Journal of Economic Dynamics & Control 35 (2011) 1–24

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

The heterogeneous expectations hypothesis: Some evidence from the lab

Cars Hommes *

CeNDEF, School of Economics, University of Amsterdam, Roetersstraat 11, NL-1018 WB Amsterdam, Netherlands

ARTICLE INFO

Article history:
Received 9 April 2010
Accepted 4 October 2010
Available online 8 October 2010

JEL classification:
E52
D83
D84
C62

Keywords:
Heterogeneous expectations
Bounded rationality
Learning
Heuristics switching

ABSTRACT

This paper surveys learning-to-forecast experiments (LtFEs) with human subjects to test theories of expectations and learning. Subjects must repeatedly forecast a market price, whose realization is an aggregation of individual expectations. Emphasis is given to how individual forecasting rules interact at the micro-level and which structure they cocreate at the aggregate, macro-level. In particular, we focus on the question wether the evidence from laboratory experiments is consistent with heterogeneous expectations.

"Recent theoretical work is making it increasingly clear that the multiplicity of equilibria…can arise in a wide variety of situations involving sequential trading, in competitive as well as finite-agent games. All but a few of these equilibria are, I believe, behaviorally uninteresting: They do not describe behavior that collections of adaptively behaving people would ever hit on. I think an appropriate stability theory can be useful in weeding out these uninteresting equilibria…But to be useful, stability theory must be more than simply a fancy way of saying that one does not want to think about certain equilibria. I prefer to view it as an experimentally testable hypothesis, as a special instance of the adaptive laws that we believe govern all human behavior" (Lucas, 1986, pp. S424–425).

1. Introduction

Individual expectations about future aggregate outcomes are the key feature that distinguishes social sciences and economics from the natural sciences. Daily weather forecasts, either by the public or by experts, do not affect the probability of rain. In contrast, overly optimistic expectations about the economic outlook may have exaggerated the strong rise in world wide financial markets in the late 1990s and, more recently, the excessive growth in housing prices in 2000–2008, while an overly pessimistic outlook by the public and by economists may have amplified the recent financial crisis and deepened the current economic crisis. Individual economic decisions today thus depend upon expectations about the future state of the global economy. A theory of individual expectations or market beliefs is therefore a crucial part of economic theory. Markets are expectations feedback systems and any dynamic economic model depends crucially

* Tel.: +31 20 5254246.
E-mail address: c.h.hommes@uva.nl

0165-1889/$ - see front matter
doi:10.1016/j.jedc.2010.10.003

upon its underlying *expectations hypothesis*. But how then should one model individuals who learn from past mistakes and adapt their behavior as more and more market realizations become available over time?

Since the seminal works of Muth (1961) and Lucas (1972), the rational expectations hypothesis (REH) has become the leading paradigm on expectation formation in economics, and rational expectations representative agent models have become the mainstream tool of analysis. In such a framework, all agents are the same and forecast rationally, using all available information. Rational expectations are by assumption *model consistent* and coincide on average with realizations, without systematic forecasting errors. The rational expectations (RE) approach has important advantages: it is simple, elegant and puts strong discipline on individual (forecasting) behavior minimizing the number of free parameters. But drawbacks of the rational agent paradigm are also well known: it is unrealistic in assuming perfect knowledge about the economy (it assumes essentially that the law of motion of the economy is common knowledge) and, even if such knowledge were available, RE requires extremely strong computing abilities of the agents to compute the equilibrium. Most importantly, RE models are at odds with empirical observations and behavior in laboratory experiments wit human subjects. For example, the decline of worldwide financial markets by almost 50% between October 2008 and March 2009 is hard to reconcile with rational behavior.

Economics, or at least a significant part of it, is currently witnessing a paradigm shift to an alternative, behavioral view, where agents are *boundedly rational*. This alternative view dates back (at least) to Simon (1957) and contains elements from psychology, e.g. through the work of Tversky and Kahneman (1974). The need for a new paradigm in economics has recently been forcefully advocated by Akerlof and Shiller (2009), Colander et al. (2009), DeGrauwe (2009) and Kirman (2010). Concerning expectations of boundedly rational agents, an alternative theory of *adaptive learning* has been developed, see, e.g. Sargent (1993) for an early and Evans and Honkapohja (2001) for a more detailed overview. Boundedly rational agents do *not* know the true law of motion of the economy, but instead use time series observations to form expectations and adapt their behavior over time by trying to learn the model parameters of their *perceived law of motion* as more observations become available. Adaptive learning sometimes enforces convergence to REE, but it may also lead to non-RE equilibria, such as the learning equilibria in Bullard (1994). Adaptive learning models are sometimes ‘‘cautious modifications of rational expectations theories’’ (Sargent, 2008, p. 26) and at other times large deviations from rationality explaining excess volatility through expectations driven fluctuations (Grandmont, 1998). Bounded rationality, however, also has important drawbacks. In particular, the ‘‘wilderness’’ of bounded rationality (Sims, 1980) creates (too) many degrees of freedom and too many free parameters. There are simply too many ways of modeling non-rational behavior. This ‘‘wilderness’’ of bounded rationality, seems particularly problematic when one allows individuals to have *heterogeneous expectations*.

A rough estimate indicates that in the past 20 years more than 1000 papers on *bounded rationality* and *learning* have appeared. Among these, in the last decade hundreds of papers on agent-based models populated with boundedly rational agents employing *heterogeneous* strategies/expectations have appeared, especially with applications in finance; see the comprehensive surveys of LeBaron (2006) and Hommes (2006) in the *Handbook of Computational Economics Volume 2*: *Agent-Based Computational Economics* (Tesfatison and Judd, 2006) and more recent surveys by Hommes and Wagener (2009) and Chiarella et al. (2009) in the *Handbook of Financial Markets: Dynamics and Evolution* (Hens and Schenk-Hoppé, 2009) as well as papers and references in the *Handbook of Economic Complexity* (Rosser, 2009).[1] Most of these papers present either stylized theoretical models or larger agent-based simulation models, with the realistic feature that these models can mimic many *stylized facts* in financial time series (Lux, 2009) and in macro-data (Delli-Gatti et al., 2008).

The empirical validation of heterogeneity is an important area of current research. For example, Baak (1999) and Chavas (2000) estimate heterogeneous agent models (HAMs) on hog and beef market data, and found empirical evidence for heterogeneity of expectations. For the beef market Chavas (2000) estimates that about 47% of the beef producers behave naively (using only the last price in their forecast), 18% of the beef producers behaves rationally, whereas 35% behaves quasi-rationally (i.e. use a univariate autoregressive time series model to forecast prices). A number of recent contributions have estimated heterogeneous agent models with fundamentalists and chartist strategies on stock prices (e.g. Boswijk et al., 2007; Amilon, 2008; de Jong et al., 2009), exchange rates (e.g. Gilli and Winker, 2003, Westerhoff and Reitz, 2003 and Franke, 2009), stock option prices (Frijns et al., 2010) and several commodities (e.g. gold prices, Alfarano et al., 2005, and oil prices, ter Ellen and Zwinkels, forthcoming). Most of these studies find significant time-variation in the fractions of agents using a mean-reverting fundamental versus a trend-following strategy. Empirical evidence for heterogeneous trading strategies in the Spanish Stock Market Exchange has been found in Vaglica et al. (2008); in particular, Lillo et al. (2008) show that the investors can be classified in different groups, including trend followers and contrarians, whose change of inventory of the stock is positively, respectively, negatively correlated with stock returns.

A related, complementary branch of empirical literature uses *survey data* to measure individual expectations; see Pesaran and Weale (2006) for a stimulating overview. An advantage of survey data analysis is that it can focus exclusively on the expectations generating process, avoiding the dilemma of testing joint hypotheses. There is quite some evidence on

[1] Some very recent contributions include, among others, Anufriev and Dindo (2010), Assenza and Berardi (2009), Bauer et al. (2009), Branch and McGough (2009), Brock et al. (2009), Anufriev and Panchenko (2009), Bekiros (2010), Dieci and Westerhoff (2010), Guse (2010), Hommes and Wagener (2010), Huang et al. (2010) and Lines and Westerhoff (2010). There is also an extensive literature on heterogeneous belief models based on heterogeneous information, see, e.g. Williams (1977), Shefrin and Statman (1994), Zapatero (1998), Basak (2000), Anderson et al. (2005), Goldbaum (2005) and Diks and Dindo (2008).

forecasting heterogeneity in survey data. For example, Frankel and Froot (1987, 1990), Allen and Taylor (1990) and Taylor and Allen (1992) already found that financial experts use different forecasting strategies to predict exchange rates. They tend to use trend extrapolating rules at short horizons (up to 3 months) and mean-reverting fundamentalists rules at longer horizons (6 months to 1 year) and, moreover, the weight given to different forecasting techniques changes over time. Vissing-Jorgensen (2003) presents evidence of heterogeneous beliefs of individual investors about the prospect of the stock market, while Shiller (2000) finds evidence that investor's sentiment changes over time, with both institutions and individual investors becoming more optimistic in response to recent significant increases of the stock market. Evidence concerning heterogeneity in survey data on exchange rate expectations can also be found in MacDonald and Marsh (1996), Elliott and Ito (1999), Prat and Uctum (2000) and Bénassy-Quéré et al. (2003). Dreger and Stadtmann (2008) show that for exchange rate forecasts at a 6 months horizon, different expectations about macroeconomic fundamentals are what drives heterogeneity. Mankiw et al. (2003) find evidence for heterogeneity in inflation expectations in the Michigan Survey of Consumers and show that the data are inconsistent with either rational or adaptive expectations, but may be consistent with a sticky information model. Capistrán and Timmermann (2009) show that heterogeneity of inflation expectations of professional forecasters varies systematically over time, and depends on the level and the variance of current inflation. Menkhoff et al. (2009) use expectations from about 300 forecasters over 15 years and show that the dispersion of exchange rate expectations varies over time and that determinants of dispersion are consistent with the chartist-fundamentalists approach: misalignments of the exchange rate and exchange rate changes explain heterogeneity. Pfajfar and Santoro (2010) measure the degree of heterogeneity in private agents' inflation forecasts by exploring time series of percentiles from the empirical distribution of survey data. They show that heterogeneity in inflation expectations is pervasive and identify three regions of the distribution corresponding to different expectations formation mechanisms: a static or highly autoregressive region on the left-hand side of the median, a nearly rational region around the median and a fraction of forecasts on the right-hand side of the median consistent with adaptive learning and sticky information. Branch (2004, 2007) estimates a simple switching model with heterogeneous expectations, along the lines of Brock and Hommes (1997), and provides empirical evidence suggesting that models which allow the degree of heterogeneity to change over time provide a better fit on exchange rate survey data.

In this paper, we discuss laboratory experiments with human subjects that can be used to validate expectations hypotheses and learning models. Lucas (1986) already stressed the importance of laboratory experiments in studying adaptive learning and its stability (see the quote at the beginning of this article). In particular, we are interested in the potential role of heterogeneity in expectations. We quote from Muth (1961, p. 321, emphasis added) on expectations heterogeneity and its aggregate effect: *"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"*.

In this paper we discuss *learning-to-forecast experiments* (LtFEs) which provide a controlled laboratory environment to study individual expectations as well as aggregate outcomes, and investigate questions such as:

- How do individuals form expectations and how do they *learn* and *adapt* their behavior?
- How do individual forecasting rules *interact* at the *micro*-level and which *aggregate outcome* do they co-create at the *macro*-level ?
- Will *coordination* occur, even when there is limited information, or will *heterogeneity* persist?
- When does *learning* enforce convergence to REE and when do boundedly rational "learning equilibria" arise?

The goals of this paper are twofold. Firstly, we summarize a number of LtFEs in *different* market environments. Secondly, we discuss how well a stylized heterogeneous expectations model explains the experimental data across different market settings. This poses a particular challenge: is there a general, perhaps even a universal theory of heterogeneous expectations, that is, can one come up with one single (heterogeneous) expectations hypothesis explaining all LtFEs across different market settings?

The paper is organized as follows. Section 2 summarizes related literature on experiments on expectations. Section 3 discusses an experiment in a cobweb market framework, while Section 4 describes an asset pricing experiment. Section 5 presents a simple forecasting heuristics switching model, where agents switch between different forecasting rules based upon their recent performance, and calibrates the model to the asset pricing experiments. In Section 6 the same switching model is calibrated to experimental data in different market settings, where the *only* difference comes from the type of expectations feedback, positive versus negative. Section 7 briefly discusses some recent experiments in a New Keynesian macro-setting, while Section 8 concludes.

2. Learning-to-forecast experiments (LtFEs)

Laboratory experiments with human subjects, with full control over the market environment and economic fundamentals, form an ideal tool to study interactions at the micro-level and how individual behavior shapes aggregate market outcomes. Duffy (2006, 2008a,b) provides stimulating and up to date surveys of "experimental macroeconomics".

Early work in this area focussed on market mechanisms, such as double auctions, and the amount of information and the presence of futures markets, ensuring that equilibrium will be reached (e.g. Smith, 1962; Plot and Smith, 1978; Plot and Sunder, 1982; Sunder, 1995). More recently, unstable market environments where equilibrium may not be reached, but instead bubbles and crashes may arise have also been designed (e.g. Smith et al., 1988; Lei et al., 2001).

In experimental work expectations often play an indirect or implicit role. However, in order to avoid joint hypothesis testing there is an expanding literature on exclusive experimental testing of the expectations hypothesis. An early example is Schmalensee (1976), who presents subjects with historical data on wheat prices and asks them to predict the mean wheat price for the next five periods. Williams (1987) considers expectation formation in an experimental double auction market, which varies from period to period by small shifts in the market clearing price. Participants predict the mean contract price for four or five consecutive periods and the participant with the lowest forecast error earns $1.00. In Dwyer et al. (1993) and Hey (1994) subjects have to predict a time series generated by an (exogenous) stochastic process such as a random walk or a simple linear first order autoregressive process. Kelley and Friedman (2002) consider learning in an Orange Juice Futures price forecasting experiment, where prices are driven by a linear stochastic process with two exogenous variables (weather and competing supply). But in these papers there is no *expectations feedback*, since market realizations are *not* affected by individual forecasts.

Here, we focus on so-called *learning-to-forecast experiments* (LtFEs), where subject's *only* task is to forecast the price of some commodity for a number, say 50–60, periods, with the realized market price in each period determined by (average) individual expectations. In LtFEs subjects' forecasting decisions are thus separated from market-trading decisions. The subjects in the experiments do not participate themselves directly in other market activities, such as trading or producing, but are forecasters (e.g. advisors to large producers or financial investors) whose earnings increase when forecasting errors decrease. At the beginning of each period, individual forecasts are collected, which feed directly into (unknown) demand and/or supply functions and computerized trading yields a market price, derived from equilibrium between aggregate demand and supply, that becomes available to the subjects at the end of the period. Demand and supply curves are derived from maximization of expected utility, profit or wealth and thus consistent with rational optimizing behavior.

