
Tomasz Makarewicz

Otto-Friedrich-Universität Bamberg

January 31, 2018

Abstract

Experimental literature on financial instability typically focuses either on subject price expectations (Learning to Forecast experiments), or directly on how the subjects trade (Learning to Optimize experiments). One exception is a study by Bao et al. (2017), who show that the subjects have problems with both tasks. In this paper I explore these experimental results, by investigating a model in which financial traders individually learn how to use forecasting and/or trading anchor and adjustment heuristics, by updating them with Genetic Algorithms.

The model replicates main outcomes of the Bao et al. (2017) experiment, and shows that both forecasters and traders coordinate on asset price oscillations, albeit the trading markets generate faster or larger price cycles. All types of agents learn similar behavior where they chase market dynamics through the observed price trends or asset return. Furthermore, agents who learn both tasks can generate large bubbles, in which the price cycles between 10% and 350% of the fundamental value. This heterogeneous agent model offers a unified framework for the forecasting and trading strands of the behavioral finance, shows that they yield symmetric insight into the causes of financial instability, and that trend chasing behavior is a robust outcome of individual learning.

JEL codes: C53, C63, C91, D03, D83, D84.

Keywords: Financial Instability, Learning to Forecast and Learning to Optimize experiments, Genetic Algorithm Model of Individual Learning.
1 Introduction

The financial meltdown of the 2007 and the resulting Great Recession have demonstrated the importance of studying the causes and effects of financial instability. One famous example is the S&P500 stock index, which fell from 1565.15 points (at close) on the 9th October 2007 to 676.53 points on the 9th March 2009. Afterward, the index surged upwards with a virtually undisturbed momentum, and as of writing of this paper (January 2017), crossed the level of 2850 points. This means that within the decade after the peak of the 2007, S&P500 first lost more than half of its value, only to roughly quadruple during the recovery, beating the maximum of the 2000’s by a factor of 180%. NASDAQ and Frankfurt stock exchange indices followed a similar path of boost and boom, while many other (including Shanghai, London, Paris and Warsaw stock indices) experienced similar drops between 2007 and 2009, but recovered in a less spectacular fashion, “only” to around the pre-crisis level. What makes such enormous price swings possible?

Modern models of asset markets make two assumptions about the behavior of financial investors. The first assumption relates the available market information with the agents’ forecasts of the asset price. The second assumption relates these forecasts to the trading positions (or demand schedules) of the agents. Traditionally, the financial literature follows Muth (1961) and embraces the Rational Expectations (RE) hypothesis: forecasting and trading are mutually model-consistent with each other and with the aggregate asset price mechanism (Hommes, 2013b). This allows for the Efficient Market Hypothesis: asset prices are an accurate reflection of their fundamentals (including the risk attitude of the individual investors). However, it is difficult to imagine that within a decade the fundamentals of one stock index can drop by a half and then quadruple. A simpler explanation of the crisis is to assume that the financial investors, albeit smart, face constraints on their cognitive abilities and can commit investment mistakes (Barberis and Thaler, 2003; Shiller, 1981). As a result, markets do not have to necessarily stabilize at the fundamental level. This leads to the central question of this paper: is the financial instability caused by non-rational expectations, or by agents’ inability to transform expectations into rational trading positions?

This question cannot be easily answered by an empirical study, since many relevant financial variables, in particular fundamental values or investors’ forecasts, are not directly observable. A remedy is to run a laboratory experiment, where the researcher can directly control the structure and parametrization of the market, and ask the subjects to elicit their beliefs (Hommes, 2011). The literature on financial experiments can be roughly divided into two main threads: Learning to Forecast, which focuses on the expectation formation of the subjects (Bao et al., 2012; Colasante et al., 2016; Heemeijer et al., 2009; Hommes et al., 2007; Marimon et al., 1993); and Learning to Optimize, where the subjects are directly tasked with trading an asset (for a selection of examples from the vast literature, see Breaban and Noussair, 2015; Dufwenberg et al., 2005; Kirchler et al., 2012; Lei et al., 2001; Noussair and Tucker, 2013; Smith et al., 1988; Weber et al., 2016). These two branches of the experimental
literature yield similar stylized facts: subjects more often than not coordinate on speculative bubbles and market crashes (Palan, 2013). In experiments with constant fundamental price, this pattern can lead to repeated off-fundamental price oscillations (see Hommes et al., 2005, for an example).

Second experimental pattern is that the subjects do not have to converge to the rational expectations solution, nevertheless, they do rely on smart behavioral rules, which can often be represented as simple anchor and adjustment heuristics (Northcraft and Neale, 1987; Sewell, 2007; Tversky and Kahneman, 1974). Moreover, subjects are smart in how they choose their heuristics. For the example of the Learning to Forecast experiments, when the subjects observe price trend, they will try to chase it. In this case, they use the last observed price as an anchor, which then they adjust by the last observed price difference (Hommes, 2011; Palan, 2013). They will furthermore adjust the degree of trend chasing depending on the particular market dynamics that they experience (Anufriev et al., 2015). Similarly, when asked to trade, subjects often adjust their previous position with the expected or last observed return of the asset (Bao et al., 2017).

To the best of my knowledge, there is a scarce number of experimental studies, which try to combine the forecasting and trading designs. A notable example is work by Nickerson et al. (2007), who enhance the Learning to Optimize design of Smith et al. (1988) by eliciting subjects’ price forecasts. The authors, in line with the Learning to Forecast experiments, find that the subjects extrapolate observed price trends.

These results are further investigated by Bao et al. (2017), who focus directly on the difference between forecasting and trading. The authors use a simple asset market, and ask their subjects to elicit forecasts (LtF treatment, where the market price depends on the average price forecast), directly trade an asset (LtO treatment, where the price depends on the average quantity decision), or perform both tasks (Mixed treatment, where subjects are rewarded for both forecasting and trading efficiency, but the price depends only on the quantity decisions). Conforming the previous literature, the authors find that all three treatments lead to substantial market instability, but the two trading treatments are relatively more unstable. Furthermore, the LtF treatment yields largely homogeneous dynamics, whereas the groups in the two trading treatments can coordinate on price oscillations with diversified amplitude and period. In other words, this experiment suggests that investors have problems with both forecasting and optimal trading, but the latter has a more destabilizing effect on the market. In response, subjects learn to coordinate on anchor-and-adjustment rules with a more aggressive adjustment factor under the trading treatments.

The most prominent strand of the theoretical literature, which tries to explain the experimental evidence, is based on the Heuristic Switching Model (henceforth HSM; see Brock and Hommes, 1998). The idea of this model is that the agents rely on simple heuristics, rules of thumb (such as fundamental forecast, naive and adaptive expectations, or trend following

---

1This also hold for experiments, in which the rational bubbles are impossible by the mathematical structure of the experimental economy, see Heemeijer et al. (2009) for an example.
rule), and they tend to switch to the rules with higher past forecasting accuracy. Under some circumstances, the investors can converge to the fundamental price, but often they coordinate on trend chasing behavior, which leads to persistent oscillatory, explosive or chaotic dynamics (see Hommes, 2013a, for a literature overview and examples). The reason is the positive feedback between expectations and trades on the one hand, and realized prices on the other hand. For example, if the investors hold optimistic beliefs about future prices of an asset, they will buy it, which drives its price up. As a result, the investor sentiment is self-fulfilling, and the agents can easily coordinate on price trends dynamics. A plethora of experimental and theoretical investigations show that this behavioral approach can explain many financial stylized facts, and outperforms the RE model in empirical studies, in particular for housing markets (Bolt et al., 2014; Dieci et al., 2017; Kouwenberg et al., 2010) and markets of different financial assets (Boswijk et al., 2007; Dieci and Westerhoff, 2010, 2012; Frijns et al., 2010; Hommes and Wagener, 2009; LeBaron, 2006, 2012; Lux, 1995, 2012; Westerhoff and Reitz, 2003). Finally, HSM is a stylized “first-order approximation” of the investors’ behavior, but can be grounded in an explicit model of individual learning (Anufriev et al., 2015).

To the best of my knowledge, there is no “trading” counterpart to the HSM literature, by which I mean a widely acknowledged model, where the agents trade financial assets with some anchor-and-adjustment heuristics. In practice, the literature on behavioral finance assumes one of the two extremes: agents trade randomly or near-randomly (see for instance Duffy and ¨Unver, 2006; Gode and Sunder, 1993); or that their only cognitive limitation lays in the forecasting problem, but then they trade consistently with their behavioral expectations, as in HSM and its derivatives (Anufriev et al., 2013a).

The goal of this paper is to provide a computational heterogeneous agent model, which can (1) replicate the stylized facts of the Bao et al. (2017) experiment, (2) explain and interpret the results of that experiment, and (3) provide a theoretical framework for models in which both forecasting and trading decision making are subject to bounded rationality. I will follow the work of Anufriev et al. (2015), who use a heterogeneous agent model of explicit learning to show that the HSM is a good approximation of subject behavior in a number of Learning to Forecast experiments. In model, agents forecast and/or trade the asset with simple anchor-and-adjustment heuristics. Forecasters use the first order heuristic as in Anufriev et al. (2015), while the trading heuristic is based on the estimated behavior of the subjects from the two trading treatments of Bao et al. (2017). Agents independently update their anchor-and-adjustment heuristics with the use of Genetic Algorithms (GA). This yields endogenous learning dynamics, in which the willingness of the GA agents to chase the market trend forms a self-reinforcing feedback with unstable price dynamics.  

The GA model allows for a comparison of the behavior of forecasters, traders, and forecaster-

---

2 Genetic Algorithms were initially used in models where the agents optimize their final strategy (such as forecast or quantity) instead of strategy rule (such as forecasting or quantity heuristic). Important early examples can be found in Arifovic (1995) and Arifovic (1996), whereas Dawid (1996) discusses the difference between social and individual learning applications of the GAs. The rule optimization approach to GAs was used for instance in Hommes and Lux (2013). See Anufriev et al. (2015) for a comprehensive literature overview.
traders of two levels of sophistication. It yields three important results. Firstly, the model does replicate the findings of Bao et al. (2017). In particular it shows that all types of markets generate unstable prices. On the other hand, traders can coordinate on price oscillations with more diversified amplitude and frequency than forecasters, where the record bubbles reach 350% of the fundamental price.

