

# The strategic environment effect in beauty contest games

Nobuyuki Hanaki, Yukio Koriyama, Angela Sutan, Marc Willinger

▶ **To cite this version:**

Nobuyuki Hanaki, Yukio Koriyama, Angela Sutan, Marc Willinger. The strategic environment effect in beauty contest games. Games and Economic Behavior, 2019, 113, pp.587-610. 10.1016/j.geb.2018.11.006 . halshs-01929113

**HAL Id: halshs-01929113**
**https://shs.hal.science/halshs-01929113**

Submitted on 13 Dec 2018

**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† Yukio Koriyama‡ Angela Sutan§ Marc Willinger¶

November 22, 2018

**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, and we use both theory and experiments to shed light on its effectiveness. Based on cognitive hierarchy and level-K models, we show theoretically that the effect is operative for interactions among three or more agents. Our experimental results show a statistically significant strategic environment effect for groups of five or more agents, establishing its robustness against the increase in the population size.

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

**JEL Classification:** C72, C91.

*We thank John Duffy, Cars Hommes, Aidas Masiliunas, Rosemarie Nagel, Charles Noussair, Ali Ihsan Ozkes, and participants at WEHIA 2015 and ASFEE 2015, seminars at GATE (Lyon), ISER (Osaka), Kyoto, Southampton, and Waseda for their comments and suggestions. A part of this research is financed by the Conseil Regional de Bourgogne under "PARI" grants, an ORA-Plus project "BEAM" (ANR-15-ORAR-0004), Investissements d'Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047, and $UCA^{JEDI}$ ANR-15-IDEX-01).

†Université Côte d'Azur, CNRS, GREDEG, France. E-mail: Nobuyuki.HANAKI@unice.fr

‡CREST, Ecole Polytechnique, University of Paris-Saclay. E-mail: yukio.koriyama@polytechnique.edu

§Univ. Bourgogne Franche-Comte, Burgundy School of Business-CEREN. E-mail: angela.sutan@bsb-education.com

¶CEE-M, Univ Montpellier, CNRS , INRA, Montpellier SupAgro, Montpellier, France. E-mail: marc.willinger@umontpellier.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). A large body of experimental and empirical research has shown that people do not behave as rationally as often assumed by economic theory.[1] This accumulated evidence has begun 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, finance, and macroeconomics.[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 deviations from rational behavior by many boundedly rational individuals will cancel each other out when we consider aggregate phenomena, thus, bounded rationality at the 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. For instance, Akerlof and Yellen (1985a,b) and Russell and Thaler (1985), show instances where the existence of non-optimizing agents whose loss from non-optimization may be very small, can nevertheless 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 influence, i.e. more than proportional to their population share, 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, Spiegler (2011) provides an overview of the growth of this type of research in the field of industrial organization, and Duffy (2017) provides a survey of experimental analyses with macroeconomic implications.

is characterized by strategic complementarity.

In this paper, we follow up on the theoretical results from Haltiwanger and Waldman (1985, 1989, 1991). In their models, they considered two types of agents, naive and sophisticated, and 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.[3] We call this phenomenon "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 amplifying the deviations from the equilibrium caused by naive agents, while in the presence of strategic substitutability they have an incentive to act in the opposite way, thus offsetting 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 is 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. Suetens

[3]Naive agents in their models are those who take some known default behavior, while sophisticated agents are those who try to maximize their payoffs knowing that there are both naive agents and other sophisticated agents like themselves.

and Potters (2007) show, based on a meta-analyses, that this finding holds not only for duopoly but also for oligopoly games.

In a similar line of research, Sutan and Willinger (2009) experimentally studied two different one-shot beauty contest games (BCGs) with interior equilibria.[4] 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 $BCG+$.[5] Unlike the experiments mentioned in the previous paragraph 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 even 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 among a relatively small number of subjects. Indeed, as Duffy (2017) 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

[4] 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).

[5] Sutan and Willinger (2009) also studied 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.

strategic environment effect.

To understand how the strategic environment effect operates for the groups of different sizes, 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), particularly models that incorporate heterogeneity in the depth of strategic thinking among players, such as the level-K (Nagel, 1995) and the cognitive hierarchy 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).[6] 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 consider,[7] we can abstract away from issues related to subjects trying to coordinate on a non-equilibrium 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.

We first show, theoretically, that the strategic environment effect should be observed in our beauty contest games for any group sizes except for $n = 2$ based on cognitive hierarchy (Camerer et al., 2004) and level-K (Nagel, 1995) models. Experimentally, we vary groups of sizes systematically for $n \in \{2, 3, 4, 5, 6, 8, 16\}$, and also consider a case of unknown size, i.e., the case where subjects are informed that the group size can be anything between 2 and the total number of participants in the session (varied between 24 to 34) with equal probability. Based on our experiment involving more than 1000 subjects, we find that there is no strategic environment effect for small group sizes, i.e. for groups of less than five interacting subjects. However, as soon as the group size is at least equal to five interacting subjects, we observe a statistically significant strategic environment effect. More

[6] Keynes perceived 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 $k$th-order distribution of beliefs about what others believe, where $k$ is the deepest level of thinking in the population of traders.

[7] Güth et al. (2002) and Nagel et al. (2016) study beauty contest games in which payoffs depend on the distance from the target number, and thus the Nash equilibrium is the unique Pareto-efficient outcome of the game. They report, however, the observed behavior in this version of the beauty contest games are not different from the contest version of the game considered in this paper.

precisely, in groups of 5 or more subjects, 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 suggest that in a BCG environment, only a few interacting players are required in order to generate larger deviations from the rational expectations prediction under strategic complementarity than under strategic substitutability. This bolsters work done on other environments in which the strategic environment effect was observed, and which are relevant for macroeconomic issues such 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 provides theoretical predictions both for finite $n$ and for the limit as $n \to \infty$, 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 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.[8] Namely, in $BCG+$, the target for player $i$, $T^i_{BCG+}$, is defined as

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

[8]Thus, in the case of 2-player games, the target number depends simply on the number chosen by the opponent. Appendix B shows that, when the target is defined by the choices made by everyone in the group including one's own, the slopes of the best reply function for $BCG+$ and $BCG-$ are different for finite values of group size $n$. Therefore, defining the target number based on the average number chosen by other players in the group allows us to study the strategic environment effect more clearly.

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

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

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.[9] Let us first define the strategic environment effect.

**Definition 1 (Strategic environment effect)** *We say that the strategic environment effect arises if the expected absolute deviation of choices from the Nash equilibrium is larger when players' actions are strategic complements than when they are strategic substitutes. In our games, let $\tilde{y}_i^+$ (resp. $\tilde{y}_i^-$) be the random variable derived from the distribution of choices normalized around the Nash equilibrium in BCG+ (resp. BCG−). Formally, we say that the* ***(strong) strategic environment effect (SEE)*** *arises if $\mathbb{E}\left[\left|\tilde{y}_i^+\right|\right] > \mathbb{E}\left[\left|\tilde{y}_i^-\right|\right]$. Alternatively, one may define SEE based on the absolute deviation of the average choices. We say that the* ***weak SEE*** *arises if $\left|\mathbb{E}\left[\tilde{y}_i^+\right]\right| > \left|\mathbb{E}\left[\tilde{y}_i^-\right]\right|$.*

In the rest of the paper, when we use the expression "strategic environment effect," we refer to strong SEE as long as there is no ambiguity. As shown in footnote 11 in the proof of Theorem 1, strong SEE implies weak SEE in our games.

Because we will be focusing on the deviation from the Nash equilibrium prediction to study the strategic environment effect, let $y_i := x_i - 60$ denote player $i$'s strategy normalized around the Nash equilibrium. Then, the target functions are:

$$\tau^i_+ = \frac{2}{3}\bar{y}_{-i},\ \tau^i_- = -\frac{2}{3}\bar{y}_{-i} \tag{3}$$

where $\bar{y}_{-i} = \sum_{j\neq i} y_j / (n-1)$ is the average of the other players' normalized strategies.

Our first result is for $n = 2$. The strategic environment effect is absent in this particular case, since it turns out that $BCG+$ and $BCG-$ are equivalent games for $n = 2$.

[9] Our explanation of the target number in the instructions given to subjects for the $BCG+$ game, which is $20 + \frac{2}{3}\frac{\sum_{j\neq i} x_j}{n-1}$, is different from the one used by Sutan and Willinger (2009), i.e., $\frac{2}{3}\left(\frac{\sum_{j\neq i} x_j}{n-1} + 30\right)$. 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 where two subjects opt to choose a focal number to share the prize between them.

**Proposition 1** *The strategic environment effect is absent for $n = 2$.*

**Proof.** See Appendix A. ■

More generally for $n \geq 2$, the following lemma turns out to be useful.

**Lemma 1** *Let $\kappa = \frac{2}{3(n-1)}$ and $Y_{ij} = \sum_{h \neq i,j} y_h$.*

*(i) In BCG+, player $i$ wins if and only if $y_i$ lies between $y_j$ and $-y_j + \frac{2\kappa}{1-\kappa} Y_{ij}$ for $\forall j \neq i$.*

*(ii) In BCG−, player $i$ wins if and only if $y_i$ lies between $y_j$ and $-y_j - \frac{2\kappa}{1+\kappa} Y_{ij}$ for $\forall j \neq i$.*

**Proof.** See Appendix A. ■

**Remark 1** *The equivalence of BCG+ and BCG− for $n = 2$ is obtained as a corollary, since the $Y_{ij}$ term disappears.*

## 2.1 Asymptotic properties

In this subsection, we analyze asymptotic properties of the cognitive hierarchy models that relax the following two assumptions used in the equilibrium analyses: (i) all players have infinite depth of reasoning, and (ii) this fact is common knowledge, and it also takes into account heterogeneity in depth-of-strategic thinking among players.

