Experiments on Expectations in Macroeconomics and Finance*

Tiziana Assenza[a,b], Te Bao[c], Cars Hommes[a], and Domenico Massaro[a]

[a] *CeNDEF, University of Amsterdam and Tinbergen Institute*

[b] *Department of Economics, Catholic University of Milan*

[c] *Department of Economics, Econometrics and Finance, University of Groningen*

July 31, 2014

Abstract

Expectations play a crucial role in finance, macroeconomics, monetary economics and fiscal policy. In the last decade a rapidly increasing number of laboratory experiments have been performed to study individual expectation formation, the interactions of individual forecasting rules and the aggregate macro behavior they co-create. The aim of this chapter is to provide a comprehensive literature survey on laboratory experiments on expectations in macroeconomics and finance. In particular, we discuss the extent to which expectations are rational or may be described by simple forecasting heuristics, at the individual as well as the aggregate level.

JEL Classification: D84, D83, E32, C92

Keywords: Expectation feedback, self-fulfilling beliefs, heuristic switching model, experimental economics

*We would like to thank the Editor John Duffy and two anonymous referees for detailed and helpful comments on an earlier draft. We gratefully acknowledge the financial support from the EU FP7 projects "Complexity Research Initiative for Systemic Instabilities" (CRISIS, grant No. 288501), "Macro-Risk Assessment and Stabilization Policies with New Early Waring Signals" (Rastanews, grant No. 320278), "Integrated Macro-Financial Modelling for Robust Policy Design" (MACFIN-ROBODS, grant No. 612796) and the INET-CIGI Research Grant "Heterogeneous Expectations and Financial Crises" (HExFiCs, grant No. INO1200026). Email addresses: Tiziana Assenza: tiziana.assenza@unicatt.it, Te Bao: t.bao@rug.nl, Cars Hommes: c.h.hommes@uva.nl and Domenico Massaro: d.massaro@uva.nl.

Contents

1 Motivation and Introduction

Modeling of expectation formation is ubiquitous in modern macroeconomic and finance where agents face intertemporal decisions under uncertainty. Households need to form expectations about future interest rate and housing prices when deciding whether to buy a new apartment and how to finance it with a mortgage loan. Similarly, fund managers need to form expectations about future stock prices to build an optimal investment portfolio. Based on their expectations, individual agents make their economic decisions, which then via market clearing, determine the realization of the aggregate variables that agents attempt to forecast. The market/economy can be modeled as an *expectation feedback system*, where agents base their expectations on available information and past realizations of the variables, and future realization of the aggregate variables depend on their expectations. It is therefore of crucial importance how the individual expectation formation process is modeled.

Economists have different views towards expectation formation. Since the seminal works by Muth (1961) and Lucas (1972), the rational expectations hypothesis (REH) has become the mainstream approach to model expectation formation. The REH assumes that agents use all available information and, on average, are able to make unbiased predictions of future economic variables, i.e. without systematic errors. When all agents form rational expectations, the economy will reach the rational expectations equilibrium (REE), the solution obtained after plugging in the REH condition to the model of the economy. In dynamic economic models, REH means that expectations are assumed to be model-consistent. REH has been applied widely because of its simplicity and strong discipline on the number of free parameters, but has been criticized by making highly demanding assumptions about agents' knowledge about the law of motion of the economy and their computing capacity.

There is an alternative, behavioral view that assumes agents are boundedly rational. This view dates back to Simon (1957), and has been advocated recently by Akerlof and Shiller (2009), Colander et al. (2009), DeGrauwe (2009), Hommes (2013) and Kirman (2010), among others. A strong motivation for this approach is that expectation formation observed from participants in real market does not seem to be rational. For example, Shiller (1990) and Case et al. (2012) find that during the housing market boom, investors in the US housing market expected housing prices to grow at an extremely high rate that cannot be supported by reasonable estimates of fundamental factors, and they also expected an even higher long run growth rate, although the growth was obviously not sustainable, and the market collapsed soon afterwards. An alternative theory of adaptive learning (Sargent, 1993, Evans and Honkapohja, 2001) has been developed, and there are also empirical works assuming agents have heterogeneous expectations (Harrison and Kreps, 1978, Brock and Hommes (1997), Branch, 2004, Xiong and Yan, 2010, Hommes, 2011). The bounded rationality approach sometimes confirms that the market will converge to the rational expectation equilibrium when the equilibrium can be found by the agents via learning, but often leads to non-RE equilibria (Bullard, 1994), or bubble-bust phenomena, e.g. when the evolutionary selection leads to agents using trend extrapolation forecasting strategies (Brock and Hommes, 1998; Anufriev and Hommes, 2012). Parallel to the literature on bounded rationality, Soros (2003, 2009) uses the terminology "reflexivity" to describe trading behavior in real financial markets, where the realized price goes up when the average expectation goes up, and the agents have a tendency to ignore information about the fundamental value of assets and instead make speculative demand based on trend following expectations, which leads to an intrinsic tendency for the market to destabilize. Understanding the way agents form expectations is not only an interesting topic for academic discussion, but also a relevant issue for policy design. For example, if agents are able to form rational expectations, and reach the rational expectations equilibrium immediately, there would be no need for policy makers to take time lags of the effect of a policy into account. On the other hand, if agents adjust their expectations adaptively, it would be very important for the policy makers to know how quickly they can learn the effect of the policy in order to decide the optimal timing for implementation.

While expectation formation plays a very important role in modern dynamic macroeconomic modeling, there are usually no empirical data on agents' expectations. Occasionally survey data on expectations on future macroeconomics variables (e.g. inflation) is available, but the surveys typically pay a fixed reward, which generates no incentive to provide an answer with careful consideration. When the data on expectations are missing, empirical works on dynamic macroeconomic models face the difficulty of "testing joint hypotheses": namely, when a model is rejected, it is not clear whether it is because of misspecification of the model, or incorrect assumption on the expectation formation rules. In this chapter, we review laboratory experiments with human subjects on expectation formation in dynamic macroeconomic and asset pricing environments. The advantages of using data from the lab include (1) the agents' expectations are explicitly elicited and properly incentivized, which makes expectation formation directly observable; (2) while it is very difficult to find the "true model" of the real macroeconomy, and hence the rational expectation equilibrium/equilibria, with empirical data, the model of the experimental economy is fully known and controlled by the experimenter; and (3) it is very difficult to get empirical data on expectations on macroeconomic variables with a large number of observations or high frequency, while in experiments, it is easier to elicit expectations for many periods within a short period of time.

This paper surveys three types of macroeconomics experiments with elicitation of expectations/forecasts: (i) experiments where agents predict time series from field data or generated by random process (e.g. random walk). These experiments show mixed result about whether the agents are able to make rational expectations on the exogenously generated time series. Some studies show that agents are able to form optimal predictions (e.g. use naive expectations to predict a random walk process), while others find that agents' forecasting behavior is better explained by models with bounded rationality (e.g. models with "regime shifting belief"). (ii) *Learning-to-*

Forecast Experiments (LtFEs), an experimental design that dates back to a series of papers by Marimon and Sunder (1993, 1994, 1995) and Marimon, Spear and Sunder (1993). The key feature of the design is that the subjects of the experiment play the role of professional forecasters, with the only task to submit their expectation about an economic variable, e.g. the market price. After collecting individual forecasts, the conditional optimal quantity decision (e.g. production, trading and saving) by the agents are calculated by a computerized program, for example derived from utility and profit maximization, which then determines the aggregate variables, e.g. the market price, via market clearing. Unlike the experiments in (i), the time series in the LtFEs are a function of the agents' expectations. The LtFEs are forecasting experiments with feedback. A general conclusion from this literature, to be discussed below, is that the agents learn to play the rational expectations equilibrium when the market is a *negative feedback system*, where the realized value of the economic variable (e.g. the price) is low when the average expectation is high (as in a traditional cobweb market), but agents fail to learn to rational expectation equilibrium when the market is a *positive feedback system*, where the realized value is high when the average expectation is high (as in a speculative asset market). (iii) Works that compare the *LtFEs* with the *Learning-to-Optimize Experiments* (LtOEs) design, where the subjects submit their quantity decisions directly. The main conclusion is that the main result of the LtFEs is a robust finding, namely, there is also a tendency for the negative feedback markets to converge to the REE, and the positive feedback markets to deviate from it with the LtOE design. When a learning to forecasting experiment is run using the learning to optimize design, the aggregate market outcome deviates further from the REE, namely, the negative feedback markets converge slower to the REE (i.e. after a larger number of periods), and the positive feedback markets experience more severe boom-bust cycles.

Within the experimental economics literature, there is a parallel literature in micro-economics and game theory about belief/expectation elicitation in games. Since in many game theoretical models, people form a belief/expectation on their opponents' action before choosing their own actions, when agents deviate from the optimal response to their opponent, it is important to understand whether it is because they do not form the right belief about their opponent, or fail to make conditional optimal decisions to their belief. Important studies in this field include Nyarko and Schotter (2002), Rustrom and Wilcox (2009), Blanco et al. (2010) and Gaecheter and Renner (2010); see the recent survey Schotter and Trevino (2014). The main conclusion is that belief elicitation provides a lot of useful information about agents' decision process, but it can be intrusive to their decisions itself. Unlike subjects in the belief elicitation experiments in game theory, subjects in the learning to forecast experiments (LtFEs) form an expectation on the aggregate market outcome, because they are unable to recognize who are their direct opponents. Game theory experiments did not elicit the forecast alone without asking for agents' decisions or actions, while many learning to forecast experiments ask for only the forecast. Moreover, many macroeconomics models also assume that the market is fully competitive and individuals do not have market power. Because of that subjects in LtFEs are typically paid according to their prediction accuracy instead of the profit of the quantity decision, so that they do not have an incentive to use their market power to manipulate the price. Furthermore, macroeconomic experiments also has larger group size (6-12 in each experimental market) than game theory experiments (2-6 in each group) in order to mimic a competitive market. Finally, another characteristic feature of many macro experiments is that subjects typically only have qualitative information about the economy, e.g. whether the feedback structure is positive or negative, but lack detailed quantitative knowledge about the law of motion of the economy.

The chapter is organized as follows. Section 2 reviews forecasting experiments of field data or exogenous stochastic processes whose realizations are not affected by forecasts. Section 3 reviews learning to forecast experiments with expectation feedbacks, where realizations of aggregate variables depend upon expectations of a group of individuals. Section 4 compares learning-to-forecast and learning-to-optimize experiments. Finally, Section 5 concludes.

2 Predicting exogenous processes and field data

Early work on lab experiments on individual forecasting behavior focused on exogenously generated time series, either from real world market data or from simple stochastic processes. In this setup there is no feedback from forecasts. It is like predicting the weather, where forecasts do not affect the probability of rain or the laws of motion of the atmosphere. An advantage of real world data is obviously its realism and relevance for economic forecasting. However in this framework, defining rational expectations is not immediately obvious as the data generating process of the real world time series data is not known. If on the other hand the time series to be forecasted is generated by a (simple) exogenous stochastic process that is known to the experimenter, deviations from rational model-consistent expectations can be more easily measured. This section reviews a number of early contributions in the experimental literature on how agents behave when forecasting an exogenous process.

Forecasting field data

One of the first contributions goes back to the late 70s by Richard Schmalensee (1976), building on earlier work in Fisher (1962). The author describes an experimental analysis on how subjects form expectations on deflated British wheat prices. In particular Schmalensee focusses on the effects of turning points on forecasting behavior. Each participant could observe twenty five realizations of the time series and she was told that the series refers to yearly actual wheat prices in a period with no big political changes. Subjects could observe plots for the time series and five years averages ($1-5$, $2-6$, ..., $21-25$). The individuals had to provide their best forecast for the next five years average (i.e. 26-30)[1]. The paper aims at investigating whether turning points in the time series are points in which important changes in expectations formation take place. In order to conduct the analysis the author applies alternative expectation formation rules including trend following and adaptive expectations rules and analyzes if there are differences in agents' behavior around turning point periods. Schmalensee finds that the adaptive model performs much better than the extrapolative model. Similar as in Fisher (1962), Schmalensee finds that turning points in the time series are special to the participants in the experiment. He introduces in his analysis a parameter that captures the speed of response in the adaptive model (in order to check whether this parameter changes in turning points) and he finds that the parameter drops during turning points.

More recently, Bernasconi et al. (2009) conduct a laboratory experiment in order to study expectations about fiscal variables using real world time series data from 15 European countries. The authors use an estimated VAR of the real world data as a benchmark for comparison. Participants in the experiment are shown graphical representations of annual data (as percentage of GDP) of gross total taxes (T_t), total public expenditure (G_t), public debt (B_t) and the change in the debt level ($\Delta B_t = B_t - B_{t-1}$). Subjects are aware they are observing data from European countries, but they do not have any information about which country they are observing or the time period. At the beginning of the experiment subjects observe the first 7 periods realization of the time series (for most of the countries it coincides with the period between 1970 and 1976), then they have to give their forecast for the next period until the end of the time series (in 1998). Once the first run ended each participant was randomly assigned to another country. The authors conduct three different treatments of the experiment. The first one is the benchmark treatment where participants are asked to predict both T_t and G_t. The second treatment is

[1]Subjects also had to provide a forecast of their confidence in their own forecast and faced higher costs (i.e. lower earnings) when their forecasts was outside their confidence interval.

labeled as the "neutral" treatment, where the time series are the same as in the baseline treatment but any economic framing is removed, and the time series are just labeled "A" and "B". The neutral treatment is useful to check whether the participants have understood the economic context of the baseline treatment or not and whether this is helpful in forecasting. Finally, the third treatment is a "control" treatment were subjects have to predict T_t only; the third treatment is useful to understand whether forecasting two variables simultaneously is too demanding for subjects, and whether asking them to predict only one variable helps to improve their forecasting performance. By comparison and analysis of expectations schemes in the experimental data and the estimated VAR benchmark the authors find that subjects violate the rational expectations hypothesis and their expectations scheme follow an "augmented–adaptive" model. That is, subjects do not always follow a pure, univariate adaptive expectations scheme, but other (e.g. fiscal) variables enter the adaptive updating rule. They also find that forecasts in the neutral case are less precise, so that economic context improves forecasting performance in this setting.

Forecasting exogenous stochastic time series

There have been quite a number of experiments on forecasting behavior of time series generated by an exogenous stochastic process. Perhaps the simplest example has been investigated in Dwyer et al. (1993), where the time series were generated by a simple random walk, that is,

$$x_t = x_{t-1} + \epsilon_t, \tag{1}$$

Where ϵ_t are IID stochastic shocks. Subjects were asked to forecast different series of "events" and were informed that past observations were informative in order to learn the exogenous generating process of the time series and that "forecasts have NO effect" on the realized values. Dwyer et al. did not find that participants were affected by systematic bias or that they were inefficiently using available information. They conclude that subjects' forecasts can be reasonably described by the rational expectation augmented or decreased by a random error and thus they find support for rational expectations. For a random walk, the rational forecast coincides with naive expectations, that is, the forecast equals the last observation. Apparently, with time series generated by a random walk, subjects learn to use the simple naive forecast, which coincides with the rational forecast.

Hey (1994) has conducted a laboratory experiment in which subjects are asked to give forecasts for a stochastic AR(1) time series[2]

$$x_t = 50 + \rho(x_{t-1} - 50) + \epsilon_t, \tag{2}$$

All with the same mean 50, but with different persistence coefficients ρ, where ϵ_t are IID stochastic shocks from a normal distribution with mean 0 and variance 5. Agents, at any time during the experiment, could choose in which form to observe past values of the time series, i.e. in table or graphical representation or both, and the time window. The main focus of the paper consists in analyzing whether agents use rational or adaptive expectations. Differently from Dwyer et al. (1993), who found that rational expectations performs pretty well, the main finding of Hey's paper is that "subjects were trying to behave rationally, but frequently in a way that appears adaptively". The authors estimated a general expectation rule and conduct F test on the coefficients. In this way, the subjects can be categorized to users of adaptive expectations, rational expectations or users of a mixture of the two. The distribution of the subjects over the rule differs depending on which time series the subjects have to predict (series 1, 2 or 3).

Beckman and Downs (2009) ran an experiments where subjects also had to forecast a random walk as in (1), but with varying levels of the noise ϵ_t . Each participant took part in four treatments, with the noise drawn from a uniform distribution of different size, and had to provide 100 predictions for each noise level. Once the experimental data were collected the authors compared them with a survey amongst professional forecasters conducted by the Philadelphia Federal Reserve Bank. The

[2]One treatment had a structural break, with the coefficient b switching from 0.1 to 0.8.

main finding of the paper is that, for both the experimental data and the survey data, as the variance of the random walk increases deviations from the theoretically correct prediction strategy increases (i.e. naive expectations) as well. Indeed a 1% increment in the random error standard deviation implies a 0.9% increase in the standard deviation of the forecast of the rational expectations rule.

Bloomfield and Hales (2002) conducted an experiment where MBA students were shown some time series generated by the random walk process. They use the data to test the "regime shifting belief" model by Barberis et al. (1998) that predicts the individuals use the number of past trend reversals to evaluate the likelihood of future reversals. Their results show support to the model. The subjects did not seem to perceive the random walk time series as randomly generated, and tended to predict more reversals if they experienced more reversals in the past. However, Asparouhova et al. (2009) found evidence against "regime shifting belief" models in favor of "law of small numbers" in Rabin (2002), namely, the subjects do not predict a continuation of the current streak when the streak is longer.

Kelley and Friedman(2002) consider learning in an Orange Juice Futures price forecasting experiment, where subjects must learn the coefficients of two independent variables in a stationary linear stochastic process. They find that learning is fairly consistent with respect to objective values, but with slight tendency toward over-response. Moreover, learning is noticeably slower than under adaptive learning. Two striking treatment effects are tendencies toward over-response with high background noise and under-response with asymmetric coefficients

Becker et al. (2009) conduct a laboratory experiment in which participants have to predict three time series subject to regime switching. First a stationary stochastic time series with integer values was generated, after which regime switches where applied by adding a constant mean in different subperiods. The main focus of the paper consists in explaining the average forecasts by means of the bound and likelihood heuristics model (B&L heuristic hereafter) by Becker and Wildburger (1996), according to which two features of the time series are most important in forecasting, namely turning points in the time series and the average variation. The authors find that after a regime switch the agents' forecasts show a higher variance and less accuracy for several periods after the structural break in the time series they observe. Hence the heuristic performs slightly better than the Rational Expectation Hypothesis. In order to explain the average forecast the B&L heuristic and the Rational Expectation Hypothesis are applied to the three different treatments of the experiment. The authors find that if the periods immediately after the break has occurred are considered as a transition phase then the B&L heuristic explains subjects' forecasting behavior even if the time series taken into account is affected by different breaks. In fact individuals have memory of the pre-break periods.

