TI 2013-206/II
Tinbergen Institute Discussion Paper

Reflexivity, Expectations Feedback and almost Self-fulfilling Equilibria: Economic Theory, Empirical Evidence and Laboratory Experiments

Cars Hommes

CeNDEF, Amsterdam School of Economics, University of Amsterdam, and Tinbergen Institute, The Netherlands.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: http://www.dsf.nl/

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Reflexivity, Expectations Feedback and Almost Self-fulfilling Equilibria: Economic Theory, Empirical Evidence and Laboratory Experiments

Cars Hommes[a]

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

August 14, 2013

Abstract

We discuss recent work on bounded rationality and learning in relation to Soros' principle of reflexivity and stress the empirical importance of non-rational, almost self-fulfilling equilibria in positive feedback systems. As an empirical example, we discuss a behavioral asset pricing model with heterogeneous expectations. Bubble and crash dynamics is triggered by shocks to fundamentals and amplified by agents switching endogenously between a mean-reverting fundamental rule and a trend-following rule, based upon their relative performance. We also discuss learning-to-forecast laboratory experiments, showing that in positive feedback systems individuals coordinate expectations on non-rational, almost self-fulfilling equilibria with persistent price fluctuations very different from rational equilibria. Economic policy analysis may benefit enormously by focussing on efficiency and welfare gains in correcting mispricing of almost self-fulfilling equilibria.

JEL Classification: D84, D83, E32, C92

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

Acknowledgment. Financial support from the Institute for New Economic Thinking (INET) grant for the project *Heterogeneous Expectations and Financial Crises* (INO1200026) is gratefully acknowledged.

1 Introduction

In his book "The Alchemy of Finance", George Soros (1987) introduced the principles of *fallibility* and *reflexivity* to describe the evolving state of financial markets and the economy. As a very successful market participant Soros argued that standard economic theory built on the paradigm of rationality is a poor description of economic reality and has been of little help to guide investment behavior. Soros articulated the crucial role of expectations and feedback in the economy and the lack of a realistic description of these phenomena by the rational expectations paradigm. Soros' view has been updated and described elegantly in his recent contribution Soros (2013) to this special issue. Here we discuss the relation between economic theory, especially the role of expectations and learning, and Soros' principles of fallibility and reflexivity emphasizing empirical and laboratory evidence.

Let me start by recalling the two principles *fallibility* and *reflexivity* and their central role in social science and economics in his own words (Soros, 2013):

"*The first is that in situations that have thinking participants, the participants view of the world never perfectly corresponds to the actual state of affairs.* ⋯ *The second proposition is that these imperfect views can influence the situation to which they relate through the behavior of the participants* ⋯ *it connects the universe of thoughts with the universe of events.* ⋯ *Reflexive feedback loops between the cognitive and manipulative functions connect the realms of beliefs and events. The participants views influence but do not determine the course of events, and the course of events influences but does not determine the participants views. The influence is continuous and circular; that is what turns it into a feedback loop.*"

Let me contrast this view with a quote from Muth's classical paper introducing rational expectations. Muth was well aware that aggregation of individual expectations into a representative rational forecast depends critically on whether or not these individual expectations are *correlated* (Muth, 1961, p.321, emphasis added):

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

Who is right, Soros or Muth? I will review some recent theory, empirical evidence and laboratory experiments that shed some light on this debate.

2 Expectations Feedback & Bounded Rationality

Soros recognizes the crucial difference between natural and social sciences: in social systems participants can think and affect actual events. Weather forecasts will not affect the probability of rain, but a forecast of the macroeconomic outlook by the president of the ECB may affect the likelihood of a recession. A dynamic economic model is an expectations feedback system, mapping individual beliefs into actions and market realizations, shaping new market expectations, etc. A simple form of an expectations feedback system is

$$p_t = F(p^e_{1,t+1}, p^e_{2,t+1}, \cdots p^e_{H,t+1}), \tag{1}$$

where today's realized market price p_t depends on the individual forecasts $p^e_{j,t+1}$ for tomorrow of all economic agents.

Traditional economics is built on the paradigm of rational expectations (RE) introduced by Muth (1961) and popularized in macroeconomics by Lucas and Prescott (1971) and others. All agents are assumed to be perfectly rational using economic theory to form their expectations. All subjective beliefs then coincide with objective *model consistent* expectations, and the model can be solved for rational expectations equilibrium (REE), which is essentially a *fixed point* of the expectations feedback system. An important motivation contributing to the popularity of RE has been the Lucas critique (Lucas, 1976), that policy conclusions based on non-RE models are potentially misleading, because changes in policy will alter individual behavior. In particular, expectations should not depend on exogenous parameters, but should take policy changes into account.

Many economists today are well aware that RE imposes unrealistically high cognitive and informational assumptions on the agents in the economy and that some form of bounded rationality is needed. But which form? RE disciplines economic modeling in an elegant and convenient way. By imposing RE, all parameters of individual forecasting are removed from the model. Allowing for non-rational expectations begs the question which errors the model should allow for. This leads to Sims' metaphor of the "wilderness of bounded rationality": if agents are non-rational, there are a million ways of how individual agents may make mistakes.