For each $k \geq 1$, a level-$k$ player holds a belief over the level of the other players. Two models which draw the most attention in the literature are (i) the level-K model (Nagel, 1995) in which each level-$k$ player assigns the entire probability on level-$(k-1)$, and (ii) the Poisson cognitive hierarchy model (Camerer et al., 2004) in which beliefs are truncated Poisson distributions. More generally, for any distribution function $(p_h)_{h=0}^{\infty}$ over $\mathbb{N}$ with $p_0 > 0$, a cognitive hierarchy model can be defined, assuming that each level-$k$ player holds a belief induced by the truncated distribution $\left(\frac{p_h}{p_0+\cdots+p_{k-1}}\right)_{h=0}^{k-1}$ for $k \geq 1$. We say that a cognitive hierarchy model is **non-degenerate** if $p_h > 0$ for $\forall h$. Our analyses here focus on non-degenerate cognitive hierarchy model. See Appendix C for the level-K model.

Below, we first consider the best response of a player holding a belief such that each of the other players' strategies is independently and identically distributed. We then compare asymptotic properties between the $BCG+$ and the $BCG-$ games.

**Proposition 2** *Suppose that player $i$ holds a belief such that $y_j$ is independently and identically distributed for $\forall j \neq i$, and let $\mu$ be the mean value of the distribution. As $n \to \infty$, player $i$'s best response converges to $\frac{2}{3}\mu$ (resp. $-\frac{2}{3}\mu$) in the BCG+ (resp. BCG−) game.*

**Proof.** See Appendix A. ■

**Theorem 1** *As $n$ goes to infinity, the strong strategic environment effect arises under any non-degenerate cognitive hierarchy model.*

**Proof.** Let $\mu_0$ be the mean value of the level-0 strategy. Suppose $\mu_0 \neq 0$, since we are interested in the deviation from the Nash equilibrium behavior. Without loss of generality, assume $\mu_0 < 0$. For a non-degenerate cognitive hierarchy model, let $z^+_{CHk}$ (resp. $z^-_{CHk}$) be the limit level-$k$ strategy as $n \to \infty$ in the $BCG+$ (resp. $BCG-$) game. With a slight abuse of notation, let $z^\pm_{CH0} = \mu_0$.[10]

Since the strong strategic environment effect is defined by the expected absolute distance from the Nash equilibrium, it is sufficient to show that $\left|z^+_{CHk}\right| \geq \left|z^-_{CHk}\right|$ for $\forall k \geq 1$ with strict inequality for $\forall k \geq 2$, implying that

$$\sum_{k=0}^{\infty} p_k \left|z^+_{CHk}\right| > \sum_{k=0}^{\infty} p_k \left|z^-_{CHk}\right| \tag{4}$$

for any non-degenerate $(p_k)_{k=0}^{\infty}$.[11] Proposition 2 implies that, for $k \geq 1$,

$$\begin{aligned} z^+_{CHk} &= \frac{2}{3}\sum_{h=0}^{k-1} \frac{p_h}{p_0 + \cdots + p_{k-1}} z^+_{CHh}, \\ z^-_{CHk} &= -\frac{2}{3}\sum_{h=0}^{k-1} \frac{p_h}{p_0 + \cdots + p_{k-1}} z^-_{CHh}. \end{aligned}$$

In particular for $k = 1$, we have $\left|z^+_{CH1}\right| = \left|z^-_{CH1}\right| = \frac{2}{3}\left|\mu_0\right|$.

Now, we claim that $\left|z^+_{CHk}\right| > \left|z^-_{CHk}\right|$ for $\forall k \geq 2$. To see that, remember $z^+_{CH0} = \mu_0 < 0$ and $z^+_{CH1} = \frac{2}{3}\mu_0$. Since $z^+_{CHk}$ is $\frac{2}{3}$ of a weighted average of $\left(z^+_{CHh}\right)_{h=0}^{k-1}$ for $k \geq 1$, we have:

$$\mu_0 < z^+_{CH1} < z^+_{CH2} < \cdots < 0.$$

[10] Even though the level-0 strategy is often taken as a mixed strategy, what only matters here is the expected value.

[11] Strong SEE implies weak SEE. As shown below, $z^+_{CHk} < 0$ for $\forall k \geq 0$ and $z^-_{CH0} < 0 < z^-_{CH1}$, which implies $\sum_{k=0}^{\infty} p_k \left|z^+_{CHk}\right| = \left|\sum_{k=0}^{\infty} p_k z^+_{CHk}\right|$ and $\sum_{k=0}^{\infty} p_k \left|z^-_{CHk}\right| > \left|\sum_{k=0}^{\infty} p_k z^-_{CHk}\right|$. Then, (4) implies that $\left|\sum_{k=0}^{\infty} p_k z^+_{CHk}\right| > \left|\sum_{k=0}^{\infty} p_k z^-_{CHk}\right|$.

Since all terms of $z^+_{CHh}$ have the same (negative) sign, the triangle inequality holds with equality:

$$\left|z^+_{CHk}\right| = \frac{2}{3}\left|\sum_{h=0}^{k-1}\frac{p_h}{p_0+\cdots+p_{k-1}}z^+_{CHh}\right| = \frac{2}{3}\sum_{h=0}^{k-1}\frac{p_h}{p_0+\cdots+p_{k-1}}\left|z^+_{CHh}\right|. \tag{5}$$

On the other hand,

$$\left|z^-_{CHk}\right| = \frac{2}{3}\left|\sum_{h=0}^{k-1}\frac{p_h}{p_0+\cdots+p_{k-1}}z^-_{CHh}\right| \leq \frac{2}{3}\sum_{h=0}^{k-1}\frac{p_h}{p_0+\cdots+p_{k-1}}\left|z^-_{CHh}\right|. \tag{6}$$

Since $p_h > 0$ for $\forall h$, the last inequality is strict unless all terms of $z^-_{CHh}$ have the same sign. Since $z^-_{CH0} = \mu_0 < 0$ and $z^-_{CH1} = -\frac{2}{3}\mu_0 > 0$, the inequality is strict for $\forall k \geq 2$. In particular,

$$\left|z^-_{CH2}\right| < \frac{2}{3}\sum_{h=0}^{1}\frac{p_h}{p_0+p_1}\left|z^-_{CHh}\right|.$$

Since $\left|z^-_{CHh}\right| = \left|z^+_{CHh}\right|$ for $h = 0, 1$, we have:

$$\frac{2}{3}\sum_{h=0}^{1}\frac{p_h}{p_0+p_1}\left|z^-_{CHh}\right| = \frac{2}{3}\sum_{h=0}^{1}\frac{p_h}{p_0+p_1}\left|z^+_{CHh}\right| = \left|z^+_{CH2}\right|$$

where the last equality holds by (5). In sum, we have $\left|z^-_{CH2}\right| < \left|z^+_{CH2}\right|$. The claim is true for $k = 2$.

For the sake of induction, suppose $\left|z^-_{CHh}\right| < \left|z^+_{CHh}\right|$ for $h = 2, \cdots, k$ with $k \geq 2$. Then,

$$\left|z^-_{CHk+1}\right| < \frac{2}{3}\sum_{h=0}^{k}\frac{p_h}{p_0+\cdots+p_k}\left|z^-_{CHh}\right| < \frac{2}{3}\sum_{h=0}^{k}\frac{p_h}{p_0+\cdots+p_k}\left|z^+_{CHh}\right| = \left|z^+_{CHk+1}\right|.$$

The first inequality holds by (6) with strict inequality, since $k \geq 2$. The second inequality holds by assumption. The last equality holds by (5). ■

**Remark 2** *The result obtained in Theorem 1 still holds even for the cases in which the level-0 strategies differ between BCG+ and BCG−, as long as their means have the same distance from the Nash equilibrium. In our games, pure strategy* 20 *(resp.* 100*) may be a focal point of BCG+ (resp. BCG−), because they appear explicitly in the definition of the target number. Since the Nash equilibirum is* 60 *in both games, the distance has the same value* 40*, and thus Theorem 1 predicts the strategic environment effect for large n under any non-degenerate cognitive hierarchy model.*

**Remark 3** *The result does not depend on the specific value of the coefficient* $\frac{2}{3}$*. For any coefficient*

*smaller than 1, the same argument holds.*

## 2.2 The strategic environment effect for finite $n$

For finite $n$, it is a rather tedious task to provide an analytical form which precisely describes the best reply of a player. A naive approximation of the best reply is the mid-point of the interval specified in Lemma 1. In the $BCG+$ game, the mid-point is:

$$\beta_n^+ := \frac{1}{2}\left(y_j - y_j + \frac{2\kappa}{1-\kappa}Y_{ij}\right) = \frac{\kappa}{1-\kappa}Y_{ij} = \frac{2\,(n-2)}{3\,(n-1)-2}\bar{y}_{ij}^+,$$

where $\bar{y}_{ij}^+$ is the average of the strategies excluding players $i$ and $j$. Similarly in the $BCG-$ game,

$$\beta_n^- := \frac{1}{2}\left(y_j - y_j - \frac{2\kappa}{1+\kappa}Y_{ij}\right) = -\frac{\kappa}{1+\kappa}Y_{ij} = -\frac{2\,(n-2)}{3\,(n-1)+2}\bar{y}_{ij}^-.$$

Therefore,

$$\left|\frac{\beta_n^+}{\beta_n^-}\right| = \frac{1+\kappa}{1-\kappa}\left|\frac{\bar{y}_{ij}^+}{\bar{y}_{ij}^-}\right| = \frac{3\,(n-1)+2}{3\,(n-1)-2}\left|\frac{\bar{y}_{ij}^+}{\bar{y}_{ij}^-}\right|.$$

Whether the strategic environment effect would appear thus depends approximately on whether the product of these two ratios is larger than one or not. Note that the first term satisfies:

$$\frac{3\,(n-1)+2}{3\,(n-1)-2} > 1 \tag{7}$$

and that the ratio is decreasing in $n$ and converges to one from above as $n \to \infty$. This inequality suggests that the strategic environment effect appears for small $n$ and fades away as $n$ increases, as long as the ratio $\left|\bar{y}_{ij}^+/\bar{y}_{ij}^-\right|$ is close to one.