These LtFEs were motivated by the bounded rationality literature and were designed to test various theories of expectations and learning in the lab. Sargent (1993) emphasizes two different requirements of rational expectations. The first requirement imposes that individuals maximize an objective function (utility, profit, wealth, etc.) subject to perceived constraints, while the second requirement imposes mutual consistency of these perceptions. Marimon and Sunder (1994) were the first to set up experiments testing individual rationality and mutual consistency either jointly or separately and used different experimental designs to distinguish between "learning-to-optimize" versus "learning-to-forecast" experiments (Marimon and Sunder, 1994, p. 134). The LtFEs focus exclusively on the role of expectations, using computerized optimal individual demand and supply schedules once individual forecasts have been made.

In LtFEs, subjects typically only have *qualitative* information about the market. They know that the price is an aggregation of individual forecasts, derived from equilibrium between demand and supply and are able to infer the type of expectations feedback, *positive* or *negative*. Positive (negative) feedback means that an increase of (average) individual forecasts leads to a higher (lower) market equilibrium price. Positive feedback is important in speculative asset markets, where higher market expectations lead to an increase of speculative demand and therefore an increase of the realized asset price. Negative feedback may be dominant in supply driven commodity markets, where an increase in expected prices leads to higher production and thus to a lower realized market price. Subjects in the LtFEs know past prices and their own past forecasts and earnings, typically in table as well as in graphic form, as illustrated by an example in Fig. 1. They do, however, *not* know the forecasts of other participants, the exact market equilibrium equation, the exact demand and supply schedules and the exact number of other demanders and/or suppliers in the market. The type of information in the

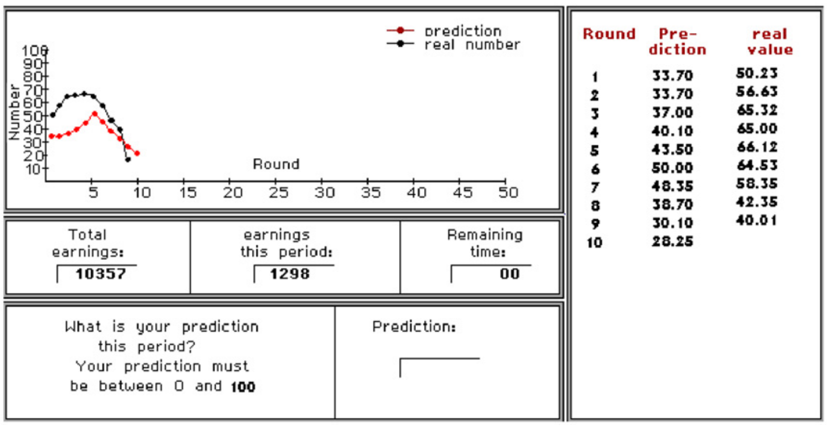

Fig. 1. Computerscreen in a learning-to-forecast experiment.

experiment is thus very similar to models of bounded rationality and adaptive learning, where agents try to learn a perceived law of motion, based upon time series observations without knowing the underlying actual law of motion of the market.

Quite a number of LtFEs have already appeared in the literature. In a series of papers, Marimon, Spear and Sunder (1993) and Marimon and Sunder (1993, 1994, 1995) studied expectation formation in inflationary overlapping generations economies. Marimon et al. (1993) find experimental evidence for expectationally driven cycles and coordination of beliefs on a sunspot two-cycle equilibrium, but only after agents have been exposed to exogenous shocks of a similar kind. Marimon and Sunder (1995) present experimental evidence that a "simple" rule, such as a constant growth of the money supply, can help coordinate agents' beliefs and help stabilize the economy. More recently, a number of LtFEs within other macroeconomic frameworks have been performed. Adam (2007) uses a simple model of sticky prices and shows that a restricted perception equilibrium explains the experimental data better than the RE benchmark solution. Pfajfar and Zakelj (2010) and Assenza et al. (2010) run LtFEs in a New Keynesian framework. Pfajfar and Zakelj (2010) find evidence for heterogeneity of expectations in their experimental data and three different types of forecasting rules: simple heuristics (e.g. trend-following rules), adaptive learning and rational expectations. The LtFEs of Assenza et al. (2010) will be briefly discussed in Section 7.

A learning-to-forecast experiment may be seen as a test bed for the expectations hypothesis in a benchmark model, assuming that all other assumptions, e.g. rational, utility and profit maximizing behavior, are satisfied. A learning-to-forecast experiment thus provides *clean data* on *individual expectations* as well as *aggregate price behavior*. Here we will discuss learning-to-forecast experiments, based on three benchmark models: (1) the cobweb model, (2) a standard asset pricing model and (3) a New Keynesian macro-model. The underlying laws of motion are of the form

$$p_t = F(p^e_{1,t}, \ldots, p^e_{H,t}) \qquad \text{cobweb,} \tag{2.1}$$

$$p_t = F(p^e_{1,t+1}, \ldots, p^e_{H,t+1}) \qquad \text{asset pricing,} \tag{2.2}$$

$$(\pi_t, y_t) = F(\pi^e_{1,t+1}, \ldots, \pi^e_{H,t+1}; y^e_{1,t+1}, \ldots, y^e_{H,t+1}). \qquad \text{New Keynesian.} \tag{2.3}$$

In the cobweb LtFE experiments in Hommes et al. (2007), the realized market price p_t in (2.1) is a (nonlinear) function of all individual one-period ahead forecasts $p^e_{h,t}$. In the asset pricing LtFE in Hommes et al. (2005a, 2008) the realized market price p_t in (2.2) is a (nonlinear) function of all two-period ahead individual forecasts $p^e_{h,t+1}$ of next period price p_{t+1}. There is another important difference between the cobweb and the asset pricing LtFEs: negative versus positive expectations feedback. Positive feedback means that the realized market price increases, when an individual price forecast increases. This feature is characteristic of speculative asset markets, where an increase of the price forecast leads to higher demand for the asset and therefore to higher market prices; mathematically it means that the map F in (2.2) is an increasing function of individual forecasts. Negative feedback prevails in supply driven markets, where a higher expected price leads to increased production and thus a lower realized market price; the map F in (2.1) underlying the cobweb experiments is decreasing in individual forecasts. In Section 6 we review the LtFE of Heemeijer et al. (2009), comparing positive versus negative feedback systems. Despite the fact that the *only* difference is the sign (positive versus negative) of the coefficient in the linear price-expectations feedback rule, the aggregate price behaviors and individual expectations turn out to be very different. Finally, in the New Keynesian macro-model expectations on two different variables interact and realized inflation π_t and realized output gap y_t in (2.3) simultaneously depend (linearly) on all two-period ahead individual forecasts of both inflation and the output gap. In Section 7 we will briefly discuss the LtFE of Assenza et al. (2010), where the dynamics of inflation and the output gap is driven simultaneously by individual expectations of both inflation and the output gap.

3. Cobweb experiments

In this section we discuss LtFEs in the classical cobweb framework, exactly the same framework employed in the seminal paper of Muth (1961) introducing rational expectations. These LtFEs may thus be seen as a direct test of the REH in the cobweb model, assuming all other modeling assumptions (e.g. producers' profit maximization and consumers utility maximization) are satisfied. Carlson (1967) already conducted hand-run experiments with subjects as cobweb suppliers. Holt and Villamil (1986) and Hommes et al. (2000) conducted individual cobweb experiments, where price fluctuations are induced by decisions of a single individual. Wellford (1989) conducted several computerized cobweb experiments, where market prices were determined by subjects' quantity decisions.

Here we focus on the LtFEs in Hommes et al. (2007) with K=6 participants in each market. The participants were asked to predict next period's price of a commodity under limited information on the structural characteristics of the market. Participants were only informed about the basic principles of the cobweb-type market. They were advisors to producers, whose only job it was to accurately forecast the price of the good for 50 subsequent periods. Pay-offs were defined as a quadratic function of squared forecasting errors, truncated at 0[2]:

$$E = \text{Max}\{1300 - 260(p^e_{i,t} - p^*_t)^2,\ 0\}. \tag{3.1}$$

[2] 1300 points correspond to 0.5 Euro, so that maximum earnings over 50 periods were 25 Euro's. Average earnings ranged from 11.5 to 21 Euro (in about 1.5 h) over different treatments.

Participants were informed that the price would be determined by market clearing and that it would have to be within the range [0,10]. Furthermore, they knew that there was a negative feedback from individual price forecasts to realized market price in the sense that if their forecast would increase, the supply would increase and consequently the market clearing price would decrease. Subjects, however, did not know how large these feedback effects would be, as they had no knowledge of underlying market equilibrium equations. Subjects thus had *qualitative* information about the market, but no quantitative details.

The realized price p_t in the experiment was determined by the (unknown) market equilibrium between demand and supply:

$$D(p_t) = \sum_{i=1}^{K} S(p^e_{i,t}), \tag{3.2}$$

with $p^e_{i,t}$ the price forecast of participant i at time t. Demand was exogenously given by a simple linear schedule:

$$D(p_t) = a - dp_t + \eta_t, \tag{3.3}$$

with η_t a small stochastic shock drawn from a normal distribution representing small random demand fluctuations. Supply $S(p^e_{i,t})$ was determined by the nonlinear schedule

$$S_\lambda(p^e_{i,t}) = \tanh(\lambda(p^e_{i,t} - 6)) + 1. \tag{3.4}$$

This increasing, nonlinear supply schedule can be derived from producer's expected profit maximization with a convex cost function. Subjects in the experiment thus do not participate themselves in production decisions, but supply is computed as if each individual producer maximizes expected profit, given his/her individual price forecast. The parameter λ tunes the nonlinearity of the supply curve and the stability of the underlying cobweb model. The resulting equilibrium price is obtained as

$$p_t = D^{-1}\left(\sum_{i=1}^{K} S_\lambda(p^e_{i,t})\right) = \frac{a - \sum_{i=1}^{K} S_\lambda(p^e_{i,t})}{d} + \varepsilon_t, \tag{3.5}$$

where $\varepsilon_t = \eta_t / d$. Given the parameters a,d and λ the aggregate realized price p_t depends on individual price expectations as well as the realization of the (small) stochastic shocks. While the parameters of the demand function and the realizations of the noise component remained unchanged across all treatments at a=13.8, d=1.5 and $\varepsilon_t = \eta_t/d \sim N(0,0.5)$, the slope parameter of the supply function was varied. Here we consider two treatments. A stable treatment had $\lambda = 0.22$, for which under naive expectations the price converges quickly to the rational expectations equilibrium. In another strongly unstable treatment, with $\lambda = 2$, under naive expectations the RE price is unstable and prices converge to a two-cycle, as illustrated in Fig. 2. Along the two-cycle producers are "irrational" in the sense that they make systematic forecasting errors, predicting a high (low) price when realized market price will be low (high).

Under rational expectations, all individuals would predict the unique price p^*, at which demand and supply intersect. Given that all individuals have rational expectations, realized prices will be given by

$$p_t = p^* + \varepsilon_t, \tag{3.6}$$

that is, small random fluctuations around the RE steady state. Given the limited market information one cannot expect that all individuals have rational expectations at the outset, but one can hope that in such a simple, stationary environment individuals would *learn to have rational expectations*. For example, if price expectations are given by the sample average of past prices, convergence to the RE price is enforced, as illustrated in Fig. 2 (right panel). The LtFE has in fact been designed to test whether individuals are able to learn from their systematic mistakes under naive expectations and coordinate on a learning algorithm enforcing convergence to the RE steady state.

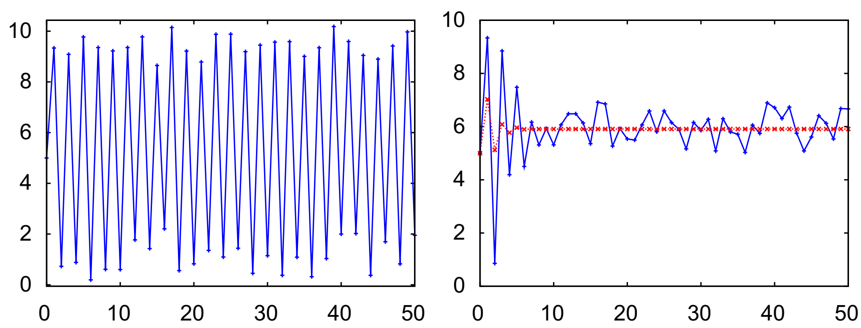

Fig. 2. Cobweb dynamics in the strongly unstable treatment ($\lambda = 2$) in two benchmark simulations. Left panel: convergence to a (noisy) two-cycle under naive expectations. Right panel: convergence to (noisy) RE equilibrium price under learning by average.

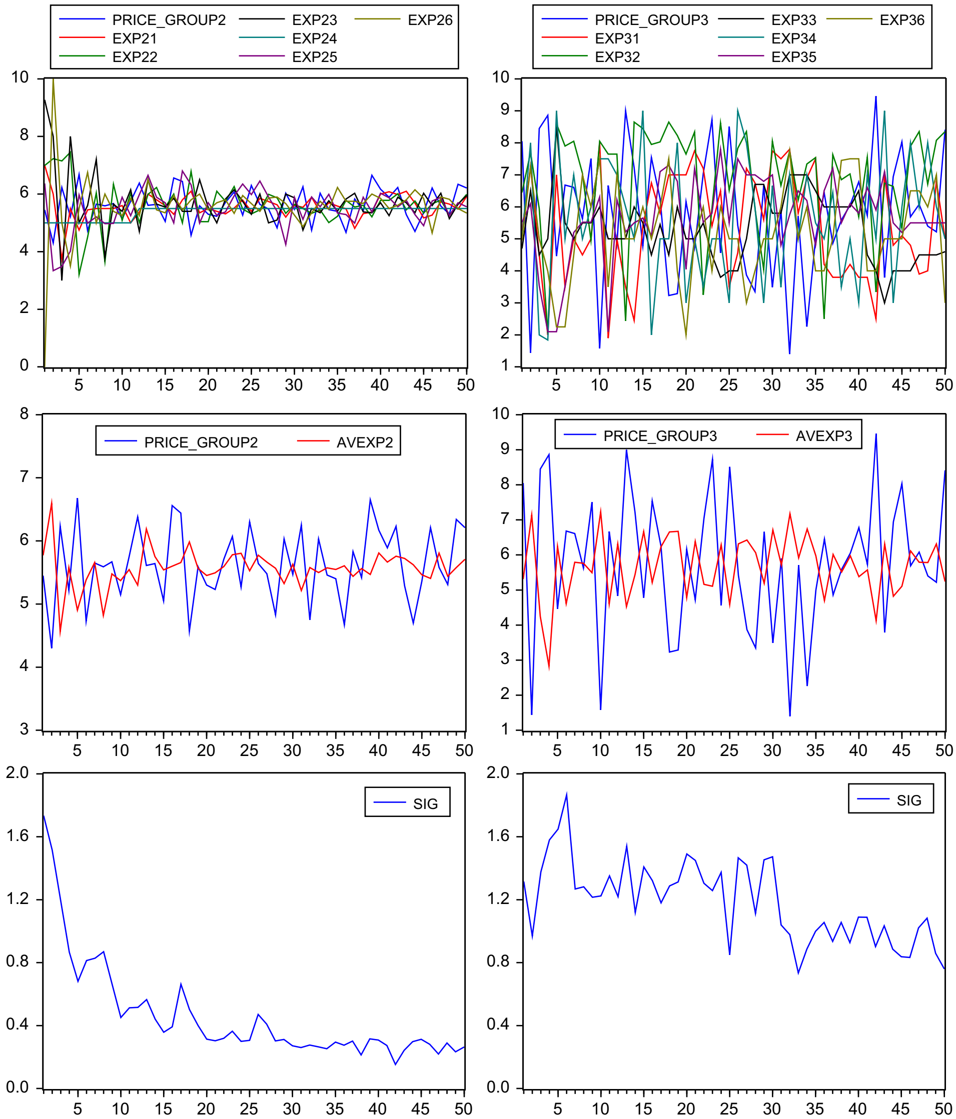

Fig. 3. Cobweb learning-to-forecast experiments, typical session in stable treatment ($\lambda = 0.22$, left panels) and strongly unstable treatment ($\lambda = 2$, right panels). Top panels: in stable treatment individual predictions coordinate on RE, while in unstable treatment heterogeneity in individual forecasting persists. Middle panels illustrate average predictions and realized prices. Bottom panels: degree of heterogeneity measured by standard deviations of individual forecasts (six individuals) averaged over the six groups in each treatment.