The second contribution of the paper is that it demonstrates that the insights of the Learning to Forecast literature are valid in the context of trading (Learning to Optimize) setup. When forecasters and traders try to learn how to forecast price of an asset and how to trade that asset respectively, both are easily pushed towards chasing the market dynamics: price trends in the case of the forecasters and asset return in the case of traders. Asset market is a positive feedback system, where optimism of traders and bullish price dynamics reinforce each other – and the GA model demonstrates that this mechanism operates both through boundedly rational forecasting and trading.

Finally, the third contribution is an important lesson about the reinforcement between trading and forecasting. As explained above, both traders and forecasters learn to chase the price dynamics. But if the agents are asked to perform both tasks at the same time, the effect of the positive feedback of the asset market is strengthened. If the agents cannot directly optimize their forecasts, this results in sharp price oscillations, where prices quickly switch between the boom and bust phases of the cycle. Conversely, if the GA traders can update both forecasting and trading heuristics, these two learning processes amplify each other’s trend chasing bias through the positive feedback nature of the asset markets. Therefore, these agents are even more likely to become overconfident and to coordinate on “super-bubbles” (which can cause prices to oscillate between 10% and 350% of the fundamental price). In sum, the GA model presented in this paper shows that under the positive feedback of the asset market, when the agents try to learn trading and forecasting at the same time – which is the most realistic representation of the empirical financial markets –, they are likely to be pushed far from the rational solution, further than was previously presumed in the literature on behavioral finance.

2 Asset market

In this section I discuss first the asset market from Bao et al. (2017), followed by a summary of the results from that experiment. Finally, I will introduce the Genetic Algorithm model of individual learning. For more details on the model and its motivation, please refer to Anufriev et al. (2015).

2.1 Experimental economy

The experimental economy of Bao et al. (2017) is based on the following asset market. Consider a set of \( I = 6 \) myopic investors, who trade at a period to period basis. At period \( t \), preferences
of agent $i$ are given by the mean-variance utility function of the form

$$U_{i,t} = z_{i,t} \rho_t - \frac{a}{2} z_{i,t}^2,$$

where $z_{i,t}$ denotes the position of agent $i$ at period $t$ (notice that short positions with $z_{i,t} < 0$ are possible), $a = 6$ denotes the risk aversion factor of the agent, $z_{i,t}^2$ corresponds to the perceived risk of the agent’s position, and finally

$$\rho_t \equiv p_t + y - R p_{t-1}$$

is the economic return of the asset given the realized price $p_t$ at period $t$, fixed dividend $y = 3.3$ and gross interest rate $R = 1 + r = 1.05$. For simplicity, suppose that the agents face no additional liquidity constraints. The problem of agent $i$ is that she needs to decide on $z_{i,t}$ before $p_t$ is actually realized. Denote her forecast of $p_t$ as $\hat{p}_{i,t}$, then the optimal demand conditional on this belief is given by

$$z^*_{i,t} \equiv \frac{1}{6} \rho_{i,t} \equiv \frac{\hat{p}_{i,t}^e + 3.3 - 1.05 p_{t-1}}{6}.$$ 

Notice that the expected asset return is of the same order of magnitude as the price trend $\Delta p_t = p_t - p_{t-1}$, especially if the realized prices are high.

Once all the agents submit their individual demands, market maker adjusts the price according to the aggregate demand with the price adjustment rule

$$p_t = p_{t-1} + \lambda \sum_i z_{i,t} + \varepsilon_t,$$

where $\lambda = 1/R = 20/21$ is a scaling factor and $\varepsilon_t \sim NID(0,1)$ is a small price shock. If the agents trade consistently with their price expectations, then under the experimental parametrization the price equation (4) reduces to

$$p_t = p^f + \frac{20}{21} (\bar{p}_t^e - p^f) + \varepsilon_t,$$

where $p^f = y/r = 66$ is the fundamental price and $\bar{p}_t^e \equiv (1/6) \sum_i \hat{p}_{i,t}^e$ is the average forecast of the agents.

Straightforward derivations prove that $p^f$ is indeed the unique fixed point of the expected price equation (5), and hence the unique stationary RE steady state. Furthermore, because the current price $p_t$ depends only on its own forecast through a simple linear equation, one can easily see that the fundamental price $p^f$ is in fact the unique RE solution, while rational bubbles are excluded in this setup. This implies that the model yields any real dynamics only if the agents fail to form perfectly rational expectations.
2.2 Experimental design and results

Bao et al. (2017) studied three treatments: Learning-to-Forecast (LtF), Learning-to-Optimize (LtO) and Mixed, each with eight groups with six subject per group (one subject associated with one computer trader). In all three treatments, subjects were explained that their task is to be a forecasting advisor or a trader to a pension fund. They were given only qualitative description of the price mechanism, and could observe only the aggregate market outcomes (past prices) and their individual decisions and payoffs, but not the decisions or payoffs of the other subjects. They were also unaware of the number of market participants, so that they would behave as price-takers. Every experimental market lasted for $T = 50$ periods.

Under the LtF treatment, every subjects played a role of a financial advisor to a computerized investor, who in turn used the subject’s forecast to compute its optimal demand (3). In practice, subjects were asked to elicit price forecasts $p_{i,t}$, and their average forecast $\bar{p}_t$ was substituted directly into the law of motion (5). Subjects observed the realized price $p_t$ and were paid according to the squared error of their forecasts.

Under the LtO and Mixed treatments, subjects played the role of financial investors. They were provided with the mathematical formula for the utility function (1). They were given a calculator that computed the expected asset return for any price forecast given the price in the previous period, as well as a table that related expected asset return and trading position to the utility of the investor. Subjects were tasked with submitting their trades $z_{i,t}$, which were substituted into the price adjustment equation (4). Subjects observed the realized price and moved to the next period. In the LtO treatment, they were paid based on the realized utility of their trades (1).

Under the Mixed treatment, subjects were also explicitly asked to elicit their price forecasts $p_{i,t}$. These forecasts had no bearing on the market prize. However, by the end of the experiment, the experimenters randomly chose (separately for each session) whether to pay the subjects based solely on their forecasting accuracy, or solely on their trading efficiency, so that the subjects had an incentive to perform both task as well as possible.

The utility function (1) is in fact a square function of the forecasting error (see Bao et al., 2017, for derivations and discussion), hence under perfect rationality, the three treatments are equivalent and should lead to exactly the same dynamics, namely perfectly rational subjects would immediately converge to the fundamental price. This was not the case in the experiment, however. Figure 1 shows sample group per each treatment.

Figure 1a displays a representative group for the LtF treatment. We observe that subjects quickly coordinated on similar forecasts. Nevertheless, the market did not converge to the fundamental, and instead mildly oscillated. In fact all eight groups looked like the group 1, which is reported in Figure 1a.

The dynamics under the LtO and Mixed treatments were statistically indistinguishable (see Bao et al., 2017, for a discussion), but also much more diversified. Most of the groups oscillated like the two sample LtO and Mixed groups reported on Figures 1b and 1c respectively, but the
Figure 1: Sample groups from Bao et al. (2017) experiment, one group per each treatment. In Figure a, the panel shows realized prices and individual forecasts. In Figures (b) and (c), the upper part of each panel shows realized prices, together with individual forecasts in panel (c), while the bottom part of each panel shows subject trades. A thick black line denotes realized price, a thin dashed black line denotes the fundamental price $p^f = 66$, dashed green lines denote individual forecasts $p_{i,t}^e$, dotted blue lines denote individual trades $z_{i,t}$ and red line with triangles denotes the average trade.
specific amplitude and frequency changed from group to group. The two trading treatments were significantly more unstable than the LtF treatment. Furthermore, under the Mixed treatment two groups coordinated on “super-bubbles”, when the price reached approximately 350% of the fundamental price, only to fall to a level close to 10% of the fundamental. In fact, two such bubbles appeared in group 8 (seen on Figure 1c).

This leads us to the three stylized facts that I want to address with my model:

S1 Trading leads to more unstable dynamics than forecasting.

S2 Trading groups generate diversified price oscillations (varying period and amplitude).

S3 Exceptionally large oscillations can emerge under trading.

3 The Genetic Algorithm model of individual learning

The agent structure of the Genetic Algorithm (GA) model is based directly on the information and market structure of the Bao et al. (2017) experiment. Consider a set of $I = 6$ agents, who are asked either to forecast, to trade, or to perform both task, in the asset market discussed in Section 2.1. The agents are unable to immediately form perfectly rational forecasts and trading positions, and instead they are forced to use anchor and adjustment type of heuristics. However, they try to use these heuristics in a smart way, and at every period they update them with a GA procedure, which constitutes the core learning mechanism of the model. Notice that the model for the GA forecasters is directly taken from Anufriev et al. (2015), where it is denoted as the GA-P2 specification. In this section, I will first provide details on the heuristics of the agents, and then explain in detail how the GA learning works.

3.1 Learning to forecast and learning to optimize

3.1.1 Forecasting heuristic

First let us pay attention to the forecasting heuristic. Suppose that an agent $i$ at period $t$ wants to directly forecast the price $p_t$. She is assumed to have a list of $H = 20$ heuristics, and each heuristic $h \in \{1, \ldots, 20\}$ yields a forecast $p_{i,t,h}$ based on the same general and simple first order rule (FOR)

$$p_{i,t,h}^e = \alpha_{i,t,h} p_{t-1} + (1 - \alpha_{i,t,h}) p_{t-1}^e + \beta_{i,t,h} (p_{t-1} - p_{t-2})$$

$$= p_{t-1}^e + \alpha_{i,t,h} (p_{t-1} - p_{t-1}^e) + \beta_{i,t,h} (p_{t-1} - p_{t-2})$$

(6)

where $p_t$ is price realized at period $t$ and $p_{t-1}^e$ is the forecast for the price at period $t - 1$ of the agent $i$. \(^3\) Coefficients $\alpha_{i,t,h} \in [0, 1]$ and $\beta_{i,t,h} \in [0, 1.1]$ are the free parameters of the heuristic $h$, and (as will be explained later) agent $i$ will update them with GAs independently from the

\(^3\)In the experiments, price forecasts were technically constrained to remain in the $[p_{t-1} - 30, p_{t-1} + 30]$ interval. In practice, subjects never hit these boundaries.
other agents. Thus the weight $\alpha_{i,t,h}$ and the trend extrapolation coefficient $\beta_{i,t,h}$ will evolve over time, and have different values for different agents.\(^4\)

The top part of equation (6) shows the FOR heuristic as a sum of adaptive expectations (with weight $\alpha$) and trend extrapolation rule (with trend coefficient $\beta$). Alternatively, the bottom part of this equation shows the FOR as a simple anchor and adjustment rule: the price forecast $p^e_{i,t,h}$ is equal to the previous forecast, but adjusted by the forecasting error (with the scaling factor $\alpha$) and by the last observed price trend (with the adjustment rate $\beta$).