To see how the ratio changes for different $n$, we run Monte Carlo simulations.[12] Table 1 shows the normalized strategies $y = x - 60$ of level 1, 2 and 3 under a Poisson cognitive hierarchy model with a mean level 2.[13] The level-0 strategy is the uniform distribution. The limit values as $n \to \infty$ are computed using the asymptotic property stated in Proposition 2.

The effect implied by (7) is most visible among the level-1 strategies, as they are defined as the best reply to the level-0 strategy, and are the same for both the $BCG+$ and the $BCG-$ games

[12]Random samples are obtained from 1 to 100 million iterative draws. For small $n$, a larger number of draws are necessary in order to avoid ambiguity caused by the discontinuity of expected payoff functions. See also discussions in the following paragraphs.

[13]The program is written in MATLAB. The code is available from authors upon request.

Table 1: Normalized strategies $x - 60$ in $BCG+$ and $BCG-$ under a Poisson CH model with mean level 2. Grid search with $\Delta = 0.1$.

| | $BCG+$ | | | $BCG-$ | | |
|---|---|---|---|---|---|---|
| $n$ | $CH1$ | $CH2$ | $CH3$ | $CH1$ | $CH2$ | $CH3$ |
| 3 | $-13.3$ | $-6.7$ | $-6.6$ | 3.4 | 3.3 | 3.2 |
| 4 | $-5.7$ | $-3.3$ | $-3.2$ | 4.8 | 4.7 | 4.6 |
| 5 | $-6.5$ | $-6.4$ | $-6.4$ | 5.9 | 5.8 | 5.7 |
| 6 | $-6.7$ | $-6.6$ | $-6.5$ | 6.3 | 3.4 | 3.3 |
| 7 | $-6.6$ | $-6.5$ | $-6.4$ | 6.6 | 1.5 | 1.4 |
| 8 | $-6.7$ | $-6.6$ | $-6.5$ | 6.7 | 1.2 | 1.1 |
| 9 | $-6.7$ | $-6.6$ | $-6.5$ | 6.8 | 0.8 | 0.7 |
| 10 | $-6.7$ | $-6.6$ | $-6.5$ | 6.9 | 0.4 | 0.3 |
| 11 | $-6.7$ | $-6.6$ | $-6.5$ | 6.9 | 0.2 | 0.2 |
| 12 | $-6.7$ | $-6.6$ | $-6.5$ | 6.9 | 0 | $-0.1$ |
| 13 | $-6.7$ | $-6.6$ | $-6.5$ | 6.9 | $-0.1$ | $-0.1$ |
| 14 | $-6.7$ | $-6.6$ | $-6.5$ | 6.9 | $-0.2$ | $-0.3$ |
| 15 | $-6.7$ | $-6.6$ | $-6.5$ | 6.9 | $-0.3$ | $-0.4$ |
| 16 | $-6.7$ | $-6.6$ | $-6.5$ | 6.9 | $-0.4$ | $-0.5$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| 100 | $-6.7$ | $-5.3$ | $-4.9$ | 6.7 | $-0.8$ | $-0.7$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| 1000 | $-6.7$ | $-5.2$ | $-4.5$ | 6.7 | $-0.8$ | $-0.2$ |
| $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots$ |
| $\infty$ | $-6.67$ | $-5.18$ | $-4.49$ | 6.67 | $-0.74$ | $-0.25$ |

(hence $\left|\bar{y}_{ij}^{+}/\bar{y}_{ij}^{-}\right| = 1$). For level 2 and higher, the strategic environment effect is observed except for $n = 4$.

Note that all strategies in the $BCG+$ game have the same (negative) sign. This is due to strategic complementarity, which implies that the best reply has the same sign as the deviation generated by the level-0 strategy. On the other hand, the strategies in the $BCG-$ game have both negative and positive signs, due to the negative slope of the best reply function which represents strategic substitution. As a consequence, players of level 2 or higher have a belief such that the aggregate deviations generated by other players would offset each other, and thus the best reply strategy is closer to the Nash equilibrium in the $BCG-$ game than in the $BCG+$ game (i.e. $\left|\bar{y}_{ij}^{+}/\bar{y}_{ij}^{-}\right| > 1$). Therefore, the strategic environment effect is expected to appear for level 2 or higher.

For small $n$, however, several effects coexist, and the critical value of $n$ for which the strategic environment effect appears is not clear. Figure 1 shows the expected payoff of a level-2 player in both $BCG+$ and $BCG-$ games for different $n$. We observe that the best reply is attained at the

point where the payoff function is discontinuous for small $n$. Since the level-1 strategy is pure, the associated mass probability causes such discontinuities. For any $n$, each player holds a belief that other players' strategies are drawn from a composite of Poisson distribution, uniform distribution, and a pure strategy. As $n$ becomes large, combinations of distributions smooth out the expected payoff for large $n$, which is visible in Figure 1. However, the discontinuity remains problematic for small $n$.

If the discontinuity is indeed the reason for these puzzling outcomes for small $n$, in particular for $n = 4$, by smoothing the strategy choices by relaxing the assumption of perfect best reply, we would obtain the strategic environment effect even for $n = 4$. As we show in the next subsection with a logistic Poisson Cognitive Hierarchy (LCH) model, this is indeed the case.

## 2.3 A better-response model

We relax the "best response" assumption by considering a logistic Poisson CH (LCH) model, where players' choices depend *proportionally* on their expected payoffs: options with higher expected payoffs are chosen with higher probabilities.[14]

In order to facilitate the presentation and numerical computations of the LCH model, we assume that players (in $BCG+$ and $BCG-$) can only choose integer numbers between 0 and 100. 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 $E_1(\pi(x))$ be the expected payoff for a level-1 player if he or she chooses integer $x$. Then, the level-1 player chooses an integer $s$ according to the probability

$$P_1^i(s) = \frac{e^{\lambda E_1(\pi(s))}}{\sum_x e^{\lambda E_1(\pi(x))}} \tag{8}$$

[14]Stahl and Wilson (1994) proposed such a model with a cognitive hierarchy structure to capture the observed behavior of subjects in symmetric 3 by 3 games. It has been shown that these 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 that assume 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. In a similar line, Bosch-Domènech et al. (2010) introduce noise in an extended level-K model that includes Nash players (i.e., Level-$\infty$) assuming that agents' choices follow level-specific beta-distributions, and find that the model captures well the data of Bosch-Domènech et al. (2002).