Beshears et al. (2013) ask subjects to forecast an integrated moving average (ARIMA) process with short-run momentum and long-run mean reversion. Subjects make forecasts at different time horizons. They find that subjects have difficulty in correctly perceiving the degree of mean reversion, especially in the case with a slow dynamic process.

3 Learning-to-forecast

In this section we consider Learning-to-Forecast Experiments (LtFE) with human subjects. The Learning-to-Forecast design has been pioneered in a series of papers by Marimon and Sunder (1993, 1994, and 1995) and Marimon, Spear and Sunder (1993), within dynamic Overlapping Generations Models; an earlier survey of LtFEs is given in Hommes (2011). Subjects have to forecast a price, whose realization depends *endogenously* on their average forecast. The key difference with the previous section is the *expectations feedback* in these systems. Subjects are forecasting within a self-referential system: their individual forecasts affect and co-create aggregate behavior, which then leads to adaptations of individual forecasts. The main goal of these experiments is to study how, within a dynamic self-referential economic system, individual expectations are formed, how these interact and which structure emerges at the aggregate level. Will agents coordinate on a common forecast and will the price converge to the rational expectations benchmark or will other, learning equilibria arise?

As already noted in Muth's classical paper introducing rational expectations, a crucial feature for aggregation of individual expectations, is whether the deviations of individual expectations from the rational forecast are *correlated* or not. To quote Muth (1961, p.321, emphasis added):

"Allowing for cross-sectional differences in expectations is a simple matter, because their aggregate affect is negligible *as long as the deviation from the rational forecast for an individual firm is* not strongly correlated with those of the others. *Modifications are necessary only if the correlation of the errors is large and depends systematically on other explanatory variables".*

Laboratory experiments are well suited to study correlation of individual expectations in a controlled self-referential environment. It turns out that the type of expectations feedback, positive or negative, is crucial. In general, the market price quickly converges to the REE in the negative feedback markets, and fails to converge in the positive feedback markets.

3.1 Asset pricing experiments

This section reviews two closely related learning to forecast experiments by Hommes et al (2005, 2008) on a speculative asset market. These experiments are based on the dynamic asset pricing model (e.g., Campbell et al., 1997), where the investor allocates his wealth between two assets. One asset is riskless, and pays a fixed gross return R, and the other is risky, paying an uncertain dividend y_t in each period, y_t is i.i.d. with mean dividend $\overline{y}$. The price of the risky asset is determined by market clearing condition, and the supply of the asset is normalized to 0. The demand for the risky asset by each individual i at period t is denoted by $z_{i,t}$. This demand function is derived from mean-variance maximization of next period expected wealth:

$$\max U_{i,t+1}(z_{i,t}) = \max \bar{E}_{i,t}\{W_{i,t+1}(z_{i,t}) - \frac{a}{2}\bar{V}_{i,t}(W_{i,t+1}(z_{i,t}))\}. \tag{3}$$

$E_{i,t}$ and $V_{i,t}$ are the subjective beliefs of agent i about the mean and variance of next period's wealth. Expected wealth can be rewritten in terms of the demand $z_{i,t}$ as

$$\max z_{i,t}\{(p_{t+1} + y_{t+1} - Rp_t) - \frac{a\sigma^2 z_{i,t}^2}{2}\}, \tag{4}$$

where we assumed homogeneous and constant beliefs about the variance of excess returns, i.e. $V_{i,t}(p_{t+1} + y_{t+1} - Rp_t) = \sigma^2$ for all agents. The optimal demand is:

$$z_{i,t}^* = \frac{E_{i,t}p_{t+1} + \overline{y} - Rp_t}{a\sigma^2}. \tag{5}$$

Imposing the market clearing condition:

$$\sum_i z_{i,t}^* = \sum_i \frac{E_{i,t}p_{t+1} + \overline{y} - Rp_t}{a\sigma^2} = z_t^s = 0. \tag{6}$$

The market clearing price then becomes:

$$p_{t+1} = \frac{1}{R}\left(\bar{p}_{t+1}^e + \overline{y}\right) + \varepsilon_t, \tag{7}$$

where $\bar{p}_{t+1}^e = \frac{\sum_i E_{i,t}p_{t+1}}{I}$ is the average prediction by the investors ($I = 6$ in the experiments) and $\varepsilon_t \sim NID(0,1)$ (i.i.d. with normal distribution) is a small noise term added to the pricing equation (representing e.g. a small fraction of noise traders). Both experiments use the parameter setting $R = 1 + r = \frac{21}{20}$ (or equivalently, the risk free interest rate is 5%) and $\overline{y} = 3$. Therefore, by substituting in the rational expectations condition, and ignoring the noise term ε_t with zero mean, the rational expectations equilibrium or the *fundamental price* of the experimental markets is $p^f = \overline{y}/r = 60$.

In both experiments, the subjects play the role of investment advisor of pension funds, and submit a price forecast for period $t+1$ repeatedly. The market price in t

is a function of the average price forecast by the subjects. The key difference between Hommes et al. (2005) and Hommes et al. (2008) is the presence of a computerized fundamental robot trader, always trading based upon the forecast that price equals fundamental. The fundamental trader acts as a "far from equilibrium stabilizing force", pushing prices back towards its fundamental value. More precisely, Hommes et al. (2005) used the pricing rule

$$p_{t+1} = \frac{1}{R}\left((1-n_t)\bar{p}^e_{t+1} + n_t p^f + \bar{y}\right) + \varepsilon_t, \tag{8}$$

With the weight assigned to the robot trader n_t given by

$$n_t = 1 - \exp\left(-\frac{1}{200}|p_{t-1} - p^f|\right). \tag{9}$$

The fraction of fundamental robot trader is 0 at the fundamental price $p^f = 60$ and becomes larger when the price deviates further from the fundamental price, with an upper limit of 0.26.[3]

Figure 1 illustrates market prices and individual forecasts in three groups in Hommes et al. (2005). Prices are very different from the rational expectation equilibrium benchmark $p^f = 60$. Within the same treatment three different types of aggregate price behavior are observed: (i) slow monotonic convergence to the fundamental price (top panel), (ii) persistent oscillations (middle panel) and (iii) dampened price oscillations (bottom panel). Another striking feature is that individual expectations are strongly coordinated, despite the fact that subjects have no information about other subjects' forecasts and only communicate through the observed price realization.

[3] The robot trader can be considered as fundamental traders who always buy when the price is below and sell when the price is above the fundamental price. Their weight n_t increases when the price deviates more from the REE, so that the price does not "explode". The intuition behind the increasing weight is that the more the price deviates from the REE, the less likely it is that the deviation will sustain. Knowing this, more fundamental trader will join because they think that mean reversion of the price becomes more likely.

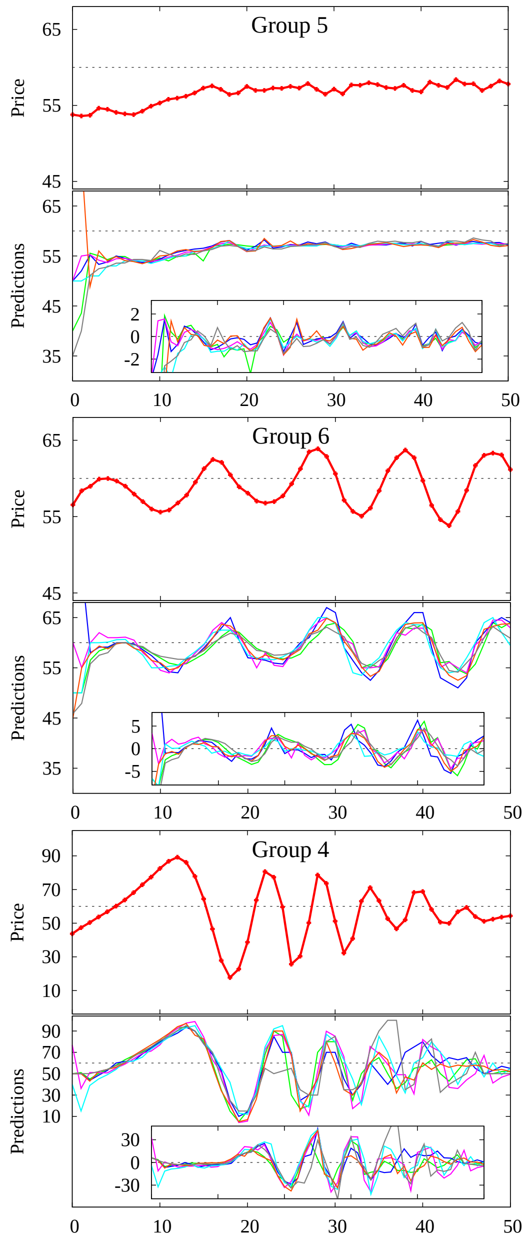


Figure 1: Market prices, individual forecasts and forecasting errors in three groups in Hommes et al. (2005). The rational expectation equilibrium (flat line) is $p^f = 60$. Three different types of aggregate price behavior are observed: slow monotonic convergence to the fundamental price (top panel), persistent oscillations (middle panel) and dampened price oscillations (bottom panel). Individual expectations are strongly coordinated.

Hommes et al. (2008) ran similar experiments *without* the presence of fundamental robot traders, that is, the realized price is generated by (7), where $\bar{p}^e_{t+1} = \frac{\sum_i E_{i,t}p_{t+1}}{I}$ is the average prediction of 6 subjects in the experiments. Figure 2 illustrates aggregate price behavior in six groups, when there are no robot traders in the market. The market price can go to very high level, sometimes more than 10 times the REE before it crashes. These bubbles are driven by strong coordination of individual expectations on trend-following behavior.

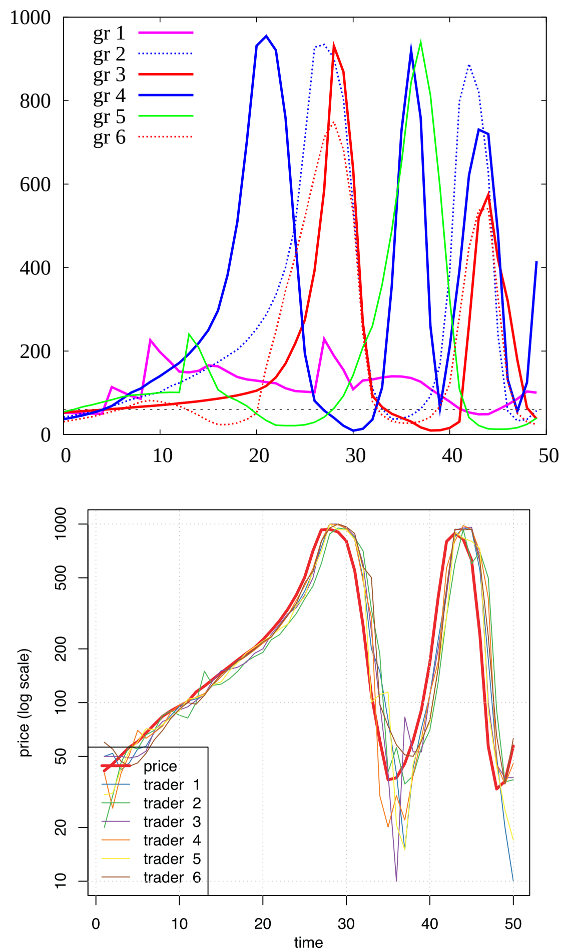


Figure 2: The market prices in six different groups (top panel) in Hommes et al. (2008). Without fundamental robot traders in the market large bubbles and crashes arise, due to strong coordination of individual forecasts (bottom panel).

Heuristics switching model

Prices in the asset pricing LtFEs clearly deviate from the RE benchmark. What would be a good theory of expectations that fits these laboratory data? In order to explain all different observed patterns (convergence, persistent oscillations, dampened oscillations and large bubbles and crashes) heterogeneity of expectations may play a key role to explain the data. Anufriev and Hommes (2012) have developed a Heuristic Switching Model (HSM), extending the heterogeneous expectations framework of Brock and Hommes (1997, 1998) to fit the experimental data. Agents choose from a list of simple "rule of thumbs" to predict, for example naive expectations, adaptive expectations or trend-following rules, and choose their forecast rule based upon its relative success. There is thus evolutionary selection of the rules: heuristics that performed better in the recent past attract more followers in the future. Hommes et al. (2005) and Heemeijer et al. (2009) estimated linear forecasting rules for the individual forecasts and showed that simple linear rules, with only one or two time lags fit individual forecasting behavior quite well. Based on these estimations they could classify the rules into simple classes, such as adaptive or trend-following expectations. Anufriev and Hommes (2012) fitted a HSM with only four heuristics:

- Adaptive expectations (ADA): $p^e_{t+1,1} = p^e_{t-1} + 0.65(p_{t-1} - p^e_{t-1,1})$.
- Weak trend rule (WTR): $p^e_{t+1,2} = p_{t-1} + 0.4(p_{t-1} - p_{t-2})$.
- Strong trend rule (STR) : $p^e_{t+1,2} = p_{t-1} + 1.3(p_{t-1} - p_{t-2})$.
- Learning, Anchoring and Adjustment heuristic (LA&A):
  $p^e_{t+1,4} = 0.5(p^{av}_{t-1} + p_{t-1}) + (p_{t-1} - p_{t-2})$.

The LAA rule is proposed by Tversky and Kahneman (1974), where p^{av}_{t-1} stands for the sample average of past market prices until period $t-1$. The key difference between the LAA and the simple trend-following rules lies in the *anchor* from which price extrapolations occur. For the simple trend-following rule, the anchor is simply the last observed price. This simple rule can easily forecast a price trend, but performs poorly at turning points. In contrast, the LAA rule use an average of the last observed price p_{t-1} and the long run sample average, p^{av}_{t-1}, which serves as a proxy of the average price level. By giving 50% weight to its long run average, the LAA heuristic predicts turning points, when the price moves away from its fundamental or long run value.

In the HSM, consistent with the incentives in the LtFEs, the performance of heuristic h, $h \in \{1, 2, 3, 4\}$ is measured by its squared prediction error of the rule in each period t:

$$U_{t,h} = -(p_t - p^e_{t,h})^2 + \eta U_{t-1,h}, \tag{10}$$

$n_{h,t}$ is the fraction of the agents using heuristic h, and $\sum_h n_{h,t} = 1$. $\eta \in [0, 1]$ is a parameter measuring memory. The updating rule for the weights given to forecast strategy h is given by a discrete choice model with asynchronous updating:

$$n_{t,h} = \delta n_{t-1,h} + (1 - \delta)\frac{\exp(\beta U_{t-1,h})}{\sum_{i=1}^{4} \exp(\beta U_{t-1,i})}. \tag{11}$$

$\beta \geq 0$ is the intensity of choice parameter. The larger the β, the more quickly agents switch to the heuristic that performs well in the recent past. $\delta \in [0, 1]$ is a parameter for inertia. With the benchmark parameter setting $\beta = 0.4, \eta = 0.7, \delta = 0.9$, the model fits each of the three patterns in the data of Hommes et al. (2005) remarkably well. Figure 3 shows the actual and simulated one-step-ahead forecasts of market prices by the HSM model in Anufriev and Hommes (2012). The results of these simulation show that convergence to the REE is driven by a coordination of subjects using the stabilizing adaptive expectations. In markets with persistent oscillations, the evolutionary selection leads to most subjects to coordinate on the LAA rule, with almost 90% of the subjects using the LAA rule. The dominating forecasting strategy changes over time for the market with dampened oscillations. Initially most subjects use the strong trend-rule, followed by a dominating LAA-rule in the middle, while most subjects eventually switch to stabilizing adaptive expectations in the final periods.

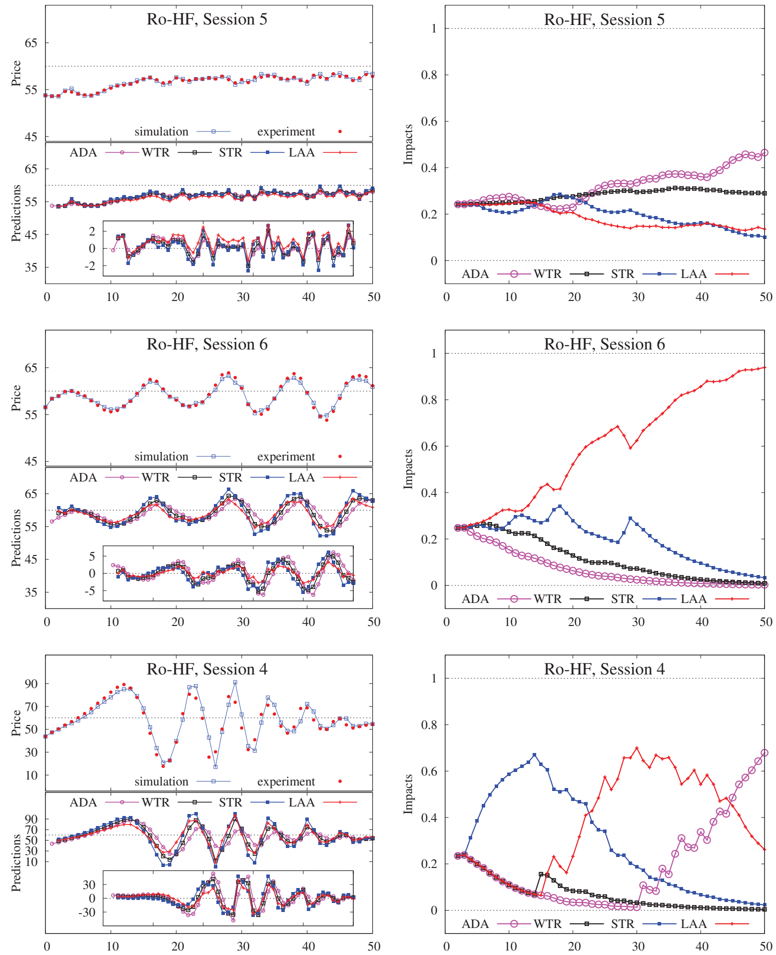


Figure 3: Left panel: top of each graph: the actual and simulated market prices (hollow squares) by the HSM model (dots) in a typical market of each treatment in Anufriev and Hommes (2012); bottom of each graph: simulated price predictions (main figure) and variances (subfigure) by the four rules in the heuristic switching model: ADA for adaptive expectations, WTR for weak trend rule, STR for strong trend rule and LAA for learning, anchoring and adjustment rule. Right panel: the simulated fraction of agents who use each kind of heuristics (hollow circle for ADA, hollow square for WTR, square for STR and + for LAA) by the Heuristic Switching Model in Anufriev and Hommes (2012).

