Simple Forecasting Heuristics that Make us Smart: Evidence from Different Market Experiments*

Mikhail Anufriev[a], Cars Hommes[b,c], and Tomasz Makarewicz[b,c]

[a]University of Technology, Sydney
[b]CeNDEF, University of Amsterdam
[c]Tinbergen Institute, Amsterdam

February 4, 2015

Abstract

We study an agent-based model in which heterogeneous individual agents use simple linear first order price forecasting rules, but adapt them to the complex evolving market environment with a smart Genetic Algorithm optimization procedure. The novelties are: (1) a parsimonious experimental micro foundation of individual forecasting behaviour; (2) an explanation of individual and aggregate behavior in four different experimental settings, (3) improved one-period and 50-period ahead forecasting of lab experiments, and (4) a characterization of the mean, the median and the empirical distribution of forecasting heuristics. The mean or median of the distribution of GA forecasting heuristics can be used in designing or validating simple Heuristic Switching Models (Anufriev and Hommes, 2012).

JEL Classification: C63, D03, D53, D83, D84, G02
Keywords: Agent-based Model, Expectation Formation, Learning to Forecast Experiment, Genetic Algorithm Model of Individual Learning.

*The authors gratefully acknowledge the financial support from NWO (Dutch Science Foundation) Project No. 40611142 "Learning to Forecast with Evolutionary Models". The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 612796. Email addresses: `Mikhail.Anufriev@uts.edu.au`, `C.H.Hommes@uva.nl`, `T.A.Makarewicz@uva.nl`.

1 Introduction

Expectations are a cornerstone of many dynamic economic models. The traditional literature (after Lucas Jr., 1972; Muth, 1961) emphasizes the Rational Expectations (RE) hypothesis, which states that in equilibrium the predictions have to be model consistent. Most economists would assume that agents are rational enough to learn to avoid systematic, correlated errors, and so RE remain a reasonable approximation of the aggregate market dynamics.[1] However, this is not confirmed by empirical market studies. A recent example of the latter comes from the housing market in the US before the 2007 economic crisis, where people systematically misjudged the long-term value of their houses (Case and Shiller, 2003; Eren et al., 2008; Goodman Jr. and Ittner, 1992). In fact economic history knows many similar examples of prolonged asset misvaluation (Reinhart and Rogoff, 2009). In a broader context of macroeconomics, inflation expectations formed by the 'Jones' are far from the RE predictions (Adam, 2007; Assenza et al., 2013; Charness et al., 2007; Pfajfar and Zakelj, 2011) and can be subject to cognitive biases (Malmendier and Nagel, 2009). Many firms similarly fail to use RE (see *e.g.* Nunes, 2010, for a discussion about the Phillips Curve and survey expectations).

The empirical deficiencies of the RE benchmark signifies the importance of learning. Recent years witnessed surge of literature trying to merge RE with learning, with the two notable examples of adaptive or statistical learning (Evans and Honkapohja, 2001; Evans and Ramey, 2006) and rational inattention (Sims, 2010). These models still face many empirical and theoretical obstacles (see Wagener, 2013; Woodford, 2013, for an overview). In addition, even if the agents know the structure of the economy, the price-expectation feedback can lead to non-trivial dynamics (Grandmont, 1985; Hommes, 2013). This picture becomes more complicated if the agents furthermore have to learn this structure itself (Bullard, 1994; Grandmont, 1998; Tuinstra and Wagener, 2007).

The alternative approach is to abandon RE altogether in favor on boundedly rational models of learning. Experimental studies on price expectations indicate that people often follow simple forecasting heuristics like chartists rules. What remains a question, however, is how exactly people invent their heuristics. This leads to the so-called 'wilderness of bounded rationality' problem: there is a myriad of possible learning mechanisms with varied restrictions on human memory and computational capabilities. These range from simple linear heuristic models (see Evans and Ramey, 2006, for a discussion of adaptive expectations), through heuristic switching type of models (Brock and Hommes, 1997) to evolutionary learning mechanisms (Arifovic et al., 2012). In this paper we present a model of individual learning of the forecasting heuristics, based on Genetic Algorithms updating. Heterogeneous individual agents use simple linear first order price forecasting rules, but adapt them to the current complex evolving environment with a smart Genetic Algorithm optimization procedure. In this sense, agents use simple forecasting heuristics that make them smart (Gigerenzer and Todd, 1999). We fit our GA model to individual and aggregate experimental data from Learning-

[1] One interesting and straightforward explication of this approach can be found in the concluding section of Blundell and Stoker (2005).

to-Forecast experiments to study the mean, the median and the distribution of individual forecasting heuristics used in different experimental laboratory settings. Our model replicates the experimental distribution of the individual heuristics in different experimental settings, and proves by Monte Carlo simulations that the popular stylized Heuristic Switching model is an accurate description of the median experimental behavior.

Learning-to-Forecast (LtF) experiments offer a simple laboratory testing ground for adaptive learning mechanisms (Hommes et al., 2005; Lucas Jr, 1986). These controlled experimental economies have a straightforward and unique fundamental (RE) equilibrium. As in real markets, subjects observe the realized prices and their own past individual predictions, but not the history of other subjects' predictions, and are not informed about the exact law of motion of the economy. Many LtF laboratory experiments contradict the RE hypothesis. The subjects can coordinate on oscillating and serially correlated time series, and the exact dynamics depend greatly on the specific feedback structure of the experimental economy. Convergence to the fundamental equilibrium happens only under severe restrictions on the underlying law of motion (Hommes, 2011). In particular, the expectations feedback structure turns out to be crucial: negative feedback (i.e. a more optimistic forecasts leads to lower market prices, as in producer driven commodity markets) systems tend to be stable and converge to the RE equilibrium rather quickly, while positive feedback (i.e. more optimistic forecasts leads to higher market prices, as in speculative asset markets) do not converge but prices rather fluctuate and oscillate around the RE fundamental price (Heemeijer et al., 2009, henceforth HHST09). Another important finding in the experiments is heterogeneity: within the same experimental group, subject tend to use different forecasting rules, see Hommes (2011).

The most successful attempt to explain the LtF experiments comes with the so-called Heuristic Switching Model (HSM; Brock and Hommes, 1997). The basic idea of the model is that agents have a (small) set of simple forecasting heuristics (rules of thumb like adaptive or trend extrapolating expectations) and gradually switch to relatively better performing rules. The HSM has successfully been used to explain different types of aggregate behavior — convergence and oscillations in various experimental settings (Anufriev and Hommes, 2012; Anufriev et al., 2013). A disadvantage of the HSM however is the small set of heuristics, which cannot fully account for individual heterogeneity. Furthermore different experiments require the HSM to utilize differently specified sets of heuristics. It is unclear why the subjects would use only these particular forecasting rules (with given parameters) and how they would learn them.

Genetic Algorithms (GA) are a prominent tool in the economic literature to model individual learning (Dawid, 1996). GA were initially applied in their social learning form to explore stylized facts from experimental data, outperforming the RE hypothesis (Arifovic, 1995). Examples include exchange rate volatility (Arifovic, 1996; Lux and Schornstein, 2005) and production level choices in a cobweb producers economy (Dawid and Kopel, 1998). More recently, Hommes and Lux (2013) investigate a model, in which agents use GA to optimize a forecasting heuristic (instead of directly optimizing a prediction) and, much like the actual subjects in LtF experiments, cannot observe each others behavior or strategies. The authors replicate the distribution (mean, variance and first order auto-correlation) of the predictions and prices of the cobweb experiments by Hommes et al. (2007) and van de Velden (2001) (henceforth HSTV07 and V01 respectively).

The main purpose of our paper is to provide a general explanation of the LtF experiments and individual forecasting heuristics using a GA based model. Agents forecast prices using a possibly large set of heuristics from a simple but general class. The agents then *independently* use GA to update and select the heuristics based on their relative success. This results in an agent-based model with explicit individual learning and endogenous heterogeneity. Monte Carlo simulations of this agent-based model provide insight in the mean, median and the distribution of forecasting heuristics.

The novelties of our paper are four. The first is the generality, but at the same time parsimony of our model. We use a heuristic space based on the linear first order rule, which is a mixture of adaptive and trend extrapolating heuristics, consistent with the estimated individual forecasting rules in (Hommes et al., 2005, henceforth HSTV05) and HHST09. This provides our model with a parsimonious *experimental micro-foundations*. A second contribution of our paper is that we apply the same GA learning model to explain *four different* LtF experiments: (1) the simple, linear positive/negative feedback system with small shocks (HHST09); (2) the linear positive/negative price-expectations feedback system with unexpected large shocks to the fundamental price (Bao et al., 2012, henceforth BHST12); (3) a stable/unstable cobweb producers economy (HSTV07; V01), used also in the GA learning model of Hommes and Lux (2013); and (4) a non-linear positive feedback asset pricing economy, where the subjects are asked for two-period ahead predictions (HSTV05).

The third contribution of the paper is that our model is able to capture the dynamics at both the *aggregate* and the *individual* level for different experimental settings. The GA model replicates the long-run behavior of the experimental prices, as well as the individual forecasting decisions. We also evaluate the out-of-sample predictive power of the model by means of a simple Sequential Monte Carlo technique. We find that depending on the experiment, our model is comparable or better than the HSM in terms of predicting both *prices* and *individual price forecasts* one period ahead. This is an important contribution to the literature on agent-based models, which usually focuses only on a model's fit to aggregate stylized facts.

Finally, the fourth contribution is that the Monte Carlo studies of the GA model enable us to characterize the emerging *median forecasting behavior*, together with its corresponding confidence bounds, in various experimental settings. The GA simulations thus (1) provide a solid motivation for describing the LtF experimental dynamics in terms of simple 'stylized' heuristics, and (2) guide the specific choice of these heuristics for a particular experimental market. This yields natural *empirical micro-foundations* for heterogeneous expectations models such as HSM.

The paper is organized as follows. In Section 2, we present the setup and findings of the LtF experiments, and briefly discuss the HSM by Anufriev and Hommes (2012). In the third section, we introduce our GA model and fit it to the experimental setup of HHST09. Section 4 investigates three other experimental settings. Finally, a concluding section gives an overview of the results and suggestions for future research. An appendix contains GA simulation details and robustness checks.

2 Learning to Forecast and Heuristic Switching

Consider a market with a number of subjects $i \in \{1, \dots, I\}$, who are asked at each period t to forecast the price of a certain good. The subjects are explicitly informed that they act as forecasting consultants for firms and are rewarded only for the accuracy of the predictions.

The feedback relationship between the prices and predictions is summarized by a law of motion of the form

$$p_t = F(p^e_{1,t}, \dots, p^e_{I,t}), \tag{1}$$

where the realized price p_t is a function of all individual forecasts $p^e_{i,t}$. The mapping $F(\cdot)$ is derived from market clearing, with aggregate supply and demand derived from optimal (*i.e.* profit/utility maximizing) choices of firms, consumers or investors, given the subjects' individual forecasts. Define the fundamental price p^f as the steady state RE outcome, the self-fulfilling prediction: $p^f = F\left(p^f, \dots, p^f\right)$. In all examples below the RE fundamental price is unique.

Unlike RE agents, subjects in the experiment have limited information about the market. They are informed that their predictions affect realized prices, but they are given only a qualitative story about this feedback. Moreover, they are not explicitly informed about the fundamental price.[2]

One important example investigated by HHST09 uses a linear version of (1):

$$p_t = A + B\left(\frac{\sum_{i=1}^{I} p^e_{i,t}}{I} - A\right) = A + B\left(\bar{p}^e_t - A\right) + \varepsilon_t, \tag{2}$$

where $\bar{p}^e_t = \frac{\sum_{i=1}^{I} p^e_{i,t}}{I}$ is the average prediction of all individuals at period t, $A = p^f$ is the fundamental price and ε_t is a small noise term. There are two important cases: $B > 0$ (positive feedback) and $B < 0$ (negative feedback). A typical example of positive feedback is a stock exchange: optimistic investors will buy more stock and due to increased demand the stock price will go up. In this sense the investor sentiments are self-fulfilling (although not perfectly if $B \neq 1$). Negative feedback arises *e.g.* in a supply driven market where producers face a lag in production. If they expect a high price in the future, they will increase production

[2] Usually it is possible to infer it from the experimental instructions. For example, in the asset pricing treatment of HHST09 the fundamental price is the ratio of mean dividend to interest rate, both variables explicitly provided to the subjects. Anecdotal evidence suggests, however, that even economics students, including graduate students, fail to realize it.

and so the market clearing price will go down.

HHST09 used two simple linear feedback maps:

$$\text{(3)} \qquad \textbf{Positive feedback:} \quad p_t = 60 + \frac{20}{21}(\bar{p}_t^e - 60) + \varepsilon_t;$$

$$\text{(4)} \qquad \textbf{Negative feedback:} \quad p_t = 60 - \frac{20}{21}(\bar{p}_t^e - 60) + \varepsilon_t,$$

where $\varepsilon_t \sim NID(0, 0.25)$. The experiment runs for 50 periods for each group of $I = 6$ subjects. The two linear treatments are symmetrically opposite. They have the same unique fundamental price $p^f = 60$ and the same absolute dampening factor $|B| = \frac{20}{21}$ (but with opposite signs). The dampening factors were chosen so that under naive expectations (*i.e.* $\bar{p}_t^e = p_{t-1}$), the fundamental price for both treatments is a (unique) stable steady state, but the system would still require some time to converge.[3]

The two feedback treatments resulted in very different aggregate price behavior, illustrated in Figure 1a and 1b. Under the negative feedback after a short volatile phase of $7-8$ periods, the price converges to the fundamental value $p^f = 60$, after which the subjects forecasts coordinate on the fundamental as well. In most of the positive feedback groups, persistent price oscillations arise (Figure 1b), where the price overshoots and undershoots p^f; if the price converges to the fundamental at all, it does so only towards the end of the experiment (which happened for two out of seven cases). In spite of the price oscillations, the subjects' forecasts coordinate within $2-3$ periods on a common value (different from the fundamental value) and remain so until the end of the experiment. In positive feedback markets, subjects' forecasts are thus strongly coordinated, but on a non-RE price. It is the almost self-fulfilling character of the near-unit root positive feedback system that allows for subject coordination on trend following behavior, which results in price oscillations (Hommes, 2013).

To describe the subjects' forecasting behavior, HHST09 use the first-order rule (FOR):

$$\text{(5)} \qquad p_{i,t}^e = \alpha_1 p_{t-1} + \alpha_2 p_{i,t-1}^e + \alpha_3 p^f + \beta(p_{t-1} - p_{t-2}),$$

for $\alpha_1, \alpha_2, \alpha_3 \geqslant 0$, $\alpha_1 + \alpha_2 + \alpha_3 = 1$, $\beta \in [-1, 1]$. Rule (5) is an anchor and adjustment rule, extrapolating a price change from an anchor, which is given by a weighted average of past price, individual forecast and the fundamental price $p^f = 60$.[4] HHST09 estimated this simple rule separately for each subject, based on their predictions from the last 40 periods, and were able to describe well the forecasting behavior of around 60% of the individuals.

HHST09 find that the individual forecasting rules varies between the subjects, even within the same treatment. The authors also report significant differences between the two treatments. Under positive feedback, subjects focus on trend extrapolation ($\beta > 0$) and the estimated weight of the fundamental price α_3 coefficients were typically insignificant. Under negative feedback, the reverse holds: trend extrapolation is barely used ($\beta \approx 0$), while the

[3]In an asset pricing market, the near unit root coefficient $\frac{20}{21} \approx 0.95$ arises from a realistic discount factor.

[4]Under RE, the FOR in (5) should be specified with $\alpha_1 = \alpha_2 = \beta = 0$, and subjects always predict the fundamental price, $p_{i,t}^e = p^f = 60$.

weight for the fundamental price ($\alpha_3 > 0$) is significant. This shows that a model with a homogeneous forecasting rule (RE, but also linear heuristics like trend extrapolation or naive expectations) may explain one of the two treatments, but not both at the same time. Moreover, a homogenous rule contradicts the significant differences between the subjects within each treatment.

This led Anufriev et al. (2013) to investigate the Heuristic Switching Model (HSM), in which the subjects are endowed with two prediction heuristics:

adaptive expectations: $p^e_{i,t} = \alpha p_{i,t-1} + (1-\alpha)p^e_{i,t-1}$ with $\alpha \in [0, 1]$,

trend extrapolation: $p^e_{i,t} = p_{i,t-1} + \beta(p_{t-1} - p_{t-2})$ with $\beta \in [-1, 1]$,

where the authors used $\alpha = 0.75$ and $\beta = 1$. Notice that both heuristics are a special case of the first-order rule (5). The idea of the HSM model is that the subjects can at any time use any of the two heuristics, but tend to switch to the rule with higher relative past performance. Under positive feedback, agents easily coordinate their predictions, for example below the fundamental, close to the first observed price, but (by the construction of the positive feedback equation) the next realized price is then slightly higher than the average prediction. The trend extrapolation heuristic captures this gradual increase of initial prices and so becomes popular among the agents. This reinforces the trend and leads to persistent price oscillations. In contrast, under negative feedback there is no possibility of coordination of individual forecasts outside the fundamental price. Under negative feedback the trend extrapolating rule performs poorly and agents switch to adaptive expectations, thus causing the price to converge to the fundamental.

HSM captures the essence of the aggregate forecasting behavior and successfully replicates the results of HHST09 in both treatments in a stylized fashion. The drawback of the model, however, is that the authors assume a limited number of only two heuristics, without explaining where these heuristics come from, that is, how the subjects are able to learn the two heuristics and the specific coefficients. Moreover, the HSM cannot account for heterogeneity of rules among subjects and hence does not explain the experiment at the individual level. To overcome these drawbacks, we will introduce a model with explicit individual learning of a class of simple forecasting heuristics through Genetic Algorithms (*c.f.* Arifovic and Ledyard, 2004).