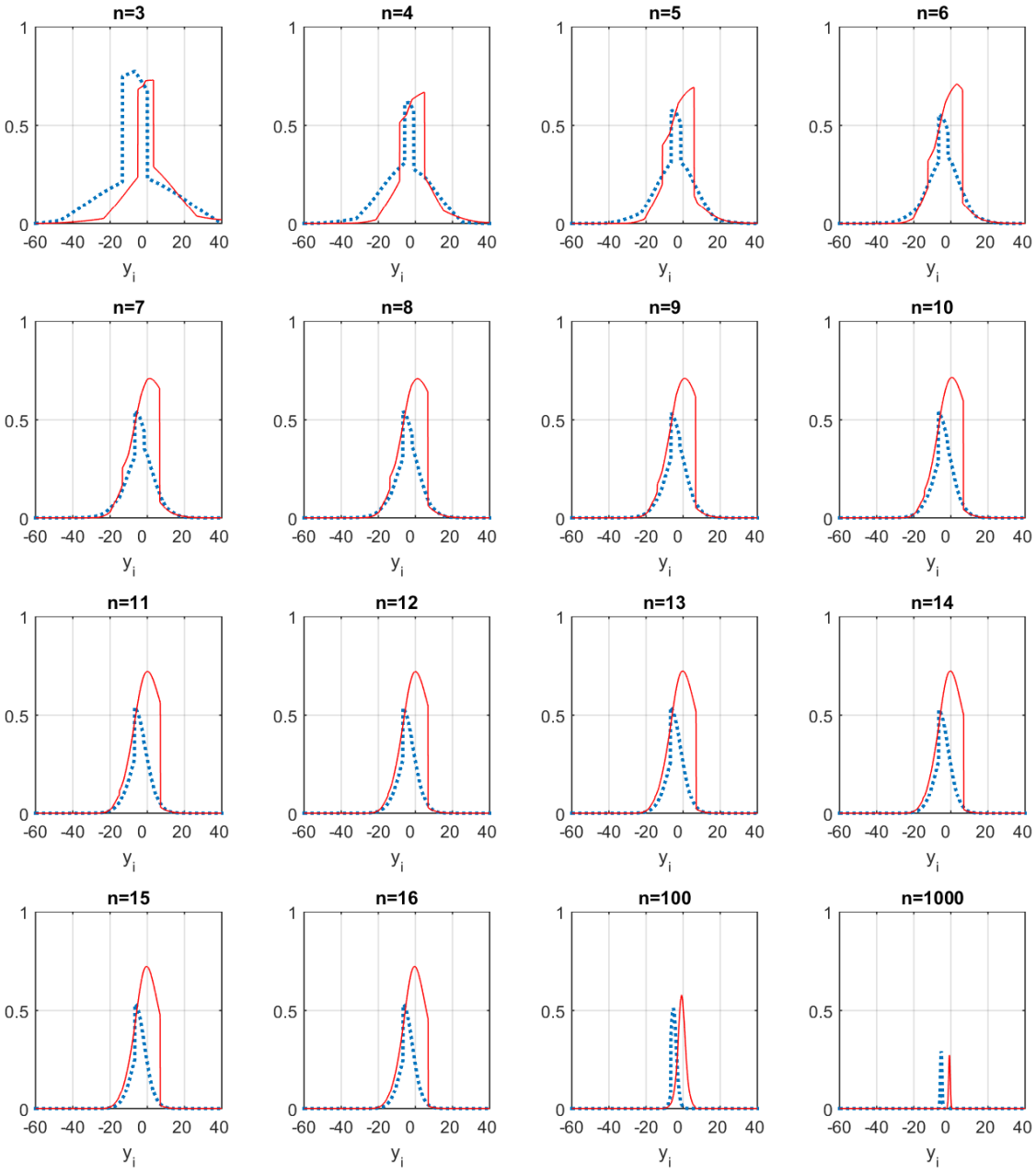



Figure 1: The expected payoff of a CH2 player for $BCG_n+$ (dashed) and $BCG_n-$ (solid).

where $\lambda$ is a parameter that governs the sensitivity of players' choices to the variation in their expected payoffs. If $\lambda = 0$, 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 $\lambda \to \infty$, the probability distribution becomes degenerate and players select the integer with the highest expected payoff with probability one, i.e., players tend to perfect best reply.[15]

A 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 (8) above. Given such computed expected payoffs for each $x$, $E_2(\pi(x))$, the level-2 player chooses integer $l$ with probability

$$P_2^i(l) = \frac{e^{\lambda E_2(\pi(l))}}{\sum_x e^{\lambda E_2(\pi(x))}}. \tag{9}$$

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

For the numerical exercise reported below, in each BCG and group size $n$, we compute the exact expected payoffs from choosing each integer for agents of level 1 to 5 for a given value of $\lambda$, assuming that an agent obtains a payoff of 1 if he wins. In case of ties, we assume that the payoff of 1 is given to one of the winners who is randomly chosen (that is, the expected payoff is $\frac{1}{\text{number of winners}}$).[16] We then aggregate the probabilities of these 6 levels of agents choosing each integer by taking the weighted average of the probabilities where the weight is based on the truncated Poisson distribution with mean 2 assuming the maximum level is 5.

Figure 2 shows the mean $|x - 60|$ in $BCG+$ (dashed blue) and $BCG-$ (solid red) for $n \in \{2, 3, 4, 5, 6, 7\}$ in LCH. We assume a mean level of 2 and $\lambda = 10.0$.[17] The strategic environment effect is observed for all the values of $n$ we have considered except for $n = 2$. Thus, as we have hypothesized, the absence of the strategic environment effect in $n = 4$ case considered in the previous subsection was driven by the strong assumption of best reply.

Before explaining our experimental design, a remark about existing experiments on 2-player beauty contest games is in order. Grosskopf and Nagel (2008) and Chou et al. (2009) studied a 2-

[15] $\lambda$ may be level dependent. In this paper, however, we assume the same $\lambda$ for all the levels above 1.

[16] The program is written in C++. The source code is available from the author upon request. We compute the exact expected payoff for each integer. In order to do so, we need to consider all the possible combinations of integers chosen by $n - 1$ other agents. This makes computation time grow as $101^{(n-1)}$ without careful programming. We have implemented this computation differently so that computing time grows roughly as ${}_{101}H_{n-1}$. Even with this implementation and making use of the parallel computation, however, it was not possible for us to go beyond $n = 7$ with a reasonable computing time.

[17] Unless $\lambda$ is too small, so that choices are not too noisy, we obtain qualitatively the same result. The results are also robust for the other mean levels, such as 1.0 and 1.5. See Appendix D.

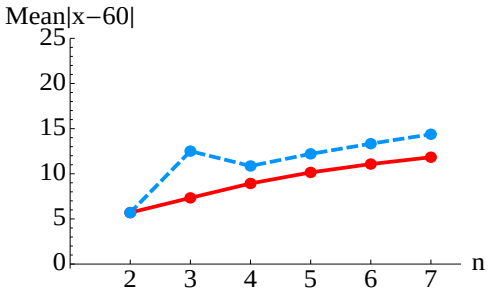



Figure 2: Mean $|x - 60|$ in $BCG+$ (dashed blue) and $BCG-$ (solid red) for $n \in \{2, 3, 4, 5, 6, 7\}$ according to LCH. We assume mean level is 2 and $\lambda = 10.0$.

player BCG for which the target number was $\frac{2}{3}mean$, where mean includes the player's own choice.[18] This 2-player BCG has the 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 this special feature about 90% of their subjects chose numbers larger than zero. Chou et al. (2009) argue that the subjects do not realize the special feature of the game. 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 $n > 3$. To better understand this rather puzzling result, Nagel et al. (2016) conducted a new experiment of a 2-player BCG with a distance payoff function: both players were paid an amount that depended on the distance of their chosen number and 2/3 of the mean. Thanks to this particular payoff function, 0 is the unique dominance solvable equilibrium which is also Pareto-efficient. Furthermore, the equilibrium solution is reached after iterated elimination of strictly (and not weakly) dominated strategies. However, in the data reported by Nagel et al. (2016), the distribution of the chosen numbers with the distance function does not differ from those of the standard 2-player BCG considered in Grosskopf and Nagel (2008). Moreover, the spikes are at exactly the same numbers for both treatments. According to the authors, in a 2-player BCG subjects are unable to find the dominant strategy of the game and do not take into account their influence on the mean. Furthermore, they probably focus on some focal numbers.

In our 2-player BCGs, as one easily could infer from the proof of Proposition 1 above, there is also a dominant strategy, i.e., to choose 60. However, this dominant strategy is far more difficult

[18] 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 the two players differed. Furthermore, in their games, the payoff declined continuously based on the distance from the target such that the equilibrium was reached by iterated elimination of strictly dominated strategies.

to find compared to the games considered in the aforementioned studies. In addition, the target for a subject is not influenced by his or her own choice. Furthermore, our focus is to investigate the strategic environment effect, which is different from these earlier studies on 2-player BCGs.

# 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. 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 9 ($n = 8$), as well as to check whether our results continue to hold when the group size becomes even larger ($n = 16$). Furthermore, we consider situations where subjects did not know the exact group size except that it could be any size between 2 and the total number of subjects involved in a session with an equal probability.[19] We call this treatment *uncertain* $n$. With the uncertain $n$ treatment, we aim to investigate the strategic environment effect when the size of the group may be larger than 16 without the subjects knowing the exact size. In many empirically relevant situations (e.g. financial markets), it is often the case that individuals do not know the exact number of agents who interact in a large group. Although we have not derived any theoretical predictions regarding the uncertain group size case, one may extrapolate from the analyses above that as long as subjects believe that the size of the group is likely to be greater than two, the strategic environment effect is expected to appear.

In our experiment, each subject chose an integer between 0 and 100, and the subject whose choice was closest to the target number won a fixed prize (8 euros). In the case of a tie, as described in the previous section, one of the winners was chosen at random to receive the prize. 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 LCH model discussed in Section 2.3. 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 played both $BCG_n+$ and $BCG_n-$ with the same $n$. In

[19] Please see Appendix F for the English translation of the instructions.

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 |
| uncertain $n$ | 145 | 134 |
| Number of participants in each session of uncertain $n$ | {25,28,29,31,32} | {23,24,26,27,34} |

Note: The number of participants of a session is the maximum possible group size for the session in the uncertain treatment.

half of the sessions, subjects played $BCG_n+$ first, and in the other half, they played $BCG_n-$ first. Subjects were informed that they would 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.

# 4 Results

Subjects were recruited from across the campus of the Burgundy School of Business between September 2015 and October 2016. In total, 1432 student subjects were involved in our experiment. Each 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. In the uncertain $n$ treatment, the number of participants in one session (thus, the maximum group size) varied between 23 and 34, while the actual group size selected varied between 13 and 31. Although we have not explicitly informed subjects about the exact number of participants in the session they have participated, subjects could have a rough estimate by seeing the number of subjects gathering for the session.

## 4.1 Between-subject analysis

We start with a comparison of the subjects' choices for *Game 1* in order to derive between-subject results. Figure 3 shows the empirical cumulative distribution (ECD) of the absolute deviation of chosen numbers, $x$'s, from the Nash equilibrium prediction ($|x-60|$) for $BCG_n-$ (in solid lines) and $BCG_n+$ (in dashed lines) for all the tested values of $n$.[20] It also reports the mean and the standard deviation of $|x-60|$. The p-value below each panel is based on a two-sample permutation test (two-tailed).[21]

One can easily see from Figure 3 that for small group sizes, i.e. for $n \in \{2,3,4\}$, $|x-60|$ does not significantly differ between $BCG_n+$ and $BCG_n-$. The difference, however, becomes statistically significant for larger and uncertain group sizes, i.e., $n \in \{5,6,8,16,\text{uncertain}\}$.

