Heterogeneity in Guessing Games: An Experiment

Liu, Tianwei

Meritco Services

10 November 2016

Heterogeneity in Guessing Games: An Experiment*

Tianwei Liu†

November 10, 2016

*This research is financed by Meritco Services. I thank Jing Chen for her thoughtful comments and suggestions. Sheng Bai provided thoughtful inputs on the experimental design.

†Division of Hedge Fund Applications, Meritco Services. 2607 Tower 1, China World Trade Center, Beijing, 100004, China. E-mail: tianwei.liu@meritco-group.com

Abstract

In interactions under strategic complementarity, naive players have a disproportionally large effect on the aggregate outcome, resulting in a nonlinear relationship between the proportion of sophisticated and naive players and the aggregate outcome. This paper studies this relationship in a beauty contest game by informing some players the game theoretic solution and systematically varying the proportion of informed players. The results show that the conditions predicted by strategic complementarity stand empirical test.

Keywords: Beauty contest, Strategic complementarity, Beliefs, Bounded rationality

1 Introduction

Canonical economic analyses typically assume rational decision makers. However, mounting empirical evidence suggests that human beings frequently and systematically behave in ways at odds with rational prescriptions (Camerer, 2003; Kahneman, 2011). Moreover, there are large individual differences in how closely people resemble the ideal agent economists have in mind (e.g, List, Haigh, & Nerlove, 2005; Palacios-Huerta & Volij, 2009).

When individuals with heterogeneous information processing ability engage in strategic interactions, how do they influence each other and how do their mutual influences shape the aggregate outcomes? The answers depend on a pair of concepts about the specific strategic environment, known as *strategic substitutability* and *strategic complementarity* (Bulow, Geanakoplos, & Klemperer, 1985; Haltiwanger & Waldman, 1985; Camerer & Fehr, 2006): Strategies are substitutes if a change in strategy by one player creates incentives for other players to change their strategies in the opposite direction. In such situations, rational individuals (partly)*correct* the "errors" of less rational individuals, thereby bring the aggregate outcome close to the predictions of rational models. On the other hand, strategies are complements if a change in strategy by one player creates incentives for other players to change in the same direction as that player. In such situations, rational individuals (partly) *mimic* the strategies of less rational individuals, thereby drive the aggregate outcome away from rational predictions.

Consistent with this idea, Fehr and Tyran (2008) found in a price setting game that price adjustment toward the new equilibrium after an anticipated monetary shock was extremely quick under strategic substitutability, yet very slow under strategic complementarity. Along the same line, Sutan and Willinger (2009) studied the beauty contest game (BCG) involving either strategic complementarity or strategic substitutability. In the standard BCG, a group of players each choose a number (real or integer) within $[0, 100]$. The player whose choice is closest to a target number – some parameter p times the average of all chosen numbers – wins a fixed prize. When $p < 1$, the game has a unique Nash equilibrium where all players choose zero, reached by iterated elimination of dominated strategies. It is easy to see that strategies in the standard BCG are complements, and experimental studies have indeed shown that choices in this game are reliably far from zero (Bosch-Domènech, Montalvo, Nagel, & Satorra, 2002). Sutan and Willinger (2009) compared two versions of modified BCGs: one called $BCG+$ where the target number is $\frac{2}{3}(mean + 30)$; the other called $BCG-$ where the target number is $(100 - \frac{2}{3}mean)$. The Nash equilibrium in both games are 60, but strategies are complements in $BCG+$, and substitutes in $BCG-$. Sutan and Willinger (2009) observed that choices were closer to 60 in $BCG-$ than in $BCG+$.

Under different strategic environments, how do the aggregate behaviors change as a function of the relative proportions of rational and irrational individuals? Haltiwanger and Waldman (1985) provided such an analysis. Assume the population is composed of two types of players: a proportion ω of sophisticated agents who always form correct beliefs about what others will do and always best respond, and a proportion $1 - \omega$ of naive agents who all play the same, fixed strategy regardless of ω. It can then be deduced that, under strategic complementarity, the absolute distance between the aggregate outcome and the rational equilibrium, denoted D, is a decreasing and *concave* function of ω. If D is twice differentiable with respect to ω, then we have: $\dfrac{\partial D}{\partial \omega} \leq 0$ and $\dfrac{\partial^2 D}{\partial \omega^2} \leq 0$. A similar analysis shows that under strategic substitution, D is a decreasing and *convex* function of ω. The interpretations of these conditions are straightforward. The first order conditions indicate that, regardless of the strategic environment, the aggregate outcome comes closer to the equilibrium prediction as the relative proportion of sophisticated agents increases. The second order conditions indicate that, under strategic complementarity, naive players have a disproportionally large impact on the aggregate outcome. Given these conditions, $D(\omega)$ is flatter at the lower end of ω and steeper at the higher end of ω, which suggests that adding a few sophisticated agents into a group of naive agents may have limited effect on the group behavior, while adding a few naive agents into a population of sophisticated agents will have a large effect on the group behavior. By the same logic, the opposite is true for strategic substitutability.

This paper puts the above analyses under empirical test. It reports an experiment on the standard BCG that examines how the aggregate behavior changes as a function of relative proportions of sophisticated and naive players, and how sophisticated players behave differently under different group compositions. Importantly, we can directly test whether the actual behaviors fit the specifications of strategic complementarity by Haltiwanger and Waldman (1985). First, let's look at how the conditions of strategic complementarity apply in the BCG.

Assume a standard BCG with a proportion ω of sophisticated players and $1-\omega$ of naive players. The target number is p times the average of all chosen numbers ($p<1$). Following Haltiwanger and Waldman (1985), the naive players choose the identical, fixed strategy C_n. In this context, C_n is a number within $[0,100]$. The number C_n is common knowledge among sophisticated players, thanks to their unlimited ability to form expectations. Because the game is symmetric, choices of sophisticated players must be identical in equilibrium. Therefore, we have $C_s = p[(1-\omega)C_n + \omega C_s]$, where C_s denotes the choice of sophisticated players.[1] This equation can be easily solved for C_s:

$$C_s = p\frac{1-\omega}{1-p\omega}C_n \tag{1}$$

The distance between aggregate outcome and the rational equilibrium is simply the average chosen number of all players, given by $M = \omega C_s + (1-\omega)C_n$. Therefore:

$$M = \frac{1-\omega}{1-p\omega}C_n \tag{2}$$

The function $M(\omega)$ at some values of C_n with $p = 2/3$ is depicted in Figure 1. It can be easily shown that $\dfrac{\partial M}{\partial \omega} \leq 0$ and $\dfrac{\partial^2 M}{\partial \omega^2} \leq 0$ for all $0<p<1$, $0 \leq \omega \leq 1$,and $0 \leq C_n \leq 100$, with strict inequality holds for both conditions when $\omega \neq 1$ and $C_n \neq 0$. Therefore, the BCG satisfies the conditions of strategic complementarity. Note that $C_s(\omega)$ is also a decreasing, concave function, meaning that the choices of sophisticated players should also decrease nonlinearly (first slow, then fast) as the proportion of their own type increases. This pattern can also be put under empirical test.

