GENDER DIFFERENCES AND STEREOTYPES IN THE BEAUTY CONTEST

María Cubel, Santiago Sanchez-Pages

Document de

**GENDER DIFFERENCES AND
STEREOTYPES IN THE BEAUTY**

María Cubel, Santiago Sanchez-Pages

GENDER DIFFERENCES AND
STEREOTYPES IN THE BEAUTY [*]

María Cubel, Santiago Sanchez-Pages

ABSTRACT: Recent literature has emphasized that individuals display different depths of reasoning when playing games. In this paper, we explore gender differences in strategic sophistication and study whether these differences are endogenous. We report results from two different experiments employing the beauty contest. In the first, large study, we show that females react very strongly to incentives to the extent that gender differences disappear when a monetary prize is awarded. In the second study, we use a within subject design to analyze how depth of reasoning varies with gender priming and the gender composition of the set of players. We corroborate that females display higher levels of sophistication and even overtake males when incentives are provided and gender is primed. On the other hand, males who believe that females are better in the game display higher sophistication when playing against females.

María Cubel
Dept. of Public Economics & IEB
Universitat de Barcelona
Av Diagonal 690
08034 Barcelona (Spain)
E-mail: cubel@ub.edu

Santiago Sanchez-Pages
Dept. of Economic Theory & Edinburgh
School of Economics
Universitat de Barcelona
Av Diagonal 690
08034 Barcelona (Spain)
E-mail: sanchez.pages@gmail.com

1 Introduction

The experimental literature provides substantial evidence of the existence of strong individual heterogeneity in strategic sophistication or depth of reasoning. Observed individual behavior departs drastically from the predictions derived under the assumption of commonly-known unbounded cognitive capabilities (Nagel, 1995; Stahl and Wilson, 1995; Ho, Camerer and Weigelt, 1998; Costa-Gomes, Crawford and Broseta, 2001; Bosch-Domenech, Montalvo, Nagel and Satorra, 2002). This seems to reflect differences in the extent in which individuals engage in mentalizing processes or "theory of mind", that is, the activity of thinking about others' thoughts, emotions and intentions (Baron-Cohen, 1991). Models of k-level thinking have been proposed in order to account for these experimental results. These models acknowledge that individuals have different cognitive levels and have non-equilibrium beliefs about the sophistication of others.[1]

But strategic sophistication is a complex concept to pin down. It is surely endogenous. The depth of reasoning displayed by an individual might depend on her innate mentalizing or cognitive abilities, but also on her beliefs about the sophistication of others, and on the incentives provided. An individual may be sophisticated enough to provide the response corresponding to the standard game theoretical prediction but that response may fail to acknowledge that the rest of the population might be incapable of that. It would be questionable to label such response as more sophisticated than a response which departs from the game theoretical prediction but which takes correctly into account the heterogeneity in strategic sophistication in the population. On the other hand, a person may be reluctant to engage in further levels of reasoning, which require extra mental effort, unless the potential rewards of doing so are high enough. Responses in low stakes settings might not then reflect the mentalizing abilities of individuals, but rather their lack of motivation to engage in the process. In short, it should be natural to expect observed strategic sophistication to depend on both beliefs about the sophistication of others and on incentives.[2]

---

[1] Level-k models of thinking were introduced by Nagel (1995) and Stahl and Wilson (1994, 1995). Later Camerer, Ho and Chong (2004) proposed the cognitive hierarchy model. Both models are anchored on the existence of non-strategic individuals, labelled Level-0 or L0, but then differ on how individuals respond to the presence of less sophisticated ones. Level-k models have been applied to a number of strategic interactions such as communication and auctions. For a thorough survey, we refer the reader to Crawford, Costa-Gomes and Iriberri (2013).

[2] See Choi (2012) and Alaoui and Penta (2013) for recent attempts to develop theoretical models capable of accounting for the endogeneity of strategic sophistication.

In this paper, we explore the heterogeneity and endogeneity of strategic sophistication in the context of gender. Gender constitutes an obvious source of observable heterogeneity across individuals. Hence, gender can bring up relevant questions in the analysis of strategic sophistication. Our main goals are to study whether gender differences in strategic sophistication exist and, more importantly, to disentangle the potential source of such differences.

Regarding individual heterogeneity, our first question is whether there exist gender differences in depth of strategic reasoning.[3] To the best of our knowledge, no previous study has attempted to study gender differences in mentalizing ability, beliefs and sensitivity to incentives in games. This question has additional importance if stereotype threat is relevant in strategic interactions. The stereotype threat literature argues that observed gender differences in performance are due not to objective differences in ability but to stereotypes that affect negatively the cognitive ability of the threatened group when triggered (Steele, 1997).

A related question is whether males and females differ in their beliefs about the sophistication of others. In the context of gender, these beliefs might be based on stereotypes. Gender stereotypes affect daily behavior in a pervasive manner and hence they might have an impact on strategic behavior. Gender priming can be used to make beliefs and stereotypes about the relative sophistication of males and females salient. If gender differences in beliefs exist, alterations in the gender composition of the group of players should modify observed strategic sophistication.

Also related to the endogeneity of depth of reasoning, our last question is whether males and females respond differently to the presence of monetary incentives. Higher stakes might frame the interaction in a different light and hence induce changes in strategic behavior as well.

We explore these questions in the p-beauty contest/guessing game (Nagel, 1995). The guessing game is well suited for our purposes for a number of reasons. First, it is a competitive game. Players must anticipate the average response of others in order to win. Only the winner obtains the prize. Incentives are easy to adjust by changing its monetary value. In addition, beliefs about the sophistication of others are extremely important, as highlighted by the models of k-level thinking. Finally, the game involves a relatively complex calculation task: subjects must think what might be the average response, including their own, and then multiply the result by the announced

---

[3]This is definitely a thorny issue, more so since the "Summers affair" in 2005. We believe that our understanding of strategic sophistication is better served by tackling such question rather than by ignoring altogether.

factor one or more times. This calculation may trigger stereotypes related to the mathematical abilities of females and make them appear as less sophisticated (Spencer, Steele and Quinn, 1999; Quinn and Spencer, 2001).[4] Krendl, Richeson, Kelley and Heatherton (2008) use fMRI to show that brain areas involved in calculation are less active in females when this stereotype is activated. These brain areas –the angular gyrus and the inferior parietal lobule– are also relevant for subjects playing the beauty-contest (Coricelli and Nagel, 2009).

We present the results of two studies. Study 1 is a large between-subject classroom experiment. Students in six different cohorts played the 2/3 guessing game. In half of these cohorts, no monetary prize was given to the winner(s), whereas a monetary prize was awarded in the other three. We find substantial gender differences in the "No-prize" treatment. Female subjects display a less sophisticated behavior than males. However, no gender differences exist in the Prize treatment because there is a significant shift down in females responses across treatments. Males behaved similarly under the two treatments.

The second study is a within-subject experiment designed to ascertain the role of beliefs and gender stereotypes in strategic thinking. It allows us to track down individual behavior when economic incentives and gender composition change. The experimental sessions of this study had two phases. During the first phase, subjects played several rounds of the p-beauty contest. There were neither gender priming nor monetary incentives in this phase as in the "No prize" treatment of Study 1. In the second phase, there was a monetary prize at stake. We also manipulated the gender composition of the groups. Subjects competed in same gender and mixed gender groups (with balanced composition). We primed gender by simultaneously moving groups of same gender subjects from one room to another and by changing the color of the instructions booklet (from white to pink or blue). The purpose of this intervention was to make gender salient. This design also allows us to explore the existence of stereotype threat in the game.

Results from Study 2 suggest that strategic sophistication is quite sensitive to beliefs and incentives. Firstly, females react strongly to incentives and priming. We corroborate the result in Study 1 showing that females display lower strategic sophistication than males when incentives are absent. However, gender differences did not disappear as in Study 1 when

---

[4]The existence of gender differences in math performance is still a much debated issue. The evidence shows that although there are no gender difference in performance in arithmetic or algebra, there exist small-to-medium differences in high-school and top-end performers (Lindberg, Hyde, Petersen and Linn, 2010; Fryer and Levitt, 2010).

we awarded a monetary prize and primed gender. The direction of these differences actually reversed. Females displayed higher strategic sophistication than males. Moreover, we found that low sophisticated subjects were responsible for these gender differences. Secondly, when comparing rounds in which gender composition changed, we observe that only males modify their responses. Males with intermediate levels of sophistication lower their answers in mixed gender groups. We thus find no evidence of stereotype threat affecting females.

We explore the reasons behind the changes in behavior due to gender composition by analyzing the responses to a questionnaire administered at the end of the session. Females who thought that females respond higher numbers changed their behavior depending on the gender composition of the group whereas the rest of females did not. Responses to these questions also suggest that males thought more deeply about the game when playing in mixed gender groups because they believed females to be better in the game and to perform well in it.