It is worth noting, following Anufriev et al. (2015), that the FOR heuristic (6) has some important special cases. In particular, it becomes naive forecast with $\alpha = 1$ and $\beta = 0$, adaptive forecast with $\beta = 0$ and pure trend following rule with $\alpha = 1$ and $\beta \neq 0$. The RE forecast can also be expressed as the FOR rule, with $\alpha = \beta = 0$ and $p^e_{i,t-1} = p^f_{i,t}$, which yields $p^e_{i,t} = p^f_{i,t}$. Notice that the RE forecasting rule has the following interpretation: the GA agents should learn not to adjust their forecast, and instead she should learn to rely solely on a correctly specified anchor (equal to the fundamental price).

### 3.2 Three approaches to trading

Under the LtO and Mixed treatments of the Bao et al. (2017) experiment, the subjects were asked to trade the asset, so they utilized a trading anchor and adjustment rule. Furthermore, the experimental instructions explained that they should trade based on the expected return of the asset, where the latter needs an explicit price forecast. Additionally, under the Mixed treatment the subjects were directly asked to elicit their forecasts. It follows that all the subjects from the Mixed treatment, and possibly some from the LtO treatment, tried to forecast the asset price in a similar fashion to the subjects from the LtF treatment. But what it is not immediately clear is how the subjects linked forecasting with trading, if at all. This leads to three possibilities.

The first possibility is that the subject disregarded altogether price forecasting when trading. This is not necessarily a poor strategy, especially for those subjects, who found it difficult to form accurate forecasts, to compute the expected return, or to fully understand the relationship between the asset return, market price and trading payoff. Under the Mixed treatment this implies that the subjects did forecast the price for the forecasting task, but they simply did not connect it with their trading decisions.

The second possibility is that the subjects realized that the price influences the asset return, and so it is useful to keep a price forecast in the trading heuristic through an asset return term. However, for any number of reasons (which may include cognitive constraints), the subjects did not directly learn the price forecasting heuristic itself, and instead, they focused on one large trading heuristic with a price forecast inside, judging it based solely on

\(^4\)For the sake of clarity, I will drop the three indices $(i,t,h)$ from the heuristic’s free parameters in non-technical discussion of this paper. The forecasting heuristic for the LtF GA agents is a function of two free parameters (price and price trend weights), as well as past market variables (realized prices and agent’s forecast). To keep the notation tractable, I will refer to these heuristics as functions of free parameters, or as functions of the past market variables, depending on the context.
its trading profitability. That heuristic does contain the expected asset return, which is based on the expected price, which in turn is based on a forecasting heuristic, but this forecasting heuristic is optimized only indirectly via the trading heuristic. It is tempting to call these subjects “unsophisticated” or “confused”. However, one should remember that if the subjects find it difficult to forecast the asset price, such an approach may be a reasonable second-best without neglecting the price forecasting altogether – in particular under the LtO treatment, where the forecasting accuracy is not directly rewarded.

The third and last possibility is that the subjects divided their task into the two components, forecasting and trading, and tried first to learn a forecasting heuristic, use it to generate a forecast, and only then place it through an asset return component into a trading heuristic, which they optimized separately from the forecasting heuristic. These subjects were likely the most “rational” in the experiment, however, such a learning may also be the most cognitively demanding.

In sum, trading subjects could pursue one of the three following trading strategies (with the description preceded by name tags that will be used henceforth):

**NO FORE** Use a trading heuristic and disregard price forecasting altogether;

**INDIRECT FORE** Use a trading heuristic with an expected asset return component, but do not learn the price forecast rule separately, and instead focus on one large heuristic;

**OPTIMIZED FORE** Learn forecasting and trading heuristics separately, and include the price forecast in the trading heuristic.

In the remainder of this subsection, these trading strategies will be linked with appropriate anchor and adjustment trading heuristics.

### 3.2.1 NO FORE traders

Consider the agents, who do not use price forecasts in their trading. **NO FORE** agent $i$ at period $t$ is endowed with $h \in \{1, \ldots, 20\}$ trading heuristics, which are based on the following simple AR rule:

$$z_{i,t,h} = \chi_{i,t,h} z_{i,t-1} + \phi_{i,t,h} \rho_{t-1} \in [-5, 5],$$

where $z_{i,t-1}$ is the trade of agent $i$ in the previous period, $\rho_{t-1}$ is the realized asset return in the previous period as in equation (2), and $\chi_{i,t,h} \in [-0.5, 1]$ and $\phi_{i,t,h} \in [0, 0.3]$ are the free parameters which the agent $i$ is trying to learn. The constraint that the trades must remain in the $z_{i,t} \in [-5, 5]$ interval was imposed in the experiment in order to exclude explosive price paths, and is roughly equivalent with the LtF treatment constraint that the price forecast must remain in the $p_{t,h}^{\epsilon} \in [p_{t-1} - 30, p_{t-1} + 30]$ interval (see Bao et al., 2017, for a discussion). Unlike under the LtF treatment, the constraint on $z_{i,t}$ was repeatedly hit by both the experimental subjects and the GA agents.
The heuristic (7) (including the bounds on parameters) is taken directly from Bao et al. (2017), who show that this rule describes well the behavior of their subjects under the LtO treatment. Furthermore, it has a simple interpretation as an anchor and adjustment rule, where the decision of agent $i$ is equal to her previous trade $z_{i,t-1}$ (scaled with factor $\chi$) adjusted with the last observed asset return (with weight $\phi$).

As explained earlier, under RE every agent forecasts $p_{i,t}^f = p' = \frac{y}{(R - 1)}$ and so $p_t = p' + \varepsilon_t = \frac{y}{(R - 1)} + \varepsilon_t$ at every period $t$. Substituting this into the optimal trade (3), we obtain that

$$z_{i,t,h}^{RE} = \frac{p_{i,t}^f + y - Rp_{t-1}}{6} = \frac{y/(R - 1) + y - R\varepsilon_t/(R - 1) + R\varepsilon_{t-1}}{6}$$

(8)

It follows that under RE, the expected trade of agent $i$ is equal to zero, while the realized trades are a small noise term. Hence, the AR trading heuristic approximates the RE solution for $\chi = \phi = 0$.

### 3.2.2 INDIRECT_FORE traders

Consider now the agents, who do use price forecast in their trade heuristic, but they do not optimize the forecasting heuristic directly. **INDIRECT_FORE** agent $i$ has $h \in \{1, \ldots, 20\}$ simple trading rules of a form

$$z_{i,t,h} = \chi_{i,t,h}z_{i,t-1} + \phi_{i,t,h}p_{t-1} + \zeta_{i,t,h} \rho_{i,t,h}^e \in [-5, 5],$$

(9)

where $\chi_{i,t,h} \in [-0.5, 1]$, $\phi_{i,t,h} \in [0, 0.3]$ and $\zeta_{i,t,h} \in [0, 0.3]$ are the free parameters of the heuristic, and

$$\rho_{i,t,h}^e = p_{i,t,h}^e + y - Rp_{t-1}$$

(10)

is the expected asset return according to heuristic $h$ of agent $i$ at period $t$, and the price expectation $p_{i,t,h}^e$ (in a similar way to the LtF agents and their FOR heuristic (6)) is constructed based on

$$p_{i,t,h}^e = \alpha_{i,t,h}p_{t-1} + (1 - \alpha_{i,t,h})p_{t-1}^e + \beta(p_{t-1} - p_{t-2}).$$

(11)

Notice, however, that in this case, the price forecasting rule (11) is in fact a part of the broader trading heuristic (7). One can think of that trading rule as a singleton function of both the
trading and forecasting free coefficients, conditional on previous trade and forecast:

\[
\begin{align*}
  z_{i,t,h} & = \chi_{i,t,h} z_{i,t-1} + \phi_{i,t,h} \rho_{t-1} + \zeta_{i,t,h} (\alpha_{i,t,h} p_{t-1} + (1 - \alpha_{i,t,h}) p_{t-2} + y - R p_{t-1}) \\
  &= \chi_{i,t,h} z_{i,t-1} + \phi_{i,t,h} \rho_{t-1} + \zeta_{i,t,h} (\alpha_{i,t,h} p_{t-1} + (1 - \alpha_{i,t,h}) p_{t-2} + y - R p_{t-1}).
\end{align*}
\]

(12)

This heuristic, together with its parametrization, is based on the estimated individual rules under the Mixed treatment from the Bao et al. (2017) experiment.

The heuristic (11) is similar to (7) in its form and interpretation. Again, it is an anchor and adjustment rule, where the previous trade \( z_{i,t-1} \) plays the role of the anchor. Unlike the NO_FORE traders, however, the INDIRECT_FORE traders have a more sophisticated adjustment mechanism, with both a backward looking component (the previous asset return weighted with \( \phi \)) and a forward looking component (the expected asset return weighted with \( \zeta \)). As a result, these agents can be more flexible, and have a more straightforward way to converge to the rational solution. Just like in the case of the LtF agents, \( \alpha = \beta = 0 \) and \( p_{e_{i,t-1}} = p_f \) yield the RE forecast. Conditional on the rational forecast, the optimal demand \( \alpha \) can be expressed with the trading heuristic (9) if \( \chi = \phi = 0 \) and \( \zeta = 1/6 \). Notice that unlike the NO_FORE agents, the INDIRECT_FORE agents can trade precisely accordingly to the RE solution, not just in expected terms.

