

The strategic environment effect in beauty contest games

Nobuyuki Hanaki, Angela Sutan, Marc Willinger

▶ **To cite this version:**

Nobuyuki Hanaki, Angela Sutan, Marc Willinger. The strategic environment effect in beauty contest games . 2016. halshs-01294915

HAL Id: halshs-01294915
https://shs.hal.science/halshs-01294915

Preprint submitted on 30 Mar 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The strategic environment effect in beauty contest games*

Nobuyuki Hanaki† Angela Sutan‡ Marc Willinger§

March 14, 2016

Abstract

Recent experimental studies have shown that observed outcomes deviate significantly more from the Nash equilibrium when actions are strategic complements than when they are strategic substitutes. This "strategic environment effect" offers promising insights into the aggregate consequences of interactions among heterogeneous boundedly rational agents, but its macroeconomic implications have been questioned because the underlying experiments involve a small number of agents. We studied beauty contest games with a unique interior Nash equilibrium to determine the critical group size for triggering the strategic environment effect. We show theoretically that the effect operates for interactions among three or more agents. Our experimental results partially support this theory, showing a statistically significant strategic environment effect for groups of five or more agents. Our findings establish that experiments involving a small number of interacting agents can provide major insights into macro phenomena and bolster previous work done on such issues as price dynamics.

Keywords: beauty contest games, iterative reasoning, strategic substitutability, strategic complementarity

JEL Classification: C72, C91.

*We thank John Duffy, Cars Hommes, Rosemarie Nagel, Charles Noussair, Ali Ihsan Ozkes, and participants at WEHIA 2015 and ASFEE 2015, seminars at Kyoto and Waseda for their comments and suggestions. A part of this research is financed by the Conseil Regional de Bourgogne under "PARI" grants and an ORA-Plus project "BEAM" (ANR-15-ORAR-0004).

†GREDEG, Université de Nice Sophia Antipolis, Skema Business School, Université de Cote d'Azur, and IUF. E-mail: Nobuyuki.HANAKI@unice.fr

‡Univ. Bourgogne Franche-Comté, ESC Dijon-CEREN, and LAMETA, University of Montpellier. E-mail: angela.sutan@escdjon.eu

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

1 Introduction

To what extent does non-rational behavior by (some) individuals affect aggregate outcomes? This question has regularly attracted the attention of leading scholars (see, for example, Becker, 1962; Conlisk, 1996; Brock and Hommes, 1997; Fehr and Tyran, 2005, and references cited therein). More recently, experimental and empirical research has shown that people do not behave as rationally as often assumed in economic theory.[1] This accumulated evidence has started to influence theoretical developments and there is now a rise in analyses based on "boundedly rational" agents in fields such as game theory, industrial organization, and finance.[2]

Despite these developments, many economists are still skeptical about the usefulness of explicitly considering the effects of bounded rationality when it comes to analyzing aggregate outcomes such as macroeconomic phenomena. One of the reasons for this skepticism is the belief held by many economists that can be summarized by an old statement from Gary Becker: "households may be irrational and yet markets quite rational" (Becker, 1962, p.8). That is, the deviation from rational behavior by many boundedly rational individuals will cancel each other out when we consider aggregate phenomena, thus, bounded rationality at individual or household level does not matter much at the aggregate level. Indeed, Gode and Sunder (1993, 1997) show that even experimental markets consisting of zero-intelligence computer traders can exhibit high allocative efficiency when these zero-intelligence traders must operate under their respective budget constraints.

However, other theoretical studies have shown that the existence of a few boundedly rational agents in a large population can have a larger-than-proportional impact on aggregate outcomes. Akerlof and Yellen (1985a,b) and Russell and Thaler (1985), for example, show instances when the existence of non-optimizing agents whose loss from non-optimization may be very small can still have a large impact on equilibrium outcomes. De Long et al. (1990) show that irrational noisy traders who take a large amount of risk can generate significant mispricing in the asset market and earn higher expected returns than rational investors. Haltiwanger and Waldman (1985, 1989, 1991) demonstrate that the behavior of boundedly rational agents can have a large, i.e., more then proportional to their population share, influence on the aggregate outcomes when the environment

[1] There are now many references for this including some very popular books such as Ariely (2008) and Kahneman (2011). An early collection of so called "anomalies" in human behavior from the point of view of economic theory can be found in Thaler (1992).

[2] For example, Camerer (2003) is a comprehensive summary of behavioral game theory, Shleifer (2000) is a nice introduction to behavioral finance, and Spiegler (2011) provides an overview of the growth of this type of research in the field of industrial organization.

is characterized by strategic complementarity.

In this paper, we follow up on the theoretical results from Haltiwanger and Waldman (1985, 1989, 1991). In these analyses, they considered two types of agents, naive and sophisticated, and they showed that the aggregate outcome deviates more from the Nash or rational expectations equilibrium in environments where agents' actions are strategic complements than in environments where they are strategic substitutes. We call this phenomena "the strategic environment effect," a term that will be used in the rest of the paper.

The underlying explanation for the strategic environment effect is the manner in which sophisticated agents best respond to the way they believe naive agents behave. In the presence of strategic complementarity, sophisticated agents have an incentive to mimic what they believe naive agents will do and therefore amplify the deviations from the equilibrium caused by naive agents, while in presence of strategic substitutability they have an incentive to act in the opposite way and thus offset the deviations from the equilibrium caused by naive agents.

Several recent experiments provide support for the strategic environment effect in various contexts. Fehr and Tyran (2008) studied price dynamics after a nominal shock in price-setting games. They found that the speed of adjustment to the new Nash equilibrium was much slower under strategic complementarities than under strategic substitutabilities. Heemeijer et al. (2009) and Bao et al. (2012) studied the strategic environment effect in the framework of "learning-to-forecast" experiments (Hommes et al., 2005). In "learning-to-forecast" experiments, the subjects' task is to repeatedly forecast the price of an asset with the knowledge that the forecasts, including their own, determine the price they are forecasting. Although subjects are not informed of the exact relationship between their forecasts and the resulting price, both Heemeijer et al. (2009) and Bao et al. (2012) observed that the price forecasts and the resulting price both converge very quickly to the rational expectations equilibrium (REE) price under strategic substitutability. In contrast, under strategic complementarity, the forecasts and the resulting price often do not converge to the REE price, but instead follow large oscillations and exhibit patterns that are reminiscent of bubbles and crashes. Potters and Suetens (2009) considered the strategic environment effect on subjects' ability to cooperate in an efficient but non-equilibrium outcome in duopoly games. They report significantly more cooperation under strategic complementarity than under strategic substitutability.

In a similar line of research, Sutan and Willinger (2009) experimentally studied two different one-

shot beauty contest games (BCGs) with interior equilibria.[3] In both experimental games, a group of 8 subjects had to simultaneously choose a number between 0 and 100. In the game called $BCG+$ which involved strategic complementarity, the winner was the subject who chose the number closest to $\frac{2}{3}(mean + 30)$ where $mean$ is the average number chosen by all other subjects (excluding oneself) of the group. In the second game called $BCG-$, which involved strategic substitutability, the winner was the one who chose the number closest to $100 - \frac{2}{3}mean$, where $mean$ is defined identically. The two games have the same unique Nash equilibrium: iterated elimination of dominated strategies predicts that all players choose 60 in both games. However, Sutan and Willinger (2009) observed significantly more subjects in $BCG-$ choosing numbers closer to 60 than in $BCG+$.[4] Unlike the above mentioned experiments, where subjects played the game repeatedly, subjects played a BCG once in the experiments by Sutan and Willinger (2009). Their results, therefore, suggest that the strategic environment effect operates when subjects are carrying out some kind of introspective strategic reasoning.

While these experimental findings quite convincingly document the existence of a strategic environment effect, i.e. larger deviations of observed outcomes from the Nash or the rational expectations equilibrium under strategic complementarity than under strategic substitutability, their robustness as well as their implications for macro phenomena are often questioned because these experimental results are based on interactions between a relatively small number of subjects. Indeed, as Duffy (2016) notes, "small numbers" is the most often raised concern when one tries to make inferences about macroeconomic phenomena based on the results obtained from a laboratory experiment. In the above-mentioned experimental studies, the sizes of groups were 2 in Potters and Suetens (2009), 4 in Fehr and Tyran (2008), 6 in Heemeijer et al. (2009) and Bao et al. (2012), and 8 in Sutan and Willinger (2009). Because these studies also differ in many other respects, it is hard to obtain a clear picture of what drives the main result, although all of them evoke what we have defined as the strategic environment effect.

In order to better understand how the strategic environment effect operates, we designed a novel experiment based on the hypothesis that the difference in the observed deviation from the

[3] In a typical guessing or beauty contest game (Nagel, 1995; Ho et al., 1998), a group of players simultaneously choose a number from within a given range, and the one who has chosen the number closest to $p \times$ mean, where $0 < p < 1$ and mean is the mean of the numbers chosen by everyone, wins a fixed prize. By changing the target number to be $p \times \text{mean} + c$ where $0 < c \leq 100$ and $0 < p < 1$ or $-1 < p < 0$, one can obtain a beauty contest game with an interior equilibrium. The first to experimentally study a beauty contest game with an interior equilibrium were Güth et al. (2002).

[4] Sutan and Willinger (2009) also study the version where $mean$ is defined by the average number chosen by all the subjects in the group including oneself. The main result of the paper, however, is robust against this change.

equilibrium prediction between two strategic environments is dependent on the size of the group of players. The intuition behind such a population size effect is rather simple. Recall that the sources of the strategic environment effect are the existence of agents who are heterogeneous in their degree of strategic sophistication and the manner in which more sophisticated subjects best-respond to the way they believe their less sophisticated counterparts will behave. In the presence of strategic substitutability of actions, the average behavior of less sophisticated agents (who are heterogeneous in their depth of strategic thinking among themselves) will not deviate much from the equilibrium, while it will in the presence of strategic complementarity of actions. However, for this reasoning to operate, the size of the group needs to be large enough so that more sophisticated players can safely consider an "average behavior" of their less sophisticated counterparts.

