Positive Expectations Feedback Experiments and Number Guessing Games as Models of Financial Markets

Joep Sonnemans
Jan Tuinstra

University of Amsterdam, and Tinbergen Institute.

Positive expectations feedback experiments and number guessing games as models of financial markets

Joep Sonnemans[a] and Jan Tuinstra[b]

[a]Center for Research in Experimental Economics and political Decision making (CREED) and Tinbergen Institute, corresponding author
[b]Center for Non-linear Dynamics in Economics and Finance (CeNDEF) and Tinbergen Institute,

University of Amsterdam, School of Economics
Roetersstraat 11, 1018WB Amsterdam, The Netherlands
J.H.Sonnemans@uva.nl; J.Tuinstra@uva.nl

August 22, 2008

Abstract

In repeated number guessing games choices typically converge quickly to the Nash equilibrium. In positive expectations feedback experiments, however, convergence to the equilibrium price tends to be very slow, if it occurs at all. Both types of experimental designs have been suggested as modeling essential aspects of financial markets. In order to isolate the source of the differences in outcomes we present several new treatments in this paper. We conclude that the feedback strength (i.e. the 'p-value' in standard number guessing games) is essential for the results. Furthermore, positive expectations feedback experiments may provide good representations of highly speculative markets while standard number guessing games model financial markets with more emphasis on dividend yield and value stocks.

1. Motivation

In a famous quote Keynes (1936) describes financial investment as a game in which players try to predict average predictions:

"...professional investment may be likened to those newspaper competitions in which the competitors have to pick out the six prettiest faces from a hundred photographs, the prize being awarded to the competitor whose choice most nearly corresponds to the average preferences of the competitors as a whole..."[1]

This beauty contest analogy is often cited in papers on higher order beliefs[2] and has inspired an increasing number of theoretical and experimental contributions to economics and finance (for a recent theoretical study see e.g. Allen et al., 2006). Most experiments focus on (variations of) the so-called number guessing game (see e.g. Nagel, 1995). In this game all players have to simultaneously submit a 'guess' from a certain interval (typically 0-100) and the winner is the player whose choice is closest to a given fraction (typically 2/3) of the average of these chosen numbers. This game has a unique Nash-equilibrium and the distance between a specific guess and the equilibrium value can be considered a measure of the belief this player has about the rationality of the population

---

[1] The quote continues with: "...so that each competitor has to pick, not those faces which he himself finds prettiest, but those which he thinks likeliest to catch the fancy of the other competitors, all of whom are looking at the problem from the same point of view. It is not a case of choosing those which, to the best of one's judgment, are really the prettiest, nor even those which average opinion genuinely thinks the prettiest. We have reached the third degree where we devote our intelligences to anticipating what average opinion expects the average opinion to be. And there are some, I believe, who practise the fourth, fifth and higher degrees." See Keynes (1936), page 156.

[2] See for example, Biais and Bossaerts (1998), Ho et al. (1998), Camerer et al. (2004) and Costa-Gomes and Crawford (2006). An alternative financial market interpretation of the guessing game is that it models the problem of leaving a market just before prices start going down, see Duffy and Nagel (1997) and Ho et al. (1998).

of players, and about the distribution of the higher order beliefs about rationality in the population.

The general findings from the experimental literature on repeated guessing games are that first period choices are not very close to the Nash equilibrium but convergence to that equilibrium is fast (typically within 4-5 periods) and stable. As a characterization of behavior of financial markets this fast convergence is surprising for at least two reasons.

First, empirical evidence suggests that asset markets are in fact not that stable. Shiller (1981, 2000), for example, shows that stock prices are excessively volatile: their variance is higher than that of the underlying fundamental value. Behavioral finance (for recent overviews see Shleifer, 2000, Barberis and Thaler, 2003) has shown that (1) many price movements are unrelated to news but are reactions to price changes (for example caused by investors using technical analyses) and that (2) prices under-react to news, causing short-term trends. Mis-pricing cannot always be arbitraged away (Shleifer and Vishny, 1997) and market prices may therefore deviate substantially from their fundamental values for a longer period of time.

Second, evidence from expectations feedback experiments (see e.g. Hommes et al., 2005a, 2008, Heemeijer et al., 2006) does not seem to be consistent with the results from guessing game experiments. Expectations feedback experiments are based upon the idea that asset markets (just like many other economic environments) are *expectations feedback systems*. Price expectations of traders determine their trading behavior which, in turn, determines the realized trading price. In an expectations feedback experiment participants have to submit their forecast of the future price of a certain asset and are paid according to their prediction accuracy. A computer program determines the optimal

trades associated with the forecasts and the resulting realized trading price. The advantage of this design over traditional experimental asset markets is that it gives a clearer picture of how people form expectations in expectations feedback environments.[3] In prediction experiments with a positive expectations feedback (that is, where an increase in average predictions leads to an increase in the realized market price) there is a remarkable tendency for participants to coordinate on a common prediction strategy but no (or only slow) convergence to the equilibrium price.

These positive feedback prediction experiments are closely related to the number guessing game, but with very different results.[4] Nevertheless, the experimental designs do differ in a number of dimensions, particularly the *feedback strength* from expectations (guesses) to realized price (target number), the *information* given to the participants, and the *incentive structure*. It is, a priori, not evident which of these design differences is responsible for the differences in outcomes. This paper reports on a series of experiments that are designed to isolate the main determinants.

Furthermore, this paper can be considered as a study about *internal validity*: "How robust are the experimental results of number guessing and expectations feedback games to specific design features?", but also as a study about *external validity*: "What kind of experiment provides the best description of the dynamics of financial markets?".

---

[3] In more traditional asset market experiments participants are also sometimes asked to submit price predictions, but it is difficult to give the appropriate incentives for providing these predictions and often they come about as a by-product to the experiment. For a more rigorous approach to expectation formation in experimental asset markets, see Haruvy et al. (2007).

[4] When we started our experimental research on expectation feedback markets we were not fully aware of the close connection with guessing games. This connection only became apparent to us when we changed from using a market-clearing environment (Hommes et al., 2005a), where participants had to predict two periods ahead, to a market-maker environment (Heemeijer, et al., 2006) where participants only have to predict one period ahead (in Section 2.2 we will discuss the differences between these two types of expectations feedback experiments in more detail).

The remainder of this paper is organized as follows. In Section 2 we will briefly review the experimental literature on number guessing games and positive expectations feedback experiments and discuss the differences in design characteristics and outcomes between these two types of experiments. The results from three new experimental studies will be discussed in Section 3 and further analyzed in Section 4. Section 5 concludes.

2. Modeling Keynes' beauty contest

2.1 Guessing games

The typical *guessing game* experiment has the following structure.[5] The game is played for T periods with a fixed group of H participants. In each period t participants simultaneously choose numbers $x_{h,t}^e$ from the interval $[l, u]$. The so-called *target number* is given as[6]

$$x_t = \alpha + \beta \overline{x}_{h,t}^e, \tag{1}$$

where $\alpha \geq 0$ and $0 < \beta < 1$ are fixed parameters[7] and $\overline{x}_{h,t}^e = \frac{1}{H} \sum_{h=1}^{H} x_{h,t}^e$ is the average number chosen in period t. The participant for which $\left| x_{h,t}^e - x_t \right|$ is smallest wins a prize in that period. If several participants have the best guess the prize is split evenly between

---

[5] Moulin (1986) was the first to discuss this game.

[6] In the literature the parameter β is often denoted *p* and the corresponding guessing game is then sometimes referred to as a "*p*-beauty contest". We depart from that convention here, since the variable *p* is used to denote prices in the remainder of this paper.

[7] Most of the guessing games restrict attention to $0 < \beta < 1$, but some of the earlier studies did consider $\beta > 1$. Nagel (1995), for example, has one treatment with $\beta = \frac{4}{3}$ and Ho et al. (1998) discuss treatments with $\beta = 1.1$ and $\beta = 1.3$. Moreover, Sutan and Willinger (2006) discuss experiments on a guessing game with *negative feedback*, i.e. $\beta < 0$, and show that it converges faster than the (positive feedback) guessing game with $\beta > 0$. On the impact of the sign of the feedback, also see Heemeijer et al. (2006).

them. The rules of the game are common knowledge and between periods participants receive feedback about the previous periods' guesses of all participants, the target number and the winning number.

From (1) it is easy to see that the Nash equilibrium of the guessing game corresponds to $x^* = \dfrac{\alpha}{1-\beta}$, provided $l \leq x^* \leq u$: if all participants choose x^* the target number indeed equals x^*. Alternatively, this equilibrium can be found by iterative elimination of dominated strategies.[8]

Finding the Nash equilibrium, for example by iterative elimination of dominated strategies, requires a (potentially high) number of steps of reasoning. The guessing game is a powerful device to study this *depth of reasoning*, as follows. So-called level-0 players randomly select a guess from the interval $[l, u]$. A level-1 player believes all other players are level-0 players, and therefore plays a best response to the expected random choice of the level-0 players, $x^1 = \alpha + \beta x^0$, where x^0 corresponds to the expected average choice of the level-0 players.[9] A level-2 player believes that all other players are level-1 players and therefore best responds to x^1, that is, $x^2 = \alpha + \beta x^1$, and so on. By looking at first period choices the guessing game can be used to classify subjects into different depth of reasoning types.

---

[8] This works as follows. Given that choices have to be in the interval $[l, u]$ the target number always lies in the interval $[\alpha + \beta l, \alpha + \beta u]$. Numbers outside this interval are dominated and can therefore be eliminated. Assuming that no participant chooses a dominated action it follows that the target number must lie in the interval $[\alpha + \beta(\alpha + \beta l), \alpha + \beta(\alpha + \beta u)]$. This implies that all numbers in the intervals $(\alpha + \beta l, \alpha + \beta(\alpha + \beta l))$ and $(\alpha + \beta(\alpha + \beta u), \alpha + \beta u)$ are dominated and can be eliminated, and so on. Eventually, this process of iterative elimination of dominated strategies leads to x^*.