One alternative approach to bounded rationality that is gaining some ground in macroeconomics is *adaptive learning*. Boundedly rational agents do not have perfect knowledge about the economy, but act as econometricians or statisticians using an econometric forecasting model and updating the parameters over time as additional observations become available; see, e.g., Sargent (1993) and Evans and Honkapohja (2001, 2013) for extensive surveys and references. The original motivation for this literature has been to study conditions under which learning converges to the RE, in the hope that learning may enforce RE without assuming perfect knowledge of the expectations feedback system. Agents are then assumed to know the structural equations of the economy, but not the parameters which need to be learned over time as additional observations become available. Many examples, however, have been provided where learning does *not* settle down to RE, but to non-rational equilibria, explaining high persistence and excess volatility, as, e.g., in the learning equilibria in Bullard (1994) or the self-fulfilling mistakes in Grandmont (1998). A *behavioral learning* approach based on simplicity and parsimony has been advocated by so-called Restricted Perception Equilibria (Branch 2006; Hommes and Zhu (2013)). Agents base their expectations on simple forecasting heuristics, such as an AR(1) rule, with the parameters pinned down by simple consistency requirements between beliefs and market realizations, for example, based on intuitive and observable quantities such as the mean and the first-order autocorrelation.

Another complementary approach to bounded rationality are *heterogeneous expectations* models as e.g. introduced in Brock and Hommes (1997, 1998) and Branch and Evans (2006). Agents endogenously switch between different forecasting rules, ranging from simple heuristics to more sophisticated strategies, based upon their relative performance. Notice that in both adaptive learning and heterogeneous switching models, the learning is *endogenous* and agents will adapt to policy changes, so that these models, at least to a first order approximation, mitigate the Lucas critique.

These recent approaches are much in the spirit of Soros' principles of fallibility and reflexivity. Agents do *not* know the correct model of the economy, but rather use some misspecified forecasting rules which may be heterogeneous across agents. This leads to a complex economic expectations feedback system. A REE may arise as a special case in which the equilibrium is exactly self-fulfilling, but often *almost self-fulfilling* behavioral learning equilibria will arise exhibiting excess volatility and deviating persistently from the rational benchmark. In what follows we will discuss the empirical relevance of *almost self-fulfilling* equilibria.

3 A behavioral asset pricing model

We consider a stylized asset pricing model with heterogeneous beliefs, as in Brock and Hommes (1998), and fit a 2-type model to S&P500 data. Investors can choose between a risk free asset paying a fixed return r and a risky asset (say a stock) paying uncertain dividends. Assume that investors have perfect knowledge of the exogenous cash flow process, and thus know the 'fundamental value' of the risky asset, but differ in their beliefs about the future price of the asset . Denote Y_t as the dividend payoff and P_t as the asset price. The market clearing pricing equation is given by:

$$P_t = \frac{1}{1+r}\bar{E}_t[P_{t+1} + Y_{t+1}], \tag{2}$$

where $\bar{E}_t[.]$ denote average expectations of the population of investors.
The dividend process follows a geometric random walk with drift:

$$\log Y_{t+1} = \mu + \log Y_t + \nu_{t+1}, \qquad \nu_{t+1} \sim IID(0, \sigma_\nu^2). \tag{3}$$

Investors are assumed to have correct, model-consistent beliefs about the exogenous dividend process, $E_{i,t}[Y_{t+1}] = (1+g)Y_t$, where $g \equiv e^{\mu + \frac{1}{2}\sigma_\nu^2}$ is the constant growth rate of dividends. This assumption has the convenient feature that the model can be written in deviations from a RE benchmark fundamental.

In the special case where *all* agents have rational expectations about prices, the price equals its RE *fundamental value* given by the discounted sum of all future expected dividends[1]:

$$P_t^* = \frac{1+g}{r-g}Y_t. \tag{4}$$

Hence, under RE the price-to-dividend ratio is constant and given by $\frac{P_t^*}{Y_t} = \frac{1+g}{r-g} \equiv \delta^*$.

Fig. 1 illustrates the S&P500 stock market index, the price-to-dividend ratio δ_t and the fundamental value. The S&P500 index clearly exhibits *excess volatility*, as pointed out already in the seminal paper of Shiller (1981). Boswijk et al. (2007) estimated a 2-type model using yearly S&P500 data from 1871-2003. More recently Hommes and Veld (2013) updated the estimation of the 2-type model using quarterly data 1950Q1-2012Q3. In deviations from the fundamental value $x_t \equiv \delta_t - \delta^*$, the

[1] This solution is known as the Gordon model.

2-type model is given by:

$$x_t = \frac{1}{R^*}(n_{1,t}E_{1,t}[x_{t+1}] + n_{2,t}E_{2,t}[x_{t+1}]), \qquad R^* \equiv \frac{1+r}{1+g}. \tag{5}$$

The asset pricing model has *positive expectations feedback*, that is, realized price deviation increases (decreases) when (average) expected deviation increases (decreases). Consider the simplest form of heterogeneity with belief types which are linear in the last observation:

$$E_{h,t}[x_{t+1}] = \phi_h x_{t-1}. \tag{6}$$

Two types, $h = 1, 2$, are sufficient to capture the essential difference in agents' behavior: fundamentalists believe the price will return to its fundamental value ($0 \leq \phi_1 < 1$) and chartists believe that the price (in the short run) will move away from the fundamental value ($\phi_2 > 1$).