Thus, by systematically varying population size, we theoretically and experimentally study the two versions of the one-shot beauty contest game with interior equilibria, $BCG+$ and $BCG-$, that were previously studied by Sutan and Willinger (2009). We focus on beauty contest games because this class of games has been an important tool 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 (CH) model (Camerer et al., 2004). In addition, a beauty contest game can be seen as a canonical model of strategic thinking in speculative markets as first brought to the attention of economists by Keynes (1936, Ch.12).[5] Furthermore, the more complex setups implemented in dynamic "learning-to-forecast" experiments mentioned above (Hommes et al., 2005; Heemeijer et al., 2009; Bao et al., 2012) essentially boil down to a version of repeated beauty contest games with noise in which subjects are not informed about exactly how the target is defined (Sonnemans and Tuinstra, 2010). Finally, given the constant sum nature of beauty contest games, we can abstract away from issues related to subjects trying to coordinate on a Pareto-efficient outcome, which has been studied in the context of oligopoly games by Huck et al. (2004), Potters and Suetens (2009), and Friedman et al. (2015) among others.

Our main hypothesis, which is set forth in Section 2, is as follows: *When the population size is small, the strategic environment effect will be weak or non-existent. However, the strategic environ-*

[5]Keynes conceived a beauty contest as an inspiring illustration of the behavior at work within the stock market: smart traders do not try to guess what the fundamental value of a stock is, but rather what every other trader believes it is, and even smarter traders try to predict what the smart traders believe others believe about the fundamental value, and so on. The implication is that asset prices are not directly related to their fundamental values but to the first kth-order distribution of beliefs about what others believe, where k is the deepest level of thinking in the population of traders.

ment effect becomes significant if the size of the population is large enough.

Based on our experiment involving more than 1000 subjects, we find a significant strategic environment effect for groups of 5 or more, but not for groups of smaller sizes. More precisely, in groups of 5 or more we observe a larger deviation from the Nash equilibrium prediction under $BCG+$ than under $BCG-$, but not in smaller size groups. Our experimental test of the strategic environment effect is quite strong because it allows for both between-subject and within-subject comparisons. Therefore, the impact of the strategic environment effect on outcomes that was reported in earlier experiments involving relatively small group sizes is robust against an increase in group size but not against a decrease.

Our findings support the fact that experimental results, even if they are based on a relatively small number of interacting players, can provide major insights into macro phenomena. This bolsters work done in earlier studies on such issues as price dynamics observed in financial markets (Hommes et al., 2005; Heemeijer et al., 2009; Bao et al., 2012) and nominal rigidity (Fehr and Tyran, 2008).

The rest of the paper is organized as follows: Section 2 establishes the main hypothesis that will be tested in this paper; Section 3 describes the experimental design; Section 4 summarizes the results of the experiment; and Section 5 offers a summary and concluding remarks.

2 Theoretical predictions

In this section, we will outline how our main hypotheses were derived. We first consider the level-K model (Stahl and Wilson, 1994; Nagel, 1995) and the cognitive hierarchy (CH) model (Camerer et al., 2004), which will be later extended to allow for imperfect best response *à la* the noisy introspection model (Goeree and Holt, 2004) and the truncated heterogeneous quantal response model (Rogers et al., 2009).

In the beauty contest game (BCG) n players ($n \geq 2$) simultaneously choose a number between 0 and 100. The player whose chosen number is closest to the target number wins a prize. In the case of a tie, one of the winners is randomly selected to receive the prize. We consider two variants of the game, $BCG+$ and $BCG-$. In $BCG+$, the players' actions are strategic complements whereas in $BCG-$ their actions are strategic substitutes. In order to equalize the slopes of the best response functions in both games, and to avoid any influence of players' choices on the target number, we set the target as the average number chosen by all players in the group excluding a player's own

choice.[6] Namely, in $BCG+$, the target for player i, T^i_{BCG+}, is defined as

$$T^i_{BCG+} = 20 + \frac{2}{3}\frac{\sum_{j\neq i} a_j}{n-1}. \tag{1}$$

where a_j is an integer chosen by player j. Similarly, the target for player i in $BCG-$, T^i_{BCG-}, is defined as

$$T^i_{BCG-} = 100 - \frac{2}{3}\frac{\sum_{j\neq i} a_j}{n-1}. \tag{2}$$

The unique Nash equilibrium in both games is that all players choose 60. The Nash equilibrium neither depends on the nature of the strategic environment nor on the number of players.[7]

2.1 Level-K and cognitive hierarchy models

Alternative predictions are obtained if the following two assumptions are relaxed: (i) all players have infinite depth of reasoning; and (ii) this fact is common knowledge. Two well-known models incorporate heterogeneity in depth-of-strategic thinking among players: the level-K model (Stahl and Wilson, 1994; Nagel, 1995), and the cognitive hierarchy model (Camerer et al., 2004).

Table 1 reports the different predictions for these two models in a specific case. We report what the level-K and the Poisson CH model predict, in terms of both the number chosen and its absolute deviation from Nash equilibrium, for each level of player for $BCG+$ and $BCG-$.[8] For both models, we assume that sophisticated players (i.e. $k > 1$) believe that level-0s randomly choose a number from $[0, 100]$ with uniform probability, which implies that the average choice made by level-0s is equal to 50. For the Poisson CH model, we furthermore assume that the mean depth of strategic thinking is equal to 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 Nash prediction in $BCG+$ and $BCG-$ for all the considered levels of k. In contrast, the Poisson CH model predicts a smaller deviation from the

[6] Thus, in the case of 2-player games, the target number depends simply on the number chosen by the opponent. See Appendix A for details.

[7] Our explanation of the target number in the instructions given to subjects for the $BCG+$ game, which is $20 + \frac{2}{3}\frac{\sum_{j\neq i} a_j}{n-1}$ is different from the one used by Sutan and Willinger (2009), i.e., $\frac{2}{3}\left(\frac{\sum_{j\neq i} a_j}{n-1} + 30\right)$. We made this change to make the explanations of the target number in $BCG+$ and $BCG-$ as symmetrical as possible. In addition, in our experiment, one of the winners was chosen randomly in the case of a tie, while in Sutan and Willinger (2009), winners received an equal share of the prize. This change was made to avoid the possibility that might especially arise in $n = 2$ games that two subjects opt to choose a focal number to share the prize between them.

[8] In the Poisson CH model, the frequencies of players with various levels (0, 1, 2, ...) in the population is assumed to follow a Poisson distribution.

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
	$\|x-60\|$	6.67	4.44	2.96	1.98
$BCG-$	x	66.67	55.56	62.96	58.02
	$\|x-60\|$	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
	$\|x-60\|$	6.67	5.19	4.49	4.18
$BCG-$	x	66.67	59.26	69.75	59.84
	$\|x-60\|$	6.67	0.74	0.25	0.16

Nash prediction for $BCG-$ for level-2 and above.[9]

The difference in predictions between the level-K model and the Poisson CH model is due to the fact that in the CH model, a player of level k best responds to the weighted average of choices made by players of lower levels, i.e. between 0 and $k-1$. Consider for instance a level-2 player. In an environment where players' choices are strategic substitutes ($BCG-$), the average choices made by level-0 and level-1 are on the opposite side of the Nash equilibrium: average choices by level-0s are below and average choices by level-1s are above the equilibrium in the example shown in Table 1. Best responding to the weighted average of these numbers essentially leads level-2 players to best respond to a number that is close to the Nash equilibrium, because in $BCG-$ the deviations from the Nash equilibrium of the choices made by level-0 players and level-1 players cancel each other out. In contrast, when players' choices are strategic complements ($BCG+$), choices made by lower levels are on the same side of the equilibrium (e.g. below the equilibrium level in the above example), and therefore cancelation of deviations by lower level players does not occur. Note that this logic does not operate in the level-K model because level-2 players are only best responding to the choices made by level-1 players. The insight gained from the CH model is essentially the same as the one offered by Haltiwanger and Waldman (1985) in their two-type model, but here the logic has been extended to include more types.

If the cancellation of deviations by lower levels is indeed the main driving force behind the convergence of choices towards the equilibrium under strategic substitutability, such a force is less likely to operate when the number of players is small. For instance, if only two players are involved in the game, a level-2 player may simply assume that the opponent is a level-0 player and best respond

[9]This difference between the prediction of the level-K and the CH models is robust, at least qualitatively, against change in the belief about the behavior of level-0 players and the Poisson parameter in the CH model, except when level-0 players are assumed to choose 60 on average.

to this assumption instead of best responding against a (weighted) average of choices expected from a large number of lower-level players.

2.2 Allowing noisy best responses

The level-K and the CH models, however, are based on the strong assumption of *perfect best response*, i.e., each player chooses a number that is a best response to his or her belief about the average of others' choices.[10] Relaxing the perfect best response assumption leads to the consideration of weaker forms of best response in which players' choices depend *proportionally* on their expected payoffs: options with higher expected payoffs are chosen with higher probabilities. It has been shown that models that rely on imperfect best response, or so-called "better response" assumptions (Rogers et al., 2009), provide a more suitable fit to the experimental outcomes than models assuming perfect best response (see, among others, McKelvey and Palfrey, 1995; Goeree and Holt, 2001; Rogers et al., 2009; Breitmoser, 2012). The assumption of perfect best response in level-K and CH models has therefore been relaxed in favor of the assumption of "better response" in the noisy introspection (Goeree and Holt, 2004) and the truncated heterogeneous quantal response (Rogers et al., 2009) models, respectively. Indeed, Goeree et al. (2014) showed that the noisy introspection model predicts the experimental outcomes much better than the level-K model in games that extend the 11-20 Money Request Game proposed by Arad and Rubinstein (2012). Similarly, Breitmoser (2012) showed that the noisy introspection model fits better the data of the beauty contest experiments compiled by Bosch-Domènech et al. (2002) than the level-K model.

Let us check whether the insight we have gained from analysis of the level-K and the Poisson CH models is robust against the introduction of imperfect best response by considering a logistic level-K (LLK) model and a logistic cognitive hierarchy (LCH) model. Both the LLK and the LCH models assume that: (i) players are heterogeneous in their depth of strategic thinking; and (ii) they do not perfectly best respond to their beliefs about choices of others. The difference between the two models is similar to the one between the level-K model and the CH model. In the LLK model, a level-k player believes that everyone else is of level-$(k-1)$. In LCH, on the other hand, a level-k player assumes that others are of a lower level, i.e., between level-0 and level-$(k-1)$. In this section, for the LCH model we assume that players of each level believe that others are distributed according

[10] However, Breitmoser (2012) showed that choosing the defined target number based on the average choices of others is generally not a best response against opponents who are randomizing in beauty contest games.

to the (truncated) Poisson distribution just as in the CH model.[11]

In order to facilitate the presentation and numerical computations of the LLK and LCH models, we assume that players (in $BCG+$ and $BCG-$) can only choose integer numbers between 0 and 100.