As for the price bubbles in Hommes et al. (2008), there is also a competing theory to explain the emergence of large bubbles. Hommes et al. (2008) discussed the possibility of "rational bubbles", as in Tirole (1982), where the price grows at the same rate as the interest rate. They find however in most markets (4 out of 6), that the growth rate of the price are much higher than the interest rate ($r = R - 1$). H*ü*sler et al (2013) explore the possibility that the bubbles in this experiment can be described as "super-exponential bubbles", with an accelerating growth rate of the price. They discuss two possibilities (1) the growth rate is an increasing function in the price deviation from the fundamental price, $\log(\frac{\overline{p}_t}{\overline{p}_{t-1}}) = a + b\overline{p}_{t-1}$, where $\overline{p}_{t-1} = p_t - p^f$. This means larger price deviation makes the investors overly optimistic/pessimistic, and expect that the deviation will grow faster than the interest rate; (2) the growth rate is an increasing function in the growth rate in the last period, $\log(\frac{\overline{p}_t}{\overline{p}_{t-1}}) = c + d\log(\frac{\overline{p}_{t-1}}{\overline{p}_{t-2}})$. This means it is the return rate instead of the price level that makes investors overly optimistic/pessimistic. They run estimations and find that specification (1) provides the best description of the experimental data.

3.2 Positive versus negative feedback

The asset pricing experiments are characterized by *positive expectations feedback*, that is, an increase of the average forecast or an individual forecast causes the realized market price to rise. Heemeijer et al. (2009) investigate how the expectations feedback structure affects individual forecasting behavior and aggregate market outcomes by considering market environments that *only* differ in the sign of the expectations feedback, but are equivalent along all other dimensions. The realized price is a linear map of the average of the individual price forecasts $p^e_{i,t}$ of six subjects. The (unknown) price generating rules in the *negative* and *positive* feedback systems were respectively:[4]

[4] Our positive and negative feedback LtFEs may be viewed as repeated guessing games or beauty contest games as introduced in Nagel (1995). Sutan and Willinger (2009) study beauty contest games with negative feedback (i.e. players actions are strategic substitutes. Moreover, positive feedback

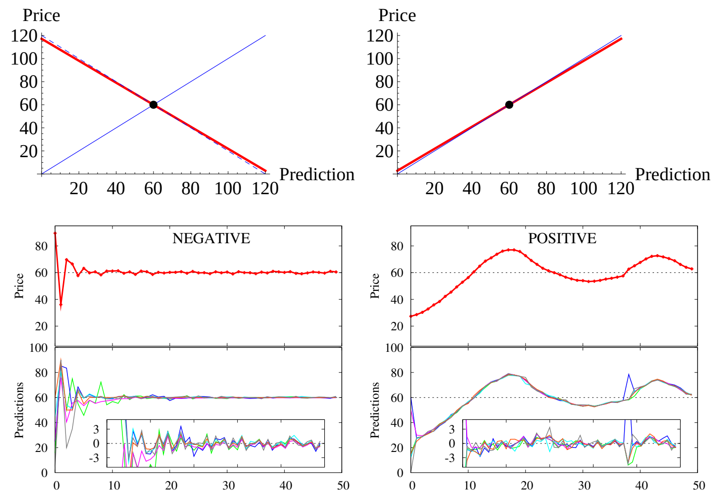


Figure 4: Negative (left panels) vs positive (right panels) feedback experiments. Linear feedback maps (top panels) share the unique RE price at the fixed point 60. The positive feedback map is very close to the diagonal and therefore has a continuum of almost self-fulfilling steady state equilibria. Realized market prices (upper part bottom panels), six individual predictions (middle parts) and individual errors (bottom parts). In the negative expectations feedback market the realized price quickly converges to the RE benchmark 60. In all positive feedback markets individuals coordinate on the "wrong" price forecast and as a result the realized market price persistently deviates from the RE benchmark 60.

$$p_t = 60 - \frac{20}{21}[(\sum_{h=1}^{6} \frac{1}{6} p^e_{ht}) - 60] + \epsilon_t, \qquad \text{negative feedback} \tag{12}$$

$$p_t = 60 + \frac{20}{21}[(\sum_{h=1}^{6} \frac{1}{6} p^e_{ht}) - 60] + \epsilon_t, \qquad \text{positive feedback} \tag{13}$$

where ϵ_t is a (small) exogenous random shock to the pricing rule. The positive and negative feedback systems (12) and (13) have the same unique RE equilibrium steady state $p^* = 60$ and *only* differ in the sign of the expectations feedback map. Both are linear near-unit-root maps, with slopes $20/21 \approx -0.95$ resp. $+20/21$[5].

is similar to strategic complements, while negative feedback is similar to strategic substitutes in terms of strategic environment. Fehr and Tyran (2005, 2008) show that, after an exogenous shock, the market price converges faster to the new fundamental price in an environment with strategic substitutes than with strategic complements.

[5] In both treatments, the absolute value of the slopes is 0.95, implying in both cases that the feedback system is stable under naive expectations. Leitner and Schmidt (2007) study a LtFE in

Fig. 4 (top panels) illustrates the dramatic difference in the negative and positive expectations feedback maps. Both have the same unique RE fixed point, but for the positive feedback map, the graph almost coincides with the diagonal so that every point is almost a steady state. Under near-unit-root positive feedback, as is typical in asset pricing models, each point is in fact an *almost self-fulfilling equilibrium*. Will subjects in LtFEs be able to coordinate on the unique RE fundamental price, the only equilibrium that is perfectly self-fulfilling?

Figure 4 (bottom panels) shows realized market prices as well as six individual predictions in two typical groups. Aggregate price behavior is very different under positive than under negative feedback. In the negative feedback case, the price settles down to the RE steady state price 60 relatively quickly (within 10 periods), but in the positive feedback treatment the market price does not converge but rather oscillates around its fundamental value. Individual forecasting behavior is also very different: in the case of positive feedback, coordination of individual forecasts occurs extremely fast, within 2-3 periods. The coordination however is on a "wrong", i.e., a non-RE-price around 30 and thereafter the price starts oscillating. In contrast, in the negative feedback case coordination of individual forecasts is slower and takes about 10 periods. More heterogeneity of individual forecasts however ensures, that, the realized price quickly converges to the RE benchmark of 60 (within 5-6 periods), after which individual predictions coordinate on the correct RE price.

In his seminal paper introducing RE, Muth (1961) considered a negative expectations feedback framework of the cobweb "hog-cycle" model. Previous LtFEs on cobweb models show that under negative expectations feedback, heterogeneity of individual forecasts around the rational forecast 60 persists in the first 10 periods, and correlated individual deviations from the RE fundamental forecast do not arise (in

an experimental foreign exchange market, which is a positive feedback system with slope +1. For all markets, the realized exchange rate is highly correlates with the (small) noise shocks. Similar to Heemeijer et al. (2009) they estimate simple linear expectations rules to subjects' forecast series and find evidence for adaptive, naive and trend-following expectations rules amplifying fluctuations.

line with Muth's observations as quoted in the introduction) and the realized market price converges quickly to the RE benchmark. In contrast, in an environment with positive expectations feedback the LtFEs show that, within 2-3 periods, individual forecasts become strongly coordinated and all deviate from the rational, fundamental forecast. As a result, in positive expectations feedback markets, at the aggregate level the market price may persistently deviate from the rational, fundamental price. Individual forecasts than coordinate on almost self-fulfilling equilibrium, very different from the perfectly self-fulfilling RE price[6]. Coordination on almost self-fulfilling equilibria has also been obtained in laboratory experiments in a Lucas asset pricing model (Asparouhova et al., 2013, 2014).

Bao et al. (2012) combines Heemeijer et al. (2009) and Hey (1994) and consider positive and negative feedback experiments, with large permanent shocks to the fundamental price level[7]. More precisely, these shocks have been chosen such that, both in the negative and positive feedback treatments, the fundamental equilibrium price p^*_t changes over time according to:

$$\begin{aligned} p^*_t &= 56, & 0 \le t \le 21, \\ p^*_t &= 41, & 22 \le t \le 43, \\ p^*_t &= 62, & 44 \le t \le 65. \end{aligned} \tag{14}$$

The purpose of these experiments was to investigate how the type of expectations feedback may affect the speed of learning of a new steady state equilibrium price, after a relatively large unanticipated shock to the economy.

Figure 5 shows for positive and negative feedback the average price behavior (top panels), realized prices in all groups (middle panels) and an example of individual

[6]See Hommes (2013, 2014) for further discussion of coordination on almost self-fulfilling equilibria in positive feedback systems and the relation to Soros' notion of reflexivity.

[7]Hommes et al. (2000) conduct individual learning-to-forecast experiments with large permanent shocks to the fundamental price in the cobweb model. More recently, Bao and Duffy (2014) conduct a learning-to-forecast experiment with both individual and group settings where the subjects have complete information about the model of the economy.

forecasts in a positive as well as a negative feedback group (bottom panels). Aggregate behaviors under positive and negative feedback are strikingly different. Negative feedback markets tend to be rather stable, with price converging quickly to the new (unknown) equilibrium level after each unanticipated large shock. In contrast, under positive feedback prices are sluggish, converging only slowly into the direction of the fundamental value and subsequently overshooting it by large amounts.

Figure 6 reveals some other striking features of aggregate price behavior and individual forecasts. The top panel shows the time variation of the median distance to the RE benchmark price over all (eight) groups in both treatments. For the negative feedback treatment, after each large shock the distance spikes, but converges quickly back (within 5-6 periods) to almost 0. In the positive feedback treatment after each shock the distance to the RE benchmark shows a similar spike, but falls back only slowly and does not converge to 0. The bottom panel shows how the *degree of heterogeneity*, that is, the median standard deviation of individual forecasts, changes over time. For the positive feedback treatment after each large shock heterogeneity decreases very quickly and converges to (almost) 0 within 3-4 periods. Under positive feedback, individuals thus coordinate expectations quickly, but they all coordinate on the "wrong", i.e., a non-RE price. In the negative feedback treatment heterogeneity is more persistent, for about 10 periods after each large shock. Persistent heterogeneity stabilizes price fluctuations and after convergence of the price to its RE fundamental individual expectations coordinate on the correct RE price.

One may summarize these results in saying that in the positive feedback treatment individuals quickly coordinate on a common prediction, but that coordination on the "wrong" non-fundamental price occurs. As a result price behavior is very different from the perfect rational expectations equilibrium price. On the other hand, in the negative feedback treatment coordination is much slower, heterogeneity is more persistent but price convergence is quick. Stated differently, positive feedback markets are characterized by quick coordination and slow price discovery, while negative feed-

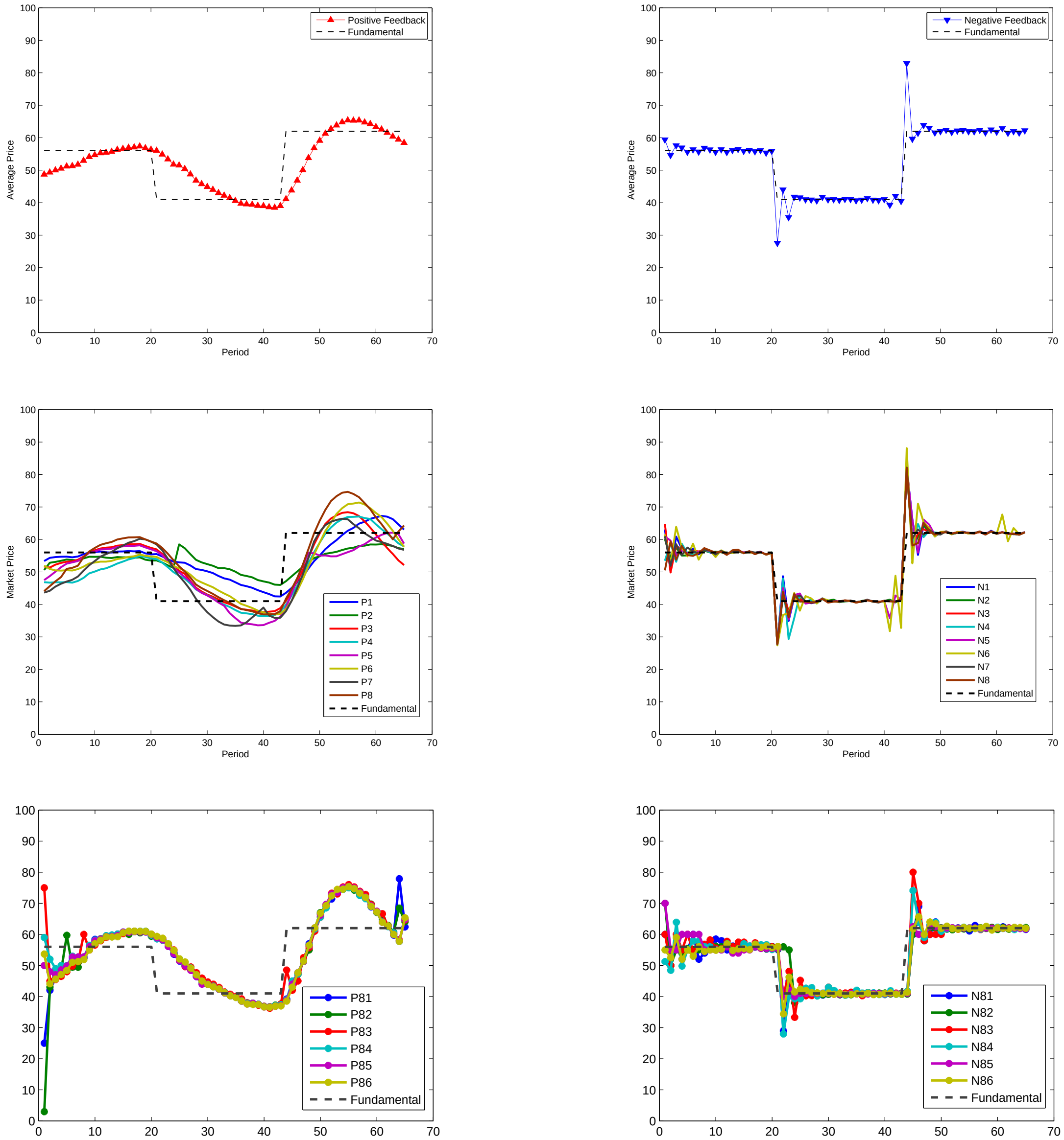


Figure 5: Positive feedback (left panels) and negative feedback (right panels) experimental data. Top panels: The average realized price averaged over all eight groups; Middle panels: the market prices for eight different groups; Bottom panels: predictions of six individuals in Group P8 (left) and Group N8 (right) plotted together with fundamental price (dotted lines).

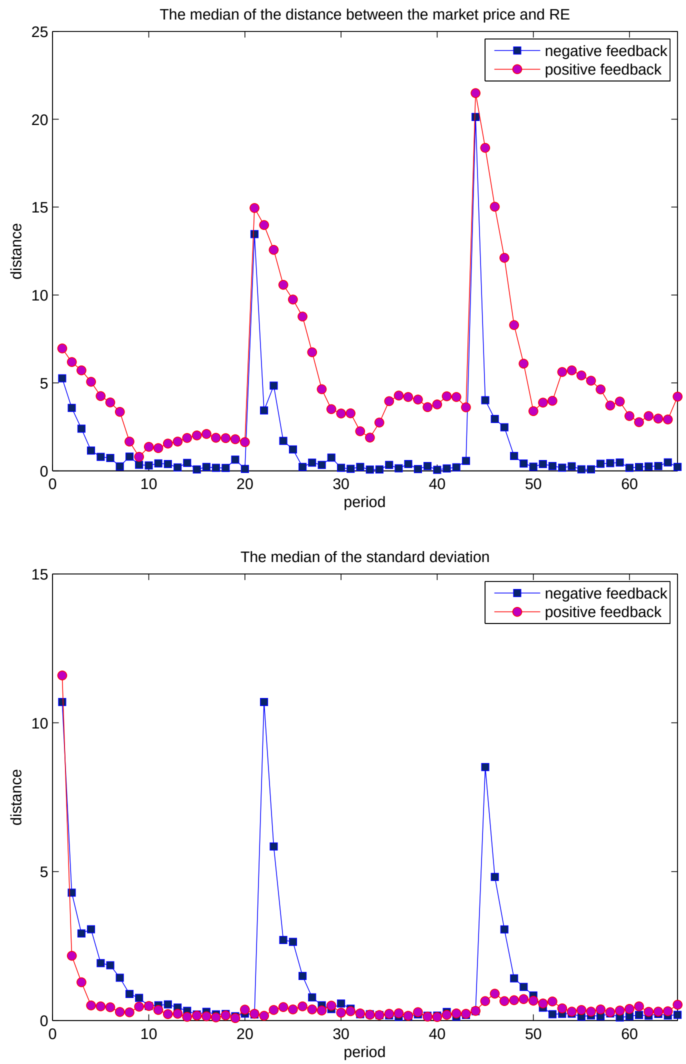


Figure 6: Positive/Negative feedback markets with large shocks. These plots illustrate price discovery (top panel) and coordination of individual expectations (bottom panel). The top panel shows the median absolute distance to RE fundamental price, while the bottom panel shows the median standard deviation of individual predictions. In positive feedback markets coordination is quick, but on the "wrong", i.e. non-RE, price.

back markets are characterized by slow coordination, more persistent heterogeneity and quick price discovery. Notice also that under positive feedback, coordination on a non-RE-fundamental price is *almost self-fulfilling*, with small individual forecasting errors. The positive feedback market is thus characterized by coordination on almost self-fulfilling equilibria with prices very different from the perfectly rational self-fulfilling equilibrium[8].