The fractions of the two types are updated with a multi-nomial logit model based on their relative performance, as in Brock and Hommes (1997), with the *intensity of choice* β measuring how quickly agents switch strategies:

$$n_{h,t+1} = \frac{e^{\beta U_{h,t}}}{\sum_{j=1}^{H} e^{\beta U_{j,t}}}. \tag{7}$$

The performance measure $U_{h,t}$ is a weighted average of past profits and past fitness, with memory parameter ω:

$$U_{h,t} = (1-\omega)\pi_{h,t} + \omega U_{h,t-1}. \tag{8}$$

Hence, consistent with empirical observations, agents tend to switch to strategies that generated higher profits in the recent past.

The econometric form of the endogenous switching model is an AR(1)- model with a time-varying coefficient:

$$R^* x_t = n_{1,t}\phi_1 x_{t-1} + (1-n_{1,t})\phi_2 x_{t-1} + \epsilon_t, \qquad R^* = \frac{1+r}{1+g}, \tag{9}$$

where ϵ_t represents an IID error term. The estimated parameter values in Hommes and in't Veld (2013) are:

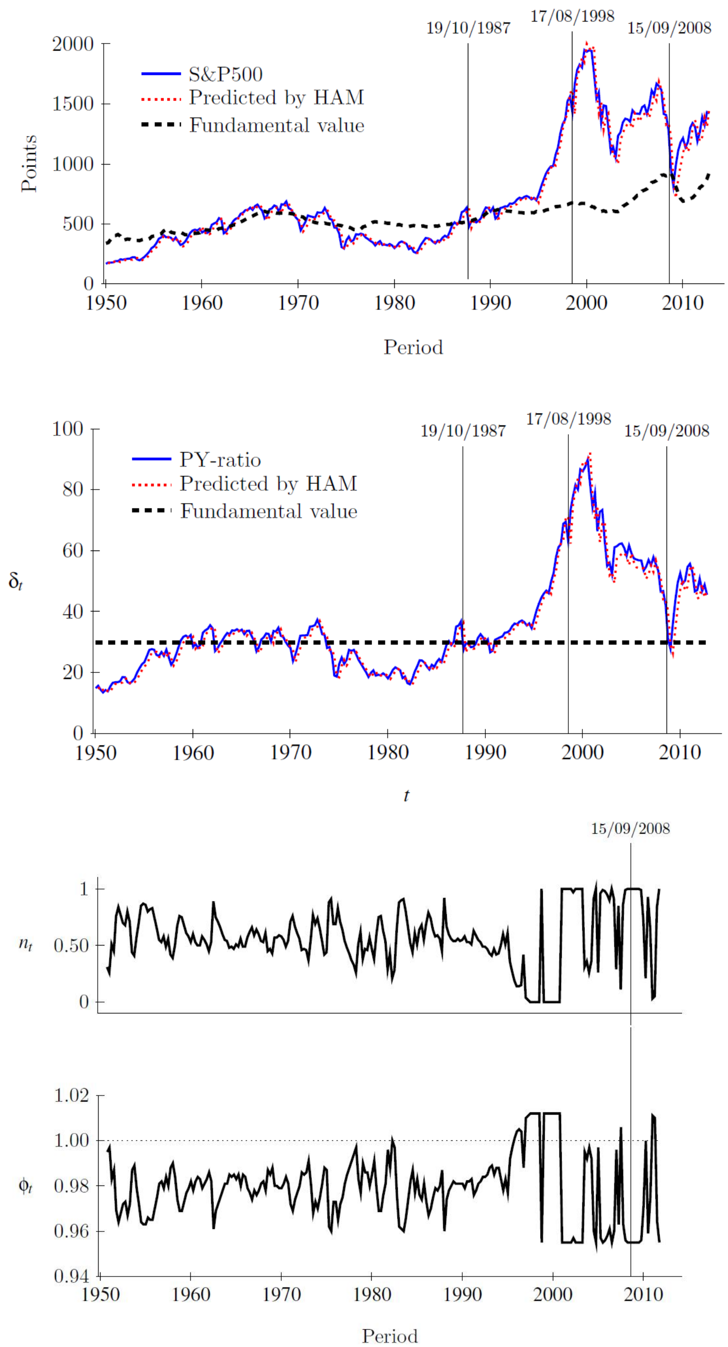

Figure 1: Time series of S&P500 and its fundamental value (top panel), price-to-dividend ratio and its (constant) fundamental (second panel), estimated fraction $n_{1,t}$ of fundamentalists (third panel) and the corresponding time varying market sentiment (bottom panel)

- $\phi_1 = 0.953$: type 1 therefore are *fundamentalists*, expecting (slow) mean reversion of the price towards its fundamental value;
- $\phi_2 = 1.035$: type 2 are *trend-extrapolators*, expecting the price deviation from fundamental to increase by 3.5% per quarter;
- $\omega = 0.816$: implying almost 20% weight is given to the most recent profit observation and about 80% to past profitability.