In both models, a level-0 player chooses an integer between 0 and 100 with uniform probability. Level-1 players assume that level-0 players choose randomly with uniform probability and compute the expected payoff for each of the integers between 0 and 100. Let $E_1(\pi(a))$ be the expected payoff for a level-1 player if he or she chooses integer a. Then, the level-1 player chooses an integer s according to the probability

$$P_1^i(s) = \frac{e^{\lambda_1 E_1(\pi(s))}}{\sum_a e^{\lambda_1 E_1(\pi(a))}} \tag{3}$$

where λ_1 is a parameter that governs the sensitivity of level-1 players' choices to the variation in their expected payoffs. If $\lambda_1 = 0$ the level-1 player chooses just like a level-0 player, i.e., $P_1^i(s) = 1/101$ whatever the number s considered. If $\lambda_1 \to \infty$ the probability distribution becomes degenerate and players select the integer with the highest expected payoff with probability one like in the level-K or the CH models, i.e., players tend to perfect best reply.

The difference in beliefs about others' choices between the LLK model and the LCH model affects players' choice for level-2 and above ($k \geq 2$). Let us consider the LLK model first. A level-k player believes that all other players are of level-$(k-1)$, and level-$(k-1)$ believes that all others are of level-$(k-2)$, etc. Thus, for example, a level-2 player computes the expected payoff of choosing an integer a, $E_2^{LLK}(\pi(a))$, based on the assumption that others will choose a number, s, according to the probability $P_1^j(s)$ defined in Eq. 3. Given the expected payoff $E_2^{LLK}(\pi(a))$, the level-2 player i chooses integer l according to the probability

$$P_2^{i,LLK}(l) = \frac{e^{\lambda_2 E_2^{LLK}(\pi(l))}}{\sum_s e^{\lambda_2 E_2^{LLK}(\pi(s))}}. \tag{4}$$

For players with a higher value of k, the choice probabilities are defined similarly in an iterative manner. The sensitivity of choices to the expected payoffs, λ_k, can differ across various ks.[12]

[11] Our LLK and LCH models are very closely related to the noisy introspection model proposed by Goeree and Holt (2004) and the truncated heterogeneous quantal response model proposed by Rogers et al. (2009), respectively. Our LLK and LCH models differ from these two models in the way the sensitivity of the choices to the expected payoffs are modeled across different levels.

[12] For example, in the noisy introspection(Goeree and Holt, 2004) or the truncated heterogeneous quantal response model (Rogers et al., 2009), λ_k increases by factor μ for each k.

However, in our analyses below, we will assume $\lambda_k = \lambda > 0$ for all $k > 0$.

Let us turn now to the LCH model. A level-k ($k \geq 2$) player believes that others are of a lower level, between level-0 and level-$(k-1)$. Let $pr(k)$ be the probability for a player to be of level-k. Just as in the Poisson CH model, we assume that $pr(k)$ follows a Poisson distribution. In the LCH model, a level-2 player believes the others are either level-0 or level-1, with probability $\frac{pr(0)}{pr(0)+pr(1)}$ and $\frac{pr(1)}{pr(0)+pr(1)}$, respectively. Similarly, a level-3 player thinks that others are either of level-0, -1, or -2 with probabilities $\frac{pr(0)}{\sum_{j=0}^{2} pr(j)}$, $\frac{pr(1)}{\sum_{j=0}^{2} pr(j)}$, and $\frac{pr(2)}{\sum_{j=0}^{2} pr(j)}$, respectively. For higher levels, their beliefs about others' levels are defined in a similar fashion.

When computing the expected payoff of choosing an integer a, a level-2 player takes this (truncated) distribution of others' levels into account. Namely, the level-2 player considers the probability of each of his or her opponents being level-0 and level-1, with respective choice probabilities $P_0^j(s) = 1/101$ and $P_1^j(s)$ defined in Eq. 3 above. Given such computed expected payoffs for each a, $E_2^{LCH}(\pi(a))$, the level-2 player chooses integer l with probability

$$P_2^{i,LCH}(l) = \frac{e^{\lambda_2 E_2^{LCH}(\pi(s))}}{\sum_a e^{\lambda_2 E_2^{LCH}(\pi(a))}} \tag{5}$$

The probability of choosing an integer for level-3 and above is defined iteratively in a similar manner.

Figures 1 to 4 show the numerical results for the LLK and the LCH models for $n \in \{2, 3, 4, 5\}$. In generating these figures, we assume $\lambda_k = 5.0$ for all k for both models. For LCH models, we also assume that the players' levels follow the Poisson distribution with the mean level equal to 2. In Figures 1 and 2, the choice probabilities for each integer, $P_k^{LLK}(a)$ and $P_k^{LCH}(a)$, respectively, for level-1, -2, and -3 players are shown for each n. In Figures 3 and 4, the cumulative distributions of the absolute deviations of choices from the Nash equilibrium prediction (60) are shown for level-1, -2, and -3 for each n for LLK and LCH, respectively. In all the figures, the outcome for $BCG+$ is shown in dashed line while that for $BCG-$ is shown in solid line. Note that for a level-1 player, the LLK and the LCH predictions are the same.

The first row, which corresponds to $n = 2$ in all figures, exhibits only one distribution in each panel, because the outcomes for $BCG-$ and $BCG+$ are identical for levels 1, 2, and 3. This confirms our insights based on the level-K and the CH models discussed above. In groups of size $n = 2$, we do not expect to observe a difference in the deviation from the Nash equilibrium between $BCG-$ and $BCG+$.

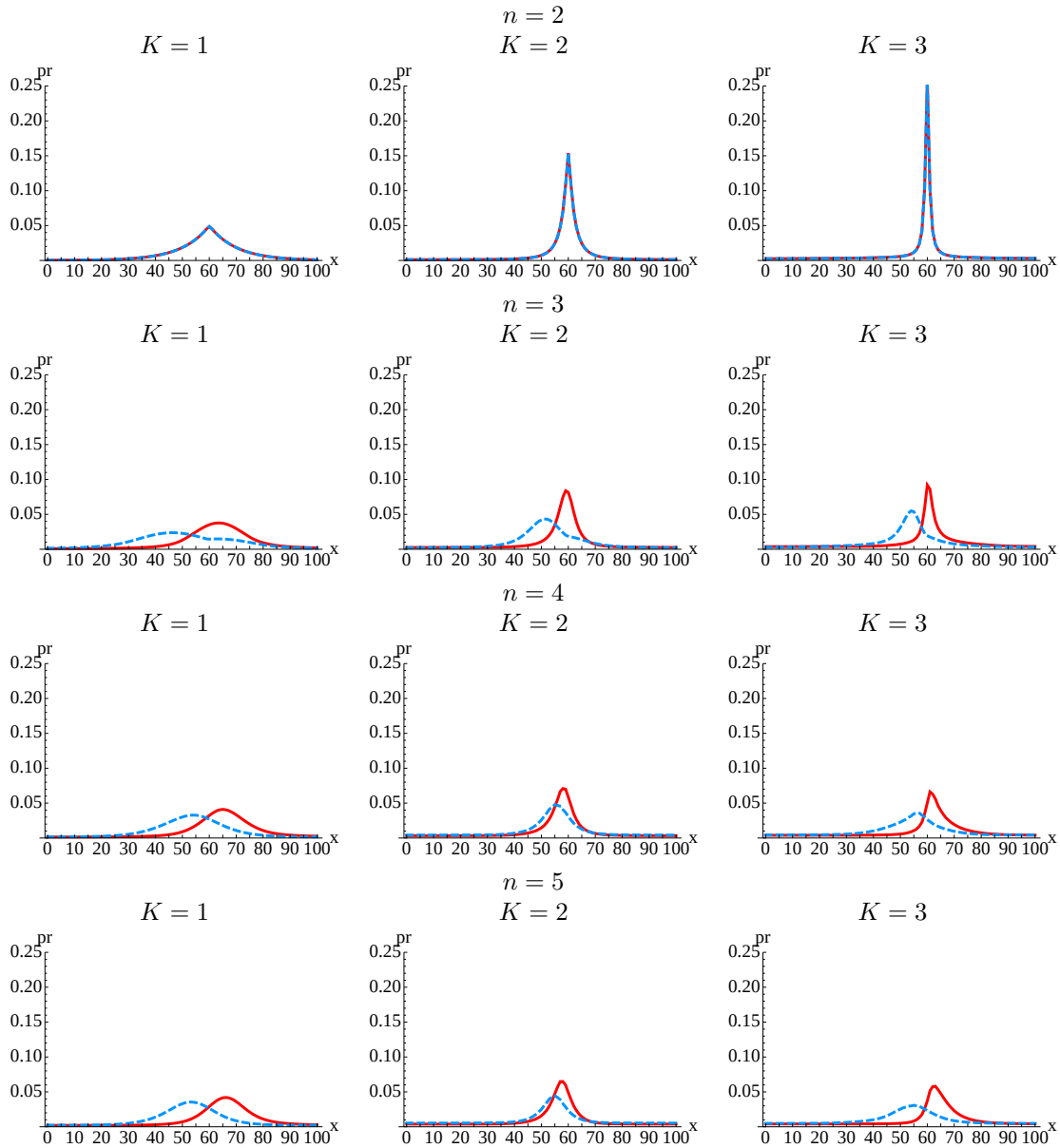


Figure 1: Probabilities for players with level-1, -2, and -3 choosing each integer in $[0, 100]$ in BCG_n- and BCG_n+ according to the logit-level-K (LLK) model for $n \in \{2, 3, 4, 5\}$. Outcomes for BCG_n- and BCG_n+ are shown in solid line and in dashed line, respectively. Note that outcomes for BCG_2- and BCG_2+ are exactly the same. For all the models, we assume $\lambda_k = 5$ for all k.