Fig. 3 shows time series of realized market prices together with the individual forecasts (top panels) as well as the average forecast (middle panels) for two typical experimental groups, one stable treatment (left panels) and one strongly unstable treatment (right panels). An immediate observation is that in the stable treatment, after a short learning phase of about 10 periods, price volatility is low and individual forecasts as well as average forecasts are very close to the RE benchmark, with price fluctuations entirely driven by the small random shocks in the experiments. Aggregate price behavior and individual forecasts are very different in the strongly unstable treatment. Realized prices exhibit large fluctuations, while individual forecasts are very volatile, even towards the end of the experiment. The bottom panels of Fig. 3 show the *degree of heterogeneity*, as measured by the standard deviations of individual forecasts (six individuals) averaged over the six groups, in the stable, respectively, the strongly unstable treatments. In the stable treatment heterogeneity quickly decreases over time, showing that individuals coordinate on a forecast close to the RE benchmark steady state. In the strongly unstable treatment heterogeneity decreases somewhat over time, but only slowly, and remains at least three times as high as in the stable treatment. Hence, in the classical cobweb framework used in Muth (1961) to introduce rational expectations, our LtFEs show that *only* in the stable cobweb case, the interaction of individual forecasting rules enforces convergence to the RE benchmark. In the unstable treatment, *heterogeneity in forecasting is persistent and leads to an aggregate effect upon prices characterized by excess volatility*.

The behavior in Fig. 3 is typical for all cobweb experiments. Hommes et al. (2007) summarize the *stylized facts* of the cobweb LtFEs as follows: (1) the sample mean of realized prices is close to the RE benchmark p^* in all treatments; (2) the sample variance of realized prices depends on the treatment: it is close to the RE benchmark in the stable treatment, but significantly higher in the unstable treatment; (3) realized market prices are irregular and do not exhibit significant linear autocorrelations.

These stylized facts across different treatments appear hard to explain by standard learning mechanisms offered by the theoretical literature. For example, naive expectations are inconsistent with the experiments, because in the unstable treatment it predicts too much regularity (convergence to a two-cycle) in aggregate price behavior. Average price expectations, which is just the simplest form of *adaptive learning* obtained when regressing prices on a constant, also are inconsistent with the experiments, because for both treatment it predicts convergence to the RE benchmark (see Fig. 2, right panel). Hommes (2009) discusses a number of other homogeneous learning algorithms and concludes that *heterogeneity* in forecasting rules is needed to explain the stylized facts of the cobweb experiments across different treatments. Apparently, the interaction of agents' individual forecasting rules washes out linear predictability in aggregate price behavior. In the stable treatment, this interaction leads to coordination on the "correct" RE benchmark steady state, but in the unstable treatment heterogeneity persists and prices are excessively volatile.

Hommes and Lux (2009) present a model of heterogenous individual learning via genetic algorithms (GAs) to explain the cobweb LtFEs.[3] Genetic algorithms require a functional specification of the forecasting rule, whose fitness-maximizing parameter values are searched for via the evolutionary algorithm. Hommes and Lux (2009) use a simple first order autoregressive rule:

$$p^e_{i,t+1} = \alpha_i + \beta_i(p_t - \alpha_i). \tag{3.7}$$

Such a first order autoregressive (AR1) rule seems a natural forecasting scheme as agents could implement it using the sample average as their estimate of α_i and the first order sample autocorrelation as the estimate of β_i.[4] Moreover, the AR1 forecasting rule (3.7) has a simple *behavioral* interpretation, with α_i representing an anchor or observed average price level around which the market price fluctuates, and β_i representing the observed persistence or anti-persistence of price fluctuations.[5]

Hommes and Lux (2009) show that the interaction of individual GA-learning rules simultaneously reproduces all stylized facts in aggregate price behavior observed in the experiments across the different treatments. Fig. 4 shows a typical price time series under GA-learning as well as time series of the two average parameters in the AR1 forecasting rule for the stable treatment (left panel) and the strongly unstable treatment (right panel). In the stable treatment the parameters converge to a neighborhood of the RE benchmark, consistent with the observed coordination of individual forecasts in the experiments, while in the strongly unstable treatment parameters continue to fluctuate and prices keep moving away from the RE benchmark, consistent with the persistent heterogeneity in the strongly unstable treatment of the experiments (cf. Fig. 3).

4. Asset pricing experiment

Before discussing the asset pricing learning-to-forecast experiments (LtFEs) in Hommes et al. (2005), we briefly discuss the underlying benchmark model.

4.1. An asset pricing model with heterogeneous beliefs

This subsection discusses a standard one-period asset pricing model, extended with heterogeneous beliefs, as in Campbell et al. (1997) and Brock and Hommes (1998). Agents can either invest in a risk free or in a risky asset. The risk free asset is in perfect elastic supply and pays a fixed rate of return r; the risky asset pays an uncertain dividend. Let p_t be the price per share (ex-dividend) of the risky asset at time t, and let y_t be the stochastic dividend process of the risky asset. Next period's wealth is given by

$$W_{t+1} = RW_t + (p_{t+1} + y_{t+1} - Rp_t)z_t, \tag{4.1}$$

where $R = 1+r$ is the gross rate of risk free return and z_t denotes the number of shares of the risky asset purchased at date t. Let E_{ht} and V_{ht} denote the "beliefs" or forecasts of trader type h about conditional expectation and conditional variance. Agents are assumed to be myopic mean-variance maximizers so that the demand z_{ht} of type h for the risky asset solves

$$\mathrm{Max}_{z_t}\left\{E_{ht}[W_{t+1}] - \frac{a}{2}V_{ht}[W_{t+1}]\right\}, \tag{4.2}$$

[3] Colucci and Valori (2006) fit simple adaptive learning rules to the individual cobweb LtFE of Hommes et al. (2000). Arifovic (1994) calibrates a GA-learning model to the cobweb experiments of Wellford (1989).

[4] See Hommes and Sorger (1998), where the parameters of an AR1 rule are updated according to sample autocorrelation learning.

[5] In similar cobweb LtFE experiments Heemeijer et al. (2009) recently estimated individual forecasting rules, and many individuals actually used forecasting rules of the simple AR1-form (3.7).

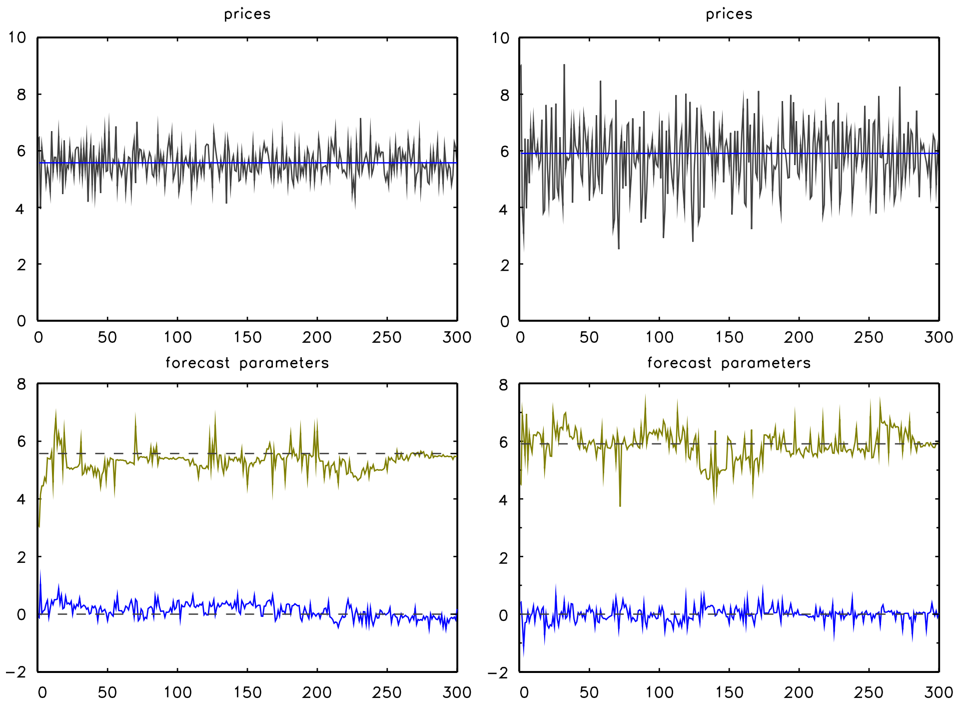

Fig. 4. Simulated prices (top panels) and average parameters α and β of the AR1 forecasting rules (3.7) under GA-learning (bottom panels). In the stable treatment (left panels) price volatility is low and parameters remain close to the RE benchmark ($\alpha = 5.57$ and $\beta = 0$), while in the unstable treatment (right panels) price volatility is high and GA-learning does not converge to the RE benchmark ($\alpha = 5.91$ and $\beta = 0$).

where a is the risk aversion parameter. The demand z_{ht} for risky assets by trader type h is then

$$z_{ht} = \frac{E_{ht}[p_{t+1}+y_{t+1}-Rp_t]}{aV_{ht}[p_{t+1}+y_{t+1}-Rp_t]} = \frac{E_{ht}[p_{t+1}+y_{t+1}-Rp_t]}{a\sigma^2}, \tag{4.3}$$

where the conditional variance $V_{ht} = \sigma^2$ is assumed to be equal for all types and constant.[6] Let z^s denote the supply of outside risky shares per investor, assumed to be constant, and let n_{ht} denote the fraction of type h at date t. Equilibrium of demand and supply yields

$$\sum_{h=1}^{H} n_{ht} \frac{E_{ht}[p_{t+1}+y_{t+1}-Rp_t]}{a\sigma^2} = z^s, \tag{4.4}$$

where H is the number of different trader types. The forecasts $E_{ht}[p_{t+1}+y_{t+1}]$ of tomorrows prices and dividends are made *before* the equilibrium price p_t has been revealed by the market and therefore will depend upon a publically available information set $I_{t-1} = \{p_{t-1}, p_{t-2}, \ldots; y_{t-1}, y_{t-2}, \ldots\}$ of past prices and dividends. Solving the heterogeneous market clearing equation for the equilibrium price gives

$$p_t = \frac{1}{1+r} \sum_{h=1}^{H} n_{ht} E_{ht}[p_{t+1}+y_{t+1}] - a\sigma^2 z^s. \tag{4.5}$$

The quantity $a\sigma^2 z^s$ may be interpreted as a *risk premium* for traders to hold all risky assets. In the experiments discussed below, z^s=0, so that (4.5) corresponds to the case of risk-neutral investors. Moreover, it will be assumed that dividends are IID, with mean $\bar{y}$, and that all traders have correct expectations about dividends, $E_{ht}[y_{t+1}] = E_t[y_{t+1}] = \bar{y}$, so that (4.5) simplifies to

$$p_t = \frac{1}{1+r} \sum_{h=1}^{H} n_{ht} E_{ht}[p_{t+1}] + \bar{y}. \tag{4.6}$$

[6] Gaunersdorfer (2000) studies the case with time-varying beliefs about variances and Chiarella and He (2002,2003) investigate the model with heterogeneous risk aversion coefficients.

4.2. Experimental design

In the asset pricing LtFEs six subjects are forecast advisors to large pension funds. Subjects only task is to forecast the price of a risky asset for 50 periods and, based on this forecast, the pension fund then decides how much to invest in the risky asset according to the mean-variance demand function (4.3). The realized asset price in the experiment is given as

$$p_t = \frac{1}{1+r}((1-n_t)\overline{p}^e_{t+1} + n_t p^f + \overline{y} + \varepsilon_t), \tag{4.7}$$

where $p^f = \overline{y}/r$ is the fundamental price, $\overline{p}^e_{t+1}$ is the average two-period ahead price forecast over six individuals and ε_t are small shocks, e.g. representing small random fluctuations in asset demand.[7] The mean-dividend $\overline{y}$ and the interest rate r are common knowledge, so that subjects can in principle use these to compute the fundamental price and use it in their forecast. The fraction n_t in (4.7) is the share of computerized fundamental robot traders, given by

$$n_t = 1-\exp\left(-\frac{1}{200}|p_{t-1}-p^f|\right). \tag{4.8}$$

The fraction of robot traders increases as the price moves further away from the fundamental benchmark. The fundamental trader thus acts as a ''far from equilibrium'' stabilizing force in the market, mimicking the feature that more traders in the market expect the price to return in the direction of the fundamental when the price deviation becomes large.[8] Subjects' earnings depend on forecasting performance and are given by a quadratic scoring rule

$$e_{i,t} = \begin{cases} 1-\left(\dfrac{p_t-p^e_{i,t}}{7}\right)^2 & \text{if } |p_t-p^e_{i,t}|<7, \\ 0 & \text{otherwise,} \end{cases} \tag{4.9}$$

so that forecasting errors exceeding 7 would result in no reward at a given period. At the end of the session the accumulated earnings of every participant were converted to euros (1 point computed as in (4.9) corresponded to 50 cents).

