Documento de Trabajo/
Working Paper

IESA 13-04

Teaching Nash Equilibrium and Strategy Dominance: A Classroom Experiment on the Beauty Contest

Virtudes Alba Fernández
Universidad de Jaén y CentrA

Pablo Brañas Garza
Universidad de Jaén, IESA-CSIC y CentrA

Francisca Jiménez Jiménez
Universidad de Jaén y CentrA

Javier Rodero Cosano
CentrA y LINEEX

Teaching Nash Equilibrium and Dominance: A Classroom Experiment on the Beauty Contest[1]

Virtudes Alba-Fernández[2] Pablo Brañas-Garza[3]
Francisca Jiménez-Jiménez[4] Javier Rodero-Cosano[5]

9th September 2004

[1]The authors wish to thank Rosemarie Nagel, Curro Martínez Mora, Quique Fatás and the participants at the *Economic Science Ass. Meeting in Tucson* (2003), specially David Reiley. We also acknowledge valuable comments from three anonymous referees. We also thank Samuel Hernández for his contribution to the data gathering and processing and Martha Gaustad for the language revision. Pablo Brañas acknowledges the hospitality of IESA-CSIC during this research. We gratefully acknowledge the financial support provided by the University of Jaén R+D program (# 20210/148).

[2]Dep. Statistics, University of Jaén & centrA, Spain

[3]Dep. Economics. University of Jaén, IESA-CSIC & centrA, Spain.

[4]Dep. Economics. University of Jaén & centrA, Spain.

[5]Economics Division, centrA & LINEEX, Spain. Corresponding author: centrA, C/Marqués de Larios, 3. 29015 Málaga. Spain. email: jrodero@fundacion-centra.org. Tel: + 34 951 03 96 10. Fax: + 34 951 03 96 01.

Abstract

The aim of this investigation is to show how the use of classroom experiments may be a good pedagogical tool to teach the Nash equilibrium (NE) concept. For our purposes, the basic game is a repeated version of the Beauty Contest Game (BCG), a simple guessing game in which repetition lets students react to other players' choices and converge iteratively to the equilibrium solution. We perform this experiment with undergraduate students with no previous training in game theory. After four rounds, we observe a clear decreasing tendency in the average submitted number in all groups. Thus, our findings prove that by playing a repeated BCG, students quickly learn how to reach the NE solution.

Keywords: Classroom Experiments — Beauty Contest Game — Teaching — Nash Equilibrium

J.E.L. Classification: A22, C99, D83

Reasons for this study

In recent years, Game Theory (GT) has become an essential part of intermediate microeconomics theory. Obviously, the key concept in Game Theory is the Nash Equilibrium (NE), making this powerful notion an inevitable part of our teaching. Educators, however, know that it is not an easy task to explain the NE in class. Students usually find difficulties in understanding the idea behind it and, consequently, the implicit features of this concept.

Generally, the teaching of GT begins with the study of strategic interactions among players considering only pure strategies. Once players understand that each player's payoffs depend not only on their own actions but also on those of other players, the next step is to explain the process of elimination of dominated strategies. To do this, only two basic assumptions are required: rationality (maximization) and common knowledge (of rationality). Both concepts let us solve (dominance solvable) games and also predict some particular behavior. However, there exist games that are not dominance solvable even if they have a NE. In this case, the NE requires an additional hypothesis: common knowledge of players' beliefs about their rivals' actions. Hence, the best response becomes the suitable mechanism for reaching the equilibrium solution. Best response is the best strategy that a player may choose given the strategies chosen by his rivals. The NE is reached when all agents play (some of) their best response. Thus, the NE is self-enforcing because no player has incentives to deviate from it. Clearly, the use of the NE eliminates circular reasoning (player 1 thinks that player 2 thinks that player 1 thinks ...) in games with at least one pure NE, not in a game with a unique mixed NE.

This paper proposes a pedagogical tool for simplifying the teaching of the NE: successive repetitions of a dominance solvable game, the Beauty Contest Game (BCG). This particular game is very useful for showing, in an intuitive

manner, the two procedures described above. Specifically, the game is played four times in the form of a classroom experiment, prior to teaching a game theory class.

In short, our findings clearly illustrate that students inexperienced in GT apply both iterated elimination of dominated strategies and best-reply behavior in each round to the previous choices. On average, we observe a recursive approximation to the Nash prediction, which we call the *learning effect*. Therefore, by repeating this static game, we ensure that students learn how to solve it in some way or another.

The rest of the paper is organized as follows. Firstly, the theoretical background of the basic Beauty Contest Game and its wide variety of experimental designs are presented in the next section. Our specific classroom experiment is described in detail in the two following sections. In the post-experimental session, we propose explaining the relevance of the Nash equilibrium and dominance concepts to the students, relating them to their own previous performance. We then analyze our findings and suggest some other variations which could also be used as classroom BCG experiments. Finally, conclusions are reached in the last section.

Theoretical Background

Since the introduction of the concept in Nash (1951), the Nash equilibrium has not only become the standard tool for the economic scientific community, but also the basis for the systematic teaching of the discipline in the modern era.

In game theory, it is common practice to assume that individuals "magically" choose a set of actions such that all the (infinitively recursive) predictions come true. So as to reach this hard-to-believe solution, theorists

can follow one of the two original Nash interpretations: either agents are perfectly rational or there exists an evolutionary equilibrium[1]. As is well known, the first interpretation relies on the assumption that all agents are able to compute the game equilibria and reach one of them. In contrast, the second interpretation presupposes the existence of a large population of ordinary people that play the game in an evolutionary framework: that is, they pick up one strategy randomly, if the outcome is good, they repeat it again or if the strategy is bad, they will disappear. After some time, the game will converge to one of the Nash equilibriums. The BCG is a good example of the underlying convergence to both interpretations. From a rational point of view, it is not credible that the subjects will solve this problem (few people can at the first try!), but if it is repeated, the solution will converge to equilibrium most of the time.