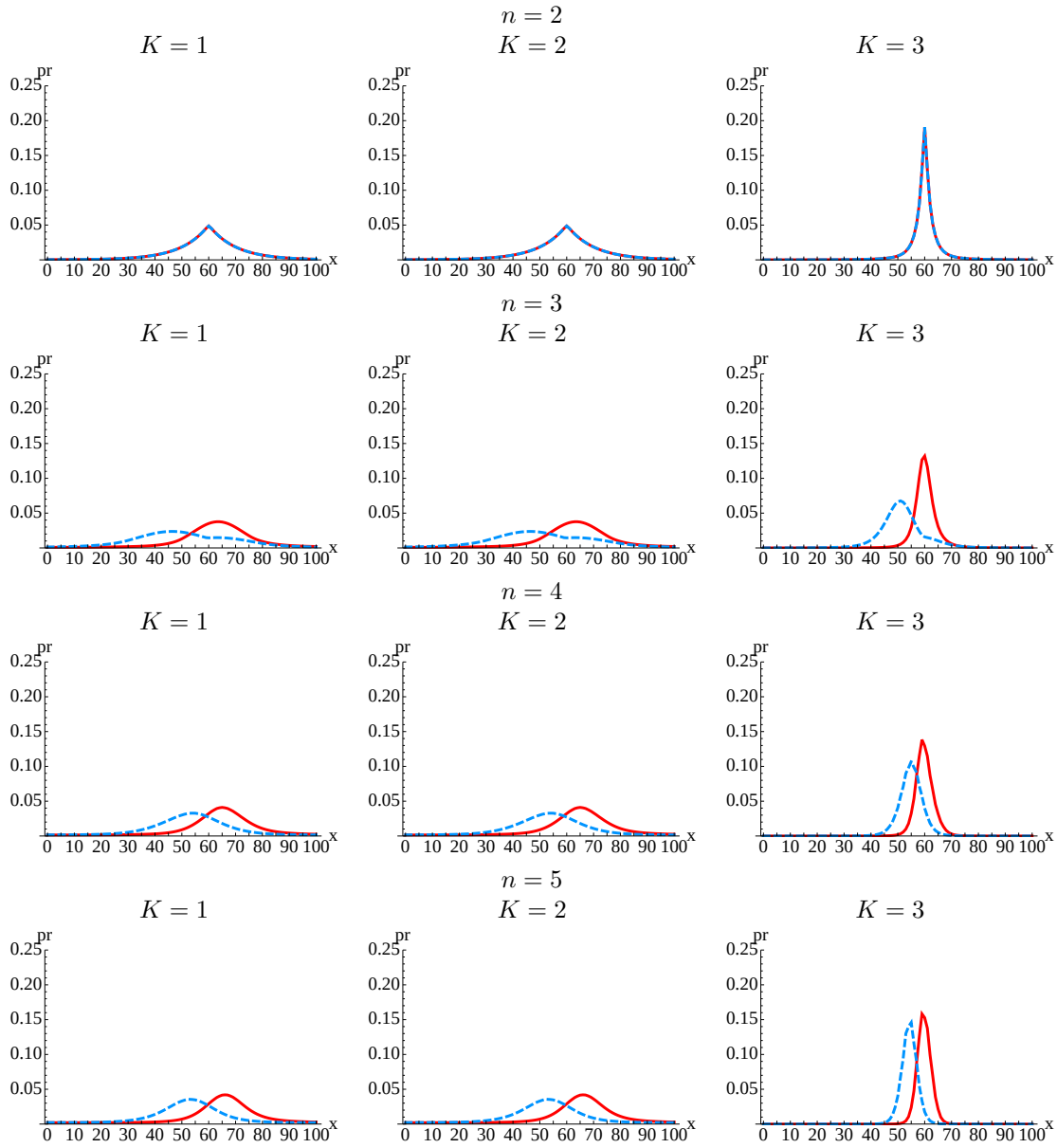


Figure 2: Probabilities for players with level-1, -2, and -3 choosing each integer in $[0, 100]$ in BCG_n- and BCG_n+ according to the logit-cognitive-hierarchy (LCH) models for $n \in \{2, 3, 4, 5\}$. Outcomes for BCG_n- and BCG_n+ are shown in solid line and in dashed line, respectively. Note that outcomes for BCG_2- and BCG_2+ are exactly the same. For all the models, we assume $\lambda_k = 5$ for all k. We assume the underlying distribution of levels follow the Poisson distribution with mean $k = 2$.

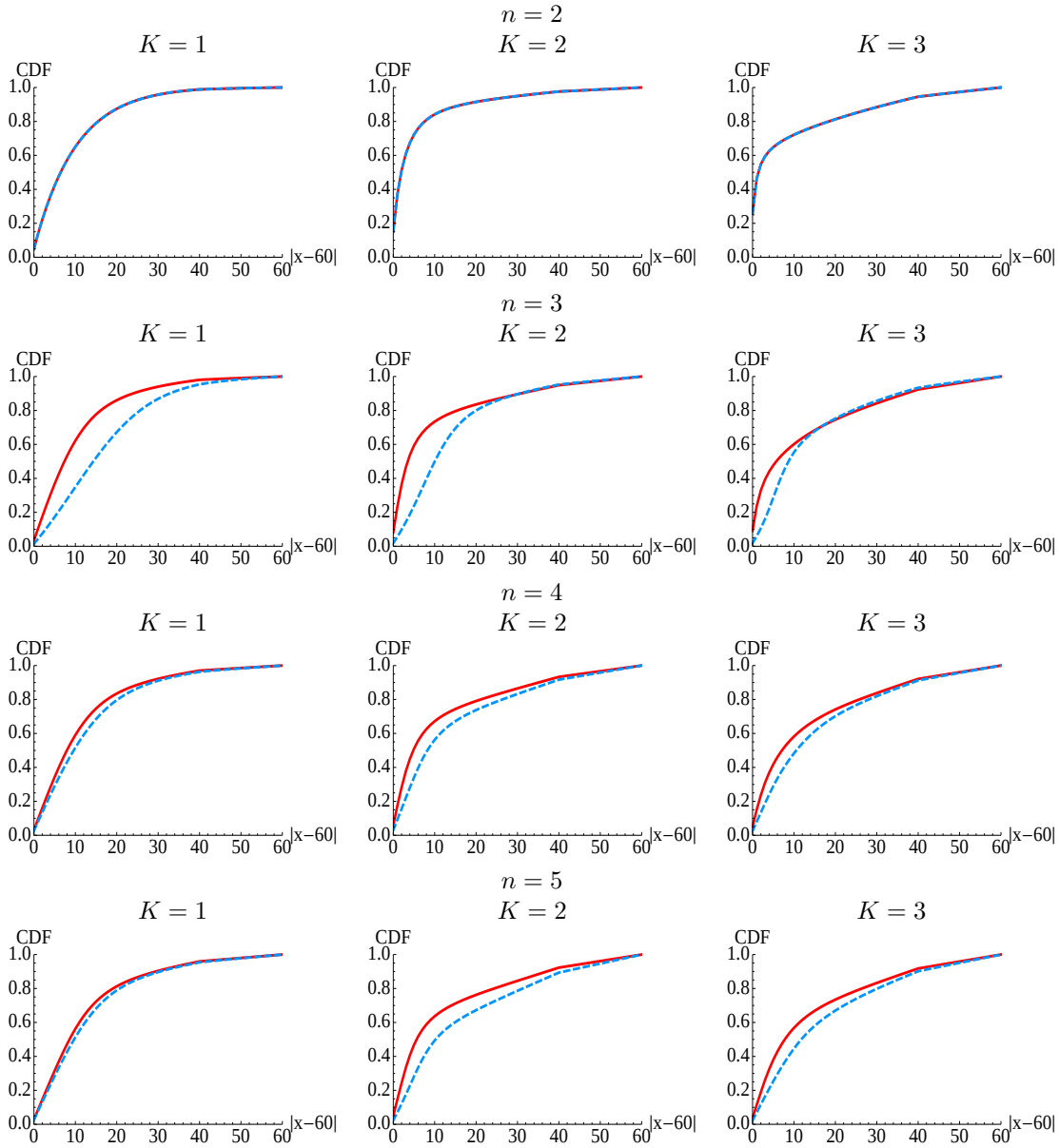


Figure 3: Cumulative distribution of the absolute deviation of the choice from the Nash equilibrium prediction (60) for players with level-1, -2, and -3 in BCG_n- and BCG_n+ according to the logit-level-K (LLK) model for $n \in \{2, 3, 4, 5\}$. Outcomes for BCG_n- and BCG_n+ are shown in solid line and in dashed line, respectively. Note that outcomes for BCG_2- and BCG_2+ are exactly the same. For all the models, we assume $\lambda_k = 5$ for all k.

