The effect of strategic environment and group size in beauty contest games

Nobuyuki Hanaki* Angela Sutan† Marc Willinger‡

March 8, 2015

Abstract

Does the effect of strategic environments (strategic substitution v.s. strategic complement) on the experimental outcomes depend on the size of group? We investigated this question by varying the group size, 2-player vs 8-player, in the two versions of beauty contest games with an interior equilibrium. We replicate main finding of a previous study that shows significantly larger deviations of chosen numbers from the equilibrium under the strategic complementarity than under the strategic substitution for 8-player games. We found, however, that such a significant effect of the strategic environment disappears in 2 player games.

Keywords: beauty contest games, iterative reasoning, strategic substitution vs strategic complementarity

JEL Classification: C72, C91.

*Aix Marseille University (Aix Marseille School of Economics), CNRS, & EHESS, E-mail: nobuyuki.hanaki@univ-amu.fr

†LESSAC, Burgundy School of Business. E-mail angela.sutan@escdjon.eu

‡LAMETA, University of Montpellier. E-mail: willinger@lameta.univ-montp1.fr

1 Introduction

In a typical guessing or beauty contest game (Nagel, 1995; Ho et al., 1998), a group of players simultaneously choose a number in a given interval, and the one who has chosen the number closest to p times the mean, where $0 < p < 1$, wins a fixed prize. This class of games has been a major workforce in the development of behavioral game theory (Camerer, 2003), in particular, models that incorporate heterogeneity in depth of strategic thinking among players such as the level-k (Stahl and Wilson, 1994; Nagel, 1995) and the cognitive hierarchy model (Camerer et al., 2004).

The initial version of the beauty contest game, which was brought to the attention of economists by Keynes (1936, Ch.12), described an issue of coordination by relying on the situation where $p = 1$. Keynes conceived it as an inspiring illustration of the behavior at work within the stock market: smart traders do not try to guess what the fundamental value is but rather what every other trader believes it is, and even more smarter traders try to predict what the smart traders believe what others believe about the fundamental value, and so on. The implication is that asset prices are not directly related on their fundamental values but on the first nth-order distribution of beliefs about what others' believe, where n is the deepest level of thinking in the population of traders.

The beauty contest game can therefore be thought as a canonical model of strategic thinking in speculative markets. For instance, it has been shown that the more complex setups implemented in dynamic "learning to forecasts" experiments (Hommes et al., 2005; Heemeijer et al., 2009; Bao et al., 2012) essentially boil down to a version of repeated guessing games with unspecified target and noise (Sonnemans and Tuinstra, 2010).

Sutan and Willinger (2009) report the results of experiments on two one-shot beauty contest games (BCGs) with an interior equilibrium.[1] In their experiments, a group of 8 subjects simultaneously choose a number between 0 and 100. In one game called $BCG+$, the winner is the one who has chosen the number closest to $\frac{2}{3}(mean + 30)$ where $mean$ is the average number chosen by the subjects in the same group. In another game called $BCG-$, the winner is the one who has chosen the number closets to $100 - \frac{2}{3}mean$. The two games have a unique Nash equilibrium: iterated elimination of dominated strategies predicts that all players choose 60 in both games. However, Sutan and Willinger (2009) observe significantly more subjects in the $BCG-$ game choosing numbers closer to 60 than in the $BCG+$ game.[2]

[1] A beauty contest game with an interior equilibrium is first studied by Güth et al. (2002).

[2] Sutan and Willinger (2009) also study the version where $mean$ is defined by the average number chosen by subjects

The main difference between the two games is the nature of the strategic interactions among subjects. $BCG+$, where the target number is $\frac{2}{3}(mean + 30)$, is characterized by the presence of strategic complementarity. A subject who expects a high (low) mean should choose a large (small) number. On the other hand, $BCG-$, where the target number is $100 - \frac{2}{3}mean$, is characterized by the presence of strategic substitutability. A subject who expects a high (low) mean should choose a small (large) number. There is also a difference between the two games in the process of iterated elimination of (weakly) dominated strategies. For both games, $BCG+$ and $BCG-$, the process of iterated elimination can start at either of the two boundaries of the strategy set. However, under $BCG+$ the process converges towards the equilibrium value by one-sided elimination of dominated strategies. In contrast, under $BCG-$, the process involves alternating elimination of high and low values, and therefore proceeds by two-sided eliminations with respect to the equilibrium value. Because the elimination process is related to the nature of the strategic interactions, we shall mainly concentrate on the latter in the remained of this paper.