**Observation 1 (between subjects)**: *The absolute deviations of the choices from the Nash equilibrium are not significantly different between $BCG_n+$ and $BCG_n-$ for $n<5$, but for $n \geq 5$, and for the case of an uncertain group size, it is significantly larger in $BCG_n+$ than in $BCG_n-$. Thus, the strategic environment effect becomes statistically significant for $n \geq 5$ but not for $n<5$.*

## 4.2 Within-subject analysis

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 $BCG_n-$.

Table 3 shows the means and the standard deviations of $\Delta^i|x-60|$ for various $n$. We have

[20] Figure 8 in Appendix E shows the histogram of chosen numbers in $BCG_n-$ and $BCG_n+$ for all values of $n$.
[21] The permutation tests are conducted using the STATA package provided by Kaiser (2007).

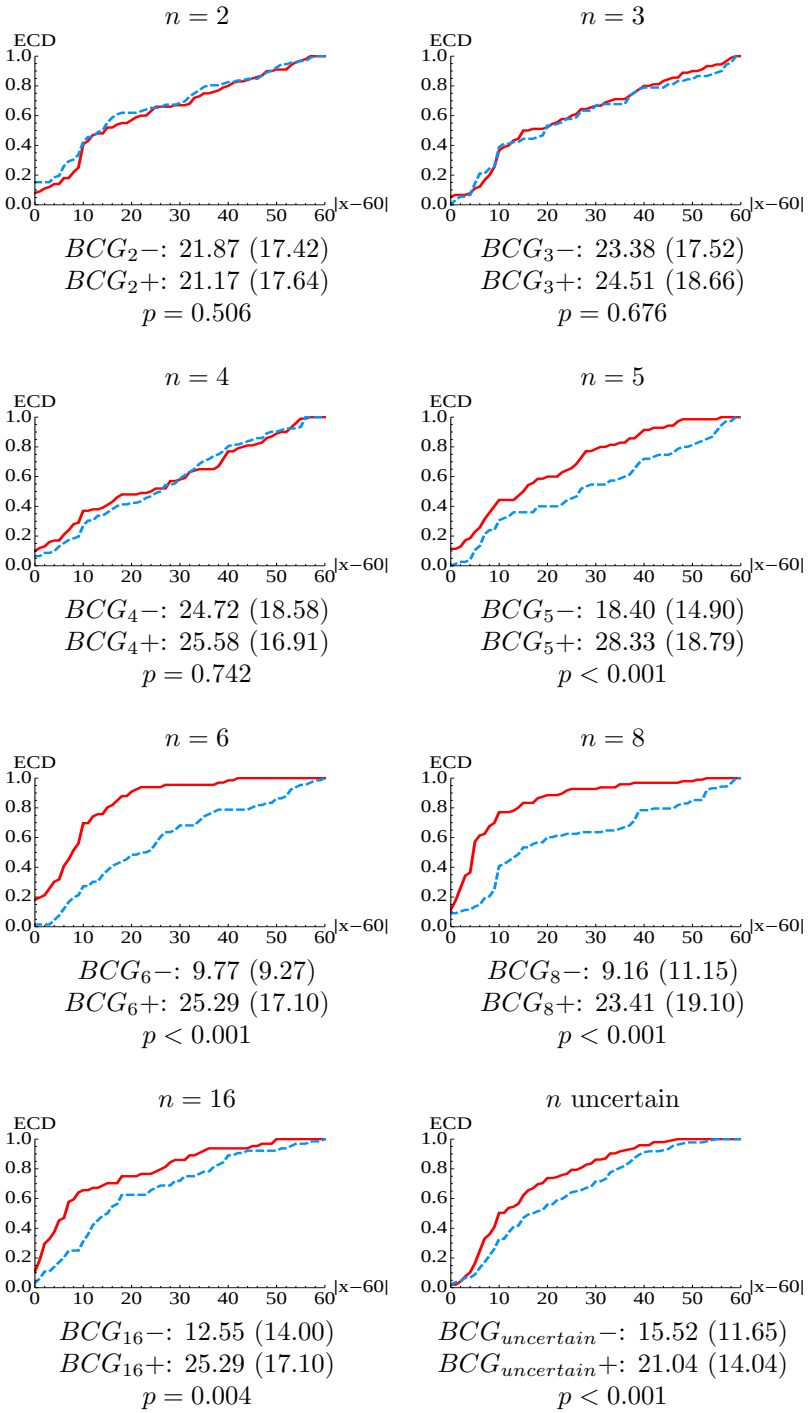



Figure 3: The empirical cumulative distribution of the absolute deviations from the Nash equilibrium predictions $|x - 60|$ for $BCG_n+$ (dashed) and $BCG_n-$ (solid). Mean $|x - 60|$ and its standard deviation in the parentheses are also reported. $p$-values are based on a two-sample permutation test (two-tailed) with the null hypothesis that $|x - 60|$s are the same between $BCG_n+$ and $BCG_n-$ for each $n$.

Table 3: Mean and standard deviation of $\Delta^i|x-60| \equiv |x^i_{BCG+}-60| - |x^i_{BCG-}-60|$ for various $n$

| $n$ | $BCG_n+ \to BCG_n-$ | $BCG_n- \to BCG_n+$ | p-value |
|---|---|---|---|
| 2 | -2.45 (22.92) | 2.83 (20.85) | 0.098 |
| 3 | 4.63 (15.77) | -2.72 (20.02) | 0.007 |
| 4 | 1.41 (19.54) | 2.60 (20.70) | 0.690 |
| 5 | 8.56 (16.39) | 6.24 (12.47) | 0.347 |
| 6 | 13.67 (19.05) | 10.15 (11.15) | 0.200 |
| 8 | 12.32 (19.79) | 7.71 (18.23) | 0.102 |
| 16 | 8.92 (17.44) | 13.06 (22.58) | 0.250 |
| uncertain | 7.07 (13.55) | 8.55 (17.09) | 0.427 |

[i] Standard deviations are in parenthesis.
[ii] p-values are based on the two-sample permutation test with the null hypothesis that $\Delta^i|x-60|$ is the same in $BCG_n+ \to BCG_n-$ sessions than in $BCG_n- \to BCG_n+$ sessions.

separated the sessions according to the order in which the two BCGs are played. Assuming that subjects "learn" to choose a number closer to the Nash equilibrium in the second game that they play than in the first game, $\Delta^i|x-60|$ tends to be larger for those subjects who played $BCG+$ first than those who played $BCG-$ first. The results of the two-sided permutation tests, reported in the last column of Table 3, show, however, that $\Delta^i|x-60|$ is significantly different between the two orderings at 5% significance level only for $n=3$. Thus, for the analysis below, we pool the data from all the sessions.

We run a linear regression with $\Delta^i|x-60|$ being the dependent variable and various group size dummies being the independent variables. The results of the regression are reported in Table 4. As one can see from the estimated coefficients for group size dummies, the strategic environment effect (measured by $\Delta^i|x-60|$) is significant and positive for $n \geq 5$ and uncertain $n$, but not for $n \in \{2,3,4\}$. This result confirms **observation 1** of the between-subject analysis.

We also tested whether the estimated strategic environment effects are significantly different within the subsamples of group sizes $n \in \{2,3,4\}$ and $n \in \{5,6,8,16,\text{uncertain}\}$. The results of the F-tests show that they are not significantly different within the two groupings. P-values are 0.647 for $n \in \{2,3,4\}$ and 0.125 for $n \in \{5,6,8,16,\text{uncertain}\}$.

**Observation 2 (within subjects)**: *For $n \geq 5$, absolute 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.* Thus, the statistically significant strategic environment effect is observed for $n \geq 5$ but not for $n < 5$.

Table 4: OLS regression $\Delta^i|x-60| \equiv |x^i_{BCG+}-60| - |x^i_{BCG-}-60|$ on various size dummies

| Group size dummy | Coefficient | Std. Err. | p-value |
|---|---|---|---|
| 2 | 0.302 | 1.322 | 0.819 |
| 3 | 0.956 | 1.365 | 0.484 |
| 4 | 2.031 | 1.322 | 0.125 |
| 5 | 7.441 | 1.521 | <0.001 |
| 6 | 11.909 | 1.594 | <0.001 |
| 8 | 9.913 | 1.350 | <0.001 |
| 16 | 10.992 | 1.619 | <0.001 |
| uncertain | 7.842 | 1.096 | <0.001 |
| No. Obs. = 1432. $R^2 = 0.141$ | | | |

[i] Dummy variables take a value of 1 for the corresponding group size, and 0 otherwise.
[ii] There is no constant term in this regression.

# 5 Discussion and conclusion

We defined the "strategic environment effect" as the tendency for 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 small numbers of interacting agents (up to 8). 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 evidence about the relevance of the strategic environment effect for addressing macro phenomena, we investigated this question theoretically and 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.

Our theoretical analyses show that the strategic environment effect is observed for any group size but 2. Equipped with these results, in our experiment, we have considered eight different group sizes: $n \in \{2, 3, 4, 5, 6, 8, 16, \text{uncertain}\}$. Except for the treatment with uncertain $n$, subjects were perfectly aware of the size of their group. For the uncertain $n$ treatment, subjects were informed that the size of the group could be anything between 2 and the number of subjects involved in the session with equal probability.

We observed a statistically significant strategic environment effect when the group size was $n \geq 5$ as well as for uncertain $n$. However, such an effect was not observed for smaller-sized groups ($n < 5$). This observation is coherent with our theoretical result that the effect should persist for large $n$.