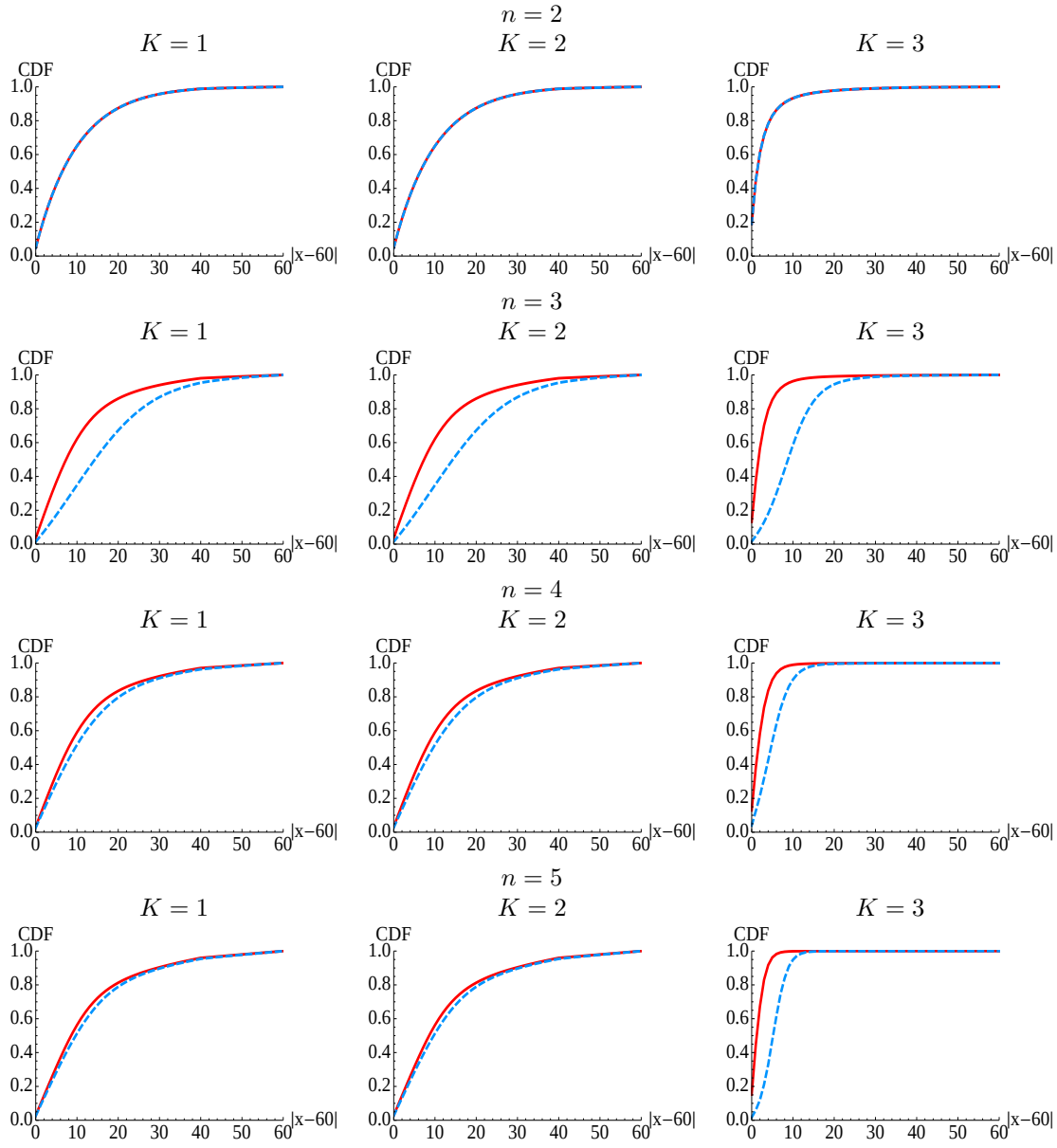


Figure 4: Cumulative distribution of the absolute deviation of the choice from the Nash equilibrium prediction (60) for players with level-1, -2, and -3 in BCG_n- and BCG_n+ according to the logit-cognitive-hierarchy (LCH) models for $n \in \{2, 3, 4, 5\}$. Outcomes for BCG_n- and BCG_n+ are shown in solid line and in dashed line, respectively. Note that outcomes for BCG_2- and BCG_2+ are exactly the same. For all the models, we assume $\lambda_k = 5$ for all k. We assume the underlying distribution of levels follow the Poisson distribution with mean $k = 2$.

For the larger groups ($n > 2$) shown in the remaining rows of Figure 1 to 4, the outcomes differ between $BCG-$ and $BCG+$. In particular, one observes that the cumulative distribution of the absolute deviations of choices from 60 (shown in Figures 3 and 4) in $BCG-$ lies on the left of that in $BCG+$ for all levels 1, 2, and 3 in both the LLK and the LCH models.

This leads to the main hypothesis to be tested in our experiments:

Hypothesis 1 *The deviation of the choices from the Nash equilibrium is the same for $BCG+$ and $BCG-$ when $n = 2$, and will be larger in $BCG+$ than in $BCG-$ for $n > 2$.*

Because it is not certain our experiments will result in such a clear-cut result with respect to n, we also put forward a weaker hypothesis:

Hypothesis 2 *The deviation of the choices from the Nash equilibrium is not significantly different between $BCG+$ and $BCG-$ when n is small, but becomes larger in $BCG+$ than in $BCG-$ when n is big enough.*

The purpose of Hypothesis 2 is to identify the critical n that distinguishes small groups from large groups. Therefore, in addition to conducting an experiment for $n = 2$ and $n = 3$ (to test Hypothesis 1), we systematically vary n in order to empirically determine the critical value of n.

Grosskopf and Nagel (2008) and Chou et al. (2009) studied a 2-player BCG for which the target number was $\frac{2}{3}mean$.[13] 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 dominant strategy in this game, i.e., to choose zero. Grosskopf and Nagel (2008) report, however, that despite this special feature about 90% of their 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 that the numbers chosen in their 2-player BCG are larger than the numbers chosen by subjects who were involved in BCG games with groups of size $n > 3$. According to Grosskopf and Nagel (2008) their result could be due to the fact that subjects tend to ignore the strength of the influence of their chosen number 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. In addition, the target for a subject is not influenced by his or her own choice. Thus, we believe the mechanism that we have outlined in this section, namely the difficulty of determining the average across choices among

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

less sophisticated players when the group size is small, is the main effect that drives the results of our experiment.

3 Experimental design

As noted in the previous section, we consider two beauty contest games, $BCG+$ and $BCG-$, while varying the size of the group n. We denote $BCG+$ and $BCG-$ games with group size of n by BCG_n+ and BCG_n-, respectively. In order to test *Hypothesis 2*, we systematically vary the group size n and consider $n \in \{2, 3, 4, 5, 6\}$. We also consider $n \in \{8, 16\}$ to check whether the result of Sutan and Willinger (2009) is robust against the small differences in experimental design between our experiment and their experiment discussed in Footnote 7 ($n = 8$), as well as to check whether our results continue to hold when the group size becomes even larger ($n = 16$). In our experiment, each subject chooses an integer between 0 and 100 and the subject whose choice is closest to the target number wins a fixed prize (8 euros). In case of a tie, as described in the previous section, one of the winners will be chosen at random to receive the money. We have opted to restrict the choice set to integers between 0 and 100, instead of real numbers, in order to make our experimental observations comparable to the predictions of the LLK and LCH models discussed in Section 2. Furthermore, because only a few subjects chose non-integers in Sutan and Willinger (2009), we expected that this restriction would not greatly influence the results.

In each experimental session, subjects play both BCG_n+ and BCG_n- with the same n. In half of the sessions, subjects play BCG_n+ first, and in the other half, they play BCG_n- first. Subjects were informed that they will play two games, called *Game 1* and *Game 2*, but they were not informed about the nature of *Game 2* when playing *Game 1*. Furthermore, no feedback regarding the outcome of *Game 1* was provided before playing *Game 2*. At the end of *Game 2*, one of the two games was chosen randomly for the payment. See the Appendix C for the English translation of the instructions that were provided to subjects in our experiment.

4 Results

Subjects were recruited from across the campus of the Burgundy School of Business between September and November 2015. In total, 1153 student subjects were involved in our experiment. Each

Table 2: Summary of treatments

Group Size	Number of Subjects in $BCG- \rightarrow BCG+$	Number of Subjects in $BCG+ \rightarrow BCG-$
2	100	92
3	90	90
4	100	92
5	70	75
6	66	66
8	96	88
16	64	64

subject participated in only one experimental session. On average, a session lasted for about 30 minutes. Experiments were not computerized and were carried out with papers and pens. Table 2 summarizes the number of subjects involved in each treatment.

4.1 Between subjects analyses

We start with a comparison of the subjects' choices for *Game 1* in order to derive between-subjects results. Figure 5 shows the empirical cumulative distribution (ECD) of chosen numbers and their absolute deviation from the Nash equilibrium choice for BCG_n- (in solid lines) and BCG_n+ (in dashed lines) for all the tested values of n.[14] The p-value below each panel is based on a two-sample permutation test (two-tailed).[15]

One can easily see from Figure 5 that for small group sizes, i.e. for $n \in \{2, 3, 4\}$, the absolute deviation of the chosen numbers from the equilibrium prediction ($|x-60|$) does not significantly differ between BCG_n+ and BCG_n-. The difference however becomes significant for larger group sizes, i.e., $n \in \{5, 6, 8, 16\}$. While the ratios of the median absolute deviation of the chosen numbers from 60 in $BCG+$ to that in $BCG-$ are 0.97, 1.25, and 1.08 in $n \in \{2, 3, 4\}$, respectively, they become 1.74, 2.88, 3.00, and 2.21 for $n \in \{5, 6, 8, 16\}$, respectively. These results give partial support to our hypotheses. Namely, we state:

Observation 1 (between subjects): *Our data reject Hypothesis 1 but accept Hypothesis 2. The deviations of the choices from the Nash equilibrium are not significantly different between BCG_n+ and BCG_n- for $n < 5$, but for $n \geq 5$ it is significantly larger in BCG_n+ than in BCG_n-.*

[14] Figure 7 in the Appendix shows the histogram of chosen numbers in BCG_n- and BCG_n+ for all values of n.
[15] The permutation tests are conducted using the STATA package provided by Kaiser (2007).