4.3. Benchmark simulations

Fig. 5 shows simulations of realized prices for a number of benchmark homogeneous expectations rules. When all individuals use the rational, fundamental forecasting rule, $p^e_{i,t+1} = \overline{y}/r = p^f$, for all i and t, the realized price $p_t = p^f + \varepsilon_t/(1+r)$ randomly fluctuates around the fundamental level p^f=60 with small amplitude, due to the small shocks. In the experiment, one cannot expect rational behavior at the outset, but aggregate prices might converge to their fundamental value through individual learning. Under naive expectations the price slowly converges towards its fundamental value. The same is true under average expectations, the simplest form of adaptive learning obtained when regressing the price on a constant, but the convergence is extremely slow. If all subjects would use the simple AR2 rule

$$p^e_{t+1} = \frac{60+p_{t-1}}{2} + (p_{t-1}-p_{t-2}), \tag{4.10}$$

price oscillations arise, as illustrated in the bottom right panel of Fig. 5. This is an example of an *anchoring and adjustment* rule, which plays an important role in psychology (Tversky and Kahneman, 1974), because it extrapolates the last price change from a reference point or *anchor* $(p^f+p_{t-1})/2$ describing the ''long-run'' price level.[9]

4.4. Experimental results

Fig. 6 shows time series of prices, individual predictions and forecasting errors in six different markets of the experiment. A striking feature of aggregate price behavior is that three different qualitative patterns emerge. The prices in groups 2 and 5 converge slowly and almost monotonically to the fundamental price level. In groups 1 and 6 persistent oscillations are observed during the entire experiment. In groups 4 and 7 prices are also fluctuating but their amplitude is decreasing.[10]

A second striking result is that, in all groups participants were able to *coordinate* their individual forecasting activity. As shown in the lower parts of the panels in Fig. 7, individual forecasts are dispersed in the first periods but then become very close to each other in all groups. The coordination of individual forecasts has been achieved in the absence of any

[7] Bottazzi et al. (forthcoming) consider asset pricing LtFEs where, in addition to a price forecast subjects must also forecast the variance of excess returns, which is then used in the mean-variance demand function (4.3) to compute the market clearing price.

[8] In the experiment n_t never exceeds 0.25, while the weight of the other traders are equal to $(1-n_t)/6$. Hommes et al. (2008) investigate price behavior in asset pricing LtFEs *without* robot traders.

[9] At this stage one could argue that the anchor of this rule, defined as the average between the last observed price and the fundamental price 60, was unknown in the experiment, since subjects were not provided explicitly with the fundamental price. However, for a number of subjects the linear estimated forecasting rule was surprisingly close to the anchor and adjustment rule (4.10).

[10] Price dynamics in group 3 is more difficult to classify. Similar to group 1 it started with moderate oscillations, then stabilized at a level below the fundamental, suddenly falling in period t=40, probably due to a typing error of one of the participants.

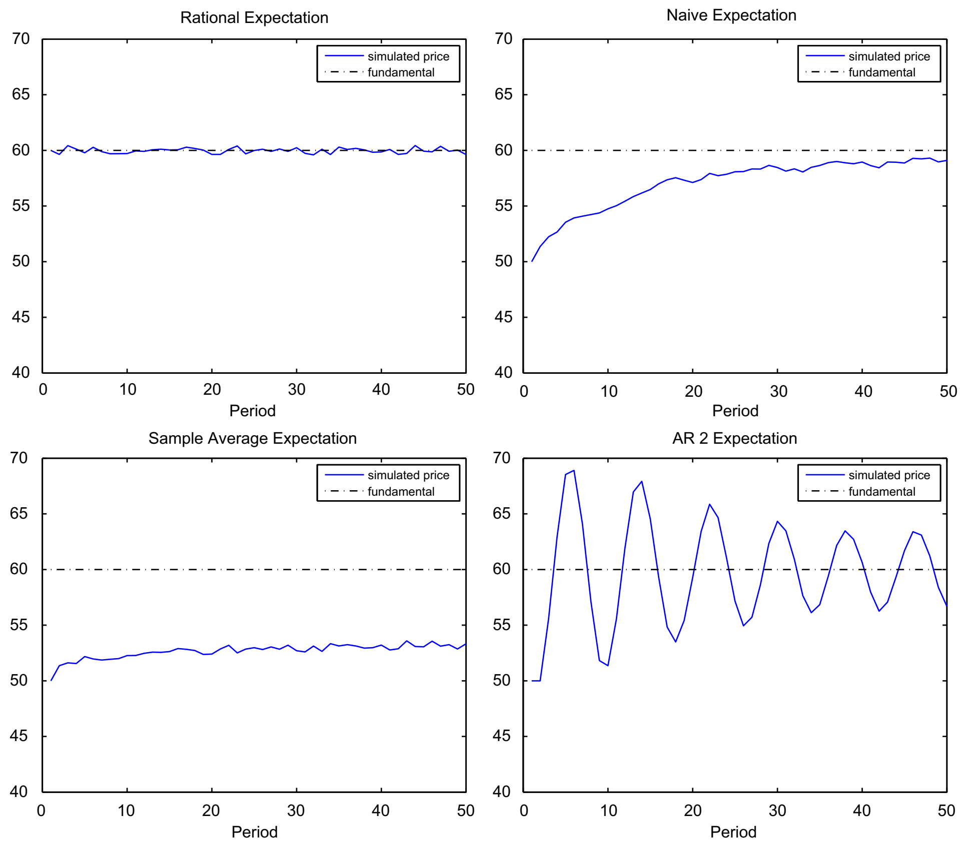

Fig. 5. Homogeneous expectations benchmarks. Top panels RE (left) and naive expectations (right). Bottom panels: average expectations (left) ad AR2 trend-following rule $p^e_{t+1} = (60 + p_{t-1})/2 + (p_{t-1} - p_{t-2})$.

communication between subjects other than through the observed realized price, and *without* any knowledge of the predictions of *other* participants.

Fig. 7 (right panel) illustrates the *degree of heterogeneity*, as measured by the standard deviation of individual forecasts, in three different groups. An immediate observation is that for each of the three groups, there is considerable *time variation* in the degree of heterogeneity. In the converging group 2 heterogeneity quickly decreases, to a level below 1 after period 15 and close to 0 after period 25. It should be noted that despite the fact that coordination is quick, prices are *not* close to the fundamental value. Hence, coordination on the "wrong", non-fundamental price occurs. In the oscillating group 1, heterogeneity fluctuates, with stronger coordination (i.e. a smaller degree of heterogeneity) during trends and weaker coordination during trend reversals. The same features, but in a more extreme form, arise in the dampened oscillation group 4. During the strong trend from periods 4–13 coordination is very strong, with the degree of heterogeneity falling from an initial level above 150 to values less than 5. Thereafter, coordination weakens and the degree of heterogeneity peaks at price trend reversals, becomes very high after period 23 (note the scale on the vertical axis) with an extremely high peak around periods 36–37. Note that, as heterogeneity increases, the asset price stabilizes and as a result coordination becomes stronger again.

To summarize, in the asset pricing LtFEs we observe the following stylized facts:

1. Participants were unable to learn the rational, fundamental price; only in some cases individual predictions moved (slowly) in the direction of the fundamental price towards the end of the experiment.
2. Although the sessions were designed in exactly the same way, three different price patterns were observed: (i) slow, (almost) monotonic convergence, (ii) persistent price oscillations with almost constant amplitude, and (iii) large initial oscillations dampening slowly towards the end of the experiment.
3. Already after a short transient, participants were able to coordinate their forecasting activity, submitting similar forecasts in every period.

One would like to have a model explaining all these stylized facts simultaneously. We have not been able to come up wit a homogeneous expectations model fitting all these experiments. The fact that qualitatively different aggregate outcomes can arise suggests that *path dependence* and *heterogeneous expectations* play a key role.

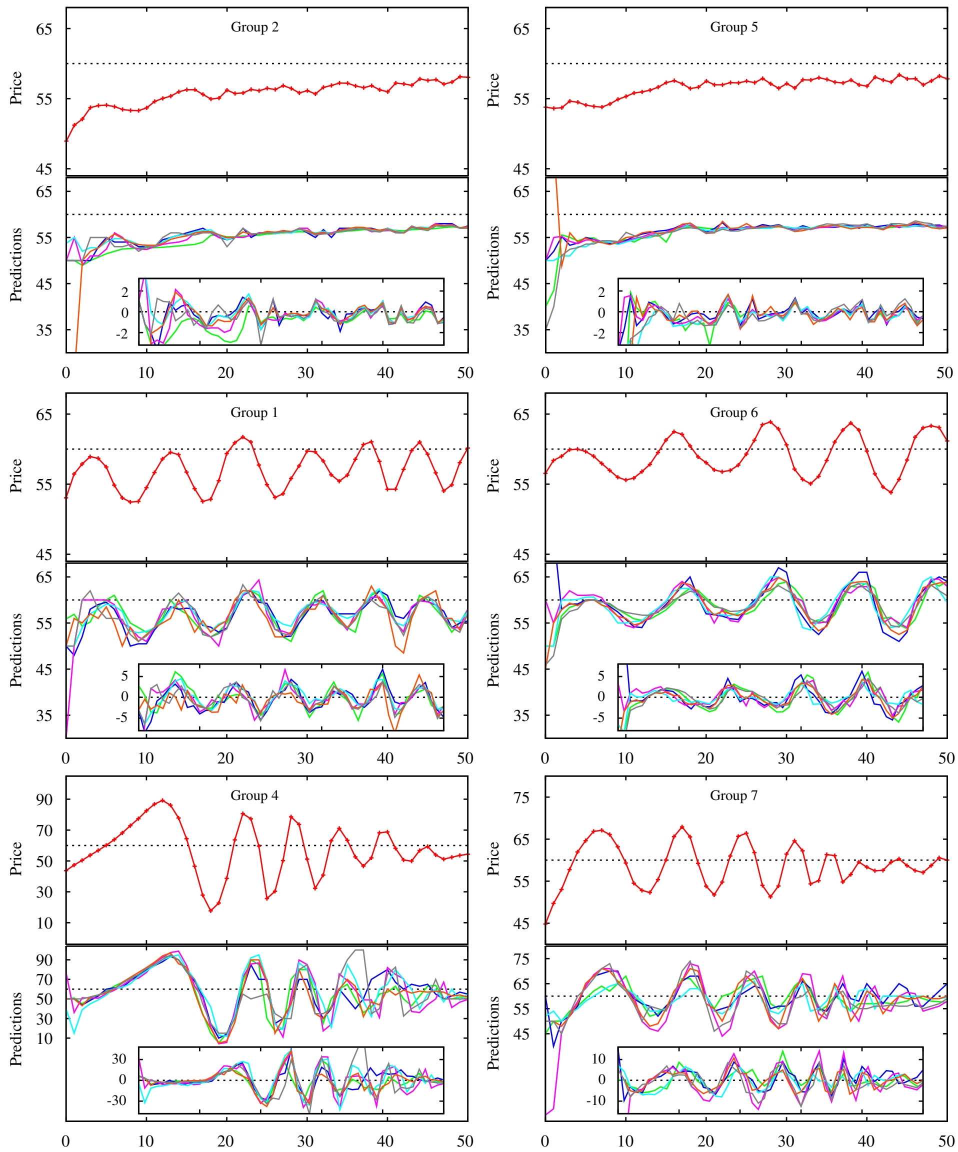

Fig. 6. Asset pricing LtFEs in six different groups. The top part of each panel shows realized market prices, the middle part shows six individual predictions and the bottom part of each panel shows the individual forecast errors. Three different patterns are observed: monotonic convergence (top panels), permanent oscillations (middle panels) and dampened oscillations (bottom panels).

5. A heterogeneous expectations model

In the last 15 years a large literature on heterogenous agent models has developed, as surveyed, e.g. by Hommes (2006) and LeBaron (2006). In particular, Brock and Hommes (1997) introduced a heterogeneous expectations model, where agents tend to switch towards forecasting strategies that have performed better in the recent past. Here we discuss a modified version, a heuristics switching model, which has recently been fitted to the asset pricing LtFEs by Anufriev and Hommes (2009, 2012).[11]

Agents can choose from a number of simple *forecasting heuristics*. To discipline the wilderness of bounded rationality, the set of behavioral forecasting rules needs to be carefully chosen. We choose forecasting heuristics similar to those obtained from

[11] The heuristics switching model is similar to other models of reinforcement learning, e.g. Erev and Roth (1998) and Camerer and Ho (1999). An important difference, however, is that our model is built in a market environment rather than the strategic environments usually studied in standard game theory. Another related work is Arifovic and Ledyard (2008), who present a new behavioral model of individual learning in repeated situations and validate the model with experimental data. Schunk (2009) introduces a dynamic model of behavioral heterogeneity in search behavior and shows that his experimental data is well explained by a model assuming dynamic updating of utility reference points.

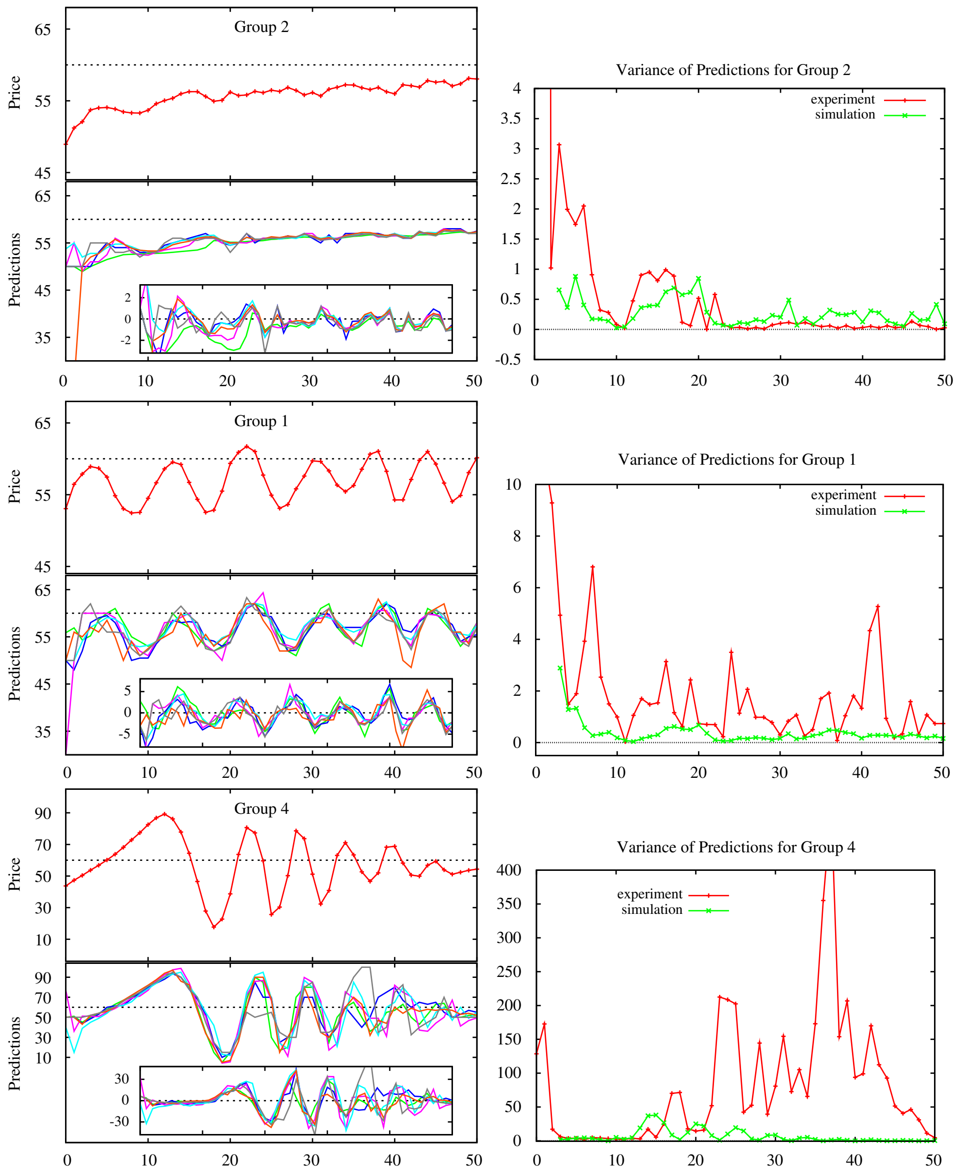

Fig. 7. Left panels: prices for laboratory experiments in three different groups. Lower parts of left panels show individual predictions and forecasting errors (inner frames). Right panels: evolution of the degree of heterogeneity as measured by the standard deviation of individual forecasts in corresponding groups.

estimations of linear models on individual forecasting data in the LtFEs in Hommes et al. (2005) and Heemeijer et al. (2009). To further discipline the wilderness of bounded rationality two forms of individual learning are introduced. First, for some heuristics *adaptive learning* takes place, that is, some parameters of the heuristics are updated over time. Second, *evolutionary selection* or *performance based reinforcement learning* takes place, that is, agents evaluate the *performances* of all heuristics, and tend to switch to more successful rules. Hence, the impact of each of the rules is evolving over time.