The existing empirical literature on the strategic environment provides not direct insight for the relationship between group composition and aggregate behavior. First, most studies (e.g., Fehr & Tyran, 2008; Sutan & Willinger, 2009) do not include a measure of individual rationality, and therefore the group composition is unknown. Second, in the studies that do measure individual rationality (e.g., Kluger & Wyatt, 2004), the relative proportion of rational and less rational participants in an experimental session is not systematically varied by the researchers, and most sessions predominantly represent one type of participants.

[1] The choice of sophisticated players can also be computed by iterated elimination of dominated strategies, starting from zero.

Therefore, to study the group composition effect, we need to identify the level of rationality of each participating individual and assemble groups with varying proportions of different types of players. To achieve these goals, the current study "manipulates" rationality (in this specific game) by offering some participants private information. Next, I will deliberate on this approach.

In the context of the BCG, rationality can be approximately interpreted as the ability to understand the iterated elimination of dominated strategies, and hence realize that everyone choosing zero is the only surviving outcome if rationality is common knowledge. Therefore, we can "create" rational agents by directly informing them the reasoning process and the solution. This is done in the written experimental instructions. The wording of the instructions was carefully considered with the goal of being as clear and simple as possible. Game theory jargons were generally avoided so that no special knowledge is necessary for understanding the material (see Appendix A.2 for an English translation of the private information).[2]

The above information is only available to a subset of the participants. The critical treatment is the relative proportion of informed and uninformed players in a group, which systematically varies across experimental sessions. The group composition (i.e., how many players are informed and how many are uninformed) is made common knowledge by public announcement.

The goal of the information manipulation is not to make the informed players more rational in general, but to enhance their knowledge of how canonical theory reasons about this particular game. The BCG is ideal for the current purpose. On the one hand, the game is complex enough that very few subjects can solve it within the time limit of a lab experiment. This is supported by the results of two-person BCGs. When the BCG is played in groups of 2 with $p = 2/3$, the lower number is always closer to 2/3 of the average. Therefore, choosing zero is the weakly dominant strategy and always wins no matter what the opponent chooses. In spite of this simplicity, Grosskopf and Nagel (2008) found that only

[2]As a pilot test, I showed these instructions, along with the rules of the game, to 10 undergraduate students from the same population as the main experiment, and let them read for 5 minutes. Then they were asked to explain what was said in the information with their own words, without looking at the instructions again. All the 10 students were able to communicate the key points (iterated dominance and choosing zero).

9.85% of college students chose zero in the two-person BCG. Since the Nash equilibrium is more difficult to compute under $n > 2$ than $n = 2$, the proportion of subjects who can solve the $n > 2$ game is likely to be even lower. On the other hand, the BCG has been widely used as *the* classical demonstration of iterated dominance in introductory game theory courses, suggesting that its reasoning, once pointed out, is simple enough for novices to comprehend. These features give us confidence that the manipulation will work as intended — that *almost all* informed players will indeed understand the reasoning and can be seen as more rational (in this particular game), while *almost no* uninformed players will be able to figure out the solution and they are therefore less rational.

The informed players in this study are apparently much less omnipotent than the sophisticated players depicted in Haltiwanger and Waldman (1985). Importantly, our informed players do not necessarily form correct beliefs about other players. Also, rationality is unlikely to be common knowledge even when all players are informed. Similarly, the uninformed players in this study might not be as naive as the agents in Haltiwanger and Waldman (1985) as well. Since the group composition is common knowledge, some uninformed players might react to the existence of informed players. Nevertheless, since the strategic environment effects are observed in a variety of contexts where the classification of rational and boundedly rational agents are much less extreme than that of Haltiwanger and Waldman (1985), we expect that the first and second order conditions of strategic complementarity will be met in our design.

The rest of paper is organized as follows. Section 2 details the experimental design. Section 3 articulates the hypotheses to be tested. Section 4 describes the results. Section 5 provides a modified model that better describes the average behavior of the informed players than the $C_s(\omega)$ specified above. Section 6 discusses and concludes the paper.

2 Experimental Design

The game was similar to that studied by Nagel (1995). Choices were limited to integers within $[0, 100]$. The parameter $p = 2/3$. The winner earned a prize of 100 Chinese yuan (About 15 US dollars at the time). In case of a tie the prize was shared by the winners. The game was repeated for 5 rounds. In this paper we only focus on choices in the first round, when learning has not taken place and choices are based solely on beliefs. Results with regard to learning will be discussed in a separate paper. I conducted 15 experimental sessions in a large classroom at Beijing Normal University between January and May 2016. The 232 participants (141 females, $Mean_{age} = 22.6$) were predominantly undergraduate and graduate students enrolled at Beijing Normal University.[3] Economics majors were excluded from participation. There were 13-16 participants in each session.[4]

Based on the proportion of players who were informed of the game theoretic reasoning and solution, there were 6 treatment conditions: the *Baseline* treatment where no player was informed, the *Few-Informed* treatment where 2 players were informed, the *Half-Informed* treatment where half of the players were informed, the *Most-Informed* treatment where all but 2 players were informed, the *All-Informed* treatment where all players were informed, and the *Lecture* treatment where all players were informed by written instructions plus a short lecture (See below for details). The summary of experimental design and the number of sessions conducted in each treatment is available in Table 1.

The experiment was implemented paper based. The procedure of the Baseline treatment is as follows. Upon arrival, participants were seated far apart to prevent communication. Participants first read the written instructions on their own (see Appendix A.1 for an English translation of the instructions). Then one of the two experimenters read the instructions aloud to ensure that the rules of the game were common knowledge. Any questions concerning the rules of the game were answered. Subjects had 4 minutes to write their choice on a paper card. Experimenters then collected the cards and record the choices.