[9] Note that level-1 players do not take into account their own effect upon the target number and furthermore believe that all other players are level-0 players.

The guessing game has been studied extensively in laboratory experiments (for overviews, see Nagel, 1999, and Camerer et al., 2003), typically with $\alpha = 0$ and very often with $\beta = \frac{2}{3}$ and $[l, u] = [0, 100]$. The first of these experiments was reported by Nagel (1995) who considered groups of 15-18 participants playing the game for four periods. Her main conclusions are: (i) First period choices are significantly different from the Nash equilibrium prediction[10] and almost all of these choices correspond to level-0 up to level-3 depth of reasoning; (ii) In subsequent periods there is rapid convergence to the Nash-equilibrium, without an increase in the depth of reasoning.

These results have been corroborated by many other experiments. Ho et al. (1998) show that convergence to the Nash equilibrium is faster when β is farther away from 1, groups are larger, and participants are experienced. Duffy and Nagel (1997) show that when the target number is based upon the median guess (maximum guess) instead of the mean guess, convergence is faster (slower). Nagel's results where also confirmed in three large scale one-shot guessing games, run through newspapers in Germany, Spain and the U.K. and involving thousands of participants (Bosch-Domènech et al., 2003).

All of the experiments discussed above use $\alpha = 0$, implying that the Nash equilibrium lies at the boundary of the action space (typically $x^* = 0$). Some authors have looked at guessing games with interior equilibria ($\alpha > 0$), particularly, Camerer and Ho (1998), Güth et al. (2002) and Kocher and Sutter (2006). The last two papers also depart from the standard winner-takes-all payoff incentive scheme and reward all participants based upon the absolute distance between their guess and the target

---

[10] Even in two-player guessing games where the Nash-equilibrium prediction, 0, strictly dominates all other choices it is only chosen by less than 10% of students and by about one third of professionals (Grosskopf and Nagel, 2008).

number.[11] Güth et al. (2002) conjecture that, because participants try to avoid extreme choices, convergence in games with interior equilibria is faster than in games with boundary equilibria. Their experiment confirms this. Moreover, the fraction of equilibrium choices in the 'interior equilibrium' game is significantly higher than in the 'boundary equilibrium' game, which they attribute partly to the payoff scheme.[12]

The guessing game has been used as a vehicle for investigating a number of other issues. Weber (2003) shows that participants still learn, albeit at a slower rate, if no feedback is given between periods. Kocher et al. (2007) and Sbriglia (2008) show that additional information, such as strategies of the winners in earlier periods, or strategies from participants in an earlier guessing game, facilitates faster convergence to the Nash equilibrium. Slonim (2005) finds that experienced players, when matched with inexperienced players, win the game more often and make choices farther away from the equilibrium. Finally, Kocher and Sutter (2005) and Sutter (2005) show that teams of players learn faster than individuals and increase convergence speed.

2.2 Positive expectations feedback experiments

Consider the following textbook asset pricing model (for reviews, see Cuthberson, 1996, Campbell et al., 1997, and Brock and Hommes, 1998). There are H traders who divide their wealth between two assets. The first asset is risk free, with fixed return $R = 1 + r$, where $r > 0$ is the interest rate. This asset is in perfect elastic supply and its price is

---

[11] Also Costa-Gomes and Crawford (2006) use a payoff scheme that depends upon the absolute distance between the guess and the target number.

[12] Morone and Morone (2008), however, argue that the results by Güth et al. (2002) are partly due to their parameterization. They show that, although first period choices are indeed closer to the equilibrium when the equilibrium is interior, speed of convergence to the equilibrium may be higher when the equilibrium is on the boundary.

normalized to one. The infinitely lived risky asset, with price p_t in period t, is in fixed

aggregate supply z^s and returns uncertain dividends y_t in period t, which are

independently and identically distributed with mean $\bar{y}$. A trader's demand depends upon

his expectation of $p_{t+1} + y_{t+1} - Rp_t$, which is the excess return of the risky asset (one unit

of the risky asset returns $p_{t+1} + y_{t+1}$, whereas the money necessary to buy that unit would

have returned Rp_t when invested in the risk free asset). Assuming trader h is a mean-

variance maximizer his demand for the risky asset in period t is given by

$$z_{h,t} = \frac{E_{ht}\left(p_{t+1} + y_{t+1} - Rp_t\right)}{a\sigma^2}, \tag{2}$$

where $E_{ht}\left(p_{t+1} + y_{t+1} - Rp_t\right)$ denotes trader h's belief about next period excess return

and $V_{ht}\left(p_{t+1} + y_{t+1} - Rp_t\right) = \sigma^2$ corresponds to his beliefs about the variance of excess

returns, which is assumed to be constant over time and the same for all traders. Finally, a

is a risk aversion parameter (again assumed to be the same for all traders).

In period t aggregate excess demand for the risky asset is given by

$$\xi_t = \sum_{h=1}^{H} z_{h,t} - z^s = \sum_{h=1}^{H} \frac{E_{h,t-1}\left(p_t + y_t - \left(1+r\right)p_{t-1}\right)}{a\sigma^2} - z^s. \tag{3}$$

In order to close the asset market model we need to specify a model for price formation.

Market clearing. Under *market clearing* the price adjusts in every period in such a way

that excess demand vanishes. That is, the price p_t in period t is implicitly determined as

the solution to $\xi_t = 0$. Hommes et al. (2005a) report on experiments in this setting.

Letting $E_{h,t}\left(p_{t+1} + y_{t+1}\right) = p^e_{h,t+1} + \bar{y}$ and assuming $z^s = 0$ they obtain

$$p_t = \frac{1}{1+r}\left[\overline{p}^{\,e}_{h,t+1} + \overline{y} + \varepsilon_t\right],$$
(4)

where $\overline{p}^{\,e}_{h,t+1} = \frac{1}{H}\sum_{h=1}^{H} p^e_{h,t+1}$ is the average price prediction and ε_t corresponds to (small)

stochastic demand and supply shocks. Note that the actual realization of *today*'s price p_t

depends upon people's belief of *tomorrow*'s price p_{t+1}. This implies that, when having to

predict p_{t+1}, traders only have information about prices up to period $t-1$. Intuitively, the

reason why investors have to predict two periods ahead is that in order to make a profit

an investor first has to buy (short sell) an asset in period t and after that sell (buy) it in

period $t+1$. Also observe that $p^f = \dfrac{\overline{y}}{r}$ corresponds to the *fundamental* value of the risky

asset (i.e. the discounted value of the stream of future dividends). If, on average, traders

predict $\overline{p}^{\,e}_{h,t+1} = p^f$ the actual price will, in expectation, equal p^f as well.[13]

Participants in the experiment by Hommes et al. (2005a) are explained that they

are the advisor of a large investor, e.g. a pension fund. Their task is to predict future

prices in a stock market and their reward depends on their prediction accuracy. They are

told the investor will take a position in the market that depends on their prediction of

future prices (see Appendix A for complete instructions) and that there are other large

investors in the market advised by other participants. They are *not* told the precise

formula used to calculate the realized price, but they know the direction of the feedback

structure (if many participants expect high (low) prices, investors will buy (sell) more

---

[13] There also exist so-called "rational bubble" solutions of the form $p_t = R^t c + p^f$, with $c \geq 0$. These bubbles grow with a fixed rate $R > 1$ per period (see e.g. Hommes et al., 2008, for a discussion).

stocks and the price will increase (decrease)). The participants receive information in each period about previous prices and their *own* predictions, both in a graph and a table.

For the experiment fourteen groups were investigated, with each group consisting of $H = 6$ participants, predicting prices for 51 periods.[14] Reported predictions had to be between 0 and 100 and two decimals could be used. The risk free rate of return, $r = 0.05$, and the mean dividend, $\bar{y} = 3$, were fixed such that the equilibrium price equals $p^f = 60$.[15] The same realization of shocks ε_t, independently drawn from $N\left(0, \frac{1}{4}\right)$, was used for all groups.

Participants could earn 1300 points each period. The number of points earned in period t by participant h was inversely related to the forecast error as follows

$$e_{h,t} = \max\left\{1300 - \frac{1300}{49}\left(p_t - p_{ht}^e\right)^2, 0\right\}, \tag{5}$$

where 1300 points is equivalent to 0.65 Dutch guilders or 0.30 euro. To avoid negative earnings, earnings in period t were zero when $\left|p_t - p_{ht}^e\right| \geq 7$. This payoff scheme was common knowledge.

