule with an anchor learned through a sample average of past prices, as the *learning anchoring and adjustment* (LAA) heuristic.

Applying results from online Appendix B to the anchoring and adjustment rule (8) we conclude that the price dynamics of the homogeneous expectations model with AA heuristic converges to the fundamental steady-state, but the convergence is oscillatory and slow, see Figure 4, panel D. For the stochastic simulation the convergence is even slower and the amplitude of the price fluctuations remains more or less constant in the last 20 periods, with an amplitude ranging from 55 to 65 comparable to that of the permanently oscillatory session 6 in the HSTV05 experiments. The small shocks  $\varepsilon_t$ , added in the experimental design to mimic (small) shocks in a real market, thus seem to be important to keep the price oscillations alive. Figure 4, panel D also shows the price dynamics of the learning anchoring and adjustment (LAA) rule (9). Simulations under homogeneous expectations given by the LAA heuristic (9) converge to the same fundamental steady-state as with the AA heuristic (8), but much slower and with less pronounced oscillations. In the presence of noise, the price fluctuations under the LAA heuristic are qualitatively similar to the fluctuations in the permanently oscillatory session 1 of the HSTV05 experiment (see, e.g., the lower left panel of Figure 5), with prices fluctuating below the fundamental most of the time.

#### A. Homogeneity versus Heterogeneity

Different homogeneous expectation models explain the three observed patterns in the HSTV05 experiments, monotonic convergence, constant oscillations, and dampened oscillations. However, a model with homogeneous expectations leaves

<sup>5</sup> For 4 out of 12 participants of sessions 1 and 6 of the HSTV05 the estimated rule was very close to (8).

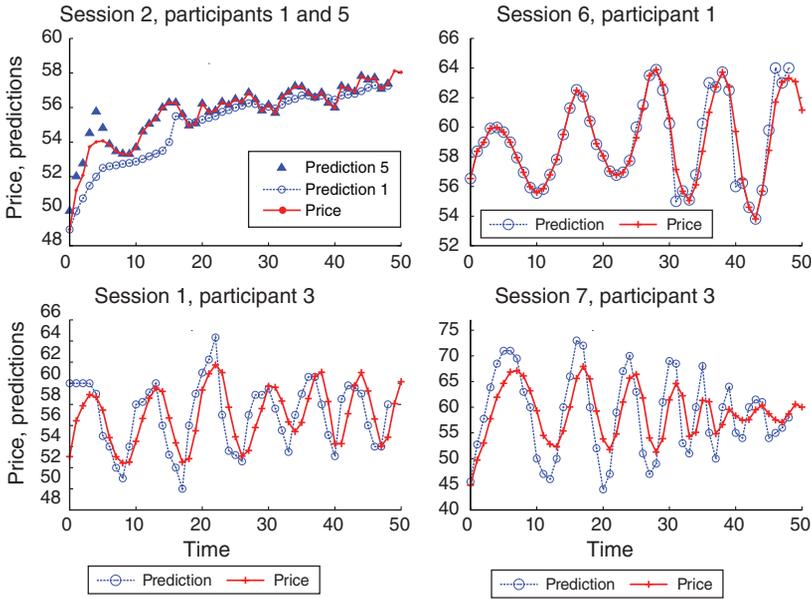


FIGURE 5. SWITCHING OF THE EXPERIMENT’S PARTICIPANTS BETWEEN SIMPLE RULES

Notes: For any point on the abscissa, representing time  $t$ , the price  $p_t$  and the forecast  $p_{i,t+2}^e$  are shown. This forecast,  $p_{i,t+2}^e$ , was made immediately after the announcement of the price  $p_t$ .

open the question *why* different patterns in aggregate behavior emerged in different experimental sessions.

The behavioral model proposed in this paper is based on the idea of *heterogeneity*, in the sense that several forecasting heuristics, those which we discussed above, could be used by the participants in every session. Participants learn which forecasting rule to use by switching between different heuristics based upon relative past performance. Individual experimental forecasting data shows some evidence that this type of learning is what happened in the experiment. Figure 5 shows time series of (lagged) individual forecasts from the HSTV05 experiment together with the realized price. The timing in the figure is important. For every time  $t$  on the horizontal axes, we show the price  $p_t$  together with the individual two-periods-ahead forecast  $p_{i,t+2}^e$  of that price by some participant  $i$ . In this way, we can infer graphically how the two-periods-ahead forecast  $p_{i,t+2}^e$  uses the last observed price  $p_t$ . For example, if they coincide, i.e.,  $p_{i,t+2}^e = p_t$ , this implies naive expectations in period  $t$ .

In session 2, subject 5 extrapolates price changes in the early stage of the experiment (see the upper left panel), but, starting from period  $t = 6$ , uses a simple *naive rule*  $p_{i,t+2}^e = p_t$ . In other words, in period 6, subject 5 switched from an extrapolative to a naive forecasting rule. Subject 1 from the same group used a “smoother,” *adaptive* forecasting strategy, always predicting a price between the previous forecast and the previous price realization. These graphs suggest individual heterogeneity in forecasting strategies in the same session.

In the oscillating group 6, subject 1 used naive expectations in the first half of the experiment (until period 24, see the upper right panel). Naive expectations however, lead to prediction errors in an oscillating market, especially when the trend reverses.

In period 25, subject 1 switches to a different, trend-extrapolating prediction strategy. Thereafter, this subject uses a trend-extrapolating strategy switching back to the naive rule at periods of expected trend reversal (e.g., in periods 27–28, 32–33, 37–38, 42–44, and 47).

Interestingly, participant 3 from another oscillating session 1 starts out predicting the fundamental price, i.e.,  $p_{t+1}^e = p^f = 60$ , in the first four periods of the experiments (see the lower left panel). But since the majority in session 1 predicted a lower price, the realized price is much lower than the fundamental, causing participant 3 to switch to a different, trend extrapolating strategy. Trend-extrapolating predictions overshoot the realized market price at the moments of trend reversal. Toward the end of the experiment, participant 3 learned to anticipate the trend changes to some extent.

Finally, in session 7 with dampened price oscillations, subject 3 started out with a strong trend extrapolation (see Figure 5, the lower right panel). Despite very large prediction errors (and thus low earnings) at the turning points, this participant sticks to strong trend extrapolation; only in the last four periods some kind of adaptive expectations strategy was used.

Figure 5, therefore, suggests that *individual heterogeneity in expectations* within the same session<sup>6</sup> plays a role in explaining the observed phenomena at the aggregate level.

### III. Heuristics Switching Model

In this section, we present a simple model with evolutionary selection between different forecasting heuristics. Before describing the model, we recall the most important “stylized facts” which we found in the individual and aggregate experimental data:

- (i) Participants tend to base their predictions on past observations following simple forecasting heuristics.
- (ii) Individual learning has a form of switching from one heuristic to another.
- (iii) In every session some form of coordination of individual forecasts occurs; the rule on which individuals coordinate may be different in different sessions.
- (iv) Coordination of individual forecasting rules is not perfect and some heterogeneity of the applied rules remains at every time period.

The main idea of the model is simple.<sup>7</sup> Assume that there exists a pool of simple prediction rules (e.g., adaptive or trend-following heuristics) commonly available to the participants of the experiment. At every time period these heuristics deliver forecasts for next period’s price, and the realized market price depends upon these

<sup>6</sup>Of course, high coordination of expectations means that the individual expectations were not spread over the entire space. Online Appendix C provides additional details (in terms of eigenvalues) of the variety of heuristics used in the same experimental session.

<sup>7</sup>Our model is built upon the Adaptive Belief Scheme proposed in Brock and Hommes (1997), but with memory in the fitness measure and asynchronous strategy updating.

forecasts. However, the *impacts* of different forecasting heuristics upon the realized prices are changing over time because the participants are learning based on evolutionary selection: the better a heuristic performed in the past, the higher its impact in determining next period's price. As a result, the realized market price and impact of the forecasting heuristics co-evolve in a dynamic process with mutual feedback. This nonlinear evolutionary model exhibits *path dependence* explaining coordination on different forecasting heuristics leading to different aggregate price behavior.