The RE trading heuristic has the following interpretation: conditional on the RE price forecast, the trading GA agent should learn to disregard the anchor (i.e., her previous trade), and instead she should focus on the adjustment part of her heuristic, in particular learn the optimal 1/6 weight on the expected asset return and disregard the previous asset return. This is the exact opposite of what the RE forecasting heuristic requires, namely disregarding the adjustment factor and focusing on the optimal anchor (which is equal to the fundamental price). This difference will have a crucial effect on the learning of the GA agents.

3.2.3 OPTIMIZED_FORE traders

The last type of the GA agents who can separate the forecasting and trading tasks. They are similar to the INDIRECT_FORE agents, but with one important difference. First, OPTIMIZED_FORE agent \( i \) has \( h \in [0, \ldots, 20] \) forecasting heuristics (6) like the LtF agents. She uses them to generate price forecast \( p_{e_{i,t}} \) (which will be explained in the following subsection). Then, she also has \( k \in [0, \ldots, 20] \) heuristics of a form

\[
\begin{align*}
  z_{i,t,k} &= \chi_{i,t,k} z_{i,t-1} + \phi_{i,t,k} \rho_{t-1} + \zeta_{i,t,k} p_{e_{i,t}}^c \\
  &= \chi_{i,t,k} z_{i,t-1} + \phi_{i,t,k} \rho_{t-1} + \zeta_{i,t,k} p_{e_{i,t}}^c + y - R p_{t-1}. \\
\end{align*}
\]

Notice the trading heuristic (13) is similar to the heuristic (9) of the INDIRECT_FORE traders. However, the OPTIMIZED_FORE traders use their heuristics in a different fashion, since they perceive their task as a two-dimensional problem. First, agent \( i \) generates price
forecast $p_{i,t}^e$ based on her forecasting heuristics $p_{i,t,h}^e = p_{i,t,h}^e(\alpha_{i,t,h}, \beta_{i,t,h} | p_{i,t-1}^e, p_{1:t})$. Second, she uses that forecast to decide on trading position based on trading heuristics

$$z_{i,t,k} = z_{i,t,k}(\chi_{i,t,k}, \phi_{i,t,k}, \zeta_{i,t,k} | z_{i,t-1}, p_{i,t-1}) .$$

On the other hand, just like in the case of **INDIRECT FORE**, the two heuristics (6) and (13) yield the RE solution when $\alpha = \beta = \chi = \phi = 0$, $\zeta = 1/6$ and $p_{i,t-1}^e = p^f$.

### 3.3 Genetic Algorithm

The model follows closely Anufriev et al. (2015). Its premise is that the agents can learn how to use different anchor and adjustment heuristics via Genetic Algorithm (GA) optimization procedure. A technical discussion of the GAs can be found in [reference], while [reference] gives a literature on the pioneering GA applications in economics. This subsection will provide an intuitive explanation of the GA, for the technical discussion refer to Haupt and Haupt (2004), while Dawid (1996) provides an overview of economic applications.

GAs are a class of numerical stochastic optimization algorithms, which are inspired by the evolutionary mechanism with which the genome of living organisms adapts to the ever-changing environment through procreation and mutation. Unlike the algorithms based on the Newton method, GAs do not operate based on (analytical or estimated) Jacobian of the value function, but only on the value of the function. As a result, GAs tend to outperform traditional algorithms when dealing with highly non-linear and multidimensional optimization problems, as well as problems with non-real arguments (such as binary or integer arguments), and are popular in engineering and related fields.

The standard binary GA algorithm operates as an evolutionary search mechanism upon the grid of arguments in the following way. Initially, a population of $H$ arguments (also denoted as chromosomes) is created at random, where the arguments are represented with so called chromosomes: binary strings of 0s and 1s. Then, at every iteration, the arguments undergo the following set of four operators:

**Procreation** The value of every argument is computed. These values are transformed (e.g., with the logistic map or power law) to a probability for each argument: the higher the value of an argument compared to other arguments (its fitness), the higher its weight. These probabilities are used to sample (with replacement) a set of new $H$ arguments (dubbed children arguments) from the old population (parent arguments).

**Crossover** The children arguments are divided into pairs, and each pair with probability $\delta_C$
will exchange a predetermined subset of bits. Typically, this means exchanging subsets of arguments.

**Mutation** With probability $\delta_M$, every bit in every argument is swapped (1s into 0s and the other way around).

**Election** The value of every child is computed. If the child argument has a higher value than its parent, it replaces the parent, otherwise the parent is retained in the population, and the child disappears (see Arifovic, 1995, 1996).

The procedure runs for a predetermined number of iterations, or until a certain convergence criterion is met.

The procreation operator is the evolutionary mechanism, with which the better arguments slowly replace the worse ones. The two evolutionary operators, crossover and mutation, allow for experimentation of the argument value. Finally, the election operator (introduced in Arifovic, 1995) screens off experimentation that leads to worse fit.

One can immediately see that this process can be reinterpreted as a learning mechanism: agents focus on more successful strategies, experiment with their parameters, and discard unsuccessful mutants. This is precisely the way, in which the agents in my model use GAs. At the beginning of every period, when they observe the new realized price, they use one iteration of the GAs to update the free parameters of their heuristics, in order to fine-tune them to the observed market dynamics. In this way, these GA agents can efficiently learn how to operate on the market, and remain relatively smart. On the other hand, one iteration of GAs is computationally simple and “cheap”, so it is reasonable to think that it remains a good approximation of the actual learning and cognitive capabilities that the subjects exhibited in the laboratory.

### 3.4 Timing of the model

There are four variants of the model, one with forecasters and three with traders. In the first variant, denoted as **LtF**, the GA agents are asked to forecast the price, and their forecasts $p_{e,t}$ are used to generate the next price via the expectation-price feedback mechanism (5). These agents use the FOR heuristic (6). In the other variants of the model, denoted as **LtO**, the GA agents are asked to specify their trading positions $z_{i,t}$, which are used to generate the next price via the price adjustment mechanism (4). The trading GA agents use one of the three trading strategies, **NO_FORE**, **INDIRECT_FORE** or **OPTIMIZED_FORE**, which corresponds to the three **LtO** variants of the model. Regardless of the variant of the model, the market operates for $t \in \{1, \ldots, 50\}$ periods like in the experiment.

---

6In each of the four model variants, the GA agents are of the same type. An interesting extension of the model would be to study markets with mixed composition of the GA agents. This is not the focus of this study, and it furthermore should at least consider $4^6 = 4096$ possible combinations of the four types for a market with six GA agents. I leave this problem for future studies.
3.4.1 Initial two periods

At the beginning of period $t = 1$, the GA agents have no information about the market. Just like the subjects in the lab, they are yet to observe the first realized price $p_1$, so their initial decision (forecast or trade) is necessarily random. In addition, the heuristics are yet undefined, since they operate on past prices and decisions. The goal of the paper is to understand the differences between the treatments of the Bao et al. (2017) experiment, therefore, following the methodology of Anufriev et al. (2015), I take the initial decisions of the GA agents (forecasts $p_{e,i,1}$ and/or trades $z_{i,1}$) as exogenous, sampled from a distribution, which is calibrated to the empirical distribution of the initial forecasts and trades from the experiment. In particular,

- in the **LtF** variant of the model, each GA agents $i$ samples one $p_{e,i,1}$;
- in the **LtO** variant of the model, each GA agent $i$ of type
  - **NO.Fore** samples one $z_{i,1}$,
  - **INDIRECT.Fore** samples one $p_{e,i,1}$ and one $z_{i,1}$,
  - **OPTIMIZED.Fore** samples one $p_{e,i,1}$ and one $z_{i,1}$.

In the **LtF** variant of the model, the initial forecasts $p_{e,i,1}$ are substituted directly into the price equation (5), while in the **LtO** variant the initial trades $z_{i,1}$ are substituted into the price equation (4). In the latter case, the price adjustment mechanism (4) requires price at period 0. Following the experimental design, it is set to $p_0 = 42$ (approximately 64% of the fundamental price).

Once $p_1$ is generated, the GA agents observe it. In the **LtO** variant of the model, the agents also learn the first realized asset return $\rho_1 = \rho(p_1, p_0)$, so that they can use it in their trading heuristics. On the other hand, the forecasting heuristic (6) also requires the initial price trend, and it is assumed that the agents observe $\Delta p_1 = p_1 - p_0 = 0$. Then, the model moves to the next period $t = 2$.

Starting from period $t = 2$, the heuristics are properly defined and the GA agents start to use them. However, the heuristic parametrization also needs to be initialized. Following Anufriev et al. (2015), agents have so little information at the beginning of the session that they will simply sample initial heuristics at random, as to start with a good spread of potential strategies. In particular, they generate their heuristics from a uniform distribution, i.e., every bit in every heuristic of every agent becomes one or zero with equal probability of 0.5.

---

7This assumption comes from Anufriev et al. (2015), who in turn impose it because $p_0$ is usually not revealed to the subjects in the LtF experiments. It is seemingly a contradiction that the LtO GA agents can use $p_0$ to compute the initial asset return, while the LtF GA agents cannot use the same variable to compute the initial price trend. On the other hand, $p_0$ plays an important role in the LtO price mechanism, unlike for the LtF model variant, hence it seems natural that the traders would be more cognizant of it. Auxiliary simulations show that in fact this detail of the model initialization plays no role, and if we set the initial observed asset return to $\rho_1 = 0$ (which roughly corresponds to $\Delta p_1 = 0$), the Monte Carlo distribution of the LtO variant remains qualitatively intact. This is because the typical initial asset return is close to zero anyway, whereas the GA agents use random asset return weight $\phi$. 