Similar to Anufriev and Hommes (2012), Bao et al. (2012) fit a heuristics switching model with four rules to these experimental data[9]. The rules are an adaptive expectation (ADA) rule:

$$p^e_{t+1,1} = p^e_t + 0.85(p_t - p^e_{t,1}). \tag{15}$$

A contrarian rules (CTR) given by:[10]

$$p^e_{t+1,2} = p_t - 0.3(p_t - p_{t-1}). \tag{16}$$

A trend extrapolating rule (TRE) given by:

$$p^e_{t+1,2} = p_t + 0.9(p_t - p_{t-1}). \tag{17}$$

The coefficients of the first three rules are the medians of the estimated individual linear rules in Bao et al. (2012). The fourth rule is again led a learning anchor and

[8] Wagener (2013) uses the same experimental data and shows weak individual rationality (i.e. unbiased forecast errors without autocorrelations) for both the negative and positive feedback treatments, but strong rationality (i.e. prices converge to the homogeneous REE price) only under negative feedback.

[9] Anufriev et al. (2013) fit a HSM with two heuristics, adaptive expectations versus a trend-following rule, to the positive-negative expectations feedback experiments of Heemeijer et al. (2009).

[10] Anufriev and Hommes (2012) used two different trend following rules in their model, a weak and a strong trend following rule, to describe asset pricing experiments with positive feedback. Because of the negative feedback treatment in Bao et al. (2012), one trend-following rule was replaced by a contrarian rule, i.e. with a negative coefficient (-0.3) which is able to detect (short run) up and down price oscillation characteristic for negative feedback markets.

adjustment heuristic (LAA) (Tversky and Kahneman, 1974):

$$p^e_{t+1,4} = 0.5(p^{av}_t + p_t) + (p_t - p_{t-1}). \tag{18}$$

As before, subjects switch between these rules depending upon their relative performance.

Figure 7 shows realized market prices and the one-period ahead simulated market prices (top panels), together with the evolution of the fractions of the four strategies of the heuristics switching model (bottom panels) for a typical group of the negative feedback (left panels) and the positive feedback treatment (right panels). The heuristics switching model matches the aggregate behavior of both positive and negative feedback quite nicely and provides an intuitive, *behavioral explanation* why these different aggregate patterns occur. In the negative feedback market, trend following strategies perform poorly and the contrarian strategy quickly dominates the market (more than 70% within 20 periods) enforcing quick convergence to the RE benchmark after each large shock. In contrast, in the positive feedback treatment, the trend following strategy performs well and dominates the market (with more than 50% trend-followers after 10 periods). The survival of trend following strategies in the positive feedback markets causes persistent deviations from the RE steady states, overreaction and persistent price fluctuations.

The difference in aggregate behavior in these experiments is thus explained by the fact that *trend-following rules are successful in a positive feedback environment* amplifying price oscillations and persistent deviations from the rational equilibrium benchmark price, while the same trend-following rules are driven out by the contrarian rule in the case of negative feedback. Coordination of individual expectations on trend-following rules and almost self-fulfilling equilibria in a positive expectations feedback environment has a large aggregate effect with realized market prices deviating significantly from the perfectly self-fulfilling RE benchmark.

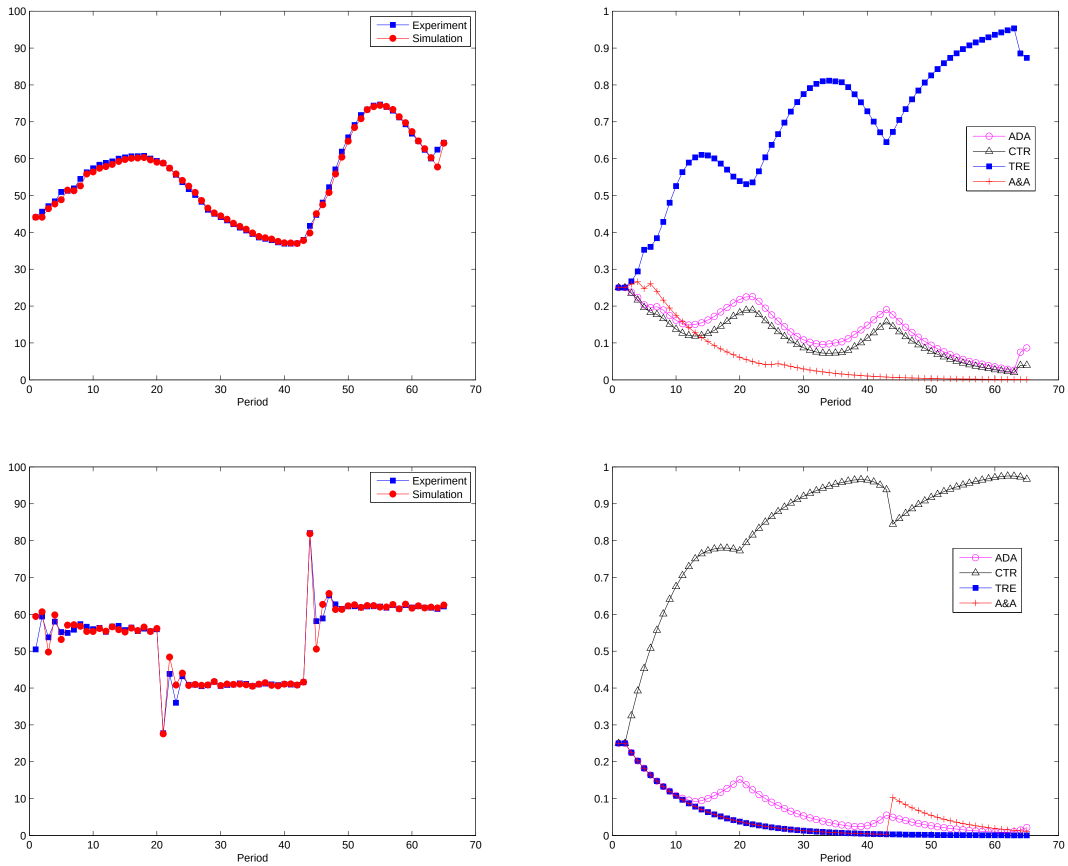


Figure 7: Experimental and simulated prices using HSM model in one typical group from the positive (top left, group $P8$) and negative feedback treatment (bottom left, group $N8$) respectively. Experimental data (blue squares) and one-step ahead simulated prices from the HSM model (red circles) almost overlap. The right panels show the evolution of the four market heuristics in the positive (top right) and negative feedback treatments (bottom right). The trend following rule dominates in the positive feedback markets, while the contrarian rule dominates in the negative feedback markets. Parameters are: $\beta = 0.4, \eta = 0.7, \delta = 0.9$, as in Anufriev and Hommes (2012).

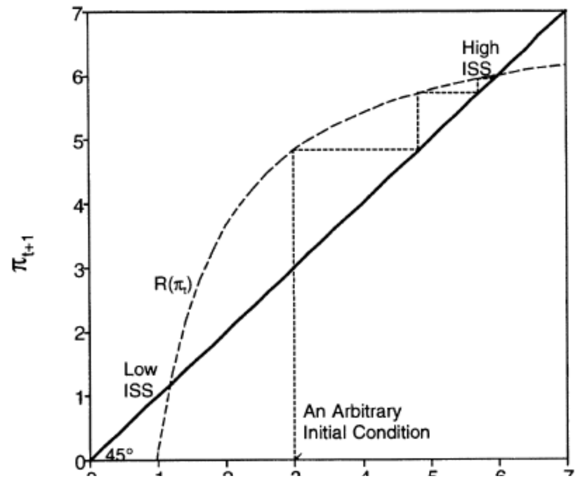
3.3 Overlapping generations economies

This section reviews the main contributions in experimental overlapping generation (OLG) economies. In a series of papers, Marimon and Sunder (1993, 1994, and 1995) and Marimon, Spear and Sunder (1993) pioneered the learning-to-optimize and learning-to-forecast design to study dynamic macroeconomic models in the laboratory, using the framework of OLG economies. [11]

Marimon and Sunder (1993) consider an OLG experimental economy in which participants are facing a monetary economy where the level of deficit is constant and financed by means of seigniorage. This is essentially a learning-to-optimize experiment, since subjects must submit supply schedules, but the subjects also take part in a forecasting contest to be able to end the experimental OLG economy in finitely many periods, without affecting the equilibria of the infinite OLG model. This OLG economy has two stationary steady states, a low inflationary stationary state (Low ISS) and a high inflationary stationary state (High ISS), as illustrated in Figure 8. Under RE the equilibrium path converges to the High ISS, while under adaptive learning the economy converges to the Low ISS. In the experiments coordination on the Low ISS occurs in all cases. These experiments thus strengthen the view that economic agents are more likely to follow adaptive learning based on observed data.

Marimon et al. (1993) design an experimental OLG economy with a RE period-2 cycle and sunspot equilibria. They use a learning-to-forecast design to study whether coordination on a 2-cycle or on sunspots can arise in the lab. As shown by Woodford (1990) sunspot equilibria are learnable and hence they cannot be ruled out a priori. Marimon et al. (1993) find that in their experimental economy the emergence of a 2-cycle does *not* arise spontaneously, but coordination on (approximate) 2-cycle equilibria may arise when they are correlated with an extrinsic sunspot signal, as illustrated in Figure 9. In the first 17 periods of this OLG economy, the generation

[11] Marimon et al. (1993) is also a pioneer work in experiments on sunspot equilibrium, followed by Duffy and Fisher (2005) and



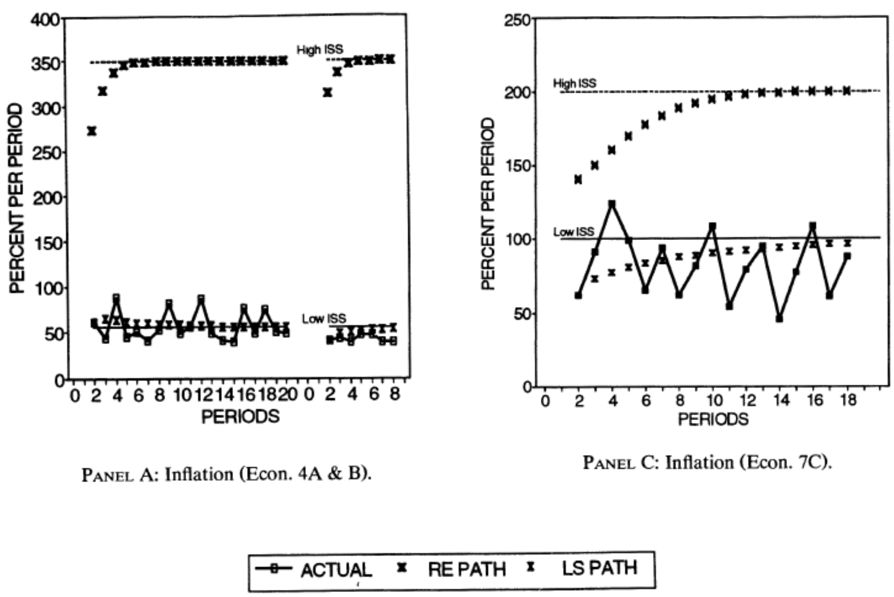


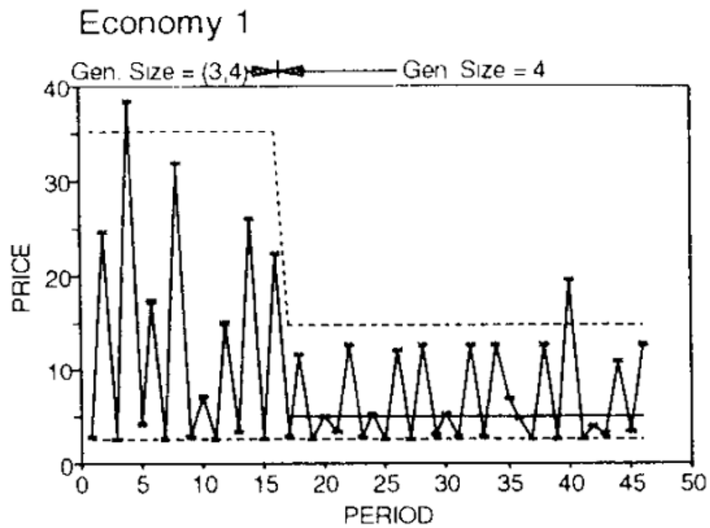
Figure 8: Experimental OLG economy in Marimon and Sunder (1993), with a low inflationary stationary state (Low ISS) and a high inflationary stationary state (High ISS). Under RE the time path converges to the High ISS, while adaptive learning converges to the Low ISS. Experimental data are consistent with adaptive learning. Figures 1 and Figure 3, panels A and C, from Marimon and Sunder, *Econometrica* 1993. Reprinted by permission of the Econometric Society.

size oscillates between 3 and 4, while after period 17 the generation size is held fixed at 4. As a result, the OLG economy has a large amplitude RE 2-cycle in the first 17 periods and a small amplitude RE 2-cycle thereafter. The extrinsic shocks in generation size in the first 17 periods facilitate coordination on an (approximate) 2-cycle. After the extrinsic shocks disappear after period 17, coordination on the RE 2-cycle remains. This OLG experimental economy thus shows the possibility of coordination on expectations driven price volatility, but only after subjects have been exposed to a sequence of extrinsic sunspot signals correlated with a real cycle.

Marimon and Sunder (1994) have studied the robustness of their earlier results and studied the effect of changes in policy specifications. They find persistence of expectations driven fluctuations in the experimental economy characterized by sunspots and the anticipation mechanism in the economy with repeated pre-announced policy regime shifts.

Marimon and Sunder (1995) investigate the Friedman prescription whether a (simple) constant money growth rule help stabilize prices (inflation rates). They find that the price volatility that is observed in the experimental data is due more to the use of adaptive learning rules than to different in monetary regimes. Inflation volatility is broadly equivalent for the two monetary rules.

Bernasconi and Kirchkamp (2000) build a very similar experiment and differently from Marimon and Sunder the authors find that: i) agents do not use first order adaptive rules; ii) agents show over-saving due to precautionary motivations; iii) the Friedman conjecture holds. Three main differences may be highlighted between the Marimon and Sunder (1995) and Bernasconi and Kirchkamp (2000). First of all subjects in Bernasconi and Kirchkamp forecast prices but they also decide savings, that is, they use a learning-to-forecast and a learning-to-optimize design; second agents do not hold by construction quasi-point forecast but they can test the results for different forecast decisions for various periods ahead before submitting their final decision. Finally, monetary policies are distinguished by labels and participants vote for them.



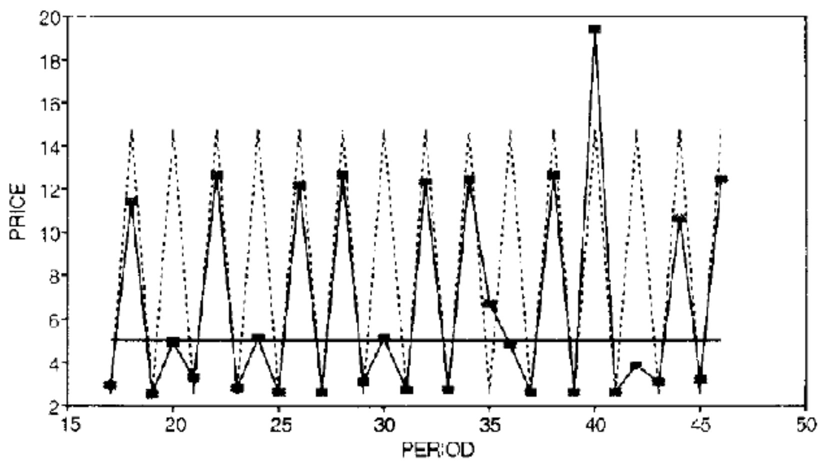


Figure 9: Experimental OLG economy in Marimon, Spear and Sunder (1993) with RE 2-cycle driven by extrinsic sunspot signal. In the first 17 periods the generation size oscillates between 3 and 4; after period 17 the generation size is fixed at 4. The OLG economy coordinates on the large amplitude RE 2-cycle correlated with the sunspot signal in the first 17 periods. After the extrinsic shocks disappear after period 17, coordination on the RE 2-cycle remains. Figure 3, Economy 1, and Figure 4 from Marimon, Spear and Sunder, *Journal of Economic Theory* 1993. Reprinted by permission of Elsevier.

Within this experimental setup Bernasconi and Kirchkamp are able to analyze the subjects' expectation formation process independently of their saving behavior. In contrast to Marimon and Sunder they find that the two implemented monetary rules are not equivalent but they show significant differences both in inflation levels and volatility, in particular they find support for Friedman's conjecture as the experiments show higher inflation volatility under a real deficit regime than under a simple money growth rule.

Finally, Hommes et al. (2012) conduct an individual learning-to-forecast experiment within a standard overlapping generation framework in which two monetary policy regimes are implemented, namely a low monetary growth rule and a high monetary growth rule. Subjects are asked to forecast inflation rates. This learning-to-forecast experiment has a more complicated structure

$$\pi_t = \theta \frac{S(\pi_t^e)}{S(\pi_{t+1}^e)}, \tag{19}$$

where S is the (non-monotonic) savings function and realized inflation π_t depends on the inflation forecasts for periods t and $t+1$ of a single subject. The authors find a wide variation in participants' forecasting ability both among participants to the same experimental session and among different treatments. The rational expectation hypothesis is not able to explain the experimental results. The authors find essentially three types of individual forecasting behavior: an accurate forecast leading to stabilizing the inflation rate; learning behavior with inflation stabilizing after a highly volatile initial phase; and finally a set of subjects never learn how to predict inflation rate with some accuracy. The Hommes et al.(2012) OLG experimental results are consistent with subjects using constant gain algorithms (e.g. adaptive expectations) as forecasting rules or average expectations when they learn to stabilize inflation. Moreover analyzing the experimental data they find evidence of agents switching among different forecasting rules on the basis of rules' forecasting performance. Hence, even if, agents do not use least squares learning, they try to improve their forecasting ability by learning eventually ending up close to the rational expectations steady state equilibrium.