### A. The Model

Let  $\mathcal{H}$  denote a set of  $H$  heuristics which participants can use for price prediction. In the beginning of period  $t$ , every rule  $h \in \mathcal{H}$  gives a two-periods-ahead point prediction for the price  $p_{t+1}$ . The prediction is described by a deterministic function  $f_h$  of available information:

$$(10) \quad p_{h,t+1}^e = f_h(p_{t-1}, p_{t-2}, \dots; p_{h,t}^e, p_{h,t-1}^e, \dots).$$

The price in period  $t$  is computed on the basis of these predictions according to equation (1) in the environment without robot traders, or to equation (3) in the environment with robots. In the latter case, the fractions of robots is determined by (4) as before. The average forecast  $\bar{p}_{t+1}^e$  in the price equations becomes, in an evolutionary model, a population-weighted average of the different forecasting heuristics

$$(11) \quad \bar{p}_{t+1}^e = \sum_{h=1}^H n_{h,t} p_{h,t+1}^e,$$

with  $p_{h,t+1}^e$  defined in (10). The weight  $n_{h,t}$  assigned to the heuristic  $h$  is called the *impact* of this heuristic. The impact is evolving over time and depends on past relative performance of all  $H$  heuristics, with more successful heuristics attracting more followers.

Similar to the incentive structure in the experiment, the performance measure of a forecasting heuristic in a given period is based on its squared forecasting error. More precisely, the *performance measure* of heuristic  $h$  up to (and including) time  $t - 1$  is given by

$$(12) \quad U_{h,t-1} = -(p_{t-1} - p_{h,t-1}^e)^2 + \eta U_{h,t-2}.$$

The parameter  $0 \leq \eta \leq 1$  represents the *memory*, measuring the relative weight agents give to past errors of heuristic  $h$ . In the special case  $\eta = 0$ , the impact of each heuristic is completely determined by the most recent forecasting error; for  $0 < \eta \leq 1$  all past prediction errors, with exponentially declining weights, affect the impact of the heuristics.

Given the performance measure, the impact of rule  $h$  is updated according to a *discrete choice model with asynchronous updating*

$$(13) \quad n_{h,t} = \delta n_{h,t-1} + (1 - \delta) \frac{\exp(\beta U_{h,t-1})}{Z_{t-1}},$$

where  $Z_{t-1} = \sum_{h=1}^H \exp(\beta U_{h,t-1})$  is a normalization factor. In the special case  $\delta = 0$ , (13) reduces to the *the discrete choice model* with synchronous updating used in Brock and Hommes (1997) to describe endogenous selection of expectations. The more general case,  $0 \leq \delta \leq 1$ , gives some persistence or inertia in the impact of rule  $h$ , reflecting the fact (consistent with the experimental data) that not all the participants update their rule in every period or at the same time (see Hommes, Huang, and Wang (2005) and Diks and van der Weide (2005)). Hence,  $\delta$  may be interpreted as the average per period fraction of individuals who stick to their previous strategy. In the extreme case  $\delta = 1$ , the initial impacts of the rules never change, no matter what their past performance was. If  $0 < \delta \leq 1$ , in each period a fraction  $1 - \delta$  of participants update their rule according to the discrete choice model. The parameter  $\beta \geq 0$  represents the intensity of choice measuring how sensitive individuals are to differences in strategy performance. The higher the intensity of choice  $\beta$ , the faster individuals will switch to more successful rules. In the extreme case  $\beta = 0$ , the impacts in (13) move to an equal distribution independent of their past performance. At the other extreme  $\beta = \infty$ , all agents who update their heuristic (i.e., a fraction  $1 - \delta$ ) switch to the most successful predictor.

*Initialization.*—The model is initialized by a sequence of initial prices, whose length  $\tau$  is long enough to allow any forecasting rule in  $\mathcal{H}$  to generate its prediction, as well as an initial distribution  $\{n_{h,0}\}$ ,  $1 \leq h \leq H$  of the impacts of different heuristics (summing to 1). Furthermore, the performance measures  $U_{h,t} = 0$  for all  $h$  and for  $t < \tau + 2$ , which is the first period when they are updated.

Given initial prices, the heuristic's forecasts for price at time  $\tau + 2$  can be computed and, using the initial impacts of the heuristics, the price  $p_{\tau+1}$  is determined. In the next period, the forecasts of the heuristics are updated, the fraction of robot traders is computed if necessary, while the same initial impacts  $n_{h,0}$  for the individual rules are used, since past performance is not well defined yet. Thereafter, the price  $p_{\tau+2}$  is computed and the initialization stage is finished. After this initialization stage the evolution of the model is well defined: first the performance measure in (12) is updated, then, the new impacts of the heuristics are computed according to (13), and the new prediction of the heuristics are obtained according to (10). Finally, the new average forecast (11) and (if necessary) the new fraction of robot traders (4) are computed, and a new price is determined by (1) or (3), respectively.

*Example with Four Heuristics.*—The evolutionary model can be simulated with an arbitrary set of heuristics. Since one of our goals is to explain the three different observed patterns in aggregate price behavior—monotonic convergence, permanent oscillations, and dampened oscillations—we keep the number of heuristics as small as possible and consider a model with only four forecasting rules. These rules, referred to as ADA, WTR, STR, and LAA and given in Table 1, were obtained as simple descriptions of typical individual forecasting behavior observed and estimated in the experiments.<sup>8</sup> As discussed in Section II, these heuristics generate qualitatively

<sup>8</sup>Each of these rules was among those estimated on individual data in HSTV05. In this paper, we combine these fixed rules in the Heuristic Switching Model and vary only learning parameters  $\beta$ ,  $\eta$ , and  $\delta$ . One could improve the

TABLE 1—HEURISTICS USED IN THE EVOLUTIONARY MODEL

ADA	Adaptive heuristic	$p_{t,t+1}^f = 0.65 p_{t-1} + 0.35 p_{t,t}^f$
WTR	Weak trend-following rule	$p_{t,t+1}^s = p_{t-1} + 0.4 (p_{t-1} - p_{t-2})$
STR	Strong trend-following rule	$p_{t,t+1}^s = p_{t-1} + 1.3 (p_{t-1} - p_{t-2})$
LAA	Anchoring and adjustment rule with learning anchor	$p_{t,t+1}^s = 0.5 (p_{t-1}^{av} + p_{t-1}) + (p_{t-1} - p_{t-2})$
AA	Anchoring and adjustment rule with fixed anchor	$p_{t,t+1}^s = 0.5 (p^f + p_{t-1}) + (p_{t-1} - p_{t-2})$

Notes: In simulations in Figures 6–8, the first four heuristics are used. The LAA heuristic is obtained from the simpler AA heuristic, by replacing the (unknown) fundamental price  $p^f$  by the sample average  $p_{t-1}^{av}$ .

different dynamics, which allows one to get some non-trivial interaction between different heuristics, so that qualitatively different patterns can be obtained.

With the four forecasting heuristics fixed, matched by experimental individual forecasting data, there are only three free “learning” parameters in the model:  $\beta$ ,  $\eta$ , and  $\delta$ . Provided that these parameters are given, the Heuristic Switching Model (HSM) with four heuristics is initialized with two initial prices,  $p_0$  and  $p_1$ , and four initial impacts  $n_{h,0}$  used in periods  $t = 2$  and  $t = 3$ .

#### IV. Empirical Validation

The Heuristic Switching Model exhibits *path-dependence* and is capable of reproducing different qualitative patterns within the same experimental environment. Anufriev and Hommes (2012) simulate 50-period-ahead forecasts (so-called “simulated paths”) of the HSM in three sessions of the HSTV05 experiment. For a specific choice of the parameters,  $\beta = 0.4$ ,  $\eta = 0.7$ , and  $\delta = 0.9$  (obtained after some trial and error simulations), but with different initial prices and impacts of heuristics, the HSM *qualitatively* reproduces price behavior in sessions 2, 1, and 7, i.e., monotonic convergence, permanent oscillations, and dampened oscillations.<sup>9</sup> In those 50-period-ahead simulations, however, the frequency and amplitude of the oscillations were difficult to match.