Define the *market sentiment* as

$$\phi_t = \frac{n_{1,t}\phi_1 + (1 - n_{1,t})\phi_2}{R^*} \tag{10}$$

Figure 1 shows the time series of the estimated fractions of fundamentalists and the market sentiment. The fraction of fundamentalists varies considerably, but gradually (due to memory) over time, with values between 0.25 and 0.9 until the 1990s, and more extreme values ranging from almost 0 to 1 after the dot com bubble. The switching model offers an intuitive explanation of the dot com bubble, as being triggered by economic fundamentals (good news about a new internet technology) strongly amplified by trend-following behavior. Estimates of the market sentiment ϕ_t vary between 0.96 and 1 until the 1990s, showing near-unit root behavior. During the dot com bubble the market sentiment ϕ_t exceeds 1 for several quarters, with the market being temporarily explosive. During the financial crisis the market is mainly dominated by fundamentalists indicating that the financial crisis has been reenforced by fundamentalists, who expected a correction of asset prices back to fundamentals.

In this behavioral asset pricing model with heterogeneous beliefs, agents switch endogenously between a mean-reversion and a trend-following strategy based upon realized profitability and aggregate behavior is very different from the rational benchmark. Strategy switching driven by (short run) profitability leads to an almost self-fulfilling equilibrium with endogenously generated bubbles triggered by shocks to fundamentals ("news") and fueled by positive feedback from trend-followers and market crashes reenforced by negative feedback from fundamentalists.

4 Laboratory Experiments

Laboratory experiments are well suited to study expectations feedback systems within a controlled environment. Hommes (2011) surveys so-called Learning-to-Forecast Experiments (LtFE), where subjects have to forecast a price, whose realization depends endogenously on their average forecast. The main goal of LtFEs is to study how individual expectations are formed, how these interact and which structure emerges at the aggregate level. Will agents learn to coordinate on a common forecast and will the price converge to the RE benchmark or will other aggregate behavior arise?

In the asset pricing LtFEs in Hommes et al. (2005) there are two assets, a risk free asset paying a fixed rate of return r and a risky asset, with price p_t, paying an uncertain dividend y_t. The asset market is populated by six large pension funds and a small fraction of fundamentalist robot traders. Six subjects are forecast advisors to each of the pension funds. Subjects' only task is to forecast the price p_{t+1} of the risky asset for 50 periods and, based on this forecast, the pension fund then computes how much to invest in the risky asset according to a standard mean-variance demand function. The fundamentalist robot trader always predict the *fundamental price* p^f and trades based upon this prediction. The realized asset price in the experiment is derived by market clearing and given by:

$$p_t = \frac{1}{1+r}\Big((1-n_t)\bar{p}^e_{t+1} + n_t\, p^f + \bar{y} + \varepsilon_t\Big), \tag{11}$$

where $\bar{p}^e_{t+1} = (\sum_{h=1}^{6} p^e_{h,t+1})/6$ is the average two-period ahead price forecast, $p^f = \bar{y}/r$ is the fundamental price, and ε_t are small shocks. Subjects do *not* know the underlying law of motion (11), but they do know the mean-dividend $\bar{y}$ and the interest rate r, so they could use these to compute the fundamental price and use it in their forecast. The fraction n_t in (11) is the share of computerized fundamental robot traders, increasing as the price moves away from the fundamental benchmark according to

$$n_t = 1 - \exp\big(-\frac{1}{200}|p_{t-1} - p^f|\big). \tag{12}$$

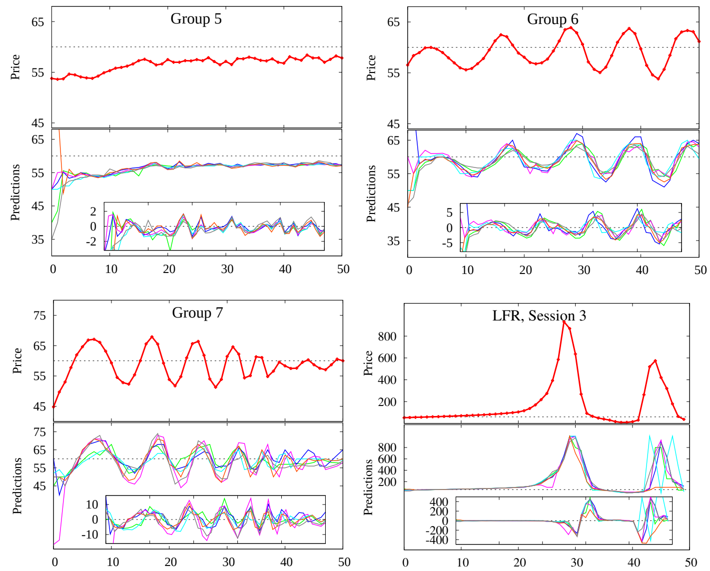

Figure 2: Laboratory experiments: realized market prices (upper part each panel), six individual predictions (middle part each panel) and individual errors (bottom part of each panel). Three asset markets with robot traders (upper + bottom left) and one asset market without robot traders (bottom right). Prices do not converge to the RE fundamental benchmark 60, but rather fluctuate. In the market without fundamental robot trader (bottom right) a long-lasting bubble arises. Individual expectations coordinate on almost self-fulfilling equilibria.