The Beauty Contest Game is a simple guessing game that makes it easy to evaluate individuals' level of reasoning. The basic BCG is as follows: a certain number of subjects are invited to play a game in which all of them must simultaneously choose a real number from an interval (generally $[0, 100]$). The winner is the player who chooses the number that is closest to p-times the mean of all the numbers chosen, where $1 > p > 0$. The winner receives a prize, the losers get nothing. Under these rules the unique Nash equilibrium is zero for all players[2]. Other parameter configurations may produce different or even multiple equilibria (see Bosch-Domènech et al., 2002, for further details).

As Stahl (1996) points out the distribution of chosen numbers lets us

[1]For reasons of space, Nash only uses the rationality explanation in the published version.

[2]This is not as obvious as it looks, Grosskopf and Nagel (2001) try a one-shot BGC with game theorists. The answers were very different from zero, but better than the usual ones!

analyze the depth of reasoning of the agents, say, level 1 includes people who expect that the other players behave randomly so they choose $p * mean$ (being $mean = 50$ if the choice distribution is uniform), level 2 contains people expecting that the other's depth of reasoning is level 1, and, choosing then $p^2 * mean, \ldots$, generalizing, at level K are people who choose $p^K * mean$ because they believe that the other people are at level $K - 1$. If K is big, $p^K * mean \simeq 0$, so if we repeat the process *ad infinitum* ($K = \infty$) we reach the theoretical solution, 0, the highest level of reasoning (figure 1 represents this process with $p = 2/3$). Random answers are called level 0 of reasoning. A level of reasoning higher than 3 is rare in BCG experiments (see Bosch-Domènech et al., 2002), although this Iterated Best Replay Behavior (IBRB) is quite commonly found (Ho et al., 1998).

[Figure 1 about here.]

Figure 2, taken from Ho et al. (1998, pg. 951), shows the convergence to the zero theoretical solution from a dominance iterated point of view[3]. Given $p = 2/3$, any number chosen between 66.6 and 100 is dominated by $66.6 = 100 * 2/3$. Hence the interval $[66.6, 100]$ corresponds to irrational behavior (what we call $R(0)$). Rational individuals will always choose a number in the $[0, 66.6]$ interval. Applying the same reasoning, $R(1)$ players will choose a number below 66.6, but above $44.4 = 2/3 * 66.6$, while $R(2)$ players will choose a number in the interval $[29.6, 44.4]$. Following this iterated reasoning level process *ad infinitum*, we reach the unique Nash equilibrium (0, with $R(\infty)$). Thus, this game is called dominance solvable. In his book, Camerer (2003, chapter 5) describes some other games that also tend to unravel and would be fun to teach in class. These include the "patent race" game, in which iteration eliminates some strategies but not others; centipede games,

[3]See Rapoport and Amaldoss (2000) for an experimental study of iterated elimination of (strongly) dominated strategies.

price competition or traveler's dilemma, in which the dominance reduces the collective payoffs to players and the "dirty faces" game, in which an increase in the number of steps of iterated deletion increases payoffs.

[Figure 2 about here.]

Nevertheless, this game makes a much more important point in practice: if a subject plays the Nash equilibrium (0), in the majority of the cases he will not win. Note that the key to this game is to take one more step of deletion of dominated strategies than the other players (but not too many)[4]. Hence, the NE is useful for knowing where the adaptive process leads, but it is not a useful normative theory as a basis for advice[5].

The original idea behind the BCG was first put forward by Keynes (1936) to express that a clever investor has to "anticipate the basis of conventional valuation a few months hence, rather than ... over a long term of years" (page 155), if he is to act in the stock market before other investors do.

The formal game model was later introduced by Moulin (1986). As we noted above, the unique equilibrium of the game is 0 for $p < 1$ and is obtained

[4]Obviously, this process of reasoning involves knowing who you are playing with. In this case, an individual who knows the zero-solution could also guess that most people will not achieve the zero solution. Thus, clever individuals will be able to link their answer to their estimation of the average rationality level. We call this rationality level ∞-plus (Keynes' 'clever investor'). Brañas-Garza et al. (2001) report some examples of ∞-plus behavior arising from internet BCG experiments. Bosch-Domènech et al. (2002, pg. 1693) found strong evidence of this behavior, reporting some comments from students who followed this strategy. Grosskopf and Nagel (2001) argue that most individuals think that other people are not fully rational, thus explaining why the equilibrium is not reached immediately.

[5]This is one of the main advantages to using the BCG for teaching purposes. Nagel (1999), using the same graph shown in Figure 2, illustrates how to explain iterated elimination of dominated strategies in a classroom setting.

by iterated elimination of weakly dominated strategies[6].

After this basic framework was laid down, some experimental researchers began to investigate BCG or "*p-beauty*" (see Ho et al., 1998). The first experimental study can be found in Nagel (1994, 1995)[7]. Other studies have been done by Bosch-Domènech et al. (2002); Bosch-Domènech and Nagel (1997a,b); Duffy and Nagel (1997) and Ho et al. (1998). See also Nagel (1998) for a survey of the literature.

A wide variety of BCG experimental designs can be found[8]. In these, sometimes participants are students, at other times they are professors or newspaper readers, some designs are one-shot, others are repeated, communication versus non-communication or laboratory versus field experiments.

Generally, BCG is run with individual subjects. Kocher and Sutter (2001), however, compare individual versus group behavior in this type of game. They find that although groups do not apply deeper levels of reasoning, they do, in fact, learn faster.

Repetition permits individuals to learn dynamically from other people's expected behavior. Our experiment is very similar to one in Ho et al. (1998). In their design, individuals are given information about previous period choices and, therefore, the learning process is based on an evolutive game. In contrast, Weber (2003b) argues that learning with the BCG can happen even without feedback. This author found that there was convergence towards the NE in the game even when subjects did not receive any information between periods.