Our findings establish a critical threshold required for population size to trigger the strategic environment effect in a BCG environment, and provide new evidence that experimental findings of the strategic environment effect are robust against an increase in the group size. Other recent studies show that it is not only in the one-shot BCG where the experimental findings on the strategic environment effect are robust against an increase in the group size. Bao et al. (2016) and Kopányi-Peuker et al. (2016) observe, in the presence of strategic complementarity, large deviations from the rational expectations equilibrium with large group sizes (21 to 32 for the former and more than 100 for the latter) in the framework of a learning-to-forecast experiment. Although similar robustness checks against increasing group sizes should be carried out in other experimental setups, our results as well as those by Bao et al. (2016) and Kopányi-Peuker et al. (2016) indicate that the experimental findings on the strategic environment effect, even if they are based on the interactions among a small number of subjects, may provide a major insight for macro phenomena.

The gap between the theoretical prediction and our experimental results for very small group sizes, namely for $n = 3$ and $n = 4$, however, leaves us with some puzzles. As one can see from the histogram of chosen numbers reported in Figure 8 in Appendix E, on one hand, for small group sizes, choices are distributed rather uniformly between 0 and 100 excepts for spikes at 50 and 60 in both BCG+ and BCG−. For larger groups, on the other hand, they are distributed more around 60 in BCG− while they continue to be rather uniformly distributed in BCG+. These observations are not fully captured by our theoretical analyses. A possible reason for not observing the strategic environment effect experimentally for these small groups size is subjects' thought processes being different in very small groups from those in larger groups. This hypothesis can be investigated by eliciting subjects' thought processes in small and in large groups, for example, based

on the experimental method recently proposed by Agranov et al. (2015) that elicits the dynamics of subjects' choice processes in an incentivized manner.

Such an analysis would also shed light on another interesting question. In the current paper, the strategic environment effect is a result of aggregating choices made by players of varying degrees of strategic sophistication. In our theoretical analyses, players are basically applying the same type of reasoning in the two strategic environments. It is, however, possible that these two strategic environments induce players to think differently about the problem itself, as suggested by a recent neuroscience study (Nagel et al., 2017) that compares brain activity when subjects are playing a critical mass game (with strategic complementarity) and an entry game (with strategic substitution). However, the two games Nagel et al. (2017) consider have different sets of Nash equilibria. There are a unique symmetric mixed strategy Nash equilibrium and many asymmetric pure strategy Nash equilibria in the market entry game, while there are two symmetric pure strategy Nash equilibria and a unique symmetric mixed strategy Nash equilibrium in the critical mass game. Furthermore, in the critical mass game, level-1 and above all behave in the similar manner (i.e., no thinking higher than level-1 is invoked), while in the entry game, their behavior differ (i.e., higher order thinkings are invoked). Thus, it is not clear whether the differences in neural activity between the two games that Nagel et al. (2017) report are due to the difference in the nature of the strategic environment (complements vs substitutes) or due to these other differences. Because the two BCGs we have considered do not differ except for the nature of the strategic environment, our setting could clarify this point.

Finally, one may ask whether our finding can be extended to other settings, such as oligopoly games, where earlier studies reported the strategic environment effect (Potters and Suetens, 2009). In oligopoly games, however, it is well known that the larger a group is, the more difficult it is to achieve the cooperative outcome (Huck et al., 2004). It is possible that such a negative group size effect on the subjects' ability to coordinate on the cooperative outcome dominates the strategic environment effect as we increase the group size. However, the precise relationship between these two effects is not known. Thus, it would surely be fruitful future research to consider this question both theoretically and experimentally.

# References


Agranov, M., A. Caplin, and C. Tergiman (2015): "Naive play and the process of choices in guessing games," *Journal of Economic Science Association*, 1, 146–157.

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., M. Hennequin, C. Hommes, and D. Massaro (2016): "Coordination on bubbles in large-group asset pricing experiments," CeNDEF Working paper 16-09, University of Amsterdam.

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.

——— (2010): "A finite mixture analysis of beauty-contest data using generalized beta distribution," *Experimental Economics*, 13, 461–475.

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. (2017): "Macroeconomics: A survey of laboratory research," in *Handbook of Experimental Economics, Volume 2*, ed. by J. H. Kagel and A. E. Roth, Princeton, NJ: Princeton University Press, chap. 1, 1–90.

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.

Kopányi-Peuker, A., C. Hommes, and J. Sonnemans (2016): "Group size and expectation formation in an asset market: a learning to forecast experiment," Paper presented at 2016 annual meeting of society for experimental finance, CeNDEF.

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.

Nagel, R., A. Brovelli, F. Heinemann, and G. Coricelli (2017): "Neural mechanisms mediating degrees of strategic uncertainty," *Social Cognitive and Affective Neuroscience*, forthcoming.

Nagel, R., C. Bühren, and B. Frank (2016): "Inspired and inspiring: Hervé Moulin and the discovery of the beauty contest game," *Mathematical Social Sciences*, forthcoming, https://doi.org/10.1016/j.mathsocsci.2016.09.001.

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.

Suetens, S. and J. Potters (2007): "Bertland colludes more than Cournot," *Experimental Economics*, 10, 71–77.

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 Proofs

## A.1 Proof of Proposition 1

**Proof.** Suppose $n = 2$. By definition, the distances to the target in $BCG+$ and $BCG-$ are respectively:

$$d_i^+ = \left| y_i - \frac{2}{3} y_j \right| \text{ and } d_i^- = \left| y_i + \frac{2}{3} y_j \right|,$$

for $i \neq j$. Player $i$ wins against player $j$ in $BCG+$ if and only if:

$$\begin{aligned} d_i^+ < d_j^+ &\Leftrightarrow \left| y_i - \frac{2}{3} y_j \right| < \left| y_j - \frac{2}{3} y_i \right| \\ &\Leftrightarrow \left( y_i - \frac{2}{3} y_j \right)^2 < \left( y_j - \frac{2}{3} y_i \right)^2 \\ &\Leftrightarrow y_i^2 < y_j^2. \end{aligned}$$

In $BCG-$, player $i$ wins against player $j$ if and only if:

$$\begin{aligned} d_i^- < d_j^- &\Leftrightarrow \left| y_i + \frac{2}{3} y_j \right| < \left| y_j + \frac{2}{3} y_i \right| \\ &\Leftrightarrow \left( y_i + \frac{2}{3} y_j \right)^2 < \left( y_j + \frac{2}{3} y_i \right)^2 \\ &\Leftrightarrow y_i^2 < y_j^2. \end{aligned}$$

The above equivalences show that the winner is the same and thus the payoffs are the same between $BCG+$ and $BCG-$, given the strategy profile $(y_1, y_2)$. Therefore, the two games are equivalent. Trivially, there is no strategic environment effect for $n = 2$. ■

## A.2 Proof of Lemma 1

**Proof.** Let $Y_{-i} = \sum_{h \neq i} y_h$. By (3), the target functions are $\tau_i^+ = \kappa Y_{-i},\ \tau_i^- = -\kappa Y_{-i}$. Hence, the distance functions are:

$$d_i^+ = |y_i - \kappa Y_{-i}|,\ d_i^- = |y_i + \kappa Y_{-i}|.$$

Hence,

$$\begin{aligned} d_i^+ < d_j^+ &\Leftrightarrow \left(y_i - \kappa y_j - \kappa Y_{ij}\right)^2 < \left(y_j - \kappa y_i - \kappa Y_{ij}\right)^2, \\ d_i^- < d_j^- &\Leftrightarrow \left(y_i + \kappa y_j + \kappa Y_{ij}\right)^2 < \left(y_j + \kappa y_i + \kappa Y_{ij}\right)^2. \end{aligned}$$

Player $i$ wins against $j$ in $BCG+$ iff:

$$\begin{aligned} & \left(y_i - \kappa y_j - \kappa Y_{ij}\right)^2 < \left(-\kappa y_i + y_j - \kappa Y_{ij}\right)^2 \\ \Leftrightarrow \quad & y_i^2 + \kappa^2 y_j^2 - 2\kappa y_i Y_{ij} + 2\kappa^2 y_j Y_{ij} < \kappa^2 y_i^2 + y_j^2 + 2\kappa^2 y_i Y_{ij} - 2\kappa y_j Y_{ij} \\ \Leftrightarrow \quad & \left(1-\kappa^2\right)\left(y_i - y_j\right)\left(y_i + y_j - \frac{2\kappa}{1-\kappa} Y_{ij}\right) < 0 \\ \Leftrightarrow \quad & \left(y_i - y_j\right)\left(y_i + y_j - \frac{2\kappa}{1-\kappa} Y_{ij}\right) < 0. \end{aligned}$$

In the last equivalence, we used $1 - \kappa^2 > 0$. Similarly in $BCG-$,

$$d_i^- < d_j^- \Leftrightarrow \left(y_i - y_j\right)\left(y_i + y_j + \frac{2\kappa}{1+\kappa} Y_{ij}\right) < 0.$$

■

## A.3 Proof of Proposition 2

**Proof.** Note that

$$\frac{2\kappa}{1 \pm \kappa} Y_{ij} = \frac{2\frac{2}{3(n-1)}}{1 \pm \frac{2}{3(n-1)}} \left(n-2\right) \bar{y}_{ij} = \frac{4\left(n-2\right)}{3\left(n-1\right) \pm 2} \bar{y}_{ij}$$

where $\bar{y}_{ij} = Y_{ij} / \left(n-2\right)$ is the average of the strategies excluding players $i$ and $j$. By the Law of Large Numbers, as $n$ goes to infinity, $\bar{y}_{ij}$ converges to the mean value $\mu$. Therefore, as $n \to \infty$, $\frac{2\kappa}{1 \pm \kappa} Y_{ij}$ converges to $\frac{4}{3}\mu$.