3 The Genetic Algorithms model

3.1 Genetic Algorithms

Genetic Algorithms (GA) form a class of numerical stochastic maximization procedures that mimic the evolutionary operations with which DNA of biological organisms adapts to the environment. GA were introduced to solve 'hard' optimization problems, which may involve noncontinuities or high dimensionality with complicated interrelations between the arguments. They are flexible and efficient and so are often used in computer sciences and engineering

(Haupt and Haupt, 2004). See *e.g.* Haupt and Haupt (2004) for technical discussion and Dawid (1996) for applications in economics.

A GA routine starts with a population of random trial solutions to the problem. Individual trial arguments are encoded as binary strings (strings of ones and zeros), or chromosomes. They are retained into the next iteration with a probability that increases with their relative performance, which is defined directly in terms of a functional value ('fitness'). This so called **procreation** operator means that with each iteration, the population of trial arguments is likely to have a higher functional value, *i.e.* be 'fitter'. On top of the procreation, GA use three evolutionary operators that allow for an efficient search through the problem space: mutation, crossover and election, where the last operator was introduced in the economic literature (Arifovic, 1995).

Mutation At each iteration, every entry in each chromosome has a small probability to mutate, in which case it changes its value from zero to one and *vice versa*. The mutation operator utilizes the binary representation of the arguments. A single change of one bit at the end of the chromosome leads to a minor, numerically insignificant change of the argument. But with the same probability a mutation of a bit at the beginning of the chromosome can occur, which changes the argument drastically. With this experimentation, GA can easily search through the whole parameter space and have a good chance of shifting from a local maximum towards the region containing the global maximum.

Crossover Pairs of arguments can, with a predefined probability, exchange predefined parts of their respective binary strings. In practice, the crossover is set to exchange bits corresponding to a subset of the arguments. For example, if the objective function has two arguments, crossover would swap the first argument between pairs of trial arguments. This allows for experimentation in terms of different mixtures of arguments.

Election The election operator is meant to screen inefficient outcomes of the experimentation phase. This operator transmits the new chromosomes (selected from the old generation and treated with mutation and crossover) into the new generation only if their functional value is greater than that of the original argument. This operator ensures that once the routine finds the global solution, it will not diverge from it due to unnecessary experimentation.

The procreation routine and the three evolutionary operators have a straightforward economic interpretation for a situation, in which the agents want to optimize their behavioral rules, *e.g.* price forecasting heuristics. The procreation means that — as in the case of HSM — people focus on better solutions (or heuristics). The mutation and crossover are experimentation with the heuristics' specifications, and finally the election ensures that the experimentation does not lead to suboptimal heuristics.

An important additional condition for a GA routine is that it requires a predefined interval for each parameter. For the above example of updating behavioral rules through GA, it means that we confine them to some predetermined, finite grid of heuristics.

3.2 Model specification

We consider a set of $I = 6$ GA agents in the price-expectation feedback economy (1). GA agents use a general forecasting rule which requires exact parameter specification, and each agent is endowed with $H = 20$ such specifications. In order to give our model robust *empirical micro-foundations*, we follow the estimations of individual rules by HHST09, as well as the simple model discussed by Anufriev et al. (2013) and focus on the first order rule (FOR).[5]

To be specific, in order to predict price at date t p_t agent $i \in \{1, \ldots, I\}$ chooses from $H = 20$ linear prediction rules given by

$$p^e_{i,h,t} = \alpha_{i,h,t} p_{t-1} + (1 - \alpha_{i,h,t}) p^e_{i,t-1} + \beta_{i,h,t}(p_{t-1} - p_{t-2}), \tag{6}$$

where $p^e_{i,h,t}$ is the prediction of price p_t, formulated by agent i conditional on using rule h at the beginning of period t, and $p^e_{i,t-1}$ is the prediction by agent i of the price p_{t-1}, which the agent submitted to the market in period $t-1$. Rule (6) is a simplified version of the general FOR (5) with $\alpha_3 = 0$.[6]

Heuristic FOR (6) depends on two parameters only, namely on $\alpha_{i,h,t}$ (price weight) and $\beta_{i,h,t}$ (trend extrapolation coefficient). These parameters are time dependent, because the agents want to fine-tune the FOR (6) for their specific market. For example, in an asset pricing market it may pay off to focus on the trend of the asset price. Agents try to find the optimal degree of trend following, by experimenting with different trend extrapolation coefficients $\beta_{i,h,t}$. This learning is embodied as a heuristic optimization with the GA procedure, and constitutes the novel individual heterogeneity of the model, compared to HSM or any homogenous expectations model.

Define $\mathbf{H}_{i,t}$ as the set of heuristics of agent i at time t. Each agent has $H = 20$ heuristics which are specified as a pair of parameters $(\alpha_{i,h,t}, \beta_{i,h,t}) \in \mathbf{H}_{i,t}$. Each pair is represented as a chromosome, a binary string of length 40, 20 bits per coefficient. This means that the coefficients have to be bound to a finite interval. Price weight simply spans a simplex $\alpha_{i,h,t} \in [0, 1]$. For the trend extrapolation coefficient $\beta_{i,h,t}$, experimentation with the model led us to report two specifications, namely with $\beta \in [-1.1, 1.1]$ (contrarian rules allowed) and $\beta \in [0, 1.1]$ (contrarian rules not allowed).[7]

The chromosomes are updated *independently* for each agent by GA evolutionary operators. We focus on the same set of parameters as Hommes and Lux (2013), see Table 1. The updating of the heuristics is based on their relative forecasting performance, specifically on mean squared

[5] *Cf.* Hommes and Lux (2013), who use GA learning of the anchor-and-adjustment rule $p^e_{i,t} = \alpha + \beta(p_{t-1} - \alpha)$.

[6] We experimented with the full FOR with an anchor equal specified as either (i) the fundamental price p^f or (ii) the average realized price so far. Neither specification could replicate the dynamics of the positive feedback. This is consistent with the fact that in the estimated rules of HHST09 under positive feedback, the anchor weight α_3 is typically insignificant. Finally, the FOR (6) is based on the two heuristics of Anufriev et al. (2013), who also used an anchor without the fundamental price for their HSM.

[7] Experimental data suggests that subjects tend to avoid contrarian strategies (HHST09 report only two subjects with such rules), thus for the sake of completeness we report both specifications.

Parameter	Notation	Value
Number of agents	I	6
Number of heuristics per agent	H	20
Number of parameters	N	2
Allowed α price weight	$[\alpha_L, \alpha_H]$	$[0, 1]$
Allowed β trend extrapolation		
Specification 1	$[\beta_L, \beta_H]$	$[-1.1, 1.1]$
Specification 2	$[\beta_L, \beta_H]$	$[0, 1.1]$
Number of bits per parameter	$\{L_1, L_2\}$	$\{20, 20\}$
Mutation rate	δ_m	0.01
Crossover rate	δ_c	0.9
Lower crossover cutoff point	C_L	20
Higher crossover cutoff point	C_H	-1 (none)
Performance measure	$U(\cdot)$	$\exp(-MSE(\cdot))$

Table 1: Parameter specification used by the Genetic Algorithms agents.

error (MSE). Let

$$MSE_{i,h,t} = (p^e_{h,i,t} - p_t)^2. \tag{7}$$

Define the normalized performance (or fitness) measure as:

$$\Pi_{i,h,t} = \frac{\exp(-MSE_{i,h,t})}{\sum_{j=1}^{H} \exp(-MSE_{i,j,t})}, \tag{8}$$

which is a logit transformation of MSE. The normalized performance measure (8) can be directly interpreted as the probability attached to each heuristic h by agent i at time t.[8]

The timing of the model is as follows. Before the market starts to operate, the agents' heuristics are initialized at random from a 'uniform' distribution: agent i samples 800 initial bits (twenty initial heuristics with two parameters, each encoded by twenty bits) independently as 0 or 1 with equal probability. In some initial periods the agents cannot use their heuristics, as these require past prices and predictions. In these initial periods the agents sample random predictions from a predefined distribution which we take as exogenous (such as the experimental initial predictions). Once the agents have enough observations to use their heuristics, the timing of the market at period t is as follows:

1. The market price p_t is realized according to (1) and agents observe it;

2. Agents *independently* update their heuristics using one GA iteration, where the GA criterion function is $\Pi_{i,h,t}$ (forecasting performance). To be specific, agent i uses four

[8] Notice that (8) is also independent between the agents, because they have different sets of heuristics. Measure (8) is somewhat different from the experimental payoff (*c.f.* Hommes and Lux, 2013). We decided to use logit transformation of the MSE to have a clear link with the HSM literature and to keep this model feature independent from the different experimental designs.

evolutionary operators:

(a) *procreation*: agent samples H child heuristics from $\mathbf{H}_{i,t}$ with $\Pi_{i,h,t}$ as the corresponding probabilities;

(b) *mutation*: each bit of each child heuristic has probability δ_m to switch its value;

(c) *crossover*: each pair of child heuristics has probability δ_c to swap last twenty bits, which corresponds to exchanging β's.

(d) *election*: each child heuristic is compared with the parent heuristic in terms of $\Pi_{i,h,t}$: a child heuristic becomes part of the set $\mathbf{H}_{i,t+1}$ if it outperforms its parent, else its parent is passed to $\mathbf{H}_{i,t+1}$.

3. With the new $\mathbf{H}_{i,t+1}$, period $t+1$ starts.

4. Each agent i picks one particular heuristic $i, h, t+1$, which is based on the hypothetical MSE of heuristics $\mathbf{H}_{i,t+1}$ in predicting the *last observed price* p_t (with probabilities $\Pi_{i,h,t}$). Agent i uses the chosen heuristic to generate her prediction $p^e_{i,t+1}$;

5. New period $t+1$ starts: the market price p_{t+1} is realized according to (1).

In the first period when the heuristics can be used, their hypothetical past performance is still undefined, and so the agents pick one with equal probabilities. For the HHST09 experiment, GA agents start to use the first-order rule in the second period (one at random, and with no observed trend *as if* $\beta = 0$) and start to update their heuristics in the third period (for a more detailed discussion of the initialization, see Appendix B).

The last step — the heuristic choice — is the same as in the HSM, but there are two important differences between HSM and our GA model. First, the set of heuristics evolves over time with $\mathbf{H}_{i,t} \neq \mathbf{H}_{i,t+1}$. As a result, we obtain a HSM in which the heuristics have time varying parameters, adapted to the specific market dynamics. Second, this learning operates through a stochastic GA procedure, and is independent between the agents. In practice thus the agents will learn different heuristics and remain heterogeneous with $\mathbf{H}_{i,t+1} \neq \mathbf{H}_{j,t+1}$.

3.3 50-period ahead simulations

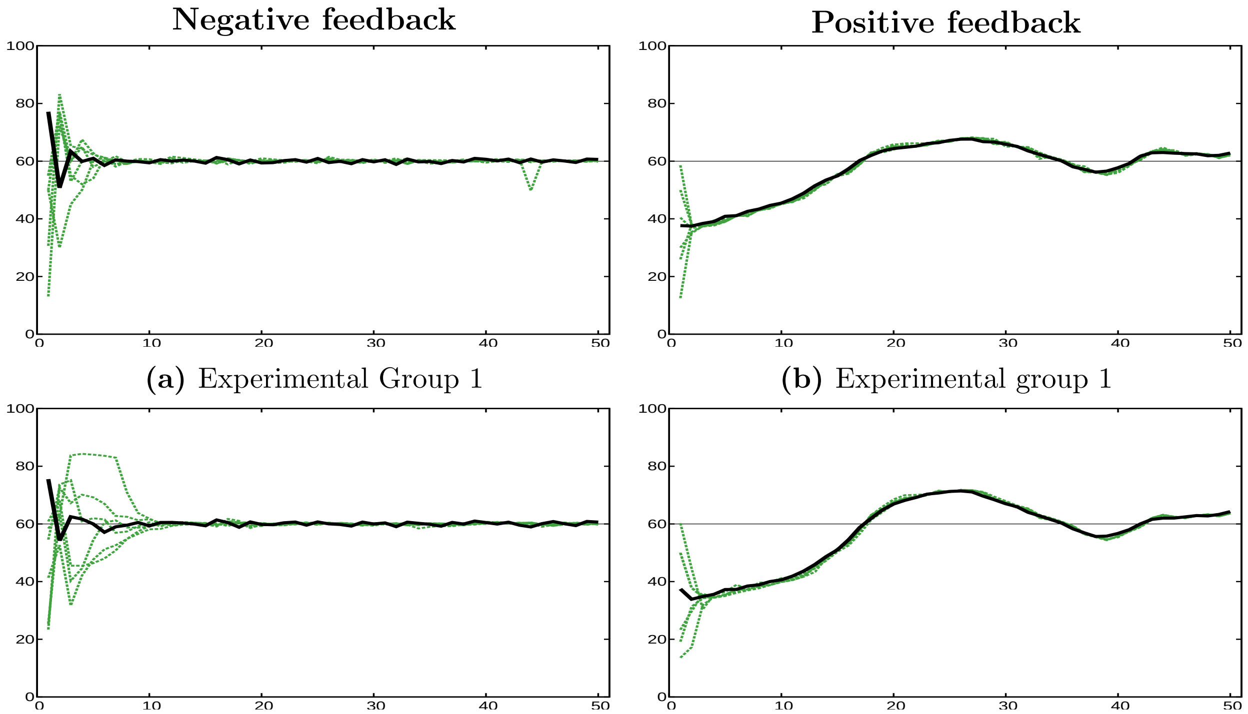
The first test for the fit of our GA model to the experimental data are 50-period ahead simulations for the HHST09 experiment.[9] We take the feedback equations (3) and (4) and simulate our model for 50 periods, without *any information* from the experiment after period 1, and hence compare the realized long-run model dynamics with the experimental data.[10]

The model requires exogenous predictions for the first period. This is important, since in the experiment the average initial prediction affected the group dynamics under the positive

[9] All simulations were written in Ox matrix algebra language (Doornik, 2007) and are available upon request.

[10] In one of the positive feedback treatment groups, one of the subjects 'out of the blue' predicted ten times higher price than both his previous forecast and the realized market price. This destabilized the whole market for a number of periods. In the following analysis, we follow Anufriev et al. (2013) and omit this group and hence focus on six positive feedback and six negative feedback treatment groups.

feedback treatment (*cf.* Anufriev et al., 2013). In the first Monte Carlo (MC) exercise, we sample initial prediction from a distribution calibrated by Diks and Makarewicz (2013), in order to obtain a general picture of the model dynamics (see Appendix B). We resample the model 1′000 times, including new initial predictions and realizations of the learning algorithm, to obtain a satisfactory MC distribution. The median of 1′000 GA simulations, with 95% confidence intervals (CI), for the model with contrarian rules $\beta \in [-1.1, 1.1]$ are shown in Figure 2. See also Figure 1 for sample experimental prices and predictions and realized 50 period ahead GA simulation of a representative group for each feedback treatment.[11]



(a) Experimental Group 1

(b) Experimental group 1

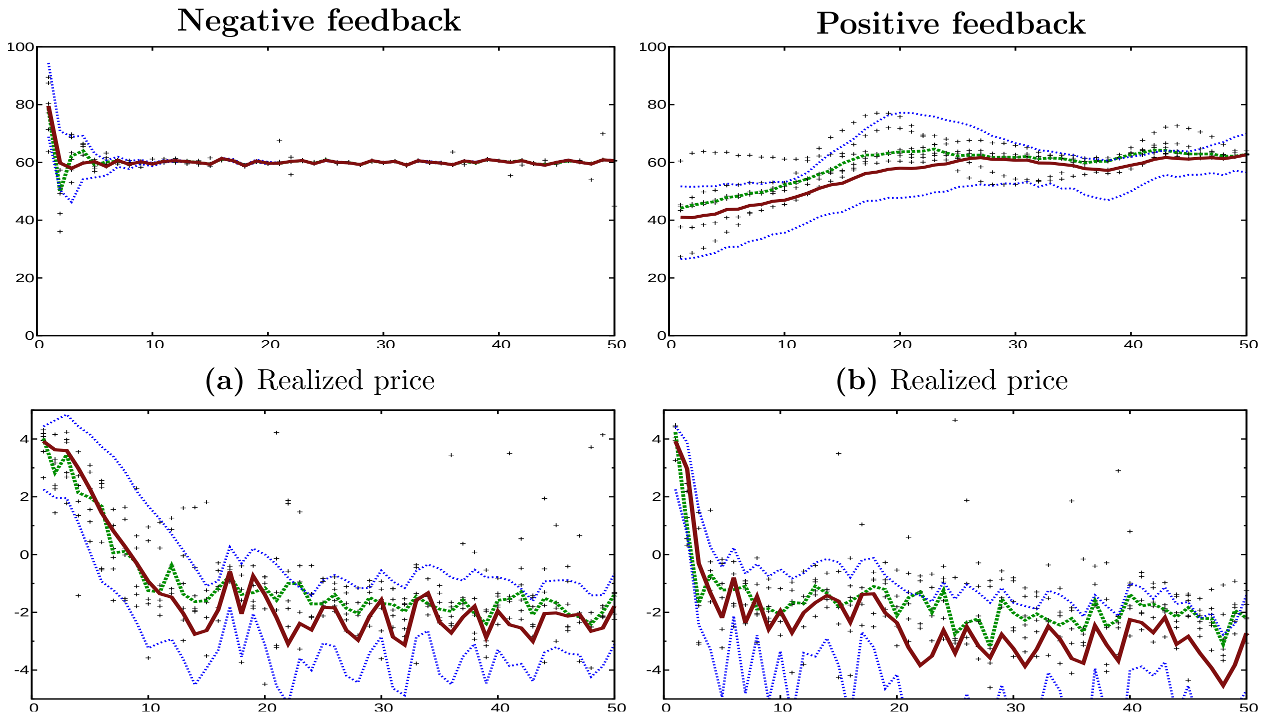
(c) Sample GA with $\beta \in [-1.1, 1.1]$

(d) Sample GA with $\beta \in [-1.1, 1.1]$

Figure 1: HHST09: experimental groups and sample 50-period ahead simulations of the GA model (with $\beta \in [-1.1, 1.1]$ and random initial predictions). Black line represents the price and green dashed lines are the individual predictions.