There exists a rationale for such difference in behaviour. Haltiwanger and Waldman (1985, 1989) theoretically demonstrate that when both naive and sophisticated agents interact, the aggregate outcome deviates from the standard equilibrium prediction much more under strategic complementarity of players' decisions than under strategic substitutability, because of the way sophisticated agents best respond to the way they believe naive agents behave. Other recent experiments (Fehr and Tyran, 2008; Heemeijer et al., 2009; Potters and Suetens, 2009; Bao et al., 2012) also report significant differences in aggregate outcomes between the two kinds of strategic environments. All these studies demonstrate that subjects' choices are much closer to the equilibrium prediction under strategic substitutability than under strategic complementarity.[3] In fact, Heemeijer et al. (2009) and Bao et al. (2012) found that subjects' choices may not converge to the equilibrium at all in the presence of a strong strategic complementarity. In this paper, we take a step forwards in trying to understand the effect of the nature of strategic interactions on such experimental outcomes.

As noted above, there are two well-known related models that incorporate heterogeneity in depth of strategic thinking among decision makers: the level-k model (Stahl and Wilson, 1994; Nagel, 1995) and the cognitive hierarchy (CH) model (Camerer et al., 2004). Note, however, while a

other than oneself in the group to obtain clearer best response dynamics. The main result of the paper, however, is robust against this change.

[3] Potters and Suetens (2009) considers the effect of strategic environment on the subjects' ability to cooperate in an efficient but non-equilibrium outcome. They found a significantly more cooperation under strategic complementarity than under strategic substitutability.

Table 1: Choices and their absolute deviation from equilibrium predicted by level-K model and Poisson cognitive hierarchy model (with the mean depth of thinking being 2).

Level-K

Game		k=1	k=2	k=3	k=4
BCG+	x	53.33	55.56	57.04	58.02
	$\lvert x-60\rvert$	6.67	4.44	2.96	1.98
BCG−	x	66.67	55.56	62.96	58.02
	$\lvert x-60\rvert$	6.67	4.44	2.96	1.98

Cognitive hierarchy model

Game		k=1	k=2	k=3	k=4
BCG+	x	53.33	54.81	55.51	55.82
	$\lvert x-60\rvert$	6.67	5.19	4.49	4.18
BCG−	x	66.67	59.26	69.75	59.84
	$\lvert x-60\rvert$	6.67	0.74	0.25	0.16

canonical level-K model does not explain the findings of Sutan and Willinger (2009), the CH model does. Table 1 shows an example of a set of choices together with their absolute deviation from the equilibrium predicted by the level-k model (left) and the Poisson CH model (right) for $BCG+$ and $BCG-$. For both models, we assume that sophisticated agents (for which $k > 1$) believe that level-0 choose 50 on average. For the Poisson CH model, in addition, we assume that the Poisson parameter (i.e. the mean depth of strategic thinking) is 2. As one can see from the left panel of the table, the level-k model predicts the same magnitude of absolute deviations from the equilibrium in $BCG+$ and $BCG-$ for all the levels. The Poisson CH model, on the other hand, predicts a much smaller deviation from the equilibrium for level above 2.[4]

The main advantage of the CH model over the level-K is the fact that agents with higher levels of strategic sophistication are best responding against the weighted average of choices made by their lower counterparts. Because of the presence of strategic substitution in $BCG-$, the choices made by lower levels (0 and 1 in the above example) are on the opposite side of the equilibrium (level 0 below and level 1 above the equilibrium). Best responding to an (weighted) average of these numbers essentially make the higher level players to best respond to a number closer to the equilibrium in $BCG-$ because the deviations of the choices made by lower level players from the equilibrium cancel out. The presence of strategic complementary in $BCG+$, on the other hand, choices made by lower levels are on the same side of the equilibrium (below the equilibrium level in the above example), and therefore cancelation of deviations cannot occur.