The fundamental trader thus acts as a "far from equilibrium" stabilizing force in the market, adding negative feedback when the asset price becomes overvalued. The negative feedback becomes stronger the more price moves away from fundamental. The overall expectations feedback system (11) has positive feedback, but the positive feedback becomes less strong (i.e. stronger mean-reverting) when price moves away from fundamental value.

Fig. 2 shows time series of prices, individual predictions and forecasting errors in three different groups with a robot trader. A striking feature of aggregate price behavior is that three different qualitative patterns emerge. The price in group 5 converges slowly and almost monotonically to the fundamental price level 60. In group 6 persistent oscillations are observed during the entire experiment, while in group 7 prices fluctuate but the amplitude is decreasing.

A second striking result is that in all groups participants were able to *coordinate* their forecasts. The forecasts, as shown in the lower parts of the panels, are dispersed in the first periods but then, within 3-5 periods, move close to each other. The coordination of individual forecasts has been achieved in the absence of any communication between subjects, other than through the realized market price, and without any knowledge of past and present predictions of *other* participants.

The fourth group in Fig. 2 shows a time series of prices, in a market *without* fundamental traders (Hommes et al., 2008). In the absence of a far from equilibrium stabilizing force due to negative feedback from the fundamental robot traders, a long-lasting asset price bubble occurs with asset prices rising above 900, i.e. more than 15 times the fundamental price, before reaching an exogenously imposed upper-bound of 1000 and a subsequent market crash.

These asset market laboratory experiments exhibit a strong degree of reflexivity. Markets do not converge to the perfectly self-fulfilling RE fundamental 60, but rather fluctuate persistently and exhibit expectations driven bubbles and crashes along almost self-fulfilling equilibria.

Positive versus negative feedback experiments

The asset pricing experiments are characterized by *positive expectations feedback*, that is, an increase of the average forecast or an individual forecast causes the realized market price to rise. Heemeijer et al. (2009) investigate how exactly the expectations feedback structure affects individual forecasting behaviour and aggregate market outcomes. Their (unknown) price generating rules were:

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

$$p_t = 60 + \frac{20}{21}[(\sum_{h=1}^{6} \frac{1}{6} p^e_{ht}) - 60] + \epsilon_t, \qquad \text{positive feedback} \tag{14}$$

where ϵ_t is a small random shock to the pricing rule. The positive and negative feedback systems (13) and (14) have the same unique RE equilibrium steady state

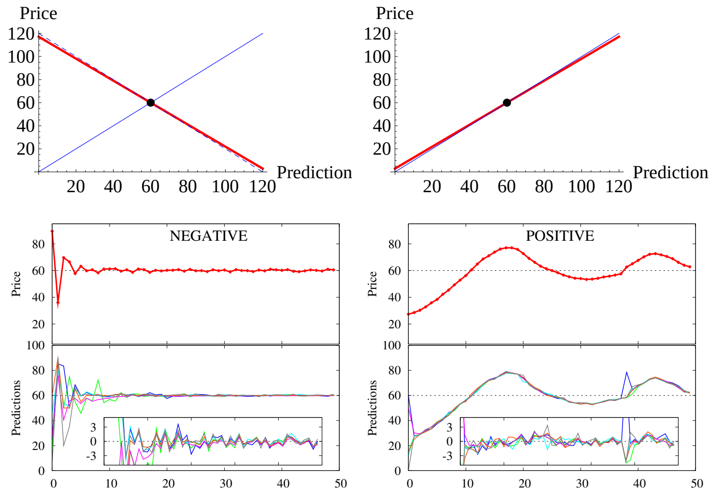

Figure 3: Laboratory experiments: realized market prices (upper part each panel), six individual predictions (middle part each panel) and individual errors (bottom part of each panel). Negative (bottom left) vs. positive (bottom right) feedback experiments. In the negative expectations feedback market the realized price quickly converges to the RE benchmark 60. In all positive feedback markets individuals coordinate on the "wrong" price forecast and as a result the realized market price persistently deviates from the RE benchmark 60.

$p^* = 60$ and *only* differ in the sign of the expectations feedback map. Both are linear near-unit-root maps, with slopes $20/21 \approx -0.95$ resp. $+20/21$. Fig. 3 (top panels) illustrates the dramatic difference in the negative and positive expectations feedback maps. Both have the same unique RE fixed point, but under near-unit-root positive feedback, as is typical in asset pricing models, each point is in fact an *almost self-fulfilling equilibrium*. Will subjects in LtFEs be able to coordinate on the unique RE fundamental price, the only equilibrium that is perfectly self-fulfilling?

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

In his seminal paper introducing RE, Muth (1961) considered a negative expectations feedback framework of the cobweb "hog-cycle" model. Our LtFEs show that under negative expectations feedback, heterogeneity of individual forecasts around the rational forecast 60 persists in the first 10 periods, and correlated individual deviations from the RE fundamental forecast do not arise (in line with Muth's observations as quoted in the introduction) and the realized market price converges quickly to the RE benchmark. In contrast, in an environment with positive expectations feedback our LtFEs show that, within 2-3 periods, individual forecasts become strongly coordinated and all deviate from the rational, fundamental forecast. As a result, in positive expectations feedback markets, at the aggregate level the market price may persistently deviate from the rational, fundamental price. Individual forecasts than coordinate on almost self-fulfilling equilibrium, very different from the perfectly self-fulfilling RE price[2].