Figure 2 shows the MC simulations of the realized prices (top panel) and the degree of coordination, that is the standard deviation of six individual forecasts (bottom panel). The model replicates the experimental outcomes well. Under negative feedback (left panels), prices are quickly pushed close to the fundamental, but individual heterogeneity of GA agents declines slowly and is visible until period 15, consistent with the experimental data. Under positive feedback, GA agents coordinate their forecasts in less than five periods, but the distribution of realized prices does not collapse into the fundamental even after 50 periods, when the 95% CI of prices is as wide a $[55, 70]$. The median price resembles the experimental oscillations, including the typical amplitude and turning points. Overall, the 95% CI for our GA model captures 65% (81%) of the experimental prices and 81% (72%) of the degree of coordination for the negative (positive) feedback treatment. This means that we are able

[11]Simulations presented in Figures 1c and 1d were among the first that we run.



(a) Realized price **(b)** Realized price

(c) SD of individual predictions **(d)** SD of individual predictions

Figure 2: HHST09: 50-period ahead Monte Carlo simulation (1000 markets) for the GA model with $\beta \in [-1.1, 1.1]$. Realized price (top) and degree of coordination (standard deviation of individual predictions; bottom; $\log_2$ scale) over time. Green dashed line and black pluses represent the experimental median and group observations; red line is the median and blue dotted lines are the 95% confidence interval for the GA model. Left panel displays the negative feedback, right panel the positive feedback.

to replicate roughly 75% of the long-run (50-period ahead) behavior of the experimental groups, both at the *aggregate* level (prices) and the *individual* level (coordination of individual predictions).

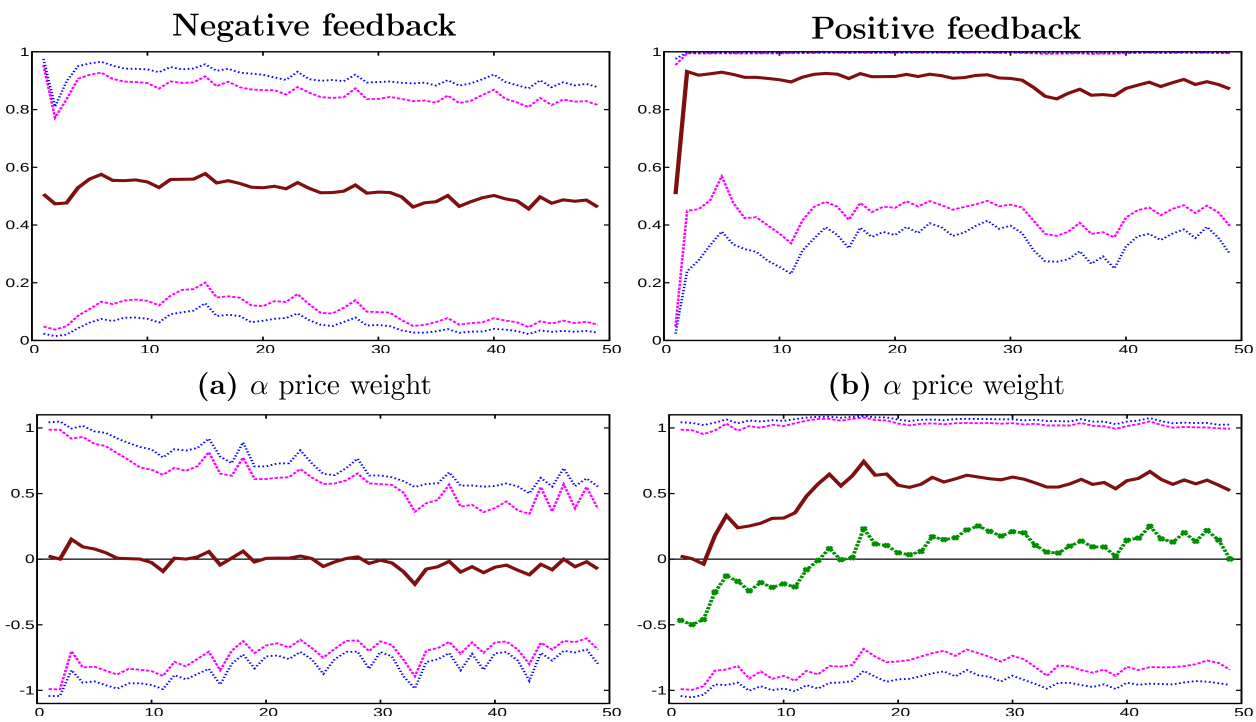
Which heuristics were learned by our GA agents? Figure 3 reports the median (with 95% and 90% CI) for the MC simulations of the price weight α and the trend extrapolation coefficient β. Large heterogeneity of individual rules persists, but there are clear differences between the two treatments. Under the positive feedback treatment, the median GA agent quickly converges towards

$$p^e_{i,t+1} \approx 0.9p_t + 0.1p^e_{i,t} + 0.6(p_t - p_{t-1}). \tag{9}$$

This median rule is close to a pure trend-following rule (*i.e.* with anchor p_t), but has a coefficient $\beta \approx 0.6$, smaller than the $\beta = 1$ coefficient that Anufriev et al. (2013) are using in their 2-type HSM. Furthermore, 72% of the GA agents never uses a negative β in the last 30 periods (see the green star-line in Figure 3d for 28% percentile); and in the last period, the chosen β has a negatively skewed distribution (see Figure 10a).

On the other hand, under negative feedback the median GA agent learns a rule close to

$$p^e_{i,t+1} \approx 0.5p_t + 0.5p^e_{i,t} \tag{10}$$



(a) α price weight

(b) α price weight

(c) β trend extrapolation

(d) β trend extrapolation

Figure 3: HHST09: 50-period ahead Monte Carlo simulation (1000 markets) for the GA model with $\beta \in [-1.1, 1.1]$. The price weight α and the trend extrapolation β *chosen* by the agents over time. Red line is the median, blue dotted lines are 95% CI, purple dashed are 90% CI for the GA model. The green star-line in (d) represents the 28% percentile of chosen β. Left panel displays the negative feedback, right the positive feedback.

(with median β trend coefficient close to 0). This median rule for negative feedback is adaptive expectations with price weight of 0.5; Anufriev et al. (2013) are using adaptive expectations with coefficient 0.75 on price in their 2-type HSM. Our learning dynamics therefore confirm the results by HHST09 and yield empirical support for the 2-type HSM by Anufriev et al. (2013), albeit with slightly different parametrization.

In the second MC study, we focus on how well our GA model can replicate long-run dynamics of a *specific* experimental group. We take initial predictions of each group and use them as initialization for 50-period simulations of the GA model. We investigate the realized prices and individual forecasts. Following Anufriev et al. (2013) we define the GA model expected individual price forecast as

$$\hat{p}^{e,GA}_{i,t} = \sum_{h=1}^{H=20} \Pi^{t}_{i,h,t} p^{e,t}_{i,h,t}. \tag{11}$$

For each sample model time path, we compute its mean squared error (MSE) in predicting the experimental data (both prices and individual price forecasts) for the last 47 periods

MSE	Negative feedback		Positive feedback	
	Prices	**Forecasts**	**Prices**	**Forecasts**
Trend extr.	3421	1696	62.84	72.45
Adaptive	4.164	16.97	95.62	108.6
Contrarian	3.446	*16.18*	108.5	122.8
Naive	112.3	136.2	69.11	79.38
RE	**2.571**	**15.21**	46.835	54.811
HSM	19.64	34.02	55.15	63.98
GA: $\beta \in [-1.1, 1.1]$	*2.884*	20.03	*44.22*	*51.98*
GA: $\beta \in [0, 1.1]$	9.392	29.51	**25.3**	**31.1**

Table 2: HHST09: 50-period ahead predictions. MSE of the experimental prices and forecasts, for Trend Extrapolation, Adaptive, Contrarian, Naive and Rational Expectations, Heuristic Switching Model and GA models (with $\beta \in [-1.1, 1.1]$ and $\beta \in [0, 1.1]$). MSE averaged over six negative feedback and six positive feedback groups.

(excluding the initialization phase)

$$MSE_X^{\text{prices}} = \frac{1}{47}\sum_{t=4}^{50}\left(p_t^{Gr\ X} - p_t^{GA}\right)^2,$$

$$MSE_X^{\text{price forecasts}} = \frac{1}{6\times 47}\sum_{i=1}^{6}\sum_{t=4}^{50}\left(p_{i,t}^{e,Gr\ X} - \hat{p}_{i,t}^{e,GA}\right)^2, \tag{12}$$

where $p_t^{Gr\ X}$ and $p_{i,t}^{e,Gr\ X}$ denote the realized price and the price forecast of subject i at period t in an experimental group X and p_t^{GA} and $\hat{p}_{i,t}^{e,GA}$ are the price and the price forecast of agent i at period t predicted by the GA model for the group X.

Table 2 reports MSE averaged over the six groups for each treatment, and over $1'024$ sample GA model paths per experimental group. We also include results for a number of benchmark models, including simple homogenous expectation rules, RE and HSM (two heuristics specification by Anufriev et al. (2013)).[12] In terms of the long-run, 50 period ahead, predictions, RE and two simple models, adaptive and contrarian expectations, perform the best under negative feedback, as they correctly predict the agents to converge to the fundamental price. Our GA model performs only slightly worse. Under positive feedback, RE, contrarian and adaptive expectations still predict convergence, in contrast to the experimental data. HSM, trend extrapolation and naive expectations perform relatively well, but surprisingly they are not better than RE. The reason is that the price oscillations predicted by these three models at the longer time horizon fall out of phase with the experimental oscillations. The best fit is achieved by our GA model, especially the one without contrarian rules $\beta \in [0, 1.1]$ (*c.f.* footnote 7). We conclude that all benchmark models are able to capture the long-run dynamics of possibly one feedback treatment, but not of two treatments at the same time. Only our GA model, in particular the specification without contrarian rules $\beta \in [0, 1.1]$, successfully

[12]For the definition of the benchmark rules, please refer to Appendix A.

MSE	Negative feedback		Positive feedback	
	Prices	**Forecasts**	**Prices**	**Forecasts**
Trend extr.	21.101	35.648	0.926	*4.196*
Adaptive	*2.3*	*14.912*	2.999	6.482
Contrarian	**2.249**	**14.856**	3.864	7.436
Naive	3.09	15.782	1.822	5.184
RE	2.571	15.21	46.835	54.811
HSM	2.999	17.106	0.889	**4.156**
GA: $\beta \in [-\mathbf{1.1}, \mathbf{1.1}]$	4.95	25.017	*0.806*	4.235
GA: $\beta \in [\mathbf{0}, \mathbf{1.1}]$	4.496	25.012	**0.802**	4.198

Table 3: HHST09: one-period ahead predictions. MSE of the experimental prices and forecasts, for the Trend Extrapolation, Adaptive, Contrarian, Naive and Rational Expectations, Heuristic Switching Model and Genetic Algorithms models (with $\beta \in [-1.1, 1.1]$ and $\beta \in [0, 1.1]$). MSE averaged over six negative feedback and six positive feedback groups.

predicts long-run behavior in both treatments.

3.4 One-period ahead predictions

A good indicator of the model's fit is the precision of its one-period ahead predictions (Anufriev and Hommes, 2012; Anufriev et al., 2013): how well the model predicts experimental outcomes in period $t+1$, conditional on the data until period t, in terms of MSE. For deterministic models such as HSM and the homogeneous expectations benchmark models, computing one period-ahead MSE is straightforward. For our GA model with its evolutionary operators, however, evaluating MSE is more complicated. Our model is both stochastic and highly non-linear: it evolves according to an analytically intractable period-to-period distribution. To address this issue, we compute the *expected* MSE using a simple Sequential Monte Carlo (SMC) approach.

Our SMC is designed in the following way. For each experimental group X, we run $1'024$ independent GA model simulations. In every simulation, we associate one GA agent with one subject, and in each period $t \geqslant 2$ every GA agent i (1) retains her heuristics from the previous period and (2) is given the *experimental* prices and the price forecasts of subject i until the previous period $t-1$. GA agents use the experimental data to update their heuristics and forecast the price p_t in the usual way, which gives us the GA's price forecasts (11) and realized prices (1) for period t. We evaluate the fit of the model to the experimental group by computing the average MSE (12) over all $1'024$ GA simulations.

The results are similar to the 50-period ahead simulations, see Table 3. Under negative feedback, RE, HSM adaptive, contrarian and naive expectations all capture the convergence of prices and forecasts to the fundamental price, slightly outperforming our GA model. Under positive feedback, these models (except for HSM) loose their predictive power and underestimate the experimental oscillatory behavior of individual forecasts. The GA model has the best fit for the positive feedback treatment and outperforms RE by a factor of 10.

These MC simulations show that our model captures both the aggregate and individual behavior in the LtF experiment reported by HHST09, both in terms of short and long-run dynamics. Moreover our model is the only one that captures the observed degree of heterogeneity of individual behavior between the experimental subjects, as measured by the coordination of the contemporary price forecasts (Figure 2c and 2c), together with the persistent heterogeneity of the forecasting heuristics.

4 Evidence from other experiments

Our GA model fits the HHST09 well. We will now move from the simple linear feedback to more complicated experimental settings. To be specific, we look at three other experiments that offer a hierarchy of challenges for the GA model of individual learning:

1. BHST12: linear feedback with large and unanticipated shocks to the fundamental price;
2. HSTV07; V01: nonlinear (cobweb) negative feedback economy, investigated with a GA model by Hommes and Lux (2013);
3. HSTV05: non-linear positive feedback economy, with two-period ahead predictions;

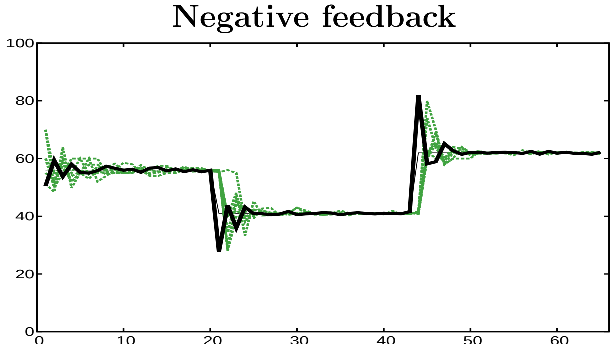
4.1 Large shocks to the fundamental price

BHST12 report a LtF experiment with the same structure as HHST09: positive and negative feedback with the linear structure given by (2) and the same dampening factor $|B| = \frac{20}{21}$. In this experiment however there are two large and unanticipated shocks to the fundamental price A: it changes from $p^f = 56$ to $p^f = 41$ in period $t = 21$ and then to $p^f = 62$ in period $t = 44$ until the last period $t = 65$.

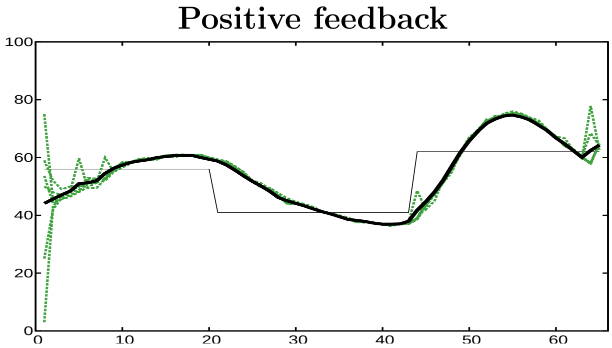
The results of BHST12 are similar to HHST09 and typical time paths are shown in Figure 4. Under negative feedback (Figure 4a), a shock to the fundamental breaks the subjects' coordination and is followed by a quick convergence to the new fundamental price. Under positive feedback (Figure 4b), shocks leave the coordination intact, and the predictions and prices move smoothly towards the new fundamental, eventually over- or undershooting it.

Figure 5 shows 65-period ahead MC simulations of prices, individual price forecasts and the degree of coordination.[13] Our model replicates well the experimental price dynamics for both treatments, as well as the impact of the shocks to the fundamental price on individual coordination. For the $\beta \in [-1.1, 1.1]$ specification, the 95% CI of our GA model contain 66% (84%) of the experimental prices and 84% (67%) of the standard deviation of individual forecasts under negative (positive) feedback. Overall, we can replicate around 75% of the experimental data with 65-period ahead simulations.

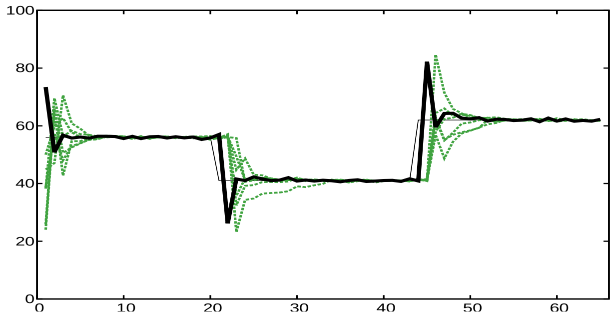
[13] We estimate the distribution of the initial predictions as in Diks and Makarewicz (2013) and sample directly from it, see Appendix B.



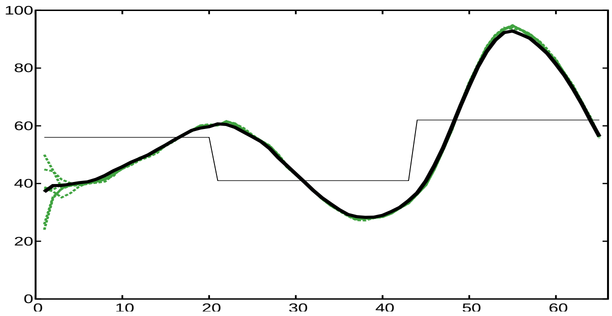
(a) Experiment group 8



(b) Experiment group 8



(c) Sample GA with $\beta \in [-1.1, 1.1]$



(d) Sample GA with $\beta \in [-1.1, 1.1]$

Figure 4: BHST12: experimental groups and sample 65-period ahead simulations of the GA model (with $\beta \in [-1.1, 1.1]$ and random initial predictions). Black line represents the price and green dashed lines are the individual predictions.