Assuming that cancellation of deviations is the main driving force behind the convergence of expectations towards the equilibrium under strategic substitutability, such cancellation mechanism is less likely to reach the target when the number of players is small. For instance, if there are only

[4]This difference between the prediction of level-K and CH models is robust against change in the belief about the behavior of level-0 and the Poisson parameter in CH model.

two players, a level-h player may simply assume that the opponent is a level-g ($g < h$) player and best respond, instead of best responding against a (weighted) average of choices expected from a large number of lower levels players. If this is indeed the case, the difference in the magnitude of deviations from the equilibrium between $BCG+$ and $BCG-$ reported for 8-player games by Sutan and Willinger (2009) should not be observed in 2-player version of these games.

The main hypothesis that underlies this paper is that the number of players affects the discrepancy between $BCG+$ and $BCG-$. When fewer subjects are involved, is it the case that the difference in chosen numbers becomes negligible? In order to answer this question we conduct a 2-player and a 8-player version of $BCG+$ and $BCG-$ examined in Sutan and Willinger (2009).

We replicate the main result of Sutan and Willinger (2009) for 8 players games, namely, our subjects in the 8-player $BCG-$ choose numbers that are closer to the equilibrium than in the 8-player $BCG+$ game. More importantly, however, we fail to find a significant difference in the 2 player version of $BCG-$ and $BCG+$. Thus, the significant effect of strategic environment on the outcomes in these BCGs depends on the group size. It exists for large groups, but not in a pair.

2 Experimental design

Our experimental set up is based on a 2×2 design: $(BCG+/BCG-)\times$(2 players/8 players). We use the notation BCG_n- and BCG_n+, where $n \in \{2, 8\}$. In each BCG game subjects simultaneously choose a number between 0 and 100. The subject whose choice is closest to the target number wins a fixed prize (8 euros). In case of a tie, the prize is divided equally among the winners. The target is $\frac{2}{3}(mean + 30)$ in $BCG+$ and $100 - \frac{2}{3}mean$ in $BCG-$. The unique Nash equilibrium is independent on the nature of the strategic environment and on the number of players. It is exactly equal to 60 when players' choices are restricted to belong to $[0, 100]$.

Grosskopf and Nagel (2008) and Chou et al. (2009) studied 2-player BCG for which the target number is $\frac{2}{3}mean$.[5] This 2-player BCG has a special feature that "whoever chooses the lower number wins." Therefore, it is relatively easy to realize the existence of a weakly dominant strategy in this game, i.e., to choose zero. Grosskopf and Nagel (2008) report, however, that, despite of this special feature, about 90% of their student subjects chose numbers larger than zero, thereby not realizing the special feature of the game (Chou et al., 2009). In addition, Grosskopf and Nagel (2008) found

[5]Costa-Gomes and Crawford (2006) also studies 2-player BCGs. But the most of the games they studied are asymmetric in that the strategy sets and/or the target numbers for two players differed.

that the numbers chosen in their 2-player BCG are larger than the number chosen by subjects who were involved in BCG games with group size above 3 players. According to Grosskopf and Nagel (2008) it could be due to the fact that subjects tend to ignore the strength of the influence their choice has on the mean and thus on the target number.

In our 2-player BCGs, unlike the one studied by Grosskopf and Nagel (2008), there is no obvious way of winning. There exists, however, a weakly dominant strategy in both of our 2-player BCGs. It is to choose 60. We expect, based on the finding by Grosskopf and Nagel (2008), however, that the majority of our subjects would not be aware of the weakly dominant strategy. Furthermore, it is plausible as noted by Grosskopf and Nagel (2008), that subjects are more likely to choose numbers that deviate from 60 in BCG_2 games than in BCG_8 games.

3 Results

Subjects from various disciplines were recruited from the universities of Dijon and Lyon during October 2010 and December 2014. In total 530 student subjects were involved in our experiment: 194 participated in the 2 player BCG games and 336 in the 8 players BCG games. Each subject participated only once. On average a session lasted for about 30 minutes overall.

Figure 1 shows the frequency distributions of the numbers chosen by the subjects in four treatments. The top panels report the results of the 2-player BCGs: BCG_2- (left) and BCG_2+ (right). The bottom panels show the results of the 8-player BCGs: BCG_8- (left) and BCG_8+ (right). The figure also reports the target number in each game as well as the number of participants in each of the treatment.