The upper panels of Figure 1 show the prices and predictions in a representative group (group 1) from Hommes et al. (2005a). Two features are apparent. First, the asset

---

[14] In some groups robot traders were added. These robot traders always predicted the fundamental price and make a trading decision based upon this prediction. The impact of robot traders in these markets is endogenous: the greater the distance between the actual price and the fundamental price the more these fundamental traders will invest, and the other way around. They therefore act as a 'stabilizing force' pushing prices in the direction of the fundamental price. Behavior in markets with robot traders is qualitatively similar to behavior in markets withouth robot traders (for details see Hommes et al., 2005a).
[15] For three of the 14 groups we chose $r = 0.05$ and $\bar{y} = 2$ resulting in an equilibrium price of $p^f = 40$, which (in contrast to the equilibrium price for the other groups) is *below* the midpoint of the interval from which predictions can be chosen. Behavior in these three groups is qualitatively the same as the behavior in the other groups (for details see Hommes et al., 2005a).

price shows persistent and significant deviations from its fundamental value (upper left panel). Secondly, the dispersion of individual predictions is remarkably small (upper right panel). Participants seem to coordinate on a common prediction strategy. Both features are robust: systematic deviations of the price from the fundamental value and coordination of prediction strategies are exhibited by all 14 groups. In 11 of these 14 groups prices and predictions oscillate around the fundamental value and in only four of these 11 'oscillating' groups these fluctuations seem to dampen towards the end of the experiment (without completely converging). In two of the remaining groups there is monotonic but slow convergence. For the final group the picture is a bit blurred due to a typing error of one of the participants in one of the periods.
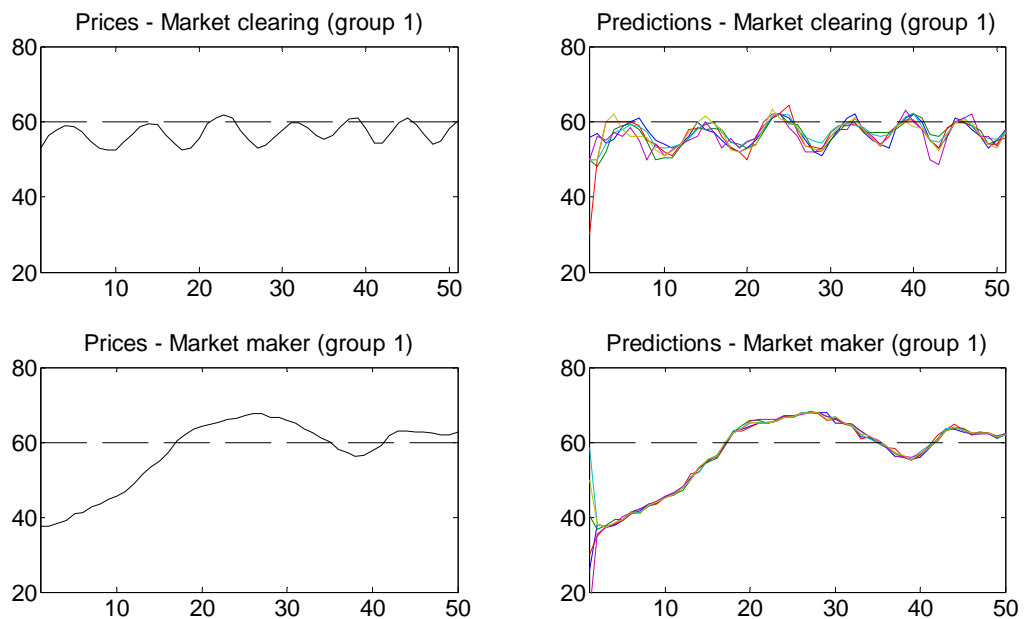

Figure 1: Upper panels show prices (left) and predictions (right) from group 1 in Hommes et al. (2005a), lower panels show prices (left) and predictions (right) from group 1 in Heemeijer et al. (2006).

Several variations of this game should be mentioned. In Hommes et al. (2008) the same model is used, without robot traders and without an upper limit for price predictions. In that setup numerous speculative bubbles and crashes emerge, with prices increasing to almost 16 times the fundamental value. Also a high degree of coordination on a common prediction strategy can be observed in this experiment. Bottazzi and Devetag (2005) study a variant of this experiment with two changes. First, participants have to give a confidence interval for the realized price, instead of a point prediction. Second, participants are rewarded on the basis of the increase in wealth their predictions generate. They find that the incidence of bubbles decreases and the heterogeneity of predictions increases, in comparison to the results by Hommes et al. (2005a, 2008). Hommes et al. (2005b) describe a repeated strategy experiment. After participating in an introductory laboratory experiment on expectation formation participants formulate a complete forecasting strategy. These strategies are programmed and markets are simulated. Participants receive feedback from the results of these simulations and can adapt their strategy. Four rounds are played. Only about 4% of the simulations converged to the equilibrium price within 50 periods. If simulations are run for 1000 periods, 10% (first round strategies) to 40% (fourth round strategies) of the simulations converge towards the equilibrium price.

Market maker. An alternative model of price formation is one where prices are set by a *market maker*. In that scenario traders report their demands (2) to a market maker who, like the well-known Walrasian auctioneer, aggregates excess demands and increases

(decreases) the price of the risky asset when there is excess demand (supply) for the risky asset (see e.g. Beja and Goldman, 1980). That is, prices change according to

$$p_t = p_{t-1} + \lambda \left(\sum_{h=1}^{H} \frac{E_{h,t-1}\left(p_t + y_t - (1+r)p_{t-1}\right)}{a\sigma^2} - z^s \right).$$ (6)

Here $\lambda > 0$ is a parameter that measures the *speed of adjustment*.

Heemeijer et al. (2006) reports on experiments with 7 groups of $H = 6$ participants each, predicting prices for 50 periods. Parameter values are fixed such that $r = 0.05$, $a\sigma^2 = 6$, $z^s = 1$, $\lambda = \frac{20}{21}$ and $E_{h,t}(y_t) = 3 + z^s$ for all h and t. This results in:

$$p_t = \frac{20}{21}\left(3 + \overline{p}_t^e\right) + \varepsilon_t.$$ (7)

Again $\varepsilon_t \sim N\left(0, \frac{1}{4}\right)$ is a random term, representing e.g. small random fluctuations in the supply of the risky asset. The equilibrium price is $p^f = 60$. No upper limit on the price predictions was enforced (with the exception of the first period, which had to be between 0 and 100). As before, payoffs were based upon the quadratic forecasting error function (5) and the exchange rate was 2600 points for 1 euro.

The lower panels of Figure 1 show the prices and predictions in a representative group (group 1) from Heemeijer et al. (2006). As in the market clearing experiment there is no apparent convergence to the fundamental steady state although fluctuations around the steady state appear to have a lower frequency. Moreover, again participants seem to coordinate their prediction strategies quite well. The other six groups show a similar pattern (for details see Heemeijer et al., 2006).

The main features of the positive expectations feedback experiments, systematic deviation of prices from fundamentals and coordination of predictions, therefore seem to be quite robust. These results are further corroborated by Leitner and Schmidt (2007), who investigate forecasting behavior in an exchange rate model and find coordination on a common prediction strategy, as well as systematic deviations of exchange rates from their underlying fundamental values.

2.3 A comparison of guessing games and positive expectations feedback experiments

The asset pricing experiment from Heemeijer et al. (2006) is closely related to the standard formulation of the guessing game. In fact, the price generating mechanism (7) is a special case of (1) with $\alpha = \dfrac{60}{21}$ and $\beta = \dfrac{20}{21}$. The results from the two types of experiments are quite different however. In guessing game experiments choices typically convergence to the steady state within a small number of periods, whereas prices and predictions in positive expectations feedback experiments keep on fluctuating, as is obvious from Figure 1. Both findings seem to be robust.

There are several differences in the designs of the two types of experiments that may be responsible for these qualitative differences. Three important differences in design are listed below.

Structure: First, in standard guessing games (where $\alpha = 0$) typically the Nash equilibrium is on the boundary of the action space ($x^* = 0$), whereas positive expectations feedback experiments (where $\alpha > 0$) typically have an interior equilibrium with a strictly positive price for the risky asset, $p^f > 0$. Obviously, oscillations around a

boundary equilibrium are by construction impossible. On the other hand, Güth et al. (2002) argue that an interior equilibrium in a guessing game leads to *faster* convergence. One explanation for this may be that in case of an equilibrium value at the boundary convergence is only possible with coordination of choices (all players choose the equilibrium number) while uncoordinated choices scattered around an internal equilibrium can still lead to an equilibrium outcome. A second structural difference is that the *feedback strength* parameter β is much smaller in most guessing games (typically, $\beta = \frac{2}{3}$) as compared to the positive expectations feedback experiments discussed above (where $\beta = \frac{20}{21} \approx 0.95$). The highest feedback strength value in guessing games that we know of (abstracting from values of β larger than 1) is $\beta = 0.9$ in Ho et al. (1998), which nevertheless leads to results that are qualitatively similar to other guessing game experiments. Finally, in asset pricing experiments a small stochastic parameter ε_t is added in every period.

__Information:__ In guessing game experiments the game that is being played is common knowledge. Moreover, participants know the number of other players and typically even receive feedback, in each round, about the chosen numbers of these other participants. In expectations feedback experiments, on the other hand, participants only have qualitative information about the underlying game. They do not have information about the number of other players in the game, nor do they see the price predictions of these other players. The reason for not providing the participants with the price formation formula is that this remains closer to the reality of real world markets: traders typically do not know how

other traders' trading decisions are related to their expectations. Also, the information in expectations feedback experiments has an economic frame, whereas in guessing games the problem is posed as an abstract game, without any reference to investment decisions, stock prices or financial markets.

__Incentives:__ Guessing games use a winner-takes-all tournament structure, whereas expectations feedback experiments reward on the basis of prediction accuracy. This incentive structure is also based upon real life: a stock market is not a winner-takes-all situation. The only guessing game experiments departing from the tournament structure are Güth et al. (2002), Kocher and Sutter (2006) and Costa-Gomes and Crawford (2006). The former suggest that the large fraction of equilibrium choices in their experiment is due to what they call the "continuous payment scheme". Moreover, Kocher and Sutter (2006) suggest that the winner-takes-all scheme might lead single players to retire mentally, or to start experimenting. On the other hand: one could argue that a tournament structure forces participants not only to predict accurately, but to predict better than others. This may inhibit satisficing behavior and force participants to think harder about the game and make prices converge faster.

In the next section we will discuss and analyze a new set of experiments to test to which of these differences in design (structure, information or incentives) the differences in results between guessing games and asset pricing experiments can be attributed.

3. Bridging the gap: New experiments

In this section we report on three new expectations feedback studies in which parameters and experimental design are varied in an attempt to bridge the gap between guessing games and positive expectations feedback experiments. Table 1 gives an overview.