Figure 6 illustrates the time evolution of the price weight α and trend extrapolation coefficient β, which were chosen by the GA agents in the 65-period ahead simulations. The median behavior is similar to that from the experiment without large shocks HHST09, discussed in the previous section. Under negative feedback, the median GA agent learns the same adaptive expectations rule $p^e_{i,t+1} \approx 0.5p_t + 0.5p^e_{i,t}$. Under positive feedback, the median GA agent converges to a heuristic

$$p^e_{i,t+1} \approx 0.95p_t + 0.05p^e_{i,t} + 0.9(p_t - p_{t-1}), \tag{13}$$

which is a trend following rule with the trend extrapolation coefficient $\beta \approx 0.9$. This trend coefficient is significantly larger than the coefficient 0.6 in rule (9) used by the median GA agent under the positive feedback from the experiment without large shocks HHST09. The 95% CI for the trend extrapolation β becomes significantly positive towards the end of the experiment (see also Figure 10b for the histogram of β's chosen in period 65). Hence, *due to the large, unanticipated shocks in the positive feedback treatment, GA agents become strong trend followers.*

Table 4 reports the MSE for the 65-period ahead simulations initialized with the experimental initial predictions (1024 simulated markets per group for the GA models). We observe that the adaptive expectations have a good fit to negative feedback treatment, while naive expectations perform well under positive feedback. Interestingly, RE are poor for both treatments: they cannot explain oscillations of the positive feedback and the short spells of volatility

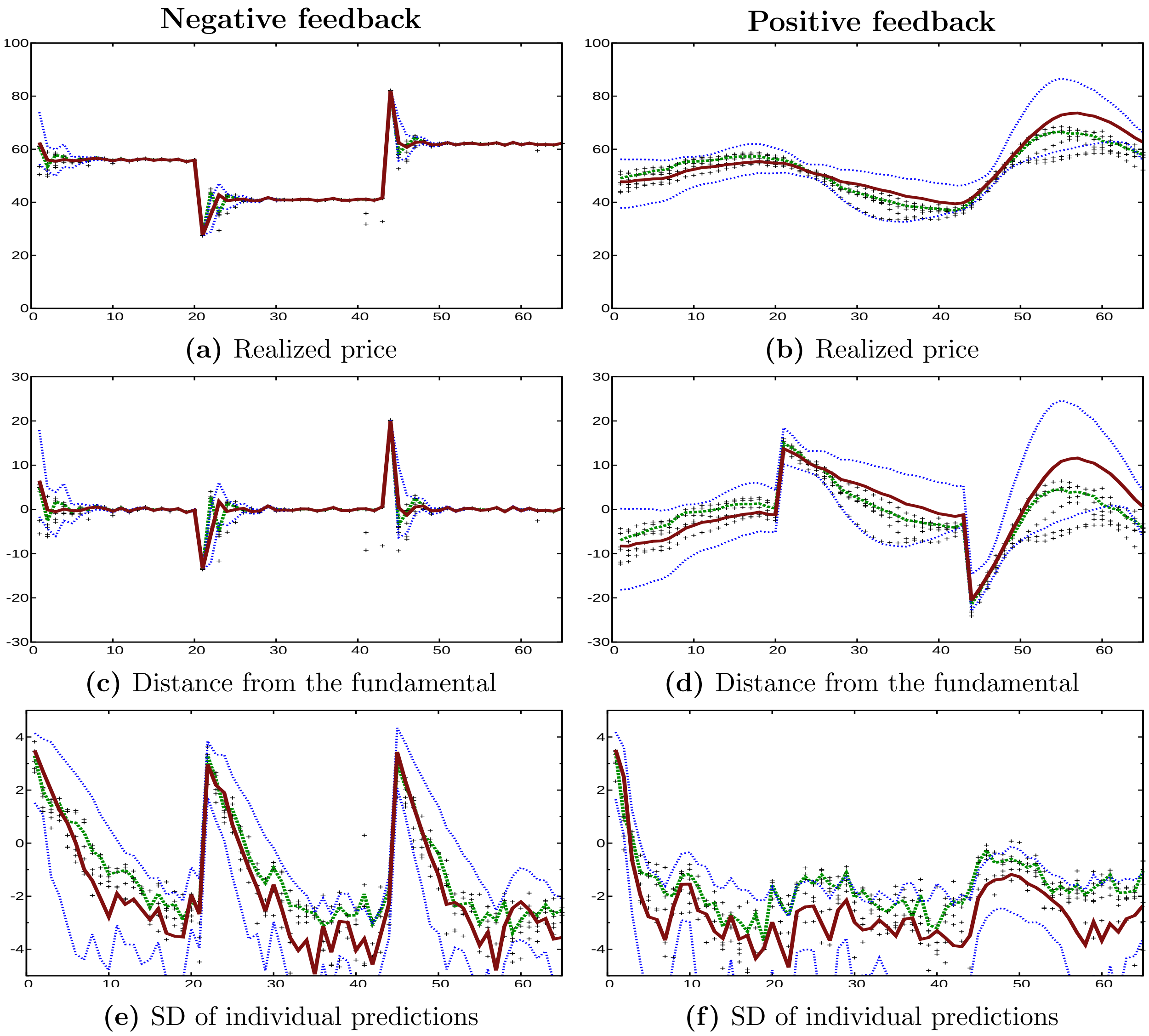


Figure 5: BHST12: 65-period ahead Monte Carlo simulation (1000 markets) for the GA model with $\beta \in [-1.1, 1.1]$. Realized price (top) and degree of coordination (standard deviation of individual predictions; bottom; $\log_2$ scale) over time. Green dashed line and black pluses represent the experimental median and group observations; red line is the median and blue dotted lines are the 95% confidence interval for the GA model. Left panel displays the negative feedback, right panel the positive feedback.

that follow the fundamental shocks under the negative feedback treatment. Also, HSM seems below average. In terms of long-run forecasting, our GA model is again second best for the negative feedback and the best for the positive feedback.

We also use the Sequential Monte Carlo (SMC) approach to compute the GA model's one-period ahead predicting power, reported in Table 5. The results are consistent with the 65-period ahead simulations. For both treatments, the GA model (especially without contrarian rules, $\beta \in [0, 1.1]$) is the best among all reported models.

4.2 Cobweb economy

HSTV07 and V01 report an LtF experiment in a Cobweb economy setting. HSTV07 inves-

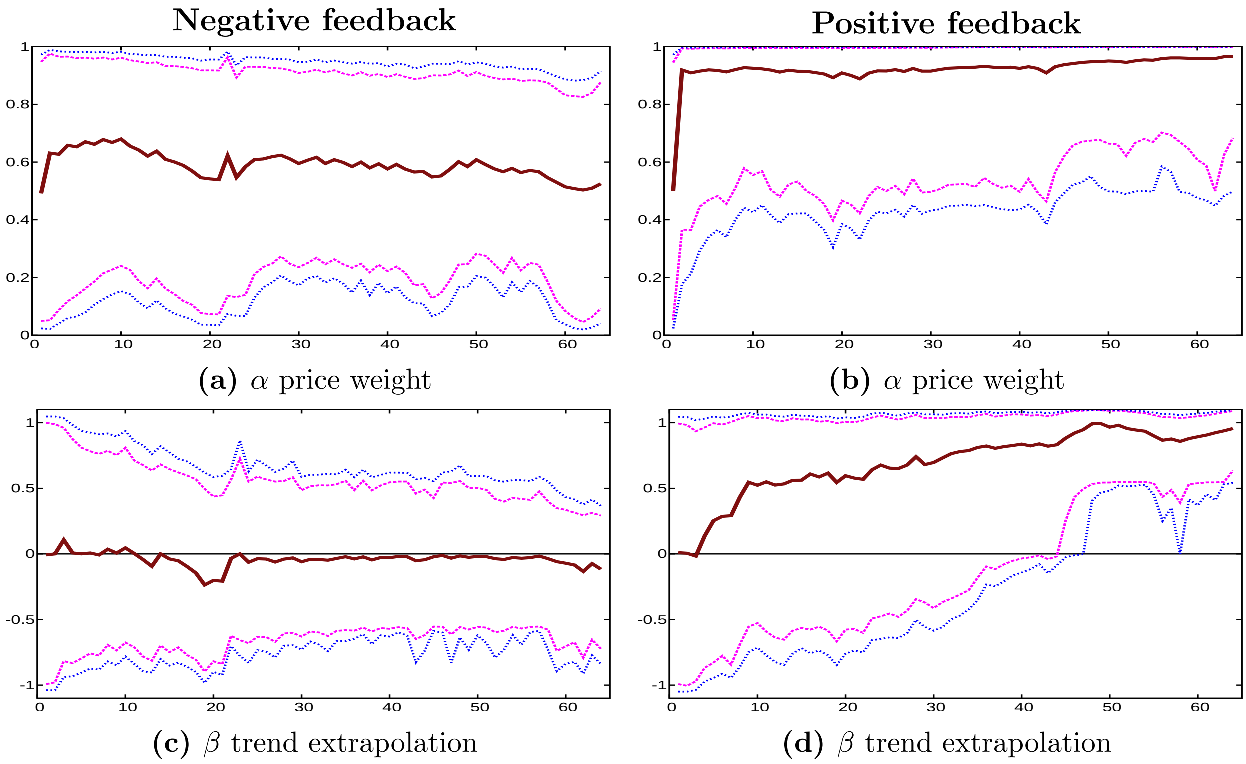


Figure 6: BHST12: 65-period ahead Monte Carlo simulation (1000 markets) for the GA model with $\beta \in [-1.1, 1.1]$. The price weight α and the trend extrapolation β *chosen* by the agents over time. Red line is the median, blue dotted lines are 95% CI, purple dashed are 90% CI for the GA model. Left panel displays the negative feedback, right the positive feedback.

	Negative feedback		Positive feedback	
MSE	**Prices**	**Forecasts**	**Prices**	**Forecasts**
Trend extr.	2736	1289	101.3	113.3
Adaptive	**3.629**	**10.75**	55	62.14
Contrarian	6.984	*14.45*	58.46	65.95
Naive	94.44	110.9	46.62	52.9
RE	13.871	20.923	55.133	60.859
HSM	73.57	87.86	90.8	101.8
GA: $\beta \in [\mathbf{-1.1}, \mathbf{1.1}]$	8.01	21.97	*43.49*	*49.44*
GA: $\beta \in [\mathbf{0}, \mathbf{1.1}]$	*6.333*	17.39	**43.49**	**49.64**

Table 4: BHST12: 65-period ahead predictions. MSE of the experimental prices and forecasts, for Trend Extrapolation, Adaptive, Contrarian, Naive and Rational Expectations, Heuristic Switching Model and GA models (with $\beta \in [-1.1, 1.1]$ and $\beta \in [0, 1.1]$). MSE averaged over eight negative feedback and eight positive feedback groups.

tigate 18 markets with six subjects each, divided into three treatments of 6 groups: with stable, unstable (on the verge of stability) and strongly unstable parametrization under the assumption of homogenous naive expectations. The experiment resulted in average prices very close to the RE fundamental price. However, the realized prices were excessively volatile, and – in contrast to positive feedback experiments – also non-persistent (with weak autocorrelation structure). Hommes and Lux (2013) study this experimental data set with a different

MSE	Negative feedback		Positive feedback	
	Prices	**Forecasts**	**Prices**	**Forecasts**
Trend extr.	114.061	121.329	1.183	2.165
Adaptive	**3.689**	**10.332**	3.776	4.618
Contrarian	5.92	*12.534*	4.737	5.559
Naive	9.979	16.81	2.411	3.286
RE	13.871	20.923	55.133	60.859
HSM	38.309	45.679	0.9996	**2.024**
Genetic Algorithm model				
$\beta \in [\mathbf{-1.1}, \mathbf{1.1}]$	10.247	21.464	*0.342*	2.059
$\beta \in [\mathbf{0}, \mathbf{1.1}]$	**4.208**	15.267	**0.341**	*2.036*

Table 5: BHST12: one-period ahead predictions. MSE of the experimental prices and forecasts, for the Trend Extrapolation, Adaptive, Contrarian, Naive and Rational Expectations, Heuristic Switching Model and Genetic Algorithms models (with $\beta \in [-1.1, 1.1]$ and $\beta \in [0, 1.1]$). MSE averaged over eight negative feedback and eight positive feedback groups.

GA model in which GA agents learn an anchor-and-adjustment (AR1) forecasting rule, *i.e.* $p^e_t = \alpha_{i,t} + \beta_{i,t} p_{t-1}$. It therefore constitutes an interesting benchmark case for our GA model.

As a first test for our model, we conduct a MC exercise in the vein of Hommes and Lux (2013). For each treatment, we compute *six* 50-period ahead simulations with different random numbers.[14] Next we compute the mean and standard deviation of the realized prices and the individual price forecasts. We repeat this procedure $1'000$ times and thus obtain a *distribution* (including 95% CI) of the realized means and variances of prices and price forecasts. We report the results in Table 6 for the two GA model specifications.

Our 50-period ahead simulations explain well the experimental data and perform significantly better than RE. The 95% CI of our GA model with $\beta \in [-1.1, 1.1]$ and $\beta \in [0, 1.1]$ replicate 12 and 11 out of 16 experimental statistics respectively (Table 6). Among the 11 cases successful for the GA model based on FOR rule (6) with $\beta \in [0, 1.1]$, 9 statistics reported by Hommes and Lux (2013) are outside 95% CI of our model. That means that we can replicate around three quarters of experimental descriptive statistics, most of which with a significantly higher precision than the GA model specification in Hommes and Lux (2013).[15]

We also check the 50-period ahead dynamics of the model conditional on the initial predictions from particular groups from HSTV07, see Table 7. Homogeneous expectation models, as well as HSM for the two unstable treatments are outperformed by RE. The dynamics of this experiment (in contrast to the linear experiments) resemble a white noise around the fundamental price. As a result, predicting the mean (as RE do) of these close-to-chaotic dynamics is better than trying to capture them with simplistic models. Only our GA model, in

[14] We estimate the distribution of the initial predictions as in Diks and Makarewicz (2013), see Appendix B.

[15] The GA model simulations are also closer to the experimental data in terms of the autocorrelation of the prices. RE always predicts zero autocorrelation, whereas benchmark models predict high autocorrelation up to the third lag. The experimental data exhibited weak autocorrelation, which is replicated by all three GA model specifications with comparable performance. See Table 12 in Appendix C for the results.

	Mean(p)	Var(p)	Mean(p^e)	Var(p^e)
	Stable			
Experiments	$5.64^{*\dagger}$	$0.36^{*\dagger}$	$5.56^{*\dagger}$	0.087^{*}
GA: AR1	5.565¶	0.326¶	5.576¶	0.1
GA: FOR $\beta \in [-1.1, 1.1]$	5.628	0.372	5.571	0.082
95% *CI*	[5.613, 5.643]	[0.359, 0.389]	[5.553, 5.59]	[0.065, 0.101]
GA: FOR $\beta \in [0, 1.1]$	5.649	0.353	5.548	0.0565
95% *CI*	[5.631, 5.667]	[0.341, 0.371]	[5.527, 5.57]	[0.043, 0.077]
	Unstable			
Experiments	$5.85^{\dagger}$	$0.63^{*\dagger}$	$5.67^{*\dagger}$	$0.101^{*\dagger}$
GA: AR1	5.817	0.647	5.645¶	0.16¶
GA: FOR $\beta \in [-1.1, 1.1]$	5.792	0.598	5.705	0.103
95% *CI*	[5.744, 5.841]	[0.525, 0.746]	[5.667, 5.739]	[0.067, 0.171]
GA: FOR $\beta \in [0, 1.1]$	5.825	0.557	5.694	0.079
95% *CI*	[5.786, 5.863]	[0.487, 0.658]	[5.67, 5.719]	[0.052, 0.122]
	Strongly unstable			
Experiments	$5.93^{\dagger}$	2.62^{*}	5.73	0.429^{*}
GA: AR1	6.2¶	2.161	5.434	0.769
GA: FOR $\beta \in [-1.1, 1.1]$	5.809	2.172	5.832	0.345
95% *CI*	[5.693, 5.908]	[1.626, 2.875]	[5.735, 5.918]	[0.182, 0.598]
GA: FOR $\beta \in [0, 1.1]$	5.962	1.487	5.807	0.206
95% *CI*	[5.876, 6.045]	[1.188, 1.834]	[5.75, 5.858]	[0.113, 0.347]
	Strongly unstable, group size 12			
Experiments	$5.937^{\dagger}$	1.783^{*}	$5.781^{*\dagger}$	$0.204^{*\dagger}$
GA: AR1	6.183¶	1.571	5.515¶	0.5¶
GA: FOR $\beta \in [-1.1, 1.1]$	5.812	1.699	5.852	0.194
95% *CI*	[5.731, 5.892]	[1.368, 2.157]	[5.779, 5.918]	[0.122, 0.338]
GA: FOR $\beta \in [0, 1.1]$	5.972	1.316	5.804	0.173
95% *CI*	[5.918, 6.026]	[1.118, 1.553]	[5.768, 5.843]	[0.111, 0.253]

Table 6: HSTV07: 50-period ahead MC results for GA simulations for four treatments, stable, unstable and strongly unstable (6 and 12 subjects). Average prices and predictions, and their variances for respectively experiment; GA simulations with AR1 rule (Hommes and Lux, 2013); GA simulations with first order rule (FOR) with or without contrarian rules (median statistics with 95% confidence intervals). $*$ and † denote experimental statistic which falls into 95% CI of GA FOR with $\beta \in [-1.1, 1.1]$ and $\beta \in [0, 1.1]$ respectively. ¶ denotes Hommes and Lux (2013) statistics which fall outside the 95% CI for GA model with $\beta \in [0, 1.1]$ when these CI contain the experimental statistics.

particular the one with $\beta \in [0, 1.1]$, keeps up with RE, and performs better than Hommes and Lux (2013) GA specification based on an AR1 rule.