5 A theory of heterogeneous expectations

The fact that qualitatively different aggregate outcomes arise suggests that *heterogeneous expectations* must play a key role to explain these experimental data. Anufriev and Hommes (2012), extending the model of Brock ad Hommes (1997), fitted a simple heuristics switching model (HSM) with four rules to the asset pricing LtFEs.

Agents choose from a number of simple *forecasting heuristics*. The forecasting

[2] In a recent paper Asparouhova et al. (2013) find similar results concerning coordination on almost self-fulfilling equilibria in their laboratory experiments based on Lucas asset pricing model.

heuristics are similar to those obtained from estimating linear models on individual forecasting experimental data. *Evolutionary selection* or *performance based reinforcement learning* based upon relative performance disciplines the individual choice of heuristics. Hence, the impact of each of the rules is evolving over time and agents tend to switch to more successful rules. The four forecasting heuristics are:

$$\begin{aligned}
\text{ADA} \quad p^e_{1,t+1} &= 0.65\, p_{t-1} + 0.35\, p^e_{1,t} &(15)\\
\text{WTR} \quad p^e_{2,t+1} &= p_{t-1} + 0.4\,(p_{t-1} - p_{t-2}) &(16)\\
\text{STR} \quad p^e_{3,t+1} &= p_{t-1} + 1.3\,(p_{t-1} - p_{t-2}) &(17)\\
\text{LAA} \quad p^e_{4,t+1} &= \frac{p^{av}_{t-1} + p_{t-1}}{2} + (p_{t-1} - p_{t-2}), &(18)
\end{aligned}$$

were $p^{av}_{t-1} = \sum_{j=0}^{t-1} p_j$ is the sample average of past prices. *Adaptive expectations* (ADA) predicts that the price is a weighted average of the last observed price p_{t-1} and the last price forecast p^e_t. The *trend-following rules* extrapolate the last price change, either with a weak (WTR) or with a strong (STR) trend parameter. The fourth rule is an *anchor and adjustment* rule (Tversky and Kahneman, 1974), extrapolating a price change from a more flexible anchor.

The fractions of the four forecasting heuristics evolve according to a discrete choice model with *asynchronous updating*:

$$n_{i,t} = \delta\, n_{i,t-1} + (1-\delta)\, \frac{\exp(\beta\, U_{i,t-1})}{\sum_{i=1}^{4} \exp(\beta\, U_{i,t-1})}. \tag{19}$$

The fitness or performance measure of forecasting heuristic i is based upon quadratic forecasting errors, consistent with the earnings in the experiments:

$$U_{i,t-1} = -\big(p_{t-1} - p^e_{i,t-1}\big)^2 + \eta\, U_{i,t-2}\ , \tag{20}$$

where $\eta \in [0,1]$ measures the strength of the agents' *memory*. In the special case $\delta = 0$, (19) reduces to the *the discrete choice model* with synchronous updating; δ represents inertia in switching as subjects change strategies only occasionally. The parameter $\beta \geq 0$ represents the intensity of choice measuring how sensitive individuals are to differences in strategy performance[3].

[3]In the simulations below the parameters are fixed at the benchmark values $\beta = 0.4, \eta = 0.7, \delta =$