To keep the model as simple as possible, Anufriev and Hommes (2009,2012) restricted attention to only four forecasting heuristics:

$$\text{ADA} \quad p^e_{1,t+1} = 0.65p_{t-1} + 0.35p^e_{1,t}, \tag{5.1}$$

$$\text{WTR} \quad p^e_{2,t+1} = p_{t-1} + 0.4(p_{t-1} - p_{t-2}), \tag{5.2}$$

$$\text{STR} \quad p^e_{3,t+1} = p_{t-1} + 1.3(p_{t-1} - p_{t-2}), \tag{5.3}$$

$$\text{LAA} \quad p^e_{4,t+1} = \frac{p^{av}_{t-1} + p_{t-1}}{2} + (p_{t-1} - p_{t-2}), \tag{5.4}$$

where $p^{av}_{t-1} = (\sum_{j=0}^{t-1} p_j)/t$ is the sample average of past prices. The first *adaptive expectations* (ADA) rule predicts that the price is a weighted average of the last observed price p_{t-1} and the last price forecast p^e_t. This ADA rule was obtained as the estimated linear rule of a number of subjects in the converging groups 2 and 5. The second and the third rules are both *trend-following rules*, with a weak trend (WTR) parameter 0.4 and a strong trend (STR) parameter 1.3, respectively.[12] These rules were obtained as the estimated linear rules for quite a number of subjects in the oscillatory markets 1, 4, 6 and 7, with 0.4 and 1.3 obtained as the smallest and largest trend extrapolating coefficients. Finally, the fourth rule is an anchor and adjustment rule, obtained from the linear AR2 rule (4.10), discussed in Section 4.3, by replacing the (unknown) fundamental price p^f by an observable proxy given by the sample average of past prices p^{av}_{t-1}. The weight coefficient of the ADA rule and the trend parameters of trend-following rules have been fixed and it appears that the simulations below are robust with respect to small changes of these parameters. The LAA rule exhibits a simple form of adaptive learning, since the anchor of the rule is time-varying, given by the average of the last observed price and the sample average of all observed past prices.

Subjects switch between the different forecasting rules based upon quadratic forecasting errors, consistent with the earnings incentives in the experiments. The fitness or performance measure of forecasting heuristic i is given by

$$U_{i,t-1} = -(p_{t-1} - p^e_{i,t-1})^2 + \eta U_{i,t-2}, \tag{5.5}$$

where the parameter $\eta \in [0,1]$ measure the strength of the agents' *memory*. Switching is described by a discrete choice model with *asynchronous updating*

$$n_{i,t} = \delta n_{i,t-1} + (1-\delta)\frac{\exp(\beta U_{i,t-1})}{\sum_{i=1}^{4}\exp(\beta U_{i,t-1})}. \tag{5.6}$$

In the special case $\delta = 0$, (5.6) reduces to the *the discrete choice model* with synchronous updating used in Brock and Hommes (1997). The more general case, $0 \le \delta \le 1$, gives some persistence or inertia in the impact of rule h, reflecting the fact (consistent with the experimental data) that not all participants update their rule in every period or at the same time (see Hommes et al., 2005b; Diks and Weide, 2005). Hence, δ may be interpreted as the average per period fraction of individuals who stick to their previous strategy. In the extreme case $\delta = 1$, the initial impacts of the rules never change; if $0 < \delta \le 1$, in each period a fraction $1-\delta$ of participants update their rule according to the discrete choice model. The parameter $\beta \ge 0$ represents the intensity of choice measuring how sensitive individuals are to differences in strategy performance. The higher the intensity of choice β, the faster individuals will switch to more successful rules. In the extreme case $\beta = 0$, the impacts in (5.6) move to an equal distribution independent of their past performance. At the other extreme $\beta = \infty$, *all* agents who update their heuristic (i.e. a fraction $1-\delta$) switch to the most successful predictor.

In all simulations below, parameters are fixed at the benchmark values $\beta = 0.4$, $\eta = 0.7$, $\delta = 0.9$[13] and the initial fractions of the four strategies are equal, i.e. n_{h0}=0.25. The simulations thus only differ in their initial prices, which have been chosen exactly as in the first two periods in the corresponding experimental group.

Fig. 8 compares the experimental data with the one-step ahead predictions made by the heuristics switching model, for one converging group (group 5), one oscillating group (group 6) and one dampened oscillating group (group 7); the other groups yield very similar results; see Anufriev and Hommes (2009). Fig. 8 suggests that the switching model with four heuristics fits the experimental data quite nicely. The one-step ahead predictions of the nonlinear switching model in Fig. 8 use past *experimental price data* to determine the forecasts and the fractions of the strategies at each period t, i.e. the model simulation uses exactly the same information that was available to participants in the experiments. An immediate observations by comparing these simulations is that the one-period ahead forecasts of the heuristics switching model can easily adapt and track the different patterns in aggregate price behavior, slow monotonic convergence, sustained oscillations as well as dampened oscillations.

The right panels in Fig. 8 show the corresponding fractions of the four strategies for each experimental group. In different groups different heuristics dominate the market after starting from the same uniform distribution. In the monotonically converging group, the impact of the different rules stays more or less equal, although the impact of adaptive expectations gradually increases and slightly dominates the other rules in the last 25 periods. In the oscillatory group the LAA rule dominates the market and its impact gradually increases to 90% towards the end of the experiment. Finally, for the group with the dampened oscillations, the one-step ahead forecast simulation leads to a rich evolutionary selection dynamics (bottom panel), with three different phases where the STR, the LAA and the ADA heuristics subsequently dominate. The STR dominates during the initial phase of a strong trend in prices, but starts declining after it misses the first turning point of the trend. The LAA does a better job in predicting the trend reversal and its impact starts increasing, dominating the market in the second phase of the experiment, with oscillating prices between periods 15–40. But the

[12] Haruvy et al. (2007) also provide experimental evidence that individual price expectations are significantly affected by past price trends. Trend-following behavior is often associated with technical trading strategies in real financial markets; see Menkhoff and Taylor (2007) for a comprehensive overview of the importance of technical analysis.

[13] These values have been obtained in Anufriev and Hommes (forthcoming) after some trial and error simulations. The simulation results, however, are fairly robust with respect to (small) changes in these parameter values. Goldbaum and Mizrach (2008) estimated the intensity of choice parameter β in mutual fund allocation decisions.

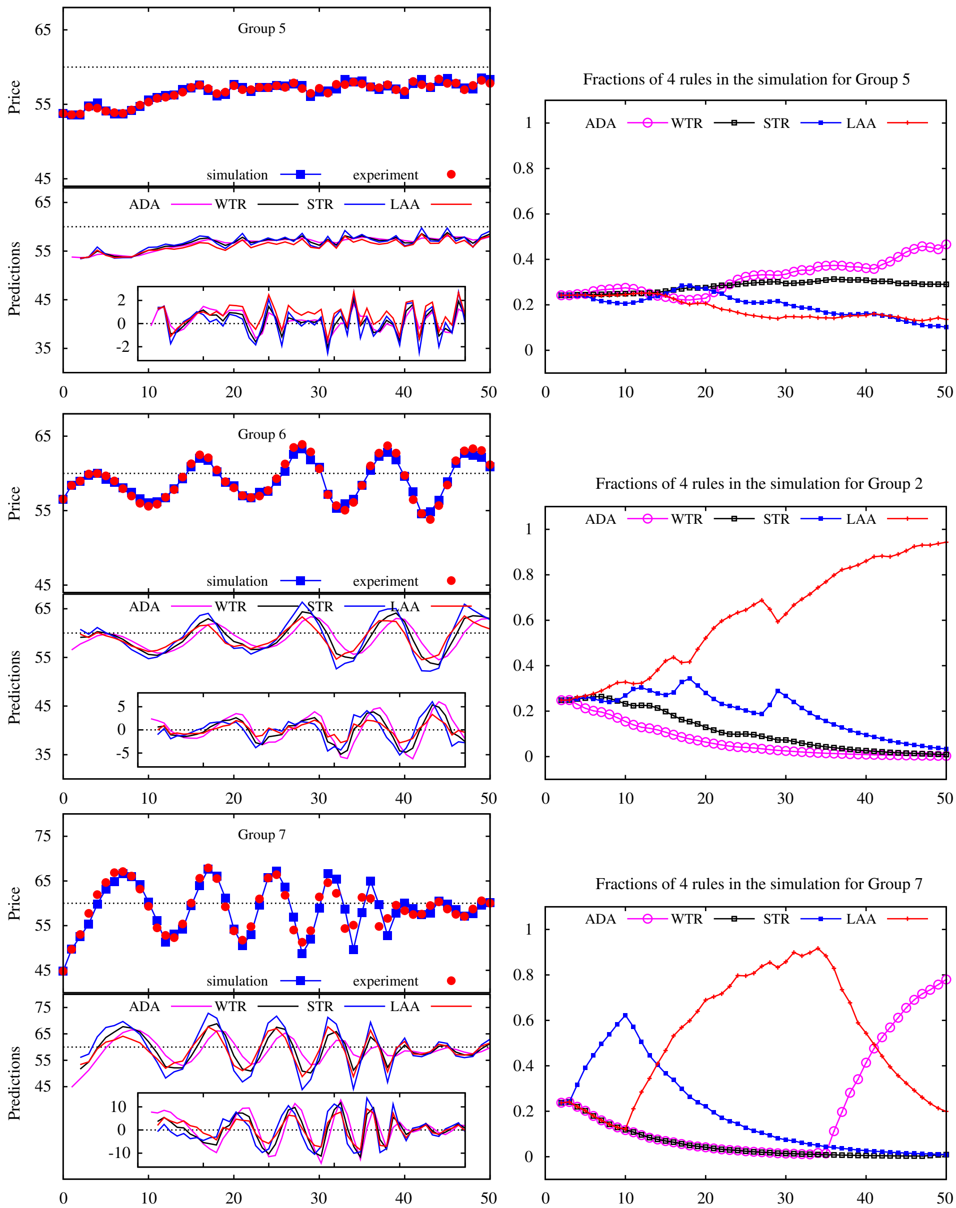

Fig. 8. Left panels: realized prices for laboratory experiments in different groups (red) with corresponding one-step ahead predictions of the heuristics switching model (blue). Lower parts of left panels show predictions and forecasting errors (inner frames) of the four heuristics. Right panels: evolution of the fractions of the four heuristics during the one-step ahead predictions of the model: adaptive expectations (ADA, purple), weak trend-followers (WTR, black), strong trend-followers (STR, blue) and anchoring adjustment heuristic (LAA, red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

oscillations slowly dampen and therefore, after period 35, the impact of adaptive expectations, which has been the worst performing rule until that point, starts increasing and adaptive expectations dominates in the last nine periods.

6. Positive versus negative feedback experiments

Aggregate price behavior in the cobweb and the asset pricing LtFEs are quite different. While in the cobweb framework the price fluctuates around its fundamental value, with a sample average of realized prices very close to the RE price, in the asset pricing experiments persistent deviations from the fundamental price with long phases of under- or over-valuations have been observed. A key difference between the cobweb and asset pricing experiments is the type of feedback: the asset pricing (cobweb) framework exhibits positive (negative) feedback, that is the realized price depends positively (negatively) on the average price forecast. In the case of positive (negative) feedback, when an individual forecast increases, the realized market price goes up (down). A natural question then is *whether the type of expectations feedback, positive versus negative, explains these differences in aggregate behavior*.

In most markets both types of feedback may play a role. Positive feedback, however, seems particularly relevant in speculative asset markets. If many agents expect the price of an asset to rise they will start buying the asset, aggregate demand will increase and so, by the law of supply and demand, will the asset price. High price expectations thus become self-confirming and lead to high realized asset prices. In markets where the role of speculative demand is less important, e.g. in markets for non-storable commodities, negative feedback may play a more prominent role. Consider, e.g. a supply driven commodity market. If many producers expect future prices to be high they will increase production which, according to the law of supply and demand, will lead to a lower realized market price.

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. In this section we discuss these LtFEs and apply the heterogeneous expectations model of Section 5 to see whether it can explain the differences in aggregate outcomes.

The distinction between positive and negative expectation feedback is related to the concepts of strategic complements versus strategic substitutes. Haltiwanger and Waldman (1985) argue that when actions are strategic complements, agents have an incentive to imitate other agents. This is the case in an asset market, where predicting a price close to the predictions of the other participants turns out to be most profitable. However, coordination of predictions enhances the impact of the irrational participants upon realized prices and convergence to the rational equilibrium price becomes unlikely. When actions are strategic substitutes, agents have an incentive to deviate from what other agents are doing. This is the case in negative feedback markets, where agents have an incentive to predict high (low) prices when the majority predicts prices below (above) the equilibrium price. The impact of irrational individuals will be limited and convergence to the equilibrium price is more likely. Coordination of predictions will only take place after convergence.

In recent experiments Fehr and Tyran (2001, 2005, 2008) study the impact of different strategic environments (strategic complementarity versus strategic substitutability) on individual rationality and aggregate outcomes. Strategic substitutability (complementarity) prevails if an increase in the action of individual i generates an incentive for j to decrease (increase) his action. Fehr and Tyran study the adjustment of nominal prices after an anticipated money shock in a price setting game with positively (complements) or negatively sloped (substitutes) reaction curves, and find much faster convergence in the case of substitutes. Sutan and Willinger (2009) investigate a new variant of beauty contest games (BCG) in which players actions are strategic substitutes versus strategic complements and find that chosen numbers are closer to rational play in the case of strategic substitutes.