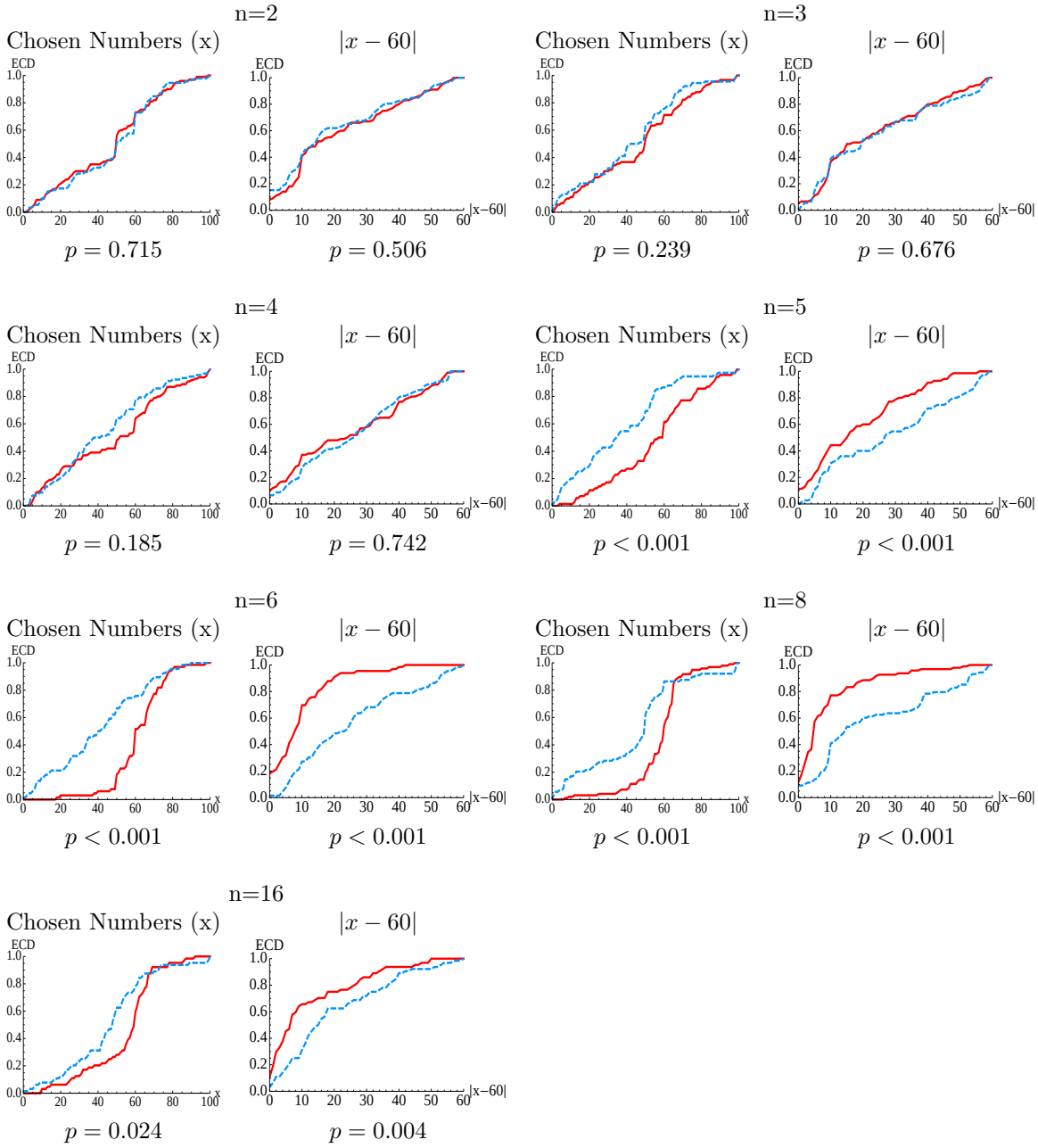


Figure 5: The empirical cumulative distribution of chosen numbers x and their absolute deviations from the Nash equilibrium predictions $|x-60|$ for BCG_n+ (dashed) and BCG_n- (solid).

4.2 Within subject analyses

We now consider the choices made by each subject in the two BCGs. Our subjects played both BCG_n+ and BCG_n- with the same group size n for each game. Some of our subjects played BCG_n+ first and then BCG_n-, while others played in the opposite order. We are primarily interested in whether the absolute deviations of the chosen numbers from the Nash equilibrium (60) are larger in BCG_n+ than in BCG_n-, *within subjects*. Thus, for each subject i we define

$$\Delta^i|x-60| \equiv |x^i_{BCG+}-60| - |x^i_{BCG-}-60|$$

where x^i_{BCG+} (x^i_{BCG-}) is the number subject i has chosen in BCG_n+ (BCG_n-). $\Delta^i|x-60|$ is the difference in the absolute difference between the numbers subject i has chosen and 60 in BCG_n+ and BCG_n-.

According to Hypothesis 1, $\Delta^i|x-60|$ should not differ significantly from zero for $n=2$, but should for $n \geq 2$. Under Hypothesis 2, $\Delta^i|x-60|$ should not be significantly different from zero for small values of n, but should be so for large enough values of n.

The results are summarized in Figure 6. The figure shows the empirical cumulative distribution (ECD) of the $\Delta|x-60|$ separately for those subjects who played BCG_n+ first (in solid black) and those who played BCG_n- first (in dashed gray). The p-values reported under the ECD plots are based on paired permutation test with the null hypothesis that $|x^i_{BCG+}-60| = |x^i_{BCG-}-60|$.

We obtain similar results for the within-subject analysis as those found for the between-subjects analyses. We fail to reject the null hypotheses that $|x^i_{BCG+}-60| = |x^i_{BCG-}-60|$ for group sizes $n \leq 4$,[16] but we reject the null hypothesis for $n \geq 5$. Thus, regardless of the order in which subjects played the two BCGs, subjects tended to deviate more from the Nash prediction in BCG_n+ than in BCG_n- for large groups ($n \geq 5$), but not for small groups ($2 \leq n < 5$). We summarize this as follows:

Observation 2 (within subjects): *Our data reject Hypothesis 1 but accept Hypothesis 2. For $n \geq 5$, deviations of the choices from the Nash equilibrium are significantly larger in BCG_n+ compared to BCG_n-*, but for $n < 5$ they are not.

[16] Except for those who played BCG_3+ first. In this particular case, we rejected $|x^i_{BCG+}-60| = |x^i_{BCG-}-60|$ in the opposite direction. Namely, $|x^i_{BCG-}-60|$ was significantly larger than $|x^i_{BCG+}-60|$. Note, however, if we had taken a one-tailed test, because our theory suggests to test $H_0 : |x^i_{BCG+}-60| \leq |x^i_{BCG-}-60|$ vs $H_1 : |x^i_{BCG+}-60| > |x^i_{BCG-}-60|$, we have not rejected the null hypothesis.

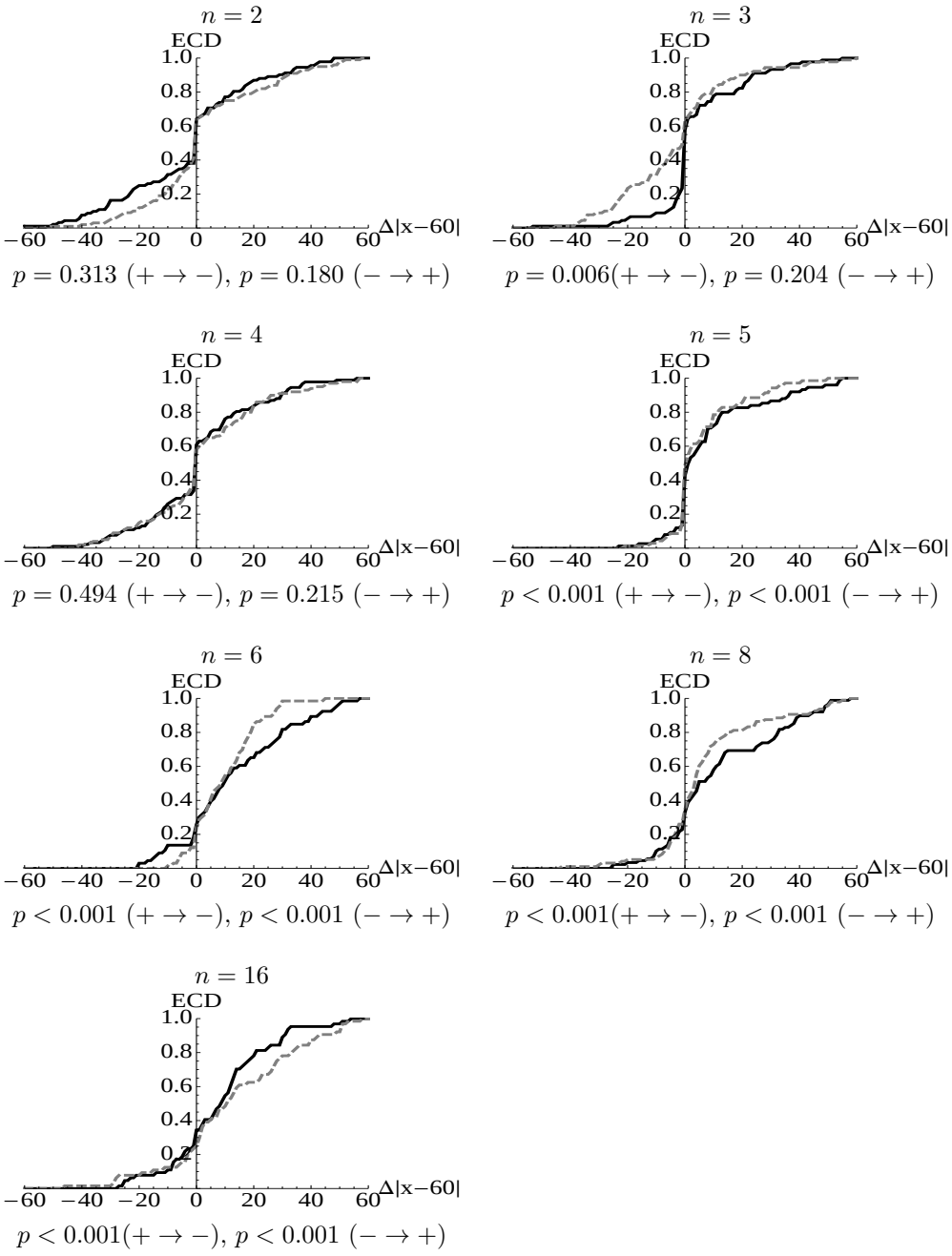


Figure 6: Empirical cumulative distribution (ECD) of $\Delta|60 - x|$. $\Delta|x - 60| \equiv |x_{BCG+} - 60| - |x_{BCG-} - 60|$. $BCG_n+ \to BCG_n-$ sessions are shown in solid black and $BCG_n- \to BCG_n+$ are shown in dashed gray. P-values are based on the paired permutation test (two-tailed). H_0 is $|x_{BCG+} - 60| = |x_{BCG-} - 60|$.

5 Discussion and conclusion

We use the term "the strategic environment effect" to refer to the tendency of agents to deviate significantly more from the Nash or rational expectations equilibrium when their actions are strategic complements than when they are strategic substitutes.

Its existence has been shown theoretically by Haltiwanger and Waldman (1985, 1989, 1991), and later confirmed experimentally by Fehr and Tyran (2008), Heemeijer et al. (2009), Potters and Suetens (2009), and Sutan and Willinger (2009).

The strategic environment effect provides promising insights into the aggregate consequences of interaction among heterogeneous boundedly rational agents, but its macroeconomic relevance has been often questioned because the above-mentioned experiments involved groups with a small number of interacting agents. This criticism has led to our research questions: Does the strategic environment effect depend on population size? If so, how?