Comparing the top panels to the bottom panels leads to the observations that the variances of the chosen numbers are greater in 2-player BCGs than in 8-player BCGs. We will come back to this difference after we first investigate our main hypothesis: the deviations of the number chosen from the equilibrium in 2-player BCGs are not significantly different between the two strategic environments, while those in 8-player BCGs are significantly different, namely, that of BCG_8+ is larger than that of BCG_8-.

One can note in Figure 1 that the difference in the average chosen numbers between BCG_8- and BCG_8+ is larger than that between BCG_2- and BCG_2+. We will make more formal comparison below.

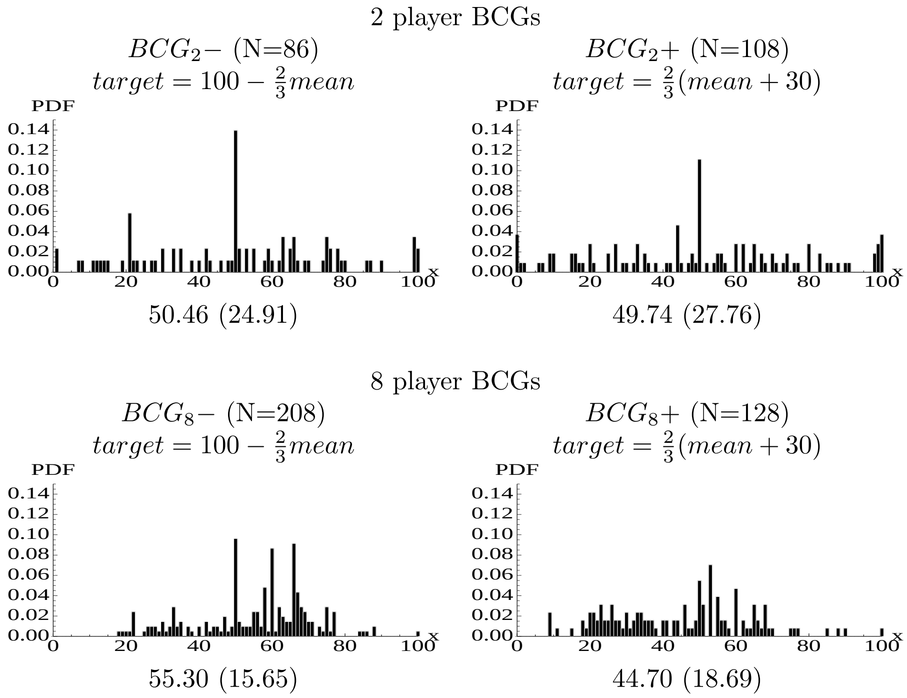

Figure 1: Frequency distribution of chosen numbers in four treatments. The numbers below each panel are the average (the standard deviation) of the numbers chosen.

Figure 2 shows the empirical cumulative distribution function (CDF) of chosen numbers (left) and their absolute deviations from the equilibrium predictions (right) for 2-player BCGs (top) and 8-player BCGs (bottom). In all the panels, the solid line represents the CDF of the $BCG-$ and the dashed line the CDF of the $BCG+$.

The top panel of the figure shows that the CDF of both the chosen numbers (left) and their deviations from the equilibrium (right) are not significantly different between the 2-player BCGs.[6] The bottom panel of the figure, on the other hand, shows that both the chosen numbers and their deviation from 60s are significantly different between two 8-player BCGs.[7] In particular, the deviations of the chosen number from 60 are larger for treatment BCG_8+ than for treatment BCG_8-. These results support our main hypothesis which we state as result 1.