By Lemma 1, in the $BCG+$ game, player $i$ wins if and only if $y_i$ lies between $y_j$ and $-y_j + \frac{2\kappa}{1-\kappa} Y_{ij}$ for $\forall j \neq i$. Since the average of these values converges to $\frac{2}{3}\mu$ for each $j$, player $i$'s best response converges to $\frac{2}{3}\mu$. Similarly, player $i$'s best response converges to $-\frac{2}{3}\mu$ in the $BCG-$ game. ■

# B The definition of the target and the slope of the best response functions

Let us 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.

# C The level-K model

Consider the (deterministic) level-K model, in which each level-$k$ player best replies against the level-$(k-1)$ strategy. As in Theorem 1, let $(p_h)_{h=0}^{\infty}$ be the level distribution among the players.

**Proposition 3** *The weak strategic environment effect arises asymptotically in the level-K model.*

**Proof.** Let $\mu_0$ ($< 0$ wolog) be the mean value of the level-0 strategy. let $z_{Lk}^{+}$ (resp. $z_{Lk}^{-}$) be the limit level-$k$ strategy as $n \to \infty$ in the $BCG+$ (resp. $BCG-$) game. By Proposition 2,

$$z_{Lk}^{+} = \left(\frac{2}{3}\right)^{k} \mu_0 \text{ and } z_{Lk}^{-} = \left(-\frac{2}{3}\right)^{k} \mu_0 \text{ for } \forall k \geq 1, \tag{10}$$

which implies $\left|z_{Lk}^{+}\right| = \left|z_{Lk}^{-}\right|$ for $\forall k \geq 1$, Then, the expected absolute deviation from the Nash equilibrium is the same in the $BCG+$ and in the $BCG-$ games:

$$\sum_{h=0}^{\infty} p_h \left|z_{Lk}^{+}\right| = \sum_{h=0}^{\infty} p_h \left|z_{Lk}^{-}\right|,$$

that is, the strong strategic environment effect is absent asymptotically.

However, (10) implies that

$$\mu_0 < z_{L1}^{+} < z_{L2}^{+} < \cdots < 0 \text{ and } \mu_0 < z_{L2}^{-} < z_{L4}^{-} < \cdots < 0 < \cdots < z_{L3}^{-} < z_{L1}^{-}.$$

Therefore, we have:

$$\left|\sum_{k=0}^{\infty} p_k z_{Lk}^{+}\right| = \sum_{k=0}^{\infty} p_k \left|z_{Lk}^{+}\right| = \sum_{k=0}^{\infty} p_k \left|z_{Lk}^{-}\right| > \left|\sum_{k=0}^{\infty} p_k z_{Lk}^{-}\right|,$$

that is, the weak SEE arises. ■

Presence of the strategic environment effect can be seen also in the logistic version of the level-K model. Figure 4 shows the relationship between mean $|x-60|$ for $BCG+$ (dashed blue) and $BCG-$ (solid red) for $n \in \{2,3,4,5,6,7\}$ for four values of $\lambda$ for the logistic level-k model (LLK). Level-0 and level-1 players in LLK and LCH models follow the identical behavioral rules. For level-2 and above, in LLK, unlike in LCH, level-$k$ players assume that all the others are level $k-1$ who are making noisy choice when computing the expected payoff of choosing each integer. As we can see from the figure, when $\lambda$ is large enough, we observe qualitatively the same strategic environment

effect as we have seen based on LCH. $\lambda$ can be smaller in LLK than in LCH to obtain the strategic environment effect. This is because in the latter, level-2 is responding against the weighted average of the choices made by level-0 and 1, which is inherently noisier than the choice made by level-2 in LLK, which responds against the choices made by level-1.

Furthermore, the results are robust against variation in the mean level as shown by Figure 5 for three mean levels, 1.0 (left), 1.5 (middle), and 2.0 (right) for two values of $\lambda$, 7.5 (top) and 10.0 (bottom).

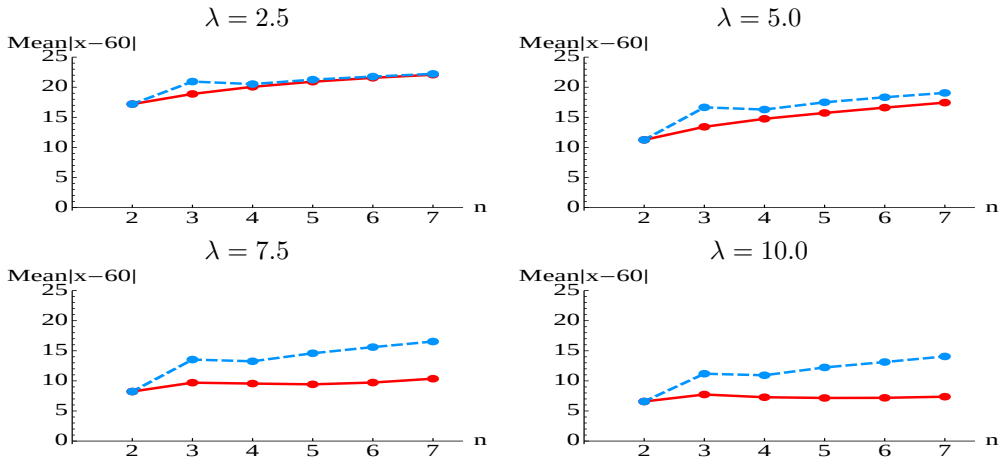



Figure 4: Mean $|x - 60|$ in $BCG+$ (dashed blue) and $BCG-$ (solid red) for $n \in \{2, 3, 4, 5, 6, 7\}$ in LLK for four values of $\lambda$. We assume mean level is 2.

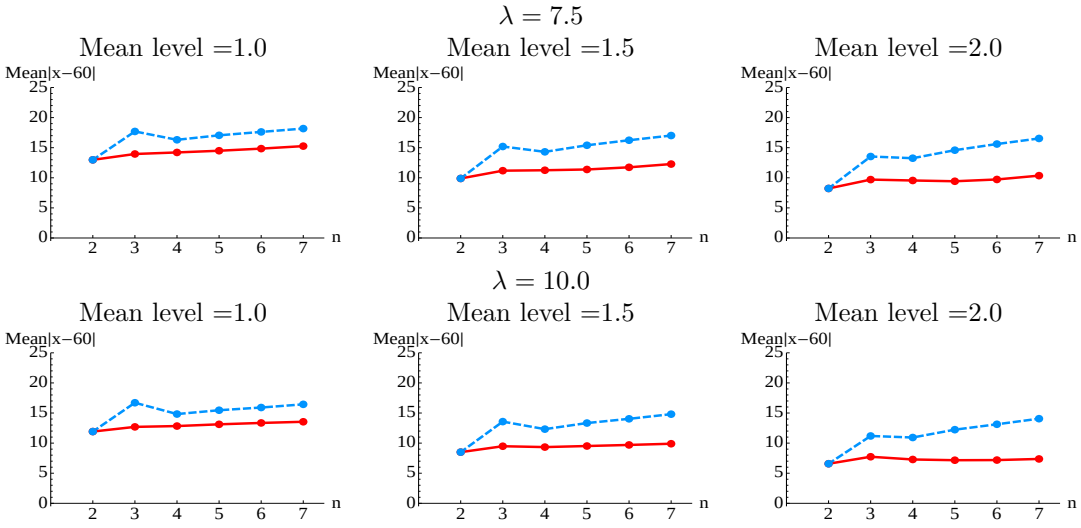



Figure 5: Mean $|x - 60|$ in $BCG+$ (dashed blue) and $BCG-$ (solid red) for $n \in \{2, 3, 4, 5, 6, 7\}$ in LLK for three mean levels, 1.0 (left), 1.5 (middle), and 2.0 (right), and two values of $\lambda$, 7.5 (top) and 10.0 (bottom)

# D Robustness of LCH simulation results with respect to $\lambda$ and the mean level

Figure 6 shows the relationship under LCH between the mean value of $|x-60|$ and $n \in \{2,3,4,5,6,7\}$ for $BCG+$ (dashed blue) and $BCG-$ (solid red) for four different values of $\lambda$, i.e. $\lambda \in \{5.0, 7.5, 10.0, 15.0\}$, while keeping the mean level to be 2.0. As one can see from Figure 6, when $\lambda = 5.0$, for LCH, the difference of the mean $|x-60|$ between two BCGs is small except for $n = 3$. However, for larger values of $\lambda$, the difference in the mean $|x-60|$ between the two BCGs continues to exist for all $n > 2$. It should also be noted that the required value of $\lambda$ to observe the strategic environment effect depends on the magnitude of the winning payoff. As noted in the text, in our computation, we normalized the winning payoff to be 1. If the winning payoff is much larger, then we would obtain the strategic environment effect with smaller values of $\lambda$.

The reason for not observing the strategic environment effect when $\lambda$ is small is simply because the choices for each level of agents are too noisy.