In the Few-Informed, Half-Informed, and Most-Informed treatments, the corresponding number of players were randomly selected to receive the private information. Public verbal announcements from the experimenter included: 1) rules of the game identical to those announced in the Baseline treatment; and 2) the fact that m out of the n players had private information "regarding the game theoretical analysis of this game". These facts should therefore (ideally) be common knowledge.

[3]The rest participants were college students from nearby universities.

[4]I aimed for 16 participants per session. So I assigned 21 slots for each session on the sign-up web page. In the event that more than 16 participants showed up, the extra participants were paid 15 yuan and dismissed. In the event that fewer than 16 participants showed up, the session started as was.

In the All-Informed treatment, the aforementioned equilibrium information was available for all players and this fact was publicly announced.[5]

In the Lecture treatment, the aforementioned equilibrium information was available for all players and this fact was publicly announced. After that, the experimenter gave a 10-minute lecture explaining the equilibrium information. On a number line from 0 to 100 projected on the screen, the experimenter explains step by step why any number larger than $100 \times \frac{2}{3}^k$ cannot win against $100 \times \frac{2}{3}^k$ at Step k, and therefore should be eliminated, leading to the conclusion that 0 is the unique equilibrium. Compared with written instructions only, the lecture may facilitate the formation of common knowledge in two ways: 1) Attending to the same instructions in the same room should strengthen one's belief that a) everyone has received the information, and b) that everyone believes everyone has received the information, and so on; 2) Players may better understand the argument, and have more confidence in other players' understanding of the argument, and have more confidence in other players' confidence in other players' understanding, and so on. Differences between the All-Informed and the Lecture treatment may highlight the role of common knowledge of rationality. [6]

For all treatments, after 5 rounds of play, the participants filled in a questionnaire containing demographic information and some open-ended questions. At the end of the session, participants were paid 25 yuan show-up fee plus any reward they won in the game. The Lecture session lasted for about 50 minutes. Sessions in other treatments lasted for about 40 minutes.

3 Hypotheses

Our main hypothesis is that the results will satisfy the theoretically deduced conditions of strategic complementarity, which can further break down into several testable hypotheses.

On the aggregate level:

[5]The information itself was not read out to the players.

[6]I do not claim that the equilibrium is common knowledge in the Lecture treatment. I only suggest that equilibrium information is *closer to* common knowledge in the Lecture treatment than in the All-Informed treatment.

Hypothesis 1: $\frac{\partial M}{\partial \omega} \leq 0$ and $\frac{\partial^2 M}{\partial \omega^2} \leq 0$. This means that, average choice decreases as the proportion of informed players increases. The decrease is slow when ω is small and fast when ω is large. It allows the possibility that average choice does not significantly decrease with ω when ω is small.

Additionally, we argue that rationality is not common knowledge even when all players are informed, and that the Lecture treatment may facilitate the formation of common knowledge of rationality. This means:

Hypothesis 2: Average choice is lower in the Lecture treatment than in the All-Informed treatment.

For uninformed players:

Hypothesis 3: Average choices of uninformed players do not change with the group composition.

For informed players:

Hypothesis 4: $\frac{\partial C_s}{\partial \omega} \leq 0$ and $\frac{\partial^2 C_s}{\partial \omega^2} \leq 0$. This means that, average choice of informed players decreases as the proportion of informed players increases. The decrease is slow when ω is small and fast when ω is large. It allows the possibility that average choice of informed players does not significantly decrease with ω when ω is small.

4 Results

Descriptive statistics for each session are shown in Table 2. Figure 2A presents the distribution of choices in the Baseline treatment. Among the 47 chosen numbers, the smallest is 16, and only 2 are larger than 67. Therefore, consistent with previous studies, choices are well away from the equilibrium, and weakly dominated strategies are rarely played. Table 3 reports descriptive statistics from the Baseline choices, as well as from some previous p-beauty contest experiments with similar settings (conducted in classroom, on college student populations). The data is generally comparable to those reported in the literature. Noticeably, average choice in the current sample is the smallest among the studies listed here. That said, average choice in this range is common among other populations and settings (Bosch-Domènech et al., 2002). Therefore:

Result 0: *Distribution of the choices in the Baseline treatment resembles those reported in other 2/3-beauty contest experiments.*

Figures 2A-F present the distributions of the choices in each treatment. There is a visible trend that choices decrease as the group contains more informed players. Using Kruskal Wallis H test, I can reject the null hypothesis that choices in all treatments are drawn from the same distribution at the .0001 level. To examine potential non-linearity, I compare each pair of treatments using Mann-Whitney U tests. Table 4 summarizes the results of this analysis. Some notable regularities are:

1. Distributions of choices in the Baseline, Few-Informed and Half-Informed treatments are not significantly different (Medians: 29, 26.5, 31, respectively).

2. As the proportion of informed players continues to increase, choices begin to decrease. Median choice drops from 31 in the Half-Informed treatment to 24 in the Most-Informed treatment ($p = .02$[7]), then to 20 in the All-Informed treatment ($p = .06$).

3. Median choice drops sharply from 20 in the All-Informed treatment to 9.5 in the Lecture treatment ($p < .001$).

The above observations also find support in the comparison of cumulative frequencies of choices (Figure 3): the Baseline, Few-Informed, and Half-Informed treatments are mostly tangled together; while the Most-Informed, All-Informed, and Lecture treatments lie progressively to the left, although not all the comparisons follow strict dominance. To summarize:

Result 1: *Aggregate choices tend to decrease as the proportion of informed players increases, but only when the proportion of informed players is large.* This result supports our

[7] All tests are two-tailed in this paper.

Hypothesis 1.

Result 2: *When all players are informed, an additional public lecture deliberating the reasoning and solution halves the median choice.* This result supports our Hypothesis 2.

Next, we move on to analyze the choices of each type of players. The average choices by player type are presented in Figure 4. We first look at uninformed players. I cannot reject the null hypothesis that choices of uninformed players in the Baseline, Few-Informed, Half-Informed, and Most-Informed treatments (Medians are 29, 26.5, 31, and 31, respectively) are drawn from the same distribution (Kruskal Wallis H test, $p = .35$). Therefore:

Result 3: *Choices of uninformed players are not affected by the group composition.* This result supports our Hypothesis 3.