Over the last years, research on the BCG has returned to Keynes'original

[6]It can also be obtained by using a best-response argument, see Nagel (1999).

[7]The main purpose of Nagel (1994, 1995) was to contrast an iterated best-reply dominance model (IBRB).

[8]See Camerer et al. (2003b) for a list of 24 different BCG experiments.

idea. Hirota and Sunder (2002) experimentally explore whether price bubbles in security markets are generated by a beauty-contest mechanism: if dividends are paid beyond investors' personal investment horizons, investors must create beliefs about others' beliefs (second-order beliefs) that depend on third-order beliefs, which in turn depend on fourth-order beliefs, and so on. They conclude that when the realization of the dividend is distant and well beyond investors' investment horizon, investors find it difficult to induce the fundamental value of securities from the future to the present (backward induction). This difficulty gives rise to price bubbles because, in this case, investors adjust their expectations on the basis of observed prices (forward induction).

Up to now, the BCG has primarily been used to study the depth of individuals' level of reasoning. However, in this paper we propose a new application of this framework: classroom experiments with a pedagogical aim[9].

The experiment

The experiment involves using a repeated version of the Beauty Contest Game with several groups of students to teach the Nash equilibrium; specifically the iterated elimination of dominated strategies. Given that the Nash equilibrium is a general topic in both intermediate microeconomics and industrial organization (and obviously in Game Theory, too), we expect this pedagogical tool to be of great help to educators.

Three fundamental issues must be taken into account when running the experiment:

[9]However, see also Nagel (1999) where she describes this application.

1. The group should be large enough to reduce the effect of any one person's guess on the average. On the other hand, a very large group may be too cumbersome for the instructor[10].

   Nagel (1995) uses 12 and 17 subjects in each group; Ho et al. (1998) reduce the group size to 7 subjects and 3 subjects (like Kocher and Sutter, 2001). Following these papers, we use small groups (5-6 subjects) and large groups (10-11 subjects)[11]. Each group plays its own game in which each subject plays a repeated BCG against the other members of the group.

2. In order to motivate students, some type of reward should be given. In our classroom experiment, the winner of each round was awarded 0.25 extra-credit points applicable towards the final exam. In the case of several winners in the same round within the same group, the prize was split among them.

   Our experiment was performed in an Intermediate Microeconomics Theory course during the spring semester of the 2002/2003 academic year with three different groups: Business majors (morning and afternoon groups, B-1 and B-2) and Business + Law majors (one single group, B+L). All groups had an identical individual answer handout, instructor and grading system (ranging from 0 to 10). Note that this subject is a requisite in both majors.

3. Last of all, another important issue is the number of rounds that experimental subjects must play in the BCG. As our aim was for students to learn the Nash equilibrium through the iterated elimination of dom-

[10] Ho et al. (1998) show that smaller groups (3 subjects in their research) need more time to converge.

[11] It is difficult to have identically sized groups in a class. Usually, the instructor does not have the exact number of students required.

inated strategies, several rounds were needed to observe this learning effect. Kocher and Sutter (2001) prove that 4 rounds are enough to approximate the theoretical prediction.

The experiment lasts about one hour and the following post-experimental session half an hour (note that it is worth running both sessions consecutively, approximately 1.5 hours).

Procedures

1. The instructor has to define the size of each group. However, it is irrelevant if one group differs slightly in size from the other groups. Nevertheless, for purposes of comparison, it is desirable to have similarly sized groups.

2. After determining group size, the instructor randomly selects one monitor for each group from the student pool. Each monitor is awarded 0.25 points. The monitors' task is to help the instructor to record and monitor the experiment in his group. Each monitor is given a calculator and a monitoring sheet (see appendix). The monitor keeps track of the individual choices in his group and calculates the mean and the winning number. The monitoring sheets are crucial to explaining the results after the experimental session as these findings will easily illustrate basic concepts of Game Theory, namely the iterated deletion of dominated strategies.

3. The third task consists of creating groups, that is, students are randomly assigned to different groups.

4. In order to make the monitor's tasks easier, the instructor should seat each group in a single row (or column).

5. The instructor then gives the monitors their instructions and explains the procedures to them. When this is done, the monitors hand out the instruction sheets and the individual answer sheets to the experimental subjects (see appendix).

6. The instructor must explain instructions out loud. It is important to avoid numerical examples as the generation of focal points is immediate in this kind of game. Any questions are answered aloud. Experimental subjects are told that speaking is absolutely forbidden[12].

7. When everyone fully understands the rules, the experiment can begin. Round 1: Experimental subjects fill in their answer sheet. The monitor then collects all guessed numbers and calculates the mean and $2/3 * mean$. By comparing this value to those reported by the experimental subjects, the monitor determines the winner. The group is told the mean, $2/3 * mean$ and the winner's guess (but not who the winner is). Nagel (1999) also gave the students the complete set of guesses [13].

8. Round 2: Following round 1, the students are informed that they will play another round, independent of the previous one. The procedure for round 2 is identical to round 1.

9. Round 3 and 4: Proceed as above.

[12]Subjects are required to keep a maximum level of confidentiality for their own sake. Note that if any subject were to know his rival's guess, he could use the best reply rule.

[13]In order to simplify the process, we did not give our students the entire set of numbers. Nonetheless, the type of information provided in our experiment affected the convergence speed: Weber (2003a,b) analyzes the effect of different feedback conditions on learning. He found that convergence towards the equilibrium occurs under all circumstances, although it is faster with feedback. See also the Variations section below.

Post Experimental Discussion

Immediately following the experimental session, the post-experimental discussion is begun. In this second part of the experiment, the theoretical background of the NE is introduced. This step involves three topics:

a. First, we present the iterated elimination of dominated strategies underlying the BCG by means of Figure (2). By explaining this process, students can understand that if rationality is common knowledge, nobody will choose a number within the interval $[66.6 - 100]$ because this subset of numbers is dominated by $[0 - 66.6]$. Following this reasoning ad infinitum —as explained previously— students fully capture this idea.