	Price Formation	*Incentives*	*Information*	*Equilibrium*	*Feedback strength*
Traditional guessing game	Market maker	Tournament	Complete	Boundary (0)	0.67
Hommes et al. (2005a) (*MC*)	**Market clearing**	**Quadratic error**	**Limited**	**Interior**	**0.95**
Heemeijer et al. (2006) (*MM*)	**Market maker**	Quadratic error	Limited	Interior	0.95
TN	Market maker	**Tournament**	Limited	Interior	0.95
TI	Market maker	Tournament	**Complete**	Interior	0.95
LOW	Market maker	**Quadratic error**	Complete	Interior	**0.67**

Table 1: Properties of the traditional number guessing game and the studies reported in the present paper. Bold entries refer to dimensions in the design that have changed in relation to the previous treatment.

Our approach is the following. Starting with the experiment from Heemeijer et al. (2006), henceforth referred to as treatment MM, we change in each new treatment one design parameter in the direction of the typical guessing game solution, in order to find the design parameter that is responsible for the difference in outcomes between guessing games and positive expectations feedback experiments.

3.1 Incentives

The pay-off function used in the first two studies is based on the quadratic forecasting error and is very flat at the optimum; small errors are barely punished. Less precision in forecasting can have large consequences in a tournament if the best competitor makes

forecasting errors of a comparable size. Tournament incentives can thus motivate participants to be more precise. On the other hand, it may demotivate those participants whose predictions are of a lower quality and who do not expect to win anyhow.

In treatment TN a tournament incentive structure (like the one typically used in guessing games) is used: in each period the participant with the smallest forecasting error receives a prize of 3 euro (in case of two or more winners the prize is split evenly). All other aspects of the design are the same as in Heemeijer et al. (2006), except that predictions of more than 1000 were not accepted by the computer program. We run 6 groups with $H = 6$ participants predicting prices for 50 periods.

Figure 2 shows predictions and prices for all periods and all groups. Recall that the equilibrium price corresponds to $p^* = 60$. The time series of predictions and prices has three important features. First, on several occasions one of the participants submits a very high price prediction. This is particularly evident in groups 1, 2 and 3. In groups 1 and 2 predictions of 999 are submitted, in group 3 a prediction of 999.9 is submitted and in groups 2 and 3 predictions of 1000 are submitted.

Also in the other three groups there are occasionally rather uncommon predictions. Obviously, these so-called "spoilers" destabilize the dynamics and inhibit convergence to the equilibrium price. We will analyze these spoilers in more depth in Section 4.4. The second feature is that there is no apparent or fast convergence to the equilibrium price. The third feature corresponds to the high degree of coordination of predictions (in the absence of spoilers). These two last features can be easily checked by inspections of the graphs for groups 4, 5 and 6, but they also hold for those periods in groups 1, 2 and 3, before the first spoiler has occurred.

Figure 2: Prices (solid line), predictions (gray line) and equilibrium price (dotted line) in 6 groups with tournament incentives and limited information. In the first 3 groups the graph is split in two and rescaled: the periods before a participant submitted a very high prediction leading to a higher price than 100, and afterwards. For clarity the equilibrium price is not displayed in the rescaled parts of the figure.

Summarizing, introduction of winner-takes-all incentives leads to an increase in "spoilers", but it does not appear to have a significant influence on convergence and coordination.

3.2 Information

In the expectations feedback experiments discussed so far the explicit price function was not available for the participants. In the typical guessing game, however, players know exactly how the target number is calculated from the reported numbers. We decided to run two sessions (six groups) in which the price function was given and explained to the participants. As in the TN treatment participants were rewarded on a winner-takes-all basis. Figure 3 shows the results.

Compared to the TN treatment the number of "spoilers" seems to be even higher. In particular, predictions of 999 or 1000 were submitted in four of the six groups, resulting in an overall decrease in the rate of convergence. For groups 2 and 3, however, it is obvious that there is little convergence, even in absence of spoilers. The "pre-spoiler" predictions in groups 4, 5 and 6 also don't show a fast convergence.

Even with a winner-take-all payment structure and complete information about the price generating mechanism the dynamics are characterized by persistent deviations from the fundamental price and coordination of individual predictions.
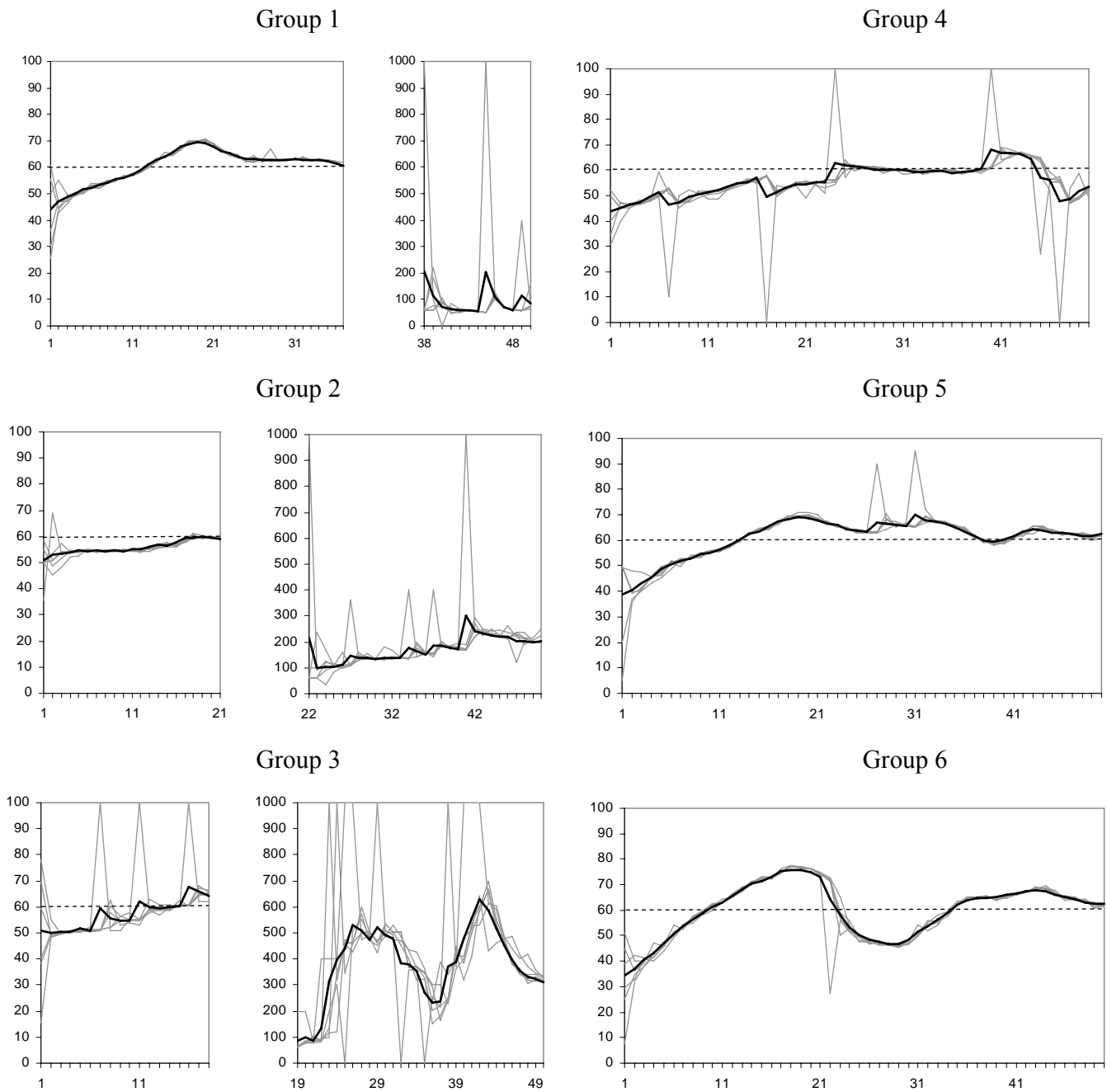
Figure 3: Prices (solid line), predictions (gray line) and equilibrium price (dotted line) in 6 groups with tournament incentives and full information. For groups 4, 5 and 6 the graph is split in two and rescaled: the periods before a participant submitted a very high prediction leading to a higher price than 100, and afterwards. In group 1 a prediction of 1000 was submitted already in period 3. For clarity the equilibrium price is not displayed in the rescaled parts of the figure.

3.3 Feedback strength

The final treatment we ran, referred to as treatment LOW, considered a change in the feedback strength. In particular, the price generating mechanism was given by:

$$p_t = \frac{2}{3}\left(30 + \bar{p}_t^e\right) + \varepsilon_t, \tag{8}$$

Note that equation (8) follows from (6) by taking parameter values $r = 0.5$, $a\sigma^2 = 6$, $z^s = 1$, $\lambda = \frac{2}{3}$ and $E_{h,t}(y_t) = 30 + z^s$ for all h and t. Also observe that (8) requires a substantial interest rate of 50%. The slope of (8) equals $\frac{2}{3}$ which corresponds to the typical value used in guessing game experiments. Participants are rewarded based upon quadratic forecasting error[16] and as in treatment TI the price function (8) was common knowledge. As in the TN and TI studies an upper limit of 1000 to the prediction was imposed after period 1.

Ho et al. (1998) provide an indication that a less steep slope could enhance convergence. They report that a higher factor (0.9 instead of 0.7) in a standard number guessing game with 7 participants causes mean choices to be farther from the equilibrium value 0. Their Figures 2A and 2C suggest that the difference is largest in the first 5 periods. It is not clear in advance whether their results will also hold in the interior equilibrium case.

Figure 4 shows results for our treatment LOW. All six groups converge very fast to the equilibrium. In group 5 one participant submitted a "spoiler" in round 25, but also prices in this group are back at the equilibrium value in only a few rounds.

---

[16] We decided to return to quadratic forecasting error incentive structure for this treatment because from treatments TN and TI it follows that the winner-takes-all incentive structure does not lead to faster convergence, but does obfuscate the results by increasing the number of spoilers.