The remainder of the paper is organized as follows. The next two sections describe the design and results in each of the two studies we report in this paper. Section 5 describes the result of the questionnaire we administered at the end of Study 2. In Section 6, we perform additional analysis under alternative measures of strategic sophistication. We conclude and discuss our results further in Section 7.

2   Related literature

To the best of our knowledge, no study has attempted to study the existence of gender differences in strategic sophistication. The reason behind this might be a genuine lack of differences, but also a conscious choice of researchers due either to ideological reasons or to the potential controversy of the topic. Very few experimental studies report evidence of gender differences as a by-product of their design either. Camerer et al. (2004) report in their Table 2 results for a beauty contest in same gender groups, but they only show summary statistics. As a by-product of a twin study, Burnham, Cesarini, Johannesson, Lichtenstein and Wallace (2009) find no gender differences in choices in the beauty contest. This is consistent with the results we obtain when gender is not primed and monetary incentives are given. Östling, Wang, Chou, and Camerer (2011) and Arad and Rubinstein (2012) report that females display slightly lower strategic sophistication in the Lowest Unique Positive Integer (LUPI) game and in the Colonel Blotto games

respectively. Let us reiterate that the main goal of these studies was not to investigate the existence of gender differences in strategic sophistication.

Several studies have explored the existence of other types of individual differences in the beauty contest. Although some differences emerge – portfolio managers and game theorists display higher strategic sophistication– the overall picture is that behavior in the beauty contest is very similar across subject pools (Bosch-Domenech, et al., 2002; Camerer et al, 2004). For instance, Kovalchik, Camerer, Gretherb, Plott and Allman (2005) find that older adults play similarly to young adults and Bühren and Bjorn (2010) find that chess grandmasters do not play differently than lay people.

A number of studies have found that depth of strategic reasoning responds strongly to the perceived sophistication of opponents. Palacios-Huerta and Volij (2009) find that when students play the centipede game against professional chess players they engage in more rounds of backward induction.[5] Similarly, Georganas, Healy and Weber (2010) find that some players increase their level of reasoning against stronger opponents but that previous play cannot predict which subjects make this adjustment. Alaoui and Penta (2013) find that subjects display less depth of reasoning when playing against opponents perceived to be less sophisticated than them. Agranov, Potamites, Schotter and Tergiman (2012) obtain a similar result in the beauty contest: undergraduate students display higher strategic sophistication when playing against graduate students than against computers.

We are aware of only one experimental study relating strategic sophistication to incentives. Alaoui and Penta (2013) find that subjects do engage in more rounds of reasoning when the prize from outguessing the opponent increases.[6] However, these authors do not explore the presence of gender differences in the response to higher stakes. Fryer, Levitt and List (2008) find that the performance of males in a GRE-style mathematical test increases relative to the performance of females when a payment per correct answer is introduced. This is the opposite result to the one we find in our two studies, but this might be due to the strategic nature of our experiment. In contrast, Frick (2011) employs data from professional distance running competitions and finds, in line with our results, that differences in the competitiveness between female and male races are significantly smaller in races where higher prizes or more prestige are at stake.

---

[5] This is not contemplated by models of k-level thinking since agents in these models do not factor the presence of individuals more sophisticated than them.

[6] Arad and Rubinstein (2012) run a treatment where they manipulate payoffs so that further levels of reasoning have no monetary cost. They find that nevertheless subjects very rarely perform more than three rounds of reasoning.

By using a competitive game, in which the player who best guesses the average response wins, our paper also relates to the literature in Economics which studies gender differences in competitive performance. Gneezy, Niederle and Rustichini (2003) and Gneezy and Rustichini (2004) have shown that females underperform in competitive environments. In the former paper, subjects had to solve mazes under different compensation schemes. Women performed worse under the tournament scheme. On the other hand, Gneezy and Rustichini (2004) find that Israeli boys but not girls run faster against another child than when they run alone. Dreber, von Essen and Ranehill (2011) fail to replicate this result for a sample of Swedish kids. This suggests that gender differences in competitiveness might be cultural. Günther, Ekinci, Schwieren, and Strobel (2010) find that competitive performance depends on the perceived bias of the task; females perform better than males when the task is perceived as female-biased. Along similar lines, Shurchkov (2012) find that females overtake men in competitions involving a verbal task and low-time pressure. Regarding the effect of gender priming, Iriberri and Rey-Biel (2013) show that omitting information about the gender of the opponent helps to mitigate the underperformance of women in competition. In contrast, we find that gender priming induces females to display higher strategic sophistication and to overtake males.

3   Study 1: Beauty in the classroom

The subjects of this study consisted of six cohorts of undergraduate students taking an Intermediate Microeconomics course at the University of Edinburgh between 2005 and 2010. As part of the course, students had to fill an online problem set containing several game-theoretic questions implemented via the website *Games and Behavior* developed by Ariel Rubinstein and Eli Zvuluny[7]. The classes were big, ranging between 116 and 170 students. Completing the problem set was compulsory and liable to a small mark penalty. Response rates were very high, a 91.83% on average. All students had received no instruction in game theory before answering the questions. The problem set was actually designed as a didactic method to introduce the basic concepts of game theory. In total, 792 students took part; 480 of them were male and 312 were female.[8]

---

[7]Available at http://gametheory.tau.ac.il/.

[8]When retrieving the data from the website, we were provided first with the list of participants' names but without their responses in order to ensure anonimity. We then assigned gender to these names and returned the list. We then received the data associ-
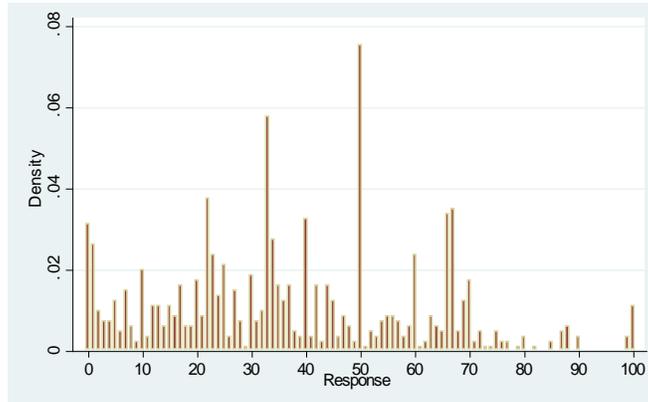
Figure 1:Histogram of responses in Study 1 (n=792).

The specific question we are interested in was a beauty contest. Students had to guess 2/3 of the average of all responses of students in the class.[9] Figure 1 contains the histogram of responses for the entire sample of 792 students. The graph shows the typical pattern of responses in beauty contests (Bosch-Domenech et al., 2002): Just above 3% or participants responded zero, the Nash equilibrium prediction. The spikes of frequency at 50, 33 and 22, according to the theory of k-level thinking, correspond to individuals with sophistication of level 0, 1 and 2 respectively.

However, this graph masks important heterogeneity. We ran two different treatments with three cohorts each: In the 2007, 2008 and 2010 cohorts (n=480), a prize of £10 (about 12 euros) was given to the student(s) who made the best guess. If there were more than one winner, the prize was divided among them. We call this the *Prize* treatment. The *No prize* treatment corresponds to the other three cohorts (n=312) in which no money was awarded to the winner. The instructor did not mention in class that the name of the winner(s) was to be announced publicly. So for the *No prize* treatment, such non-monetary reward was not made explicit.[10]

---

ating responses to the gender of the responder.

[9]The exact phrasing was: "Each of you (the students in this course) have to choose an integer between 0 and 100 in order to guess 2/3 of the average of the responses given by all students in the course. Each student who guesses 2/3 of the average of all responses rounded up to the nearest integer, will receive a prize to be announced by your teacher (or alternatively will have the satisfaction of being right!)."

[10]This does not rule out that students could seek prestige or status among their closer peers by winning.

Table 1 shows the aggregate results for the two treatments in Study 1 and compares them with the aggregate results of other experiments on the beauty contest. The studies in italics correspond to subject pools composed by non-students.[11] The first clear thing to observe is that the mean and median response for the *Prize* treatment are in line with those in previous experiments. We can then safely conclude that despite being implemented online, this treatment is comparable to lab experiments.

	Mean	Median	Std dev	Group size
Study 1 - Prize	**36.1**	**33**	**23**	**110-170**
Study 1 - No prize	**39.2**	**37**	**23.3**	**103-156**
Nagel (1995)	37.2	33	20	14-16
Ho et al. (1998)	38.9	NA	24.7	7
Camerer (2003)	32.5	NA	18.6	20-32
Kovalchik et al. (2005)	37	NA	17.5	33
Kocher and Sutter (2005)	34.9	32	NA	17
Buhren and Bjorn (2010)	32.1	29.6	22.2	6,112
Agranov et al. (2012)	36.4	33	21	8

Table 1: Aggregate results in Study 1 and in various beauty contest games.

Second, the *No prize* treatment shows the highest average and median responses. If lower responses indicate deeper reasoning, the lack of monetary prizes should induce subjects to think less about the game and hence they should respond higher numbers than in the *Prize* treatment. Under that assumption, the accumulated distribution of responses for the *No prize* treatment should first order stochastically dominate the accumulated distribution of responses for the *Prize* treatment.[12] Figure 2 corroborates this: the cumulative distribution of responses of the *No prize* treatment is below the one of the *Prize* treatment.