   At this moment it is a good idea to show students the first round results of the groups involved in the session. Alternatively, a simple histogram of the pooled data may be used so as to cover up any embarrassing answers (for more details, see the Results section). Students may compare their own chosen numbers with those reported in Figure 2. Moreover, it could be fun and interesting to show other experimental results so as to offer students the opportunity to compare themselves to others (for example, high school students, 80-year olds, corporate CEOs or even game theorists). Nagel, 1999 and Camerer et al., 2003b report a high number of BCG experiments with several contexts and subject pools[14].

b. Once common knowledge rationality is clearly defined, it is time to explain *best response* behavior. Students observe that their own performance depends on their beliefs about other players' actions. Furthermore, they understand that other players' behavior also depends

[14] In general, these studies reflect a substantial regularity across very different groups, except some highly trained ones.

on their beliefs about others and so on.

As an example, we present some preliminary results of their own responses. To do this, some monitoring sheets (see appendix) can be used to illustrate the average behavior of the groups. Since the average chosen number decreases through successive rounds, students realize that GT approximately predicts their performance.

c. We now have all the necessary ingredients to "cook up" the *Nash equilibrium* concept. Let $\overline{x}$ be player i's forecast of the mean of all numbers submitted. Player i chooses x_i with the objective of $min_{x_i \in [0,100]} \mid x_i - p\overline{x} \mid$. All players behave identically. Common knowledge of beliefs, together with mutual best response, imply that in any equilibrium, we must have $x = px$. Hence, the unique equilibrium must be that $x = 0$, the number which is equal to p times itself.

   Since students have experienced the recursive "dynamics" of the NE, we explain what the abstract concept implies and how this can be reconciled with the real world. In the theoretical model, we suppose that everything is instantaneously adjusted; the BCG must be explained as if it were a slow motion picture: repetition lets individuals reach the theoretical solution step by step. Perfectly rational agents do not need this, but nobody -not even economists- is fully rational. This is the second remarkable point: repetition could be a good substitute for rationality. In fact, the modern view in game theory is that equilibrium arises from adaptation, evolution, communication or imitation. Since these processes take time, equilibration should not occur instantly in one-shot games.

   Nagel (1998), Stahl (1996) and, most recently, Camerer et al. (2003b) present very precise formal models of limited rationality that accurately describe the experimental results of these games. Specifically,

Camerer et al. (2003a) develop a non-equilibrium model for one-shot games with decision rules for players doing different steps of thinking[15]. Their prediction for dominance solvable games will converge to the Nash equilibrium as a limit result.

Finally, as an anecdote, we mentioned *A Beautiful Mind*, the movie about John Nash's life. The film tries to explain the NE to ordinary people using an allegorical sequence. In this scene, John is in a pub with some friends when a group of girls walk in. If all the guys try to pick up only the prettiest girl, all the girls would leave and all the guys would be out of luck. The BCG can be seen as a similar symbolic game: individuals have to compete to outguess the rest of the participants, but without cooperation, costly errors can occur. Knowing game theory concepts could reduce these costs[16].

Results

The experimental BCG was played in three different sessions using two samples of first-year Business undergraduates (B-1 and B-2) and one sample of first-year Business+Law undergraduates (B+L). Table 1 shows the distribution of subjects participating in the experimental sessions by group. Our entire sample consists of 139 subjects assigned to 6 groups per session (5 in

[15]The link between thinking steps and the iterated deletion of dominated strategies lies in the fact that players who are not thinking strategically randomize, while players doing k steps of thinking accurately predict what lower-level players do and best-respond accordingly.

[16] It is important to remind students about the importance of their intellectual achievement: John Nash won the Nobel Prize thanks to this simple discovery. See Varian (2002) for a nice explanation about Nash's brilliant and beautiful contribution. In contrast to what viewers of the film might think, it is clearly not a new strategy to pick up girls. Myerson (1999) gives an interesting historical perspective to this important innovation.

B+L). All groups played four rounds of the BCG.

[Table 1 about here.]

Although our interest specifically lies in the process of learning the NE, we have considered it meaningful to show first period behavior. In this way, we are able to relate our experimental results to the theoretical reasoning levels shown in Figure 2. Figure 3 shows the relative distribution of first-round iterated steps. As we can see, the behavior of subjects in session B-1 and session B+L are quite similar, while the subjects in session B-2 seem to have a higher level of reasoning than the other two[17].

[Figure 3 about here.]

At an aggregate level, Table 2 shows the average chosen number and the standard deviation across sessions and rounds.

[Table 2 about here.]

As can be seen, the mean of the chosen numbers decreases throughout successive rounds in all cases[18]. However, we observe that the path to convergence is not identical. Figure 4 illustrates each trend in the experimental session.

[Figure 4 about here.]

[17] Camerer et al. (2003a) estimate the average number of thinking steps at 1.61 across 24 one-shot beauty contest games. They find quite low estimates $(0-1)$ when $p > 1$ and high estimates $(3-5)$ in games where the equilibrium is within the interval.

[18] Although the initial average observed in our experiment is similar to that of Kocher and Sutter (2001), our final average is very different! After running 4 rounds as well, their final averages are around 7, while ours are clearly higher (from 9.7 to 19.1).

At first sight, sessions B-1 and B+L behave similarly, but the trend in B-2 shows a flatter slope. As our main interest is to study the *learning effect*, we analyze the average variation more accurately round after round. We define *speed of convergence* as: $(\mu_i - \mu_{i-1})/\mu_{i-1}$ where μ_i is the average guessed number in round i ($i = 1, \ldots, 4$). This speed of convergence is shown in Table 3.

[Table 3 about here.]

Note that all the sessions share a similar initial speed of convergence close to 20%. However, this speed varies dramatically within and between groups from round 2 to 3 and from round 3 to 4 . Thus, the learning effect is clearly observable in all rounds and sessions.