Group 1

Group 4

Group 2

Group 5

Group 3

Group 6

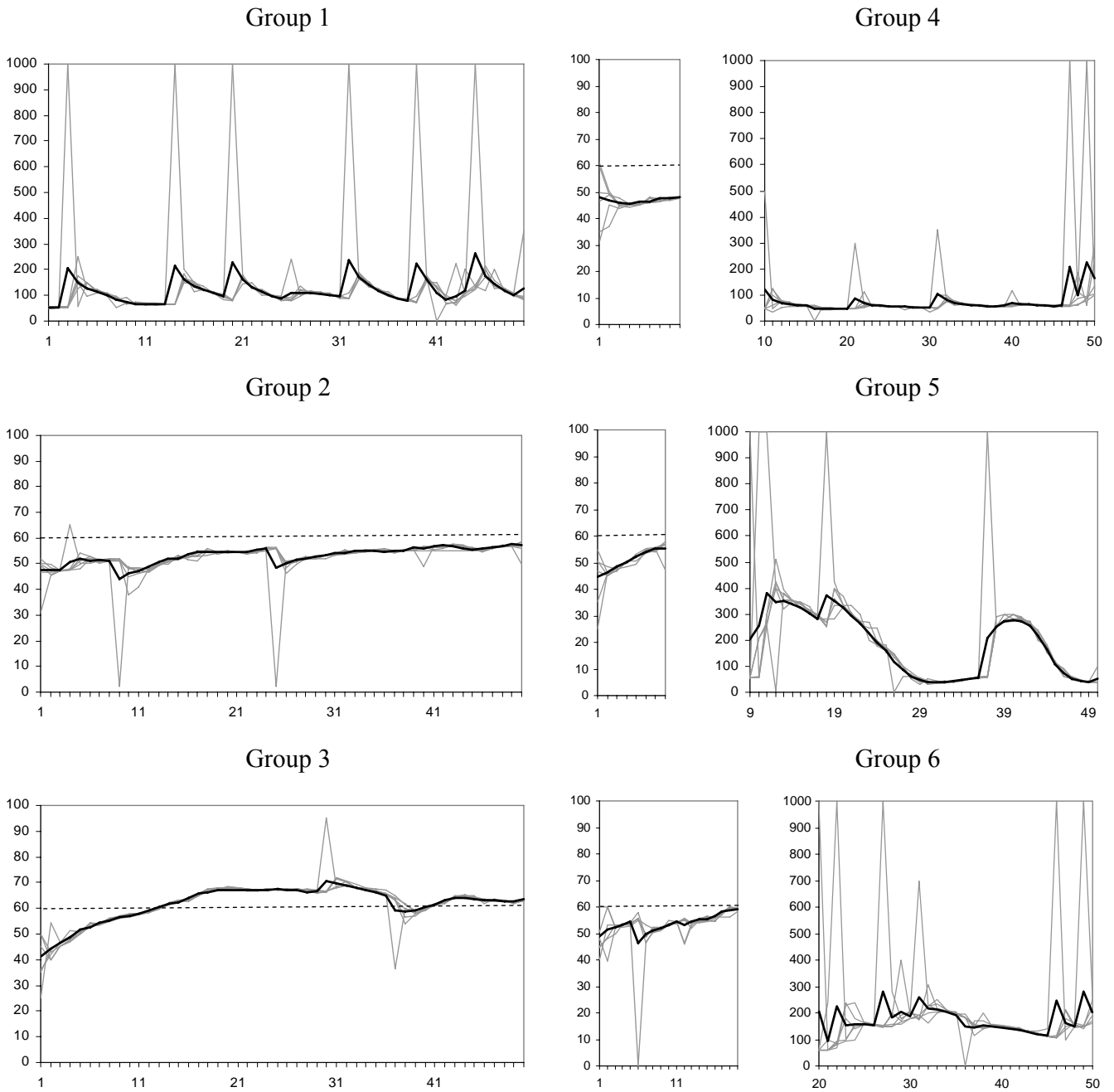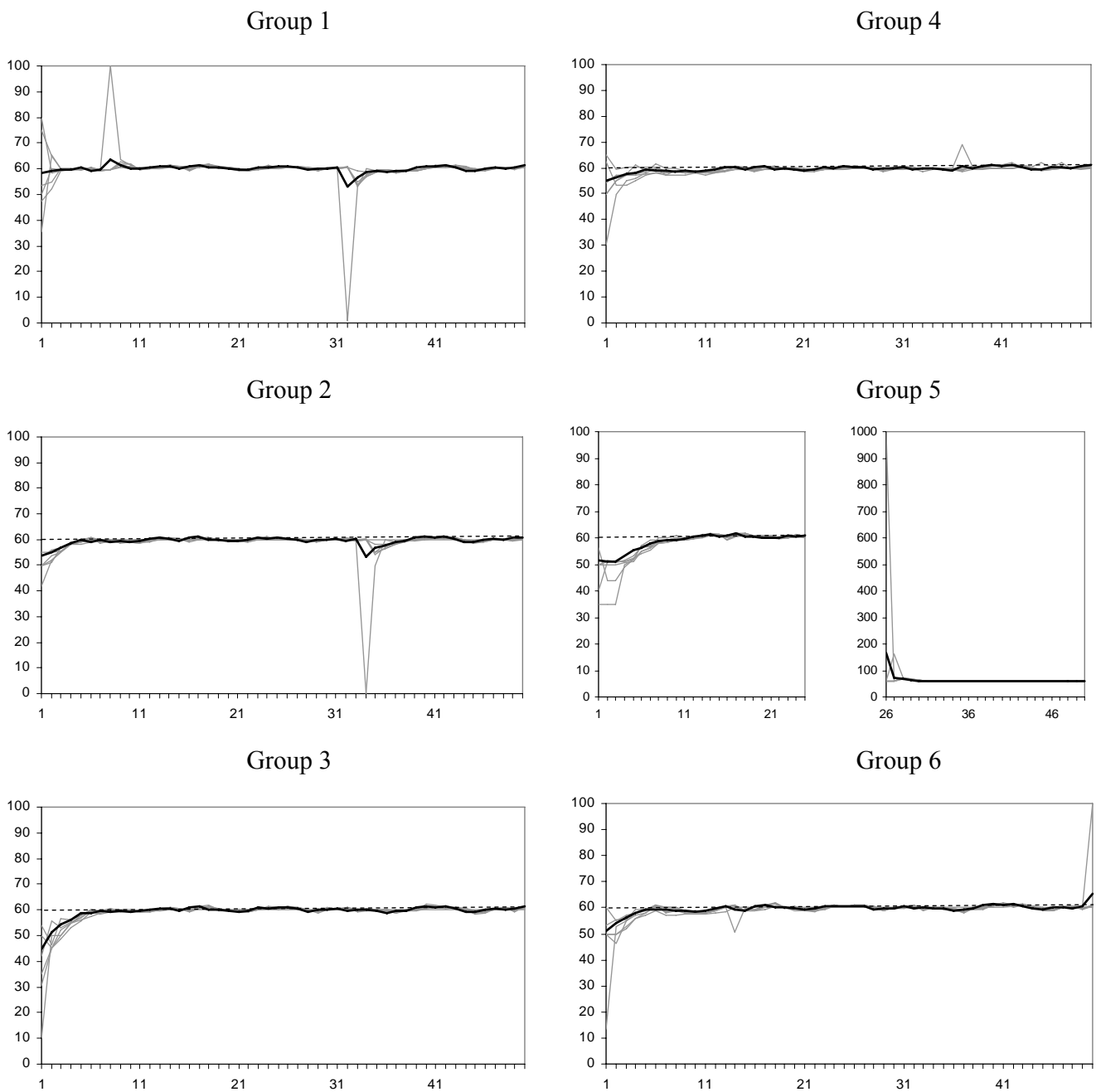**Figure 4:** Prices (solid line), predictions (gray line) and equilibrium price (dotted line) in 6 groups with tournament incentives and low feedback strength. For group 5 the graph is split in two and rescaled: the periods before a participant submitted a very high prediction leading to a higher price than 100, and afterwards.

4. Further analysis

The results described in the previous section suggest that providing more information to participants or changing to a 'winner-takes-all' incentive structure does not change the convergence and coordination properties of the positive expectations feedback experiments, but that a decrease in feedback strength is an important determinant for convergence. In this section we will try to provide some additional evidence to substantiate that claim. Moreover, in Section 4.4 we will analyze the increased incidence of spoilers in the 'winner-takes-all' treatments.

4.1 Convergence

Figures 1-4 suggest that convergence only occurs in treatment LOW. Figure 5 confirms this conjecture. It measures in each period and for each of the five experiments the median over all participants of the absolute distance between the individual prediction and the equilibrium price.

Not all data are used in creating this picture. In particular, in each period we only took those groups into account for which no spoilers had emerged yet. Obviously, a period in which one participant enters a spoiler leads to a large divergence from the equilibrium price for all predictions in that group in that period, but often there is also an effect in the following periods: the general price level increases leading to larger absolute prediction errors. Therefore from each experimental group we have only included those observations from periods before the first spoiler is submitted. Alternative procedures yield similar qualitative results. Table 5 in Appendix B gives a precise description of the data we used. Since the number of spoilers in treatments TN and TI is substantial the

number of periods with a positive number of observations in these treatments is restricted to 37 and 29, respectively.