3.4 New Keynesian DSGE

This section surveys learning to forecast experiments (LtFEs) framed in New Keynesian (NK) macro environments. In its basic formulation, the NK model consists of an IS curve derived from households' intertemporal optimization of consumption expenditures, representing the demand side of the economy, and a Phillips curve derived from firms' intertemporal optimization under monopolistic competition and nominal rigidities, representing the supply side of the economy. The IS curve and the Phillips curve are respectively specified as

$$y_t = E_t^* y_{t+1} - \varphi(i_t - E_t^* \pi_{t+1}) + g_t \tag{20}$$

$$\pi_t = \lambda y_t + \beta E_t^* \pi_{t+1} + u_t \ , \tag{21}$$

where y_t denotes the output gap, i_t the interest rate, π_t the inflation rate, while g_t and u_t are exogenous shocks. The terms $E_t^* y_{t+1}$ and $E_t^* \pi_{t+1}$ denote subjective (possibly non-rational) expected values of the future output gap and inflation respectively.[12] The model is closed by specifying a policy rule for the nominal interest rate.

Experimental implementations of the NK model have been targeted at shedding lights on two important issues, namely:

- the nature of the expectation formation process and its impact on aggregate dynamics in a more complicated framework, with inflation and output depending on expectations of *both* variables;
- the effectiveness of alternative monetary policies in stabilizing experimental economies in which the expectational terms in (20) – (21) are replaced by subjects' average (or median) forecasts.

[12] Detailed derivations of the NK model can be found in Woodford (2003) among others.

This section will discuss the LtFEs presented in Pfajfar and Zakelj (2011) (and companion paper Pfajfar and Zakelj (2014)), Assenza et al. (2012) (and the revised and extended paper Assenza et al. (2014b)) and Kryvtsov and Petersen (2013), which contributed on the issues mentioned above within experimental economies described by equations (20) – (21), and touch upon the LtFEs described in Adam (2007). [13]

The small scale NK model described by the aggregate demand equation (20) and the aggregate supply equation (21) is widely used for policy analysis and its popularity is based on its ability to replicate a number of stylized facts. However, the implementation of such model in an experimental setup is more complicated, because subjects have to submit two period ahead forecasts for two variables, inflation as well as the output gap. In order to simplify subjects' cognitive task, Pfajfar and Zakelj (2011) only ask for inflation expectations and assume $E_t^* y_{t+1} = y_{t-1}$.[14] This scenario corresponds to a situation in which subjects have naive expectations about the output gap, or to an extreme case of habit persistence. To deal with the same issue, Assenza et al. (2012) elicit forecasts of the endogenous variables from different groups of subjects, one group forecasting inflation and another group forecasting output gap. In more recent experiments Kryvtsov and Petersen (2013) ask subjects to forecast both inflation and the output gap.

Formation of individual expectations

The information set available to participants in the experiments of Assenza et al. (2012) and Pfajfar and Zakelj (2011) include realizations of inflation, the output gap and the interest rate up to period $t-1$. Subjects also have information about their past forecast, but they do not observe the forecasts of other individuals.

[13] Other contributions to LtFEs are Arifovic and Sargent (2003) and Cornand and M'baye (2013), which are discussed in detail in the survey on experiments on monetary policy and central banking by Cornand and Heinemann (2014).

[14] The analysis performed in Pfajfar and Zakelj (2014) uses the experimental data collected in Pfajfar and Zakelj (2011).

Overall, Assenza et al. (2012, 2014b) find that the predictions of the model with homogeneous rational expectations can hardly describe the experimental outcomes. The authors find evidence for heterogeneity in individual expectations and estimate simple forecasting *first-order heuristics* using the time series of individual predictions.[15] A stylized fact that emerges from the analysis of Assenza et al. (2012) is that individual learning takes the form of switching from one heuristic to another. The authors use the Heuristic Switching Model (HSM) developed by Anufriev and Hommes (2012) to explain both individual forecasting behavior and aggregate macro outcomes observed in the experiments. Using the same set of heuristics of Anufriev and Hommes (2012) (described in detail in Section 3.1), Assenza et al. (2012, 2014b) show that the HSM explains how the different macro patterns observed in the experiment, i.e., convergence to the target equilibrium level, inflationary/deflationary spirals, persistent oscillations and dampened converging oscillations, emerge out of a self-organization process of heterogeneous expectations driven by their relative past performance. Convergence to equilibrium is explained by coordination on adaptive expectations, inflationary/deflationary spirals arise due to coordination on strongly extrapolating trend-following rules, persistent oscillations arise after coordination on an anchor and adjustment rule and dampened converging oscillations arise when initially dominating (weak) trend-following rules are finally driven out by adaptive expectations.

Pfajfar and Zakelj (2014) focus on the analysis of individual data on inflation expectations collected in Pfajfar and Zakelj (2011).[16] The authors fit 12 alternative models of expectation formation to individual prediction series and find significant

[15] In total 216 subjects participated in the experiment of Assenza et al. (2012) divided in 18 experimental economies with 12 subjects each (6 subjects forecasting inflation and 6 subjects forecasting the output gap). Each participant submitted forecasts for 50 consecutive periods.

[16] In total 216 subjects participated in the experiment of Pfajfar and Zakelj (2011) divided in 24 independent groups of 9 subjects each. Each participant submitted inflation forecasts for 70 consecutive periods.

heterogeneity in subjects' forecasting strategies. The paper develops a new test for rational expectations that checks the consistency of expectation formation rules with the actual laws of motion, explicitly allowing for the possibility of heterogeneous expectations. In other words, the test allows for the possibility that the perceived law of motion (PLM) of a rational agent may differ from that implied by the assumption of homogeneous rational expectations, and include additional state variables as a result of the presence of heterogeneous forecasters. Using this test, Pfajfar and Zakelj (2014) find that for 30 – 45% of subjects it is not possible to reject rationality. Moreover, 20 – 25% of subjects' forecasting strategies are described by adaptive learning algorithms. The authors also find evidence for simple heuristics. Roughly 25 – 35% of subjects can be described by trend extrapolation rules and an additional 10 – 15% by adaptive expectations or by a sticky information type of model. Finally, Pfajfar and Zakelj (2014) find evidence for switching between forecasting models. The authors study "unrestricted" switching, i.e., they re-estimate all alternative models in each period and for each individual they select the best performing model in each period, finding that switching between alternative models better describe subjects' behavior.

The experimental setup of Adam (2007) is closely related to the NK framework described in this section. The author implements a sticky price environment where inflation and output depend on expected inflation. As in Pfajfar and Zakelj (2011) and Assenza et al. (2012), the information set available to subjects include past realizations of endogenous variables through period $t-1$ and, in each experimental economy, a group of 5 subjects is asked to provide one and two step ahead forecasts of inflation for 45 – 55 periods. The results show cyclical patterns of inflation around its steady state. Although the rational forecast for inflation should condition on lagged output, Adam finds that in most of the experimental sessions, the forecast of the "representative subject", i.e., the average forecasts entered by subjects in any given period, uses a simple AR(1) model. He shows that such behavior can result in a *restricted perception equilibrium* (RPE) in which the autoregressive inflation model outperforms the rational forecast model. Adam further notes that mis-specified forecasting rules provide a source of inflation and output persistence, explaining therefore the observed persistence of inflation cycles.

Kryvtsov and Petersen (2013) focus on measuring the strength of the expectation channel for macroeconomic stabilization. The main difference between the experimental setup developed by Kryvtsov and Petersen and the ones described above in this section consists in the information set available to subjects in the experiment. In fact, Kryvtsov and Petersen provide subjects with full information about the only exogenous shock process, i.e., g_t in the IS equation,[17] and about the model underlying the experimental economy. Moreover, information about histories of past outcomes and shocks, as well as detailed model description, is available at a small time cost. This setup allows estimating forecasts as function of the observed shock history, i.e., g_t in the IS equation, which is then used to quantify the contribution of expectations to macroeconomic stabilization via counterfactual analysis. Krystov and Petersen show that a model with a weak form of adaptive expectations, attributing a significant weight on $t-1$ realizations of inflation and the output gap, fits best both the magnitude and the timing of aggregate fluctuations observed in the experiment.

The effectiveness of monetary policies

An important contribution of LtFEs casted in the NK framework is to analyze the effectiveness of alternative monetary policy rules in stabilizing the variability of inflation in a setting where expectations about endogenous variables are potentially non-rational and heterogeneous across subjects.

Pfajfar and Zakelj (2011) close the NK model described by equations (20) – (21) with a *forward-looking* interest rule of the form

$$i_t = \phi_\pi (E_t^* \pi_{t+1} - \bar{\pi}) + \bar{\pi} \ ,$$

[17]In Kryvtsov and Petersen (2013) the cost-push shock $u_t = 0$ in every period t.

where the monetary authority responds to deviations in subjects' inflation expectations from the target $\bar{\pi}$, set at 3%. The authors vary the value of ϕ_π, measuring the strength of policy reaction, in different treatments. Pfajfar and Zakelj (2011) also consider an alternative *contemporaneous* policy rule of the form

$$i_t = \phi_\pi(\pi_t - \bar{\pi}) + \bar{\pi} \ ,$$

where the central bank responds to deviations of current inflation from the target. The different treatments implemented in Pfajfar and Zakelj (2011) are summarized in Table 1. Figure 10 displays the experimental outcome of Pfajfar and Zakelj (2011).

Treatment	Parameter
1 - Forward-looking rule	$\phi_\pi = 1.5$
2 - Forward-looking rule	$\phi_\pi = 1.35$
3 - Forward-looking rule	$\phi_\pi = 4$
4 - Contemporaneous rule	$\phi_\pi = 1.5$

Table 1: Treatments in Pfajfar and Zakelj (2011)

A cyclical behavior of inflation and the output gap around their steady states can be observed in all treatments of Pfajfar and Zakelj (2011). Among the forward-looking policy rules, a reaction coefficient $\phi_\pi = 4$ (Treatment 3) results in lower inflation variability compared to reaction coefficients $\phi_\pi = 1.35$ (Treatment 2) and $\phi_\pi = 1.5$ (Treatment 1). The authors report that there is no statistical difference between Treatments 1 and 2. When comparing the results in Treatments 4 and 1, Pfajfar and Zakelj find that the inflation variance under the contemporaneous rule is significantly lower than under the forward-looking rule with the same reaction coefficient $\phi_\pi = 1.5$. The intuition provided by the authors for this result is that the variability of the interest rate is generally lower under the contemporaneous rule.

Assenza et al. (2014b) implement different versions of the contemporaneous in-

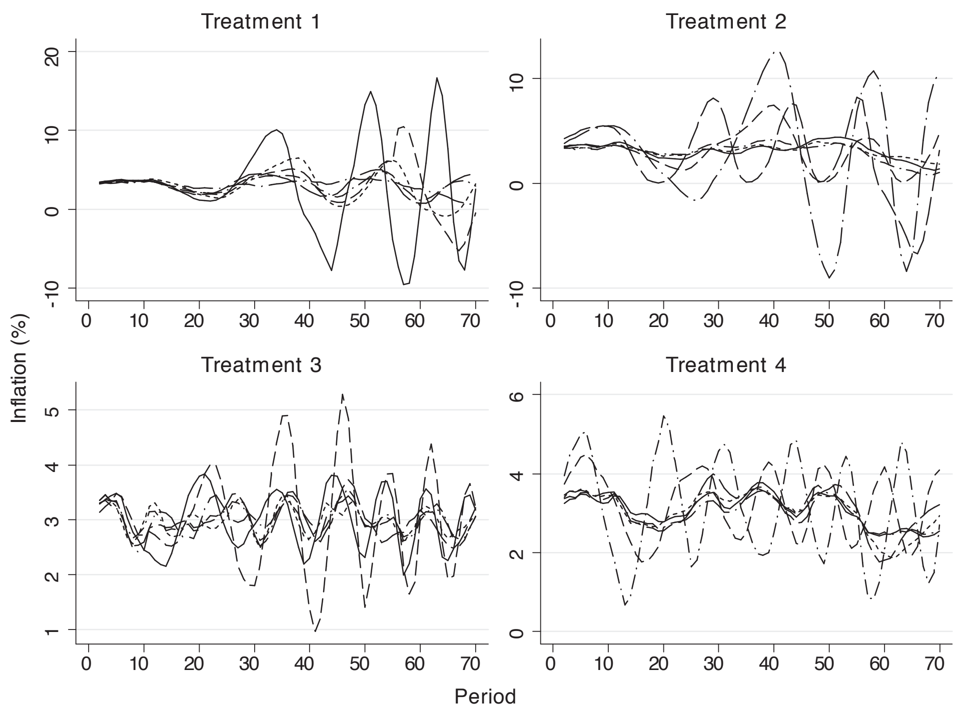


Figure 10: Realized inflation by treatments in Pfajfar and Zakelj (2011). Each line represents one of the 24 experimental economies.

terest rate rule which takes, as in Pfajfar and Zakelj (2011), the following form

$$i_t = \phi_\pi(\pi_t - \bar{\pi}) + \bar{\pi} \ .$$

In particular the authors set an inflation target $\bar{\pi} = 2\%$ and consider a case in which $\phi_\pi = 1$, so that the Taylor principle does not hold and thus policy does not play a stabilizing role, and compare it with the case where the Taylor principle does hold, setting $\phi_\pi = 1.5$. In this particular setting, the Taylor principle corresponds to setting $\phi_\pi > 1$. Moreover, since $\bar{\pi} = 2\%$ could be a focal point for subjects' forecasts, the authors run an additional treatment with $\bar{\pi} = 3.5\%$ to check the robustness of the policy rule obeying the Taylor principle to alternative target values. The different treatments implemented in Assenza et al. (2014b) are summarized in Table 2. Figure 11 illustrates the experimental results of Assenza et al. (2014b). The evidence

Treatment	Policy	Target
a - Contemporaneous rule	$\phi_\pi = 1$	$\bar{\pi} = 2\%$
b - Contemporaneous rule	$\phi_\pi = 1.5$	$\bar{\pi} = 2\%$
c - Contemporaneous rule	$\phi_\pi = 1.5$	$\bar{\pi} = 3.5\%$

Table 2: Treatments in Assenza et al. (2014b)

presented in Assenza et al. (2014b) suggests that a monetary policy that reacts aggressively to deviations of inflation from the target (treatments b and c) stabilizes macroeconomic fluctuations and leads the economy to the desired target. The specific value of the target seems to have little influence on the stabilizing properties of the policy rule. On the other hand, when the interest rate reacts weakly to inflation fluctuations, Assenza et al. (2014b) observe convergence to non-fundamental equilibria (treatment a, Groups $1-3$) or exploding behavior (treatment a, Groups $4-6$).

Overall, the results of Assenza et al. (2012, 2014b) are in line with those of Pfajfar and Zakelj (2011). Treatment 4 in Pfajfar and Zakelj (2011) uses the same contemporaneous policy rule adopted in Assenza et al. (2012, 2014b) with reaction

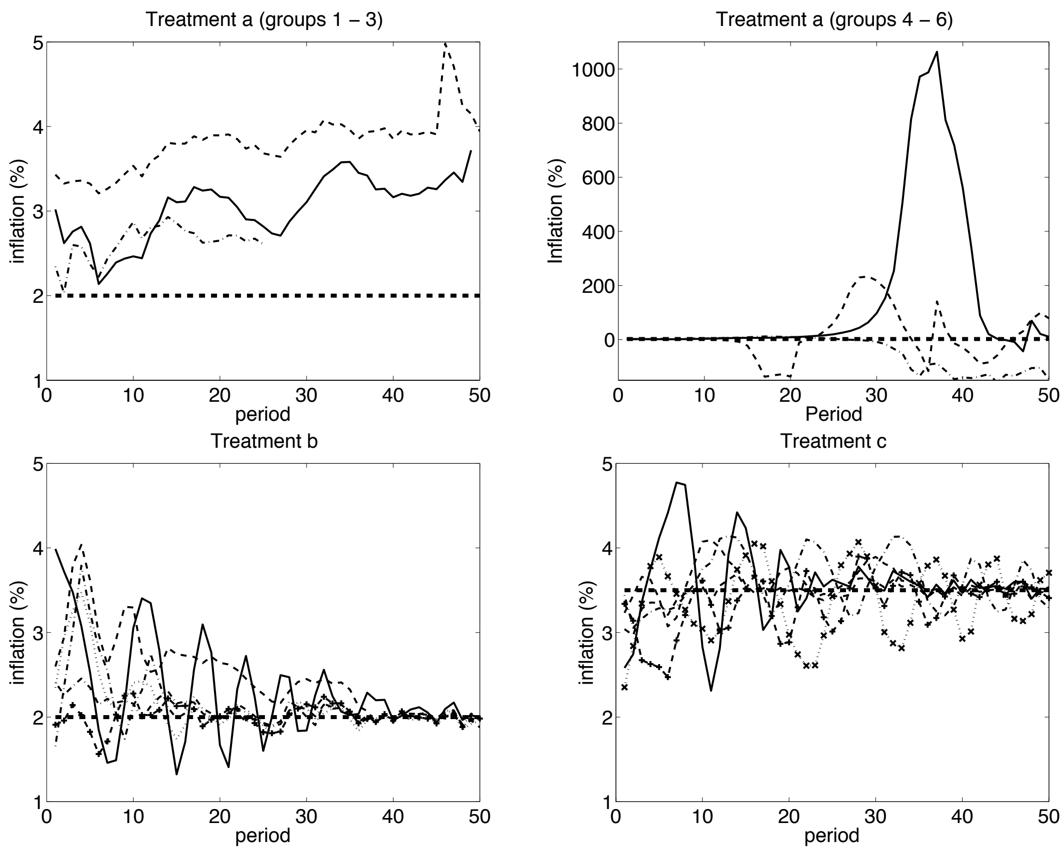


Figure 11: Realized inflation by treatments in Assenza et al. (2014b). Each (thin) line represents one of the 18 experimental economies. Dashed thick lines depict the inflation targets.

coefficient $\phi_\pi = 1.5$. A qualitative comparison of the outcomes of the two experiments shows sustained inflation oscillations around the steady state in Pfajfar and Zakelj (2011), while in Assenza et al. (2012) inflation seems to converge to the target value, at least in the late stages of the experiment. The different behavior might be due to the differences in the two experimental setups. While Pfajfar and Zakelj (2011) only elicit inflation expectations, assuming that expectations of the future output gap are given by lagged output, i.e., $E_t^* y_{t+1} = y_{t-1}$, Assenza et al. (2012) elicit forecasts of both future inflation and the future output gap in accordance with the NK model. Moreover, Pfajfar and Zakelj (2011) assume AR(1) processes for the exogenous shocks g_t and u_t, while Assenza et al. (2012) use IID shocks with the consequence that the rational expectation fundamental solution is an IID process and any observed fluctuations in aggregate variables are endogenously driven by individual expectations. The experimental results in Assenza et al. (2012) suggest therefore that a policy rule reacting more than point to point to deviations of inflation from the target can stabilize endogenous expectations-driven fluctuations in the aggregate variables.