---

[11] Camerer (2003) uses CEOs; Kovalchik et al. (2005) 80 year olds; Buhren and Bjorn (2010) employs chess players, from amateurs to Grand masters.

[12] Given two cumulative distributions $F(x)$ and $G(x)$, F first order stochastically dominates G if $F(x) \leq G(x)$ for every x.
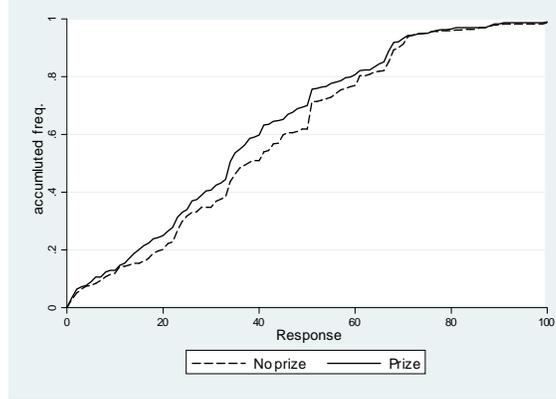
Figure 2: Cumulative distributions of responses by treatment.

Result 1.1 The distribution of responses under *Prize* differs from the distribution under *No-prize* (Mann-Whitney, $p = 0.041$; Median test, $p = 0.047$).

A closer look at the data shows that males and females respond very differently to the presence of a monetary prize. Table 2 shows the aggregate results for males and females in the two treatments. The striking result there is that average and median responses for females in the *No prize* treatment are much higher than any other. The other interesting observation is that males' responses do not seem to differ much across the two treatments.

	Mean	Median	Std dev	Obs.
Male & No-prize	37.6	35	23.9	243
Female & No-prize	41.9	42	23.3	148
Male & Prize	35.7	33	22.7	237
Female & Prize	36.4	34	23.5	164

Table 2: Aggregate results by gender and treatment.

Result 1.2 The distribution of responses of females differs between *Prize* and *No-prize* treatments (Mann-Whitney, $p = 0.026$; Median test, $p = 0.054$), and differs between males and females in the *No prize* treatment (Mann-Whitney, $p = 0.049$; Median test, $p = 0.029$).
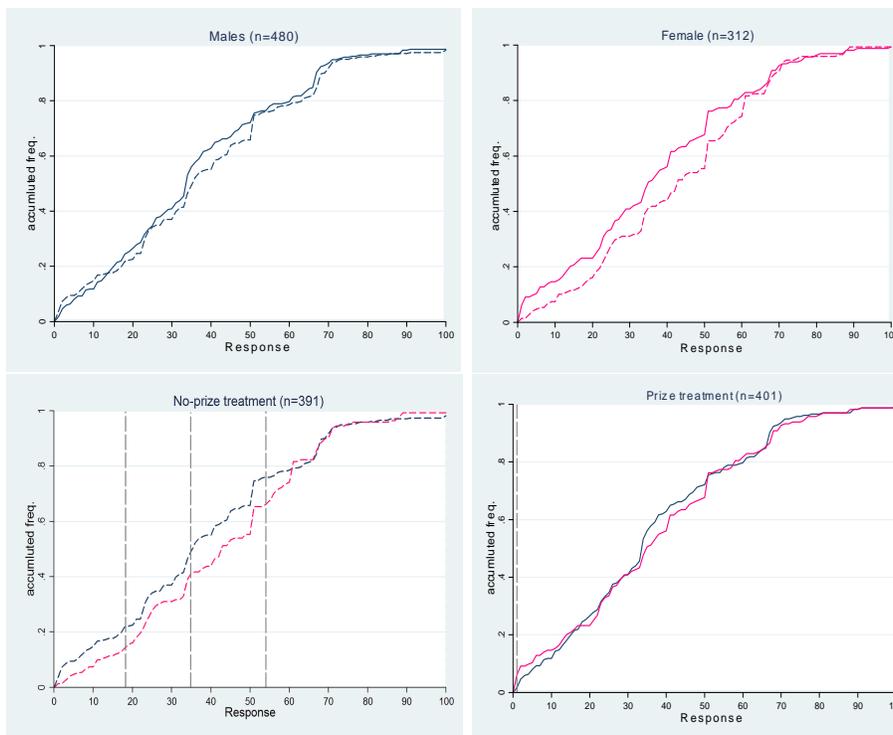
Figure 3: Cumulative distributions of responses by gender and treatment.

Figure 3 breaks down the cumulative distributions of responses by gender across treatments (upper panels) and by treatments across genders (lower panels). The cumulative distribution of responses in the *Prize* treatment is given by the solid line. The graphs show clearly that the distribution of female responses under *No prize* first order stochastically dominates the distribution of female responses in the *Prize* treatment, whereas dominance for males is unclear. In order to compare distributions of responses, we employ the test of stochastic dominance introduced by Davidson and Duclos (2000).[13] This tests allows us to associate stochastic dominance to a particular range of responses and, hence, to a certain degree of strategic sophistication. The Davidson and Duclos (2000) test yields that the distribution of female responses under the *No prize* treatment first order stochas-

---

[13] This test compares distributions at pre-determined points. A distribution is said to first stochastically dominate another if for all comparison points for which differences between the two distributions are statistically significant the sign of these difference is identical. In our analysis, we compare distributions at all possible responses.

tically dominates the distribution under the *Prize* treatment. In addition, the distribution of male responses is not statistically different across the two treatments. Although a higher proportion of males respond numbers between 30 and 50 compared to the *No prize* treatment, the difference is too small to be significant.

The lower left panel shows the cumulative distributions of male and female responses under the *No prize* treatment. The dominance test allows us to conclude that the distribution of female responses first order stochastically dominates the distribution of male responses in the *No prize* treatment. Furthermore, dominance comes from the ranges marked by the vertical dashed lines, the responses between 0 and 18 and between 35 and 54. According to the k-level theory, these are the responses roughly corresponding to very sophisticated and relatively unsophisticated subjects respectively. More females then seem to populate the medium sophistication range of responses (between 19 and 34) and the quite unsophisticated range (55 and above), although no gender differences exist in the proportion of truly irrational responses (68 and above). The lower right panel shows a very different picture: when a monetary prize is given, the only significant differences emerge for responses 0 and 1. This difference is due to a *higher proportion of females* responding these numbers. Hence, according to Davidson and Duclos (2000) criterion, dominance reverses, suggesting that in the presence of incentives females are at least as sophisticated as males.

We conclude from Study 1 that females drastically change their responses depending on the presence or not of financial incentives. Females show deeper levels of reasoning (as suggested by their lower responses) when a monetary prize was given. On the other hand, males do not respond to the monetary incentive and display a similar behavior in the two treatments. This would suggest that the conclusion of Camerer and Hogarth (1999) whereby monetary incentives have a small effect in experimental games might not necessarily apply to female populations. If financial incentives constitute a cue indicating that the beauty contest is a competitive game, it is to be expected that females react to this contextual information more strongly than men (Croson and Gneezy, 2009).

Nevertheless, the fact that gender differences disappear when a monetary prize is awarded suggest that males consider that a non-monetary prize is at stake in the *No prize* treatment. This non-monetary prize could be the utility of winning. This factor has been already studied in auctions (Cox et al., 1992). In contests, Sheremeta (2010) finds that about a third of subjects are willing to spend a positive amount of money in order to win a zero value prize. Furthermore, Sheremeta (2010) is able to elicit the valuation

of winning in his experiment in monetary terms. One might be tempted to conclude that in our experiment, since males in the *No prize* treatment behave similarly to females in the *Prize* treatment (Mann-Whitney, $p = 0.648$), males' value of winning would be roughly equivalent to £10. Note however that we find no evidence of males increasing their depth of reasoning (lowering their responses) when a monetary prize is given. This might suggest that for males 1) psychological and monetary rewards crowd out each other, so a utility function incorporating both types of gains should not be additive; or 2) that rewards of any type have very strong decreasing returns. We choose not to elaborate more on this since the correct specification of such utility function is not the focus of our design.

Note that Study 1 does not allow us to check whether beliefs about the sophistication of others differ between males and females. Nor whether they respond differently to gender priming. To that end, we present next the result of a laboratory experiment using a within-subjects design.

4   Study 2: Beauty in the lab

The second study was conducted with undergraduate students at the University of Barcelona in May of 2012 and March of 2013. Subjects were recruited in the standard way. We conducted two sessions of 48 participants each. The gender composition of each group of participants was balanced, so that half were female and the other half were male. The experiment was implemented with pen and paper, no feedback was provided at any time. Experimenters answered privately any questions that subjects had. Sessions lasted between 40 and 50 minutes.