We then look at informed players. Figure 4 shows a general trend that informed players tend to choose smaller numbers as the proportion of their own type increases. Indeed, I can reject the null hypothesis that choices of informed players in the Few, Half, Most and All-Informed treatments are drawn from the same distribution (Kruskal Wallis H test, $p < .01$). Choices of informed players do not change much from Few to Half-Informed treatment (Median: 30 to 31.5, $p = .98$). As the proportion of informed players increases further, median choice decreases from 31.5 in the Half-Informed treatment to 22.5 in the Most-Informed treatment ($p = .03$), and then to 20 in the All-Informed treatment ($p = .13$), or 9.5 in the Lecture treatment ($p < .001$).

Result 4: *Choices of informed players tend to decrease as the proportion of informed players increases, but only when the proportion of informed players is large.* This result supports our Hypothesis 4. There is a caveat, however, that there are only 6 observations of informed players in the Few-Informed treatment.

There are 32 observations of both informed and uninformed players in the Half-Informed treatment, which allows a direct comparison between the two types of players. No difference is found in this comparison (Mann-Whitney U test, $p = .99$). Moreover, neither of these two distributions significantly differs from the Baseline treatment ($ps > .5$). This indicates that telling as many as 8 out of 16 players how the game should be played has completely no effect on how the game was actually played. Since the informed players do react to the group composition, we can rule out the possibility that the majority of informed players simply are not influenced by the information. Therefore, the informed players in the Half-Informed treatment seemed to *choose* to play high numbers after taking the group composition into consideration. I will further discuss this point in Section 5.

Result 5: *In the Half-Informed treatment, choices of informed and uninformed players do not differ.*

5 Estimates of Average Informed Play

Recall that under the assumptions of Haltiwanger and Waldman (1985), the choice of sophisticated players follows Equation (1): $C_s = p\dfrac{1-\omega}{1-p\omega}C_n$. This model provides insights into the general pattern of choices changing as a function of group composition, which leads to our hypotheses. However, it does not adequately describe the actual behaviors of our informed players. Most saliently, the model predicts that everyone will choose zero in the All-Informed treatment, which is far away from our observations.

We now modify this model to make it better reflect the realities in this experiment. First, we retain the assumption that all informed players best respond to their beliefs, but acknowledge the possibility that informed players may doubt if other informed players will act like themselves. They may not believe that everyone has understood the private information, or they may not believe that no one would consider the possibility that someone might not have understood the information. To account for these potential doubts, in all treatments except for the Lecture treatment, we assume that all informed players have the same belief that only a proportion π ($0 \leq \pi \leq 1$) of the informed players will actually behave like informed players.[8] To avoid introducing new parameters, the rest $1 - \pi$ informed players are

[8]Because the Lecture treatment was designed to reduce these doubts, π should be larger in the Lecture treatment than in other treatments, which we shall show later.

assumed to act like uninformed players. In other words, informed players give ω a *discount* and behave as if only $\pi\omega$ players are informed.

Second, informed players are unlikely to always form correct expectations about what the uninformed players will choose. In this regard, we assume that informed players all have the same belief about the average choice of uninformed players, denoted as C'_n, and that C'_n is common knowledge among informed players. Importantly, C'_n does not necessarily reflect the actual average choice of uninformed players.

With these two modifications, the best response is now given by: $C_s = p[(1-\pi\omega)C'_n + \pi\omega C_s]$. Solving the equation for C_s yields:

$$C_s = p\frac{1-\pi\omega}{1-p\pi\omega}C'_n \tag{3}$$

Table 5 reports the least squares estimates for π and C'_n. The assumptions that all informed players have the same beliefs about π and C'_n are highly simplified. These parameters may be better interpreted as the means of the distributions that describe the corresponding beliefs. Accordingly, we look at how well the model predicts the *average choice* at each value of ω as an approximation for goodness of fit. As visualized in Figure 5, the predicted choices by the model are very close to the actual averages except for the Few-Informed treatment.[9]

The estimated values of the two parameters are worth discussing. The estimated value of C'_n implies that an average informed player believes the average choice of uninformed player is 56.6. This belief may seem unusually high. However, we note that the parameter C'_n represents the latent belief derived from choices, given best response, not participants' stated beliefs. Stated beliefs usually reveal higher levels of reasoning than latent beliefs (Costa-gomes & Weizsäcker, 2008). Nevertheless, the high estimated latent belief seems to imply that informed players may have *underestimated* the average sophistication of uninformed players. This is a potential explanation for why informed players choose similar numbers as uninformed players in the Few and Half-Informed treatments.

The estimated value for π implies that informed players expect that 3/4 of all informed players will act in a rational way. This further confirms that our information manipulation is successful. The Lecture treatment was designed to facilitate formation of common knowledge

[9]The less precise prediction for the Few-Informed treatment is quite understandable since there are only 6 data points at this value of ω and therefore the estimation gives it a small weight.

of rationality. Assume that C'_n stays the same, we can estimate π in the Lecture treatment with least squares. This estimate is 0.88. Comparing the Lecture treatment with the All-Informed treatment, a 16% increase in π leads to a 60% decrease in average choice. This again shows that under strategic complementarity, behaviors are very sensitive to changes of rational expectations when the average level of rationality is already quite high.

6 Discussion

Strategies are complements in the beauty contest game. If a player believes other players will choose high numbers, she should choose high numbers as well. Therefore, limitedly rational players should have a disproportionally large impact on the aggregate behavior in the BCG. By informing a subset of players the game theoretic solution and systematically varying the proportion of informed players across sessions, this paper shows that the conditions predicted by strategic complementarity stand empirical test.

Other researchers have also studied the BCG with players of heterogeneous strategic sophistication. Slonim (2005) studied the competitions between experienced and inexperienced players. Experienced players are those who have already played the game for several rounds. The results showed that, inexperienced players do not behave differently whether their opponents are experienced or not, while experienced players tend to choose higher numbers when they face inexperienced than experienced new opponents. These results are related to our findings that uninformed players are not sensitive to who they play with but informed players are. However, there are important distinctions between the two studies: Our informed players are more sophisticated in the sense that they have a better idea how the game *should be* played. The experienced players in Slonim (2005) are more sophisticated in the sense that they have a better idea how the game *are actually* played.