16
Once the agents have their heuristics initialized, each agent $i$ samples one heuristic (with equal weights of $1/H = 0.05$) to perform her decision. In particular,

- in the LtF variant of the model, each GA agents $i$ samples at random one $h$ heuristic from her 20 heuristics with equal probabilities, and uses it to generate her forecast $p_{i,2}^e = p_{i,2,h}^e(p_1, p_{i,1}^e, \Delta p_1 = 0)$,

- in the LtO variant of the model with GA agents of type

  NO FORE each GA agent $i$ samples at random one $h$ heuristic from her 20 heuristics with equal probabilities, and uses it to generate her trade $z_{i,2} = z_{i,2,h}(p_1, \rho_1)$,

  INDIRECT FORE each GA agent $i$ samples at random one $h$ heuristic from her 20 heuristics, and uses it to generate her trade and forecast

  $$
  \begin{pmatrix}
  z_{i,2} \\
  p_{i,2}^e
  \end{pmatrix} = z_{i,2,h}(p_1, \rho_1, p_{i,1}^e, \Delta p_1 = 0),
  $$

  OPTIMIZED FORE each GA agent $i$ samples at random one $h$ heuristic from her 20 forecasting heuristics, and uses it to generate $p_{i,2}^e = p_{i,2,h}^e(p_1, \Delta p_1 = 0)$. Next, agent $i$ samples at random one $k$ heuristic from her 20 trading heuristics with equal probabilities, and uses it to generate her trade $z_{i,2} = z_{i,2,h}(p_1, \rho_1, p_{i,2}^e(p_{i,2}))$.

Once all the agents have submitted their decisions, the second price $p_2$ is realized according to the price mechanism (5) or (4) for the LtF and the LtO variants of the model respectively.

### 3.4.2 Periods after $t = 2$

In period $t = 3$, each GA agent can now evaluate her forecasting and/or trading heuristics from the previous period, since she observes the realized price $p_2$. This is the moment, when the agents start learning with the GAs. In particular, they focus on the following criteria functions $V(\cdot)$, which measure the hypothetical performance of the heuristics in the previous period:

- in the LtF variant of the model, each GA agents $i$ evaluates each of her heuristic $h$ with a simple squared error measure

  $$(14) \quad V_F(i, h, t - 1) \equiv V_F(p_{i,t-1,h}^e(\alpha_{i,t-1,h}, \beta_{i,t-1,h})) = \left( p_{i,t-1,h}^e(\alpha_{i,t-1,h}, \beta_{i,t-1,h}) - p_{t-1} \right)^2. $$

- in theLtO variant of the model, each GA agent $i$ of type

  NO FORE evaluates each of her heuristics $h$ with a simple measure

  $$(15) \quad V_{NO}(i, t - 1, h) \equiv V_{NO}(z_{i,t-1,h}(\chi_{i,t-1,h}, \phi_{i,t-1,h})) = \left( z_{i,t-1,h}(\chi_{i,t-1,h}, \phi_{i,t-1,h}) - \frac{\rho_{t-1}}{6} \right)^2, $$

17
which shows the squared deviation from the optimal demand at period \( t - 1 \) (3)
and is proportional to the realized utility of agent \( i \) (1).

**INDIRECT_FORE** evaluates each of her heuristics \( h \) with

\[
V_{IND}(i, t - 1, h) = V_{IND}(z_{i,t-1,h}(x_{i,t-1,h}, \phi_{i,t-1,h}, \zeta_{i,t-1,h}, \alpha_{i,t-1,h}, \beta_{i,t-1,h}))
\]

\[
= \left( z_{i,t-1,h}(x_{i,t-1,h}, \phi_{i,t-1,h}, \zeta_{i,t-1,h}, \alpha_{i,t-1,h}, \beta_{i,t-1,h}) - \frac{p_{t-1}}{6} \right)^2 ,
\]

which has the same interpretation as a loss of potential utility for agent \( i \).

**OPTIMIZED_FORE** evaluates her forecasting and trading heuristics separately, namely,
she evaluates each of her forecasting heuristics \( h \) with the squared error measure
(similarly to the LtF agents)

\[
V_{OPT}^F(i, t - 1, h) = V_{OPT}^F(p_{i,t-1,h}(\alpha_{i,t-1,h}, \beta_{i,t-1,h}))
\]

\[
= \left( p_{i,t-1,h}(\alpha_{i,t-1,h}, \beta_{i,t-1,h}) - p_{t-1} \right)^2 ,
\]

and each of her trading heuristics \( k \) with

\[
V_{OPT}^T(i, t - 1, k) = V_{OPT}^T(z_{i,t-1,k}(x_{i,t-1,k}, \phi_{i,t-1,k}, \zeta_{i,t-1,k}))
\]

\[
= \left( z_{i,t-1,k}(x_{i,t-1,k}, \phi_{i,t-1,k}, \zeta_{i,t-1,k}) - \frac{p_{t-1}}{6} \right)^2 ,
\]

where these two measures have the same interpretation as for the other types of agents.

Notice that the measures \( V(\cdot) \) are computed treating \( p_{i,t-2}^c \) and/or \( z_{i,t-2} \) as constants. They are then used to evaluate the relative performance of the heuristics in comparison with each other. GA agents are assumed to focus on the logit transformation

\[
\Pi_{i,t-1,u} = \frac{\exp - V_{i,t-1,u}(\cdot)}{\sum_{l \in U} \exp - V_{i,t-1,l}(\cdot)} \in [0, 1),
\]

of the relevant criterion function for all heuristics \( u \in U \), where \( U \) denotes (i) the set of \( H \) forecasting heuristics of the LtF GA agents, (ii) the set of \( H \) trading heuristics of the **NO_FORE** and **INDIRECT_FORE** GA agents, and (iii) separately the set of \( H \) forecasting heuristics and the set of \( K \) trading heuristics for the **OPTIMIZED_FORE** GA agents.

Notice that \( \Pi(\cdot) \) can be reinterpreted as probabilities, which assign weight to each heuristic of agent \( i \) depending on how well this heuristic performs in comparison with her other heuristics.

Hence, starting at period \( t = 3 \), every period \( t \) has the following timing:

1. GA agents observe the realized market price from the previous period \( p_{t-1} \), and compute the relevant auxiliary variables such as \( \Delta p_{t-1} \) for the forecasting agents and \( \rho_{t-1} \) for the trading agents.
2. Every GA agent $i$ updates her heuristics with one iteration of the GA operators (procreation, mutation, crossover, election), where the value function for the GAs is the logit transformation (19) of the relevant criterion function $V(\cdot)$. On the one hand, this implies the $V_F(i, h, t - 1)$ criterion (14) for the LtF GA agents, the $V_{NO}(i, t - 1, h)$ criterion (15) for the NO_FORE LtO GA agents and the $V_{IND}(i, t - 1, h)$ criterion (16) for the INDIRECT_FORE LtO GA agents. On the other hand, the OPTIMIZED_FORE LtO GA agents follow this procedure twice. First, each agent $i$, based on the logit transformation (19) of the $V_{OPT}^F(i, t - 1, h)$ criterion (17), updates her forecasting heuristics with one iteration of the GAs. She then repeats this step, updating her trading heuristics with GAs based on the logit transformation of the $V_{OPT}^T(i, t - 1, h)$ criterion (18). Remark that the two GA update steps are performed independently from each other, as well as from all the other agents.

3. Each GA agent $i$ samples one relevant heuristic from the list of the updated rules, and uses it to generate her decision:

- each LtF agent $i$ samples one forecasting heuristic (6) and submits the corresponding price forecast $p_{i,t}$ to the market maker;
- each LtO NO_FORE agent $i$ samples one trading heuristic (7) and submits the corresponding trading position $z_{i,t}$ to the market maker;
- each LtO INDIRECT_FORE agent $i$ samples one trading heuristic (9), remembers the corresponding price forecast $p_{i,t}$, and submits her final trading position $z_{i,t}$ to the market maker;
- each LtO OPTIMIZED_FORE agent $i$ samples one forecasting heuristic (6), computes her price forecast $p_{i,t}$, then samples one trading heuristic (13) and submits the corresponding trading position $z_{i,t}$ (based on $p_{i,t}$) to the market maker.

To sample their heuristics, GA agents use probability weights equal to the logit transformation of the relevant criterion $V(\cdot)$ (exactly like in the GA learning step 2 of this enumeration).

4. The market maker collects the decisions of the GA agents, and uses them to generate the next price $p_t$: based on the individual price forecasts $p_{i,t}$ and the price mechanism (5) in the LtF variant of the model, and based on the individual trading positions $z_{i,t}$ and the price adjustment mechanism (4) in the LtO variant of the model. Notice that the forecasts of the LtO INDIRECT_FORE and OPTIMIZED_FORE GA traders do not directly influence the market maker.

5. Once the new price $p_t$ is established, the algorithm goes back to its first step, and new period $t + 1$ starts.

The market operates according to this algorithm for a pre-specified number of periods $T$, which in this paper is set to $T = 50$ as in the experiment of Bao et al. (2017).
4 Results

4.1 Simulation setup

The GA model is highly non-linear and allows for no analytical solution. Instead, I will focus on a simple Monte Carlo (MC) exercise with which I will study the dynamics of the model in the experimental setting of Bao et al. (2017). Two issues are worth mentioning at this point. First, the goal of this paper is to shed some light on the learning that took place in the experiment, and so I will not study the long-run behavior of the model. Instead, I will focus on 50 period long simulations as in the experiment, using the experimental data as a natural benchmark for the model results.

Second, it would be interesting to compute a measure of how well the model fits the experimental data. An example is a one-period ahead mean squared prediction error of the model as in Anufriev et al. (2015), who show that the GA model outclasses other forecasting models in explaining the LtF experiments. This is possible, since there exists a well established behavioral literature on expectation formation, including the Heuristic Switching Model, which offered a natural benchmark for the LtF GA model. To the best of my knowledge, there exists no direct counterpart behavioral model of how people learn to trade financial assets, in spite of a vast empirical and experimental literature of what are the trading biases of investors. Therefore, computing a fitness measure for my model would yield little additional insight. I will instead focus on the stylized facts from the experiment by Bao et al. (2017).

There are four specifications of the model: one LtF and three LtO. For each, I will run 1000 independent simulations. Each simulation is based on different initial conditions (initial forecasts and/or trades, and initial heuristics), as well as different pseudo-random numbers for the learning process. The simulation executable is available on request.

4.2 Market (in-)stability

Figure 2 shows sample simulations of the four specifications of the GA model. These sample paths are representative for the MC simulations of the realized prices, as seen in Figure 3. In all four cases the market does not truly converge to the RE solution, however, the three trading specifications (Figures 3b-3d) lead to substantially more pronounced price oscillations, but also to much wider 95% Confidence Intervals (CI). This result replicates the first stylized fact from the experiment S1.