Both Pfajfar and Zakelj (2011) and Assenza et al. (2014b) further investigate the relationships between expectations, monetary policy and macroeconomic stability. Using panel data regressions, Pfajfar and Zakelj (2011) show that a higher proportion of agents using trend extrapolating rules increase the volatility of inflation. In contrast, having more agents that behave according to the adaptive expectations models has a stabilizing effect on the experimental economies. Moreover, the regression results show that monetary policy also has an impact on the composition of different forecasting rules in each treatment. Pfajfar and Zakelj find that the percentage of destabilizing trend extrapolation rules and the variability of inflation are lowest in Treatment 3, where the strength of the positive expectational feedback is the lowest. Interestingly, the explanation of how different macro patterns emerge out of a process of self-organization of heterogeneous expectations provided by the HSM in

Assenza et al. (2014b) delivers comparable insights. Assenza et al. (2014b) find that macroeconomic instability arise due to coordination on strongly extrapolation trend following rules while convergence to equilibrium is associated with coordination on adaptive expectations. Moreover, Assenza et al. (2014b) show that an aggressive policy rule can avoid almost self-fulfilling coordination on destabilizing trend-following expectations by reducing the degree of positive feedbacks in the system.

Kryvtsov and Petersen (2013) assume a policy rule of the form

$$i_t = \phi_\pi E^*_{t-1}\pi_t + \phi_y E^*_{t-1} y_t$$

where the reaction coefficients assume values $\phi_\pi = 1.5$ and $\phi_y = 0.5$ in the Benchmark treatment and $\phi_\pi = 3$, $\phi_y = 1$ in the Aggressive Monetary Policy treatment. They find that inflation and output gap predominantly exhibit stable cyclical behavior, with inflation and the output gap displaying less volatility and less persistence in the Aggressive Monetary Policy treatment.[18] The experiment of Kryvtsov and Petersen is designed to identify the contribution of expectations to macroeconomic stability achieved by systematic monetary policy. The authors find that, despite some non-rational component in individual expectations, monetary policy is quite powerful in stabilizing the experimental economies confirming thus the results of Pfajfar and Zakelj (2011) and Assenza et al. (2012), and they report that monetary policy accounts for roughly a half of business cycle stabilization.

4 Learning to optimize

"Learning to optimize experiment"(LtOE) refers to the experiments where subjects submit their economic decisions (i.e. consumption, trading, production) directly, without elicitation of their forecasts of the market price. It is a literature with a longer history than learning to forecast experiments. Some examples of this approach include

[18] Kryvtsov and Petersen (2013) elicit expectations on both inflation and the output gap and they assume an AR(1) process for the exogenous driving process.

Smith et al.(1988), Lim et al.(1994), Arifovic (1996), Noussair et al. (2007) and Crockett and Duffy(2013). There have been many surveys on studies in this approach already (Noussair and Tucker, 2013). In this section, we limit our attention to a few experiments with parallel learning to forecast and learning to optimize treatments (based on the same model of the experimental market) in order to compare them[19]. These experiments are helpful in answering the robustness question: "will the results of the learning to forecast experiments change if the subjects make a quantity decision directly (instead of making a forecast only)?"

There are two potential sources that may lead to different results due to LtFE vs. LtOE design: the nature of the task and the payoff structure. In terms of the nature of the task: (1) the subjects in LtFEs are aided by computer to make calculations, which should facilitate the learning of the REE; (2) on the other hand, since the price determination equation is in the end a function of quantity decisions, it should be easier for the subjects in the LtOE design to understand how their decisions are translated to market price, which helps the subjects to learn to play rationally. In terms of payoff structure, subjects are typically paid accordingly to forecasting accuracy in the LtFE design, and profitability of the quantity decision in the LtOE design. We speculate that the market price should be closer to the REE when the subjects are paid according to the forecasting accuracy than profit. The reason is two folded: (1) when the subjects are paid according to forecasting accuracy, predicting the REE is the unique symmetric Nash equilibrium of the "prediction game" (the payoff of every subject is maximized when they predict the REE). When they are paid according to the profit, they may earn higher payoff if they deviate from the REE. For example, in a finite player cobweb market, the subjects can earn higher payoff if they play the collusive equilibrium instead of the REE (competitive equilibrium). (2) Some studies show that emotion can influence the optimality of individual decisions and

[19] See also Roos and Luhan (2013), for a recent LtOE and LtFE in an experimental macroeconomy with monopolistic firms and labor unions.

price stability in asset market experiments (Breaban and Noussair, 2013). Moreover, the emotion of subjects can be heavily driven by past gains and losses of their trading behavior (quantity decisions). On the other hand, it seems the accuracy of predictions should have less influence on subjects' emotional status. Therefore, payoff based on prediction accuracy should generate less fluctuation in emotions, and more stable market prices.

According to the experimental results, the answer to this question is:

(1) Holding other things equal, there is indeed a difference between aggregate price behavior in the learning to forecast and the learning to optimize markets.

(2) The markets in the learning to optimize treatment deviate more from the rational expectation equilibrium than the markets in the learning to forecast treatments. More specifically, in negative feedback markets, it takes longer for the market price to converge to the REE in the learning to optimize treatment than in the learning to forecast treatment. In positive feedback markets, there are larger bubble-crash patterns in the asset price in the learning to optimize treatment than the learning to forecast treatment.

To conclude, the aggregate market outcome is closer to the rational expectations equilibrium in the LtFE design than the LtOE design. This suggests researchers interested in testing dynamic macroeconomic model in the rational expectations benchmark case should probably use the LtFE design. Meanwhile, the larger deviation in the LtOE design is probably a result of agents' failure to solve the optimization problem after they form their expectations. Since there is already a large literature on bounded rationality in expectation formation, modeling bounded rationality in solving the optimization problem given one's own expectations may be a good direction for future research.

4.1 Cobweb Market

This section discusses the experiment by Bao et al. (2013). The model behind this experiment is a traditional "cobweb" economy as studied by Muth (1961) when he proposed the famous rational expectation hypothesis. Before this experiment, there have been pure learning to forecast experiments on cobweb markets (Hommes et al., 2000, 2007). But those experiments use non-linear models. The model used in this experiment is similar to the one used in the negative feedback treatment in Heemeijer et al. (2009). Heemeijer et al. (2009) found that the price converge quickly to the REE in negative feedback markets like the cobweb economy. One of the main target of Bao et al. (2013) is to investigate whether this result still holds if the subjects submit a production quantity instead of a price forecast. The model is about a non-storable commodity. Let p_t be the price of the good at period t. D is the linear demand function for the good that is decreasing in p_t, $D(p_t) = a - bp_t$, where $a = 63, b = \frac{21}{20}$. The subjects play the role of the advisors of firms that produce the good. The supply of firm h in period t is denoted by $S_{h,t}$. Let $p^e_{h,t}$ be the price forecast made by firm h in period t. The supply function may be rewritten as $S(p^e_{h,t})$. It should be the solution of the expected profit maximization problem:

$$\max \pi^e_{h,t} = \max[p^e_{h,t} q_{h,t} - c(q_{h,t})]. \tag{22}$$

Each firm has a quadratic cost function $c(q) = \frac{Hq^2}{2}$, where H is the number of firms in the market. Taking the first order condition of the expected profit, it is not difficult to find

$$S^*(p^e_{h,t}) = \frac{p^e_{h,t}}{H}. \tag{23}$$

The total supply of the good equals the sum of supplies of individual firms. If every firm makes supply based on equation (23), the total supply on the market will coincide with the average price forecast, $\sum_h S^*(p^e_{h,t}) = \overline{p^e_t}$.

The market price of the good is determined by the market clearing condition (supply equals demand):

$$p_t = D^{-1}(\sum_h S_{h,t}) + \epsilon_t, \tag{24}$$

Plugging in the parameters, the price determination equation becomes:

$$p_t = \max\left\{\frac{20}{21}(63 - \overline{p_t^e}) + \epsilon_t,\ 0\right\}, \tag{25}$$

where $\epsilon_t \sim N(0,1)$. Imposing the rational expectations assumption $p_t^e = Ep_t = E(\max\left\{\frac{20}{21}(63 - \overline{p_t^e}) + \epsilon_t,\ 0\right\})$, and noting that the expected value of ϵ_t is zero, the rational expectations equilibrium (REE) price of this economy is $p_t^e = p^* = 30.73, p_t = 30.73 + \epsilon_t$. The optimal supply in the REE is 5.12.

Five treatments were designed in the experiment, and we focus on the first three of them.

1. Treatment 1: the LtFE treatment. Subjects (firms) only make a price forecast $p_{h,t}^e$ in each period t. Their implicit quantity decision, $S(p_{h,t}^e)$ will be calculated based on equation (23) by the experimental computer program. They are paid according to the prediction error, namely $|p_t - p_{h,t}^e|$. The larger the prediction error, the smaller the payoff.

2. Treatment 2: the LtOE treatment. Subjects (firms) make quantity decision $S_{h,t}$ directly, and there is no assistance from the computer. Each subject is paid according to the profit his firm makes each period as defined by equation (22), namely, revenue minus cost.[20]

3. Treatment 3: the LtFE+LtOE treatment. Each subject makes both a price forecast $p_{h,t}^e$ and a quantity decision $S_{h,t}$ in each period. The market price is again

[20]Note that the difference between Treatment 1 and 2 can be a result of the joint force of difference in tasks and in payoff structures. In order to better isolate the effect of each of them, there was another treatment (Treatment 5) in Bao et al. (2013) where the subjects make forecasts but are paid according to profits. The result of that treatment is just between Treatment 1 and 2. It seems that the task nature and payoff structure play equally important roles in causing the treatment difference.

determined by the production decisions submitted by the firms as in Treatment 2. Subjects are paid according to an equal weighted linear combination of the payoff functions used in the LtFE and LtOE treatments.

If agents are able to form rational expectations, the experimental results should be exactly the same in all of the three treatments. From the results in former learning to forecast experiments on cobweb markets, we know this kind of markets have a tendency to convergence to the REE reliably. In a way, one can argue that the convergence to the REE should take fewer periods in Treatments 1 than 2, because subjects are helped with a computer program to calculate the conditionally optimal quantity, but not in treatment 2. But one could also argue that because people make quantity decisions for their economic activities on a daily basis, but less experience with making a forecast, therefore, the quantity decision task should be more familiar, and easier for them. In a way, subjects are faced with a situation that resembles the situation in a theoretical RE model in Treatment 3, where they first make a forecast, and then make a quantity decision. Therefore, they should learn faster from theoretical point of view, because doing the two tasks at the same time should make them think more about how the economy works.

It turns out that in terms of the number of periods it takes for the market price to converge to the REE, the convergence is fastest in Treatment 1, and slowest in Treatment 3. Figure 12 plots the average market price in the three treatments. It can be seen that the market price is most stable in Treatment 1, and most unstable in Treatment 3. The market price deviates to the largest extent from the REE in treatment 3, and it also takes longer for even the average market price to get close to the REE. The authors declare convergence to have occurred in the first period for which the difference between the market price and the REE price is less than 5 and stays below 5 forever after that period. If a market fails to converge, the number of periods before convergence is counted as 50. The result shows that the median number of periods before convergence is only 3 in Treatment 1, 13 in Treatment 2, but 50 in Treatment 3. Related to the arguments in the paragraph above, the finding suggests that the learning to forecast experiment indeed provide the highest speed of convergence due to that fact the subjects are helped by computers. The subjects seem to be cognitively overloaded by the complexity of the decision problem, which leads to lower speed to find the REE. The authors followed Rubinstein (2007) and used decision time as a proxy to measure cognitive load. They found that subjects in Treatment 3 indeed took significantly longer time to make each decision.

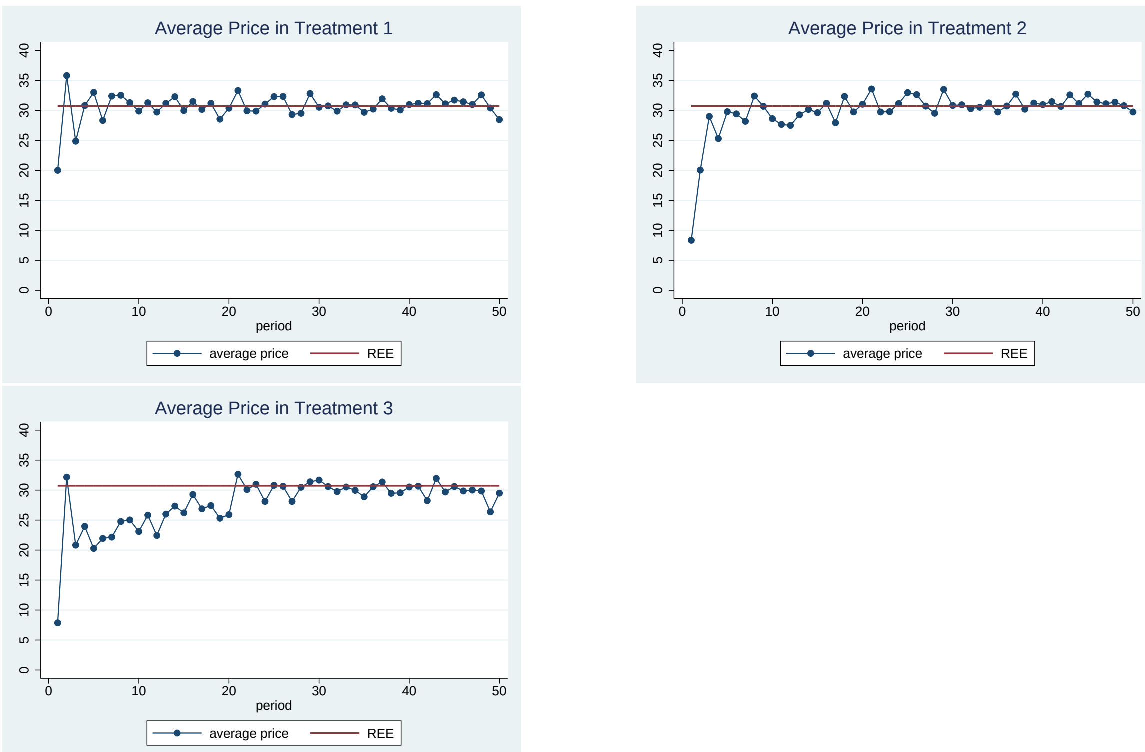


Figure 12: The average market price and the REE price in each of the three treatments of the LtFE and LtOE in Bao et al (2013).

4.2 Asset Pricing Market

Similar to Bao et al. (2013), Bao et al. (2014) set up an experiment with comparable treatments of learning to forecast (Treatment 1), learning to optimize (Treatment

2) and forecast+optimize (Treatment 3) for an experimental asset market similar to Hommes et al. (2005, 2008), and the positive feedback treatment of Heemeijer et al. (2009)[21] The purpose is again to see whether the price deviation in positive feedback markets in the previous learning to forecast experiments is a robust finding under the learning to optimize or forecasting+optimizing design.

In the LtFE treatment subjects submit a price forecast $p^e_{i,t+1}$ and are paid according to forecasting accuracy. This treatment is a replication of the positive feedback treatment of Heemeijer et al. (2009). The pricing rule in the LtFE is given by

$$p_{t+1} = 66 + \frac{20}{21}\left(\bar{p}^e_{t+1} - 66\right) + \varepsilon_t, \tag{26}$$

where 66 is the fundamental price ($\overline{y} = 3.3$ and $r = 0.05$) $\bar{p}^e_t = \frac{1}{6}\sum_{i=1}^{6} p^e_{i,t+1}$ is the average prediction of price p_{t+1} and ε_t a small IID noise term[22] In the second, LtOE treatment subjects submit the amount of asset they want to buy/sell, $z_{i,t}$ directly, and are paid according to the trading profit. The price adjustment thus takes the form of

$$p_{t+1} = p_t + \frac{20}{21}\sum_{i=1}^{6} z_{i,t} + \varepsilon_t, \tag{27}$$

where $z_{i,t}$ is the demand of subject i (a quantity choice between -5 and $+5$). In Treatment 3 subjects submit both a forecast and a trading quantity, and are paid randomly with equal chance according to their forecast accuracy or trading profit.

[21]Bostian and Holt (2013) developed web-based experimental software for LtO classroom experiments for asset bubbles in an environment with a constant fundamental value.

[22]Notice that this is a one-period ahead LtFE, in contrast to the two-period ahead asset pricing experiments in Hommes et al. (2005, 2008). The reason is that for one-period ahead the payoff table in the corresponding LtOE is two-dimensional, depending upon quantity and realized return. For a two-period ahead LtFE, the corresponding payoff table for the LtOE would depend on three variables due to an extra time lag. Therefore, the LtFE treatment in Bao et al. (2014) is not directly comparable to Hommes et al. (2005, 2008), but more comparable to the positive feedback treatment in Heemeijer et al. (2009), which is more stable than Hommes et al. (2005, 2008). Instead of very large bubbles and crashes, the asset price in many markets in Heemeijer et al. (2009) shows a mild upward trend, which typically overshoots the REE.

Since the learning to forecast experiments by find many bubble-crash patterns in the market price, a natural question to ask is whether these patterns are still there in the learning to optimize, or forecast+optimize design. The results show that the bubble-crash pattern is not only robust, but even stronger in the learning to optimize, and forecast+optimize treatment.