Using the strategy method, Agranov et al. (2012) studied the BCG with undergraduate students playing against a varying mixture of random-choosing computers and graduate students. They found that, players systematically lower their choices as the group contains more graduate students, because they believe the graduate students are more sophisticated than the random-choosing algorithm. So why haven't our uninformed players choose lower numbers as they know there are more players who have private information "regarding the game theoretical analysis of this game"? I argue for two reasons: First, the gap in strategic sophistication between graduate students and random-choosing computers is more obvious than that between informed and uninformed players. Second, the within-subjects design of Agranov et al. (2012) allows players to systematically adjust their choices based on the group composition, which is impossible in our current between-subjects design.

Although the group compositions are discussed in terms of relative "proportions" in this paper and in the analysis of Haltiwanger and Waldman (1985), I do not claim that the group size does not matter. In the context of the BCG, Ho, Camerer, and Weigelt (1998) showed that 7-player groups converge to equilibrium faster than 3-player groups. More recently, Hanaki, Sutan, and Willinger (2016) found that choices in $BCG+$ and $BCG-$ only differ when group $n \geq 5$, but not when $n < 5$. Future studies can investigate whether and how the relationship between the group composition and the aggregate behavior depends on the group size.

The number of informed and uninformed players is always publicly announced in this experiment. It might be informative to run a treatment where the group composition is hidden from the participants. Although we do not have that data, given the prevalence of ambiguity aversion (Ellsberg, 1961), I suspect that informed players, not knowing the group composition, will choose numbers *no smaller* than our informed players in the Half-Informed treatment. Therefore, it is possible that if we tell every player the equilibrium solution, but do not tell them how many others are also informed, the group may act as if no one has any information. The test of this hypothesis is left for future work.

References

Agranov, M., Potamites, E., Schotter, A., & Tergiman, C. (2012). Beliefs and endogenous cognitive levels: An experimental study. *Games and Economic Behavior*, *75*(2), 449–463.

Bosch-Domènech, A., Montalvo, J. G., Nagel, R., & Satorra, A. (2002). One, two, (three), infinity, ... : Newspaper and lab beauty-contest experiments. *American Economic Review*, *92*(5), 1687–1701.

Bulow, J. I., Geanakoplos, J. D., & Klemperer, P. D. (1985, June). Multimarket Oligopoly: Strategic Substitutes and Complements. *Journal of Political Economy*, *93*(3), 488-511.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press.

Camerer, C., & Fehr, E. (2006). When Does "Economic Man" Dominate Social Behavior? *Science*, *47*.

Costa-gomes, M., & Weizsäcker, G. (2008). Stated Beliefs and Play in Normal Form Games. *Review of Economic Studies*, *75*(3), 729–762.

Ellsberg, D. (1961). Risk, ambiguity, and the savage axioms. *The Quarterly Journal of Economics*, *75*(4), 643-669.

Fehr, E., & Tyran, J.-R. (2008). Limited rationality and strategic interaction: The impact of the strategic environment on nominal inertia. *Econometrica*, *76*(2), 353–394.

Grosskopf, B., & Nagel, R. (2008). The two-person beauty contest. *Games and Economic Behavior*, *62*(1), 93–99.

Haltiwanger, J., & Waldman, M. (1985). Rational expectations and the limits of rationality: An analysis of heterogeneity. *The American Economic Review*, *75*(3), 326-340.

Hanaki, N., Sutan, A., & Willinger, M. (2016, March). *The Strategic Environment Effect in Beauty Contest Games* (GREDEG Working Papers No. 2016-05). Groupe de REcherche en Droit, Economie, Gestion (GREDEG CNRS), University of Nice Sophia Antipolis. Retrieved from https://ideas.repec.org/p/gre/wpaper/2016-05.html

Ho, T. H., Camerer, C., & Weigelt, K. (1998). Iterated dominance and iterated best response in experimental 'p-beauty' contests. *American Economic Review*, *88*(4), 44–69.

Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.

Kluger, B. D., & Wyatt, S. B. (2004). Are judgment errors reflected in market prices and allocations? Experimental evidence based on the Monty Hall problem. *Journal of Finance*, *59*(3), 969–997.

List, J. A., Haigh, M. S., & Nerlove, M. (2005). A simple test of expected utility theory using professional traders. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(3), 945–948.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *American Economic Review*, *85*(5), 1313–1326.

Palacios-Huerta, I., & Volij, O. (2009). Field Centipedes. *American Economic Review*, *99*(4), 1619–1635.

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

Sutan, A., & Willinger, M. (2009). Guessing with negative feedback: An experiment. *Journal of Economic Dynamics and Control*, *33*(5), 1123 - 1133. (Complexity in Economics and Finance)