In the LtFEs of Heemeijer et al. (2009), the (unknown) price generating rules in the *negative* and *positive* feedback systems were, respectively:

$$p_t = 60 - \frac{20}{21}\left[\sum_{h=1}^{6} \frac{1}{6} p^e_{ht} - 60\right] + \varepsilon_t \quad \text{negative feedback,} \tag{6.1}$$

$$p_t = 60 + \frac{20}{21}\left[\sum_{h=1}^{6} \frac{1}{6} p^e_{ht} - 60\right] + \varepsilon_t \quad \text{positive feedback,} \tag{6.2}$$

where ε_t is a random shock to the pricing rule. First we will consider positive and negative feedback systems with small IID shocks ε_t, $\varepsilon_t \sim N(0,0.25)$, and later on with large permanent shocks.

A common feature of the positive and negative feedback systems (6.1) and (6.2) is that both have a RE equilibrium steady state $p^*=60$ (when the shocks ε_t have mean 0). The *only* difference between (6.1) and (6.2) is therefore the sign of the slope of the linear map, $20/21 \approx +0.95$, resp. $-20/21 \approx -0.95$.[14]

Fig. 9 shows realized market prices as well as individual predictions in two typical groups. A striking feature is that aggregate price behavior is very different in the positive versus negative feedback cases. In the negative feedback case, the price quickly settles down to the RE steady state price 60, while in the positive feedback case, the market price oscillates slowly around its fundamental value. Individual forecasting behavior is also different for the different feedback treatments: in the case of positive feedback, coordination of individual forecasts occurs extremely quickly, within 2–3 periods. The coordination, however, is on a "wrong" non-fundamental price. In contrast, in the negative feedback case coordination of individual forecasts is slower and takes about 10 periods. More persistence in heterogeneity of individual forecasts, however, ensures that, after 10 periods, the realized market price is very close to the RE benchmark of 60.

Can the heterogeneous expectations model of Section 5 explain these different outcomes in individual and aggregate behavior? Fig. 10 shows realized market prices together with the simulated prices (left panels), and the corresponding evolution of the fractions of the four strategies (right panels) of the heuristics switching model with the same benchmark parameters as before, i.e. $\beta = 0.4$, $\eta = 0.7$, $\delta = 0.9$. The model matches aggregate price behavior in both the negative and

[14] 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.

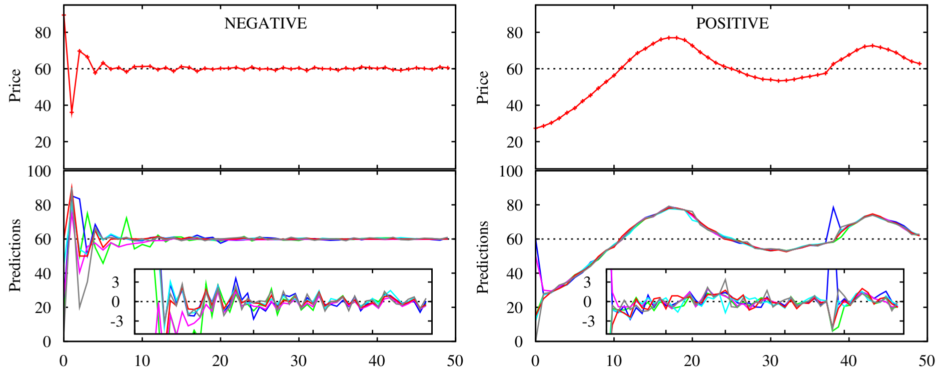

Fig. 9. Negative (left panel) versus positive (right panel) feedback learning-to-forecast experiments with small IID shocks; prices (top panels), individual predictions (bottom panels) and forecast errors (small panels).

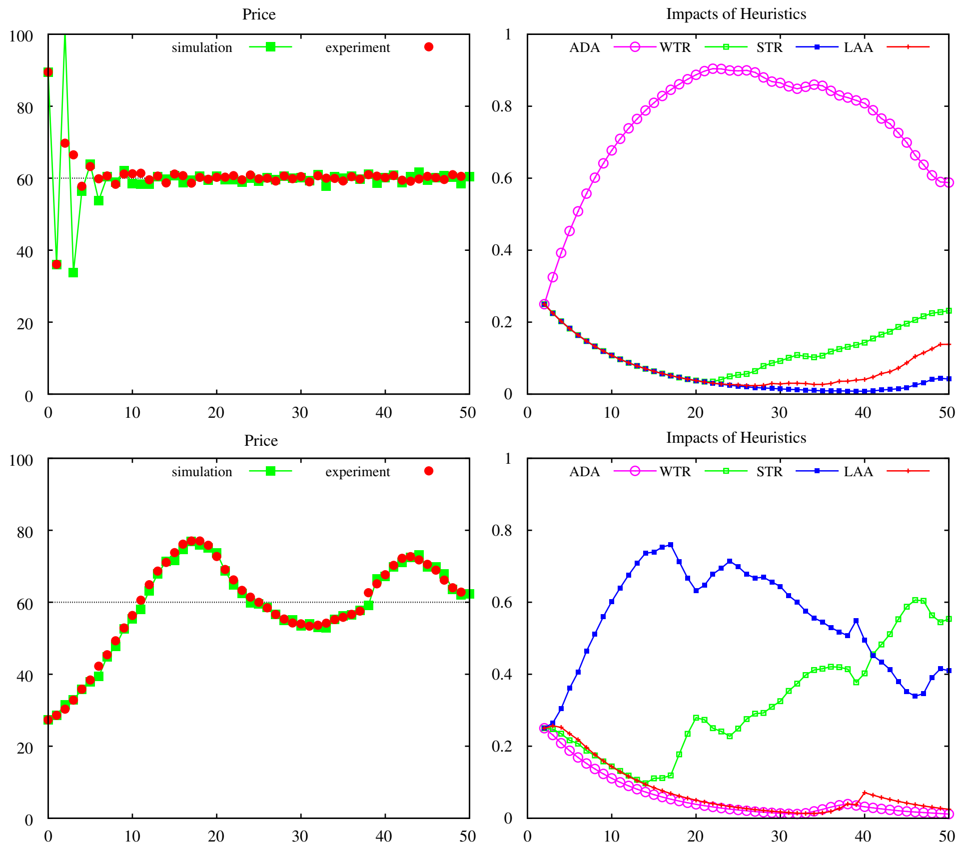

Fig. 10. Positive feedback (bottom panels) and negative feedback (top panels) markets with small shocks. Realized and simulated prices (left panels) and corresponding evolution of fractions of four strategies in heuristics switching model.

positive feedback treatment. Furthermore, the time series of the fractions of the different forecasting heuristics (Fig. 10, right panels) provide an intuitive explanation of why aggregate behavior is different. In the negative feedback treatment, the adaptive expectations strategy performs best and starts dominating quickly, capturing more than 90% of the market within 20 periods, thus enforcing convergence towards the fundamental equilibrium price. In contrast, in the positive feedback treatment the impact of the strong trend-following rule (STR) quickly increases and it captures more than 75% of the market after 15 periods. Thereafter, the impact of the STR rule gradually declines, while the fraction of weak trend-followers (WTR) gradually increases due to the fact that the STR rule makes (somewhat) larger mistakes (especially at the turning points) than the WTR-rule.

The difference in aggregate behavior is thus explained by the fact that trend-following rules are more successful in a positive feedback environment reenforcing price oscillations and persistent deviations from the fundamental equilibrium benchmark price, while the trend-following rules are outperformed and driven out by adaptive expectations in the case of negative feedback.

Bao et al. (2010) recently ran similar LtFEs with large permanent shocks ε_t to the price generating mechanisms (6.1) and (6.2). These shocks have been chosen such that, both in the negative and positive feedback treatments, the fundamental equilibrium price p^* changes over time according to

$$\begin{aligned} p_t^* &= 56, \quad 0 \le t \le 21, \\ p_t^* &= 41, \quad 22 \le t \le 43, \\ p_t^* &= 62, \quad 44 \le t \le 65. \end{aligned} \tag{6.3}$$

The purpose of these experiments is to investigate how the type of expectations feedback may affect the *speed of learning* of a new steady state equilibrium price. Fig. 11 shows realized market prices together with simulated market prices (left panels), and the evolution of the fractions of the four strategies of the heuristics switching model (right panels) for typical groups of the negative feedback (top panels) and the positive feedback treatment (bottom panels). The heuristics switching model is exactly the same as in Anufriev and Hommes (2009), in the case of the asset pricing experiments (see Section 5), with the same benchmark parameters, i.e. $\beta = 0.4$, $\eta = 0.7$, $\delta = 0.9$. As in the case of small shocks, there is a striking difference between positive and negative feedback markets. In the negative feedback market, after each large shock the price quickly (within five periods) settles down to the new RE benchmark, while in the positive feedback market the price slowly oscillates with persistent deviations from the RE benchmark. The heuristics switching model matches both patterns 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 adaptive expectations strategy quickly dominates the market (more than 50% within 10 periods) enforcing quick convergence to the RE benchmark after each large shock. In contrast, in the positive feedback treatment, trend-following strategies perform well, the weak trend-following rule dominates in the first 20 periods, while the strong trend-following rule starts dominating after the first large shock in period 22.

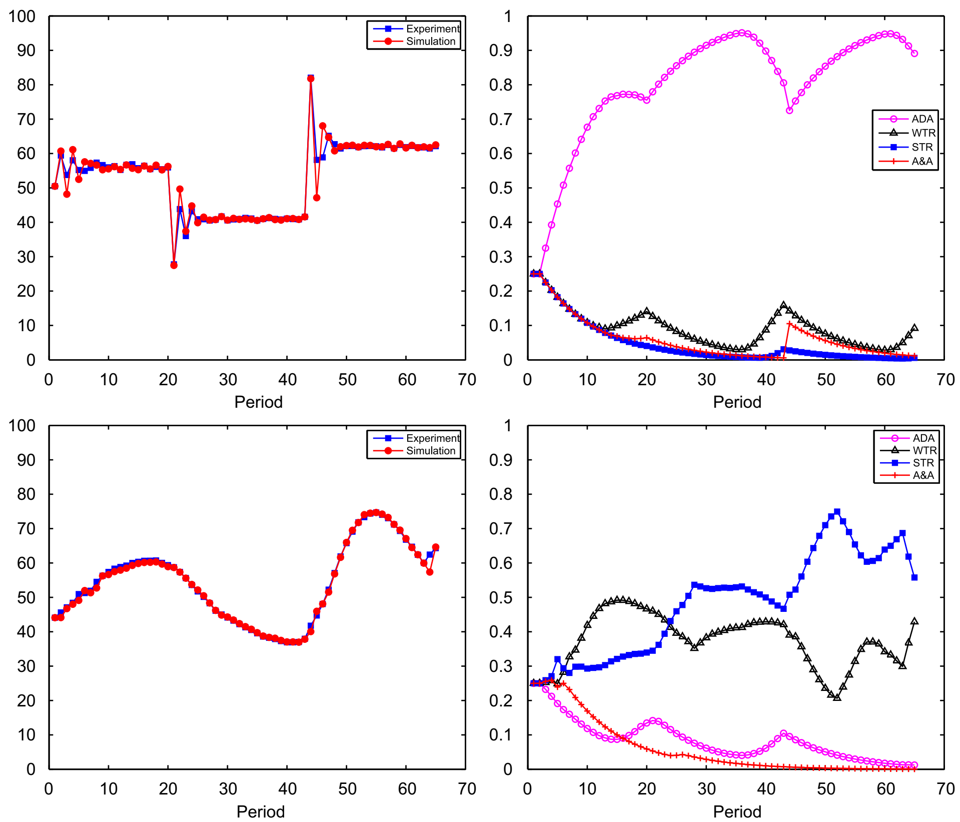

Fig. 11. Positive feedback (bottom panels) and negative feedback (top panels) markets with large permanent shocks to the equilibrium steady state level. Realized and simulated prices (left panels) and corresponding evolution of fractions of the four strategies in the heuristics switching model.

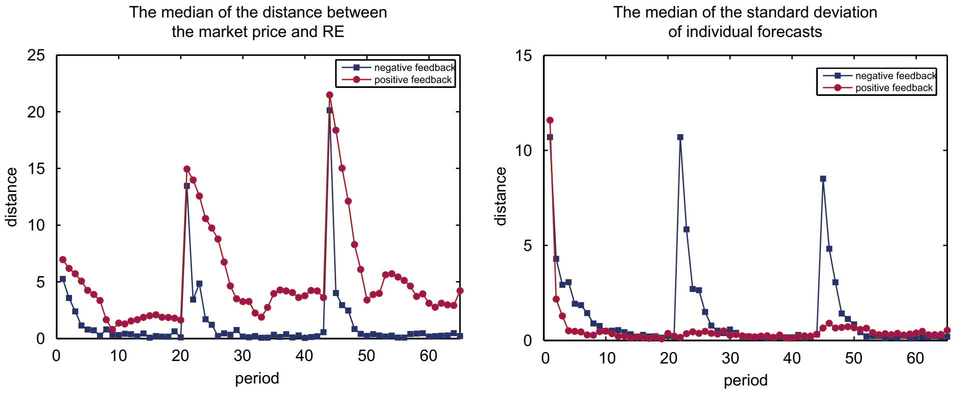

Fig. 12. Positive/negative feedback LtFEs with permanent large shocks. Left panel: median absolute distance to RE fundamental price over eight groups for both treatments. Right panel: median standard deviation of individual predictions over eight groups for both treatments.

Fig. 12 reveals some other striking features of aggregate price behavior and individual forecasts. The left 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 0, while for the positive feedback treatment after each shock the distance to the RE benchmark shows a similar spike but does not converge to 0 and only decreases slowly. The right 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, while in the negative treatment heterogeneity is more persistent for about 10 periods after each large shock. 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. On the other hand, in the negative feedback treatment coordination is slow, 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 feedback markets are characterized by slow coordination, more persistent heterogeneity and quick price discovery. See Bao et al. (2010) for a more detailed discussion.