After this descriptive approximation, we statistically analyze the results obtained. More specifically, we are interested in testing if students are induced to modify their behavior towards the Nash equilibrium by repeatedly playing a BCG a set number of times. Therefore, we expect the average number in a round to be different from the following round and, furthermore, that the difference between one round and the following round be greater than zero, that is, that it will decrease. In order to verify this, we formulate the following null hypothesis:

$$H_0 : \mu_{ij} - \mu_{kj} = 0, i < k \text{ and } i, k = 1, ..., 4$$

and alternative hypothesis:

$$H_1 : \mu_{ij} - \mu_{kj} > 0, i < k \text{ and } i, k = 1, ..., 4$$

where i, k represent two consecutive rounds and j represents the session number (1 for B-1, 2 for B-2 and 3 for B+L, respectively).

The statistical analysis is developed in depth in the appendix. Briefly, we check if each population follows the normal distribution. If one of them does not follow the normal distribution, we use a parametric and a non-parametric test to contrast our null hypothesis. Our main results show that:

- In session B-1 the learning effect is observed in the entire experiment.
- In session B-2 the learning effect is observed in the last two rounds.
- In session B+L the learning effect is also observed in all rounds.

To sum up, these results indicate that this is a powerful tool to aid students in reaching the Nash equilibrium.

Variations of the Experiment

This section describes several variations of the classroom experiment proposed in our paper. As stated above, there exist multiple possible variations of the beauty contest game that could be used as a classroom experiment. Some of the most important of these are discussed below.

The two-person beauty contest game

This version of the game is interesting because, unlike the general BCG in which there are more than two players, in a two-person game the player who picks the lowest number wins (if $p < 1$). Thus, the subject who plays the Nash equilibrium strategy (0) always wins. Therefore, 0 is not only a weakly dominant strategy, but also the best response for all the choices of the other person. For this reason, this variant is a good example of the attractive power of the equilibrium point. As Camerer et al. (2003a) point out, this special

game can be solved by only one step of weak dominance. In their model, all players using one or more thinking steps will choose zero. However, it is important to recall that the most important point of the experiment is to show students the strength of the equilibrium concept, even for complicated games.

However, Grosskopf and Nagel (2001) show that usually subjects do not chose the zero in the one-shot game and then, do not reach the equilibrium immediately. Even more, under full information students need some repetitions to realize that zero is the Nash equilibrium. Then, authors conclude saying that "convergence toward equilibrium is driven by imitation and adaptation rather than self-initiated rational reasoning".

Nonetheless, an interesting option could be to perform both the $n = 2$ and $n > 2$ experiments in the same class. Some groups can play the first version, while the others play the second one. It would then be interesting to compare both versions in the post-experimental session and show students that the NE sometimes fails to describe the reality if the convergence process is too long (see Grosskopf and Nagel, 2001).

The entire class experiment

If the class is not large enough, an interesting alternative is to perform the experiment using the whole class as a single group. The main advantage to this is that no student needs to sit out as a monitor. Also, the individual effect on the population average would decrease slightly. In contrast, if the group is too large, the process can become a bit cumbersome for the teacher.

The most straightforward way to run this BCG is for the instructor (or any "ad hoc" volunteer) to write the submitted numbers on the board while another volunteer calculates the mean and $p * mean$. Note that all guessed

numbers should be collected beforehand.

Web-based experiments

In his web page at Virginia University[19], Charles Holt provides some excellent internet software for playing the BCG (among many other games). Not only can students play the BCG on this web site but also check their own results and performance on line as well as comparing their outcomes with other previous experiments.

However, it should be noted that running the experiment in this way requires having a computer room with internet access and banning its use in the classroom. Excepting this "minor" problem, Veconlab is an amazing tool for teaching. Another point to take into consideration, in our case, is the fact that Veconlab is only available in English at this time

Other variations

There are several additional variations on the game. A good idea would be to run some of these simultaneously with different groups in the same class in order to compare the different outcomes arising from different rules.

- When $p > 1$, Nagel (1998) shows that this variation of the game is not trivial. She considered two cases: (1) under $p > 1$ and the interval $[0, 100]$, there are two equilibria: everybody chooses either zero or 100. The latter is the perfect equilibrium[20]. (2) For $p > 1$ and the interval $[100, 200]$, the only equilibrium is 200; the upper bound of the interval.

[19] http://www.people.virginia.edu/ cah2k/

[20] Nevertheless, there are no dominated strategies in this version and thus no process of iterated elimination of dominated strategies leading to equilibrium.

The iterated deletion process starts from the lower bound $100p$, $100p^2$ and so on[21].

- Using the median instead of the mean, for example, can help to illustrate to our students that the game does not change and the convergence towards the equilibrium is basically the same under different conditions (see Nagel et al., 1999).

- By using large and small groups, identical results may be found in the one-shot game, but faster convergence to the equilibrium in larger groups. Duffy and Nagel (1997) studied the influence of a single player on aggregated performance. They use three treatments changing the order statistic from mean to maximum or median for $p = 1/2$, so the extreme choices have very different weights.

- Giving students different feedback. When comparing feedback-free BCG with 'normal' BCG we may observe that convergence is slower when subjects are not informed about the numbers chosen by their partners. The lack of information constrains subjects' learning to their own experience (see Weber, 2003a and Weber, 2003b). In our case, we gave students very little information (only aggregated data). Given that one of the goals of the experiment was to show students the attractive power of the equilibrium, we preferred a low information variant[22].

[21] In this case, the number of eliminations is finite. That is the reason why these games are called finite threshold games (FT). See also Ho et al. (1998).

[22] As Weber (2003b) points out, even if you do not give any feedback, convergence will be achieved, although the pace will be slower. It will take more time to make the calculation, although the rounds will be faster. It is up to the instructor to choose the desired mix of information/convergence speed.