In this section, we investigate the forecasting performance of the HSM model *quantitatively* both in-sample and out-of-sample. We fit the model to 20 experimental sessions of 4 different treatments in HSTV05 and HSTV08, providing an important generality test of the model. The HSM model fits all 20 sessions quite well. The section is divided in three parts. We, first, illustrate the one-period-ahead forecasts of the model visually. Then we proceed to the rigorous evaluation of the in-sample performance of the model, and, finally, look at out-of-sample forecasts made by the model.

fit of the model even further if estimation would include both learning parameters and behavioral rules.

<sup>9</sup>The initial distribution of agents over the four heuristics, i.e., initial impacts  $\{n_{1,0}, n_{2,0}, n_{3,0}, n_{4,0}\}$ , played a crucial role for these price patterns. When the initial impacts of heuristics are distributed almost uniformly, the monotonic convergence of session 2 is reproduced. Oscillations of session 1 are obtained when both the WTR and STR heuristics have relatively high initial weights (35 percent each). The dampened price oscillation in session 7 are reproduced when the STR rule has a large initial impact (66 percent).

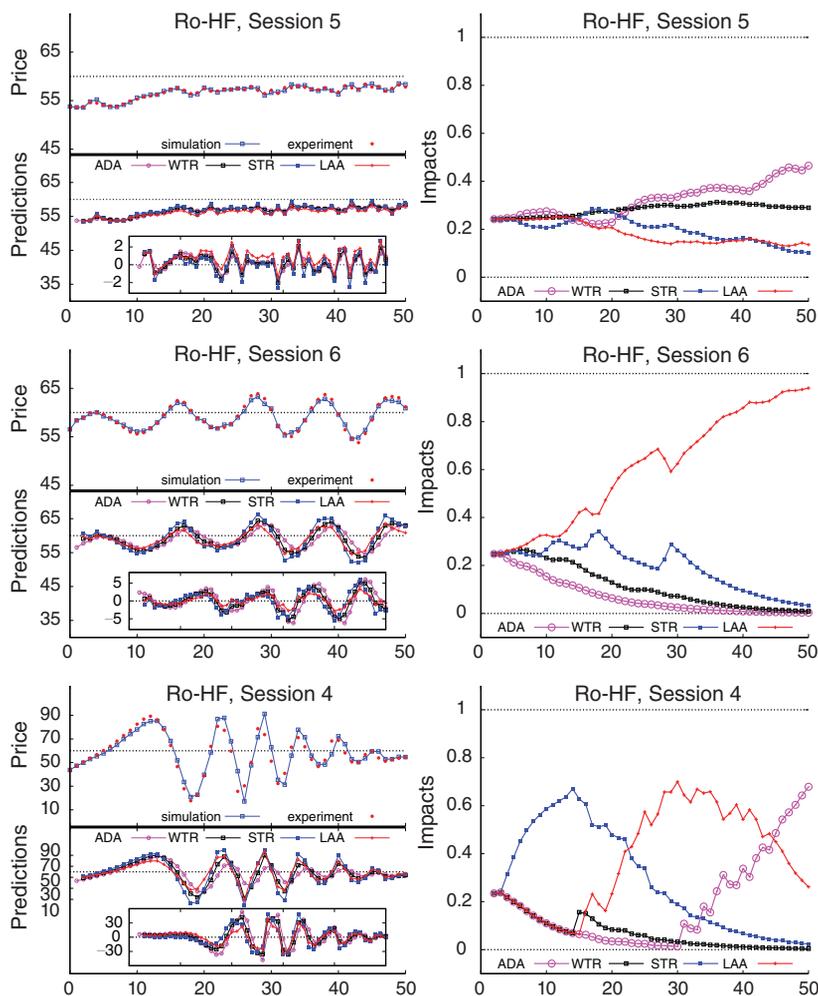


FIGURE 6. LABORATORY EXPERIMENTS AND THE HSM IN THREE SESSIONS FROM **Ro-HF** HSTV05 EXPERIMENT

*Notes:* Three sessions from different experiments are shown. The upper parts of left panels show prices for experiments with corresponding one-step-ahead predictions of the HSM. The lower parts show predictions and forecasting errors (inner frames) of the four heuristics. The right panels show the evolution of the impacts of the four heuristics.

### A. One-Period-Ahead Simulations

The left panels in Figure 6 suggest that the switching model with four heuristics fits the experimental data from the **Ro-HF** HSTV05 experiments quite nicely. The upper parts of the panels compare the experimental data with the one-step-ahead predictions of the HSM for the benchmark parameter values  $\beta = 0.4$ ,  $\eta = 0.7$ , and  $\delta = 0.9$ . We stress that, at this stage, no fitting exercise has been performed. In each session and every period experimental data (dots) are quite close to the prediction given by the HSM (line with squares).

In all these simulations, the initial prices are chosen to coincide with the initial prices in the first two periods in the corresponding experimental group, while the

initial impacts of all heuristics are equal to 0.25. At each step, in order to compute the heuristics' forecasts and update their impacts, past *experimental price data* are used, which is exactly the same information that was available to participants in the experiments. The one-period-ahead forecasts can follow easily the monotonically converging patterns as well as the sustained or dampened oscillatory patterns, starting out from a uniform initial distribution of forecasting rules.

The lower parts of the left panels in Figure 6 show that the heuristics forecasts are correlated, so that the model reproduces coordination of expectations of the participants. The frames of the lower panels display the prediction errors of the four heuristics. These errors are of the same order as in the experiment (see Figure 3). Behavioral heuristics applied to non-equilibrium dynamics do not eliminate the forecasting errors in line with experimental evidence.

The right panels of Figure 6 show the transition paths of the impacts of each of the four forecasting heuristics. In the case of monotonic convergence (session 5, upper panel), the impacts of all four heuristics remain relatively close, although the impact of adaptive expectations gradually increases and slightly dominates the other rules in the last 25 periods. In contrast to the previous case, in experimental session 6 the two initial prices exhibited an increasing trend. Consequently the trend-following rules and the learning anchor and adjustment heuristic increase their impacts (middle panels). However, the trend rule misses the turning points and its impact gradually decreases. The LAA heuristic with a more flexible anchor predicts price oscillations better than the static STR and WTR rules. The impact of the LAA heuristic gradually increases, rising to more than 80 percent after 40 periods. The HSM thus explains coordination of individual forecasts on a LAA rule enforcing persistent price oscillations around the long run equilibrium level.

Finally, in the simulations for session 4 with dampened oscillations (lower panels), the initial price trend is so strong that the STR rule clearly dominates during the first 20 periods of simulation. To some extent this rule reinforces a trend in prices, but the presence of other heuristics and the impact of the robot traders lead to a trend reversal around period 12. The STR misses this turning point and the LAA rule starts to increase its share, eventually overcoming the STR rule. The mixture of the LAA and STR rules produces time series which gives some space to other rules as well, and the price oscillations slowly dampen. As a result, the ADA rule, which was the worst until period 30, starts to increase its impact. Eventually it overcomes the LAA rule explaining price stabilization.

The dynamics in sessions 5, 6, and 7 of the **Ro-HF** HSTV05 experiment were qualitatively similar to the dynamics in sessions 2, 1, and 4, respectively. It is not surprising, therefore, that the fit by the HSM and the dynamics of impacts is similar, see online Appendix D. Figure 7 illustrates the fit of the HSM for the remaining session 3 of the **Ro-HF** experiment. The price in this session started with moderate oscillations, then stabilized at a level below the fundamental price 60, but suddenly fell in period  $t = 41$  likely due to a typing error by one of the participants.<sup>10</sup> Our nonlinear switching model, of course, misses the typing error, but nevertheless matches the overall pattern before and after the unexpected price drop.