Result 1 *The deviation of outcomes from the equilibrium prediction between $BCG+$ and $BCG-$ is lower for small groups ($n = 2$) than for large groups ($n = 8$).*

[6] $p = 0.75$ for the chosen numbers and $p = 0.244$ for their deviations from 60, Mann-Whitney (MW) test, two-tailed.
[7] $p < 0.0001$ for both, MW test, two-tailed

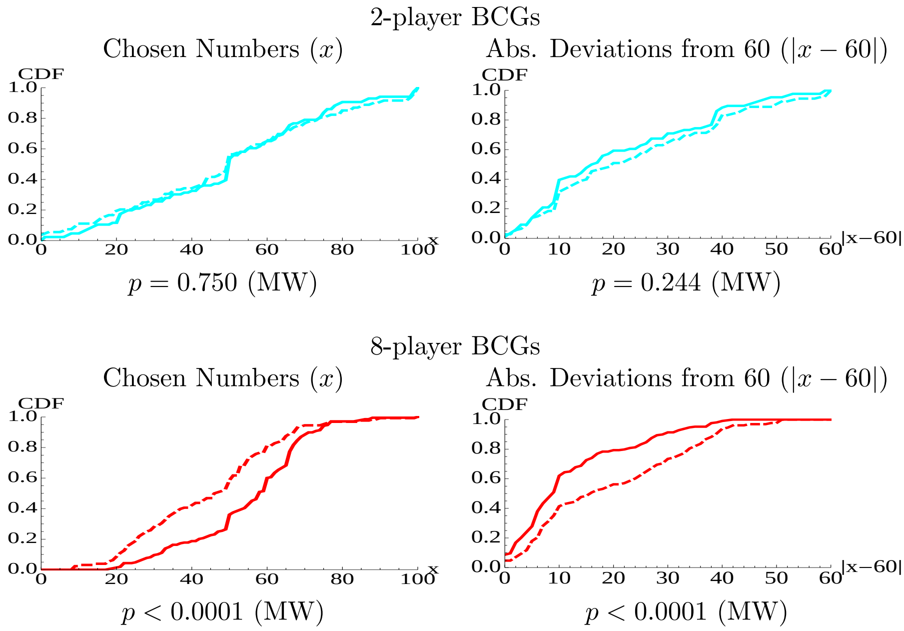

Figure 2: Empirical cumulative distribution of chosen numbers (left) and absolute deviation of chosen numbers from the equilibrium prediction (right) for 2-player (top) and 8-player (bottom) BCGs. For each panel, *BCG−* is shown in solid, *BCG+* is shown in dashed. P-values from Mann-Whitney (MW) test, two-tailed, are reported.

Now we investigate the effect of group size within each game. Figure 3 compares the CDFs of the absolute deviations of chosen numbers from 60 between 2-player (blue) and 8-player (red) BCGs. In both *BCG+* (left panel) and *BCG−* (right panel), the deviations of the chosen number from the equilibrium tend to be larger for the 2-player game than for the 8-player game. This we state as our result 2.

Result 2 *The deviation of outcomes from the equilibrium predictions is greater for small groups ($n = 2$) than for large groups ($n = 8$) in regardless of strategic environments.*

This result complements the one reported by Grosskopf and Nagel (2008) who reported, in a beauty contest game where target number is $\frac{2}{3}mean$, that subjects submitted larger numbers in 2 player game than in game with more than 3 players. The larger deviations of the chosen numbers from the equilibrium in 2-player games than in 8-player games can be a result of larger strategic uncertainty subjects face in 2-player games than in 8-player games. In 8-player games, the law of large number may operate in thinking about the average choice of others in the group, it is not

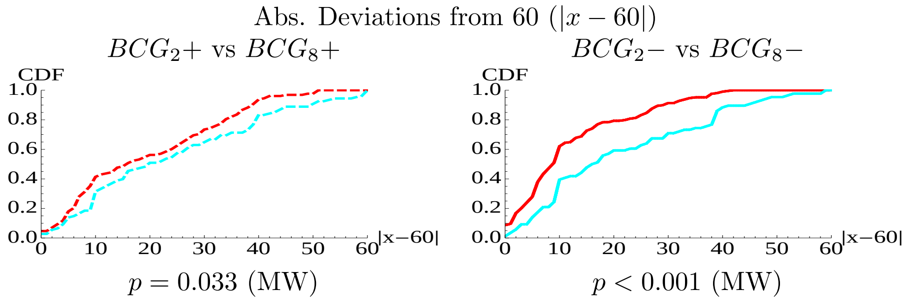

Figure 3: Comparison of the empirical cumulative distribution the absolute deviations of chosen numbers from the equilibrium prediction between 2-player (blue) and 8-player (red) BCGs for $BCG+$ (right) and $BCG-$ (left). P-values from Mann-Whitney (MW) test, two-tailed, are reported.

the case in 2-player games. This greater strategic uncertainty subjects face in 2-player games may manifests as a larger variance in the numbers of chosen by them as we have seen in Figure 1, and also as a larger deviation from the equilibrium we have just seen.