Final Remarks

In this paper we have formulated an interesting classroom experiment: a repeated version of the Beauty Contest Game. Our experience is that this sort of game is a good way to introduce the complexity of the equilibrium concept. Not only does it permits students to see the mechanism in action, but also to comprehend the difference between a theoretical abstraction and real world dynamics towards equilibrium.

The BCG is one of those tricky games worth playing: although it is plainly simple in its formulation and the solution is always obvious *ex-post*, you realize the difficulties of outsmarting other people while having a good time doing it.

The classroom experiment needs some careful preparation to make conveniently mixed groups. However, once everything is ready to go, if the monitors are well trained, rounds should run quite smoothly and increasingly fast. Although for our purposes four rounds are enough, it would be easy to play as many as ten rounds if necessary. Continuous repetitions would be very rewarding for students, as this would let most of them reach the theoretical solution on their own before the post-experimental session.

Once the BCG is done and the post-experimental session is concluded, it is possible to run other experiments so students realize how game theory training helps them to solve interesting (and sometimes lucrative) puzzles.

Appendix

Experimental Instruments

All these documents were originally written in Spanish.

Instructions

You are going to participate in a microeconomic experiment. You have been randomly assigned to a 6-10 person group. In this game you will have to make decisions repeatedly in four rounds. Your partners will be the same throughout the four stages. Moreover, your group has been assigned a monitor that will oversee the procedure.

The rules of the game in each period are as follows: you should choose a (integer or decimal) number in the interval $[0, 100]$. You are allowed to choose zero and one hundred. Once the monitor has collected your group's choices, the winner will be determined. The winning number is the number closest to 2/3 of the average of all the numbers chosen by your group:

$$x = \frac{2}{3} \cdot \frac{\sum_{i=1}^{n} x_i}{n}.$$

The winner of each round will receive a prize of 0.25 extra points valid towards the Micro II final exam. If two or more people are equally close to x, the prize will be split equally among them. A person so lucky as to guess the right number all four times will get a whole point!

First you must write down the group and code you were given at the top of the attached sheet. Then in each round, you must select your number. When everyone has finished, the monitor will collect your responses. Afterwards, you will be given back your sheet with some additional information: the average of your group, 2/3 of this average and the winning number. This procedure will be repeated four times.

If you have any questions, please raise your hand and the instructor will come to you. You are not allowed to speak during the experiment.

Monitoring Sheet

Monitor: _____ Group: _____ Session: _____

	Full Name	Round 1	Round 2	Round 3	Round 4
Subject 1					
Subject 2					
Subject 3					
Subject 4					
Subject 5					
Subject 6					
Subject 7					
Subject 8					
Subject 9					
Subject 10					

Average				
2/3 of average				

Individual Answer Sheet

Code: _____ Group: _____

First Round

Group Average: _____

Your Choice: [ ] 2/3 of the average _____

Winning Number: _____

Second Round

Group Average: _____

Your Choice: [ ] 2/3 of the average _____

Winning Number: _____

Third Round

Group Average: _____

Your Choice: [ ] 2/3 of the average _____

Winning Number: _____

Fourth Round

Group Average: _____

Your Choice: [ ] 2/3 of the average _____

Winning Number: _____

Statistical Analysis

Let X_{ij} be the number chosen by a student at round i in session j, where $i = 1, ..., 4$ and $j = 1, 2, 3$. Let $F_{ij}(X)$ be the distribution function associated to each variable X_{ij}.

First, we verify if each $F_{ij}(X)$ follows a normal distribution with mean μ_{ij} and standard deviation σ_{ij}. To do so, we use the Kolmogorov-Smirnof goodness-of-fit test. Table 4 shows the results:

[Table 4 about here.]

As we can see, individual choices adjust to a normal distribution except in the third round of the first session (B-1) of the game.

Using a parametric test for the equality of means between paired rounds, the results are the following:

[Table 5 about here.]

Observe that it is not possible to apply this test to the third round of the first session (B-1) because it does not follow a normal distribution.

From the p-values shown in Table 5 we can say that the difference between successive average choices is statistically positive in all cells except for the first two rounds in the second session (B-2). In this last case, the null hypothesis of mean equality is accepted.

In order to include the third round of the first session in the statistical analysis, we apply the Wilcoxon non-parametric test using the same hypothesis as above. Results are summarized in the following table:

[Table 6 about here.]

Given these p-values, only the means of the first two rounds in the second session are equal, while the null hypothesis is rejected for the remainder of cases.

References

Antoni Bosch-Domènech, José García Moltalvo, Rosemarie Nagel, and Albert Satorra. One, two, (three), infinity, . . . : Newspaper and lab beauty-contest experiments. *American Economic Review*, 92(5):1687–1701, 2002.

Antoni Bosch-Domènech and Rosemarie Nagel. El juego de adivinar el número x: Una explicación y la proclamación del vencedor. *Expansión*, pages 42–43, 16th June 1997a.

Antoni Bosch-Domènech and Rosemarie Nagel. Guess the number: Comparing the FT's and Expansión results. *Financial Times*, page 14, 30th June 1997b.

Pablo Brañas-Garza, Francisca Jiménez-Jiménez, Virtudes Alba-Fernández, and Javier Rodero-Cosano. Measuring competition among students through experimental beauty contest games. Technical Report wp-0103, Dep. of Economics, University of Jaén, April 2001.

Colin F. Camerer. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton University Press, 2003.

Colin F. Camerer, Teck-Hua Ho, and Juin-Kuan Chong. A congnitive hierarchy theory of one-shot games and experimental analysis. Working paper, Caltech, 2003a. http://www.hss.caltech.edu/ camerer/QJErev8.pdf.

Colin F. Camerer, Teck-Hua Ho, and Juin-Kuan Chong. Models of thinking,

learning, and teaching in games. *American Economic Review*, 93(2):192–195, 2003b.

John Duffy and Rosemarie Nagel. On the robustness of behavior in experimental beauty contest games. *Economic Journal*, 107:1684–1700, 1997.