There are also important differences between the threeLtO model specifications. First, the INDIRECT_FORE model (Figure 3b), where the agents use price forecast, but fail to
optimize it directly, leads to faster oscillations with lower amplitude. The shape of the 95% CI of the prices are less regular than in other simulations (Figures 3c and 3d), which shows that different markets under this model specification tend to generate different amplitude phases.

The two other LtO specifications, NO_FORE and OPTIMIZED_FORE, are more regular (Figures 3c and 3d), and can generate two large bubbles within 50 periods. During the first bubble, the upper bounds of the 95% CI of both models lay above 250, that is above 400% of the fundamental price $p_f = 66$. This is followed by a significant market crash, where the median prices for the two specifications drop to around a third of the fundamental, while many simulations hit the zero price bound. The shape of the 95% CI (and that of the sample simulations) shows that, unlike for the INDIRECT_FORE specifications, these markets tend to generate oscillations of similar frequency.

Nevertheless, it is clear that when the GA agents include the expected asset return in their trading heuristic, and their price forecast heuristic is directly optimized, this further destabilizes the market. At the peak of the first bubble, the median price of the OPTIMIZED_FORE GA market reaches level of around 234 (Figure 3d), which is approximately two thirds larger than 140, the maximum of the median of the NO_FORE GA markets, and equal to around 350% of the fundamental price $p_f = 66$. Both models predict a second bubble, but this one is visibly smaller under the NO_FORE model specification. In contrast, some markets of the OPTIMIZED_FORE GA model hit the price level of 250 for the second time, and the price at the median market is around twice as large as under the NO_FORE specification. This replicates the third stylized fact of the experiment S3.

To sum up, when we ask the GA agents to learn trading heuristics, it is much more
complicated for them to converge to a stable regime than under the forecasting specification of model. Furthermore, the model is sensitive to what are the trading heuristics that the GA agents use. This is in line with the second experimental stylized fact S2.

**Result 1.** *The 50 period ahead simulations of the GA model replicate all the three stylized facts from the Bao et al. (2017) experiment.*

### 4.3 Individual learning

What is the link between the observed market dynamics and the realized individual learning in the GA model? As expected, this depends on the specification of the model, but large heterogeneity prevails also within each specification.

#### 4.3.1 Forecasting heuristics

In line with the existing literature on the LtF setting, the mechanism behind price oscillations in the LtF model is price chasing behavior. Figure 4 shows the Monte Carlo evolution of the two parameters from the forecasting heuristic (6), price weight $\alpha$ and trend extrapolation coefficient $\beta$. In the LtF model specification, agents learn to chase the price trend with median forecasting rule close to

$$p_{i,t}^c \approx 0.9p_{t-1} + 0.1p_{i,t-1} + 0.7(p_{t-1} - p_{t-2}),$$

---

**Figure 3:** 50-period ahead MC simulation (1000 markets) for the four specifications of the GA model: evolution of the asset price. Red thick line is the median and blue dotted lines are the 95% confidence interval for the GA model.
though the median trend coefficient $\beta$ oscillates to some degree, taking values between 0.6 and 0.9. Nevertheless, the 95% CI remain of both coefficients remain wide, indicating a significant heterogeneity between particular markets. This outcome replicates the central result by Anufriev et al. (2015), and offers a clear interpretation. The asset market features a positive feedback, since the price mechanism (5) leads to almost self-fulfilling predictions (with a near unit root coefficient of $20/21 \approx 0.952$). If the GA agents experiment with trend following heuristics while observing price increase, they predict higher prices in the future. This increases asset demand, and thus the realized price itself. The GA agents therefore are further encouraged to extrapolate price trends, which closes the self-reinforcing feedback of investor optimism and pessimism.

The LtO GA agents from the OPTIMIZED_FORE specification, who use an independently optimized price forecast in their trading heuristic, converge to a similar behavior. In particular, the median agent learns forecasting heuristic

$$p_{i,t}^e \approx 0.9p_{t-1} + 0.1p_{t-1}^e + 0.9(p_{t-1} - p_{t-2}),$$

which in comparison with the LtF heuristic has a higher trend coefficient (0.9 instead of 0.7). As we will see later in this section, these GA agents learn to put a positive weight on the expected asset return in their trading heuristic. It follows that we observe the same mechanism as in the LtF case, however, the trading amplifies the feedback and hence results in even stronger trend chasing.

In both model specifications, the trend chasing behavior is only lightly anchored in the past forecasts of the GA agents (with weight $1 - \alpha \approx 0.1$ for the median agent in both specifications). The agents start with a flexible anchor-and-adjustment heuristic, but learn to disregard the anchor, and instead, focus on aggressive adjustment strategies.

In contrast to the previous model specifications, the LtO GA agents in the INDIRECT_FORE model, i.e., the agent who use expected asset return in their trading heuristic, but do not directly optimize the price forecast, seem to by and large disregard price trends. The median GA agent learns here to use

$$p_{i,t}^e \approx 0.2p_{t-1} + 0.8p_{t-1}^e + 0.1(p_{t-1} - p_{t-2}),$$

which is a heuristic that adjusts the anchor quite conservatively. The simplest interpretation is that the GA agents find it difficult to come up with a good price forecasting heuristic, which is natural given that (i) they cannot optimize it directly and (ii) the MC simulations indicate fast oscillations that may be difficult to follow to begin with.

### 4.4 Trading heuristics

The asset market becomes even more unstable when the GA agents try to learn how to directly trade. Figure 5 shows the MC evolution of the parameters of the trading heuristics for the
Price weight $\alpha$  

Trend weight $\beta$  

LtF  

INDIRECT FORE  

OPTIMIZED FORE  

Figure 4: 50-period ahead MC simulation (1000 markets) for the LtF, INDIRECT_FORE and OPTIMIZED_FORE specifications of the GA model: evolution of forecasting heuristics parameters, price weight $\alpha$ (left panels) and trend extrapolation weight (right panels). Red thick line is the median and blue dotted lines are the 95% confidence interval for the GA model.

three LtO model specifications. Unlike for the case of forecasting, the learning patterns here are not always clear-cut, and large heterogeneity prevails (in the sense of wide 95% CI of the used heuristic parameters).

This is clearly visible for the case of the INDIRECT_FORE model, where the 95% CI simply cover whole allowed intervals for all the three parameters, and the median parameters are almost exactly at the halves of their respective intervals. In particular, the median INDIRECT_FORE GA agents learns to trade according to

$$z_{i,t} \approx 0.25z_{i,t-1} + 0.15\rho_{t-1} + 0.15\rho_t^e.$$
Previous trade weight $\chi$  
Past return weight $\phi$  
Expected return weight $\zeta$

**Figure 5:** 50-period ahead MC simulation (1000 markets) for the LtF, INDIRECT_FORE and OPTIMIZED_FORE specifications of the GA model: evolution of trading heuristics parameters, previous trade weight $\chi$ (left panels), past asset return weight $\phi$ (middle panels) and expected asset return weight $\zeta$ (right panels). Red thick line is the median and blue dotted lines are the 95% confidence interval for the GA model.

Unlike in all the other cases, this median behavior has no interpretation on its own. Instead, one should say that the INDIRECT_FORE GA agents can converge to any trading heuristic. It follows that the GA agents in this model specification can generate any market dynamics: stable paths, or oscillations of varying degree of amplitude and frequency. However, recall that these agents also learn to forecast by simply repeating their previous forecast with forecasting heuristic (22). This implies that, unlike for example the OPTIMIZED_FORE GA agents, the INDIRECT_FORE agents do not extrapolate price trends, which mitigates the maximal amplitude of the asset price cycle.

The NO_FORE GA agents, who disregard price forecasting, learn a more tangible behavior. In particular, the median agent in this model specification converges gradually to trading heuristic

$$ z_{i,t} \approx -0.2 z_{i,t-1} + 0.174 \rho_{t-1}. $$

The anchor element in this heuristic comes with a negative sign, so the GA agents learn some degree of mean-reversion behavior. On the other hand, the high coefficient on the past asset return (recall that under RE the coefficient on the expected asset return should be $1/6 \approx 0.167$) means that these GA agents emphasize the adjustment element of their heuristics and learn
to strongly follow the market dynamics. In particular, if the asset price rises (decreases) sufficiently, the asset return becomes positive (negative). The median NO FORE agent increases (decreases) her position by more than the RE factor, which feeds back into the price growth (decrease). This reinforces the GA agents to chase the asset return. On the other hand, the mean-reverting $\chi < 0$ and the lack of forward looking component in the trading heuristic imply that, if the market loses some of its momentum, the GA agents will then adjust their positions more slowly. This explains why the second bubble in the MC simulations (Figure 3b) tends to be smaller than the first one.

The OPTIMIZED_FORE GA agents follow a similar pattern as the NO_FORE. The median agent in this model specification converges to a heuristic of a form

$$z_{i,t} \approx -0.3 z_{i,t-1} + 0.14 \rho_{t-1} + 0.11 \rho_{t}^e.$$

As we seen, in comparison with the NO_FORE model specification, the median GA agent chooses even more negative weight on the anchor (with $\chi \approx -0.3 < -0.2$), and a slightly lower coefficient on the past asset return. The latter is offset by the additional positive weight on the expected asset return, which in turn is based on trend following price heuristic (21). This causes the market in this model specification to be so strikingly unstable.

It is worth noting that the median GA agent in the OPTIMIZED_FORE model specification learns the $\zeta$ coefficient on the expected return which is similar to the rational solution, but not exactly on the spot, since $0.14 < 1/6 \approx 0.1667$. Another deviation from RE is that the median agent chooses non-zero weights on the previous trade and asset return. This result is a striking example of how learning and limitations on cognitive abilities an lead to sub-optimal behavior for the case of even the most primitive aspects of our decision making, and how it forms a self-reinforcing feedback system with the market dynamics. In fact Bao et al. (2017) found that only a quarter of their subjects under the Mixed treatment used the optimal trading rule (3).