<sup>10</sup>In fact, one of the participants predicted 5.25 for period 42, whereas his/her five previous predictions all were between 55.00 and 55.40. Perhaps the intention was to type 55.25.

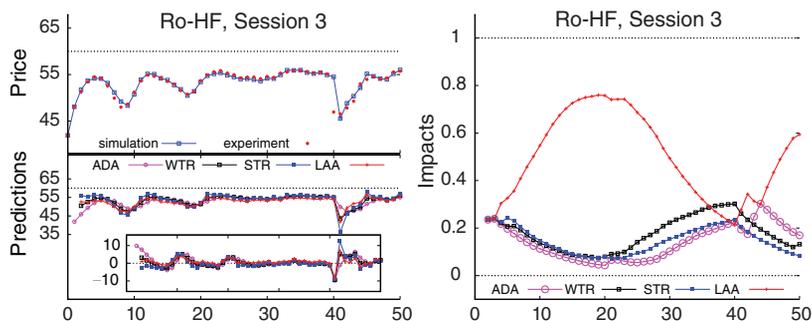


FIGURE 7. LABORATORY EXPERIMENTS AND THE HSM IN SESSION 3 OF THE **Ro-HF** HSTV05 EXPERIMENT

*Notes:* The upper part of the left panel shows prices for a laboratory experiment with corresponding one-step-ahead predictions of the HSM. The lower part shows predictions and forecasting errors (inner frames) of the four heuristics. The right panel shows the evolution of the impacts of the four heuristics.

We demonstrated that the HSM fits all sessions of the **Ro-HF** environment quite nicely. How robust are these results? Figure 8 shows the fit and the evolution of impacts for typical sessions from the three other treatments with modified experimental environments. The upper panels displays session 8 of the **Ro-LF** HSTV05 experiment with a fundamental price of 40 instead of 60. Slowly increasing price oscillations are dominated by the STR rule, but when the frequency of trend-reversals increases, the LAA rule takes over again. The middle panel illustrates the HSM for session 12 of the **NoRo** HSTV05 experiment without robot traders. Large amplitude price oscillations, with prices almost reaching their maximum 100 and minimum 0, arise, but prices stabilize toward the end of the experiment. The evolution of the impacts of the four forecasting heuristics is similar to the case of dampening price oscillations considered before (see the lower panels of Figure 6), with the STR rule initially dominating the market, followed by the LAA rule taking over around period 25 during the price oscillations, and the ADA rule finally dominating toward the end, when the price stabilizes.

Finally, the lower panel displays session 2 of the **LFR** HSTV08 experiment with large forecasting range. The asset price oscillates with very large amplitude, with a long lasting trend of more than 25 periods with prices rising close to their maximum 1,000, after which a crash follows to values close to their minimum 0, and the market starts oscillating. The evolution of the impacts of the 4 forecasting heuristics is similar as before, with the STR rule taking the lead, increasing its share to more than 90 percent after 25 periods during the long-lasting price bubble. When the upper forecasting bound is met, the trend cannot be sustained any longer and the share of the STR rule declines quickly. The LAA rule takes over and large amplitude oscillations are observed. The fact that our nonlinear switching model nicely captures all patterns in these different groups shows that the model is robust concerning changes of the asset pricing market environment.<sup>11</sup>

<sup>11</sup> Online Appendix D contains analogous plots for all other groups of the HSTV05 and HSTV08 experiments.

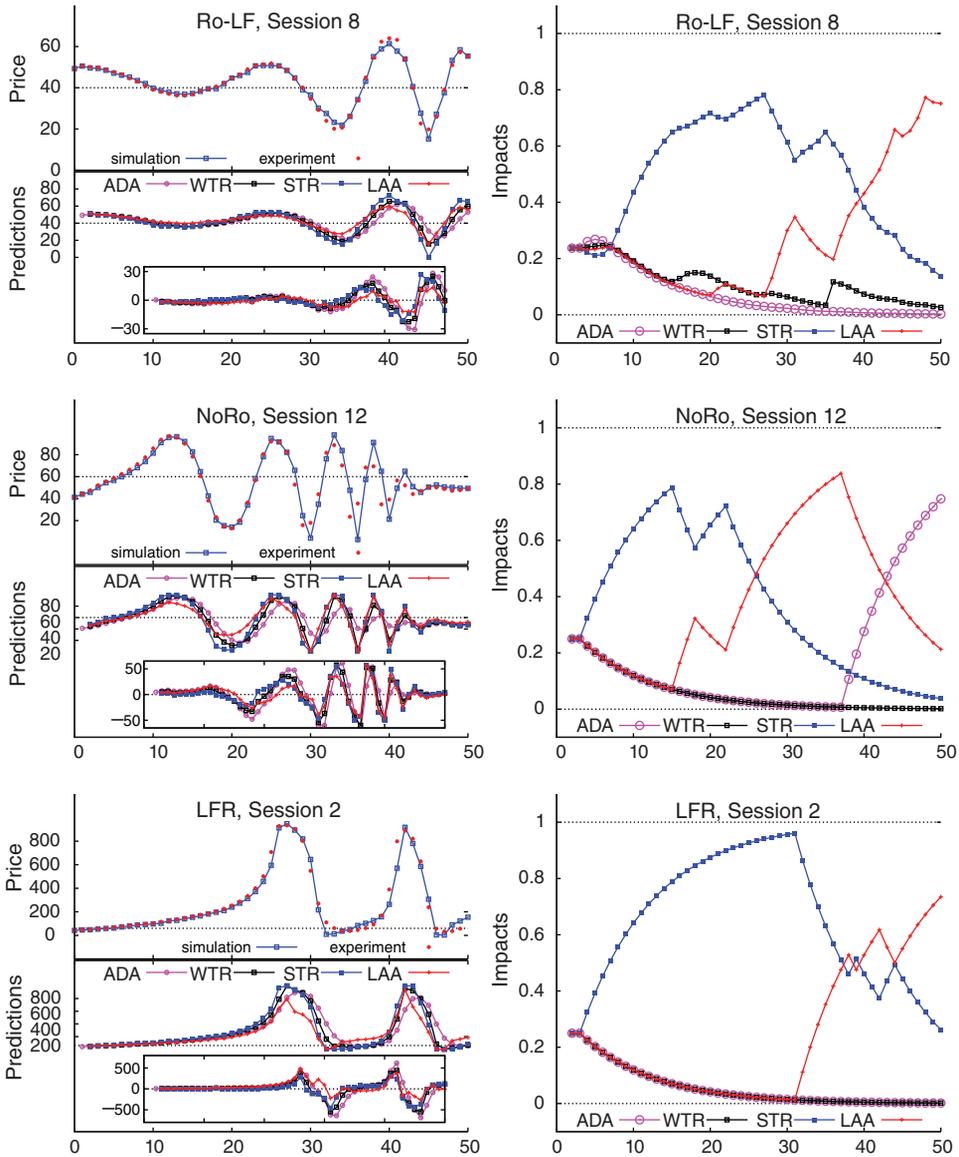


FIGURE 8. LABORATORY EXPERIMENTS AND THE HSM IN SESSIONS OF THREE DIFFERENT EXPERIMENTS

Notes: Three sessions from different experiments are shown: **Ro-LF** HSTV05 (top panels), **NoRo** HSTV05 (middle panels), and **LFR** HSTV08 (bottom panels). The upper parts of left panels show prices for experiments with corresponding one-step-ahead predictions of the HSM. The lower parts show predictions and forecasting errors (inner frames) of the four heuristics. The right panels show the evolution of the impacts of the four heuristics.

To summarize, the one-step-ahead simulations of the HSM match both converging and oscillating prices closely and produce clear differences in the evolution of impacts. In the case of monotonic convergence and constant oscillations we observe a self-confirming property: more coordination on one rule (ADA in the case of convergence and LAA in the case of oscillations) leads to price dynamics which reinforces this rule. For the sessions with the dampened oscillations, one step

ahead forecast produces a rich evolutionary selection dynamics. These sessions go through three different phases where the STR, the LAA, and the ADA heuristics subsequently dominate. The STR dominates during the initial phase of a strong trend in prices, but starts declining after it misses the first turning point of the trend. The LAA does a better job in predicting the trend reversal, but if it fails to keep the amplitude of oscillations constant it may be taken over by a stabilizing ADA rule.