7. A New Keynesian macro-model

As a final example, we briefly discuss the LtFEs in a standard New Keynesian macro-model of Assenza et al. (2010); see also Pfajfar and Zakelj (2010) for closely related LtFEs. The New Keynesian model is described by the following equations:

$$y_t = y^e_{t+1} - \varphi(i_t - \pi^e_{t+1}) + g_t, \tag{7.1}$$

$$\pi_t = \lambda y_t + \beta \pi^e_{t+1} + u_t, \tag{7.2}$$

$$i_t = \overline{\pi} + \phi_\pi(\pi_t - \overline{\pi}), \tag{7.3}$$

where y_t and y^e_{t+1} are, respectively, the actual and the expected output gap i_t is the nominal interest rate, π_t and π^e_{t+1} are, respectively, the actual and expected inflation rates, $\overline{\pi}$ is the target inflation rate, φ, λ, β and ϕ_π are positive coefficients, and g_t and u_t are (small) white noise shocks. The coefficient ϕ_π measures the response of the nominal interest rate i_t to changes in the inflation rate π_t. Eq. (7.1) is the IS curve in which the actual output gap y_t depends on the expected output gap y^e_{t+1} and on the real interest rate $i_t - \pi^e_{t+1}$. Eq. (7.2) is the expectations-augmented New Keynesian Phillips curve according to which actual inflation depends on the actual output gap and expected inflation. Finally, Eq. (7.3) is the monetary policy rule implemented by the Central Bank in order to keep inflation at its target level $\overline{\pi}$.

In the LtFEs of Assenza et al. (2010), two different groups of six subjects have to provide two-period ahead forecasts of the inflation rate, respectively, the output gap for 50 periods. Realized inflation and realized output gap are determined by the (average) individual expectations of two different groups of 6 individuals. Subjects only obtain qualitative information about the macroeconomy, but they do not know the underlying law of motion (7.1)–(7.3).

Fig. 13 shows time series of realized inflation, the output gap and the interest rate, together with individual forecasts of inflation and the output gap. Both inflation and the output gap exhibit dampened oscillations eventually converging to the RE benchmark steady state. An interesting feature of these experiments is that an aggressive monetary policy described by a Taylor type interest rate rule that adjust the interest rate more than one point for one in response to inflation (with a coefficient $\phi_\pi = 1.5$) is able to stabilize heterogeneous expectations.

Fig. 14 shows simulated time series of inflation (right panel) and output gap (left panel), together with the fractions of the forecasting rules (bottom panels) of the heuristics switching model of Anufriev and Hommes (2009, forthcoming) with

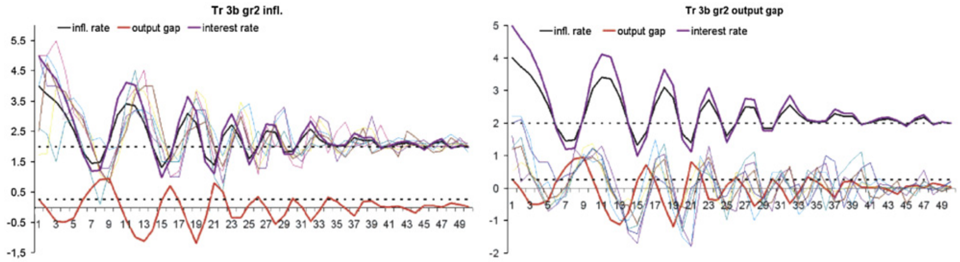

Fig. 13. Learning-to-forecast experiment in a New Keynesian macro-framework. Left panel: time series of realized inflation, individual inflation forecasts, interest rate and realized output gap. Right panel: time series of realized inflation, interest rate, realized output gap and individual forecasts of output gap. Parameter values: $\varphi=1$, $\lambda=0.3$, $\beta=0.99$, $\overline{\pi}=2$ and $\phi_\pi=1.5$.

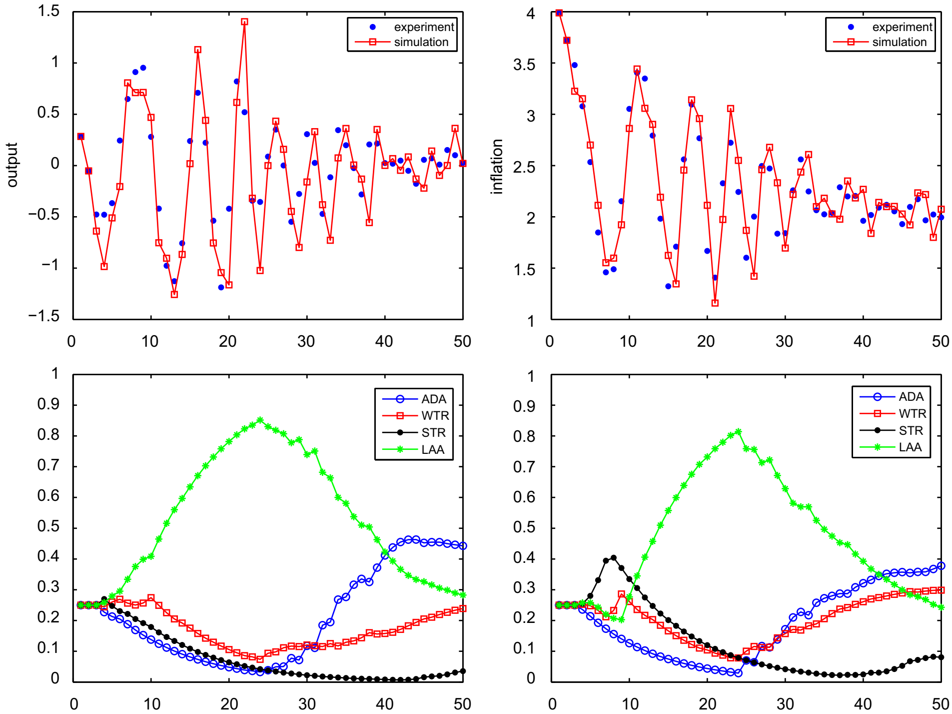

Fig. 14. Simulations of heuristics switching model in the New Keynesian macro-framework. Left panels: output (top) and fractions of output expectations (bottom). Right panels: inflation (top) and fractions of inflation expectations (bottom). Learning parameters: $\beta=0.4$, $\eta=0.7$ and $\delta=0.9$.

the benchmark parameters as before. The same heuristics switching model is used for both inflation and output gap forecasts and the model fits the experimental data quite well. The patterns of the weight of the forecasting heuristics are quite similar for inflation and output, as one would expect since both time series are qualitatively similar. The Learning Anchor and Adjustment (LAA) rule dominates most of the time, with a peak of about 80% after 25 periods and the adaptive expectations rule dominating in the last 10 periods when the economy stabilizes. There is a slight difference in the initial phase, with the strong trend rule STR dominating inflation forecasting in periods 5–10, picking up the stronger trend in inflation.

8. Concluding remarks

Learning-to-Forecast Experiments (LtFEs) can be used to test different theories of expectations and learning in standard model settings. LtFEs are tailor-made experiments to test the expectations hypothesis, with all other model assumptions computerized and under control of the experimenter. Different types of aggregate behavior have been observed in different market settings. To our best knowledge, *no homogeneous expectations model* fits the experimental data across different market settings. Quick convergence to the RE benchmark *only* occurs in stable (i.e. stable under naive expectations) cobweb markets with negative expectations feedback, as in Muth's (1961) seminal rational expectations paper. In all other market settings persistent deviations from the RE fundamental benchmark seem to be the rule rather than the exception.

The laboratory experiments suggest that *heterogeneity* is a crucial aspect of a theory of expectations for at least two reasons. First, only a *heterogeneous expectations model* can explain the *path dependence* and the different aggregate outcomes in the same market environment, as observed for example in the asset pricing LtFEs. Second, a single model of heterogeneous expectations can explain different aggregate outcomes across different market settings. Indeed a simple heuristics switching model based on evolutionary selection and reinforcement learning provides an intuitive theory of individual learning and an explanation of the different observed types of aggregate price behavior. In positive feedback markets a trend following strategy does relatively well and reenforces price oscillations and persistent deviations from the RE benchmark. In contrast, when negative expectations feedback dominates, trend-following strategies perform poorly and are driven out by adaptive expectations which enforces stable price behavior.[15]

An important challenge to a research program in behavioral economics and finance based on bounded rationality is to come up with a plausible and general theory of heterogeneous expectations. The fact that the same simple heuristics switching model (with fixed parameters) fits experimental data in different economic environments suggests that a general heterogeneous expectations hypothesis may explain individual expectations and the aggregate behavior their interaction co-creates across different market settings. More research in this area is called for to test the robustness of these findings.

Acknowledgments

This work was presented as a plenary lecture at the 15th International Conference on Computing in Economics and Finance (CEF), Sydney, July 15–17, 2009 and, subsequently at the Econophysics Colloquium, Erice, October 25–31, 2009 and the International Research Forum on Monetary Policy, Federal Reserve Board, Washington, March 26–27, 2010. I would like to thank the participants of these conferences, in particular Jasmina Arifovic, for stimulating discussions and comments. The survey is based on longlasting cooperation and benefitted from many ideas of my coauthors Joep Sonnemans, Jan Tuinstra, Henk van de Velden, Peter Heemeijer, Buz Brock, Mikhail Anufriev, Thomas Lux, Tiziana Assenza, Domenico Massaro and Te Bao. Special thanks go to Mikhail Anufriev, Te Bao, Domenico Massaro and Thomas Lux for providing some of the simulations and figures. Detailed comments of an anonymous referee on an earlier draft have been extremely helpful. Financial support from the EU 7th framework collaborative project "*Monetary, Fiscal and Structural Policies with Heterogeneous Agents (POLHIA)*", Grant no. 225408, is gratefully acknowledged. None of the above can be blamed for errors in this paper.

References

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

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

Allen, H., Taylor, M.P., 1990. Charts, noise and fundamentals in the London foreign exchange market. The Economic Journal 100 (400), 49–59 Conference Papers.

Amilon, H., 2008. Estimation of an adaptive stock market model with heterogeneous agents. Journal of Empirical Finance 15, 342–362.

Anderson, E.W., Ghysels, E., Juergens, J.L., 2005. Do heterogeneous beliefs matter for asset pricing? Review of Financial Studies 18, 875–924

Anufriev, M., Dindo, P., 2010. Wealth-driven selection in a financial market with heterogeneous agents. Journal of Economic Behavior & Organization 73, 327–358.

Anufriev, M., Panchenko, V., 2009. Asset prices, traders' behavior and market design. Journal of Economic Dynamics and Control 33, 1073–1090.

Anufriev, M., Hommes, C.H., 2009. Evolutionary selection of individual expectations and aggregate outcomes. CeNDEF Working Paper 09-09, University of Amsterdam.

Anufriev, M., Hommes, C.H., forthcoming. Evolution of market heuristics. Knowledge Engineering Review.

Arifovic, J., 1994. Genetic algorithm learning and the cobweb model. Journal of Economic Dynamics and Control 18 (1), 3–28.

Arifovic, J., Ledyard, J., 2008. A behavioral model for mechanism design: individual evolutionary learning. Working Paper Simon Fraser University, January.

Assenza, T., Berardi, M., 2009. Learning in a credit economy. Journal of Economic Dynamics and Control 33, 1159–1169.

Assenza, T., Heemeijer, P., Hommes, C.H., Massaro, D., 2010. Individual expectations and aggregate macro behavior. CeNDEF Working Paper, University of Amsterdam, July.

Baak, S.J., 1999. Tests for bounded rationality with a linear dynamic model distorted by heterogeneous expectations. Journal of Economic Dynamics and Control 23, 1517–1543.

Bao, T., Hommes, C.H., Sonnemans, J., Tuinstra, J., 2010. Individual expectation, limited rationality and aggregate outcome. CeNDEF working paper 10-07, June.

Basak, S., 2000. A model of dynamic equilibrium asset pricing with heterogeneous beliefs and extraneous risk. Journal of Economic Dynamics and Control 24, 63–95.

Bauer, C., DeGrauwe, P., Reitz, S., 2009. Exchange rate dynamics in a target zone—a heterogeneous expectations approach. Journal of Economic Dynamics and Control 33, 329–344.

[15] These findings are in line with Schunk (2010) who uses laboratory panel data from an intertemporal choice task and shows that (i) a small set of both rational and rule of thumb types explains almost all observed decisions, (ii) there is evidence of type stability, i.e. individuals stick to a strategy for a while and (iii) the type distribution is affected by the decision environment.