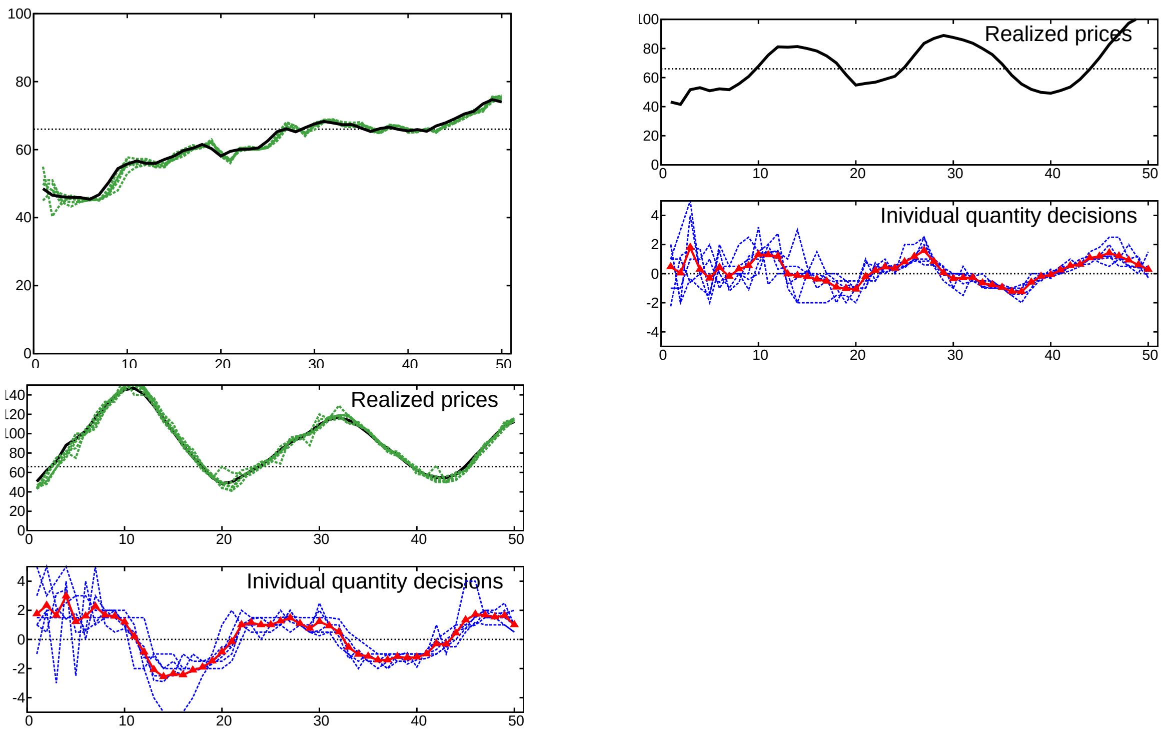


Figure 13: The price (squares) and individual expectations if applicable (lines) in a typical market (Market 1) against the REE price (REE=66, dashed line) in each of the three treatments in Bao et al (2014).

Figure 13 shows the market price in a typical market in each of the three treatments. The market price is most stable in Treatment 1, that steadily and slowly goes up. There is some mild oscillation in the market price in Treatment 2. Treatment 3 is the only treatment where the market price can go above 100 (in 3 out of 6 markets). In the market that generates the largest bubble, the price reaches 215 at the peak, which is more than 3 times the fundamental price (REE) of the asset! The deviation of the asset price from the REE is 10.8% in Treatment 1, 23% in Treatment 2 and 36% in Treatment 3 in terms of Relative Absolute Deviation (RAD) defined by Stöckl et al. (2010).

The results of this study confirms that the price deviation from the REE in the learning to forecast experiments with positive feedback, as in Hommes et al (2005, 2008) and Heemeijer et al. (2009), are robust against the experimental design. The learning to forecast design provides the result that is closest to the rational expectations equilibrium in the laboratory.

Besides this experiment, there are also learning to optimize experiments with elicitation of subjects' price forecasts. For example, Haruvy et al. (2007) study the asset market with double auction trading mechanism as in Smith et al. (1988). The subjects trade an asset that lives for 15 periods, and the fundamental value of the asset is determined by the sum of the remaining (expected) dividend of the asset at each point of time. The typical finding with this kind of asset market is that the subjects fail to trade according to the fundamental value of the asset. There is bubble-crash pattern where the market price first goes higher than the fundamental value, and then crashes till the end of the experiment, when the fundamental price goes to 0. Haruvy et al. (2007) ask the subjects to provide their price prediction for every future period at the beginning of each period (namely, in period 1, to predict prices in each of period 1-15, in period 2, to predict for each of period 2-15, and so on). They find that the subjects did not predict the fundamental value of the asset. When the subjects play in the market for the first time, they tend to predict that the past trend in the price will continue, and when they get experienced, they are able to predict the downturn of the price in the end, but still overestimate the number of periods before the downturn happens. Peterson (1993) elicits one-period-ahead forecast in a similar experimental market. He finds that the subjects also fail to form rational expectations, but there is evidence of learning over the periods. More recently, Akiyama et al. (2013) set up a forecast only (FO) treatment where both experienced and inexperienced subjects are invited to make one-period ahead forecast for a similar market populated with other subjects. They find that the initial price prediction by the experienced subjects is significantly closer to the fundamental price than the inexperienced subjects, but the difference becomes smaller in later periods.

4.3 Price-Quantity setting under monopolistic competition

Assenza et al. (2014b) present results from 50-rounds learning-to-optimize experimental markets in which firms decide repeatedly both on price and quantity of a perishable good. The experiment is designed to study price-quantity setting behavior of subjects acting as firms in monopolistic competition. In particular, each firm i in a $N = 10$ firms market faces in every period a demand curve of the form

$$q_i = \alpha - \beta p_i + \theta \bar{p} \ ,$$

where q_i represent demand for the good produced by firm i, p_i is the price set by firm i and $\bar{p} = \frac{1}{N}\sum_{i=1}^{N} p_i$ is the average market price. All firms face common constant marginal costs c.[23] Subjects are endowed with qualitative information about the market structure but they do not know the functional form of the demand for their product. Assenza et al. (2014b) are interested in understanding whether subjects in the experiment converge to the monopolistically competitive (MC) outcome without knowledge of the demand function and production set in advance. Moreover, they analyze the price-quantity setting strategies used by subjects in response to signals from the firms' internal conditions, i.e., individual profits, excess demand/supply, and the market environment, i.e., aggregate price level. Assenza et al. implement two treatments, differing in the information sets available to subjects. In Treatment 1 subjects observe the average market price, their own price, production, sales, profits and *excess supply* up to period $t-1$. In Treatment 2 subjects have the same informational

[23] Davis and Korenok (2011) use a similar experimental monopolistically competitive market setup to investigate the capacity of price and information frictions to explain real responses to nominal price shocks.

structure, but in addition firms can also observe *excess demand*, i.e., the portion of demand they were not able to satisfy given their price-quantity decisions and the average market price. Comparison between Treatment 1 and 2 allows to assess the impact of alternative information sets and, ultimately, different market structures. Finally, given that expected market price represents an important variable in firms' decisions on how much to produce and at which price to sell, the authors also elicit expectations of the average market price.

Overall, Assenza et al. report convergence of average prices and quantities to (a neighborhood of) the MC equilibrium in both treatments. Figure 14 reports the median (over 4 markets per treatment) of the absolute difference between the MC equilibrium and the realized prices and quantities.

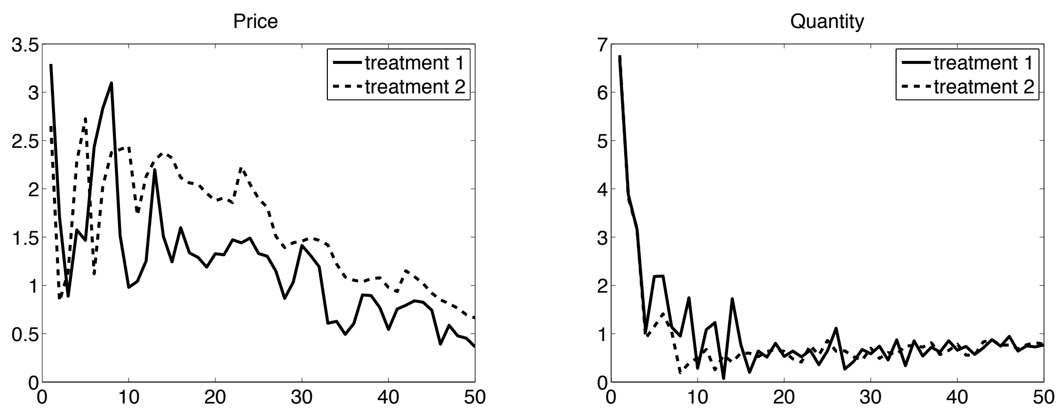


Figure 14: **Left panel:** Median of the absolute difference between average price and MC equilibrium price. **Right panel:** Median of the absolute difference between the average quantity and the MC equilibrium quantity.

In the case of prices, the authors report a significant difference between treatments, with an observed higher degree of convergence in Treatment 1. Quantities show no significant difference in the degree of convergence between treatments. Although average prices and quantities show a tendency towards equilibrium, Assenza et al. (2014b) find substantial heterogeneity among individual price and quantity decisions.

Figure 15 shows the median of the standard deviations of individual decisions for each period over the four markets of each treatment. A low standard deviation implies a high level of coordination among the subjects.

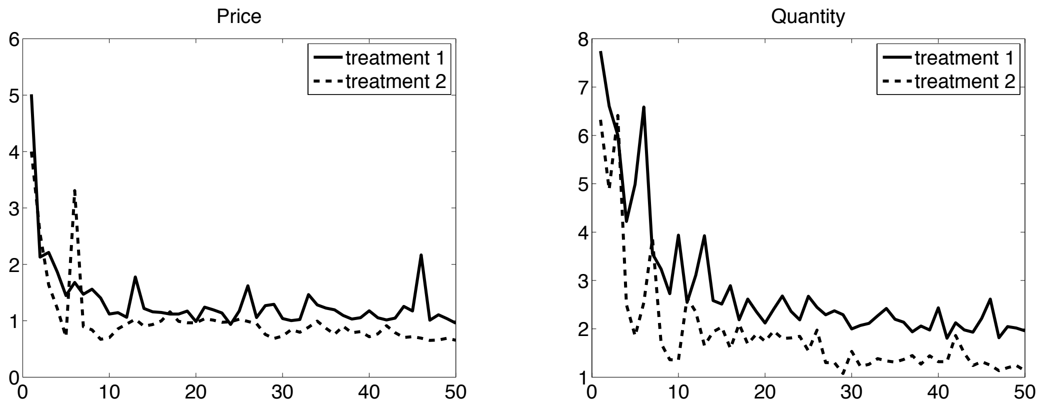


Figure 15: **Left panel:** Median of the standard deviations of individual prices. **Right panel:** Median of the standard deviations of individual quantities.

For both price and quantity, the authors report a statistically significant higher degree of coordination among individual decisions in Treatment 2 than in Treatment 1. In order to gain further insights on aggregate market behavior and explain individual price and quantity setting decisions, the authors estimate the following behavioral model

$$\begin{aligned}
\bar{p}^e_{i,t} &= c + \alpha_1 \bar{p}_{t-1} + \alpha_2 \bar{p}^e_{i,t-1} + \alpha_3 \bar{p}_{t-2} + \varepsilon_t \\
p_{i,t} &= c + \beta_1 p_{i,t-1} + \beta_2 \bar{p}^e_{i,t} + \beta_3 \Pi_{i,t-1} + \beta_4 S_{i,t-1} + u_t \\
q_{i,t} &= c + \gamma_1 q_{i,t-1} + \gamma_2 p_{i,t} + \gamma_3 \bar{p}^e_{i,t} + \gamma_4 S_{i,t-1} + \eta_t \ ,
\end{aligned}$$

where $\bar{p}$ refers to realizations of the aggregate price, the variable $\bar{p}^e_i$ refers to individual forecasts of the aggregate price, the variable p_i refers to individual prices, S_i refers to individual excess supply/demand and q_i denotes individual quantities. The variable Π_i is a profit-feedback proxy defined as $\Pi_i = \Delta p_i \cdot \mathrm{sign}(\Delta \pi_i)$, where π_i are individual profits and the Δ is the first order difference operator. Assenza et al. identify three types of behavioral strategies on the basis of estimated price setting rules: *market followers*, i.e., subjects for which $\beta_3 = \beta_4 = 0$, *profit-adjusters*, i.e., subjects for which $\beta_3 > 0$ and $\beta_4 = 0$, and *demand-adjusters*, i.e., subjects for which $\beta_3 = 0$ and $\beta_4 > 0$. Overall, 46% of the subjects are market followers, 28% are profit-adjusters and 26% are demand-adjusters.

The authors investigate the impact of each behavioral type on market dynamics by means of simulations. The main findings can be summarized as follows: (a) profit-adjusters play a key role in pushing the experimental market towards the MC equilibrium; (b) the anchor term, i.e., the first three components in the price setting rule, is important to determine the (long run) equilibrium price and its stability; (c) demand-adjusters move their price in the direction that reduces excess supply, acting as loss-minimizes.

Although firms display a higher level of coordination in Treatment 2, Fig. 14 shows that, in the same treatment, the difference between market price and the MC equilibrium is higher. Assenza et al. attribute the different market behavior in the two treatments to the different information sets available to firms. In particular, the limited information in Treatment 1 provides subjects with an incentive to "explore" the demand function by experimenting with different prices. In order to support this hypothesis, the authors construct a proxy for individual exploration via price experimentation and confirm their conjecture. The lower level of exploration in Treatment 2 leads firms to rely on the information conveyed by market prices in their price setting decisions. This results in a higher level of coordination among firms, and at the same time it slows down convergence to the MC equilibrium. In fact, inertia in price-setting behavior of a significant share of firms (i.e., demand-adjusters with less incentive to explore and market followers) prevents profit-adjusters from pushing the market towards the MC equilibrium. In fact in this scenario profit-adjusters may realize low profits if they deviate too much from the average price, even if they move in a direction corresponding to a positive slope in the profit function, causing the market to lock in sub-optimal regions. This explains the stylized fact of lower degree of convergence to equilibrium in Treatment 2.

5 Concluding remarks

This chapter surveys laboratory experiments on expectation formation in macroeconomics and finance. We summarize three important findings of this literature focusing on key differences in the experimental designs: (1) exogenous time series versus endogenous expectations feedback, (2) positive versus negative feedback and (3) learning-to-optimize versus learning-to-forecast.

To our best knowledge, there is no systematic study comparing forecasting exogenously given time series and forecasting time series within an endogenous expectations feedback environment. The following comparison may be instructive however. In one of the treatments of Hey (1994), subjects must forecast an exogenous stochastic AR(1) process $x_t = 50 + \rho(x_{t-1} - 50) + \varepsilon_t$, with mean 50 and persistence coefficient $\rho = +0.9$. Using individual forecast series, Hey estimates simple trend-extrapolating rules of the form

$$x_t^e = \alpha_1 x_{t-1} + \alpha_2 (x_{t-1} - x_{t-2}), \tag{28}$$

and for most subjects finds a coefficient α_1 not significantly different from 1. The coefficient α_2 varies across subjects and assumes positive as well as negative values for different subjects. For about one third of the subjects the coefficient is significantly different from 0, either positive or negative, and Hey presents examples of -0.27 and $+0.21$[24]. This means that for a simple exogenous AR(1) process, subjects learn a simple trend-extrapolating rule, but they disagree about the sign of the coefficient. Some subjects are trend-followers, while others are contrarians and go against the trend. Heemeijer et al. (2009) run LtFEs with endogenous expectations feedback

[24]Hey (1994) does not report all estimates.

and a linear feedback map

$$x_t = 60 + \rho(\overline{x}_t^e - 60) + \varepsilon_t, \tag{29}$$

where $\overline{x}_t^e$ is the average forecast of a group of individuals, ε_t is a (small) noise term and the slope coefficient of the linear feedback map is either positive ($\rho = 0.95$) or negative ($\rho = -0.95$). Heemeijer et al. (2009) also estimate first-order linear rules to the individual forecast series. In the positive feedback treatment for most subjects they find a positive trend-coefficient, ranging from 0.27 to 0.94. Apparently, in self-referential positive feedback systems, subjects learn to coordinate on (strong) trend-extrapolating rules. In the negative feedback treatment, the estimated trend-coefficient in most cases is not significant, and in the few significant cases the trend-coefficient is negative.

This brings us to the second important finding, the difference in aggregate behavior between positive and negative expectations feedback systems. A frequently heard argument in macroeconomics and finance is that at the aggregate level expectations should be rational, because on average individual errors wash out at the aggregate level. The evidence from the laboratory experiments however shows that this is only true under negative expectations feedback, but not under positive feedback. LtFEs show that under positive feedback (small) individual errors may become strongly correlated, individual expectations may coordinate on prices very different from the rational expectations benchmark and prices do not converge but rather fluctuate around the fundamental. Surprisingly, oscillating prices already arise in positive feedback systems with a slope coefficient *less than* 1, e.g. 0.95. Most adaptive learning algorithms, including the simple naive expectations rule, would predict convergence to equilibrium in this case, but the experiments show that at the aggregate level prices may oscillate persistently. For an intuitive explanation we refer the reader once more to the graph of the positive feedback map (Figure 4): for a near-unit root linear feedback map, the graph almost coincides with the diagonal so that every point is almost a steady state. As a consequence, at any moment in time, any price forecast is almost self-fulfilling. Subjects may then easily coordinate on a dynamic price pattern very different from the unique RE price steady state with small forecasting errors.

In the case of 2-period ahead forecasts, as is common in temporary equilibrium models in macro and finance such as the NK DSGE and asset pricing frameworks, the LtFE takes the form (cf. Eq. 7)

$$x_t = 60 + \rho(\overline{x}^e_{t+1} - 60) + \varepsilon_t, \tag{30}$$

where $\overline{x}^e_{t+1}$ is the average 2-period ahead forecast of a group of individuals. In this setup price volatility strongly increases with large bubbles and crashes (as in Figure 2). In this type of temporary equilibrium framework the coefficient ρ often represents a discount factor close to 1, so that the system exhibits strong positive feedback.

Finally, let us discuss differences between the learning-to-forecast and learning-to-optimize designs. Price oscillations, with bubbles and crashes, have been observed in many experimental studies within a learning-to-forecast design. This design fits with models where consumption, savings, production and investment quantity decisions are optimal, given subjective forecasts. Recent experimental studies with a learning-to-optimize design show that convergence to RE may be even slower and instability under positive feedback may be even stronger. For subjects in laboratory experiments, learning-to-optimize seems even more difficult than learning-to-forecast. This experimental evidence calls for relaxing the rationality assumption in utility, profit and portfolio optimization and more realistic modeling of boundedly rational, heterogeneous decision heuristics in macroeconomics and finance.