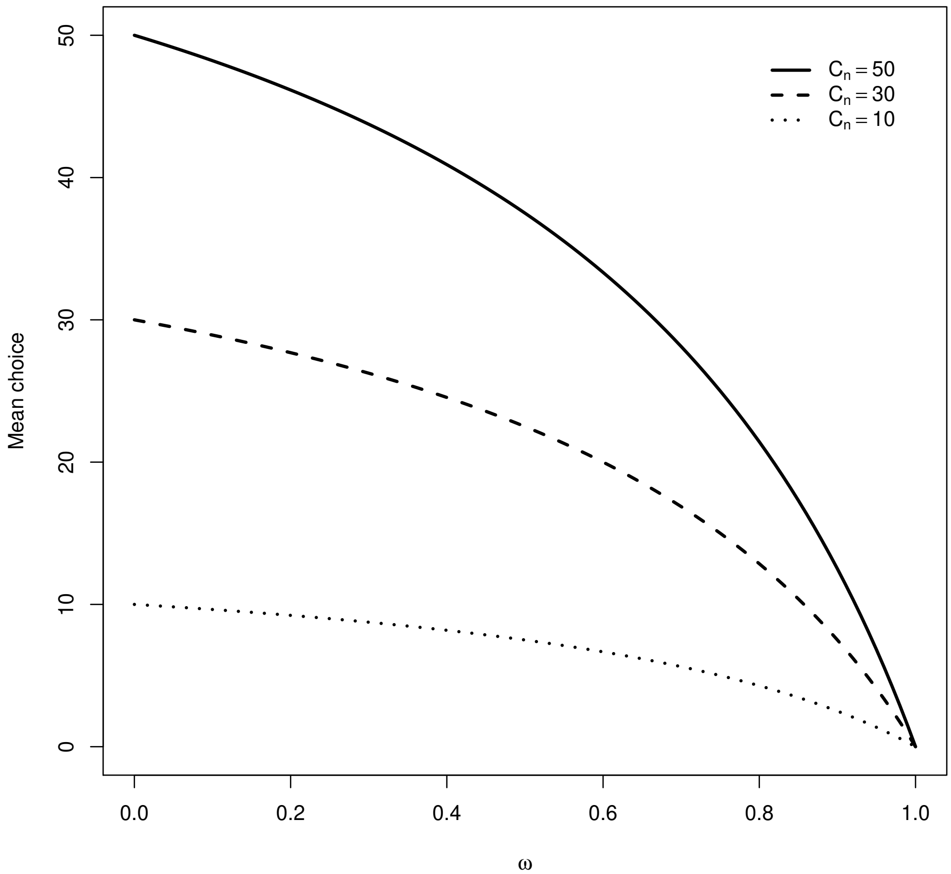

Figure 1: $M(\omega)$ with $p = 2/3$

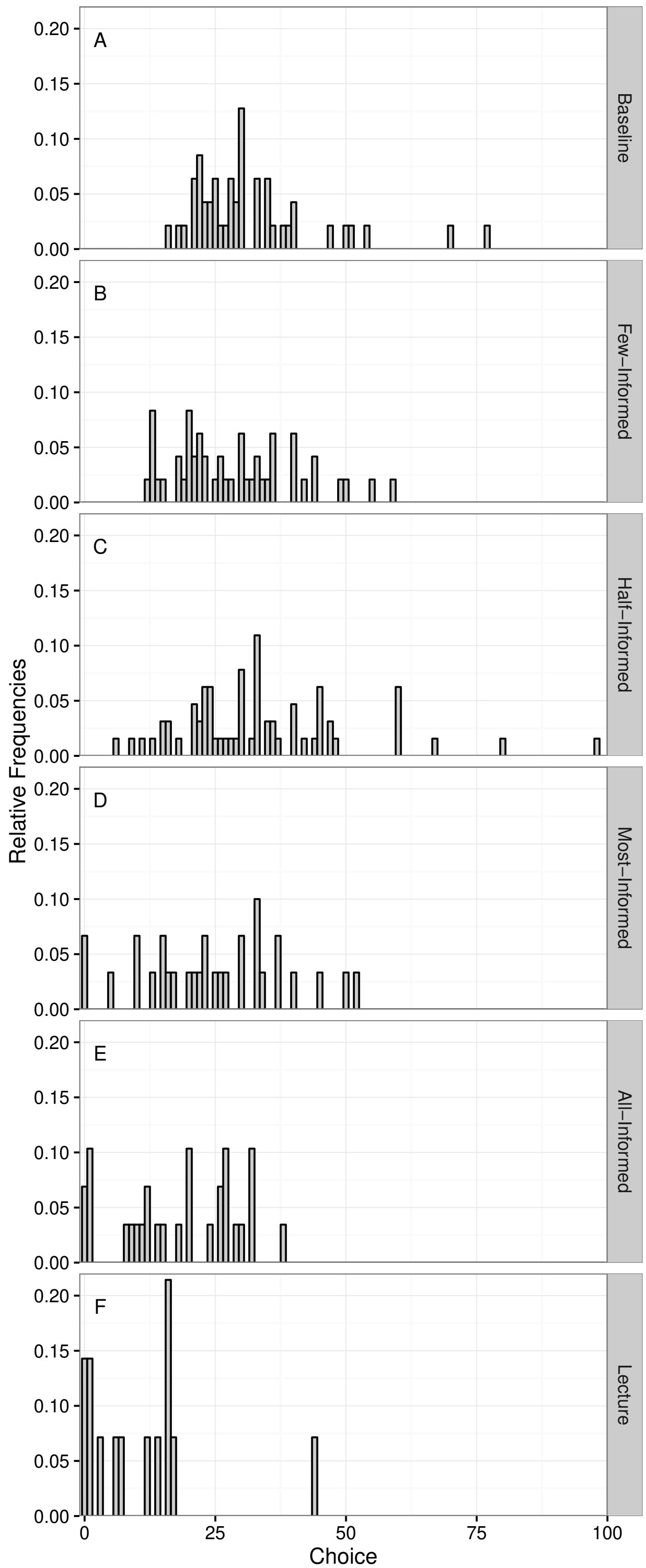

Figure 2: Distributions of choices in each treatment

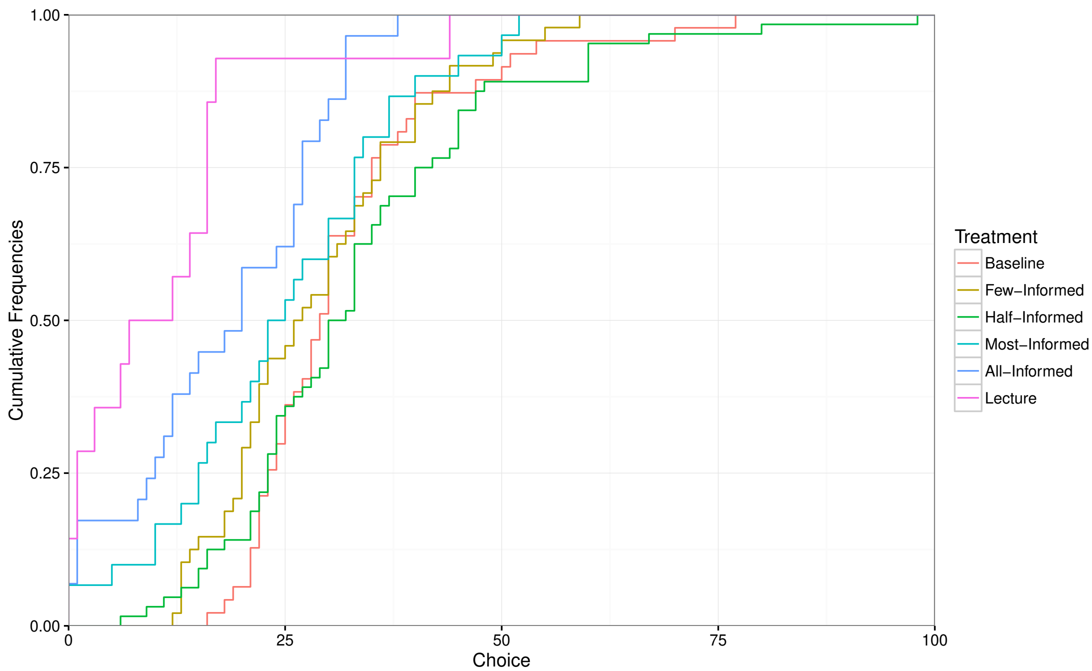