5 Comparison of the individual learning between model specifications

What drives the differences between the four model specifications? Recall that in each specification, the GA agents use different heuristics based on different parameter space. Therefore, in order to compare the median behavior of the agents, it is useful to focus on a generalized version of their forecasting and trading heuristics given by

$$z_{i,t} = c_0 y + c_1 z_{i,t-1} + c_2 \rho_{t-1}^e + c_3 \Delta \rho_{t-1} + c_4 \rho_{t-1}.$$

This simple linear rule, just like its special cases, is an example of an anchor-and-adjustment heuristic: every GA agent $i$ starts with the dividend $y$, two anchor terms $z_{i,t-1}$ (previous
trading position) and $\rho_{i,t-1}$ (previous asset return forecast), and adjusts these by extrapolating the last observed price trend $\Delta p_{t-1} = p_{t-1} - p_{t-2}$. Finally, the last observed price $p_{t-1}$ can be interpreted as an additional adjustment term, but in practice its weight is relatively small with $c_3 \approx 0$, and can thought of as a residual term (cf., Bao et al. 2017).

The five $c_k$ coefficients of the generalized heuristic (26) are directly related to the coefficients of the forecasting heuristic (6) of the LtF GA agents, and the three trading heuristics (7), (9) and (13) of the LtO GA agents (see Appendix A for derivation). In practice, the GA agents in the four model specifications learn to use this rule differently, which explains the differences in the asset price dynamics. Table 1 shows the results for the median GA agents from four model specifications.

<table>
<thead>
<tr>
<th>Heuristic arguments</th>
<th>Dividend</th>
<th>Trade</th>
<th>Exp. return</th>
<th>Price trend</th>
<th>Price</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$y$</td>
<td>$z_{i,t-1}$</td>
<td>$\rho_{i,t-1}$</td>
<td>$\Delta p_{t-1}$</td>
<td>$p_{t-1}$</td>
</tr>
<tr>
<td>LtF</td>
<td>0.15</td>
<td>0.1</td>
<td>*</td>
<td>0.0992</td>
<td>−0.0075</td>
</tr>
<tr>
<td>LtF specifications</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NO_FORE</td>
<td>0.174</td>
<td>−0.2</td>
<td>0</td>
<td>0.1827</td>
<td>−0.087</td>
</tr>
<tr>
<td>INDIRECT_FORE</td>
<td>0.1726</td>
<td>0.15</td>
<td>0.12</td>
<td>0.0465</td>
<td>−0.009</td>
</tr>
<tr>
<td>OPTIMIZED_FORE</td>
<td>0.239</td>
<td>−0.3</td>
<td>0.011</td>
<td>0.2345</td>
<td>−0.012</td>
</tr>
</tbody>
</table>

Table 1: Coefficients of the generalized trading heuristic (26) learned by median GA agents under four different model specification. Remark that under the LtF specification, the GA agents always trade optimally conditional on their expected asset return, hence $z_{i,t-1} = \rho_{i,t-1}/6$, which is signified by a star in the Table.

There is a clear relationship between stability of the model specification and the learned behavior, in terms of both the anchor and the adjustment elements of the GA agents' heuristics. We observe in Table 1 that the more unstable the specification, the higher adjustment rate associated with price trend following. In particular, the LtF and INDIRECT_FORE GA agents put a weight of less than 0.1 on the previous price trend, which is half as much as the two other type of LtO agents. This is the reason why the former specifications do not yield the “super-bubbles” as the latter specifications do. Furthermore, the price chasing weight of the OPTIMIZED_FORE GA agents is bigger by around a quarter in comparison with that of the NO_FORE agents, which explains why only the OPTIMIZED_FORE GA agents can coordinate on a cycle of two "super-bubbles".

Second observation is that in the more stable markets the GA agents learn a more conservative anchoring scheme. In particular, the NO_FORE and OPTIMIZED_FORE GA agents, in comparison with the LtF and INDIRECT_FORE GA agents, use a negative coefficient on their previous trade. In other words, they learn mean-reverting behavior in terms of their anchor. This effect has an interesting interpretation. It is often said that the real economic agents are aware of the “no Ponzi scheme” condition. For example, actual finan-
cial investors know that (real) asset prices cannot grow forever, which excludes out rational bubbles (typically through a transversality condition). Obviously the GA agents cannot be attributed with a comprehension of the “no Ponzi” rule, but the mean reversion element of their heuristic is “as-if” they did. This explains why, in spite of the extreme trends, the price in these two model specifications oscillate instead of growing exponentially.

The second part of the anchor, the previous expected asset return, is largely ignored by the **OPTIMIZED FORE** GA agents. In other words, they adjust their price beliefs in an aggressive fashion in comparison to the **LtF** and **INDIRECT FORE** GA agents, who put an approximately tenfold higher weight on that variable.

The literature on Learning to Forecast experiments and models demonstrates that people use simple anchor-and-adjustment forecasting heuristics, and that there exists a feedback between the market environment and types of heuristics that are learned by the agents. For the particular example of the positive feedback economies (such as asset markets), people learn to follow price trends, which leads to self-fulfilling price oscillations. My GA model shows that when the agents are asked to directly learn how to trade an asset, the learning dynamics push them towards the same type of behavior, despite it having a different frame. In other words, Learning to Optimize results in agents who are “as-if” price trend chasing, just like under the Learning to Forecast regime. If we furthermore ask the agents to do both, these two learning patterns amplify each other, resulting in either more extreme, or more complicated price dynamics.

**Result 2.** The GA model confirms the insights of the Learning-to-Forecast literature, and shows that they can be generalized to trading behavior: in asset markets, positive feedback between agent decisions and prices makes the agents learn to chase the price trends.

Why does the trading lead to less stable prices and more trend chasing, as seen in Table 1? The simplest interpretation is based on the Rational Expectations solution to the trading and forecasting problems. As explained in the previous section, both trading and forecasting heuristics of the GA agents have a special “as-if” RE. Recall that these special cases have a significantly different structure. Under the **LtF** model specification, the GA agents should disregard the adjustment part of their heuristic (the price trend), and rely solely on the anchor equal to the fundamental price, that is set $\alpha = \beta = 0$ and find $p^e_{t,t} = p^f$. In other words, under this scenario the rational agents are trying to learn a constant.

The trading agents face an opposite task. Recall that the price equation (4) contains small shocks. As a result, the rational demand changes from period to period, as seen in equation (8). This implies that the traders have to learn two things. First, they need to converge to the fundamental forecast (or whatever is the rational forecast in a more general setting). On the top of that, they need to learn how to react to this forecast. Since their trading heuristic is a linear function of the expected asset return, they need to learn the optimal slope of that function (in this model equal to $\zeta^* = 1/6$).

---

11Recall that since the **LtF** GA agents trade optimally conditional on their expected asset return, the coefficient on $z_{t-1}$ can also be interpreted as a coefficient on $\rho_{t-1}^e$.  

28
As we know from any econometric application, estimating slopes tends to be more challenging than estimating constants. This problem becomes significantly more difficult, when the previous market conditions (i.e., the sample available to the agents) depends endogenously on the behavior of the agent (i.e., on what she has previously learned, see Anufriev et al. (2013b) for a discussion and an example in the setting of Industrial Organization). What complicates the learning even more in the context of this model is that the asset market is a positive feedback environment.

Suppose that in some LtF market, agents learned not to follow price trends, and converged to the neighborhood of the fundamental forecast. In words, they understand that \( \alpha = \beta = 0 \) is the rational heuristic, but they still need to learn the actual \( p^f \). If they keep on experimenting with the forecast level, trying to learn the exact rational solution, the average forecast will most likely alternate around and very close to the fundamental. To be specific, suppose that every agent \( i \) forecasts the fundamental with a small noise, which represents the experimentation with the forecast level. Formally, \( p^e_{i,t} = p^f + e_{i,t} \), where \( e_{i,t} \sim NID(0, v^2) \) is a normally distributed error with some small variance \( v^2 \). Given that \( \sigma^2 = 1 \), the price \( p_t \) becomes normally distributed as

\[
p_t = p^f + \frac{20}{21} \left( \frac{\sum_{i=1}^{6} p^e_{i,t}}{6} - p^f \right) + \varepsilon_t
\]

\[
= p^f + \frac{10}{63} \sum_{i=1}^{6} e_{i,t} + \varepsilon_t.
\]

This is a positive feedback system, because when the average forecast overshoots the fundamental, the realized price will be also above the fundamental. However, the price \( p_t \) remains independent from the previous one \( p_{t-1} \) on the merit of the price equation itself. As a result, the agents have to actively pursue price trends for the prices to become serially correlated.

The LtO market is very different. Now the agents have to learn to forecast the fundamental price, as well to figure out the optimal slope in their trading heuristic. Consider a market, in which the LtO OPTIMIZED FORE GA agents experiment with their forecast as the LtF agents in the previous example, and then use the rational trading heuristic with the exception of the weight on the expected asset return. Formally, they use \( \chi = \phi = 0 \), \( p^e_{i,t} = p^f + e_{i,t} \) and \( \zeta_i = \frac{1}{6} + \hat{e}_{i,t} \), where \( \hat{e}_{i,t} \sim NID(0, \hat{v}^2) \) is a small random error. Then

\[
p_t = p_{t-1} + \frac{20}{21} \sum_{i=1}^{6} \zeta_i \left( p^f + e_{i,t} + y - R p_{t-1} \right) + \varepsilon_t
\]

\[
= p^f + \frac{20}{21} \left( p^f + y - R p_{t-1} \right) \sum_{i=1}^{6} \hat{e}_{i,t} + \sum_{i=6}^{6} e_{i,t} \hat{e}_{i,t} + \varepsilon_t.
\]

Unlike in the LtF case, the random learning mistakes \( \hat{e}_{i,t} \) make the current price \( p_t \) dependent on the previous price \( p_{t-1} \). For example, suppose that the initial price is below the fundamental with \( p_1 < p^f \) (as typically happens in experimental sessions), which implies that the term
If the average agent is overconfident (underconfident) with \( \sum_{i=1}^{6} \hat{e}_{i,t} > 0 \) \((\sum_{i=1}^{6} \hat{e}_{i,t} < 0)\), then the next realized price \( p_2 \) is likely to be higher (lower) than the fundamental, just like in the \text{LtF} case. In this case, however, the deviation of the price \( p_2 \) from the fundamental level scales up with the previous deviation of \( p_1 \). In other words, experimentation with \( \zeta \) causes the price mechanism itself to amplify any initial volatility, and makes learning much more difficult. If so, the agents have even more incentive to abandon the fundamental solution and try to learn some other behavior, in comparison with the \text{LtF} model specification.