### B. Forecasting Performance

A measure of model fit are the mean squared errors of simulated one-step ahead predictions. Table 2 compares these mean squared errors (MSEs) for several experimental sessions<sup>12</sup> for nine different models: the RE fundamental prediction, five homogeneous expectations models (naive expectations, and each of the four heuristics of the switching model), and three heterogeneous expectations models with four heuristics, namely, the model with fixed fractions (corresponding to  $\delta = 1$ ), the switching model with benchmark parameters  $\beta = 0.4$ ,  $\eta = 0.7$ , and  $\delta = 0.9$ , and, finally, the “best” switching model fitted by minimization of the MSE over the parameter space (the last three lines in Table 2 show the corresponding optimal parameter values).<sup>13</sup> The MSEs for the benchmark switching model are shown in bold and, for comparison, for each session the MSEs for the best among the four heuristics are also shown in bold. The best among the nine models for each session is shown in italic and it is always the best fitted HSM.

An immediate observation from Table 2 is that, for all sessions, the fundamental prediction rule is by far the worst. This is due to the fact that in all experiments realized prices deviate persistently from the fundamental benchmark. Another observation is that, all models explain the monotonically converging groups very well, with very low MSE. The homogeneous expectations models with naive, adaptive or WTR expectations fit the monotonic converging group 5 particularly well. However, only adaptive expectations outperform the benchmark switching model. For the permanent as well as the dampened oscillatory sessions 6 and 8, the flexible LAA rule is the best homogeneous expectations benchmark, but the benchmark switching model has an even smaller MSE. In session 8 from the **Ro-LF** HSTV05 experiment with low fundamental price, the WTR heuristic is the best among homogeneous rules, but the HSM also outperforms this model. In the sessions without robots, as in the **NoRo** HSTV05 treatment and especially in the **LFR** HSTV08 treatment, all homogeneous models produce very large MSEs. The models often under- or overestimate the trend and also miss the turning points of trend-reversal. The MSEs of the switching model are much smaller in the sessions of the **NoRo** HSTV05 treatment and comparable with the best homogeneous model (STR heuristic) in the **LFR**

<sup>12</sup>The MSEs are computed over 47 periods for  $t = 4, \dots, 50$  in the sessions of HSTV05 experiment and over 46 periods, for  $t = 4, \dots, 49$ , in the sessions of HSTV08 experiments. We skip the errors of the first four periods in order to minimize the impact of the initial conditions (i.e., the initial impacts of the heuristics) for the HSM. Period  $t = 4$  is the first period when the prediction is computed with both the heuristics' forecasts and the heuristics' impacts being updated based on experimental data. For comparison, in all other models we compute errors also from  $t = 4$ .

<sup>13</sup>Analogous tables for the remaining sessions can be found in online Appendix D.

HSTV08 treatment. When the learning parameters of the HSM model are chosen to minimize the MSE for the corresponding session, the model produces the smallest MSE over all alternative specifications,<sup>14</sup> though an improvement of fit with respect to the benchmark HSM is not as large as when the benchmark HSM is compared with the best homogeneous model. This suggests that the good fit of the switching model is fairly robust w.r.t. the model parameters.

The values of the parameters in Table 2 show that either  $\eta$  or  $\delta$  should be high in order to explain the experimental data. High values of  $\delta$  imply that inertia in individual choices played a big role in the experiment. In other words, subjects did not exhibit an immediate reaction to the difference in relative performances of the rules, but only switched occasionally. High values of  $\eta$  also generate persistence but through another channel. When memory  $\eta$  is high it means that the performances of different rules are slowly updated and the rules which used to be good in the past are not discarded immediately. It might not be easy for a subject to remember past performances, and lower values of  $\eta$  may reflect this cost of memory. Finally, the intensity of choice,  $\beta$ , reflects how fast the switching reaction is on differences in performances of different rules. In the benchmark case,  $\beta$  is quite low but in some sessions (5, 6, and 8) the fit improves quite a lot when higher values of  $\beta$  are used. It seems that in these experimental sessions subjects were able to coordinate on the common forecasting rules faster than our stylized model predicts in the benchmark case.

We stress that the evolutionary Heuristic Switching Model is able to make the best out of different heuristics. Indeed, in different sessions (and treatments) of the learning to forecast experiment *different* homogeneous expectations models produce the best fit of the price dynamics. However, the switching model produces an even better fit than the best homogeneous model.

### C. Out-of-Sample Forecasting

Let us now turn to the out-of-sample validation of the model. In order to evaluate the out-of-sample forecasting performance of the model, we first perform a grid search to find the parameters of the model minimizing the MSE for periods  $t = 4, \dots, 43$  and then compute the forecasting errors of the fitted model for the remaining periods (7 in the case of the HSTV05 experiments and 6 in the case of the HSTV08 experiments). The results are reported in the upper part of Table 3 for each session. The first two lines show the in-sample MSE of the fitted model and the corresponding values of the parameters. The next lines give the values of the forecasting errors. In the middle part of the table, we report the results of the same procedure performed for the switching model with benchmark parameters. Finally, we compare our structural behavioral learning model with a simple non-structural model with three parameters. To this purpose we estimate an AR(2) model to the data up to period  $t = 43$  and show in the bottom part of Table 3 the in-sample MSE and out-of-sample prediction errors.

<sup>14</sup>This holds for all 20 sessions of HSTV05 and HSTV08 experiments, except session 5 in the LFR HSTV08 treatment, where the STR heuristic outperforms the HSM. See online Appendix D.

TABLE 2—MSE OF THE ONE-STEP AHEAD FORECAST FOR 7 SESSIONS OF DIFFERENT LTFES

Specification	Ro-HF				Ro-LF	NoRo	LFR <sup>a</sup>
	5	6	4	3	8	12	2
Fundamental prediction	10.82	9.32	300.99	49.54	129.31	571.59	138,314.09
Naive heuristic	0.05	2.45	141.06	2.38	31.82	257.72	15,896.48
ADA heuristic	<b>0.04</b>	4.61	210.33	3.59	59.28	358.26	28,323.14
WTR heuristic	0.14	1.13	92.22	<b>1.84</b>	15.48	203.06	8,995.96
STR heuristic	0.66	0.81	90.59	3.40	<b>10.27</b>	182.35	<b>2,909.13</b>
LAA heuristic	0.48	0.66	66.26	1.92	16.84	<b>159.24</b>	16,065.67
4 heuristics ( $\delta = 1$ )	0.17	0.66	67.66	1.55	11.08	141.08	7,459.75
4 heuristics (Figs. 6–7)	<b>0.11</b>	<b>0.29</b>	<b>41.34</b>	<b>1.48</b>	<b>3.21</b>	<b>85.51</b>	<b>2,931.13</b>
4 heuristics (best fit)	<i>0.03</i>	<i>0.16</i>	<i>35.35</i>	<i>1.46</i>	2.23	<i>76.44</i>	<i>2,815.42</i>
$\beta \in [0, 10]$	10.00	10.00	0.13	0.39	10.00	0.58	0.92
$\eta \in [0, 1]$	0.95	0.05	0.92	0.49	0.96	0.88	0.70
$\delta \in [0, 1]$	0.28	0.76	0.11	0.87	0.82	0.45	0.93

Note: Nine different model specifications are compared.

<sup>a</sup>For the LFR HSTV08 experiment MSE is computed over 46 periods.