Clearly, there is not much difference between treatments MC, MM, TN and TI, but there is a significant difference between treatment LOW and the other treatments.[17]
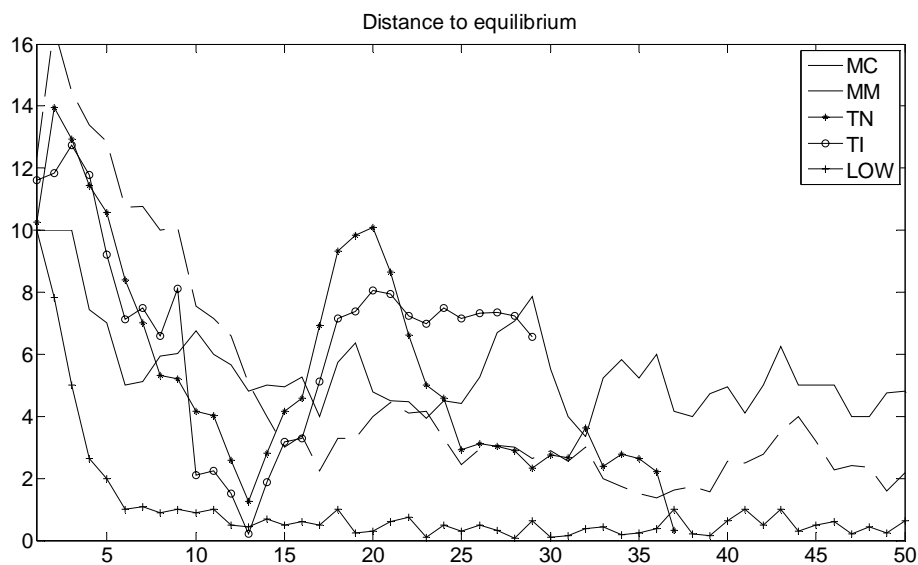

Figure 5: Distance from equilibrium. For each treatment the median of the numbers $\left| p_{h,t} - p^f \right|$ is plotted. Note that only information from "pre-spoiler" periods is used. Consequently not every point represents the same number of participants.

4.2 Prediction accuracy

Figure 6 shows, for each period, the median over all participants of the absolute value of the individual prediction error, using the same data as above. In this case the difference between treatments MM, TN, TI and LOW is not significant at a 5% level. The different designs therefore do not seem to have an impact upon prediction accuracy.

---

[17] Allowing for a small initial learning phase of 5 periods and further restricting data to the first 29 observations there is no significant difference between treatments MM, MC, TN and TI (Friedman test, 5% significance level), but there is if treatment LOW is taken into account.

Accuracy is, however, significantly lower in treatment MC, possibly reflecting the difficulty of predicting two periods ahead.
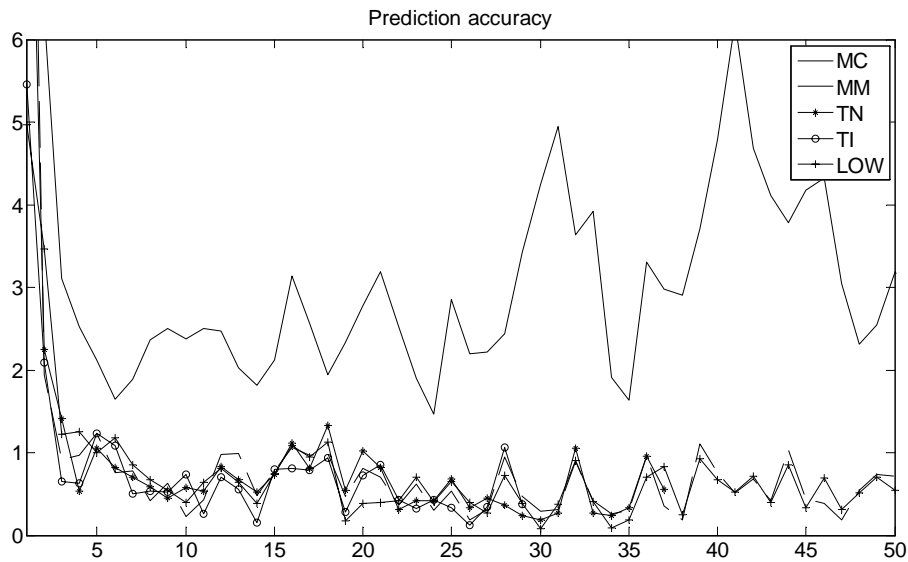
Figure 6: Prediction accuracy for participants in the different treatments. For each treatment the median of the numbers $\left| p_{h,t} - p_t \right|$ is plotted.

4.3 Coordination of expectations

Finally, Figure 7 gives, for the different treatments, the medians (over groups) of the standard deviation of predictions. A low value of this standard deviation implies a high level of coordination of predictions. Clearly, coordination seems to be much harder to obtain in the MC treatment (although the null hypothesis that there is no difference between treatments MM, TN, TI and LOW also has to be rejected at the 5% level).
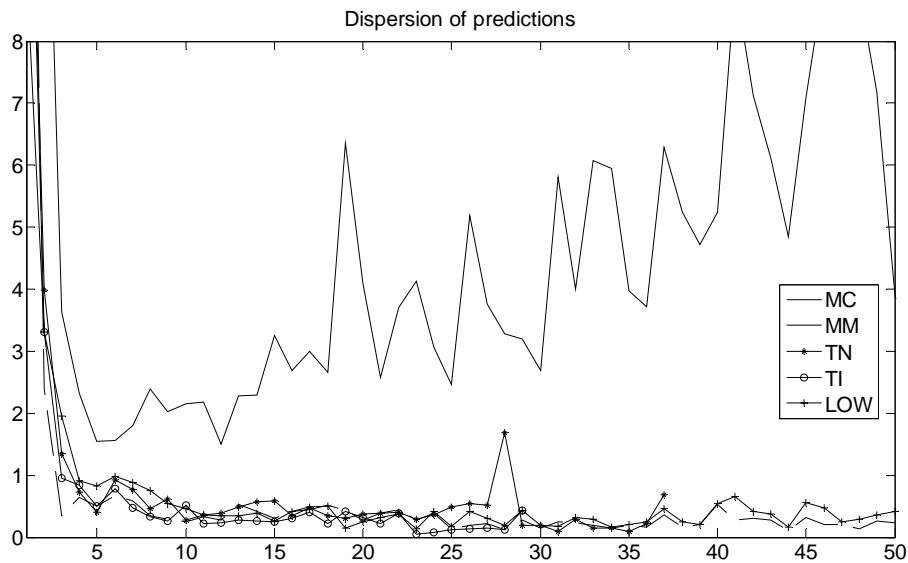
Figure 7: Dispersion of individual predictions. For each treatment the median over groups of the standard deviation of predictions is plotted.

4.4 Spoilers

For each treatment we classified the participants and periods that corresponds to spoilers (detailed information can be found in Appendix B). It is obvious that in the winner-takes-all treatments TN and TI more participants submit spoilers and spoilers are submitted in more periods. In particular, 14 participants in treatment TN and 18 participants in treatment TI (corresponding to 39% and 50% of participants in these treatments) were responsible for submitting a total of 35 and 37 spoilers, respectively (corresponding to a percentage of 1.9% and 2.1% of the total number of choices in those treatments, respectively). In the quadratic forecast error treatments MC, MM and LOW the percentage of participants submitting spoilers is 1.2%, 5.6% and 11.1% respectively, submitting spoilers in 0.0%, 0.2% and 0.3% of the total number of choices in those

treatments, respectively. The winner-takes-all payoff scheme therefore seems to have a substantial effect on the incidence of spoilers.

By construction, the average payoff of a participant in treatments TN and TI is 25 euro (50 rounds x 3 euro / 6 participants). The average total earnings of participants that submit spoilers at least once in treatments TN and TI is 95% and 99% of 25 euro, respectively, which is not significantly different at a 5% level, implying that they do not structurally perform better or worse than the other participants.

There might be several reasons why participants submit these spoilers. Participants may, for example, be frustrated with their earnings thus far, and/or they may want to increase the probability of winning in period t by disrupting the price dynamics in period $t-1$. Each period 3 euro is to be won by one of the participants in a group thus on average participants earn $3/6 = 0.50$ euro per period. The average earnings per period for "spoiling" participants before their first spoiler equal 49 cents in TN and 54 cents in TI. Moreover, the median number of periods before their first "spoiler" that a spoiling participant did not earn anything is equal to 3 for the TN treatment and equal to 3.5 for the TI treatment (notice that on average each participants should earn something (at least) every 6 periods). This suggests that 'frustration' is not the driving force behind the increased number of spoilers in treatments TN and TI.

Finally, spoilers did not prove to be exceptionally profitable: 5 of the 35 spoilers in treatment TN and 7 of the 37 spoilers in treatment TI led to positive payoffs in the next period, which is consistent with what one would expect (one success in every six periods). Moreover, in treatment TN (TI) for 5 (10) participants relative earnings after spoiling were higher than before, and for 7 (8) it was lower.

5. Conclusions

In our earlier papers on (positive) expectations feedback experiments we found very slow or no convergence to the equilibrium price. Number guessing games are very much related to expectations feedback experiments but typically show fast convergence to the Nash equilibrium. This striking discrepancy was the reason for designing three additional treatments where we searched for the driving force behind this difference. We found that changing the incentive structure to a tournament or presenting more information to the participants did not significantly change the results for the expectations feedback games; this robustness can be considered as a positive result on internal validity. On the other hand, treatment LOW shows that presenting the number guessing game in the context of a financial market, with an interior Nash equilibrium and changing the incentive structure to a quadratic error rule resulted in very fast convergence like in the traditional number guessing game, which again can be considered as a positive result on internal validity.

We found that the main driving force behind differences in convergence is the value of the feedback strength β : a change from 0.95 to 0.67 causes convergence. The next question concerns external validity: which value of the feedback strength gives a better description of the price dynamics on actual financial markets? The price function in treatment LOW is given as $p_t = 20 + \frac{2}{3} \bar{p}_t^e + \varepsilon_t$, which implies that the price is mainly driven by dividends. For treatments MM, TN and TI, on the other hand, the price function is given by $p_t = 2.85 + \frac{20}{21} \bar{p}_t^e + \varepsilon_t$, which corresponds to a financial market where prices are mainly driven by other trader's expectations and not so much by the underlying stream of dividends. This may be similar to the internet-bubble at the turn of

the century where many investors focused on capital gains and the growth in stock prices could not be explained by dividends. In hindsight we know that asset prices at that time where far away from the fundamental value. Treatment LOW (as well as the typical number guessing game) may be more relevant for the quiet periods at asset markets, when investors focus on value-stocks instead of growth-stocks. The expectations feedback experiments may be a more relevant description for speculative markets and markets for high-growth stocks.[18]

References

Allen, F., Morris, S. & Shin, H.S. (2006). Beauty contests, bubbles and iterated expectations in asset markets. *Review of Financial Studies*, 19, 719-752.