Figure 3: Cumulative distributions of choices for each treatment

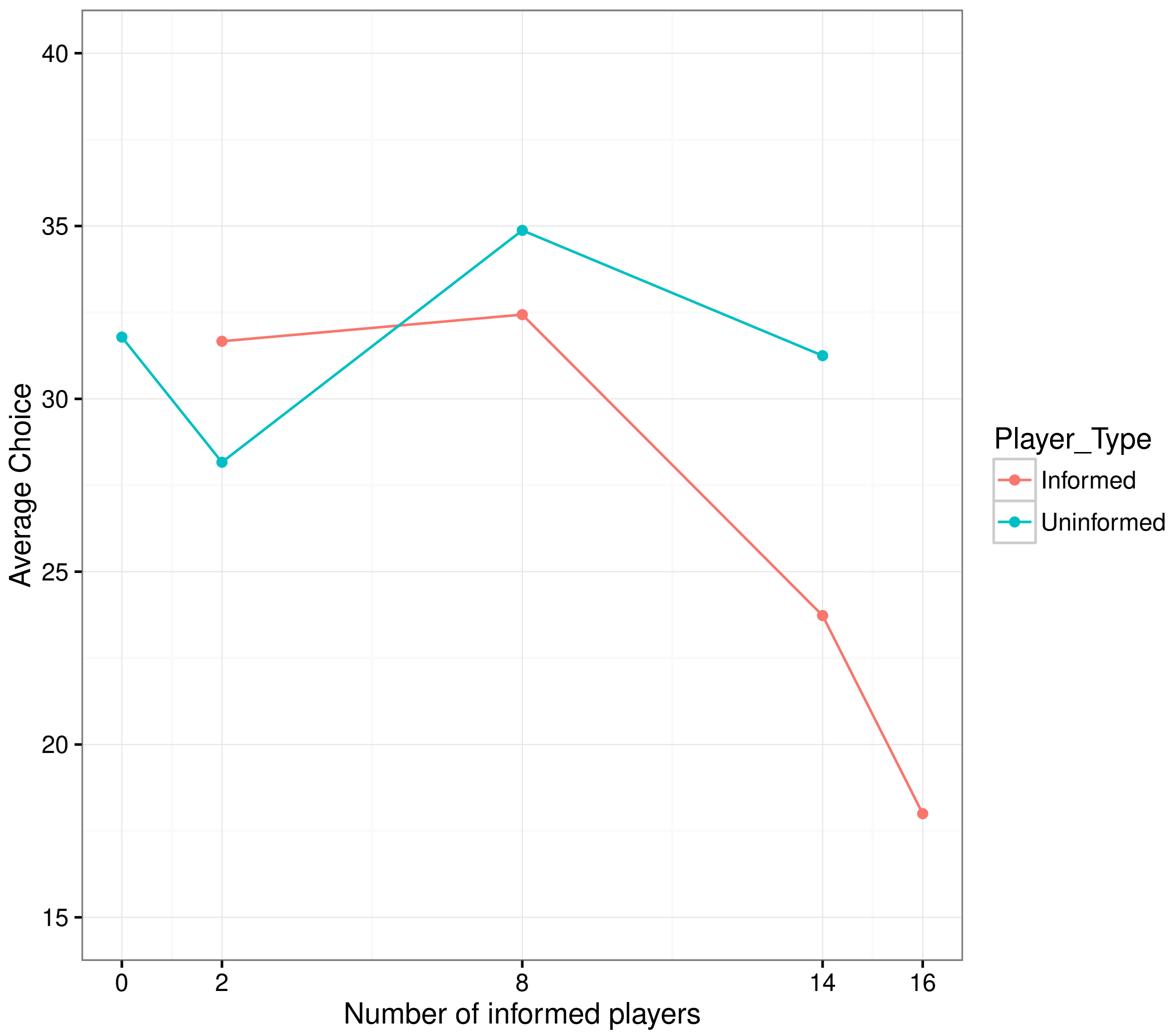

Figure 4: Average choice by player type as a function of number of informed players in a group. Lecture treatment not included.