For the converging groups of the **Ro-HF** treatment (e.g., session 5) the prediction errors typically increase with horizon but remain very low and comparable with the errors computed in-sample. This is not surprising given that the qualitative property of the data (i.e., monotonic convergence) does not change in the last periods, and that the adaptive heuristic, which generates such convergence, takes the lead already around period 25. In the oscillating groups (session 6 of the **Ro-HF** and session 8 of the **Ro-LF**), the out-of-sample errors generated by the switching model are varying with the time horizon and typically larger than the in-sample error. Even if the switching model with leading LAA heuristic captures oscillations qualitatively, the prediction errors can become large when the oscillations predicted by the model have a different frequency than the oscillations in the experimental session, so that the prediction goes out of phase. The prediction errors for sessions with damping oscillations (session 4 of the **Ro-HF** and session 12 of the **NoRo**) are not very high when compared with the in-sample errors. This is because the price in these sessions converges toward the end of the experiment. At the end of the in-sample period, i.e., at  $t = 43$ , the switching model already selects the ADA heuristic, which generates the same behavior as in the experiment.

Comparing the forecasting errors of the switching and AR(2) models, we conclude that the former model is better than the latter, on average. More specifically, in the converging session 5 and oscillating session 6 of the **Ro-HF** treatment the out-of-sample performances are very similar, but in all other sessions (including those which are not shown in Table 3) the different variations of the switching model outperform the AR(2) model out-of-sample. The nonlinearity and behavioral foundation allows the HSM to adapt faster than the AR(2) model to changes in the price dynamics. In session 8 of the **Ro-LF**, the oscillations become more pronounced toward the end of the experiment. The HSM model captures this effect, while AR(2) does not. As a result, the AR(2) model produces errors larger than 7 in absolute value, which would lead to 0 earnings in the experiment, see equation (2), whereas the HSM does not have such large errors.

TABLE 3—OUT-OF-SAMPLE PERFORMANCE OF THE HEURISTIC SWITCHING MODEL AND THE AR(2) MODEL

Specification	Ro-HF				Ro-LF	NoRo	LFR
	5	6	4	3	8	12	2
<b>Best fit switching model</b>							
MSE <sub>40</sub>	0.03	0.17	39.10	1.66	1.19	89.19	2,856.01
( $\beta, \eta, \delta$ )	(10.0, 1.0, 0.3)	(10.0, 0.0, 0.7)	(0.1, 0.9, 0.1)	(0.3, 0.5, 0.9)	(0.8, 1.0, 0.9)	(0.1, 0.9, 0.4)	(0.5, 0.7, 0.9)
1 p ahead	-0.06	0.16	-2.26	-0.92	3.57	2.02	-30.69
2 p ahead	0.15	-0.44	1.18	-0.00	-1.88	-0.95	22.76
3 p ahead	0.31	-1.76	11.05	0.53	-6.53	-0.09	-49.30
4 p ahead	0.50	-2.73	13.58	0.69	-10.38	1.24	-28.81
5 p ahead	0.44	-3.17	8.85	0.23	-8.03	3.92	45.79
6 p ahead	0.50	-2.68	3.53	-0.56	0.33	4.47	163.32
7 p ahead	0.72	-0.91	-0.41	-0.59	8.95	3.36	NA
<b>Benchmark switching model</b>							
MSE <sub>40</sub>	0.10	0.28	47.42	1.68	2.34	99.65	2,939.01
( $\beta, \eta, \delta$ )	(0.4, 0.7, 0.9)	(0.4, 0.7, 0.9)	(0.4, 0.7, 0.9)	(0.4, 0.7, 0.9)	(0.4, 0.7, 0.9)	(0.4, 0.7, 0.9)	(0.4, 0.7, 0.9)
1 p ahead	-0.29	0.72	-3.15	-0.82	5.01	-0.16	-43.61
2 p ahead	0.31	1.48	-3.10	0.20	2.35	0.57	-8.75
3 p ahead	0.31	1.15	3.45	0.78	0.10	3.08	-52.77
4 p ahead	0.45	-0.09	5.82	0.93	-3.43	4.86	-29.69
5 p ahead	-0.05	-2.07	4.82	0.41	-3.25	7.27	57.36
6 p ahead	0.13	-3.76	4.61	-0.46	-0.12	7.95	189.63
7 p ahead	0.74	-3.75	4.07	-0.60	2.81	6.97	NA
<b>AR(2) model</b>							
MSE <sub>44</sub>	0.23	0.27	65.34	1.94	1.56	159.52	4,638.42
( $\beta, \eta, \delta$ )	(6.8, 0.8, 0.1)	(18.9, 1.6, -1.0)	(26.4, 1.3, -0.7)	(14.8, 1.2, -0.4)	(8.0, 1.9, -1.1)	(22.7, 1.2, -0.6)	(42.4, 1.7, -0.8)
1 p ahead	-1.00	-1.02	-3.65	-2.05	2.16	-2.61	60.30
2 p ahead	-0.44	-1.63	-1.49	-1.69	-8.18	-0.47	282.72
3 p ahead	-0.48	-2.18	7.63	-1.52	-22.84	6.07	292.09
4 p ahead	0.40	-1.15	11.65	-0.64	-38.05	12.20	164.78
5 p ahead	-0.18	-0.16	9.25	-0.82	-44.28	15.46	49.54
6 p ahead	-0.87	-0.09	5.60	-2.22	-37.64	13.91	-34.08
7 p ahead	-0.47	0.34	2.57	-2.54	-17.49	10.44	NA

The largest out-of-sample errors are observed in the sessions of the **LFR** treatment, see also Table 3 from Appendix D. However, in all the sessions the out-of-sample errors of the HSM are comparable and often even smaller than the in-sample errors. Furthermore, the errors of the HSM are smaller than the analogous prediction errors of the AR(2) model.

To further investigate one- and two-periods-ahead, out-of-sample forecasting performance, we fit the model on a moving sample and compute an average of the corresponding squared prediction errors. The results of this exercise for the same three models (best fit, benchmark, and AR(2)) are presented in Table 4. The smallest forecasting errors within the same class (e.g., one-period ahead in session 4) are shown in *italic*. It turns out that our nonlinear switching model predicts the experimental data better than the simple AR(2) model for both one- and two-periods-ahead predictions in almost all sessions.<sup>15</sup> In many cases the difference in performance is substantial.

<sup>15</sup>The AR(2) model outperforms both the benchmark and the best-fitted HSMs in session 7 of the **Ro-HF** treatment (one period ahead), session 14 of the **NoRo** treatment (two periods ahead), session 3 of the **LFR** treatment (one period ahead) and session 5 of the **LFR** treatment (both one and two periods ahead).

TABLE 4—OUT-OF-SAMPLE PERFORMANCE OF THE HEURISTIC SWITCHING MODEL AND THE AR(2) MODEL

Specification	Ro-HF				Ro-LF	NoRo	LFR
	5	6	4	3	8	12	2
Best fit switching model							
Average MSE <sub>40</sub>	0.04	0.17	45.19	1.75	1.90	84.75	2,507.10
1 p ahead	0.03	0.15	5.75	0.38	15.00	4.24	2,258.12
2 p ahead	0.09	0.92	12.17	0.40	26.87	6.06	4,727.68
Benchmark switching model							
Average MSE <sub>40</sub>	0.10	0.29	50.74	1.78	3.24	103.53	2,548.28
1 p ahead	0.15	0.41	7.69	0.42	9.82	5.74	2,291.63
2 p ahead	0.23	1.90	16.22	0.56	9.51	17.72	5,322.47
AR(2) model							
average MSE <sub>44</sub>	0.24	0.25	65.75	1.99	3.66	159.92	4,362.90
1 p ahead	0.44	0.56	13.46	1.43	35.02	13.93	11,880.66
2 p ahead	0.40	1.49	44.65	2.33	203.34	63.11	48,065.41

Note: The averaged (over moving window of 40 periods) in-sample MSE and averaged (over 7 forecasts) 1- and 2-periods ahead squared forecasting errors are shown.

## V. Conclusion and Discussion

The time evolution of aggregate economic variables, such as stock prices, is affected by market expectations of individual agents. Neoclassical economic theory assumes that individuals form expectations rationally, thus enforcing prices to track economic fundamentals and leading to an efficient allocation of resources. Laboratory experiments with human subjects have shown however that individuals do not behave fully rational, but instead follow simple heuristics. In laboratory markets prices may show persistent deviations from fundamentals similar to the large swings observed in financial stock prices.