Each session had two phases. In the first phase, there were neither monetary incentives nor gender priming.[14] We assigned students randomly in two groups of equal size and with an even gender composition. Therefore, there were two groups of 24 subjects (12 male and 12 female) with two instructors (one female and one male) in each room. Each group occupied a separate room. The two rooms were located in two different corridors. During this phase, subjects were asked to guess a fraction p of the average response in their room. They played nine rounds of this guessing game with different values of p in each round, i.e. $p = \left(1, \frac{2}{3}, \frac{11}{10}, \frac{1}{3}, \frac{3}{2}, \frac{1}{5}, \frac{6}{5}, \frac{1}{2}, \frac{4}{3}\right)$. Instructions were provided through white paper booklets where they also had to record their answers. Subjects did not write neither their name or their gender in these booklets. Each participants was assigned a number.

---

[14]For showing up, subjects received a voucher worth a refreshment in the campus café.

The second phase of the session was composed by two rounds where subjects had to guess the $\frac{2}{3}$ of the average response in the room. In this phase, monetary incentives and gender priming were introduced. The difference between the two rounds was the gender composition of the group. In the first round (the SG round), students were regrouped so that there were two gender homogenous groups: one formed only by female subjects and the other formed by male subjects only. We did this regrouping by moving the male students in one room and the female students in the other through different corridors so they could not see each other. They were guided in this process by an instructor of their same gender who made all efforts to avoid information transmission among participants. For the second and last round (the MG round), subjects were mixed again by gender following the same process. Subjects who moved were those who had stayed in the previous round. During this second phase, gender priming was further implemented by distributing pink paper booklets to female subjects and blue paper booklets to male subjects. In addition, the gender of the pair of instructors present in each room matched the gender composition of the subjects in it. This means that when groups were gender balanced there was one female and one male instructor in each room. And when groups were of the same gender, the gender of the two instructors in the room coincided with the gender of the group. The payoffs of the session were determined by selecting randomly one the two rounds of the second phase. We provided two prizes of 40 euros (around 55\$) each, one prize for the winner(s) in each room (if the MG round was selected) or within each gender group (if the SG round was selected). At the end of the session, students filled up a short questionnaire aimed at measuring beliefs and potential stereotypes.

4.1 Results within rounds

4.1.1 First phase: no gender priming, no incentives

Table 3 depicts the aggregate results for the round with no priming and no incentives with $p = \frac{2}{3}$ (NPNI round henceforth). It shows that the distribution of male responses has a lower mean and median than the one of females.

	Mean	Median	Std dev
Males	28.5	25	15.7
Females	34.1	30	21.3

Table 3: Aggregate results by gender in the NPNI round.

By comparing the cumulative distributions of responses for male and females we can state the first result of Study 2, which is in line with the results in the *No-prize* treatment of Study 1.

Result 2.1 In the NPNI round, the distribution of female responses first order stochastically dominates the distribution of male responses.

Figure 4 illustrates this result. Females show a lower level of sophistication than males in the absence of incentives and gender priming. The Davidson and Duclos (2000) test allows us to establish that this gender difference emerges in the interval of responses from 42 to 56 which corresponds to a low level of strategic sophistication. This dominance result holds in all the other rounds of the first phase.[15]
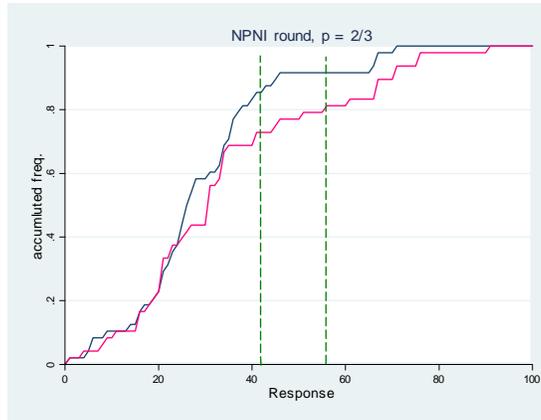

Figure 4: Cumulative distributions of responses by gender in NPNI round.

Next, we assess the sophistication of subjects by considering their responses in the rounds of the first phase (no incentives and no gender priming) with $p \neq 1$. For each p, we assign level $k = \{1, 2, 3, \infty\}$ of sophistication to each individual response x_i when k minimizes $d = \left(x_i - 50p^k\right)^2$. We follow Coricelli and Nagel (2009) to classify responses as low level if $k = 1$ (high otherwise). A subject is considered of low (high) sophistication if at least 5 out of 8 responses are of low (high) level. The rest of subjects are considered random and discarded from the analysis. This classification does not only

---

[15]The distribution of females responses first order stochastically dominates the one of males for all rounds with $p < 1$ and the reverse holds in all rounds with $p > 1$.

reflect how close an individual plays with respect to the equilibrium prediction. It also indirectly incorporates beliefs about the sophistication of the opponent: Coricelli and Nagel (2009) show through fMRI that, compared to a treatment in which subjects play against computers, subjects classified as highly sophisticated display activation in areas of the brain associated with theory of mind. On the contrary, low sophistication subjects do not display additional activity. Nevertheless, we consider an alternative classification in Section 5. All our results still go through.

As Table 4 shows, 77 % of the classified individuals in the sample are classified as low sophisticated; from these, 67% are female. The percentages of High and Low subjects out of the whole pool (70.8% and 20.8% respectively) are very similar to the ones obtained in previous studies.[16] However, these figures mask important gender differences. The small fraction of females who exhibit high strategic sophistication stands out. Statistical tests confirm that the distribution of levels of sophistication differs across genders ($\chi^2 = 8.39$, $p = 0.003$; Fisher's exact $p = 0.005$).

	Low	High	Total
Males	26	15	41
Females	42	5	47
Total	68	20	88

Table 4: Sophistication by gender (first phase).

A word of caution is in order here: This classification employs choices under no incentives. Hence, it is likely that the number of subjects that we classify as low sophisticated is an upper bound. From Study 1 we know that females display higher strategic sophistication when a monetary prize is awarded. That could account for the low number of highly sophisticated females. However, as we will see below, this classification has considerable predictive power in the remaining rounds of the study.

4.1.2   Second phase: incentives, gender priming

Now we move to the second phase of the experimental session where we introduce gender priming and incentives. As mentioned above, the second phase had in two rounds. In the SG round, participants were separated in two rooms in such a way that they were competing against individuals

---

[16] Coricelli and Nagel (2009) obtain 50% and 35% respectively (n=20). Brañas-Garza, García-Muñoz and Hernan (2012) obtain percentages of 78% and 13% (n=191).

of their same gender. In the MG round, we regrouped them again so the gender composition in each room was balanced. In both rounds, pecuniary incentives came into play. One of the rounds was randomly chosen at the end of the session to determine payoffs.

The MG round In this round, half of the participants were male and half were female in each room. Table 5 shows that the distribution of responses of females has a lower mean and median than the one of males.

	Mean	Median	Std dev
Males	27.1	25	18.1
Females	20.7	15	18.2

Table 5: Aggregate results by gender in the MG round.

Females appear to display higher strategic sophistication than males in this round. The comparison of the distributions of responses across genders in the MG round shows indeed that they are statistically different (Mann-Whitney, $p = 0.036$; Median test, $p = 0.014$). Furthermore, the dominance test provides a clear ordering between them.

Result 2.2 The distribution of male responses first order stochastically dominates the distribution of female responses in the MG round.

Figure 5 illustrates this result. The dominance test establishes that there is a higher number of females than males who chose responses in the interval between 12 and 29 (marked by the vertical dashed lines). This suggests that more females display relatively high levels of sophistication.
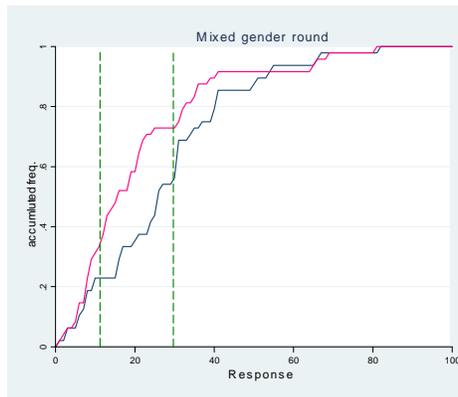

Figure 5: Cumulative distributions of responses by gender, MG round.