As it is useful to provide robust experimental evidence about the relevance of the strategic environment effect for addressing macro phenomena, we investigated this question experimentally by studying variants of the beauty contest game that have the same unique interior solution, and by systematically varying the size of the group of interacting subjects. The two different beauty contest games we have considered involved either strategic substitutes or strategic complements. Both games are dominance solvable and have the same interior equilibrium that is reached after the same number of iterated eliminations of weakly dominated strategies. We have considered seven different group sizes: $n \in \{2, 3, 4, 5, 6, 8, 16\}$. In all games, the n subjects simultaneously chose an integer between 0 and 100. The winner is the subject who has chosen the number closest to the target. Under strategic complementarity, $BCG+$, the winning number for player i was defined as $20 + \frac{2}{3} mean_{-i}$. Under strategic substitutability, $BCG-$, the target for player i was set as $100 - \frac{2}{3} mean_{-i}$ where $mean_{-i}$ is the average of the numbers chosen by all players in the group except player i. The two games have the same Nash equilibrium, everyone choosing 60. The logit-level-K (LLK) and logit-cognitive-hierarchy (LCH) models predict that the deviation of the chosen number from the Nash equilibrium is larger in $BCG+$ than in $BCG-$ for $n > 2$.

We found that deviations from the Nash equilibrium prediction are larger in a statistically significant manner in $BCG+$ games than in $BCG-$ games when the group size exceeds $n \geq 5$. But such a difference is not observed for smaller-sized groups ($n < 5$). The ratio of the median absolute

deviation of the chosen number from the Nash equilibrium in $BCG+$ to that in $BCG-$ is close to one for $n < 5$, while it is more than two for $n > 5$.

Our findings establish the critical threshold required for population size to trigger the strategic environment effect and provide new evidence that experiments involving a small number of interacting agents can provide major insights into macro phenomena. Our results also support earlier experimental findings about the equilibrating force of strategic substitutes in contrast to strategic complements which tend to amplify deviations. However, the equilibrating force of the strategic substitutability operates properly only if the size of the group is large enough, i.e. $n \geq 5$ in our experimental setting. The reason is that under strategic substitutability, outcomes are closer to the Nash equilibrium prediction only if more-sophisticated players can safely average the numbers chosen by the other less-sophisticated players when selecting their own number. The self-correcting force at work at the aggregate level under strategic substitutability is absent under strategic complementarity, thus giving rise to larger deviations. However, even under strategic substitutability, a minimum group size is required for averaging. In the simple beauty contest games that we have considered, we found that a group of five subjects is enough to observe a significant strategic environment effect.

References

Akerlof, G. A. and J. L. Yellen (1985a): "Can small deviations from rationality make significant differences in economic equilibria?" *American Economic Review*, 75, 708–720.

——— (1985b): "A near-rational model of the business cycle, with wage and price inertia," *Quarterly Journal of Economics*, 100, 823–838.

Arad, A. and A. Rubinstein (2012): "The 11-20 money request game: A level-k reasoning study," *American Economic Review*, 102, 3561–3573.

Ariely, D. (2008): *Predictably Irrational: The Hidden Forces That Shape Our Decisions*, New York, NY: HarperCollins.

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.

Becker, G. S. (1962): "Irrational Behavior and Economic Theory," *Journal of Political Economy*, 70, 1–13.

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

Breitmoser, Y. (2012): "Strategic reasoning in p-beauty contests," *Games and Economic Behavior*, 75, 555–569.

Brock, W. A. and C. H. Hommes (1997): "A Rational Route to Randomness," *Econometrica*, 65, 1059–1095.

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.

Conlisk, J. (1996): "Why bounded rationality?" *Journal of Economic Literature*, 34, 669–700.

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.

De Long, J. B., A. Shleifer, L. H. Summers, and R. J. Waldmann (1990): "Noise trader risk in financial markets," *Journal of Political Economy*, 98, 703–738.

Duffy, J. (2016): "Macroeconomics: A survey of laboratory research," forthcoming in Kagel and Roth eds. *Handbook of Experimental Economics, Vol. 2*, Department of Economics, University of California, Irvine.

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

——— (2008): "Limited Rationality and Strategic Interaction: The Impact of the strategic environment on nominal inertia," *Econometrica*, 76, 353–394.

Friedman, D., S. Huck, R. Oprea, and S. Weidenholzer (2015): "From imitation to collusion: Long-run learning in a low-information environment," *Journal of Economic Theory*, 155, 185–205.

Gode, D. K. and S. Sunder (1993): "Allocative Efficiency of Markets with Zero-Intelligence Traders: Market as a Partial Substitute for Individual Rationality," *Journal of Political Economy*, 101, 119–137.

——— (1997): "What makes markets allocationally efficient?" *Quarterly Journal of Economics*, 112, 603–630.

Goeree, J. K. and C. A. Holt (2001): "The Little Treasures of Game Theory and Ten Intuitive Contradictions," *American Economic Review*, 91, 1402–1422.

——— (2004): "A model of noisy introspection," *Games and Economic Behavior*, 46, 365–382.

Goeree, J. K., P. Louis, and J. Zhang (2014): "Noisy Intropection in the "11-20" Game," University of Zurich.

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.

——— (1991): "Responders versus non-responders: A new perspective on heterogeneity," *Economic Journal*, 101, 1085–1102.

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.

Huck, S., H.-T. Normann, and J. Oechssler (2004): "Two are few and four are many: number effects in experimental oligopolies," *Journal of Economic Behavior and Organization*, 53, 435–446.

Kahneman, D. (2011): *Thinking, Fast and Slow*, New York, NY: Farrar, Straus and Giroux.

Kaiser, J. (2007): "An exact and a Monte Carlo proposal to the Fisher-Pitman permutation tests for paired replicates and for independent samples," *The Stata Journal*, 7, 402–412.

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

McKelvey, R. D. and T. R. Palfrey (1995): "Quantal response equilibria for normal form games," *Games and Economic Behavior*, 10, 6–38.

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.

Rogers, B. W., T. R. Palfrey, and C. F. Camerer (2009): "Heterogeneous quantal response equilibrium and cognitive hierarchies," *Journal of Economic Theory*, 144, 1440–1467.

Russell, T. and R. Thaler (1985): "The relevance of quasi rationality in competitive markets," *American Economic Reivew*, 75, 1071–1082.

Shleifer, A. (2000): *Inefficient Markets: An introduction to behavioral finance*, New York, NY: Oxford University Press.

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.

Spiegler, R. (2011): *Bounded Rationality and Industrial Organization*, New York, NY: Oxford University Press.

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.

Thaler, R. H. (1992): *The Winner's Curse: Paradoxes and Anomalies of Economic Life*, New York, NY: Free Press.

A The definition of the target and the slope of the best response functions

Let's consider BCG_n+ and BCG_n- where the winner is the one closer to the following target based on the mean including the player. That is for BCG_n+, the target, T, is

$$T = 20 + \frac{2}{3}\frac{\sum_j x_j}{n}$$

and for BCG_n-, it is

$$T = 100 - \frac{2}{3}\frac{\sum_j x_j}{n}$$

where x_j is the number chosen by subject j in the same group.

Ignoring the inequality of the winning condition, a bit of algebra will lead us to have the following best response functions. In BCG_n+

$$x(x_{-i}) = \frac{60n}{3n-2} + \frac{2n-2}{3n-2}\frac{\sum_{j\neq i} x_j}{n-1}$$

and in BCG_n-

$$x(x_{-i}) = \frac{300n}{3n+2} - \frac{2n-2}{3n+2}\frac{\sum_{j\neq i} x_j}{n-1}$$

Here if $n \to \infty$ then, we have $\frac{2n-2}{3n-2} = \frac{2n-2}{3n+2} = \frac{2}{3}$ so that the slope of the best response functions will be the same between BCG_n+ and BCG_n-, but with a small group size n, they are quite different. For any n, the absolute value of the slope is smaller in BCG_n- than in BCG_n+.

Thus, to be able to study the effect of the group size n, as well as the difference in the nature of the strategic interaction between BCG_n+ and BCG_n- without them influencing the slope of the best response functions, one needs to define the target based on the average choice by the other players in the group as we have done in this paper.

B Histogram of numbers chosen

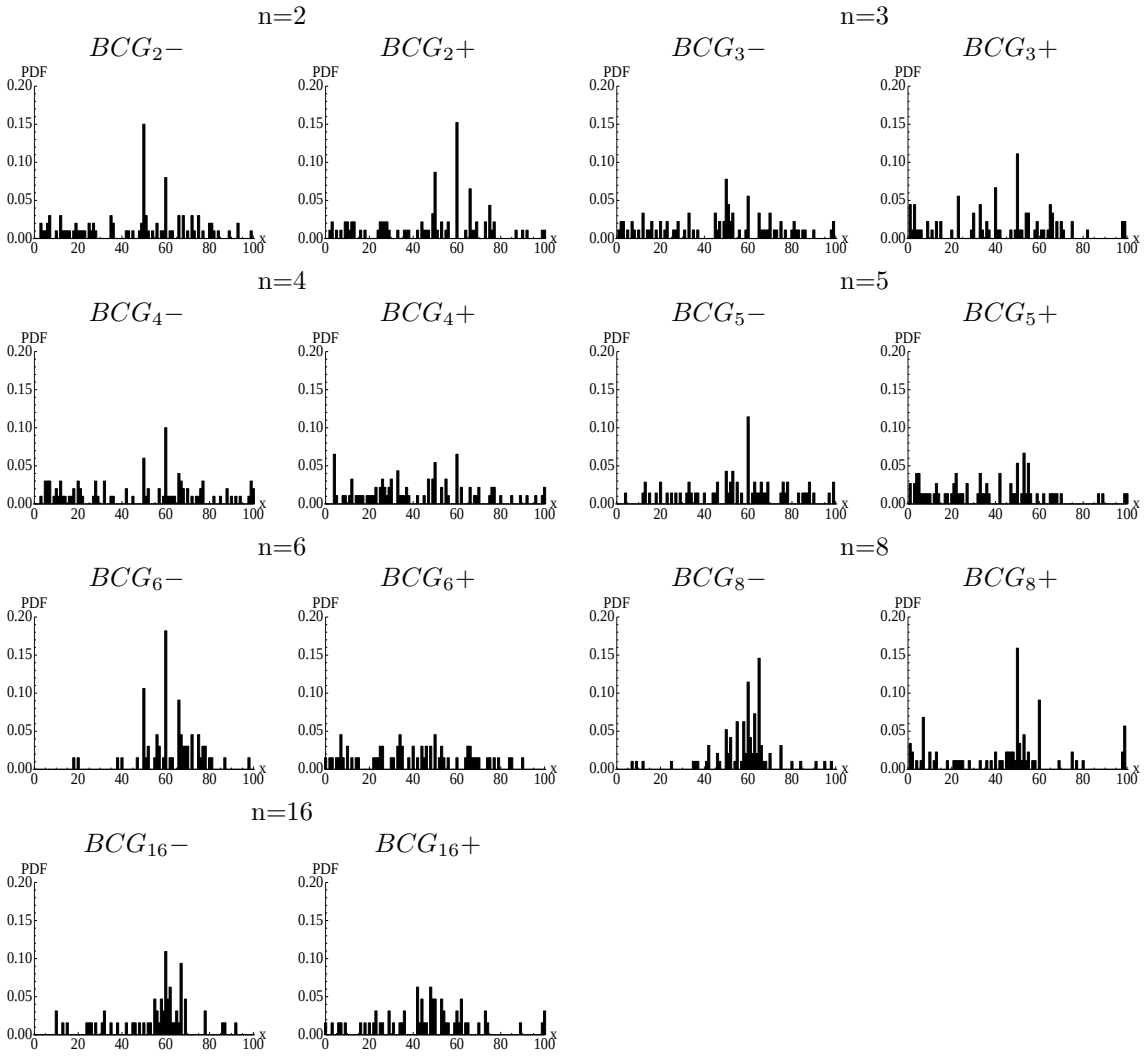


Figure 7: Histogram of numbers chosen by subjects in BCG_n- and BCG_n+ for various group size n. Here we consider only the first game each subject played.