Our results show that performance-based evolutionary selection among simple forecasting heuristics can explain coordination of individual forecasting behavior leading to three different aggregate outcomes observed in identically organized treatments of recent laboratory market forecasting experiments: slow monotonic price convergence, oscillatory dampened price fluctuations, and persistent price oscillations. Furthermore, we show that the same model also fits other learning-to-forecast experiments reasonably well.

In our Heuristic Switching Model, forecasting strategies are selected every period from a small population of plausible heuristics, such as adaptive expectations and trend-following rules. Individuals adapt their strategies over time, based on the relative forecasting performance of the heuristics. As a result, the nonlinear evolutionary switching mechanism exhibits *path dependence* and matches individual forecasting behavior as well as aggregate market outcomes in the experiments. We showed that none of the homogeneous expectation models constituting the HSM fit *all* different observed patterns. However, different heuristics fit the different patterns observed in different sessions. The HSM not only provides a simultaneous explanation for the aggregate outcomes in all different sessions, but also outperforms the best homogeneous model in almost every session. Our results are in line with recent work on agent-based models of interaction, and contribute to a behavioral explanation of market fluctuations.

The HSM is, to our best knowledge, the first learning model explaining *different* time series patterns in the *same* laboratory environment. Our approach is similar to many game-theoretical learning models (often referred as “reinforcement learning”), see, e.g., Arthur (1991), Erev and Roth (1998), and Camerer and Ho (1999). These models were developed as a response to experimental evidence that individuals often do not play or learn to play equilibrium predictions such as Nash equilibrium, or its refinements. Our market environment is quite different from the strategic environments studied in game theory, however. For example, agents in our framework do not have well-defined strategies, other than predicting a number between 0 and 100 in two decimals. Other important differences are that the strategies used in our switching model are state dependent and that individuals do not know the “payoff matrix,” which moreover is changing over time in a path-dependent manner.

The model presented in this paper is one of the first models *simultaneously* explaining *individual micro behavior* as well as *aggregate macro behavior*. Recently, Fuster, Laibson, and Mendel (2010) proposed a somewhat similar model of “natural expectations,” which are the weighted average of a so-called intuitive, extrapolative expectations and rational expectations. Their model can successfully explain a number of stylized facts of macroeconomic time series. Our heuristics correspond to different plausible parametrization of their intuitive expectations, but our approach to discipline-bounded rationality is different. Instead of fixing the intuitive rule on the basis of a misspecified estimation of the true model as in Fuster, Laibson, and Mendel (2010), we fix the heuristics’ parameters on the basis of experimental estimations, but allow agents to change their heuristics dynamically. Of course, fundamental expectations can also be included into the set of the rules. In a related paper, Hommes and Lux (forthcoming) explain learning to forecast experiments in the classical cobweb (hog-cycle) market environment, using a heterogeneous expectations model, where individual agents use a Genetic Algorithm (GA) learning. The GA-learning model captures all stylized facts of the cobweb markets, both at the individual and at the aggregate level, across different experimental treatments.

The cobweb market environment is an example of a negative feedback system, where higher expectations lead to lower prices. In contrast, the environment considered in this paper has positive feedback, because higher expectations lead to higher price realizations. The type of feedback affects the aggregate behavior significantly. For instance, Heemeijer et al. (2009) report learning-to-forecast experiments where subjects had to predict the price for the same period (i.e., they submitted one-period-ahead forecasts) in environments with different feedback. In the sessions with negative feedback the realized market price converged to the fundamental level very quickly and stayed there for the rest of the experiment. In contrast, in the sessions with positive feedback the price exhibited persistent fluctuations and long lasting deviations from the fundamental level. Anufriev, Hommes, and Philippe (forthcoming) apply the HSM with only two heuristics to these experimental data. The two heuristics are adaptive expectations and a trend-following rule, and the HSM coordinates on one of these heuristics depending on the feedback. Under negative feedback the adaptive rule performs better and the model endogenously coordinates on this rule. Under positive feedback, the trend-following rule performs better and the HSM selects this heuristic thus reinforcing price oscillations.

Anufriev, Hommes, and Philipse (forthcoming) estimate the model and show that it outperforms several other models, including those with homogeneous expectations. In this way, it is shown that in some situations (with one-period-ahead predictions, in particular), the HSM with two heuristics only can explain aggregate dynamics under different feedback structures of expectations.

In recent work, Assenza et al. (2011) fit the HSM to learning-to-forecast experiment in a macroeconomic, New Keynesian framework with feedback from output and inflation expectations. A more systematic investigation in which market environments other than the asset pricing framework our nonlinear heuristics switching model can explain individual forecasting behavior as well as aggregate price behavior is beyond the scope of the current paper. We conjecture, however, that the simple HSM analyzed in this paper works well in market environments with some structure and persistence in price fluctuations, either in the form of converging prices or in the presence of price oscillations and (temporary) trends.

In future work, we intend to apply our HSM to financial market data with more irregular and persistent price time series in order to see whether they can pick up temporary trends in asset prices and improve financial forecasting. For this application the set of heuristics should be augmented by more sophisticated rules, those which are applied in financial markets. Allen and Taylor (1990) and Frankel and Froot (1990) provide evidence based on survey data, that simple extrapolation heuristics similar to those we employ, are used in real markets. An important challenge is to see whether strategy switching models can improve financial forecasting.

## REFERENCES

- Adam, Klaus.** 2007. "Experimental Evidence on the Persistence of Output and Inflation." *Economic Journal* 117 (520): 603–36.
- Adam, Klaus, Albert Marcet, and Juan Pablo Nicolini.** 2011. "Stock Market Volatility and Learning." London School of Economics (LSE) Centre for Economic Performance (CEP) Discussion Paper 1077.
- Allen, Helen, and Mark P. Taylor.** 1990. "Charts, Noise, and Fundamentals in the London Foreign Exchange Market." *Economic Journal* 100 (400): 49–59.
- Anufriev, Mikhail, and Cars Hommes.** 2012. "Evolutionary Selection of Individual Expectations and Aggregate Outcomes in Asset Pricing Experiments: Dataset." *American Economic Journal: Microeconomics*. <http://dx.doi.org/10.1257/mic.4.4.35>.
- Anufriev, Mikhail, and Cars H. Hommes.** 2012. "Evolution of Market Heuristics." *Knowledge Engineering Review* 27 (2): 255–71.
- Anufriev, Mikhail, Cars H. Hommes, and Raoul H. S. Philipse.** Forthcoming. "Evolutionary Selection of Expectations in Positive and Negative Feedback Markets." *Journal of Evolutionary Economics*. <http://dx.doi.org/10.1007/s00191-011-0242-4>.
- Anufriev, Mikhail, and Valentyn Panchenko.** 2009. "Asset Prices, Traders' Behavior and Market Design." *Journal of Economic Dynamics and Control* 33 (5): 1073–90.
- Arthur, W. Brian.** 1991. "Designing Economic Agents that Act Like Human Agents: A Behavioral Approach to Bounded Rationality." *American Economic Review* 81 (2): 353–59.
- Assenza, Tiziana, Peter Heemeijer, Cars H. Hommes, and Domenico Massaro.** 2011. "Individual Expectations and Aggregate Macro Behavior." University of Amsterdam Center for Nonlinear Dynamics in Economics and Finance (CeNDEF) Working Paper 2011-01.
- Bao, Te, Cars H. Hommes, Joep Sonnemans, and Jan Tuinstra.** 2012. "Individual Expectation, Limited Rationality, and Aggregate Outcome." *Journal of Economic Dynamics & Control* 36 (8): 1101–20.
- Boswijk, H. Peter, Cars H. Hommes, and Sebastiano Manzan.** 2007. "Behavioral Heterogeneity in Stock Prices." *Journal of Economic Dynamics and Control* 31 (6): 1938–70.
- Branch, William A., and George W. Evans.** 2006. "Intrinsic Heterogeneity in Expectation Formation." *Journal of Economic Theory* 127 (1): 264–95.