4 Discussion and conclusion

Does the effect of strategic environments (strategic substitution vs strategic complement) on the experimental outcomes depend on the size of group? We investigated this question varying the group size, 2-player vs 8-player, in the two versions of beauty contest games with an interior equilibrium first studied by Sutan and Willinger (2009). In all the games, a group of subjects simultaneously choose a number between 0 and 100. The winner is the subject who has chosen the number closest to the target. The target in one game with strategic complementarity, $BCG+$, is set as $\frac{2}{3}(mean + 30)$. In another game with strategic substitution, $BCG-$, the target is set as $100 - \frac{2}{3}mean$ where $mean$ is the average of the numbers chose by everyone in the group. The both game have the same equilibrium in which everyone chooses 60.

We have replicated the finding by Sutan and Willinger (2009) that the deviations of chosen number from the equilibrium is higher under the environment with the strategic complementarity than the one with the strategic substitution when the size of group is large (8 players). More importantly, however, we found that the significant effect of the strategic environment disappears in pairs.

We believe our finding shed a light on understanding further the reasons behind the importance of differences in strategic environments, in particular, the way experimental outcomes deviate from the equilibrium prediction. Under strategic substitution, outcomes become closer to the equilibrium prediction precisely because more sophisticated agents can safely average among the behavior of less sophisticated agents (whose deviations from the equilibrium cancel each other out) in deciding their own choices. Such a self-correcting force (at the aggregate level) is not present under strategic complementarity. The self-correcting force under strategic substation, however, operates only when the group is large enough.

References

Bao, T., C. Hommes, J. Sonnemans, and J. Tuinstra (2012): "Individual expectations, limited rationality and aggregate outcomes," *Journal of Economic Dynamics and Control*, 36, 1101–1120.

Camerer, C. F. (2003): *Behavioral Game Theory: Experiments in Strategic Interaction*, New York: Russell Sage Foundation.

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

Chou, E., M. McConnell, R. Nagel, and C. R. Plott (2009): "The control of game form recognition in experiments: understanding dominant strategy failures in a simple two person "guessing" game," *Experimental Economics*, 12, 1386–4157.

Costa-Gomes, M. A. and V. P. Crawford (2006): "Cognition and Behavior in Two-Person Guessing Games: An Experimental Study," *American Economic Review*, 96, 1737–1768.

Fehr, E. and J.-R. Tyran (2008): "Limited Rationality and Strategic Interaction: The Impact of the strategic environment on nominal inertia," *Econometrica*, 76, 353–394.

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

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

Haltiwanger, J. and M. Waldman (1985): "Rational Expectations and the Limits of Rationality: An Analysis of Heterogeneity," *American Economics Review*, 75, 326–340.

——— (1989): "Limited Rationality and Strategic Complements: The implications for macroeconomics," *Quarterly Journal of Economics*, 104, 463–484.

Heemeijer, P., C. Hommes, J. Sonnemans, and J. Tuinstra (2009): "Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation," *Journal of Economic Dynamics and Control*, 33, 1052–1072.

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

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

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

Nagel, R. (1995): "Unraveling in Guessing Games: An Experimental Study," *American Economics Review*, 85, 1313–1326.

Potters, J. and S. Suetens (2009): "Cooperation in experimental games of strategic complements and subsitutes," *Review of Economic Studies*, 76, 1125–1147.

Sonnemans, J. and J. Tuinstra (2010): "Positive expectations feedback experiments and number guessing games as models of financial markets," *Journal of Economic Psychology*, 31, 964–984.

Stahl, D. O. and P. W. Wilson (1994): "Experimental evidence on players' models of other players," *Journal of Economic Behavior and Organization*, 25, 309–327.

Sutan, A. and M. Willinger (2009): "Guessing with negative feedback: An experiment," *Journal of Economic Dynamics and Control*, 33, 1123–1133.