Barberis, N. & Thaler, R. (2003). A survey of behavioral finance. In: G.M. Constantinides, G.M., Harris, M. & Stulz, R.D. (eds.), *Handbook of the Economics of Finance 1*, 1053-1128, Elsevier.

Beja, A. & Goldman, M.B. (1980). On the dynamic behavior of prices in disequilibrium. *Journal of Finance*, 35, 235-248.

Biais, B. & Bossaerts, P. (1998). Asset prices and trading volume in a beauty contest. *Review of Economic Studies*, 65, 307-340.

Bosch-Domènech A., Montalvo J.G., Nagel R. & Satorra A. (2002). One, two, (three), infinity, ...: Newspaper and lab beauty-contest experiments. *American Economic Review*, 92, 1687-1701.

Bottazzi, G. & Devetag, M.G. (2005). Expectations structure in asset pricing experiments. In: T. Lux, S. Reitz & E. Samanidou (eds.), *Nonlinear Dynamics and Heterogeneous Agents. Lecture Notes in Economics and Mathematical Systems 550,* 11-26. Berlin: Springer Verlag.

Brock, W.A. & Hommes, C.H. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic Dynamics & Control*, 22, 1235-1274.

Camerer, C.F. & Ho, T.-H. (1998). The effect of incentives in experimental p-beauty contest. Caltech Mimeo.

---

[18] Our results are related to those of Hirota and Sunder (2007) who present an asset market experiment with two treatments. In the long-horizon treatment participants are in the market until the asset matures and prices indeed converge to fundamental values and are mainly determined by dividends, just as in the guessing games experiment. In the short-horizon treatment, on the other hand, participants leave the experiment before the asset has matured and prices typically do not converge to their fundamental value. In the latter case dividend payments play a minor role in the determination of asset prices, just as in our expectations feedback experiments.

Camerer, C.F., Ho, T.-H. & Chong, J.-K. (2003). Models of thinking, learning and teaching in games. *American Economic Review Papers and Proceedings*, 93, 192-195.

Camerer, C.F., Ho, T.-H. & Chong, J.-K. (2004). A cognitive hierarchy model of games *Quarterly Journal of Economics*, 119, 861-898.

Campbell, J.Y., Lo, A.W. & MacKinlay, A.C. (1997). *The Econometrics of Financial Markets*, Princeton, NJ: Princeton University Press.

Costa-Gomes, M.A. & Crawford, V.P. (2006). Cognition and behavior in two-person guessing games: An experimental study. *American Economic Review*, 96, 1737-1768.

Cuthbertson, K. (1996). *Quantitative Financial Economics: Stocks, Bonds and Foreign Exchange*. Chichester, UK: Wiley.

Duffy, J. & Nagel, R. (1997). On the robustness of behaviour in experimental 'beauty contest' games. *Economic Journal*, 107, 1684-1700.

Grosskopf, B. & Nagel, R. (2008). The two-person beauty contest. *Games and Economic Behavior*, 62, 93-99.

Güth, W., Kocher, M. & Sutter, M. (2002). Experimental 'beauty contests' with homogeneous and heterogeneous players and with interior and boundary equilibria. *Economics Letters*, 74, 219–228.

Haruvy, E., Lahav, Y. & Nouissair, C.N. (2007). Traders' expectations in asset markets: Experimental evidence. *American Economic Review*, 97, 1901-1920.

Heemeijer, P., Hommes, C., Sonnemans, J. & Tuinstra, J. (2006). Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation. *CeNDEF Working paper 06-05 University of Amsterdam.*

Hirota, S. & Sunder, S. (2007). Price bubbles sans dividend anchors: Evidence from laboratory stock markets. *Journal of Economics Dynamics & Control*, 31, 1875-1909.

Ho, T.-H., Camerer, C. & Weigelt, K. (1998). Iterated dominance and iterated best response in experimental "p-beauty contests". *American Economic Review*, 88, 947-969.

Hommes, C., 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. (2005b) A strategy experiment in dynamic asset pricing. *Journal of Economic Dynamics and Control*, 29, 823-843.

Hommes, C., Sonnemans, J., Tuinstra, J. & van de Velden, H. (2008). Expectations and bubbles in asset pricing experiments. *Journal of Economic Behavior and Organization*, in press.

Keynes, J.M. (1936) *The General Theory of Employment, Interest and Money*. New York: Macmillan.

Kocher M.G. & Sutter, M. (2005). The decision maker matters: Individual versus group behaviour in experimental beauty-contest games. *Economic Journal*, 115, 200-223.

Kocher M.G. & Sutter, M. (2006). Time is money – Time pressure, incentives, and the quality of decision-making. *Journal of Economic Behavior & Organization*, 61, 375-392.

Kocher, M.G., Sutter, M. & Wakolbinger, F. (2007). The impatct of naïve advice and observational learning in beauty contest games. *Tinbergen Institute Discussion Paper TI 2007-015/1*

Leitner, J. & Schmidt, R. (2007). Expectation formation in an experimental foreign exchange market. *Central European Journal of Operations Research*, 15, 167-184.

Morone, A. & Morone, P. (2008). Boundary and interior equilibria: What drives convergence in a 'beauty contest'? *MPRA Paper 9584.*

Moulin, H. (1986). *Game Theory for Social Sciences.* New York: New York Press.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, 85, 1313-1326.

Nagel, R. (1999). A survey on experimental beauty contest games: Bounded rationality and learning. In: Games and Human Behavior. D. Budescu, I. Erev & R. Zwick (eds.). Lawrence Erlbaum Associates, 105-142.

Sbriglia, P. (2008). Revealing the depth of reasoning in p-beauty contest games. *Experimental Economics,* in press.

Shiller, R.J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, 71, 421-36.

Shiller, R.J. (2000) *Irrational Exuberance*. Princeton, NJ: Princeton University Press.

Shleifer, A. (2000). *Inefficient Markets: An Introduction to Behavioral Finance.* Oxford: Oxford University Press.

Shleifer, A. & Vishny, R.W. (1997). The limits of arbitrage. *Journal of Finance*, 52, 35-55.

Slonim, R.L. (2005). Competing against experienced and inexperienced players. *Experimental Economics*, 8, 55-75.

Sutan, A. & Willinger, M. (2006). Why do we guess better in negative feedback situations? An experiment on beauty contest games with negative feedback. *SSRN Working Paper.*

Sutter, M. (2005). Are four heads better than two? An experimental beauty-contest game with teams of different size. *Economics Letters*, 88, 41-46.

Weber, R.A. (2003). 'Learning' with no feedback in a competitive guessing game. *Games and Economic Behavior*, 44, 134-144.

Appendix A:

Procedure and instructions
We present the procedure and a translation of the instructions for treatment MM (the instructions in treatment MC differ only in some phrasing). Boxes are included where the instructions are different in the other studies.

Procedure
A short welcoming message was read aloud from paper, after which the participants were randomly assigned to a cubicle in the computer lab. In each cubicle there was a computer, some experimental instructions on paper and some blank paper with a pen. The two treatments had different instructions. When all the participants were seated, they were asked to read the instructions on their desks. After a few minutes, they were given the opportunity to ask questions regarding the instructions, after which the experiment started. When the 50 time periods were completed, the participants were asked to remain seated and fill in the questionnaire, which was subsequently handed out to them. After a reasonable amount of time, the participants were called to the ante-room one by one to hand in the questionnaire and receive their earnings, in cash. The participants left the computer lab after receiving their earnings.

The experimental instructions the participants read in their cubicles consisted of three parts, totalling five pages. The first part contained general information about the market the experiment was about to simulate. The second part contained an explanation of the computer program used during the experiment. The third part displayed a table relating the absolute prediction error made in any single period to the amount of credits earned in that period. The conversion rate between credits and euros, being 2600 credits to 1 euro. (In **treatments TN** and **TI** a tournament was implemented and the tabel was omitted).

Experimental instructions
The shape of the artificial market used by the experiment, and the role you will have in it, will be explained in the text below. Read these instructions carefully. They continue on the backside of this sheet of paper.
General information
You are an advisor of a trader who is active on a market for a certain product. In each time period the trader needs to decide how many units of the product he will buy, intending to sell them again the next period. To take an optimal decision, the trader requires a good prediction of the market price in the next time period. As the advisor of the trader you will predict the price P(t) of the product during 50 successive time periods. Your earnings during the experiment will depend on the accuracy of your predictions. The smaller your prediction errors, the greater your earnings.
About the market
The price of the product will be determined by the law of supply and demand. Supply and demand on the market are determined by the traders of the product. Higher price predictions make a trader demand a higher quantity. A high price prediction makes the trader willing to buy the product, a low price prediction makes him willing to sell it. There are several large traders active on this market and each of them is advised by a participant of this experiment. Total supply is largely determined by the sum of the

individual supplies and demands of these traders. Besides the large traders, a number of small traders is active on the market, creating small fluctuations in total supply and demand.

About the price

The price is determined as follows. If total demand is larger than total supply, the price will rise. Conversely, of total supply is larger than total demand, the price will fall.