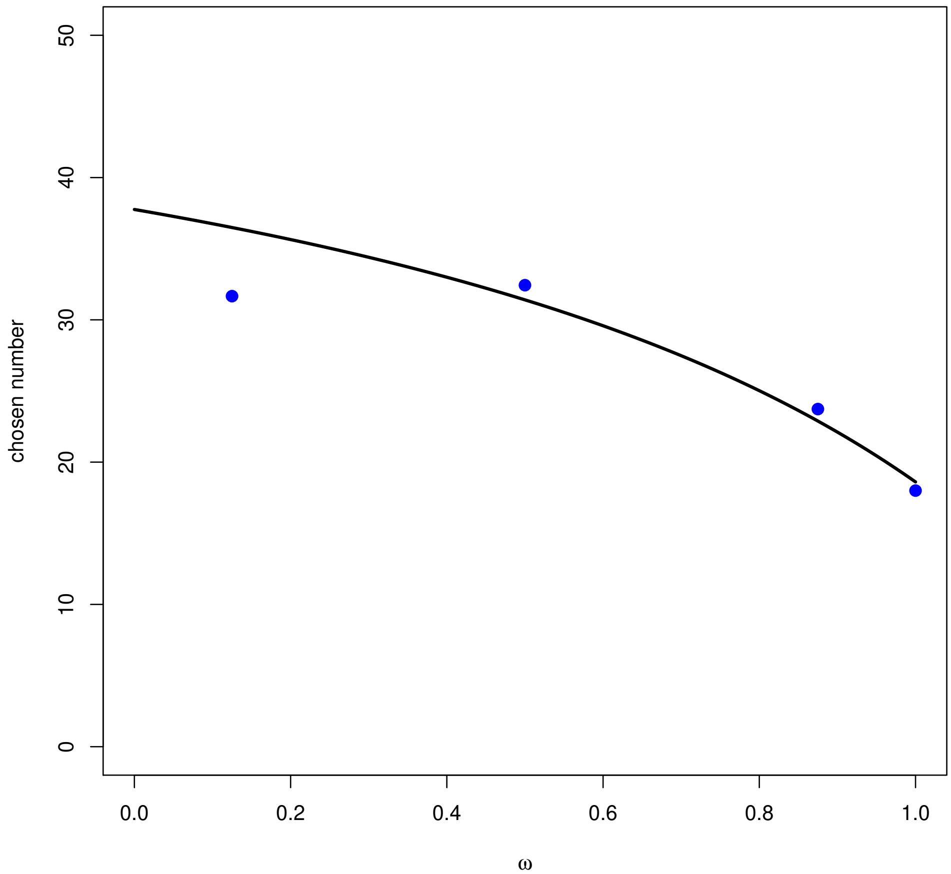

Figure 5: Predicted choice and actual average choice of informed players

Table 1: Summary of experimental design for each treatment

Treatment	N of sessions	Player type	Written instructions	Verbal announcement
Baseline	3	Uninformed	RG[a]	RG
Few-Informed	3	Uninformed	RG 2/16 players have AI[b]	RG 2/16 players have AI
		Informed	RG 2/16 players have AI AI	RG 2/16 players have AI
Half-Informed	4	Uninformed	RG 8/16 players have AI[c]	RG 8/16 players have AI
		Informed	RG 8/16 players have AI AI	RG 8/16 players have AI
Most-Informed	2	Uninformed	RG 14/16 (or 12/14) players have AI	RG 14/16 (or 12/14) players have AI
		Informed	RG 14/16 (or 12/14) players have AI AI	RG 14/16 (or 12/14) players have AI
All-Informed	2	Informed	RG All players have AI AI	RG All players have AI
Lecture	1	Informed	RG All players have AI AI	RG All players have AI AI

[a] RG = Rules of the game and other baseline instructions

[b] AI = Additional information (the equilibrium analysis)

Table 2: Descriptive statistics for each session

Session	Treatment	N	N informed	Mean	Median	Std. Dev.
Base1	Baseline	15	0	30.2	29	10.3
Base2	Baseline	16	0	31.6	29	13.3
Base3	Baseline	16	0	33.5	29.5	13.8
Few1	Few-Informed	16	2	26.3	22.5	10.9
Few2	Few-Informed	16	2	27.6	22.5	13.8
Few3	Few-Informed	16	2	31.9	32.5	10.0
Half1	Half-Informed	16	8	30.6	33	9.8
Half2	Half-Informed	16	8	32.1	30	16.0
Half3	Half-Informed	16	8	41.8	34.5	24.5
Half4	Half-Informed	16	8	30.1	30	10.2
Most1	Most-Informed	16	14	25.2	24	12.5
Most2	Most-Informed	14	12	24.2	24	14.9
All1	All-Informed	16	16	19.1	25	11.8
All2	All-Informed	13	13	16.7	15	10.9
Lect1	Lecture	14	14	10.9	9.5	11.6

Table 3: Descriptive statistics for Baseline choice and other beauty contest experiments

	Mean	Median	Std. dev.	Group size
Baseline treatment	31.8	29	12.4	15-16
Nagel (1995)	37.2	33	20	14-16
Ho et al. (1998)	40	35	24.8	7
Kocher and Sutter (2005)	34.9	32	-	17-18
Agranov et al. (2012)	35.1	33	21	8
Luccasen (2013)	33.5	30	17.2	18
Cubel and Sanchez-Pages (2016)	36.1	33	23	110-170

Table 4: Comparisons of choices under each pair of treatments

Treatment	1	2	3	4
Baseline	29			
Few-Informed	26.5	26.5		
Half-Informed	31			
Most-Informed		24	24	
All-Informed			20	
Lecture				9.5

Note: Entries are medians from each treatment. Treatments that appear in the same column are not different from each other at the .05 level of significance.

Table 5: Parameter estimates for informed play

Parameter	Estimate	Std. Err.
C'_n	56.63	4.37
π	0.76	0.06
π (Lecture)	0.88	0.04

Appendix A Experimental Instructions

A.1 Instructions for the Baseline Treatment:

Welcome to the experiment. Please read the following instructions carefully.

1. Do not communicate:

Throughout the experiment, please do not talk to other participants or communicate in any other way. Please do not make comments no matter you win or not. This is a competition game. Communicating with others will not help you win, and will compromise the reliability of our data. If you have any questions, please raise your hand and we will come to assist you.

2. Rules of the Game:

Unlikely some experiments you might have participated, there is *no deception* in this game. Therefore, all the information provided to you is real.

You will play a game that repeats for *5 rounds*. In each round of the game, everyone chooses an *integer* between 0 and 100 (including 0 and 100). Please write your choice on an answer card given to you. We will collect all the cards, and calculate the average of all chosen numbers. The average multiplied by 2/3 is called the *target number*. The player whose choice is closest to the target number (i.e., 2/3 of the average) wins 100 yuan. Other players win nothing. In case there are multiple winners, the 100 yuan reward will be split evenly among them. For example, if 3 players choose the same number and this number is closest to the target number, then each of the 3 players wins 33.33 yuan. The same game will repeat for 5 rounds, with 100 yuan reward for each round. After each round, we will announce all chosen numbers, the average, the target number and the winner's choice on the screen in front.

3. Payment:

At the end of the experiment, you will receive 25 yuan show-up fee, plus all the reward you win in the game. We will pay you via Alipay transfer.

4. ID:

Each player has a unique ID, which is written in the upper-left corner of their answer cards. We use this ID to identify players. Please remember your ID.

5. Confidentiality:

The data will only be used for research purposes.

We will conduct similar experiments in the near future. So please do not mention the details of this experiments to other people. We have no control over this matter but we trust you.

A.2 Equilibrium information (Half-Informed treatment)

You have additional information. In this group, 8 players have this additional information, and the other 8 players do not.

Additional information:

Because choices are restricted to numbers between 0 and 100, 2/3 of the average must be between 0 and 66.67. Therefore, 67 must be closer to 2/3 of the average than any number within $[68, 100]$. Therefore, any number within $[68, 100]$ cannot win against 67. Therefore, *a rational player will not choose a number larger than 67.*

One step further, if all players are rational, then all chosen numbers will be between 0 and 67, and 2/3 of the average has to be between 0 and 44.67. Therefore, any number larger than 45 cannot win against 45. Therefore, *a rational player who believes all other players are also rational will not choose a number larger than 45.*

This thinking process can go on infinitely, until all numbers are eliminated except 0. So, if all players:

1) are rational;

2) believe all players are rational;

3) believe all players believe all players are rational;

...

ad infinitum,

The game has only one stable way of play: *every player chooses 0.*

This is the additional information you have. Remember, all the inferences are based on the corresponding premises. How to use this information in the actual game depends on your own judgment.