The next MC exercise is the one-period ahead forecasting of the model with SMC approach for the 18 groups from HSTV07. Table 8 gives the summary results. It is apparent that the less stable the treatment, the worse fit has any model. As for the 50 period ahead forecasts, the clear winners are RE and our GA model, which are able to explain the data well also for

Treatments	Stable		Unstable		Strongly unstable	
MSE	Prices	Forecasts	Prices	Forecasts	Prices	Forecasts
Trend extr.	13.3	71.1	16.33	89.59	16.55	89.07
Adaptive	0.117	0.339	7.206	3.272	16.45	7.822
Contrarian	0.093	0.308	1.746	0.834	13.95	5.282
Naive	1.076	1.724	14.67	16.18	16.55	18.55
RE	*0.048*	**0.248**	0.364	**0.385**	**2.257**	**1.844**
HSM	0.178	0.422	7.446	3.431	16.46	7.885
GA: AR1	0.05742	0.3759	0.3552	0.6596	2.838	2.64
GA: $\beta \in [-\mathbf{1.1}, \mathbf{1.1}]$	0.088	0.356	*0.346*	0.631	3.445	3.261
GA: $\beta \in [\mathbf{0}, \mathbf{1.1}]$	**0.043**	*0.275*	**0.223**	*0.449*	*2.376*	*2.114*

Table 7: HSTV07: 50-period ahead predictions. MSE of the experimental prices and forecasts, for Trend Extrapolation, Adaptive, Contrarian, Naive and Rational Expectations, Heuristic Switching Model and GA models (FOR with $\beta \in [-1.1, 1.1]$ and $\beta \in [0, 1.1]$). MSE averaged over six groups for each treatment (stable, unstable, strongly unstable).

Treatments	Stable		Unstable		Strongly unstable	
MSE	Prices	Forecasts	Prices	Forecasts	Prices	Forecasts
Trend extr.	1.176	1.997	2.122	3.719	5.856	14.39
Adaptive	0.108	0.328	0.434	0.549	2.784	2.863
Contrarian	0.102	0.318	0.414	*0.497*	*2.929*	2.729
Naive	0.196	0.448	0.577	0.788	3.095	3.731
RE	**0.048**	**0.248**	*0.364*	**0.385**	**2.257**	**1.844**
HSM	0.212	0.474	0.52	0.732	3.065	3.691
GA: AR1	0.054	0.36	0.51	0.674	5.36	3.432
GA: $\beta \in [-\mathbf{1.1}, \mathbf{1.1}]$	0.13	0.393	0.866	0.795	5.547	3.25
GA: $\beta \in [\mathbf{0}, \mathbf{1.1}]$	*0.07*	*0.31*	**0.25**	0.531	3.079	*2.358*

Table 8: HSTV07: one-period ahead predictions. MSE of the experimental prices and forecasts, for the Trend Extrapolation, Adaptive, Contrarian, Naive and Rational Expectations, Heuristic Switching Model and Genetic Algorithms models (with $\beta \in [-1.1, 1.1]$ and $\beta \in [0, 1.1]$). MSE averaged over six groups for each treatment (stable, unstable, strongly unstable).

the strongly unstable treatment.[16] Our specification again prevails over the AR1 GA model of Hommes and Lux (2013).

We conclude that the cobweb experiments result in unstable, non-persistent prices, and simpler models like homogenous heuristics, but also HSM, miss-identify here any structure. As a result, their point predictions are so poor that it is better to predict the mean price, as in RE. Only our GA model (with $\beta \in [0, 1.1]$) comes close to RE in terms of this task. The GA model furthermore allows to explain the excess volatility observed in the experimental markets, which RE cannot account for. Finally, it is clear that the use of experimental micro-

[16]Notice that the scale of the prices in this experiment is $[0, 10]$ in contrast with the two previous settings, where the prices belonged to $[0, 100]$ intervals. The highest possible MSE in the linear experiments is 100 times higher than in the cobweb experiment.

foundations has an advantageous effect: our GA model, with an empirically motivated anchor and adjustment rule (6), has a better fit to the data than the AR1 specification used by Hommes and Lux (2013).

4.3 Two-period ahead asset pricing

HSTV05 report an experiment based on a *2-period ahead* non-linear positive feedback market; an asset-pricing model with a robotic fundamental trader, in which the *current* price depends on the subjects' average expectations about the price in the *next* period:

$$p_t = \frac{(1-n_t)\bar{p}^e_{t+1} + n_t p^f + y + \varepsilon_t}{1+r}, \tag{14}$$

where y is the mean dividend, r is the interest rate, $p^f = y/r$ is the fundamental price and n_t is the relative weight of the robotic trader. The latter acts as a stabilizing force on the market, by trading based on the fundamental price p^f and becoming more active if the price deviates from the fundamental according to

$$n_t = 1 - \exp\left(-\frac{1}{200}|p_{t-1} - p^f|\right). \tag{15}$$

There were two treatments with different fundamental price: seven markets were based on $p^f = 60$ and three on $p^f = 40$. Three different aggregate outcomes were observed: (i) monotonic convergence to the fundamental price (2 groups), (ii) dampened oscillations (3 groups) and (iii) volatile price oscillations (5 groups).

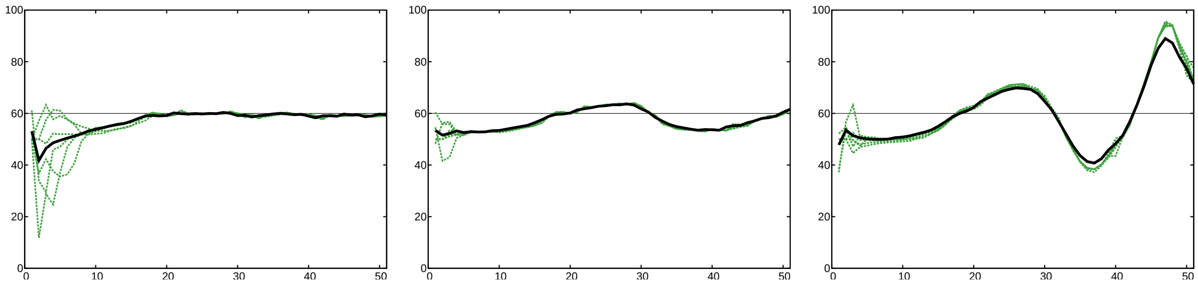
Notice that in this experiment the subject's decisions are based on a different information set than in the previous one-period ahead experiments. The two period ahead version of our GA model is be based on the following specification of the learned heuristic:

$$p^e_{i,h,t+1} = \alpha_{i,h} p_{t-1} + (1-\alpha_{i,h}) p^e_{i,t} + \beta_{i,h}(p_{t-1} - p_{t-2}). \tag{16}$$

Once p_t is realized, the agents can evaluate their rules based on the hypothetical performance of predicting p_t *two periods ago*. GA agents focus on $(p_t - p^e_{i,h,t})^2$, where $p^e_{i,h,t}$ is function of p_{t-2}, p_{t-3} and $p^e_{i,t-2}$. This specification is the most straightforward translation of the baseline one-period ahead forecasting heuristic (6).[17] In the baseline simulations, we look at the range of trend specified as before (with $\beta \in [-1.1, 1.1]$ and $\beta \in [0, 1.1]$). HSTV05 report that many of their subjects use very strong trend extrapolation, thus for the sake of completeness we will also report the results of our model with $\beta \in [-1.3, 1.3]$ and $\beta \in [0, 1.3]$.

In the seven groups with the fundamental price $p^f = 60$, HSTV05 observe groups which have converged to this fundamental, as well as groups with oscillations of different amplitude and frequency. Figure 7 displays three typical simulated markets of the GA model (with

[17] As for the linear positive feedback, specifications with a a positive weight for the fundamental price were overly stable, *c.f.* footnote 6.



(a) Convergence (b) Unclear (c) Oscillations

Figure 7: HSTV05: sample 50-period ahead simulations for GA model with $\beta \in [-1.3, 1.3]$ with different initial predictions and learning. The green lines are individual predictions, the black line is the realized price and the purple dashed line is the fundamental price.

MSE	**Prices**	**Forecasts**
Trend extr.	178.2	174.9
Adaptive	*96.12*	*145.9*
Contrarian	157	146.8
Naive	**95.29**	**144.6**
RE	96.0328	145.998
GA: $\beta \in [\mathbf{-1.1, 1.1}]$	103.9	155.8
GA: $\beta \in [\mathbf{0, 1.1}]$	114.9	169.1
GA: $\beta \in [\mathbf{-1.3, 1.3}]$	139.4	201.5
GA: $\beta \in [\mathbf{0, 1.3}]$	226.5	318.5

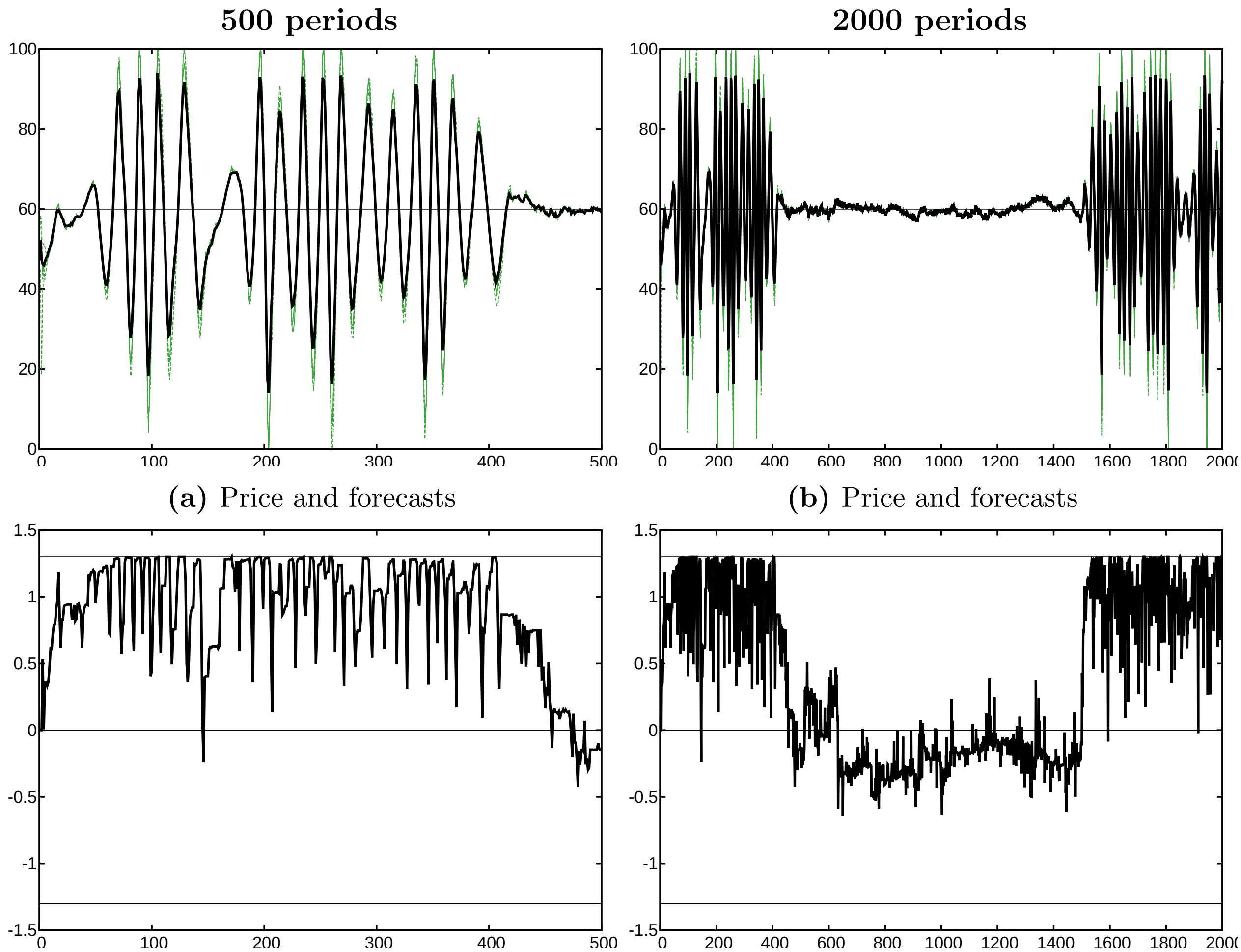
Table 9: HSTV05: 50-period ahead predictions. MSE of the experimental prices and forecasts, for Trend Extrapolation, Adaptive, Contrarian, Naive and Rational Expectations, Heuristic Switching Model and GA models with differently specified trend extrapolation coefficient β. MSE averaged over all experimental groups.

$\beta \in [-1.3, 1.3]$) for the HSTV05 economy (with $p^f = 60$).[18] The GA agents can either converge to the fundamental price (Figure 7a), or coordinate on unruly oscillations (Figure 7c). Furthermore, sometimes an intermediate outcome occurs. Figure 7b shows a sample simulation, in which the price seemingly stabilizes at the fundamental value around 20, but then resumes to oscillate mildly.

To further stress the volatile behavior of this market structure, we report one long run simulation for the GA model with $\beta \in [-1.3, 1.3]$. The top panel of Figure 8 displays its first 500 (Figure 8a) and 2′000 (Figure 8b) periods. Oscillations of different amplitude are persistent and can reappear even if the market settles on the fundamental price for some time, as seen in Figure 8b before period 800 and after 1500. This means that under the GA learning the fundamental price is not a unique attractor. Furthermore *clustered volatility* is clearly observed, with phases of stable price behavior interchangeable with highly volatile price fluctuations.

To explain this outcome, we take a closer look at the trend extrapolation coefficient β chosen by the GA agents. Figure 9 shows results for MC 50-period ahead simulations for two

[18]See Appendix B for initialization.



(a) Price and forecasts

(b) Price and forecasts

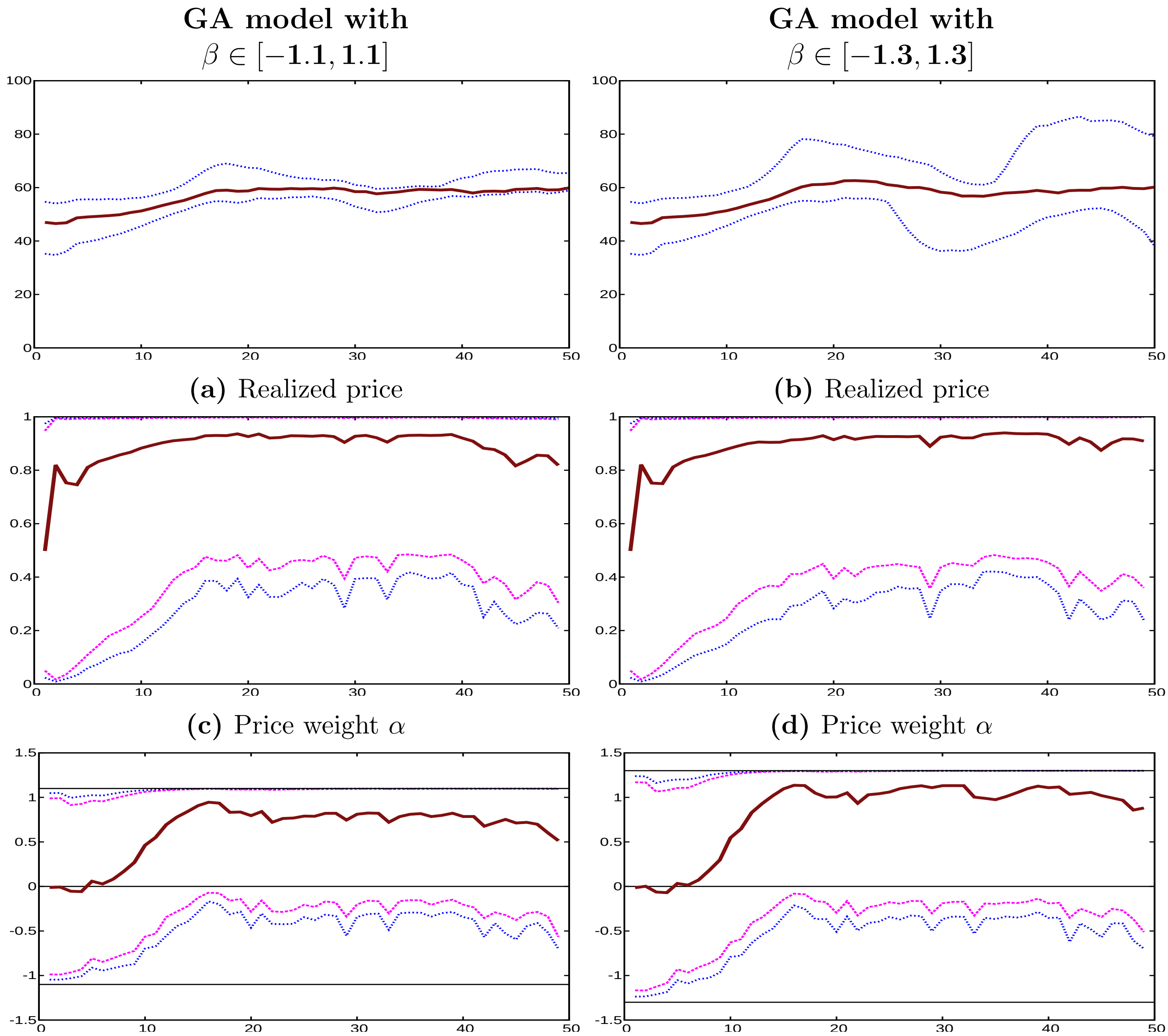
(c) Average chosen β

(d) Average chosen β

Figure 8: HSTV05: sample $2'000$-period ahead simulation (right panel) and its first 500 periods (left panel) of the GA model with $\beta \in [-1.3, 1.3]$ with fundamental price $p^f = 60$ and random initial predictions. Top panel are the individual predictions (green lines) together with the realized price (black line) and the fundamental price (the purple dashed line). Bottom panel is the average trend extrapolation coefficient β chosen by six GA agents.