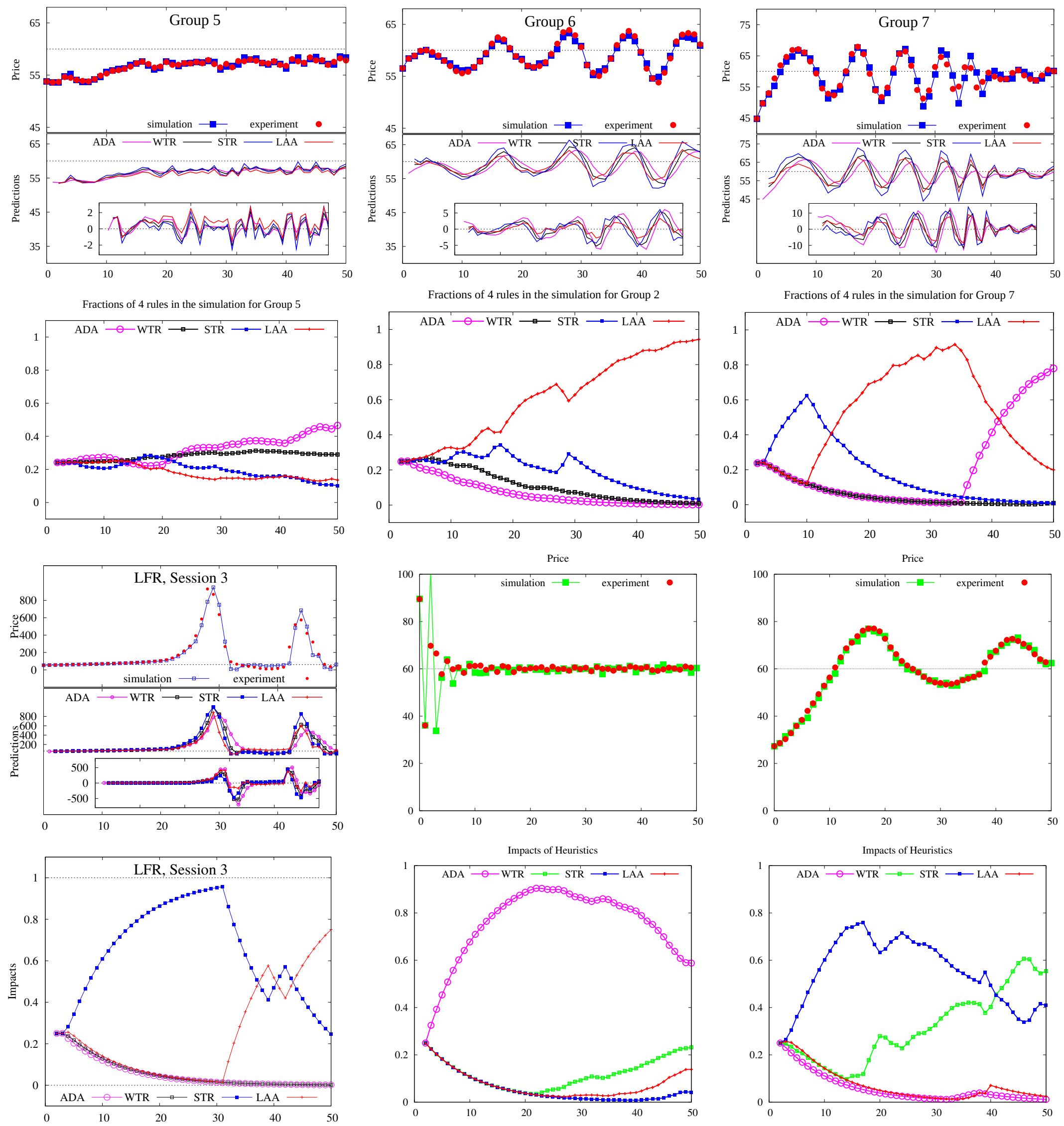

Figure 4: Simulated prices in laboratory experiments in different groups (red) with corresponding one-step ahead predictions of the heuristics switching model (blue), predictions and forecasting errors (inner frames) of four heuristics and time series of fractions of each of the four heuristics adaptive expectations (ADA, purple), weak trend followers (WTR, black), strong trend followers (STR, blue) and anchoring adjustment heuristic (LAA, red). Two top panels correspond to three groups with robot traders; two bottom panels correspond to group without robot trader and large bubble (left panels) and negative and positive feedback groups. In the negative feedback market the adaptive expectations (ADA) rule dominates and enforces quick convergence to the RE fundamental price 60. In the positive expectations feedback market, the strong (STR) and the weak (WTR) trend following rules perform well and reinforce price oscillations. In all positive feedback groups individual expectations coordinate on a non-RE almost self-fulfilling equilibrium.

Fig. 4 compares the experimental data with the *one-step ahead predictions* made by the HSM. The one-step ahead simulations use exactly the same information available to participants in the experiments. The one-period ahead forecasts easily follow the different patterns in aggregate price behavior in all groups. The second and bottom panels show the corresponding fractions of the four heuristics for each group. In different groups different heuristics are dominating the market, after starting off from an equal 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 from the start and its impact increases to about 90% towards the end of the experiment. For the group with the dampened oscillations, one step ahead forecast produces a rich evolutionary selection dynamics, 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, because of its more slowly time varying anchor and its impact starts increasing. The LAA takes the lead in the second phase of the experiment, with oscillating prices, and its share increases to almost 90% after 35 periods. But the 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 the group in the last 9 periods. In the asset market without a fundamental trader subjects coordinate on the strong trend-following strategy, thus explaining the large bubble in the experiment.

The HSM also matches aggregate price behaviour in both the negative and positive feedback experiments very well. The time series of the fractions of the different forecasting heuristics provide an intuitive explanation of how individual learning leads

0.9, as in Anufriev and Hommes (2012).

to different aggregate price behavior. In the negative feedback treatment, the adaptive expectations strategy performs best and within 20 periods it captures more than 90% of the market, thus enforcing convergence towards the RE fundamental equilibrium price. In contrast, in the positive feedback treatment the strong and weak trend-following rules dominate the market, amplifying price fluctuations. The difference in aggregate behavior is thus explained by the fact that *trend following rules are successful in a positive feedback environment* reinforcing price oscillations and persistent deviations from the fundamental equilibrium benchmark price, while the trend-following rules are driven out by adaptive expectations in the case of negative feedback. Self-confirming coordination on trend-following rules in a positive expectations feedback environment has an aggregate effect with realized market prices deviating significantly and persistently from the RE benchmark.