Much work remains to be done on the empirical validation of individual expectations and aggregate behavior in experimental economic feedback systems. An important question, for example, is the robustness of these results in large groups. The fact that expectations at the aggregate level may persistently deviate from rationality has important policy implications. Policy analysis is often based upon RE models. But if RE fails the empirical test of simple laboratory environments, can we trust macroeconomic and financial policy analysis based on the rational paradigm? This survey indicates a potentially successful strategy for policy to manage self-referential systems of heterogeneous boundedly rational agents. In order to stabilize macroeconomic or financial expectations feedback systems, policy should add negative feedback to the system and weaken the positive feedback so that coordination on destabilizing trend-following behavior becomes less likely and the system is more likely to coordinate on stabilizing adaptive expectations.

References

[1] Adam, K., (2007), Experimental Evidence on the Persistence of Output and Inflation, Economic Journal, 117, 603-635.

[2] Akerlof, G.A., and R.J. Shiller (2009): *Animal Spirits: How Human Psychology Drives the Economy, and Why It Matters for Global Capitalism"*. Princeton University Press.

[3] Akiyama, E., Hanaki, N., and Ishikawa, R. (2013). How Do Experienced Traders Respond to Inflows of Inexperienced Traders? An Experimental Analysis, SSRN working paper.

[4] Anufriev, M.,and Hommes, C.H., (2012), Evolutionary Selection of Individual Expectations and Aggregate Outcomes in Asset Pricing Experiments, American Economic Journal: Microeconomics 4(4), 35-64.

[5] Anufriev, M., Hommes, C.H. and Philipse, R., (2013), Evolutionary Selection Of Expectations In Positive And Negative Feedback Markets, Journal of Evolutionary Economics 23, 663-688.

[6] Arifovic, J. (1996), The Behavior of the Exchange Rate in the Genetic Algorithm and Experimental Economies, *Journal of Political Economy* 104: 510-541.

[7] Arifovic, J. and T. Sargent (2003), Laboratory Experiments with an Expectational Phillips Curve, in D. Altig and B. Smith (eds.), The Origins and Evolution of Central Banking: Volume to inaugurate the Institute on Central Banking of the Federal Reserve Bank of Cleveland, Cambridge University Press, pp 23 - 56.

[8] Asparouhova, E., Hertzel, M., and Lemmon, M. (2009). Inference From Streaks in Random Outcomes: Experimental Evidence on Beliefs in Regime Shifting and The Law of Small Numbers. Management Science, 55(11), 1766-1782.

[9] Asparouhova, E., Bossaerts, P., Eguia, J., and Zame, W. (2014). Asset Prices and Asymmetric Reasoning, Working Paper of Department of Economics, University of Bristol.

[10] Asparouhova, E., Bossaerts, P., Roy, N. and Zame, W. (2013), 'Lucas' in the Laboratory, Caltech and UCLA Working Paper, May 2013.

[11] Assenza, T., Grazzini, J., Hommes, C.H. and Massaro, D. (2014), PQ Strategies in Monopolistic Competition: Some Insights from the Lab, CeNDEF Working Paper, University of Amsterdam.

[12] Assenza, T., Heemeijer, P., Hommes, C.H. and Massaro, D. (2012), Individual Expectations and Aggregate Macro Behavior, CeNDEF Working Paper, University of Amsterdam.

[13] Assenza, T., Heemeijer, P., Hommes, C.H. and Massaro, D. (2014b), Managing Self-Organization Of Expectations Through Monetary Policy: A Macro Experiment, CeNDEF Working Paper, University of Amsterdam.

[14] Bao, T., and Duffy, J. (2014), Adaptive vs. Eductive Learning: Theory and Evidence, SOM Research Report, University of Groningen.

[15] Bao, T., Duffy, J., and Hommes, C.H. (2013). Learning, Forecasting and Optimizing: An Experimental Study. European Economic Review, 61, 186-204.

[16] Bao, T., Hommes, C.H., Sonnemans, J. and Tuinstra, J. (2012), Individual Expectations, Limited Rationality and Aggregate Outcomes, Journal of Economic Dynamics and Control 36, 1101-1120.

[17] Bao, T., Hommes, C. H., and Makarewicz, T. A. (2014). Bubble Formation and (In) efficient Markets in Learning-to-Forecast and-Optimize Experiments, CeNDEF Working Paper (No. 14-01). Universiteit van Amsterdam.

[18] Barberis, N., A. Shleifer, R. Vishny. (1998). A Model of Investor Sentiment. Journal of Financial Economics. 49(3) 307-C343

[19] Becker, O., and Leopold-Wildburger, U. (1996). Some New Lotka-Volterra-Experiments. In Operations Research Proceedings 1995 (pp. 482-486). Springer Berlin Heidelberg.

[20] Becker, O., Leitner, J. and Leopold-Wildburger, U. (2009), Expectations Formation and Regime Switches, *Experimental Economics* 12, 350-364.

[21] Beckman, S.R. and Downs, D.H. (2009), Forecasters As Imperfect Information Processors: Experimental And Survey Evidence, *Journal of Economic Behaviour & Organization* 32, 89-100.

[22] Bernasconi, M., Kirchkamp, O. and Paruolo, P. (2009), Do Fiscal Variables Affect Fiscal Expectations? Experiments with Real World and Lab Data. *Journal of Economic Behaviour & Organization* 70, 253-265.

[23] Beshears, J., Choi, J.J., Fuster, A., Laibson, D. and Madrian, B.C. (2013), What Goes Up Must Come Down? Experimental Evidence on Intuitive Forecasting, American Economic Review 103, Paper and Proceedings, 570-574.

[24] Blanco, M., Engelmann, D., Koch, A. K., and Normann, H. T. (2010). Belief Elicitation in Experiments: Is There a Hedging Problem?. Experimental Economics, 13(4), 412-438.

[25] Bloomfield, R., and J. Hales, (2002). Predicting the Next Step of a Random Walk: Experimental Evidence of Regime-Shifting Beliefs, Journal of Financial Economics 65, 397- 414.

[26] Branch, W.A. (2004). The Theory of Rationally Heterogeneous Expectations: Evidence from Survey Data on Inflation Expectations, *Economic Journal*, 114, 592-621.

[27] Breaban, A., and Noussair, C. N. (2013). Emotional State and Market Behavior, CENTER Working Paper No. 2013-031, University of Tilburg.

[28] Brock, W.A. and Hommes, C.H. (1997), A Rational Route To Randomness, Econometrica 65, 1059-1095.

[29] Bullard, J., (1994). Learning Equilibria. Journal of Economic Theory 64, 468-485.

[30] Case, K. E., Shiller, R. J., and Thompson, A. (2012). What Have They Been Thinking? Home Buyer Behavior in Hot and Cold Markets (No. w18400). National Bureau of Economic Research.

[31] Colander, D., P. Howitt, A. Kirman, A. Leijonhufvud, and P. Mehrling (2008): "Beyond DSGE Models: Toward an Empirically Based Macroeconomics". *American Economic Review*, 98(2), 236–240.

[32] Cornand, C. and Heinemann, F. (2014), Experiments on Monetary Policy and Central Banking, Chapter in this volume.

[33] Cornand, C. and M'baye, C. K. (2013), "Does Inflation Targeting Matter? An Experimental Investigation", Working paper GATE 2013-30.

[34] Crockett, S., and J. Duffy, (2013) "An Experimental Test of the Lucas Asset Pricing Model," University of Pittsburgh Working Paper.

[35] Davis D. and Korenok, O. (2011), "Nominal Price Shocks in Monopolistically Competitive Markets: An Experimental Analysis", Journal of Monetary Economics 58, 578–589.

[36] DeGrauwe, P. (2012), Lectures on Behavioral Macroeconomics, Princeton University Press.

[37] Duffy, J., (2008). "Experimental Macroeconomics", Entry in: S. Durlauf and L. Blume, (Eds.), *The New Palgrave Dictionary of Economics*, 2nd ed., New York: Palgrave Macmillan.

[38] Duffy, J., and E.O. Fisher, (2005). Sunspots in the Laboratory. American Economic Review, 95, 510-529.

[39] Dwyer, G.P., A.W. Williams, R.C. Battalio and T.I. Mason, (1993), "Tests of Rational Expectations in a Stark Setting," *Economic Journal*, 103, 586-601.

[40] Evans, G.W., Honkapohja, S., (2001). Learning and Expectations in Macroeconomics. Princeton University Press, Princeton.

[41] Fehr, E., and J.R. Tyran, (2005). "Individual Irrationality and Aggregate Outcomes," Journal of Economic Perspectives, 19, 43-66.

[42] Fehr, E., and J.R. Tyran, (2008). "Limited Rationality and Strategic Interaction: the Impact of the Strategic Environment on Nominal Inertia", *Econometrica*, 76(2), 2008, 353-394.

[43] Fisher, F.M., (1962). A Priori Information and Time Series Analysis. North-Holland, Amsterdam.

[44] Gaechter, S., and Renner, E. (2010). The Effects of (Incentivized) Belief Elicitation in Public Goods Experiments. Experimental Economics, 13(3), 364-377.

[45] Harrison, J. M., and Kreps, D. M. (1978). Speculative Investor Behavior in A Stock Market With Heterogeneous Expectations. The Quarterly Journal of Economics,92(2), 323-336.

[46] Haruvy, E., Lahav, Y., and Noussair, C. N. (2007). Traders' Expectations in Asset Markets: Experimental Evidence. American Economic Review, 97(5), 1901-1920.

[47] Heemeijer, P., Hommes, C.H., Sonnemans, J., Tuinstra, J., (2009). Price Stability and Volatility In Markets With Positive And Negative Expectations Feedback. Journal of Economic Dynamics & Control 33, 1052-1072.

[48] Hey, J. (1994), Expectations formation: Rational or adaptive or ...?, *Journal of Economic Behavior & Organization* 25, 329-349

[49] Bostian, A.J.A. and Holt, C.A. (2013), Price Bubbles With Discounting: A Web-Based Classroom Experiment, *Journal of Economic Education* 40, 27-37.

[50] Hommes, C.H., (2011). The Heterogeneous Expectations Hypothesis: Some Evidence for the Lab. Journal of Economic Dynamics & Control 35, 1-24.

[51] Hommes, C.H., (2013), Reflexivity, Expectations Feedback and Almost Self-Fulfilling Equilibria: Economic Theory, Empirical Evidence and Laboratory Experiments, *Journal of Economic Methodology* 20, 406-419.

[52] Hommes, C.H., (2013), Behavioral Rationality and Heterogeneous Expectations in Complex Economic Systems, Cambridge University Press, Cambrdige.

[53] Hommes, C.H., (2014), Behaviorally Rational Expectations and Almost Self-Fulfilling Equilibria, *Review of Behavioral Economics* 1, 75-97.

[54] Hommes, C., Sonnemans, J., and van de Velden, H. (2000). Expectation Formation in a Cobweb Economy: Some one Person Experiments. In Interaction and Market Structure. Springer Berlin Heidelberg.

[55] Hommes, C.H., Sonnemans, J.H., Tuinstra, J. and van de Velden, H, (2005). "Coordination of Expectations in Asset Pricing Experiments". *Review of Financial Studies*, 18 (3), 955-980.

[56] Hommes, C.H., Sonnemans, J.H., Tuinstra, J.,and van de Velden, H., (2008). "Expectations and Bubbles in Asset Pricing Experiments". *Journal of Economic Behavior and Organization* 67, 116-133.

[57] Hommes, C.H., Sonnemans, J. and Velden, H. van de, (2000), Expectation Formation in a Cobweb Economy: some one Person Experiments, In: Delli Gatti, D., Gallegati, M., and Kirman, A.P. (Eds.) *Interaction and Market Structure*, Springer Verlag Berlin (2000), pp.253-266

[58] H*ü*sler, A., Sornette, D., and Hommes, C. H. (2013). Super-Exponential Bubbles in Lab Experiments: Evidence for Anchoring Over-Optimistic Expectations on Price. Journal of Economic Behavior and Organization, 92, 304-C316.

[59] Kelley, H. and D. Friedman, (2002), "Learning to Forecast Price," *Economic Inquiry*, 40, 556–573.

[60] Kirman, A. (2010), Complex Economics: Individual and Collective Rationality. Routledge, Oxford.

[61] Leitner, J. and Schmidt, R. (2007), Expectations Formation in an Experimental Foreign Exchange Market, *Central European Journal of Operations Research* 15, 167-184.

[62] Lim, S.S., E.C. Prescott and S. Sunder (1994), "Stationary Solution to the Overlapping Generations Model of Fiat Money: Experimental Evidence," Empirical Economics 19, 255-77.

[63] Lucas, R., (1972). Expectations and the Neutrality Of Money. Journal of Economic Theory 4, 103-124.

[64] Marimon, R., S.E. Spear and S. Sunder, (1993), "Expectationally Driven Market Volatility: an Experimental Study," *Journal of Economic Theory*, 61, 74–103.

[65] Marimon, R. and S. Sunder, "Indeterminacy of Equilibria in a Hyperinflationary World: Experimental Evidence", *Econometrica*, 61(5), 1993, 1073-1107.

[66] Marimon, R. and S. Sunder, "Expectations and Learning under Alternative Monetary Regimes: An Experimental Approach", *Economic Theory*, 4, 1994, 131-62.

[67] Marimon, R. and S. Sunder, (1995). "Does a constant money growth rule help stabilize inflation?" *Carnegie-Rochester Conference Series on Public Policy*, 43, 111-156.

[68] Muth, J.E., (1961). Rational Expectations and the Theory of Price Movements. Econometrica 29, 315-335.

[69] Nagel, R. (1995), Unraveling in Guessing Games: An Experimental Study, *American Economic Review* 85, 1313-1326.

[70] Noussair, C., Plott, C., and Riezman, R. (2007). Production, trade, prices, exchange rates and equilibration in large experimental economies. European Economic Review, 51(1), 49-76.

[71] Noussair, C. N., and Powell, O., (2010). "Peaks and Valleys: Price Discovery in Experimental Asset markets with Non-Monotonic Fundamental", *Journal of Economic Studies*, 37(2), 152-180.

[72] Noussair, C. N., and Tucker, S. (2013). Experimental Research on Asset Pricing. Journal of Economic Surveys, 27(3), 554-569.

[73] Nyarko, Y., and Schotter, A. (2002). An Experimental Study of Belief Learning Using Elicited Beliefs. Econometrica, 70(3), 971-1005.

[74] Pesaran, M. H. (1987). The limits to rational expectations. Oxford: Blackwell.

[75] Pfajfar, D. and B. Zakelj, (2011) "Inflation Expectations and Monetary Policy Design: Evidence from the Laboratory" *CentER Discussion Paper 2011 - 091*, Tilburg University.

[76] Pfajfar, D. and B. Zakelj, (2014) Experimental Evidence on Inflation Expectation Formation. *Journal of Economic Dynamics and Control*, 44, 147-168.

[77] Sargent, T.J., (1993). Bounded Rationality in Macroeconomics. Oxford University Press Inc., New York.

[78] Peterson, S. P. (1993). Forecasting Dynamics and Convergence to Market Fundamentals: Evidence from Experimental Asset Markets. Journal of Economic Behavior and Organization, 22(3), 269-284.

[79] Potters, J. and S. Suetens., (2009), Cooperation in Experimental Games of Strategic Complements and Substitutes, *Review of Economic Studies*, 76(3), 1125-1147.

[80] Rabin, M., (2002), Inference by Believers in the Law of Small Numbers, Quarterly Journal of Economics 117, 775-816.

[81] Roos, M.W.M. and Luhan, W.J. (2013), Information, Learning and Expectations in an Experimental Model Economy, *Economica* 80, 513-531.

[82] Rubinstein, A., (2007) "Instinctive and Cognitive Reasoning: A Study of Response Times," Economic Journal 117, 1243-1259.

[83] Schmalensee, R., (1976), "An Experimental Study of Expectation Formation," *Econometrica*, 44, 17–41.

[84] Schotter, A. and Trevino, I. (2014), Belief Elicitation in the Lab, *Annual Review of Economics* 6.

[85] Shiller, R.J., (1990), “Speculative Prices and Popular Models,” *Journal of Economic Perspectives*, 4, 55–65.

[86] Shiller, R.J., (2000), Irrational Exuberance, Princeton University Press, Princeton.

[87] Smith, V.L., G.L. Suchanek and A.W. Williams, (1988), “Bubbles, Crashes and Endogenous Expectations in Experimental Spot Asset Markets,” *Econometrica*, 56, 1119-1151.

[88] Sonnemans, J.,Tuinstra, J., (2010) Positive Expectations Feedback Experiments and Number Guessing Games as Models of Financial Markets, *Journal of Economic Psychology*, 31(6), 964-984.

[89] Soros, G. (2003). The Alchemy of Finance. John Wiley and Sons.

[90] Soros, G. (2009). The Crash Of 2008 and What It Means: The New Paradigm for Financial Markets. PublicAffairs.

[91] Stöckl, T., Huber, J., and Kirchler, M. (2010). Bubble Measures in Experimental Asset Markets. *Experimental Economics*, 13(3): 284-298.

[92] Sutan, A. and Willinger, M., (2009), Guessing with Negative Feedback: An Experiment, *Journal of Economic Dynamics and Control* 33, 1123-1133.

[93] Tirole, J. (1982). On The Possibility of Speculation under Rational Expectations. Econometrica, 50(5), 1163-1181.

[94] Tversky, A. and Kahneman, D. (1974), Judgment Under Uncertainty: Heuristics And Biases, *Science* 185, 1124-1131.

[95] Wagener, F.O.O. (2013), Experiments and Expectations, *Annual Reviews of Economics*, forthcoming.

[96] Williams, A.W., (1987), "The Formation of Price Forecasts in Experimental Markets," *Journal of Money, Credit and Banking*, 19, 1-18.

[97] Woodford, M., (1990). Learning to Believe in Sunspots. Econometrica 58, 277-307.

[98] Woodford, M., (2003), Interest and Prices: Foundations of a Theory of Monetary Policy, Princeton University Press, Princeton, NJ.

[99] Xiong, W., and Yan, H. (2010). Heterogeneous Expectations and Bond Markets. Review of Financial Studies, 23(4), 1433-1466.