C Instructions

This section of the appendix presents an English translation of the instructions used in our experiment with $n = 4$ where $BCG{+}_4$ was played before $BCG{-}_4$. The other treatments are different from the one presented in terms of the the number of players in the room and the examples shown. We have taken a particular care to give isomorphic examples for all values of n. The examples shown in various treatments are summarized at the end of this section.

General rule

This is an experiment about decision making.

You will play two games. Instructions for the second game will be given to you after you finish playing the first game.

This experiment allows you to earn real money. The payment rule is explained in each game.

One of the two games will be randomly selected at the end of the experiment and your payments in that game will be given to you.

Game 1

You interact with 3 other randomly selected people in this room.

You have to choose an integer between 0 and 100. The other 3 persons will also do the same. The rule to earn money is the following: the person choosing the number closest to his or her TARGET will earn 8 euros, the others will earn 0. If the numbers chosen by several persons are equally close to their own TARGETS, one of them will be randomly chosen to earn the 8 euros.

The TARGET for you is defined as follows:

$$\text{TARGET} = 20 + \frac{2}{3}(\text{average of the numbers chosen by the 3 other persons})$$

Example : You choose 3 and the 3 other persons choose 0, 1, 2. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$20 + \frac{2}{3}[(0+1+2)/3] = 20.66$	20.66-3 = 17.66 **The lowest**
Player with choice 0	$20 + \frac{2}{3}[(1+2+3)/3] = 21.33$	21.33-0=21.33
Player with choice 1	$20 + \frac{2}{3}[(0+2+3)/3] = 21.11$	21.11-1=20.11
Player with choice 2	$20 + \frac{2}{3}[(0+1+3)/3] = 20.88$	20.88-2=18.88

You win because the difference between the number you chose and your TARGET is the lowest.

When everybody has made his/her choice, the choices will be randomly divided into groups of 4 and the winners will be determined. The result will be communicated to you after you play Game 2 (and you will be able to take your money if you won). Good luck!

Your choice (between 0 and 100): _____

Game 2

Now you still interact with three randomly selected people in this room, but the TARGET is different.

You have to choose an integer between 0 and 100. The other 3 persons will do the same. The rule to earn money is the following: the person choosing the number closest to his or her TARGET will earn 8 euros, the others will earn 0. If the numbers chosen by several persons are equally close to their own TARGETS, one of them will be randomly chosen to earn the 8 euros.

The TARGET for you is defined as follows:

$$\text{TARGET} = 100 - \frac{2}{3}(\text{average of the numbers chosen by the 3 other persons})$$

Example : You choose 3 and the 3 other persons choose 0, 1, 2. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$100 - \frac{2}{3}[(0+1+2)/3] = 99.33$	99.33-3 = 96.33 **The lowest**
Player with choice 0	$100 - \frac{2}{3}[(1+2+3)/3] = 98.66$	98.66-0=98.66
Player with choice 1	$100 - \frac{2}{3}[(0+2+3)/3] = 98.88$	98.88-1=97.88
Player with choice 2	$100 - \frac{2}{3}[(0+1+3)/3] = 99.11$	99.11-2=97.11

You win because the difference between the number you chose and your TARGET is the lowest.

When everybody has made his/her choice, the choices will be randomly divided into groups of 4 and the winners will be determined (and you will be able to take your money if you won). Good luck!

Your choice (between 0 and 100): _____

C.1 Examples used in other treatments

BCG$+_2$: You choose 2 and the other person chooses 0. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 2)	$20 + \frac{2}{3}0 = 20$	20-2 = 18 **The lowest**
Player with choice 0	$20 + \frac{2}{3}2 = 21.33$	21.33-0=21.33

BCG$-_2$: You choose 2 and the other person chooses 1. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 2)	$100 - \frac{2}{3}1 = 99.33$	99.33-2 = 97.33 **The lowest**
Player with choice 1	$100 - \frac{2}{3}2 = 98.66$	98.66-1=97.66

BCG$+_3$: You choose 3 and the other two people both choose 1. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$20 + \frac{2}{3}[(1+1)/2] = 20.66$	20.66-3 = 17.66 **The lowest**
Player with choice 1	$20 + \frac{2}{3}[(1+3)/2] = 21.33$	21.33-1=20.33

BCG$-_3$: You choose 3 and the other two people both choose 1. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$100 - \frac{2}{3}[(1+1)/2] = 99.33$	99.33-3 = 96.33 **The lowest**
Player with choice 1	$100 - \frac{2}{3}[(1+3)/2] = 98.66$	98.66-1=97.66

BCG$+_5$: You choose 3 and the 4 other persons choose 0, 1, 1, and 2. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$20 + \frac{2}{3}[(0+1+1+2)/4] = 20.66$	20.66-3 = 17.66 **The lowest**
Player with choice 0	$20 + \frac{2}{3}[(1+1+2+3)/4] = 21.66$	21.66-0=21.66
Player with choice 1	$20 + \frac{2}{3}[(0+1+2+3)/4] = 21$	21-1=20
Player with choice 2	$20 + \frac{2}{3}[(0+1+1+3)/4] = 20.83$	20.83-2=18.83

BCG$-_5$: You choose 3 and the 4 other persons choose 0, 1, 1, and 2. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$100 - \frac{2}{3}[(0+1+1+2)/4] = 99.33$	99.33-3 = 96.33 **The lowest**
Player with choice 0	$100 - \frac{2}{3}[(1+1+2+3)/4] = 98.83$	98.83-0=98.83
Player with choice 1	$100 - \frac{2}{3}[(0+1+2+3)/4] = 99$	99 -1 = 98
Player with choice 2	$100 - \frac{2}{3}[(0+1+1+3)/4] = 99.16$	99.16-2=97.16

BCG$+_6$: You choose 3 and the 5 other persons choose 0, 0, 1, 1, and 2. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$20 + \frac{2}{3}[(0+0+1+1+2)/5] = 20.53$	20.53-3 = 17.53 **The lowest**
Player with choice 0	$20 + \frac{2}{3}[(0+1+1+2+3)/5] = 20.93$	20.93-0=20.93
Player with choice 1	$20 + \frac{2}{3}[(0+0+1+2+3)/5] = 20.8$	20.8-1=19.8
Player with choice 2	$20 + \frac{2}{3}[(0+0+1+1+3)/5] = 20.66$	20.66-2=18.66

BCG$-_6$: You choose 3 and the 5 other persons choose 0, 0, 1, 1, and 2. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$100 - \frac{2}{3}[(0+0+1+1+2)/5] = 99.46$	99.46-3 = 96.46 **The lowest**
Player with choice 0	$100 - \frac{2}{3}[(0+1+1+2+3)/5] = 99.06$	99.06-0=99.06
Player with choice 1	$100 - \frac{2}{3}[(0+0+1+2+3)/5] = 99.2$	99.2 -1 = 98.2
Player with choice 2	$100 - \frac{2}{3}[(0+0+1+1+3)/5] = 99.33$	99.33-2=97.33

BCG$+_8$: You choose 3 and all the other persons choose 1. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$20 + \frac{2}{3}[(1+1+1+1+1+1+1)/7] = 20.66$	20.66-3 = 17.66 **The lowest**
Player with choice 1	$20 + \frac{2}{3}[(1+1+1+1+1+1+3)/7] = 20.85$	20.85-1=19.85

BCG$-_8$: You choose 3 and all the other persons choose 1. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$100 - \frac{2}{3}[(1+1+1+1+1+1+1)/7] = 99.33$	99.33-3 = 96.33 **The lowest**
Player with choice 1	$100 - \frac{2}{3}[(1+1+1+1+1+1+3)/7] = 99.14$	99.14-1=98.14

BCG+$_{16}$: You choose 3 and 7 other persons choose 1, and remaining 8 have chosen 0. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$20 + \frac{2}{3}[(7 \times 1 + 8 \times 0)/15] = 20.31$	20.31-3 = 17.31 **The lowest**
Player with choice 1	$20 + \frac{2}{3}[(6 \times 1 + 8 \times 0 + 3)/15] = 20.4$	20.4-1=19.4
Player with choice 0	$20 + \frac{2}{3}[(7 \times 1 + 7 \times 0 + 3)/15] = 20.44$	20.44-0=20.44

BCG−$_{16}$: You choose 3 and 7 other persons choose 1, and remaining 8 have chosen 0. The TARGETs and the differences to the TARGETs are:

Players	TARGET	Difference to the TARGET
You (your choice: 3)	$100 - \frac{2}{3}[(7 \times 1 + 8 \times 0)/15] = 99.68$	99.68-3 = 96.68 **The lowest**
Player with choice 1	$100 - \frac{2}{3}[(6 \times 1 + 8 \times 0 + 3)/15] = 99.6$	99.6-1=98.6
Player with choice 0	$100 - \frac{2}{3}[(7 \times 1 + 7 \times 0 + 3)/15] = 99.55$	99.55-0=99.55