Figure 7 shows the same relationship as in Figure 6 but varying the mean level. It shows the results for three mean levels, 1.0 (left), 1.5 (middle), and 2.0 (right) for two values of $\lambda$, 7.5 (top) and 10.0 (bottom). We consider these three values for the mean level following Camerer et al. (2004) who report that the median estimated mean level, based on 24 beauty contest data sets, is 1.61. As one can easily imagine from the analysis in Section 2, the strong strategic environment effect is driven mainly by the way level 2 and above agents behave. Thus, as the mean level declines, the difference in the mean $|x-60|$ between the two BCGs become smaller. The strategic environment effect, however, is observed even with mean level 1.0.

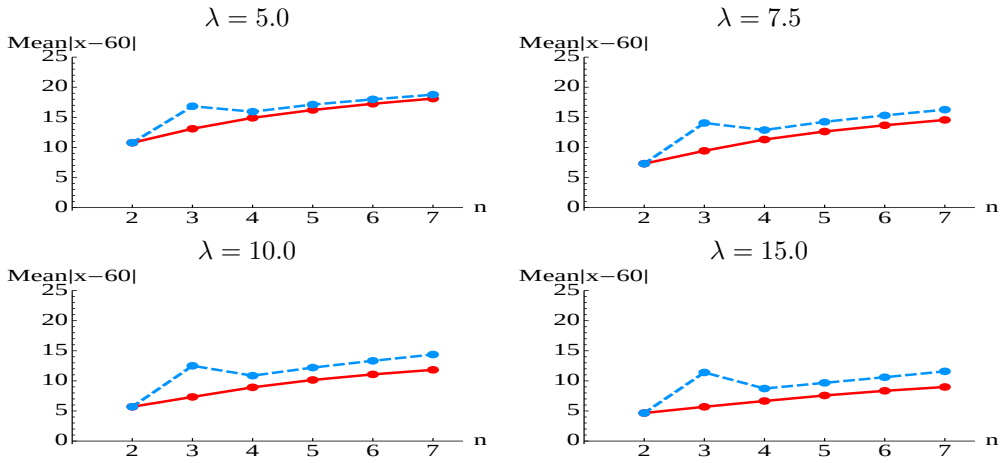



Figure 6: Mean $|x - 60|$ in $BCG+$ (dashed blue) and $BCG-$ (solid red) for $n \in \{2, 3, 4, 5, 6, 7\}$ in LCH for four values of $\lambda$. We assume the mean level is 2.

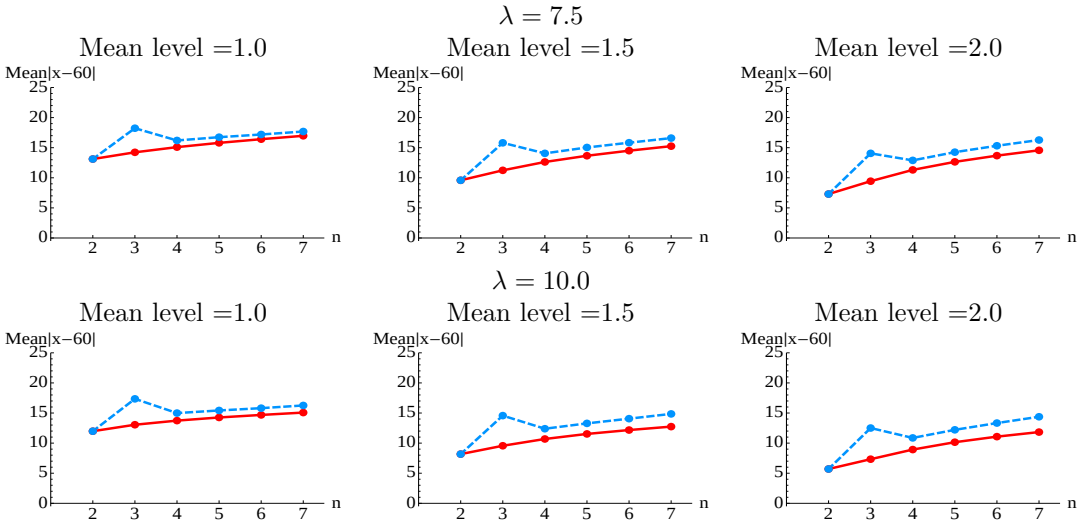



Figure 7: Mean $|x - 60|$ in $BCG+$ (dashed blue) and $BCG-$ (solid red) for $n \in \{2, 3, 4, 5, 6, 7\}$ in LCH for three mean levels, 1.0 (left), 1.5 (middle), and 2.0 (right), and two values of $\lambda$, 7.5 (top) and 10.0 (bottom)

# E Histogram of numbers chosen

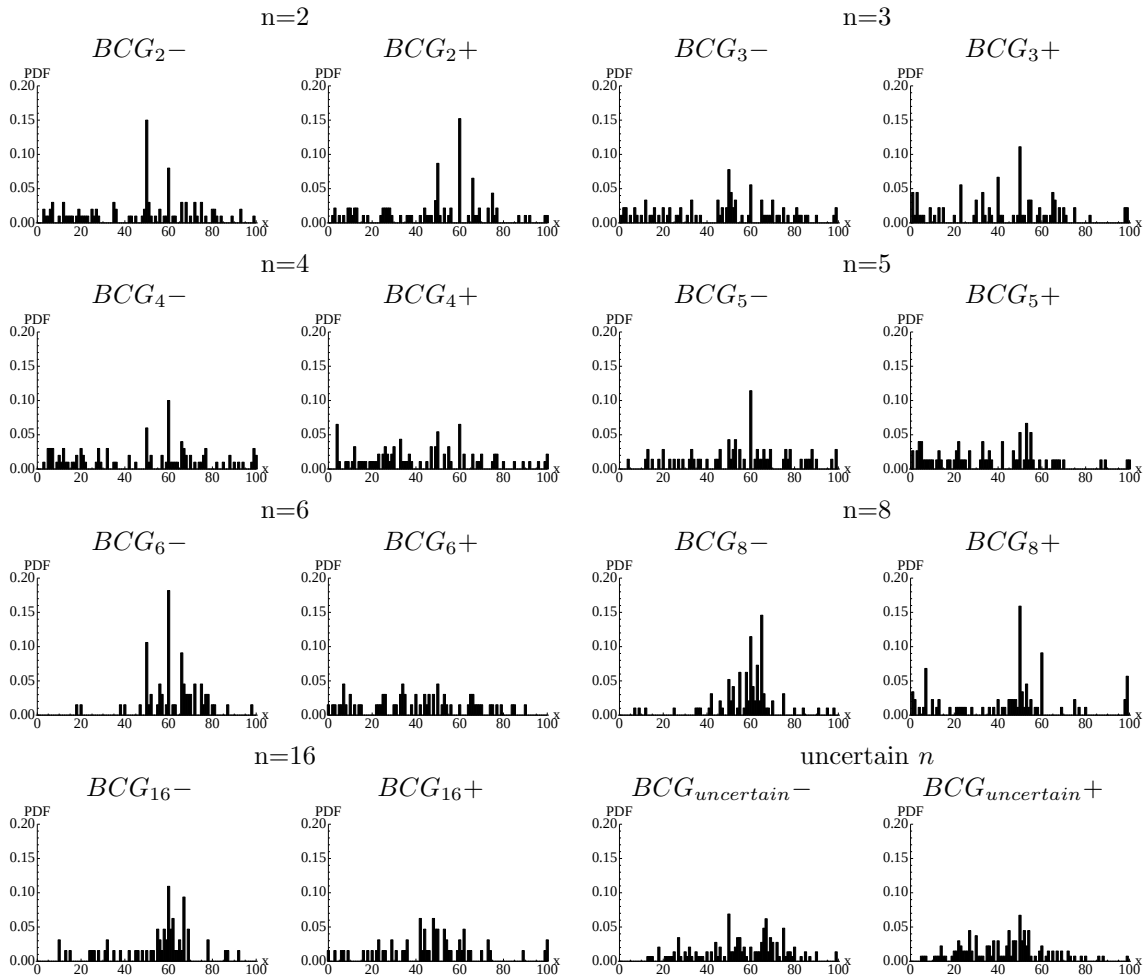



Figure 8: Histogram of numbers chosen by subjects in $BCG_n-$ and $BCG_n+$ for various group size $n$. Here we consider only the first game each subject played.

# F 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 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.

The instructions for the uncertain group size treatment were in the same format. The main difference is the first sentence and the last paragraph of the instructions of each game. For the first sentence in the uncertain group size treatment, instead of stating "You interact with $n-1$ other randomly selected people in this room" as in the other treatments, it reads "You do not know with how many people in this room you will interact. You can interact with some of the people in this room (1, 2, 3, 4 etc) or with all of them. " The last paragraph in the uncertain group size treatment reads, "When everybody has made his/her choice, we will randomly choose with how many people you will interact and the choices will be randomly assembled in groups of that size, and the winners will be computed (and you will be able to take your money if you won). Every group size has equal probability to be chosen." This should be contrasted with the paragraph of the other treatments that reads "When everybody has made his/her choice, the choices will be randomly divided into groups of $n$ and the winners will be determined (and you will be able to take your money if you won)." The examples used in the instructions of the uncertain group size treatment are the same as in the $n = 4$ treatment.

## 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 people 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 people 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:

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

Example : You choose 3 and the 3 other people 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<br>**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 people 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 people 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:

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

Example : You choose 3 and the 3 other people 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<br>**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): _____

## F.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<br>**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<br>**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<br>**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<br>**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 people 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<br>**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 people 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<br>**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 people 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<br>**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 people 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<br>**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 people 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<br>**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 people 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<br>**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 people choose 1, and the 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<br>**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 people choose 1, and the 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<br>**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 |