Brit Grosskopf and Rosemarie Nagel. Rational reasoning or adaptative behavior? evidence from two-persons beauty contest games. Harvard NOM Research Paper 01-09, Harvard Business School, 2001.

Shinichi Hirota and Shyam Sunder. Stock market as a "beauty contest" investor beliefs and price bubbles sans dividend anchors. Working Paper 02-42, Yale International Center for Finance (ICF), November 2002.

Teck-Hua Ho, Colin F. Camerer, and Keith Weigelt. Iterated dominance and iterated best response in experimental "p-beauty contests". *American Economic Review*, 88(4):947–969, September 1998.

John M. Keynes. *The General Theory of Interest, Employment and Money.* MacMillan, London, 1936.

Martin G. Kocher and Matthias Sutter. The 'decision maker' matters: Individual versus group behavior in experimental 'beauty-contest' games. WP, University of Innsbruck, 2001.

Hervé Moulin. *Game Theory for Social Sciences.* New York Press, N. Y., 1986.

Roger B. Myerson. Nash Equilibrium and the history of Economic Theory. Unpublished Paper, 1999.

Rosemarie Nagel. *Reasoning and Learning in Guessing Games and Ultimatum Games with Incomplete Information: An Experimental Study.* Unpublished dissertation, University of Bonn, Bonn, 1994.

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

Rosemarie Nagel. A survey on experimental "beauty-contest games": Bounded rationality and learning. In D. Budescu, I. Erev, and R. Zwick, editors, *Games and Human Behavior, Essays in Honor of Amnon Rapoport*, pages 105–142. Lawrence Erlbaum Associates, New Jersey, 1998.

Rosemarie Nagel. A keynesian beuty contest in the classroom. *Classroom Expernomics*, 8, 1999. http://www.marietta.edu/ delemeeg/expernom/f99.html#nagel.

Rosemarie Nagel, Antoni Bosch-Domènech, Albert Satorra, and José García-Montalvo. One, two, (three), infinity: Newspaper and lab beauty-contest experiments. Working Paper 438, UPF, Barcelona, November 1999.

John F. Nash. Non-cooperative games. *Annals of Mathematics*, 1951.

Amnon Rapoport and Wilfred Amaldoss. Mixed strategies and iterative elimination of strongly dominated strategies: an experimental investigation of states of knowledge. *Journal of Economic Behavior & Organization*, 42: 483–521, 2000.

Dale O. Stahl. Rule learning in a guessing game. *Games and Economic Behavior*, 16:303–330, 1996.

Hal R. Varian. What, exactly, was on John Nash's Beautiful Mind? *Economic Scene*, April 11, 2002.

Roberto A. Weber. Learning and transfer of learning with no feedback: An experimental test across games. Working paper, Carnegie Mellon, 2003a.

Roberto A. Weber. 'Learning' with no feedback in a competitive guessing game. *Games and Economic Behavior*, 44:134–144, 2003b.

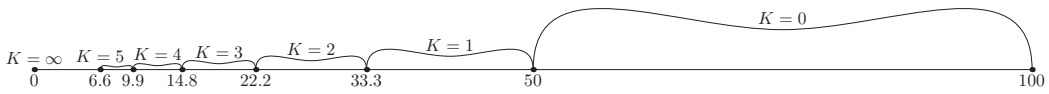


Figure 1: An example of different reasoning levels

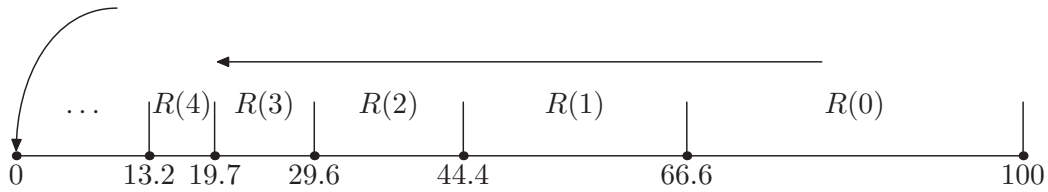
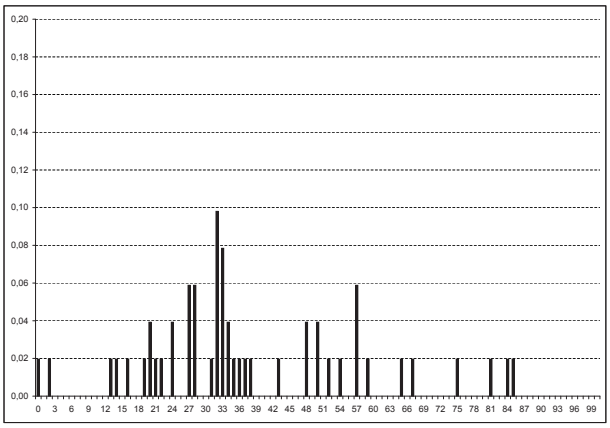
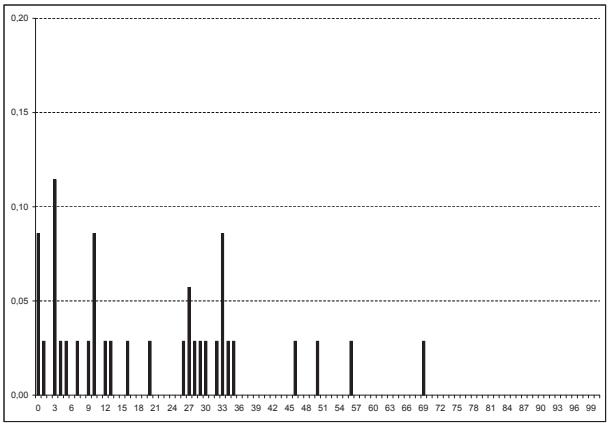