About the price <u>treatments TI and LOW</u>

The price in each period depends upon your prediction and the prediction of the other 5 participants. Let $GV(t)$ be the average prediction in period t, than:

$price(t) = 2.85 + 0.95 \times GV(t)$ **(treatment TI)** $\qquad\qquad$ (1)

$price(t) = 20 + 2/3 \times GV(t)$ **(treatment LOW)** $\qquad\qquad$ (1)

This is the price when only the large traders (who are advised by the six particpants) would be influencing the price. The small traders on the market cause a small change of the price, sometimes negative, sometimes positive and on average zero. We will indicate this amount in period t by $k(t)$ and it will be almost always between -1 and 1. The value of $k(t)$ is not related to this value in other periods. The realized price in period t will be:

$P(t) = price(t) + k(t)$. $\qquad\qquad$ (2)

We will give and example. The predictions of the 6 participants in period 1 are 14, 80, 76, 30, 57 and 23. The average prediction is:

$$GV(1) = \frac{14 + 80 + 76 + 30 + 57 + 23}{6} = 46.67$$

and this gives the price:

$price(1) = 2.85 + 0.95 \times 46.67 = 47.30$ **(treatment TI)**

$price(1) = 20 + 2/3 \times 46.67 = 51.11$ **(treatment LOW)**

The influence of the small traders in this first period $k(1)$ equals 0.13 and the realized price will be:

$P(1) = 47.30 + 0.13 = 47.43$ **(treatment TI)**

$P(1) = 51.11 + 0.13 = 51.24$ **(treatment LOW)**

This example and the formulas (1) and (2) show that the realized price will be near the average predicted price; if the average predicted price (GV) is low, than the realized (P) will be low and if GV is high, P will be high.

About predicting the price

The only task of the advisors in this experiment is to predict the market price P(t) in each time period as accurately as possible. The price (and your prediction) can never become negative and lies always between 0 and 100 euros in the first period. The price and the prediction in period 2 through 50 is only required to be positive. The price will be predicted one period ahead. At the beginning of the experiment you are asked to give a prediction for period 1, V(1). When all participants have submitted their predictions for the first period, the market price P(1) for this period will be made public. Based on the prediction error in period 1, P(1) - V(1), your earnings in the first period will be calculated. Subsequently, you are asked to enter your prediction for period 2, V(2). When all participants have submitted their prediction for the second period, the market price for

that period, P(2), will be made public and your earnings will be calculated, and so on, for 50 consecutive periods. The information you have to form a prediction at period t consists of: All market prices up to time period t-1: {P(t-1), P(t-2), ..., P(1)}; All your predictions up until time period t-1: {V(t-1), V(t-2), ..., V(1)}; Your total earnings at time period t-1.

About the earnings

Your earnings depend only on the accuracy of your predictions. The better you predict the price in each period, the higher will be your total earnings. On your desk is a table listing your earnings for all possible prediction errors.

For example, your prediction was 13.42. The true market price turned out to be 12.13. This means that the prediction error is: $13.42 - 12.13 \approx 1.30$. The table then says your earnings are 1255 credits (as listed in the second column).

About the earnings <u>treatments TN and TI</u>

All participants start with 5 euros and whether they will earn more will depend on the quality of their predictions. In every period the participant in your group with the smallest prediction error wins 3 euro and the others earn nothing. If more than one participant have the smallest error, the prize is split. For example, if the realized price is 34.1 and two participants predicted 31.9 and one 36.3 (and the other predictions are less accurate), all three have made a prediction error of 2.2 and the earn 3/3=1 euro each, and the other participants in the group earn nothing.

When you are done reading the experimental instructions, you may continue reading the computer instructions, which have been placed on your desk as well.

Computer instructions

The way the computer program works that will be used in the experiment, is explained in the text below. Read these instructions carefully. They continue on the backside of this sheet of paper.

The mouse does not work in this program. To enter your prediction you can use the numbers, the decimal point and, if necessary, the backspace key on the keyboard.

Your prediction can have two decimal numbers, for example 30.75. Pay attention not to enter a comma instead of a point. Never use the comma. Press enter if you have made your choice.

The available information for predicting the price of the product in period t consists of: All product prices from the past up to period t-1; Your predictions up to period t-1; Your earnings until then.

Ronde	Voor-spelling	Werke-lijke prijs
1	25.96	37.66
2	38.00	37.48
3	37.38	38.41
4	37.68	39.07
5	38.98	40.89
6	41.00	41.10
7	41.80	42.61
8	43.50	43.40
9	43.80	44.68
10	45.30	45.46
11	46.40	46.95
12	47.80	48.93
13	50.20	51.50
14	53.20	53.48
15	55.00	54.91
16	56.50	57.39
17	59.40	60.27
18	63.00	62.01

Prijs — voorspelling / werkelijke prijs — Ronde

Totale verdiensten: 21722   Verdiensten deze periode: 1274   Ronde: 19

Wat is uw voorspelling voor de volgende periode?   Voorspelling:

Een nieuwe ronde is begonnen.

The main experimental computer screen. The Dutch labels translate as follows: "prijs" = price; "voorspelling" = prediction; "werkelijke prijs" = market price; "ronde" = round; "totale verdiensten" = total earnings; "verdiensten deze periode" = earnings this period; "Wat is uw voorspelling voor de volgende periode?" = What is your prediction for the next period?; "Een nieuwe ronde is begonnen" = A new round has started.

The computer screen. The instructions below refer to this figure.

In the upper left corner a graph will be displayed consisting of your predictions and of the true prices in each period. This graph will be updated at the end of each period.

In the rectangle in the middle left you will see information about the number of credits you have earned in the last period and the number you have earned in total. The time period is also displayed here, possibly along with other relevant information.

On the right hand side of the screen the experimental results will be displayed, that is, your predictions and the true prices for at most the last 20 periods.

At the moment of submitting your price prediction, the rectangle in the lower left side of the figure will appear. When all participants have subsequently submitted their predictions, the results for the next period will be calculated.

When everyone is ready reading the instructions, we will begin the experiment. If you have questions now or during the experiment, raise your hand. Someone will come to you for assis-tance.

Appendix B

MC	# (first time)
3-1	1 (42)
1/84	**1/4200**
MM	**# (first time)**
5-2	2 (8)
6-6	1 (25)
2/42	**3/2100**
Low	**# (first time)**
1-1	2 (8)
2-6	1 (34)
5-4	1 (26)
6-2	1 (50)
4/36	**5/1800**

Table 2: "Spoilers" in treatments MC, MM and LOW. The first column gives the identity of the participant (i-j refers to participant j in group i). The second column gives the number of spoilers and (between brackets) the period of the first spoiler for this participant

TN	# (first time)	Payoff	Payoff before (# no pay)	Score
1-4	1 (49)	0.72	18 - 0.75 (2)	0
1-6	3 (38)	1.14	21 - 1.14 (2)	1
2-2	3 (22)	0.78	3 – 0.29 (17)	1
2-3	1 (23)	0.78	13.5 – 1.23 (1)	0
2-6	2 (27)	1.38	19.5 – 1.5 (2)	1
3-3	1 (23)	0.84	9 – 0.82 (4)	0
3-4	1 (25)	1.08	18 - 1.5 (5)	1
3-5	1 (24)	0.96	12 - 1.04 (1)	0
3-6	12 (19)	0.72	6 – 0.67 (4)	1
4-3	1 (45)	0.82	17.5 – 0.80 (4)	0
4-4	5 (7)	1	7 - 2.33 (2)	0
5-1	1 (27)	0.54	6 – 0.64 (9)	0
5-5	1 (31)	1.80	27 - 1.80 (0)	0
6-5	1 (22)	0.78	3 - 0.29 (4)	0
14/36	**35/1800**	**0.95**	**0.98 (med. 3)**	**5/35**

Table 3: Analysis of "spoilers" in treatment TN. The first two columns as above. The

third column gives payoffs during the whole experiment relative to expected payoffs of

25 euro, the fourth column gives the payoffs before the first spoiler absolutely, as well as

relative to expected payoffs, and (between brackets) the number of periods before the

first spoiler that nothing was earned. The last column gives the number of times earnings

were strictly positive in the period following a spoiler.

TI	# (first time)	Payoff	Payoff before (# no pay)	Score
1-3	4 (3)	0.90	0-0 (0)	0
1-4	4 (14)	1.32	3-0.46 (9)	3
1-6	2 (26)	1.32	18-1.44 (7)	0
2-2	1 (25)	0.84	15-1.25 (9)	0
2-5	1 (9)	1.08	3-0.75 (3)	0
3-3	2 (30)	0.8	9.50 – 0.66 (0)	0
4-1	2 (21)	0.84	15-0.75 (0)	0
4-2	2 (10)	1.20	12-2.67 (0)	0
4-3	2 (47)	0.66	13.5-0.59 (16)	1
4-5	1 (50)	1.26	31.50-1.29 (6)	0
5-1	2 (10)	0.70	3-0.67 (6)	1
5-2	3 (9)	1.24	6-1.5 (3)	0
5-5	1 (12)	0.96	12-2.18 (1)	0
5-6	1 (18)	0.66	3-0.35 (6)	0
6-2	1 (22)	1.38	15-1.43 (3)	0
6-4	3 (6)	0.72	0-0 (5)	0
6-5	3 (20)	0.66	10.50-1.10 (4)	0
6-6	2 (29)	1.26	15-1.03 (0)	1
18/36	**37/1800**	**0.99**	**1.08 (med. 3.5)**	**6/37**

Table 4: Analysis of "spoilers" in treatment TI. For explanation of variables see Table 3.

Treatment						
MC	**Periods**	1—41	42—51			
	Groups	1-14	1,2,4-14			
MM	**Periods**	1—7	8—24	25—50		
	Groups	1-7	1-4,6,7	1-4,7		
TN	**Periods**	1—6	7—18	19—21	22—26	27—37
	Groups	1-6	1-3,5,6	1,2,5,6	1,5	1
TI	**Periods**	1—2	3—5	5—8	9	10—29
	Groups	1-6	2-6	2-5	3,4	3
LOW	**Periods**	1—7	8—25	26—33	34—49	50
	Groups	1-6	2-6	2-4,6	3,4,6	3,4

Table 5: Data used in the construction of Figures 5—7. Note that each group consists of six participants.