GA model specifications, with $\beta \in [-1.1, 1.1]$ and $\beta \in [-1.3, 1.3]$. If the agents are allowed to experiment with higher β, the median price has a very similar oscillatory shape. The difference is seen in the 95% CI: both specifications are likely to generate two price bubbles within 50 periods, but the model with $\beta \in [-1.3, 1.3]$ has larger potential oscillations (Figure 9b), and the second bubble can be even bigger than the first (unlike in the linear positive feedback). In both specifications, the median GA agent converges to a strong trend extrapolation rule, close to $p^e_{i,t+1} = p_{t-1} + (p_{t-1} - p_{t-2})$, which is consistent with the behavior of our model in the previous experiments. Nevertheless, the 95% CI of the chosen trend coefficient remain wide and the distribution of this variable in period 50 (Figures 10c and 10d) is close to bi-modal, with a relatively large mass centered around zero (*i.e.* weak or no trend extrapolation).

This finding is in line with the trend extrapolation observed in the 2000 periods sample GA simulation. The bottom panel of Figure 8 shows the *average* β chosen by the six GA agents. Despite continuing instability, a clear pattern is that the average β remains close to zero in the stable phase of the simulation, but stays close to the upper limit of 1.3 in volatile



(a) Realized price **(b)** Realized price

(c) Price weight α **(d)** Price weight α

(e) Trend extrapolation β **(f)** Trend extrapolation β

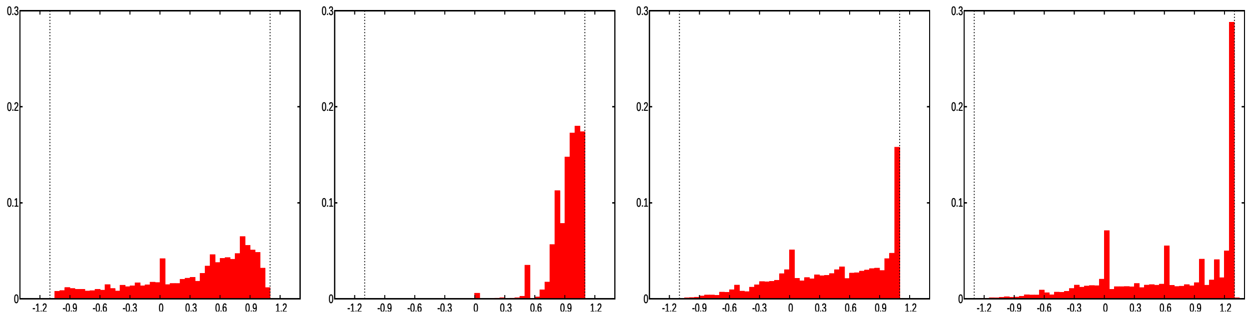
Figure 9: HSTV05: 50-period ahead Monte Carlo simulation (1000 markets) for the GA model with $\beta \in [-1.1, 1.1]$ (left panel) and $\beta \in [-1.3, 1.3]$ (right panel). The price weight α and the trend extrapolation β *chosen* by the agents over time. Red line is the median, blue dotted lines are 95% CI, purple dashed are 90% CI for the GA model.

times.

We interpret this finding in the following way. If the price is sufficiently stable and close to the fundamental value, the robotic fundamental trader is powerful enough to mitigate additional price deviations. This discourages GA agents from extrapolating an insignificant trend, and so price stability becomes self-reinforcing. However, if the trend in prices is sufficiently large, the stabilizing effect of the robotic trader can be counter-weighted by the GA agents periodically coordinating on strong trend extrapolation. The non-linear two-period ahead price feedback amplifies the realized price oscillations (which become self-reinforcing), but also allows their specific shape to be diversified. As a result, there is still space for the GA agents to experiment with the specific strength of trend following. In the two period ahead feedback system, our GA model thus entails two types of attractors: fundamental price and clustered volatility. This corresponds well to the diversified dynamics observed in the experiment.

MSE	Prices	Forecasts
Trend extr.	17.4527	55.0898
Adaptive	44.125	*25.3157*
Contrarian	59.3905	30.8646
Naive	31.6864	**20.8416**
RE	96.0328	145.998
HSM (4 heuristics)	*6.798*	—
GA: $\beta \in [\mathbf{-1.1, 1.1}]$	42.224	74.95
GA: $\beta \in [\mathbf{0, 1.1}]$	**5.934**	30.341
GA: $\beta \in [\mathbf{-1.3, 1.3}]$	21.192	53.238
GA: $\beta \in [\mathbf{0, 1.3}]$	16.29	42.125

Table 10: HSTV05: one-period ahead predictions. MSE of the experimental prices and forecasts, for Trend Extrapolation, Adaptive, Contrarian, Naive and Rational Expectations, 4-type Heuristic Switching Model (source: Anufriev and Hommes, 2012) and GA models with differently specified trend extrapolation coefficient β. MSE averaged over all experimental groups.



(a) HHST09 ($\beta \in [-1.1, 1.1]$) **(b)** BHST12 ($\beta \in [-1.1, 1.1]$) **(c)** HSTV05 ($\beta \in [-1.1, 1.1]$) **(d)** HSTV05 ($\beta \in [-1.3, 1.3]$)

Figure 10: Positive feedback treatments: HHST09, BHST12 and HSTV05 with $p^f = 60$: 50-period ahead predictions. Distribution of trend extrapolation coefficient β chosen by the agents in the last period $t = 50$ across the whole MC sample for each treatment, and two β specifications for HSTV05.

We note that the GA model does not capture coordination on relatively fast price oscillations, as was observed in some session of HSTV05. In order to improve the GA model's fit to the observed subject learning in this setup, one could experiment with higher order rules, but we leave this for future investigations.[19]

Even though our GA model leaves space for improvement, it is the only model which is comparatively good in predicting the experimental results of HSTV05 both in the long- and the short-run. Table 9 reports the MSE of 50-period ahead simulations initialized with the experimental initial predictions. These long-run predictive power is relatively poor for all models. The best three models are naive, adaptive and RE, though our model (with 1.1 as the upper bound for trend extrapolation) yields similar results. Table 10 shows the MSE of one-period ahead predictions for our GA model and other benchmark models. The

[19]The HSM based on four heuristics (adaptive, two different trend and anchor and adjustment expectations) has a good one-period ahead fit to these faster price oscillations (Anufriev and Hommes, 2012).

GA model is now among the best, especially in terms of predicting the experimental prices. Surprisingly, the models that did well in 50-period ahead predictions are poor now, while trend extrapolation is comparable with our model. Anufriev and Hommes (2012) investigated the HSTV05 experiment with a *four-heuristics* HSM, which is a richer model than the two-heuristic HSM we used as a benchmark for the previous experiments. Interestingly, only our GA model (specifically with $\beta \in [0, 1.1.])$ is able to compete with this richer HSM in terms of predicting the experimental prices.

5 Conclusions

In this paper we present a model in which agents independently use Genetic Algorithms (GA) to optimize a simple forecasting heuristic. The GA learning model is able to replicate individual, as well as aggregate behavior in four different Learning-to-Forecast experiments.

In these experiments, the realized market price depends on the forecasts of a group of subjects. This mimics many well studied economic environments, such as asset pricing markets or cobweb economies. These experiments can be used to study how human subjects try to adapt to the price-predictions feedback in a controlled environment. An important insight is that the market converges to the rational expectations equilibrium only if the relationship between the average price expectation and the realized price is negative (Heemeijer et al., 2009). In positive expectations feedback settings, subjects may coordinate on extrapolating observed price trends, which reinforces price oscillations (Hommes et al., 2005).

The most successful attempt to replicate these dynamics comes from the Heuristic Switching Models (Anufriev and Hommes, 2012). The main intuition of this approach is that among different prediction heuristics, agents gradually switch to heuristics having a relatively good past performance. On the other hand, Heuristic Switching Models cannot explain the full degree of observed individual heterogeneity, nor do they explain how the agents could learn their heuristics.

Genetic algorithms have been applied to solve complex optimization problems (Haupt and Haupt, 2004). Similar as in Hommes and Lux (2013) we use GAs to capture learning of a simple forecasting heuristics to study evolution of human decisions rules. Our GA agents use a linear first-order heuristic (a mixture of adaptive and trend extrapolating expectations) to forecast prices. They independently optimize the parametrization of their heuristics with Genetic Algorithms, thus learning to fine-tune their forecasting rules of thumb to the specific market conditions. This gives a parsimonious agent-based model of explicit learning-to-forecast with strong empirical motivation for the particular forecasting behavior.

We use our Genetic Algorithms model to investigate four Learning-to-Forecast experiments. The simple linear setting of the experiment reported by Heemeijer et al. (2009) enables us to set up the model. The experiment reported by Bao et al. (2012) adds large and unanticipated shocks to the basic linear structure of Heemeijer et al. (2009). The third experiment, reported by van de Velden (2001) and Hommes et al. (2007) focuses on a non-linear cobweb economy, and is an important benchmark already investigated by Hommes and Lux (2013). Finally, the asset pricing experiment reported by Hommes et al. (2005) introduces two-periods ahead feedback between the predictions and the realized prices.

We evaluate the out of sample one-period ahead and 50 period ahead prediction accuracy of our model in comparison with benchmark models: rational expectations, a number of simple homogenous expectations models (including adaptive and naive expectations) and the Heuristic Switching Model. To our best knowledge, this paper is the first to present an explicit evaluation of a full fledged agent-based model's fit to the experimental data to the individual and the aggregate level. For the difficult task of one-period ahead model predictions of the subjects' forecasts, we use a Sequential Monte Carlo technique. This is a novelty in the literature, which would rather focus on explaining the aggregate experimental outcomes.

Across the four different experiments, we observe a clear pattern in how different models can predict subject behavior. Rational and adaptive expectations tend to explain relatively well the negative type of price-expectations feedback, while naive and trend extrapolation expectations fit well the positive feedback markets. However, the explanatory power of these simple homogeneous benchmarks is limited to only a certain type of the Learning-to-Forecast experiments. This is where the strength of our Genetic Algorithms model lies, which is able to account for both the *aggregate* outcomes and the *individual* behavior across *different* experiments. When agents face a negative feedback type of economy, a median GA agent will increasingly rely on adaptive expectations, enforcing convergence of the market to the fundamental equilibrium[20].

In contrast, positive feedback induces the agents to follow the observed price trend and median forecasting behavior converges to a trend extrapolation rule, which amplifies price oscillations. The more 'complex' the positive feedback (in terms of shocks to the fundamental solution, or non-linear law of motion of the price), the stronger trend extrapolation will be chosen by the median agent and the more volatile will be the price fluctuations. In addition, our Genetic Algorithms model identifies two co-existing self-reinforcing regimes in the two-period ahead asset pricing market of HSTV05. If this market has already stabilized on the fundamental price, the agents have no space for trend extrapolation and learn self-fulfilling adaptive expectations, which enables the robotic trader to further keep the prices stable. However, once the agents coordinate on strong trend extrapolation heuristics, they are able to induce price oscillations, which in turn bolsters their learning of trend following. This *path-dependence* feature of GA learning is nicely illustrated by the bi-modal distributions of trend-extrapolation coefficients, with peaks around 0 and at the extreme 1.1 (or 1.3) (Figures 10c and 10d). Long-run simulation shows that our model can shift between these two self-reinforcing regimes, which corresponds well with the diversified experimental findings reported by HSTV05.

The strength of our model lies in its simplicity, parsimony and generality. Heterogeneous

[20]Only if the negative feedback is extremely strong and leads to strong overshooting, the market price may not converge but rather exhibit near-to-chaotic dynamics around the fundamental solution.

agents are using a simple, linear first-order forecasting heuristics, and adapt its (two) parameters to the current environment using a smart optimization procedure. Agents learn to use forecasting heuristics that make them smart (Gigerenzer and Todd, 1999) and the learning process selects different heuristics in different market environments. Furthermore the agent-based structure of the model allows for replicating the *individual* as well as *aggregate* behavior observed in these experiments, by a realistic account of heterogeneity and learning. The same agent-based GA model can be used to investigate settings with more complicated interactions between individual agents. This can include economies with heterogeneous preferences, unequal market power, information networks, decentralized price setting, etc.. In any of these cases, heterogeneous price expectations may have important consequences for market efficiency and price dynamics. Our Genetic Algorithms model gives a realistic explanation of how such heterogeneity between the agents emerges from their individual learning and, for each environment, which heuristics make them smart.

References

Adam, K. (2007). Experimental evidence on the persistence of output and inflation. *The Economic Journal*, 117(520):603–636.

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

Anufriev, M., Hommes, C. H., and Philipse, R. H. (2013). Evolutionary selection of expectations in positive and negative feedback markets. *Journal of Evolutionary Economics*, 23(3):663–688.

Arifovic, J. (1995). Genetic algorithms and inflationary economies. *Journal of Monetary Economics*, 36(1):219 – 243.

Arifovic, J. (1996). The behavior of the exchange rate in the genetic algorithm and experimental economies. *Journal of Political Economy*, 104(3):pp. 510–541.

Arifovic, J., Bullard, J., and Kostyshyna, O. (2012). Social learning and monetary policy rules. *The Economic Journal*.

Arifovic, J. and Ledyard, J. (2004). Scaling up learning models in public good games. *Journal of Public Economic Theory*, 6(2):203–238.

Assenza, T., Heemeijer, P., Hommes, C., and Massaro, D. (2013). Individual expectations and aggregate macro behavior. Technical report, Tinbergen Institute Discussion Paper.

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

Blundell, R. and Stoker, T. M. (2005). Heterogeneity and aggregation. *Journal of Economic Literature*, 43(2):pp. 347–391.

Brock, W. A. and Hommes, C. H. (1997). A rational route to randomness. *Econometrica*, 65(5):pp. 1059–1095.

Bullard, J. (1994). Learning equilibria. *Journal of Economic Theory*, 64(2):468 – 485.

Case, K. E. and Shiller, R. J. (2003). Is there a bubble in the housing market? *Brookings Papers on Economic Activity*, 2003(2):pp. 299–342.

Charness, G., Karni, E., and Levin, D. (2007). Individual and group decision making under risk: An experimental study of bayesian updating and violations of first-order stochastic dominance. *Journal of Risk and Uncertainty*, 35:129–148.

Dawid, H. (1996). *Adaptive learning by genetic algorithms: Analytical results and applications to economic models*. Springer-Verlag New York, Inc.

Dawid, H. and Kopel, M. (1998). On economic applications of the genetic algorithm: a model of the cobweb type. *Journal of Evolutionary Economics*, 8:297–315. 10.1007/s001910050066.

Diks, C. and Makarewicz, T. (2013). Initial predictions in learning-to-forecast experiment. In *Managing Market Complexity*, volume 662 of *Lecture Notes in Economics and Mathematical Systems*, pages 223–235. Springer Berlin Heidelberg. 10.1007/978-3-642-31301-1_18.

Doornik, J. (2007). *Object-oriented matrix programming using Ox*. Timberlake Consultants Press, London, 3rd edition.

Eren, S., Jiménez Martín, S., Heiland, F., and Benítez-Silva, H. (2008). How well do individuals predict the selling prices of their homes? Documentos de trabajo (FEDEA) 10:1-32.

Evans, G. and Honkapohja, S. (2001). *Learning and expectations in macroeconomics*. Princeton University Press.

Evans, G. and Ramey, G. (2006). Adaptive expectations, underparameterization and the lucas critique. *Journal of Monetary Economics*, 53(2):249–264.

Gigerenzer, G. and Todd, P. M. (1999). *Simple heuristics that make us smart*. Oxford University Press.

Goodman Jr., J. L. and Ittner, J. B. (1992). The accuracy of home owners' estimates of house value. *Journal of Housing Economics*, 2(4):339 – 357.

Grandmont, J.-M. (1985). On endogenous competitive business cycles. *Econometrica*, 53(5):pp. 995–1045.

Grandmont, J.-M. (1998). Expectations formation and stability of large socioeconomic systems. *Econometrica*, 66(4):pp. 741–781.

Haupt, R. and Haupt, S. (2004). *Practical Genetic Algorithms*. John Wiley & Sons, Inc., New Jersey, 2nd edition.

Heemeijer, P., Hommes, C., Sonnemans, J., and Tuinstra, J. (2009). Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation. *Journal of Economic Dynamics and Control*, 33(5):1052–1072.

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

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

Hommes, C. and Lux, T. (2013). Individual expectations and aggregate behavior in learning to forecast experiments. *Macroeconomic Dynamics*, 17(2):373–401.

Hommes, C., Sonnemans, J., Tuinstra, J., and van de Velden, H. (2007). Learning in cobweb experiments. *Macroeconomic Dynamics*, 11(Supplement S1):8–33.

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

Lucas Jr, R. (1986). Adaptive behavior and economic theory. *Journal of Business*, pages S401–S426.

Lucas Jr., R. E. (1972). Expectations and the neutrality of money. *Journal of Economic Theory*, 4(2):103–124.

Lux, T. and Schornstein, S. (2005). Genetic learning as an explanation of stylized facts of foreign exchange markets. *Journal of Mathematical Economics*, 41(1–2):169 – 196. Special Issue on Evolutionary Finance.

Malmendier, U. and Nagel, S. (2009). Learning from inflation experiences. Available online at `http://faculty-gsb.stanford.edu/nagel/documents/InflExp.pdf`.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica*, 29(3):pp. 315–335.

Nunes, R. (2010). Inflation dynamics: The role of expectations. *Journal of Money, Credit and Banking*, 42(6):1161–1172.

Pfajfar, D. and Zakelj, B. (2011). Inflation expectations and monetary policy design: Evidence from the laboratory. Technical report, CentER Discussion Paper Series.