A closer look to the data reveals that the participants who are responsible for this result are those classified as low sophisticated in the first phase. The distributions of responses differ across genders for low sophisticated subjects only (Mann-Whitney, $p = 0.005$; Median test, $p = 0.002$). The dominance result that we found for the whole sample is stronger and applies to a larger interval of responses for low sophisticated individuals (left panel of Figure 6 below). Dominance is not statistically significant for highly sophisticated individuals (right panel of Figure 6 below).
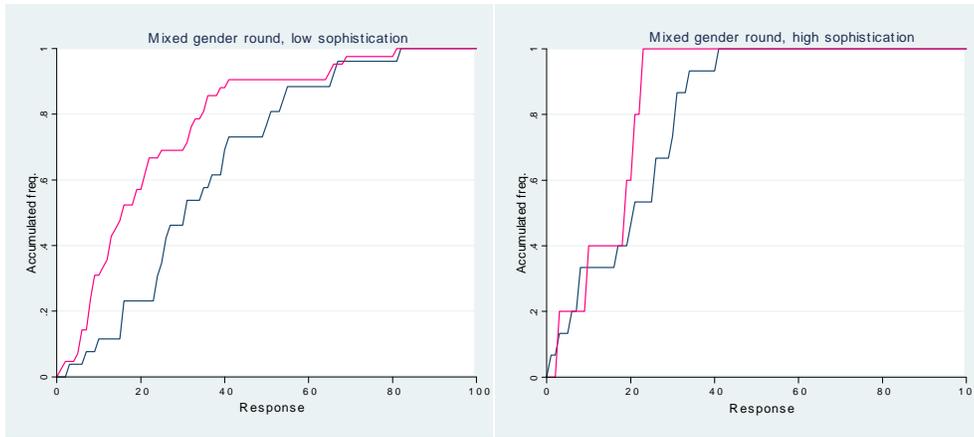

Figure 6: Cumulative distributions of responses by gender and sophistication in the MG round.

Result 2.3 The distribution of male responses first order stochastically dominates the distribution of female responses only for low sophisticated individuals in the MG round.

The SG round In the SG round, participants played against their same gender opponents. Table 6 shows the mean and median responses for males and females in this round. The first noticeable result again is that the average male response is higher than the average female response.

	Mean	Median	Std dev
Males	31.1	29	21.2
Females	20.1	17	13.5

Table 6: Aggregate results by gender in the SG round.

The distributions of responses across genders are statistically different (Mann-Whitney, $p = 0.012$; Median test, $p = 0.014$) and the dominance result is even stronger than in the MG round as Figure 7 illustrates. The interval of significant dominance ranges from 13 to 67.

Result 2.4 The distribution of male responses first order stochastically dominates the distribution of female responses in the SG round.
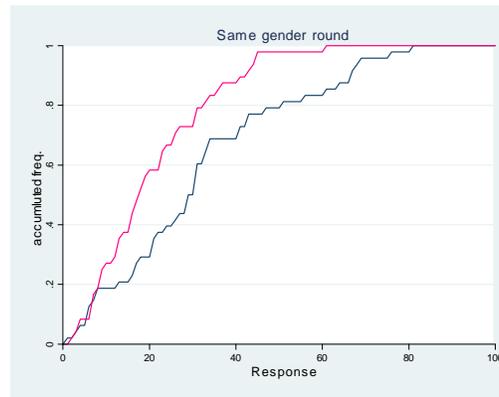

Figure 7: Cumulative distributions of responses by gender in the SG round.

Again, if we split participants by their level of sophistication we observe that gender differences in responses only remain significant for low sophisticated individuals (Mann-Whitney, $p = 0.002$; Median test, $p = 0.002$). There are no gender differences among high sophisticated individuals. The dominance of the distribution of male responses becomes stronger when we only look at low sophisticated participants, as Figure 8 below illustrates.

Result 2.5 The distribution of male responses first order stochastically dominates the distribution of female responses only for low sophisticated individuals in the SG round.
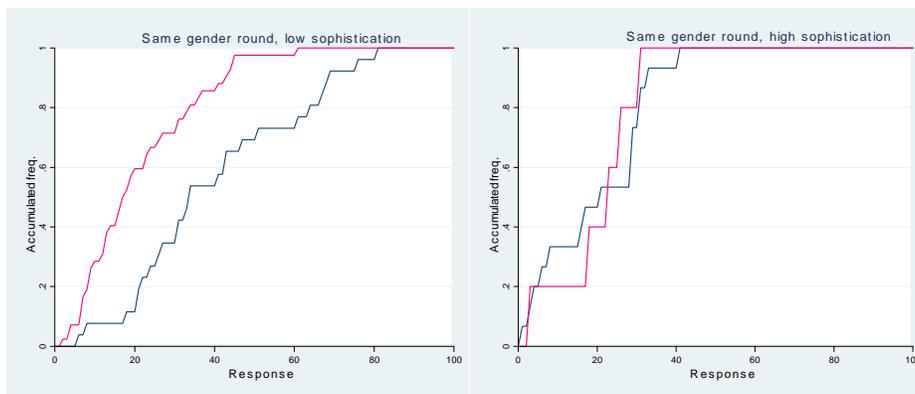
Figure 8: Cumulative distributions of responses by gender and sophistication in the SG round.

To summarize, the main findings of the analysis per round are:

1. Results for the NPNI round confirm that males show higher levels of sophistication than females in the absence of incentives and gender priming. This is in line with Study 1.

2. This result reverses when gender priming and incentives are introduced: Females show higher levels of sophistication than males.

3. The gender differences in the SG and MG rounds are due to low sophisticated individuals. There are no gender differences among high sophisticated individuals.

4.2   Results across rounds

Let us now exploit our within-subject design and compare individual responses first across the NPNI and the MG rounds, and second, across the MG and the SG rounds. The latter comparison involves a change in the gender composition of groups. This was designed to trigger beliefs and stereotypes about the relative strategic sophistication of males and females. This intervention also allows us to explore the presence of stereotype threat. The stereotype threat literature suggests that members of the threatened group suffer negative feelings and thoughts which in turn worsen their performance. We then expect the responses of members of the threatened group (if any) to shift up when playing in mixed gender groups compared to when they play in same gender groups. This might be the case for female if they perceive that the calculation involved in the guessing game favors males.

4.2.1 Comparison of the NPNI and the MG rounds

Table 7 below compares responses across the NPNI and the MG rounds and across genders. We observe that females change their behavior considerably. Their mean and median responses are much lower in the mixed gender round. However, men barely change their answers across rounds.

	Mean	Median	Std dev
Male, NPNI	28.5	25	15.7
Female, NPNI	34.1	30	21.3
Male, MG	27.1	25	18.1
Female, MG	20.7	15	18.2

Table 7: Aggregate results by gender across NPNI and MG rounds.

Result 2.6 The distributions of responses in the NPNI and MG rounds differ for females (Wilcoxon sign-rank, $p < 0.001$; Sign-test $p < 0.001$).

The Q-Q plot in Figure 9 illustrates this result. Male quantiles are distributed around the 45 degree line while female quantiles are above this line. This indicates that the distribution of female responses shifts down when there is gender priming and incentives, whereas the distribution for males does not shift in any clear direction between rounds.
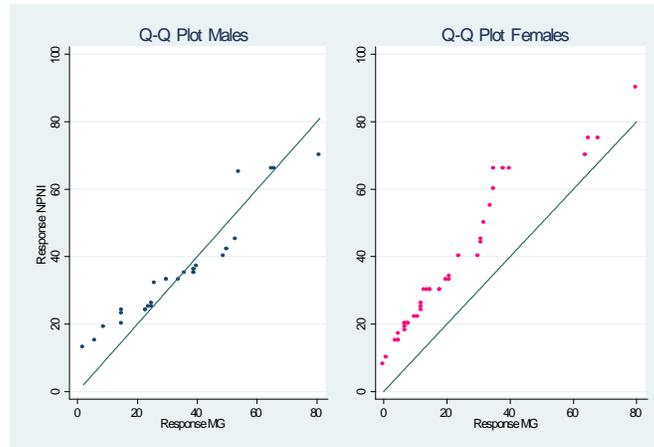

Figure 9: Q-Q plots by gender comparing NPNI and MG rounds.

This corroborates our results in Study 1. Females change their behavior when incentives are introduced whereas males behave similarly with and without incentives. Note that the difference in play between these two rounds cannot be accounted for merely by the decrease in responses that takes place in repeated guessing games with no feedback (Weber, 2003). In contrast with Study 1 gender differences do not disappear when incentives are introduced. This suggests that the additional strategic sophistication that females display in the MG round is due to gender priming. Alternatively, it might be that the bigger size of the prize compared to the one in Study 1 (40 vs 12 euros) is inducing females to think more deeply about the game. However, that would imply that males are extremely insensitive to the size of monetary incentives.

Further analyses show that low sophisticated females are the ones lowering their responses (Wilcoxon sign-rank, $p < 0.001$; Sign-test $p < 0.001$) whereas highly sophisticated females do not change their responses significantly. However, due to the small fraction of highly sophisticated females this result is inconclusive. We come back to this in section 6.2 where we use an alternative measure of strategic sophistication.

4.2.2 Comparison of the SG and the MG rounds

Let us now compare the SG and the MG rounds. Recall that in the second phase of the session we manipulated the gender composition of the group of participants. The purpose of this manipulation was to explore the role of beliefs about the sophistication of others. We expect gender priming to make individual beliefs about the relative sophistication of males and females salient. If this is the case, responses should change with the gender composition of the group. In the following analysis, we assume that if an individual believes that a change in the gender composition shifts up the distribuion of levels of sophistication in the group, his/her choice would decrease[17]. On the other hand, if stereotype threat is relevant in the beauty contest, the change in the gender composition of the group of opponents should arouse anxiety and negative thoughts in the threatened group. If these negative feelings impair the mentalizing process, the threatened group should display lower levels of sophistication when playing in mixed gender groups. In that case, we expect the distribution of responses of the threatened group to shift down when playing against the other group.