6 Conclusions

The main conclusion to be drawn from the theoretical, empirical and experimental work discussed below may be formulated as follows. In positive feedback markets aggregate behavior is *not* well described by perfectly self-fulfilling rational expectations equilibrium. Instead, under positive feedback individuals tend to coordinate their expectations on *almost self-fulfilling equilibria*, very different from the exact rational self-fulfilling equilibria, and characterized by excess volatility and persistent price fluctuations.

The main finding, consistent with Soros principles of fallibility and reflexivity, is that under positive expectations feedback almost self-fulfilling equilibria provide a much better fit to individual and aggregate behavior in lab experiments and empirical data than the perfectly self-fulfilling REE in traditional models. Parsimony is an important and attractive feature as simplicity makes coordination on an almost self-fulfilling equilibrium more likely as a description of aggregate behavior. Agents make mistakes, as their belief is only an approximation of complex reality. But in equilibrium, the mistake becomes self-fulfilling and it is not easy for agents to improve upon their individual forecasting.

Macroeconomics assigns a central role for expectations in economic modeling. For example, in his standard work on monetary policy in modern New Keynesian macro, Woodford (2003) emphasizes the key role of "managing expectations" for monetary policy. Policy analysis should focus more on managing almost self-fulfilling equilibria. The bounded rationality models discussed here, at least to a first order approximation, take into account the Lucas critique, as expectations and learning will adapt to policy changes. Economic policy analysis may benefit enormously by focussing on efficiency and welfare gains in correcting mispricing of almost self-fulfilling equilibria.

Mapping F is nonlinear complex system and/of agent-based model.

References

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

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

[3] Bao, T., Hommes, C.H., Sonnemans, J. and Tuinstra, J. (2012), Individual expectations, limited rationality and aggregate outcomes, *Journal of Economic Dynamics and Control* 36, 1101-1120.

[4] Boswijk, H.P., Hommes, C.H. and Manzan, S. (2007), Behavioral heterogeneity in stock prices, *Journal of Economic Dynamics and Control* 31, 1938-1970.

[5] Branch, W.A., 2006. Restricted perceptions equilibria and learning in macroeconomics, in: Colander, D. (Ed.), Post Walrasian Macroeconomics: Beyond the Dy-

namic Stochastic General Equilibrium Model. Cambridge University Press, New York, pp. 135-160.

[6] Branch, W.A., Evans, G.W., 2006. Intrinsic heterogeneity in expectation formation. Journal of Economic Theory, 127, 264-295.

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

[8] Bullard, J., 1994. Learning equilibria. Journal of Economic Theory 64, 468-485.

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

[10] Evans, G.W., Honkapohja, S., 2013. Learning as a rational foundation for macroeconomics and finance, forthcoming in Rethinking Expectations: The Way Forward for Macroeconomics, Roman Frydman and Edmund S. Phelps (eds.), Princeton University Press, 2013.

[11] Grandmont, J.-M., (1998), Expectation formation and stability in large socioeconomic systems, *Econometrica* 66, 741-781.

[12] Heemeijer, P., Hommes, C.H., Sonnemans, J., Tuinstra, J., 2009. Price stability and volatility in markets with positive and negative expectations feedback. Journal of Economic Dynamics & Control 33, 1052-1072.

[13] Hommes, C.H., 2011. The heterogeneous expectations hypothesis: some evidence for the lab. Journal of Economic Dynamics & Control 35, 1-24.

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

[15] Hommes, C.H., Sonnemans, J., Tuinstra, J., and van de Velden, H., (2005) Coordination of expectations in asset pricing experiments, *Review of Financial Studies* 18, 955-980.

[16] Hommes, C.H., Sonnemans, J., Tuinstra, J. and Velden, H. van de, (2008), Expectations and bubbles in asset pricing experiments, *Journal of Economic Behavior & Organization* 67, 116-133.

[17] Hommes, C.H. and Veld, D., in't, (2013), Behavioral heterogeneity and the financial crisis, CeNDEF Working Paper, University of Amsterdam, July 2013.

[18] Hommes, C.H. and Zhu, M. (2013), Behavioral learning equilibria, CeNDEF Working Paper, University of Amsterdam, June 2013.

[19] Lucas, R., 1976, Econometric Policy Evaluation: A Critique", in Brunner, K. and Meltzer, A., (Eds), The Phillips Curve and Labor Markets, Carnegie-Rochester Conference Series on Public Policy 1, New York: American Elsevier, pp. 1946.

[20] R.E. Lucas and E.C. Prescot, (1971), Investment under uncertainty. Econometrica 39, 659-681.

[21] Muth, J.E., 1961. Rational expectaions and the theory of price movements. Econometrica 29, 315-335.

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

[23] Shiller, R.J., (1981) Do stock prices move too much to be justified by subsequent changes in dividends?, *American Economic Review* 71, 421-436.

[24] Shiller, R.J., 2000, Irrational Exuberance, Crown Business.

[25] Soros, G. (1987), The Alchemy of Finance, Hoboken NJ: Wiley & Sons, Inc.

[26] Soros, G. (2013), Fallibility, Reflexivity and the Human Uncertainty Principle, this issue.

[27] Tversky, A. and Kahneman, D. (1974), Judgment under uncertainty: heuristics and biases, *Science* 185, 1124-1131.

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