Reinhart, C. M. and Rogoff, K. (2009). *This time is different: eight centuries of financial folly*. princeton university press.

Sims, C. A. (2010). Rational inattention and monetary economics. *Handbook of Monetary Economics*, 3:155–181.

Tuinstra, J. and Wagener, F. (2007). On learning equilibria. *Economic Theory*, 30:493–513. 10.1007/s00199-005-0059-1.

van de Velden, H. (2001). *An experimental approach to expectation formation in dynamic economic systems*. PhD thesis, Tinbergen Institute and Universiteit van Amsterdam.

Wagener, F. (2013). Expectations in experiments. Technical report, Tinbergen Institute Discussion Paper.

Woodford, M. (2013). Macroeconomic analysis without the rational expectations hypothesis. Technical report, National Bureau of Economic Research.

Appendices

A Definition of forecasting rules

Table 11 gives the exact specification for all the prediction rules used in the one-period ahead forecasting exercises for the four experiments. The full HSM specification can be found in Anufriev et al. (2013).

B Initialization of the model

In this appendix we discuss the initialization of the GA model for the 50-period Monte Carlo simulations, which we use to show that our model replicates experimental stylized facts. Initialization is crucial, since in the experiments the initial individual predictions influenced later outcomes, such as appearance and characteristics of oscillations, or dynamics of coordination. Two examples can be given for HHST09. Under negative feedback, the individual price forecasts coordinated only after the price itself has already converged; in our simulations we want to start with a similar degree of non-coordination between the agents, to show that disappears in the same way as happened in the experiment. Anufriev et al. (2013) suggest that under positive feedback, price oscillations require the subjects to start relatively far from the fundamental price, as was also the case for their HSM. Therefore, proper distribution of initial predictions of the experimental subjects is a crucial aspect of model calibration; without a realistic initialization, the model will not fit the data well.

Diks and Makarewicz (2013) investigate this issue in a systematic fashion for the case of the HHST09 experiment. They argue that the initial subject predictions can be regarded as a sample from a common distribution, which they next estimate. We use their methodology to

Rule	**Prediction**
	Homogeneous rules
Trend extr.	$p^e_t = p_{t-1} + (p_{t-1} - p_{t-2})$
Adaptive	$p^e_t = 0.75p_{t-1} + 0.25p^e_{t-1}$
Contrarian	$p^e_t = p_{t-1} - 0.5(p_{t-1} - p_{t-2})$
Naive	$p^e_t = p_{t-1}$
RE	$p^e_t = p^f$
	Heterogeneous rules
HSM	two heuristic model (trend extrapolation vs. adaptive expectations)
GA model	$p^e_{i,t} = \alpha_{i,t}p_{t-1} + (1-\alpha_{i,t})p^e_{i,t-1} + \beta_{i,t}(p_{t-1} - p_{t-2})$
$\beta \in [\mathbf{-1.1}, \mathbf{1.1}]$	$\alpha_{i,t} \in [0,1]$ and $\beta_{i,t} \in [-1.1, 1.1]$
$\beta \in [\mathbf{0}, \mathbf{1.1}]$	$\alpha_{i,t} \in [0,1]$ and $\beta_{i,t} \in [0, 1.1]$

Table 11: Specification of the forecasting rules p^e_t for one-period ahead forecasting environment.

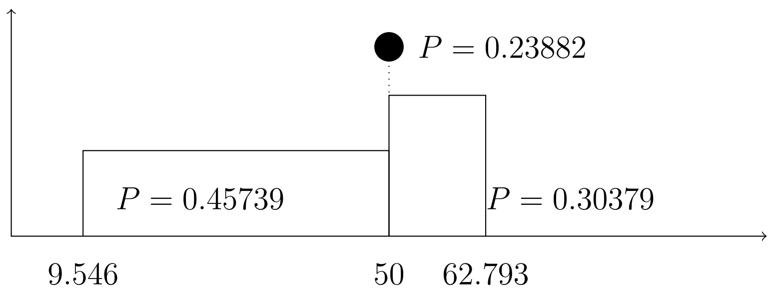


Figure 11: Density function of winged focal point distribution for HHST09. Initial prediction will be equal to $p^e_{i,1} = 50$ with probability $P = 0.30379$ (mass point); with probability $P = 0.45739$ it will fall into the left wing, where its value is drawn from $Uniform(9.546, 50)$; with probability $P = 0.23882$ it will fall into the right wing, where its value is drawn from $Uniform(50, 62.793)$. The size of the wings is scaled to their masses and lengths.

calibrate the initial period of our model to all the other experiments, that we investigate for our GA model. In each MC simulation, we sample the initial predictions from the distribution calibrated to the respective experimental data.

HHST09

For this experiment we use the estimated Winged Focal Point (WFP) reported by Diks and Makarewicz (2013), which is given by

$$p^e_{i,1} = \begin{cases} \varepsilon^1_i \sim U(9.546, 50) & \text{with probability } 0.45739, \\ 50 & \text{with probability } 0.30379, \\ \varepsilon^2_i \sim U(50, 62.793) & \text{with probability } 0.23882. \end{cases} \tag{17}$$

With WFP we replicate the observed behavior of the subjects in the first period. Around 1/3 would predict 50, a mid-point of the suggested interval for the initial price forecast $[0, 100]$. Others were evenly spread around this focal point, with more people choosing < 50 and almost nobody choosing > 60. Hence the distribution is a composite of a unit mass at 50 and two 'wings', uniform distributions preading from the focal point. ee Figure 11 for a visualization of the density function for this distribution.

BHST12

We reestimate WFP model for the data reported by BHST12 using the same methodology as reported by Diks and Makarewicz (2013). This leads to WFP specified as

$$p^e_{i,1} = \begin{cases} \varepsilon^1_i \sim U(16.406, 50) & \text{with probability } 0.32296, \\ 50 & \text{with probability } 0.35159, \\ \varepsilon^2_i \sim U(50, 70.312) & \text{with probability } 0.32296. \end{cases} \tag{18}$$

HSTV07; V01

In the case of the cobweb economy experiment, the subjects were asked to predict prices in the $[0, 10]$ interval. Interestingly, the initial predictions still have the WFP form, with a large proportion equal to the midpoint 5 and the rest (not necessarily rounded to a full integer) distributed around this new focal point. To account for that, we reestimate the WFP and obtain

$$p^e_{i,1} = \begin{cases} \varepsilon^1_i \sim U(1.875, 5) & \text{with probability } 0.17983, \\ 5 & \text{with probability } 0.36344, \\ \varepsilon^2_i \sim U(5, 7.5) & \text{with probability } 0.45673. \end{cases} \tag{19}$$

HSTV05

In this experiment, the predictions are two-period ahead, hence the subjects would have to give *two* initial predictions, $p^e_{i,1}$ and $p^e_{i,2}$. First period forecasts are similar to those from the other experiments. As for the second period, one can notice that 2/3 of the subjects, who would predict $p^e_{i,1} = 50$ the focal point in the first period, would do the same in the second period; otherwise they would again draw predictions resembling WFP, but with a substantially small weight on the focal point 50. Hence we follow Diks and Makarewicz (2013) and get the following estimations for the first period:

$$p^e_{i,1} = \begin{cases} \varepsilon^1_i \sim U(4.712, 50) & \text{with probability } 0.31306, \\ 5 & \text{with probability } 0.45536, \\ \varepsilon^2_i \sim U(50, 64.062) & \text{with probability } 0.23158. \end{cases} \tag{20}$$

Define the auxiliary draw

$$p^{aux}_{i,2} = \begin{cases} \varepsilon^1_i \sim U(3.125, 50) & \text{with probability } 0.44958, \\ 5 & \text{with probability } 0.018761, \\ \varepsilon^2_i \sim U(50, 67.227) & \text{with probability } 0.53166. \end{cases} \tag{21}$$

Thus, the second period predictions are given by

$$p^e_{i,2} = \begin{cases} p^{aux}_{i,2} & \text{always if } p^e_{i,1} \neq 50, \\ p^{aux}_{i,2} & \text{with probability } 1/3 \text{ if } p^e_{i,1} = 50, \\ 50 & \text{with probability } 2/3 \text{ if } p^e_{i,1} = 50. \end{cases} \tag{22}$$

C Price autocorrelation in the cobweb experiment

Table 12 gives the first three autororrelations of the experimental groups in HSTV07 and the 50-period ahead simulations of the GA and benchmrak models.

Treatments	**Stable**			**Unstable**			**Strongly unstable**		
Autocorrelation	ρ_1	ρ_2	ρ_3	ρ_1	ρ_2	ρ_3	ρ_1	ρ_2	ρ_3
Experiment	−0.1878	0.06323	−0.12	−0.2948	0.01363	−0.01114	−0.1973	0.211	0.02144
Trend extr.	−0.9661	0.9423	−0.9209	−0.9655	0.9404	−0.9159	−0.9639	0.9403	−0.918
Adaptive	−0.5996	0.3446	−0.3078	−0.9628	0.9235	−0.8927	−0.964	0.94	−0.9176
Contrarian	−0.257	−0.3006	0.1604	−0.4556	−0.4895	0.8202	−0.4756	−0.4704	0.8974
Naive	−0.9043	0.837	−0.7911	−0.967	0.9394	−0.9143	−0.9639	0.9403	−0.918
RE	0	0	0	0	0	0	0	0	0
HSM	−0.6528	0.4224	−0.3438	−0.9561	0.9153	−0.8816	−0.9639	0.9399	−0.9175
GA: AR1	−0.1161	0.008603	−0.1253	−0.1686	−0.005697	−0.1028	−0.2346	−0.09282	−0.02373
GA: $\beta \in [-1.1, 1.1]$	−0.1102	−0.3232	0.002674	−0.2201	−0.2013	0.0362	−0.2478	−0.3148	0.2432
GA: $\beta \in [0, 1.1]$	−0.2955	0.1059	−0.171	−0.3882	0.1405	−0.1848	−0.6206	0.4428	−0.356

Table 12: HSTV07: 50-period ahead predictions. First three autocorrelations in the experimental groups, and in the 50-period ahead simulations of the Trend Extrapolation, Adaptive, Contrarian, Naive and Rational Expectations, Heuristic Switching Model and GA models (FOR with $\beta \in [-1.1, 1.1]$ and $\beta \in [0, 1.1]$). Autocorrelations averaged over six groups for each treatments (stable, unstable and strongly unstable).

D Formal definition of Genetic Algorithms Not for publication

In this appendix we present a formal definition of the Genetic Algorithms (GA) version, which served as the cornerstone of our model. It closely follows the standard specification suggested by Haupt and Haupt (2004) and used by Hommes and Lux (2013).

D.1 Optimization procedures: traditional and Genetic Algorithms

Consider a maximization problem where the target function $\mathcal{F}$ of N arguments $\theta = (\theta^1, \ldots, \theta^N)$ is such that a straightforward analytical solution is unavailable. Instead, one needs to use a numerical optimization procedure.

Traditional maximization algorithms, like the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm, iterate a candidate argument for the optimum of the target function $\mathcal{F}$ by (1) estimating the curvature around the candidate and (2) using this curvature to find the optimal direction and length of the change to the candidate solution. This so called 'hill-climbing' algorithm is very efficient in its use of the shape of the target function. On the other hand, it will fail if the target function is 'ill-behaved': non-continuous or almost flat around the optima, has kinks or breaks. Here the curvature cannot be reliably estimated. Another problem is that the BFGS may perform poorly for a problem of large dimensionality.

The Genetic Algorithms are based on a fundamentally different approach and therefore can be used for a wider class of problems. The basic idea is that we have a population of arguments which compete *only* in terms of their respective function value. This competition is modeled in an evolutionary fashion: mutation operators allow for a blind-search experimentation, but the probability that a particular candidate will survive over time is relative to its functional value. As a result, the target function may be as general as necessary, while the arguments can be of any kind, including real numbers, integers, probabilities or binary variables. The only constraint is that each argument must fall into a predefined dense interval a_n, b_n.

D.2 Binary strings

A Genetic Algorithm (GA) uses H chromosomes $g_{h,t} \in \mathbb{H}$ which are binary strings divided into N genes $g^n_{h,t}$, each encoding one candidate parameter $\theta^n_{h,t}$ for the argument θ^n. A chromosome $h \in \{1, \ldots, H\}$ at time $t \in \{1, \ldots, T\}$ has predetermined length L and is specified as

$$g_{h,t} = \{g^1_{h,t}, \ldots, g^N_{h,t}\}, \tag{23}$$

such that each gene $n \in \{1, \ldots, N\}$ has its length equal to an integer L_n (with $\sum_{n=1}^{N} L_n = L$) and is a string of binary entries (bits)

$$g^n_{h,t} = \{g^{n,1}_{h,t}, \ldots, g^{n,L_n}_{h,t}\}, \quad g^{n,l}_{h,t} \in \{0, 1\} \text{ for each } j \in \{1, \ldots, L_n\}. \tag{24}$$

The relation between the genes and the arguments is straightforward. An integer θ^n is simply encoded by (24) with its binary notation. Consider now an argument θ^n which is a probability. Notice that $\sum_{l=0}^{L_n-1} 2^j = 2^{L_n} - 1$. It follows that a particular gene $g^n_{h,t}$ can be decoded as a normalized sum

$$\theta^n_{h,t} = \sum_{l=1}^{L_n} \frac{g^{n,l}_{h,t} 2^{l-1}}{2^{L_n} - 1}. \tag{25}$$

A gene of zeros only is therefore associated with $\theta_n = 0$, a gene of ones only – with $\theta_n = 1$, while other possible binary strings cover the $[0, 1]$ interval with an $\frac{1}{2^{L_n}-1}$ increment. Any desired precision can be achieved with this representation. Since $2^{-10} \approx 10^{-3}$, the precision close to one over trillion (10^{-12}) is obtained by a mere of 40 bits.

A real variable θ^n from an $[a_n, b_n]$ interval can be encoded in a similar fashion, by a linear transformation of a probability:

$$\theta^n_{h,t} = a_n + (b_n - a_n) \sum_{l=1}^{L_n} \frac{g^{n,l}_{h,t} 2^{l-1}}{2^{L_n} - 1} \tag{26}$$

where the precision of this representation is given by $\frac{b_n-a_n}{2^{L_n}-1}$. Notice that one can approximate an unbounded real number by reasonably large a_n or b_n, since the loss of precision is easily undone by a longer string.

D.3 Evolutionary operators

The core of GA are evolutionary operators. GA iterates the population of chromosomes for T periods, where T is either large and predefined, or depends on some convergence criterion. First, at each period $t \in \{1, \ldots, T\}$ each chromosome has its fitness equal to a monotone transformation of the function value $\mathcal{F}$. This transformation is defined as $V(\mathcal{F}(\theta_{h,t})) \equiv V(h_{k,t}) \rightarrow \mathbb{R}^+ \cap \{0\}$. For example, a non-negative function can be used directly as the fitness. If the problem is to minimize a function, a popular choice is the exponential transformation of the function values, similar to the one used in the logit specification of the Heuristic Switching Model (Brock and Hommes, 1997).

Chromosomes at each period can undergo the following evolutionary operators: procreation, mutation, crossover and election. These operators first generate an offspring population of chromosomes from the parent population t and therefore transform both populations into a new generation of chromosomes $t + 1$ (notice the division of the process).

D.3.1 Procreation

For the population at time t, GA picks subset $\mathbb{X} \subseteq \mathbb{H}$ of χ chromosomes and picks $\kappa < \chi$ of them into a set $\mathbb{K}$. The probability that the chromosome $h \in \mathbb{X}$ will be picked into $\mathbb{K}$ as its

zth element (where $z \in \{1, \dots, \kappa\}$) is usually defined by the power function:

$$Prob(g_z = g_{h,t}) = \frac{V(g_{h,t})}{\sum_{j \in \mathbb{X}} V(g_{j,t})}. \tag{27}$$

This procedure is repeated with differently chosen $\mathbb{X}$'s until the number of chromosomes in all such sets $\mathbb{K}$'s is equal to H. For instance, the *roulette* is procreation with $\chi = H$ and $\kappa = 1$: GA picks randomly one chromosome from the whole population, where each chromosome has probability of being picked equal to its function value relative to the function value of all other chromosomes. This is repeated exactly H times.

So called *tournaments* are often used for the sake of computational efficiency. Here, $\chi << H$. For instance, GA could divide the chromosomes into pairs and sample two offspring from each pair.

Procreation is modeled as the basic natural selection mechanism. We consider subsets of the original population (or maybe the whole population at once). Out of each such a subset, we pick a small number of chromosomes, giving advantage to these which perform better. We repeat this procedure until the offspring generation is as large as the old one. Thus the new generation is likely to be 'better' than the old one.

D.3.2 Mutation

For each generation $t \in \{1, \dots, T\}$, after the procreation has taken place, each binary entry in each new chromosome has a predefined δ_m probability to mutate: ones turned into zeros and *vice versa*. In this way the chromosomes represent different numbers and may therefore attain better fit.

The mutation operator is where the binary representation becomes most useful. If the bits, which are close to the beginning of the gene, mutate, the new argument will be substantially different from the original one. On the other hand, small changes can be obtained by mutating bits from the end of the gene. Both changes are equally likely! In this way, GA can easily evaluate arguments which are both far away from and close to what the chromosomes are currently encoding. As a result, GA efficiently converges to the maximum, but are also likely *not* to get stuck on a local maximum. This is clearly independent of the initial conditions, which gives GA additional advantage over hill-climbing algorithms (like BFGS), where a good choice of the initial argument can be crucial to obtain the global maximum.