---

[17] Agranov et al (2012) make a similar assumption when they make undergraduate students play against graduate students.

Table 8 shows that the average and median responses of females do not significantly differ between the SG and the MG rounds. On the other hand, male responses appear to be higher when males play against males compared to when they play in mixed gender groups.

	Mean	Median	Std dev
Male, SG	31.1	29	21.2
Female, SG	20.1	17	13.5
Male, MG	27.1	25	18.1
Female, MG	20.7	15	18.2

Table 8: Aggregate results by gender across SG and MG.

To have a more clear picture, we provide in Figure 10 the Q-Q plot of male responses to both rounds, SG and MG. It is evident that distribution of male responses shifts down when they play in mixed gender groups compared to when they play only against males.
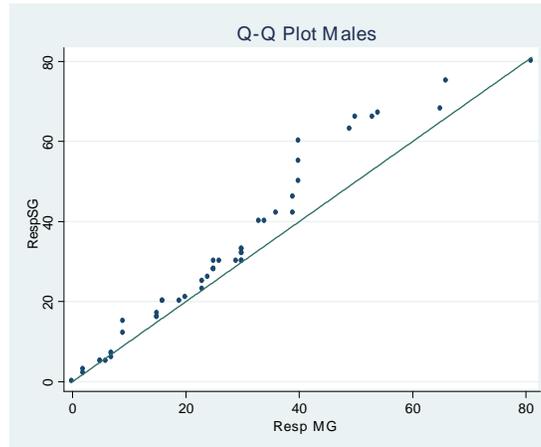
Figure 10: Q-Q plot for males comparing MG and SG rounds.

Result 2.7 The distributions of responses in the SG and the MG rounds differ for males (Wilcoxon sign-rank, $p = 0.083$; Sign-test $p = 0.099$).

Changes in gender composition seem to affect only males. When playing in mixed gender groups, they show more strategic sophistication than when they play only against males. Females do not alter their behavior so there is no evidence of stereotype threat among them. If anything, stereotypes

might be affecting males. However, the actual mechanism is unclear at this point. Males might be enjoying a "stereotype lift" (Walton and Cohen, 2003), because they think that females are less sophisticated. That enhanced sense of self-efficacy would lead them to think more about the game when playing in mixed gender groups. Alternatively, males could be experiencing "stereotype arousal" (O'Brien and Crandall, 2003), and react to stereotype threat by making additional mentalizing effort. Finally, males might be lowering their responses simply because they believe females respond lower numbers. Next, we explore this and related issues by analyzing the responses to the questionnaire that we administered at the end of the session.

Before that, let us summarize the results of the analysis across rounds:

1. Results confirm that females react strongly to the presence of incentives whereas males do not.

2. Males react to the gender composition of the group. Males display higher strategic sophistication when playing in mixed gender groups compared to when they play only against males.

5   Beliefs and stereotypes

In this section, we explore whether (ex-post) stated beliefs correlate with behavior. This is relevant for two reasons: On the one hand, strategic sophistication is endogenous to the perceived sophistication of others (Georganas et al, 2010; Agranov et al, 2012; Alaoui and Penta, 2013). On the other hand, if subjects have gender stereotypes, these could affect their choices.

Next, we analyze responses to the questions in the personal questionnaire that subjects had to fill at the end of the session and before winners were announced. The objective of these questions was to capture beliefs or gender stereotypes which could affect answers in the game. Although these ex-post beliefs were not elicited via a scoring rule, they still may offer a complementary view on observed behavior.

We focus mostly on the responses to two questions. First, "Which gender responds higher numbers? (when $p = \frac{2}{3}$)" (Q1), and second, "Which gender is better at this game?" (Q2).[18] These questions try to capture different factors which might be important to understand the behavior of participants.

---

[18] The rest of questions were "Do males/females compete better in single/mixed gender groups?" and " Did you change your response when there was a monetary prize/the composition of the group changed?".

Q1 is designed to obtain information on beliefs about the behavior of others. Q2 is designed to detect stereotype threat and beliefs about the relative strategic sophistication of males and females. The responses to these two questions were significantly correlated (Spearman $\rho = -0.369$, $p < 0.001$). This implies that participants understand the basics of the game and associate a better performance with lower responses.
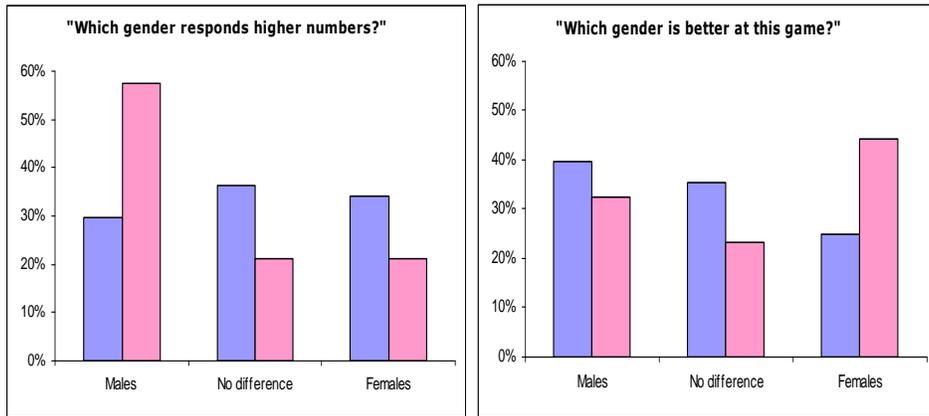

Figure 11: Responses by gender to Q1 and Q2 (females/pink, males/blue).

The left panel of Figure 11 shows the histogram of responses to Q1. The majority of females in the sample think that males respond higher numbers, while males' responses to this question are more balanced. This difference in responses is statistically significant ($\chi^2 = 7.32$ $p = 0.026$; Fisher's exact $p = 0.028$). The right panel of Figure 11 shows that females have more polarized views than males on which gender is better at the game. Both females and males, show a higher proportion of answers in favour of their own gender. But this effect is stronger for females. However, the are no statistically significant gender differences in the responses to this question.

Next, let us explore whether there exist differences in the responses to these questions by level of sophistication. The left panel of Figure 12 shows that there are clear differences between low and high sophisticated subjects in the response to Q1. These differences are statistically significant ($\chi^2 = 12.66$ $p = 0.002$; Fisher's exact $p = 0.002$). Low sophisticated subjects believe that men tend to respond higher numbers than females. Highly sophisticated subjects have the opposite beliefs. Responses to Q2 show a less clear-cut pattern. The majority of high sophisticated subjects think that males are better at the game. Low sophisticated individuals do not hold the

opposite belief. They have more polarized views. Still, the distributions of responses are statistically different between the two sophistication levels ($\chi^2 = 9.95$ $p = 0.041$; Fisher's exact $p = 0.042$).

Recall that the high sophisticated group is mainly populated by male subjects. Hence, their responses to Q1 and Q2 may not be that surprising. The gender composition of the low sophisticated group is more diverse. However, males and females in this group seem to share the belief that males chose higher numbers.
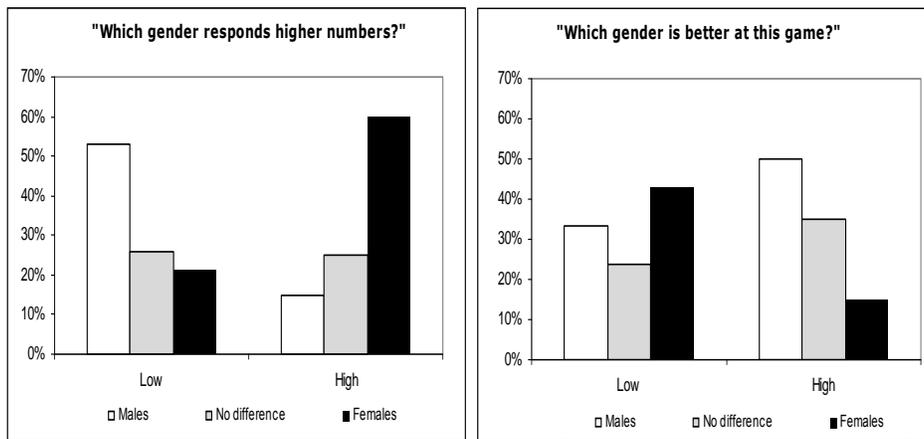
Figure 12: Responses by sophistication level to Q1 and Q2.

Given these differences in responses across genders and sophistication levels, the next question is whether they can explain differences in behavior. The left panel of Figure 13 compares the distribution of responses in the SG and MG rounds for subjects who state that individuals of the same gender respond higher numbers in the beauty contest. It shows that these subjects lowered their responses in the MG round compared to the SG round. The right panel performs the same comparison for the rest of subjects. They tend to increase their answers in the MG round.