Bekiros, S.D., 2010. Heterogeneous trading strategies with adaptive fuzzy Actor-Critic reinforcement learning: a behavioral approach. Journal of Economic Dynamics and Control 34, 1153–1170.
Bénassy-Quéré, L.S., Larribeau, S., MacDonald, R., 2003. Models of exchange rate expectations: How much heterogeneity? International Financial Markets, Institutions and Money 13, 113–136.
Boswijk, H.P., Hommes, C.H., Manzan, S., 2007. Behavioral heterogeneity in stock prices. Journal of Economic Dynamics and Control 31, 1938–1970.
Bottazzi, G., Devetag, G., Pancotto, F., forthcoming. Does volatility matter? Expectations of price return and variability in an asset pricing experiment. Journal of Economic Behavior & Organization.
Branch, W.A., 2004. The theory of rationally heterogeneous expectations: evidence from survey data on inflation expectations. The Economic Journal 114, 592–621.
Branch, W.A., 2007. Sticky information and model uncertainty in survey data on inflation expectations. Journal of Economic Dynamics and Control 31.
Branch, W.A., McGough, B., 2009. A New Keynesian model with heterogeneous expectations. Journal of Economic Dynamics and Control 33, 1036–1051.
Brock, W.A., Hommes, C.H., 1997. A rational route to randomness. Econometrica 65, 1059–1095.
Brock, W.A., Hommes, C.H., 1998. Heterogeneous beliefs and routes to chaos in a simple asset pricing model. Journal of Economic Dynamics and Control 22, 1235–1274.
Brock, W.A., Hommes, C.H., Wagener, F.O.O., 2009. More hedging instruments may destabilize markets. Journal of Economic Dynamics and Control 33, 1912–1928.
Camerer, C.F., Ho, T.H., 1999. Experienced-weighted attraction learning in normal form games. Econometrica 67, 827–874.
Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton.
Capistrán, C., Timmermann, A., 2009. Disagreement and biases in inflation expectations. Journal of Money, Credit and Banking 41, 365–396.
Carlson, J., 1967. The stability of an experimental market with a supply response lag. Southern Economic Journal 33, 305–321.
Chavas, J.P., 2000. On information and market dynamics: the case of the U.S. beef market. Journal of Economic Dynamics and Control 24, 833–853.
Chiarella, C., Dieci, R., He, X., 2009. Heterogeneity, market mechanisms, and asset price dynamics. In: Hens, T., Schenk-Hoppé, K.R. (Eds.), Handbook of Financial Markets: Dynamics and Evolution. Elsevier, pp. 277–344.
Chiarella, C., He, X., 2002. Heterogeneous beliefs, risk and learning in a simple asset pricing model. Computational Economics 19, 95–132.
Chiarella, C., He, X., 2003. Heterogeneous beliefs, risk and learning in a simple asset pricing model with a market maker. Macroeconomic Dynamics 7, 503–536.
Colander, D., Föllmer, H., Haas, A., Goldberg, M.D., Juselius, K., Kirman, A., Lux, T., Sloth, B., 2009. The Financial crisis and the systemic failure of academic economics. University of Copenhagen, Department of Economics, Discussion Paper No. 09-03. Available at SSRN: 〈http://ssrn.com/abstract=1355882〉.
Colucci, D., Valori, V., 2006. Ways of learning in a simple economic setting: a comparison. Chaos, Solitons and Fractals 29, 653–670.
DeGrauwe, P., 2009. What's wrong with modern macroeconomics. Top-down versus bottom-up macroeconomics. In: CESifo Economic Studies Conference Paper, November.
Delli-Gatti, D., Gaffeo, E., Gallegati, M., Giulioni, G., Pallestrini, A., 2008. Emergent Macroeconomics. An Agent-Based Approach to Business Fluctuations. Springer-Verlag, Milan.
Dieci, R., Westerhoff, F., 2010. Heterogeneous speculators, endogenous fluctuations and interacting markets: a model of stock prices and exchange rates. Journal of Economic Dynamics and Control 34, 743–764.
Diks, C., Dindo, D., 2008. Informational differences and learning in an asset market with boundedly rational agents. Journal of Economic Dynamics and Control 32, 1432–1465.
Diks, C.G.H., van der Weide, R., 2005. Herding, a-synchronous updating and heterogeneity in memory in a CBS. Journal of Economic Dynamics and Control 29, 741–763.
Dreger, C., Stadtmann, G., 2008. What drives heterogeneity in foreign exchange rate expectations: insights from a new survey. International Journal of Finance and Economics 13, 360–367.
Duffy, J., 2006. Agent-based models and human-subject experiments. In: Tesfatsion, L., Judd, K.L. (Eds.), Handbook of Computational Economics, first ed., vol. 2, pp. 949–1011 (chapter 19).
Duffy, J., 2008a. Experimental macroeconomics. In: Durlauf, S., Blume, L. (Eds.), The New Palgrave Dictionary of Economics second ed. Palgrave Macmillan, New York.
Duffy, J., 2008b. Macroeconomics: a survey of laboratory research, In: Kagel, J., Rot, A.E. (Eds.), Handbook of Experimental Economics, vol. 2. Working Paper 334, University of Pittsburgh.
ter Ellen, S., Zwinkels, R.C.J., forthcoming. Oil price dynamics: a behavioral finance approach with heterogeneous agents. Energy Economics.
Elliott, G., Ito, T., 1999. Heterogeneous expectations and tests of efficiency in the yen/dollar forward exchange rate market. Journal of Monetary Economics 43, 435–456.
Erev, I., Roth, A.E., 1998. Prediction how people play games: reinforcement learning in games with unique strategy equilibrium. American Economic Review 88, 848–881.
Evans, G.W., Honkapohja, S., 2001. Learning and Expectations in Macroeconomics. Princeton University Press, Princeton.
Fehr, E., Tyran, J.-R., 2001. Does money illusion matter? American Economic Review 91, 1239–1262.
Fehr, E., Tyran, J.-R., 2005. Individual irrationality and aggregate outcomes. Journal of Economic Perspectives 19, 43–66.
Fehr, E., Tyran, J.-R., 2008. Limited rationality and strategic interaction: the impact of the strategic environment on nominal inertia. Econometrica 76, 353–394.
Franke, R., 2009. Applying the method of simulated moments to estimate a small agent-based asset pricing model. Journal of Empirical Finance 16, 804–815.
Frankel, J.A., Froot, K.A., 1987. Short-term and long-term expectations of the yen/dollar exchange rate: evidence from survey data. Journal of the Japanese and International Economies 1, 249–274 also published as NBER working paper 2216, April 1987.
Frankel, J.A., Froot, K.A., 1990. The rationality of the foreign exchange rate. Chartists, fundamentalists and trading in the foreign exchange market. American Economic Review 80 (2), 181–185 AEA Papers and Proceedings.
Frijns, B., Lehnert, T., Zwinkels, R.C.J., 2010. Behavioral heterogeneity in the option market. Journal of Economic Dynamics and Control 34, 2273–2287.
Gaunersdorfer, A., 2000. Endogenous fluctuations in a simple asset pricing model with heterogeneous beliefs. Journal of Economic Dynamics and Control 24, 799–831.
Goldbaum, D., 2005. Market efficiency and learning in an endogenously unstable environment. Journal of Economic Dynamics and Control 29, 953–978.
Goldbaum, D., Mizrach, B., 2008. Estimating the intensity of choice in a dynamic mutual fund allocation decision. Journal of Economic Dynamics and Control 32, 3866–3876.
Grandmont, J.-M., 1998. Expectation formation and stability in large socio-economic systems. Econometrica 66, 741–781.
Guse, E.A., 2010. Heterogeneous expectations, adaptive learning, and evolutionary dynamics. Journal of Economic Behavior & Organization 74, 42–57.
Haltiwanger, J., Waldman, M., 1985. Rational expectations and the limits to rationality: an analysis of heterogeneity. American Economic Review 75, 326–340.
Haruvy, E., Lahav, Y., Noussair, C.N., 2007. Traders' expectations in asset markets: experimental evidence. American Economic Review 97, 1901–1920.
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 and Control 33, 1052–1072.
Hens, T., Schenk-Hoppé, K.R. (Eds.), 2009. Handbook of Financial Markets: Dynamics and Evolution. Elsevier.
Holt, C.A., Villamil, A.P., 1986. A laboratory experiment with a single person cobweb. Atlantic Economic Journal 14, 51–54.

Hommes, C.H., 2006. Heterogeneous agent models in economics and finance. In: Tesfatsion, L., Judd, K.L. (Eds.), Handbook of Computational Economics, Volume 2: Agent-Based Computational Economics. North-Holland, pp. 1109–1186 (Chapter 23).
Hommes, C.H., 2009. Bounded rationality and learning in complex markets. In: Barkley Rosser Jr., J. (Ed.), Handbook of Economic Complexity. Edward Elgar, Cheltenham, pp. 87–123.
Hommes, C.H., Huang, H., Wang, D., 2005b. A robust rational route to randomness in a simple asset pricing model. Journal of Economic Dynamics and Control 29, 1043–1072.
Hommes, C.H., Lux, T., 2009. Individual expectations and aggregate behavior in learning-to-forecast experiments. CeNDEF Working Paper, University of Amsterdam.
Hommes, C.H., Sonnemans, J., Tuinstra, J., van de Velden, H., 2005a. Coordination of expectations in asset pricing experiments. Review of Financial Studies 18, 955–980.
Hommes, C., Sonnemans, J., Tuinstra, J., van de Velden, H., 2007. Learning in cobweb experiments. Macroeconomic Dynamics 11 (S1), 8–33.
Hommes, C.H., Sonnemans, J., Tuinstra, J., van de Velden, H., 2008. Expectations and bubbles in asset pricing experiments. Journal of Economic Behavior & Organization 67, 116–133.
Hommes, C.H., Sonnemans, J., van de Velden, H., 2000. Expectation formation in an experimental cobweb economy. In: Delli Gatti, D., Gallegati, M., Kirman, A. (Eds.), Interaction and Market Structure: Essays on Heterogeneity in Economics. Lecture Notes in Economics and Mathematical Systems, vol. 484. Springer-Verlag, pp. 253–266.
Hommes, C.H., Sorger, G., 1998. Consistent expectations equilibria. Macroeconomic Dynamics 2, 287–321.
Hommes, C.H., Wagener, F.O.O., 2009. Complex evolutionary systems in behavioral finance. In: Hens, T., Schenk-Hoppé, K.R. (Eds.), Handbook of Financial Markets: Dynamics and Evolution. Elsevier, pp. 217–276.
Hommes, C.H., Wagener, F.O.O., 2010. Does eductive stability imply evolutionary stability? Journal of Economic Behavior & Organization 75, 25–39
Huang, W., Zheng, H., Chia, W.-M., 2010. Financial crises and interacting heterogeneous agents. Journal of Economic Dynamics and Control 34, 1105–1122.
de Jong, E., Verschoor, W.F.C., Zwinkels, R.C.J., 2009. Behavioural heterogeneity and shift-contagion: evidence from the Asian crisis. Journal of Economic Dynamics and Control 33, 1929–1944.
Kelley, H., Friedman, D., 2002. Learning-to-forecast price. Economic Inquiry 40, 556–573.
Kirman, A., 2010. Complex Economics: Individual and Collective Rationality. Routledge.
LeBaron, B., 2006. Agent-based computational finance. In: Tesfatsion, L., Judd, K.L. (Eds.), Handbook of Computational Economics, Volume 2: Agent-Based Computational Economics. North-Holland, Amsterdam, pp. 1187–1233 (Chapter 24).
Lei, V., Noussair, C.N., Plott, C.R., 2001. Nonspeculative bubbles in experimental asset markets: lack of common knowledge of rationality vs. actual irrationality. Econometrica 69, 831–859.
Lillo, F., Moro, E., Vaglica, G., Mantegna, R.N., 2008. Specialization and herding behavior of trading firms in a financial market. New Journal of Physics 10, 043019.
Lines, M., Westerhoff, F., 2010. Inflation expectations and macroeconomic dynamics: the case of rational versus extrapolative expectations. Journal of Economic Dynamics and Control 34, 246–257.
Lucas, R.E., 1972. Econometric testing of the natural rate hypothesis. In: Eckstein, O. (Ed.), The Econometrics of Price Determination Conference. Board of Governors of the Federal Reserve System and Social Science Research Council, Washington DC, pp. 50–59.
Lucas, R.E., 1986. Adaptive behavior and economic theory. Journal of Business 59 (4), S401–S426.
Lux, T., 2009. Stochastic behavioral asset pricing models and the stylized facts. In: Hens, T., Schenk-Hoppé, K.R. (Eds.), Handbook of Financial Markets: Dynamics and Evolution. Elsevier.
MacDonald, R., Marsh, I., 1996. Currency forecasters are heterogeneous: confirmation and consequences. Journal of International Money and Finance 15, 665–685.
Mankiw, N.G., Reis, R., Wolfers, J., 2003. Disagreement about inflation expectations. In: Gertler, M., Rogoff, K. (Eds.), NBER Macroeconomics Annual 2003.
Marimon, R., Spear, S.E., Sunder, S., 1993. Expectationally driven market volatility: an experimental study. Journal of Economic Theory 61, 74–103.
Marimon, R., Sunder, S., 1993. Indeterminacy of equilibria in a hyperinflationary world: experimental evidence. Econometrica 61, 1073–1107.
Marimon, R., Sunder, S., 1994. Expectations and learning under alternative monetary regimes: an experimental approach. Economic Theory 4, 131–162.
Marimon, R., Sunder, S., 1995. Does a constant money growth rule help stabilize inflation? Carnegie-Rochester Conference Series on Public Policy 43, 111–156
Menkhoff, L., Taylor, M.P., 2007. The obstinate passion of foreign exchange professionals: technical analysis. Journal of Economic Literature 45, 936–972.
Menkhoff, L., Rebitzky, R.R., Schröder, M., 2009. Heterogeneity in exchange rate expectations: evidence on the chartist-fundamentalists approach. Journal of Economic Behavior & Organization 70, 241–252.
Muth, J.F., 1961. Rational expectations and the theory of price movements. Econometrica 29, 315–335.
Pesaran, H.M., Weale, M., 2006. Survey expectations. In: Elliott, G., Granger, C.W.J., Timmermann, A. (Eds.), Handbook of Economic Forecasting. North-Holland, Amsterdam, pp. 715–776.
Pfajfar, D., Santoro, E., 2010. Heterogeneity, learning and information stickiness in inflation expectations. Journal of Economic Behavior & Organization 75, 426–444.
Pfajfar, D., Zakelj, B., 2010. Inflation expectations and monetary policy design: evidence from the laboratory. Working Paper Tilburg University, March.
Plot, C.R., Smith, V.L., 1978. An experimental examination of two exchange institutions. Review of Economic Studies XLV, 133–153.
Plot, C.R., Sunder, S., 1982. Efficiency of experimental security markets with insider information: an application of rational expectations models. Journal of Political Economy 90, 663–698.
Rosser Jr., J.B., 2009. Handbook of Economic Complexity. Edward Elgar, Cheltenham.
Sargent, T.J., 1993. Bounded Rationality in Macroeconomics. Clarendon Press.
Sargent, T.J., 2008. Evolution and intelligent design. American Economic Review 98, 5–37.
Schmalensee, R., 1976. An experimental study of expectation formation. Econometrica 44, 17–41.
Schunk, D., 2009. Behavioral heterogeneity in dynamic search situations: theory and experimental evidence. Journal of Economic Dynamics and Control 33, 1719–1738.
Schunk, D., 2010. Heterogeneous agents in intertemporal choice: theory and experimental evidence. Working Paper University of Zurich, May.
Shefrin, H., Statman, M., 1994. Behavioral capital asset pricing theory. The Journal of Financial and Quantitative Analysis 29, 323–349.
Shiller, R.J., 2000. Measuring bubble expectations and investor confidence. Journal of Psychology and Financial Markets 1, 49–60.
Simon, H.A., 1957. Models of Man. Wiley, New York, NY.
Sims, C.A., 1980. Macroeconomics and reality. Econometrica 48, 1–48.
Smith, V.L., 1962. A experimental study of competitive market behavior. Journal of Political Economics 70, 111–137.
Smith, V.L., Suchanek, G.L., Williams, A.W., 1988. Bubbles, crashes and endogenous expectations in experimental spot asset markets. Econometrica 56, 1119–1151.
Sunder, S., 1995. Experimental asset markets: a survey. In: Kagel, J.H., Roth, A.E. (Eds.), Handbook of Experimental Economics. Princeton University Press, pp. 445–500.
Sutan, A., Willinger, M., 2009. Guessing with negative feedback: an experiment. Journal of Economic Dynamics and Control 33, 1123–1133.
Taylor, M.P., Allen, H., 1992. The use of technical analysis in the foreign exchange market. Journal of International Money and Finance 11, 304–314.
Tesfatison, L., Judd, K.L., 2006. Handbook of Computational Economics, Volume 2: Agent-Based Computational Economics. North-Holland.
Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. Science 185, 1124–1131.
Williams, A.W., 1987. The formation of price forecasts in experimental markets. Journal of Money, Credit and Banking 19, 1–18.

Vaglica, G., Lillo, F., Moro, E., Mantegna, R.N., 2008. Scaling laws of strategic behavior and size heterogeneity in agent dynamics. Physical Review E 77, 036110.
Vissing-Jorgensen, A., 2003. Perspective on behavioral finance: does 'irrationality' disappear with wealth? Evidence from expectations and actions. In: Gertler, M., Rogoff, K. (Eds.), NBER Macroeconomics Annual. MIT Press.
Wellford, C.P., 1989. A laboratory analysis of price dynamics and expectations in the cobweb model. Discussion Paper No. 89-15, Department of Economics, University of Arizona.
Williams, J., 1977. Capital asset prices with heterogeneous beliefs. Journal of Financial Economics 5, 219–239.
Zapatero, F., 1998. Effects of financial innovation on market volatility when beliefs are heterogeneous. Journal of Economic Dynamics and Control 22, 597–626.