In the \text{LtF} setting it is easy to coordinate on price trend chasing due to the positive feedback of the market (Hommes, 2013b). This effect becomes stronger under the \text{LtO} setting. For example, consider the \text{OPTIMIZED_FORE} model specification. Suppose that the typical GA agent observes a positive price trend, and extrapolates it with \( \tilde{p}_t > p_{t-1} \), and that she does not care for the previous asset return (with \( \phi = 0 \)), but is overconfident with \( \zeta > 1/6 \). As seen in equation (28), the price grows faster than for the rational solution \( \tilde{\zeta} = 1/6 \), and yields even higher realized asset return \( \rho_t \), which in turn legitimizes the overconfidence of the GA agents.

This positive feedback does not require overconfidence alone. The GA agent can compensate lower weight on the expected asset return \( \zeta \) with a positive weight on the previous asset return with \( \phi > 0 \), since the realized asset return will remain positive as long as the prices continue to grow. As we saw in the previous section, the GA agents in the \text{OPTIMIZED_FORE} learn to put high weights on both parameters in their trading heuristic (25), together with a strong trend chasing weight in their forecasting heuristic (21). If they do use such weights, the price again grows even faster than under the optimal trading from the \text{LtF} specification, which in turn justifies higher asset return extrapolation on the top of strong price trend extrapolation.

Either case seals an enhanced positive feedback, where extrapolation of (i) price trend and (ii) asset return (last observed or expected) fuels large price oscillations, which in turn fuel chasing of price trend and asset return. Since the agents tend to start far from the fundamental, they are likely to get locked in this regime, and coordinate on strong oscillatory price paths.

This explains why the \text{LtO} model specifications tend to generate more unstable price dynamics. In the \text{NO_FORE} case, the GA agents coordinate on price trend following in the sense of high extrapolation of the previous asset return \( \rho_{t-1} = p_{t-1} + y - p_{t-2} \approx \Delta p_{t-1} \). The \text{OPTIMIZED_FORE} GA agents double this down with price trend chasing and overconfident extrapolation of expected and previous asset return. Finally, the \text{INDIRECT_FORE} GA agents cannot link their price forecast with the realized prices, making their learning much more difficult. This confusion leads to more complicated dynamics, with bursts of trend extrapolation resulting in oscillations that lose momentum more quickly then those in the other two \text{LtO} markets – hence these oscillations become faster, but have lower amplitude.

**Result 3.** If the investors are asked directly to trade, they face an “enhanced” positive feedback. If the agents extrapolate price trends and put a relatively high weight on the previous
or expected asset return, they buy more of the asset. This makes the price grow, which justifies the behavior of the agents: price trend seeking and confident trading. As a result, traders can generate higher price oscillations than forecasters, who react only to price trends, but not to the dynamics of the asset return.

6 Conclusions

Financial instability remains one of the most important policy issues after the Great Recession of the 2008. Literature on behavioral finance demonstrates that the Efficient Market Hypothesis does not have to hold, and instead, asset markets can easily generate off-equilibrium cycles of booms and busts. This is confirmed by experimental evidence, which shows that laboratory subjects often coordinate on non-rational price bubbles. The reason is the positive feedback character of asset markets: if the agents are optimistic about the asset’s future profitability, they invest more in that asset, which in turn increases its price. This results in self-reinforcing dynamics between unstable market dynamics and financial agents learning to chase price trends.

The stylized findings of the experimental literature, however, are still subject to one puzzle: what exactly causes the off-equilibrium learning dynamics of the subjects? Are the financial markets unstable, because the agents fail to learn to forecast asset prices in a rational fashion, because they fail to translate their forecasts into optimal trading strategies – or maybe both? The literature on this topic is scarce, but the existing studies suggest that in fact subjects find both tasks difficult (Bao et al., 2017; Nickerson et al., 2007).

In this paper I study a heterogeneous agent model based on Anufriev et al. (2015), in which agents are asked to forecast price of an asset, trade it, or both, in the setting of the Bao et al. (2017) experimental economy. The agents do not know the rational solution to that market, and instead rely on simple anchor-and-adjustment heuristics in the spirit of Tversky and Kahneman (1974). These general heuristics require specific parametrization, and the agents individually learn one by updating their rules with Genetic Algorithms. The forecasting agents utilize a general adaptive expectation rule with additional price trend extrapolation component. There are three types of traders. The first type uses a heuristic, in which the previous position is adjusted by the last observed asset return. The second and third type add to this the expected asset return, where the second type learns only how to use the trading rule, whereas the third type in addition directly optimizes her forecasting heuristic.

The model yields three contributions. Firstly, it replicates the stylized results of the Bao et al. (2017) experiment. All types of agents learn to chase trend price trends, which results in significant price oscillations. However, the trading agents can coordinate on more diversified oscillations, including “super-bubbles”, in which the price repeatedly shifts between 10% and 350% of the fundamental price (as in the most extreme groups in the Bao et al. (2017) experiment). Secondly, the paper offers a unified framework for Learning to Forecast and Learning to Optimize strands of the behavioral literature, and demonstrates that the insights
of former are relevant also for the latter. The positive feedback aspect of the asset markets pushes both forecasters and traders to a symmetric behavior, where in response to price instability they learn to extrapolate it, reinforcing price oscillations. Final contribution of the model is that it demonstrates that the trend chasing biases of forecasters and traders amplify each other. When the agents have to perform both tasks, market dynamics become particularly unstable and agents adopt even higher degree of overconfidence, in comparison to the case of markets with only forecasters or only traders. These results directly contradict the presumptions of the literature based on Muth (1961) by demonstrating that the asset price oscillations are a robust outcome of individual learning for realistic financial markets.
References


A Generalized anchor-and-adjustment trading/forecasting heuristic

The goal of this appendix is to show that the three trading heuristic of the trading GA agents, as well as the trade based on the forecasting heuristic (6) of the forecasting GA agents, can be represented as a trading rule of a form

\[ z_{i,t} = c_0 y + c_1 z_{i,t-1} + c_2 \rho_{i,t-1} + c_3 p_{t-1} + c_4 (p_{t-1} - p_{t-2}). \]

The LtF agents only forecast the price, which is substituted into the optimal demand (3). Hence, the forecasting heuristic (6) corresponds to the following trading rule:

\[ z^F_{i,t} = \left(1 - \alpha\right)z_{i,t-1} + \alpha y + \frac{\alpha(1 - R)}{6} p_{t-1} + \frac{\beta - (1 - \alpha)R}{6} \Delta p_{t-1}. \]

Notice that since the LtF GA agents always trade optimally conditional on their asset return forecast, an alternative representation of (30) is given by

\[ z^F_{i,t} = \left(1 - \alpha\right)z_{i,t-1} + \alpha y + \frac{\alpha(1 - R)}{6} p_{t-1} + \frac{\beta - (1 - \alpha)R}{6} \Delta p_{t-1}. \]

Thus the median heuristic (34) yields

\[ z^F_{i,t} \approx 0.15y + 0.1z_{i,t-1} - 0.0075p_{t-1} + 0.0992\Delta p_{t-1}. \]

The NO.FORE traders adjust their previous trading position with the observed asset return, and so heuristic (7) can be rewritten as

\[ z^{TN}_{i,t} = \chi z_{i,t-1} + \phi \rho_{t-1} \]

\[ = \chi z_{i,t-1} + \phi (p_{t-1} + y - R\rho_{t-2}) \]

\[ = \phi y + \chi z_{i,t-1} + (1 - R)\phi p_{t-1} + R\phi \Delta p_{t-1}. \]

Thus the median heuristic (34) yields

\[ z^{TN}_{i,t} \approx 0.174y - 0.2z_{i,t-1} - 0.087p_{t-1} + 0.1827\Delta p_{t-2}. \]

The INDIRECT.FORE and OPTIMIZED.FORE GA traders both adjust their previous position with previous and expected return (though they learn the expected return in a different way). Thus, using \( p^E_{i,t} = \alpha p_{t-1} + (1 - \alpha)p^E_{i,t-1} + \beta (p_{t-1} - p_{t-2}) \), we have that heuristics
(9) and (13) can be both represented as

\[ z_{TF}^{i,t} = \chi z_{i,t-1} + \phi \rho_{t-1} + \zeta \rho_t^e \]
\[ = \chi z_{i,t-1} + \phi (p_{t-1} + y - Rp_{t-2}) + \zeta \left( p_{t,1}^e + y - Rp_{t-1} \right) \]
\[ = \chi z_{i,t-1} + \left( \phi + \zeta \alpha + \zeta \beta - \zeta R \right) p_{t-1} - (\phi R + \zeta \beta) p_{t-2} + \zeta (1 - \alpha) p_{t-1}^e + (\phi + \zeta) y \]
\[ = \chi z_{i,t-1} + \zeta (1 - \alpha) \left( p_{t,1}^e + y - Rp_{t-2} \right) + (\phi + \zeta \alpha + \zeta \beta - \zeta R) p_{t-1} \]
\[ - (\phi R + \zeta \beta - \zeta (1 - \alpha) R) p_{t-2} + (\phi + \zeta - \zeta (1 - \alpha)) y \]
\[ = (\phi + \alpha \zeta) y + \chi z_{i,t-1} + \zeta (1 - \alpha) \rho_{i,t-1}^e + (1 - R) (\phi + \alpha \zeta) p_{t-1} \]
\[ + (\phi R + \beta \zeta - (1 - \alpha) R \zeta) \Delta p_{t-1}. \]

(35)

It follows that the median trading rule (23) of the INDIRECT_FORE GA agents can be represented as

(36) \[ z_{TI}^{i,t} = 0.1726 y + 0.15 z_{i,t-1} + 0.12 \rho_{i,t-1}^e - 0.009 p_{t-1} + 0.0465 \Delta p_{t-1}, \]

while the median trading rule (25) of the OPTIMIZED_FORE GA agents can be rewritten as

(37) \[ z_{TO}^{i,t} = 0.239 y - 0.3 z_{i,t-1} + 0.011 \rho_{i,t-1}^e - 0.012 p_{t-1} + 0.2345 \Delta p_{t-1}. \]