Figure 13: Q-Q plot across SG and MG rounds by responses to Q1.
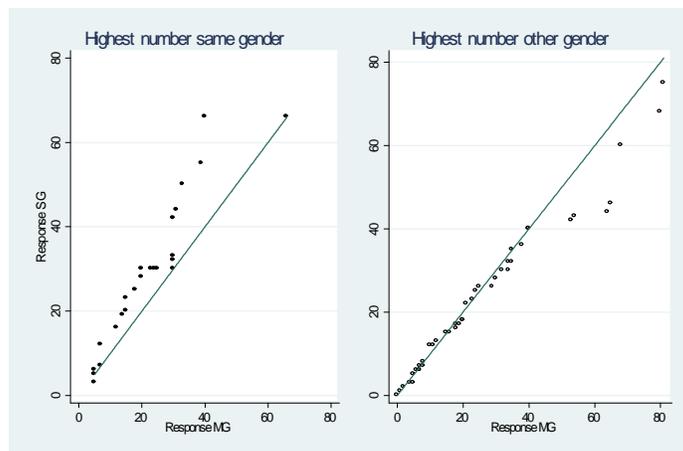
Result 3.1 The median responses of subjects who think their same (the other) gender respond higher numbers is higher (lower) in the SG round than in the MG round (Sign-test $p = 0.002$ and $p = 0.044$ respectively).

This result is relevant since it shows that ex-post stated beliefs are consistent with behavior.[19] The difference between rounds is not surprising for males since we already found that males change their behavior according to the gender composition of the group (Result 2.7). Females, on the contrary, did not show significant differences in behavior between the MG and the SG round. Hence, Result 3.1 does not fit well with that finding.

An explanation to this can be obtained by analyzing the behavior of females who believe that females respond higher numbers. They do change their behavior across the SG and the MG rounds (Wilcoxon sign-rank, $p = 0.031$; Sign-test $p = 0.070$). Thus, Q1 allows us to qualify our previous result stating that females do not respond to the gender composition of the group. This begs the question of whether these females are different from the rest in any dimension. A careful look to the data reveals no differences in any of the other characteristics we have information about. Thus, we can only conclude that females who react to the gender composition of the group act purely because of their beliefs about the responses of other females.

---

[19] Costa-Gomes and Weizsäcker (2008) find that behavior is not consistent with stated beliefs in half of the normal-form games they study. Rey-Biel (2009) obtains a somewhat higher proportion of consistent choices in constant-sum games. Because we do not elicit beliefs in an incentive compatible manner, we cannot contribute to this debate.

Regarding males, recall that we closed Section 4 with an open question. We had observed that the distribution of responses of males in the MG round shifted down compared to the SG round. We stated three candidate explanations there: 1) males experienced stereotype lift; 2) males exerted extra effort in order to compensate their stereotype threat; and 3) they responded lower numbers just because they expected females to do so. The analysis of Q1 and Q2 can help us to shed light on this. Under explanation 1) males who changed behavior should be those who think that males are better at the game. The opposite would occur under explanation 2). Finally, we should be able to check explanation 3) in the responses to Q1.

As expected from Result 3.1, males who think that men respond higher numbers than females change their behavior between the SG and the MG rounds (Wilcoxon sign-rank, $p = 0.022$; Sign-test $p = 0.032$). The same result applies for males who think that females are better at the game. (Wilcoxon sign-rank, $p = 0.053$; Sign-test $p = 0.054$). This would seem to discard explanation 1) because is not the case that males who display higher sophistication in the MG round believe females play worse in the game. These results thus give support to explanations 2) and 3). However, the significant correlation between the answers to Q1 and Q2 and the fact that the p-values are lower when using Q1 seem to suggest that males' change in behavior is due mostly to beliefs about female behavior.

The responses to the question "Do females compete better in single/mixed gender groups?" (Q3) provide a different view. Responses to this question are not significantly correlated to responses in Q1 and Q2. But as it turns out, males who change their behavior between the SG and the MG rounds are the ones who believe that females compete better in mixed gender groups (Wilcoxon sign-rank, $p = 0.030$; Sign-test $p = 0.032$). This together with their answer to Q2 suggest that males who display more sophistication in the MG round might be thinking more deeply about the game because they experience stereotype arousal.

6   Robustness checks

A key instrument in the analysis so far has been the definition of sophistication. The classification we followed, based on Coricelli and Nagel (2009), presents two main limitations. On the one hand, this classification is coarse. On the other hand, it does not take explicitly into account beliefs about the sophistication of others. In order to address both issues, we modify our definition of sophistication in two ways.

Firstly, we expand the classification of strategic sophistication by considering three levels (low, medium and high) instead of two. Secondly, we measure strategic sophistication as the quadratic distance to the winning response. We refer to this measure as accuracy. It has the advantage of accounting for depth of reasoning and correctness of beliefs on others' responses simultaneously.

6.1 Robustness check (I): A finer classification

Under the classification of Coricelli and Nagel (2009), many of the subjects in our sample are classified as low sophisticated. Recall that low sophisticated individuals are defined as those with a majority of L0 and L1 responses. In order to refine this classification, we expand the set of types and consider L0 as low sophisticated and L1 responses as medium sophisticated. We then classify subjects according to the mode sophistication of their responses in the first phase. In cases were there is more than one mode, the subject remains unclassified and is excluded from the analysis. Hence, subjects who were classified before as highly sophisticated remain in this category.[20]

	Low	**Medium**	**High**	Total
Males	8	14	17	39
Females	18	15	6	39
Total	26	29	23	78

Table 9: Sophistication by gender under alternative classification.

Table 9 shows sophistication levels by gender under this classification. Sophistication levels are not equally distributed across genders ($\chi^2 = 9.14$ $p = 0.01$; Fisher's exact $p = 0.012$). Still the number of low sophisticated females is larger than of males. The opposite holds for high sophisticated individuals. Recall that this classification employs the answers to the rounds where there were no priming and no gender incentives.

All our previous results by level of sophistication hold under this new classification. Moreover, this finer classification allows us to further explore the behavior of medium and low sophisticated subjects. This yields an additional finding.

Result 4.1 The distributions of responses in the SG and MG rounds differ for medium sophisticated males (Wilcoxon sign-rank, $p = 0.020$; Sign-

---

[20] One more female is classified as high sophisticated than under the previous classification. This is because she was classified as random before.

test $p = 0.022$) and highly sophisticated females (Wilcoxon sign-rank, $p = 0.052$; Sign-test $p = 0.062$).

Result 4.1 is illustrated in Figure 14. The distribution of responses of medium sophisticated males and high sophisticated females shifts down when playing in mixed gender groups. We see this result to be in line with the findings in Alaoui and Penta (2013). They find that subjects who react to changes in the perceived sophistication of the opponent are those capable of at least one level of depth of reasoning. Similarly, we find that males who react to the gender composition of the group of opponents are sophisticated enough to entertain beliefs about the sophistication of others. High sophisticated males, on the other hand, might not be reacting to gender composition because they follow equilibrium play more closely. This is substantiated by the fact that the great majority of males classified as high sophisticated believe that females respond higher numbers (see left panel of Figure 12). Despite that, they do not alter their behavior significantly when the gender composition of the group changes.
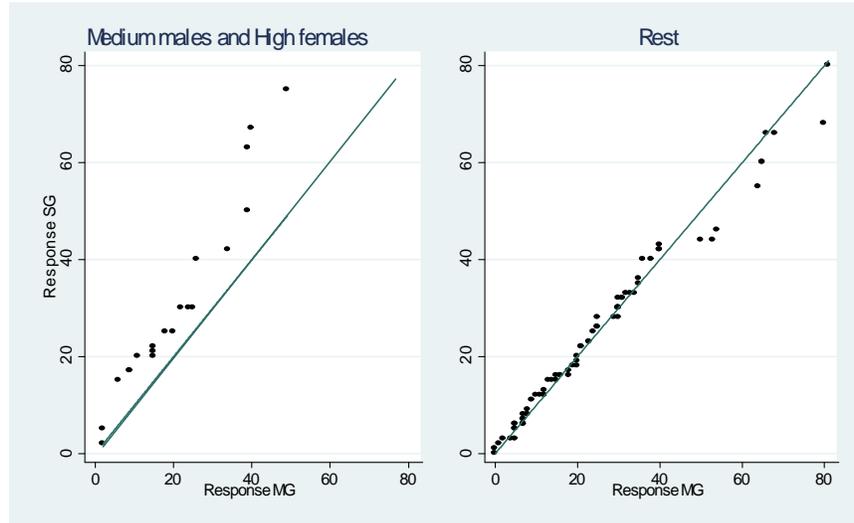

Figure 14: Q-Q plot comparing MG and SG rounds ($p = \frac{2}{3}$).