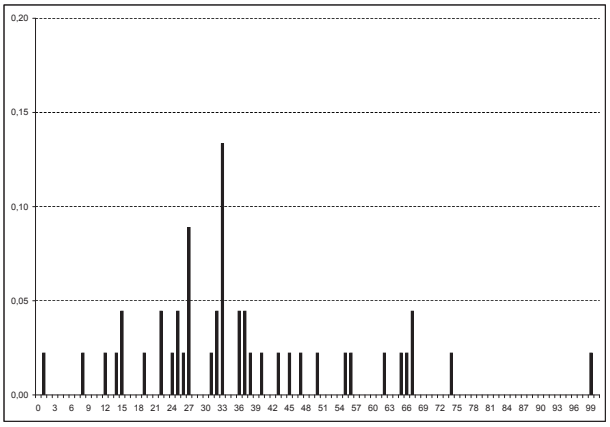
Figure 2: Iterated reasoning of individuals by eliminating dominated strategies



Group B-1



Group B-2



Group B+L

Figure 3: Relative frequencies in the first round

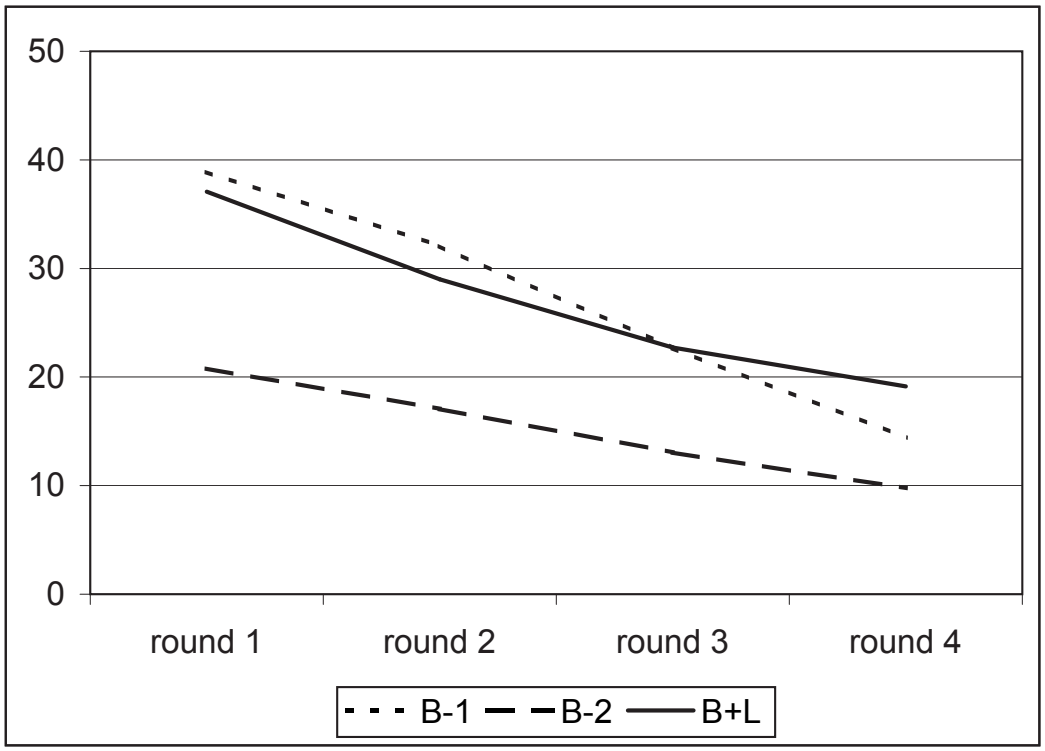


Figure 4: Evolution of session averages

Number of Subjects per group & session

	Session B-1	Session B-2	Session B+L
Group 1	10	6	10
Group 2	11	6	8
Group 3	9	6	9
Group 4	9	5	8
Group 5	10	6	10
Group 6	10	6	—
Session Pop.	59	35	45

Table 1: Distribution of group size

	Session B-1		Session B-2		Session B+L	
Round	Average	St. Dev.	Average	St. Dev.	Average	St. Dev.
1	38.885	19.348	20.782	17.611	37.067	18.891
2	32.087	14.235	17.060	12.362	29.008	11.554
3	22.660	16.750	13.042	7.318	22.712	11.577
4	14.419	13.360	9.748	7.503	19.142	13.944

Table 2: Average number and standard deviation per round and session

	B-1	B-2	B+L
$(\mu_2 - \mu_1)/\mu_1$	0.17	0.18	0.22
$(\mu_3 - \mu_2)/\mu_2$	0.29	0.24	0.22
$(\mu_4 - \mu_3)/\mu_3$	0.36	0.25	0.16

Table 3: Speed per round and session

	B-1		B-2		B+L	
Round	Z-statistic	p-value	Z-statistic	p-value	Z-statistic	p-value
1	1.116	0.165	0.937	0.344	1.135	0.152
2	1.063	0.208	0.713	0.690	0.681	0.742
3	1.579	0.014(*)	0.384	0.998	0.792	0.556
4	1.285	0.074	0.709	0.696	0.867	0.440

(*) significant at 5%

Table 4: Kolmogorov-Smirnof goodness-of-fit results

	B-1		B-2		B+L	
	t-stat.	p-value	t-stat.	p-value	t-stat.	p-value
Round 1 vs. 2	2.315	0.012	1.252	0.109	2.951	0.002
Round 2 vs. 3	-	-	2.107	0.021	3.148	0.001
Round 3 vs. 4	-	-	2.859	0.003	1.758	0.042
Round 2 vs. 4	7.586	0	-	-	-	-

Table 5: Parametric test for mean equality

	B-1		B-2		B+L	
	Z-stat.	p-value	Z-stat.	p-value	Z-stat.	p-value
Round 1 vs. 2	-2.378	0.008	-1.077	0.1414	-2.467	0.006
Round 2 vs. 3	-4.612	0	-2.001	0.022	-2.871	0.002
Round 3 vs. 4	-5.225	0	-3.459	0	-3.45	0

Table 6: Wilcoxon non-parametric test for mean equality