D.3.3 Crossover

Let $0 \leqslant C_L, C_H \leqslant \sum_{n=1}^{N} L_n = L$ be two predefined integers. The crossover operator divides the population of chromosomes into pairs. If $C_L < L - C_H$, it exchanges the first C_L and the last C_H bits between chromosomes in each pair with a predefined probability δ_c. Otherwise, the crossover operator exchanges $max\{C_L, C_H\}$ bits in each pair of chromosomes with this predefined probability δ_c. This operator facilitates experimentation in a different way than the mutation operator. Typically, it is set to exchange whole arguments, that is there are $0 \leqslant \nu_L \leqslant \nu_H \leqslant N$ such that $C_L = \sum_{n=1}^{\nu_L} L_n$ and $C_H = \sum_{n=\nu_H}^{N} L_n$. This allows the chromosomes to experiment with different compositions of the individual arguments, which on their own are already successful.

D.3.4 Election

The experimentation done by the mutation and crossover operators does not need to lead to efficient binary sequences. For instance, a chromosome which actually decodes the optimal argument should not mutate at all. To counter this effect, it is customary to divide the creation of a new generation into two stages. First, the chromosomes procreate and undergo mutation and crossover in some predefined order. Next, the resulting set of chromosomes is compared in terms of fitness with the parent population. Thus, offspring will be passed to the new generation only if it *strictly* outperforms the parent chromosome. In this way each generation will be at least as good as the previous one, what in many cases facilitates convergence.

E Parametrization of the forecasting heuristic Not for publication

In this appendix, we will address two issues. First, following HHST09 we will look on the importance of the anchor, for the said experiment and the two-period ahead Hommes et al. (2005) setting. Second, we study the proper degree of allowed trend extrapolation, based on he linear feedback from HHST09.

E.1 Is the anchor important for HHST09?

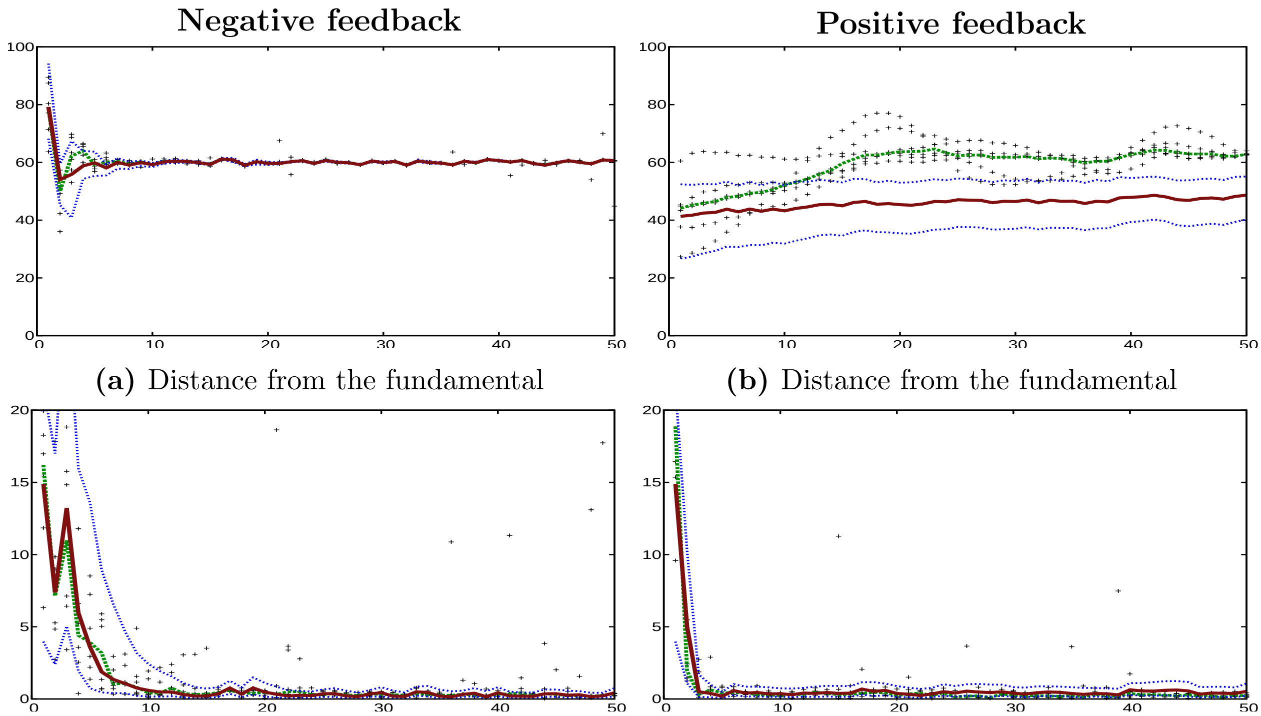
HHST09 show that most of their subjects (around 60%) use First-Order prediction rule with heterogeneous parameter specification:

$$p^e_{i,t} = \alpha_1 p_{t-1} + \alpha_2 p^e_{i,t-1} + \alpha_3 60 + \beta(p_{t-1} - p_{t-2}) \tag{28}$$

where the fundamental price 60 serves as an anchor,[21] the three α_i span a simplex and β is the trend extrapolation coefficient. Our rule (6) is a special case of (28) with the restriction that $\alpha_3 = 0$, which implies that fixed anchor is not used by the agents.

Experimental literature suggests that in general anchors and focal points are important in explaining human behavior. However, HHST09 report that the anchor weight α_3 is typically significant for the subjects under negative feedback treatment, while most of the subjects under positive feedback treatment would not use it. Furthermore, under negative feedback

[21] Notice that what is the anchor, can be a matter of interpretation. One may think of the (6) rule as an anchor-based rule as well, since it can be rewritten as a rule that adjusts the previous price forecast with the latest observed price and trend.



(a) Distance from the fundamental

(b) Distance from the fundamental

(c) Predictions standard deviation

(d) Predictions standard deviation

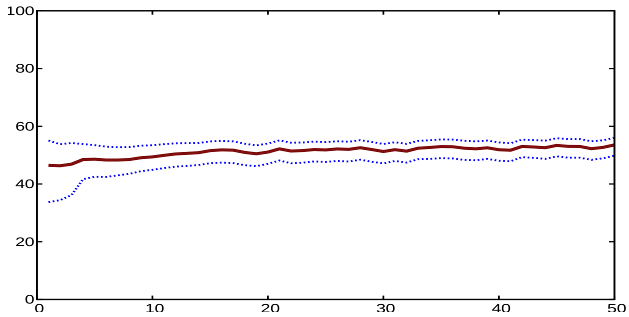
Figure 12: HHST09: 50-period ahead Monte Carlo simulation (1000 markets) for the GA model with anchored-FOR and $\beta \in [-1.1, 1.1]$. Realized price and coordination over time. Green dashed line and black pluses represent the experimental median and group observations; red line is the median and blue dotted lines are the 95% confidence interval for the GA model. Left panel displays the negative feedback, right the positive feedback.

prices and predictions converge to the vicinity of 60, which in practice makes the coefficients α sample-unidentifiable; and could also make redundant the anchor itself. When designing our GA model, we therefore investigated whether the anchor has any additional explanatory power.

To simplify econometric issues, the authors specify the anchor as the fundamental price 60, which however was not directly observed by the subjects. It is more plausible that they used the average price so far as an anchor, $p^a_t = p^t \equiv \sum_{s=1}^{t} p_s$. We will use thus anchored-FOR specified as

$$p^e_{i,t} = \alpha_1 p_{t-1} + \alpha_2 p^e_{i,t-1} + \alpha_3 \left(\sum_{s=1}^{t} p_s \right) + \beta(p_{t-1} - p_{t-2}). \tag{29}$$

We consider the Monte Carlo (MC) simulations exactly as in the first part of Section 3.3, but for the GA model based on (29) with $\beta \in [-1.1, 1.1]$. The results are presented on Figure 12. We observe for the positive feedback that, in contrast to our restricted model without an anchor, the GA model based on FOR as in (29) does not predict oscillations at all, but rather a sluggish convergence towards the fundamental. This is seen in the stable median price, bounded by relatively narrow 95% CI. This specification misses most of the dynamics observed in half of the experiment. We conclude that there is no evidence for a need of an anchor, specified as a long-run average of the observed prices, in our GA model.



(a) $\beta \in [-1.1, 1.1]$

(b) $\beta \in [-1.3, 1.3]$

Figure 13: HSTV05 with $p^f = 60$: 50-period ahead Monte Carlo simulation (1000 markets) for the GA model with anchored-FOR and $\beta \in [-1.1, 1.1]$ and $\beta \in [-1.3, 1.3]$. Realized price over time: red line is the median and blue dotted lines are the 95% confidence interval for the GA model.

E.2 Anchor and HSTV05

The HSTV05 non-linear, two-period ahead LtF asset pricing market resulted in much more unruly oscillations than those observed in the simple linear experiment HHST09 under positive feedback. One could therefore think that some kind of a long-run anchor might have been important for the subjects, even though they would not use it in one-period ahead forecasting setting. Furthermore, in the experiment the oscillations typically unraveled around the fundamental price, which again suggests that the subjects tried to extrapolate the trend around it. To address this issue, we run the 50-period ahead MC simulation like in Section 4.3, but with FOR (16) replaced by the anchored-FOR rule (29) adapted for the two-period ahead setting, assuming that the fundamental price is $p^f = 60$.

Results for two specifications (with allowed trend extrapolation $\beta \in [-1.1, 1.1]$ and $\beta \in [-1.3, 1.3]$) are presented on Figure 13. Just as in the case of HHST09, we find that the GA model with anchored-FOR rule generates sluggish convergence towards the fundamental price from below. Indeed, in contrast to HHST09, the 95% CI of the GA model's prices do not include the fundamental $p^f = 60$ even after 50 periods. This indicated that adding an anchor to the GA model would decrease its fitness to the experimental data.

E.3 Degree of trend extrapolation

Recall that the GA requires a predefined finite interval for the optimized parameters. In the case of our GA model based on (6), the price weight is confound to $\alpha \in [0, 1]$, but *prima facie* there is no 'natural' bound for the trend extrapolation $\beta \in [\beta_L, \beta_H]$, since *a priori* we do not know the degree of trend extrapolation that people consider while forecasting prices. As mentioned in Section 3, we argue that the model performs well if we specify the (6) rule to use 1.1 as the upper bound to the trend.

It turns out (not surprisingly) that the allowed trend extrapolation interval has little effect on the behavior of our GA model under the negative feedback. However, the larger the interval $\beta \in [\beta_L, \beta_H]$ is, the bigger the amplitude of the price fluctuations generated under the positive

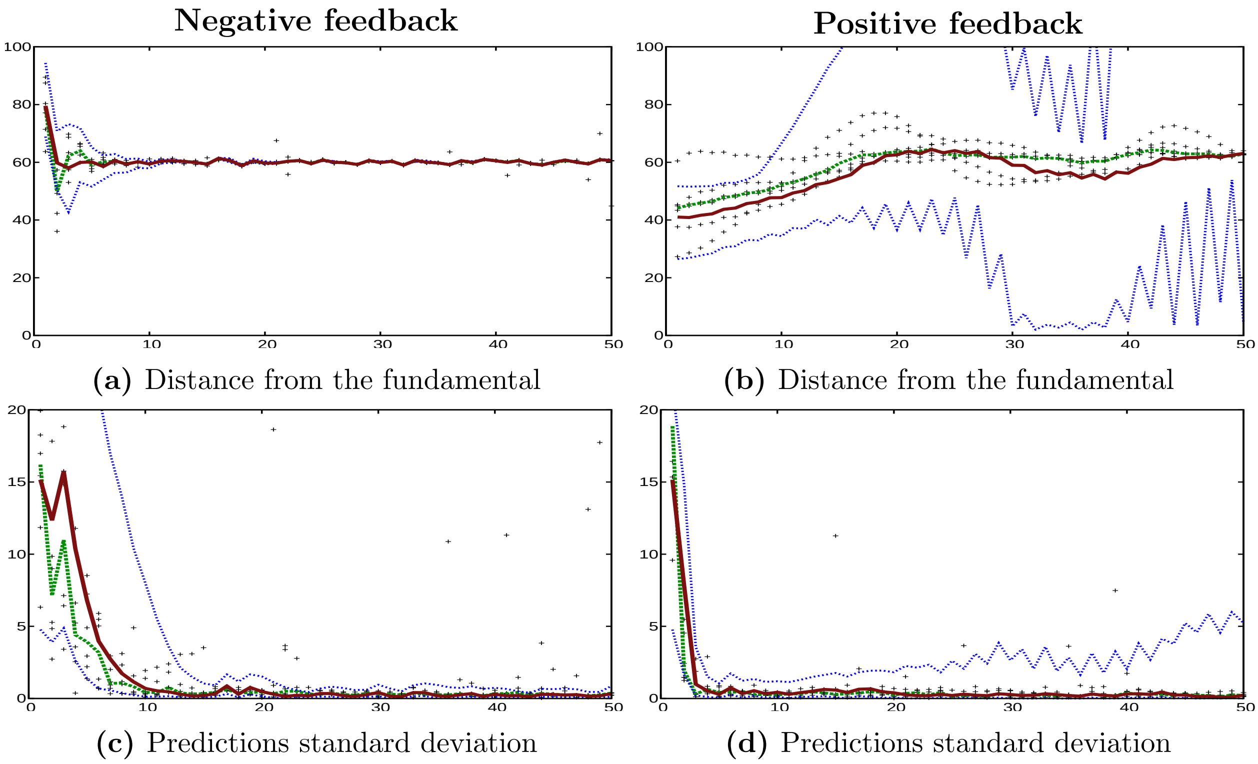


Figure 14: HHST09: 50-period ahead Monte Carlo simulation (1000 markets) for the GA model with $\beta \in [-1.5, 1.5]$. Realized price and coordination over time. Green dashed line and black pluses represent the experimental median and group observations; red line is the median and blue dotted lines are the 95% confidence interval for the GA model. Left panel displays the negative feedback, right the positive feedback.

feedback. Thus we experimented with different β's, trying to calibrate the model to the the experimental oscillations. We used the same Monte Carlo experiments as in the first part of Section 3.3.

Allowing for a high trend extrapolation $\beta \in [-1.5, 1.5]$ results in a model with huge possible oscillations and little predictive power, see Figure 14. On the other hand, specification with $\beta \in [-0.5, 0.5]$ has narrow CI, but predicts small oscillations, see Figure 15. We found the model with $\beta \in [-1.1, 1.1]$ to be the best trade-off between fit and explanatory power of the experiment.

This result reflects the experimental findings. HHST09 find that under positive feedback, four out of twenty estimated rules had $\beta > 0.9$ and further five rules had $\beta > 0.75$. Nevertheless, HHST09 in their estimations impose a restriction that $\beta \in [-1, 1]$. Our GA model suggests that such a restriction is inconsistent with the degree of experimental price oscillations.

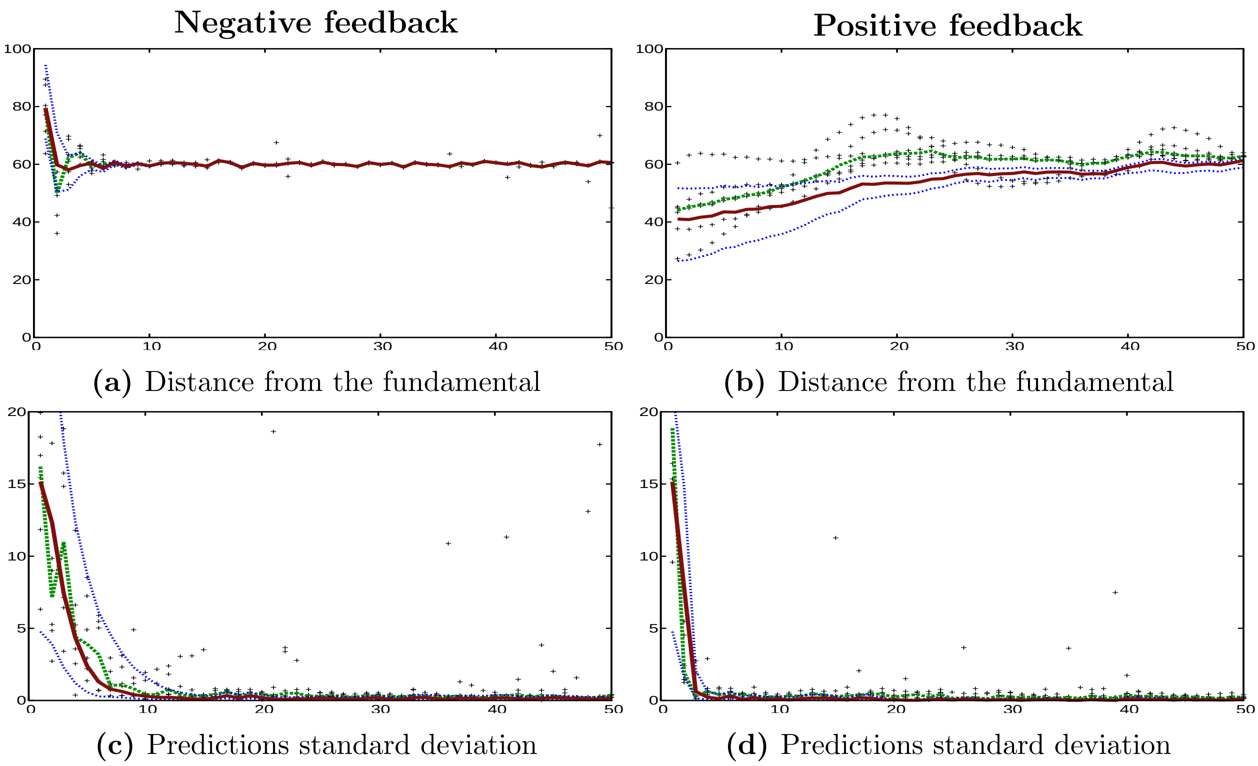


Figure 15: HHST09: 50-period ahead Monte Carlo simulation (1000 markets) for the GA model with $\beta \in [-0.5, 0.5]$. Realized price and coordination over time. Green dashed line and black pluses represent the experimental median and group observations; red line is the median and blue dotted lines are the 95% confidence interval for the GA model. Left panel displays the negative feedback, right the positive feedback.