- Branch, William A., and George W. Evans.** 2011. "Learning about Risk and Return: A Simple Model of Bubbles and Crashes." *American Economic Journal: Macroeconomics* 3 (3): 159–91.
- Bray, Margaret M., and Nathan E. Savin.** 1986. "Rational Expectations Equilibria, Learning, and Model Specification." *Econometrica* 54 (5): 1129–60.
- Brock, William A., and Cars H. Hommes.** 1997. "A Rational Route to Randomness." *Econometrica* 65 (5): 1059–95.
- Brock, William A., and Cars H. Hommes.** 1998. "Heterogeneous Beliefs and Routes to Chaos in a Simple Asset Pricing Model." *Journal of Economic Dynamics and Control* 22 (8–9): 1235–74.
- Bullard, James.** 1994. "Learning Equilibria." *Journal of Economic Theory* 64 (2): 468–85.
- Camerer, Colin F., and Ernst Fehr.** 2006. "When Does 'Economic Man' Dominate Social Behavior?" *Science* 311 (5757): 47–52.
- Camerer, Colin F., and Teck-Hua Ho.** 1999. "Experienced-Weighted Attraction Learning in Normal Form Games." *Econometrica* 67 (4): 827–74.
- Campbell, John Y., Andrew W. Lo, and A. Craig MacKinlay.** 1997. *The Econometrics of Financial Markets*. Princeton: Princeton University Press.
- Carceles-Poveda, Eva, and Chryssi Giannitsarou.** 2008. "Asset Pricing with Adaptive Learning." *Review of Economic Dynamics* 11 (3): 629–51.
- DeBondt, Werner F. M., and Richard H. Thaler.** 1989. "Does the Stock Market Overreact?" *Journal of Finance* 40 (3): 793–805.
- de Jong, Eelke, Willem F. C. Verschoor, and Remco C. J. Zwinkels.** 2009. "Behavioural Heterogeneity and Shift-Contagion: Evidence from the Asian Crisis." *Journal of Economic Dynamics and Control* 33 (11): 1929–44.
- Diks, Cees, and Roy van der Weide.** 2005. "Herding, A-Synchronous Updating, and Heterogeneity in Memory in a CBS." *Journal of Economic Dynamics and Control* 29 (4): 741–63.
- Duffy, John.** 2008. "Experimental Macroeconomics." In *New Palgrave Dictionary of Economics*, 2nd ed., edited by Steven Durlauf and Larry Blume. New York: Palgrave Macmillan.
- Erev, Ido, and Alvin E. Roth.** 1998. "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria." *American Economic Review* 88 (4): 848–81.
- Evans, George W., and Seppo Honkapohja.** 2001. *Learning and Expectations in Macroeconomics*. Princeton: Princeton University Press.
- Fama, Eugene F.** 1970. "Efficient Capital Markets: A Review of Theory and Empirical Work." *Journal of Finance* 25 (2): 383–417.
- Franke, Reiner, and Frank Westerhoff.** 2011. "Structural Stochastic Volatility in Asset Pricing Dynamics." Bamberg Economic Research Group (BERG) Working Paper 78.
- Frankel, Jeffrey A., and Kenneth A. Froot.** 1990. "Chartists, Fundamentalists, and Trading in the Foreign Exchange Market." *American Economic Review* 80 (2): 181–85.
- Friedman, Milton.** 1953. *Essays in Positive Economics*. Chicago: University of Chicago Press.
- Fuster, Andreas, David Laibson, and Brock Mendel.** 2010. "Natural Expectations and Macroeconomic Fluctuations." *Journal of Economic Perspectives* 24 (4): 67–84.
- Gaunersdorfer, Andrea, Cars H. Hommes, and Florian O. O. Wagener.** 2008. "Bifurcation Routes to Volatility Clustering under Evolutionary Learning." *Journal of Economic Behavior and Organization* 67 (1): 27–47.
- Heemeijer, Peter, Cars Hommes, Joep Sonnemans, and Jan Tuinstra.** 2009. "Price Stability and Volatility in Markets with Positive and Negative Expectations Feedback: An Experimental Investigation." *Journal of Economic Dynamics and Control* 33 (5): 1052–72.
- Hey, John D.** 1994. "Expectations Formation: Rational or Adaptive or . . . ?" *Journal of Economic Behavior and Organization* 25 (3): 329–49.
- Hommes, Cars H.** 2006. "Heterogeneous Agent Models in Economics and Finance." In *Handbook of Computational Economics: Agent-Based Computational Economics*. Vol. 2, edited by K. L. Judd and L. Tesfatsion, 1109–59. Amsterdam: Elsevier.
- Hommes, Cars H.** 2011. "The Heterogeneous Expectations Hypothesis: Some Evidence from the Lab." *Journal of Economic Dynamics and Control* 35 (1): 1–24.
- Hommes, Cars H., Hai Huang, and Duo Wang.** 2005. "A Robust Rational Route to Randomness in a Simple Asset Pricing Model." *Journal of Economic Dynamics and Control* 29 (6): 1043–72.
- Hommes, Cars H., and Thomas Lux.** Forthcoming. "Individual Expectations and Aggregate Behavior in Learning to Forecast Experiments." *Macroeconomic Dynamics*.
- Hommes, Cars H., Joep Sonnemans, Jan Tuinstra, and Henk van de Velden.** 2005. "Coordination of Expectations in Asset Pricing Experiments." *Review of Financial Studies* 18 (3): 955–80.

- Hommel, Cars H., Joep Sonnemans, Jan Tuinstra, and Henk van de Velden.** 2008. "Expectations and Bubbles in Asset Pricing Experiments." *Journal of Economic Behavior and Organization* 67 (1): 116–33.
- Kahneman, Daniel.** 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review* 93 (5): 1449–75.
- Lei, Vivian, Charles N. Noussair, and Charles R. Plott.** 2001. "Nonspeculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality." *Econometrica* 69 (4): 831–59.
- Lucas, Robert E., Jr.** 1972. "Expectations and the Neutrality of Money." *Journal of Economic Theory* 4 (2): 103–24.
- Marimon, Ramon, Stephen E. Spear, and Shyam Sunder.** 1993. "Expectationally Driven Market Volatility: An Experimental Study." *Journal of Economic Theory* 61 (1): 74–103.
- Muth, John F.** 1961. "Rational Expectations and the Theory of Price Movements." *Econometrica* 29 (3): 315–35.
- Peterson, Steven P.** 1993. "Forecasting Dynamics and Convergence to Market Fundamentals: Evidence from Experimental Asset Markets." *Journal of Economic Behavior and Organization* 22 (3): 269–84.
- Pfajfar, Damjan, and Blaz Zakelj.** 2010. "Inflation Expectations and Monetary Policy Design: Evidence from the Laboratory." Tilburg University Center for Economic Research (CentER) Discussion Paper 2011-091.
- Samuelson, Paul A.** 1965. "Proof That Properly Anticipated Prices Fluctuate Randomly." *Management Review* 6 (2): 41–49.
- Sargent, Thomas J.** 1993. *Bounded Rationality in Macroeconomics*. New York: Oxford University Press.
- Shiller, Robert J.** 1981. "Do Stock Prices Move Too Much to Be Justified by Subsequent Changes in Dividends?" *American Economic Review* 71 (3): 421–36.
- Simon, Herbert A.** 1957. *Models of Man: Social and Rational*. New York: John Wiley.
- Smith, Vernon L.** 1962. "An Experimental Study of Competitive Market Behavior." *Journal of Political Economy* 70 (2): 111–37.
- Smith, Vernon L., Gerry L. Suchanek, and Arlington W. Williams.** 1988. "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets." *Econometrica* 56 (5): 1119–51.
- Tversky, Amos, and Daniel Kahneman.** 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185 (4157): 1124–30.
- Williams, Arlington W.** 1987. "The Formation of Price Forecasts in Experimental Markets." *Journal of Money, Credit, and Banking* 19 (1): 1–18.