6.2   Robustness check (II): Accuracy

As a second robustness check, we use an alternative measure of sophistication that we refer to as accuracy. We measure the quadratic distance to the

winning response in each round and take the average. This measure of sophistication is similar to strategic IQ in Coricelli and Nagel (2009). This measure accounts for depth of reasoning and for the correctness of beliefs on others' responses directly. We also compute the quadratic distance (inverse of accuracy) for the SG and MG rounds.

The distributions of average accuracies for females and males are different (Mann-Whitney, $p = 0.001$; Median test, $p = 0.002$). As Figure 15 shows, female players appear to be less accurate than male players. Notice though, that this average measure of accuracy corresponds to the rounds with no priming and no incentives. Average distance to the winning response is clearly higher for low sophisticated individuals. In a nutshell, there is a close relationship between our previous measure of sophistication and accuracy.
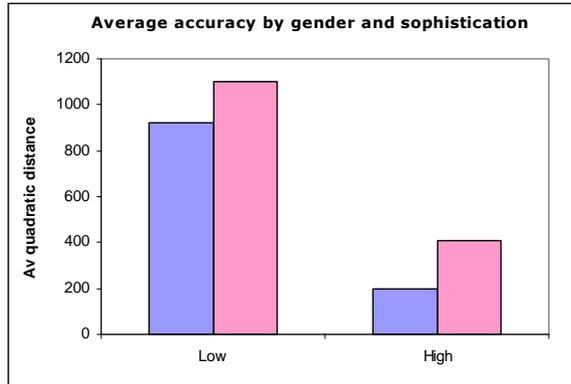
Figure 15: Average accuracy by gender and sophistication.

The distributions of accuracies for males and females also differ in the SG and MG rounds (Mann-Whitney, $p = 0.032$ and $p = 0.009$ respectively). Figure 16 shows that females display more sophistication (lower quadratic distance) than males when gender is primed. Moreover, gender differences in accuracy in this two rounds are not statistically significant for those subjects with above average accuracy in the first phase (Mann-Whitney, $p = 0.465$ in SG, $p = 0.199$ in MG). All these results are in line with those in section 4.1.2.

Figure 16: Average accuracy by gender and sophistication.
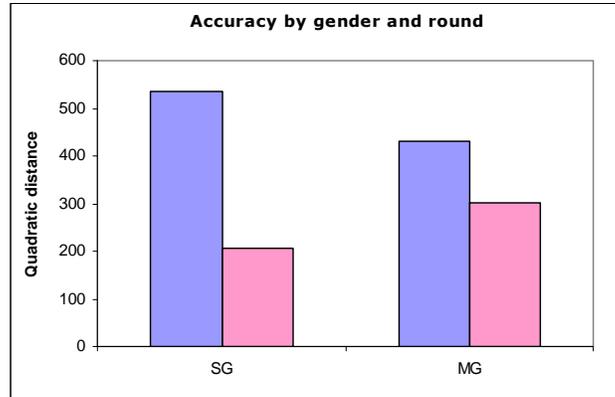
Recall that we observed that females thought more deeply about the game when we introduced incentives. On the other hand, males reacted to the composition of the group by showing more sophistication in the MG roundthan in the SG round. We revisit these results using our measure of accuracy.

Result 4.2 The distributions of accuracies across the NPNI and MG rounds do not differ for females in the top 30th percentile (Wilcoxon sign-rank, $p = 0.266$) but differ for the rest (Wilcoxon sign-rank, $p < 0.001$). The distributions do not differ across the SG and MG rounds.

The left panel of Figure 17 illustrates this result. The distribution of female accuracy shifts up when we introduced incentives. When we restrict attention to the top 30th percentile ($n = 16$) we find that these females did not react to incentives. This confirms the preliminary results that we obtained in section 4.2.1.

The right panel illustrates the increase in accuracy of males when they play against females. The shift, however, is not enough to be statistically significant. This shows that there is not a one to one correspondence between accuracy and sophistication defined as closeness to equilibrium play. Males show more sophistication but not more accuracy when they play in mixed gender groups compared to when they play only against males. They adjusted their responses to the fact that female respond lower numbers but not enough to become more accurate.
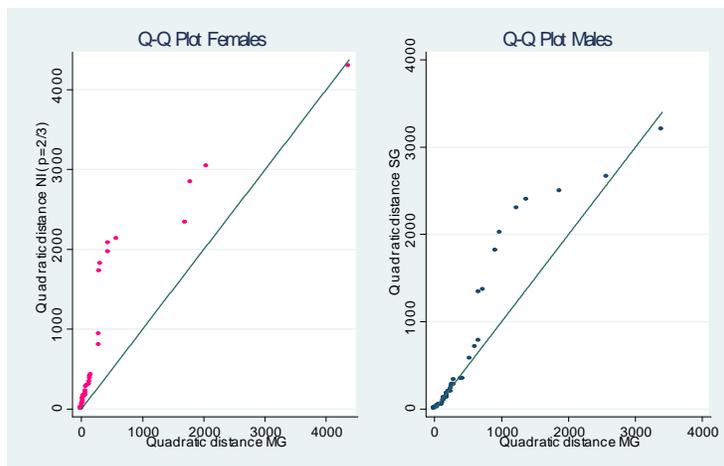
Figure 17: Q-Q plot of accuracy by gender comparing rounds.

Finally, we explore the relationship between average accuracy and beliefs. Figure 18 shows that males who believe that females respond higher numbers are more accurate (have a lower average distance to the winning response) than the rest of subjects. In addition, males who think that females are better in the game are much less accurate than the rest of males.

Figure 18: Average accuracy and stereotypes by responses to Q1 and Q2.

7   Discussion and conclusions

There is substantial experimental evidence of individual heterogeneity in depth of strategic reasoning. An individual's observed strategic sophisti-

cation depends on her mentalizing abilities, but also on her beliefs about the sophistication of others and on the size of monetary incentives. Gender might be relevant to all these three factors.

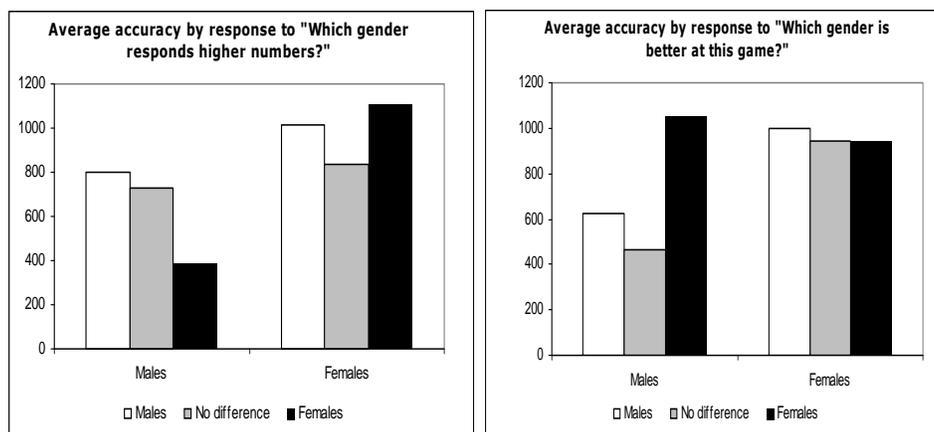In this paper, we explored the existence and endogeneity of gender differences in observed strategic sophistication. We used the beauty contest game as experimental device. We chose this game because it is competitive, because incentives and gender stereotypes can be easily manipulated and because it involves a relatively complex calculation task.

We reported results from two studies. Study 1 was a large between-subject experiment. The main result of this study is that gender differences in behavior appear only when no monetary prize was awarded. We interpret this result in line with Croson and Gneezy (2009): The presence of monetary incentives frames the interaction as competitive. Females, being more sensitive to contextual information, think more deeply about the game. An alternative, but not incompatible, explanation might be that males derive utility from winning regardless of whether a monteray ptize is at stake.

Study 2 was a within-subject experiment. In this study, we manipulated incentives and gender priming. The main results of this study were two: In line with Study 1, females display higher sophistication when incentives are introduced and gender is primed. However, gender differences do not disappear as in Study 1; they actually reverse. Females appear to be more sophisticated than males. This differential effect is due to gender priming.

The second result emerges out of the manipulation of the gender composition of the group. We observed that only males modified their responses then. Males displayed higher depth of reasoning when playing in mixed gender groups than when playing only against males. We thus find no evidence of stereotype threat affecting females. Responses to a questionnaire that participants filled at the end of each session suggest that males react to gender composition because they think females are better in the game. We conjecture that males are reacting to stereotype arousal by engaging further in the mentalizing process.

As a robustness check, we modified the definition of sophistication in two ways. First, we expanded the classification in levels of sophistication. We observed that medium sophisticated males are the ones reacting to the gender composition of the group. Highly sophisticated males do not change their responses despite the majority of them believe that females respond higher numbers. This suggests that beliefs about the sophistication of others affect choices only when the individual is sufficiently sophisticated but not so much to adhere to equilibrium play as a rule. Second, we measured strategic sophistication as the quadratic distance to the winning response. We found

that the most accurate females did not respond to our manipulation of incentives nor to gender priming.

Our results confirm previous findings in the experimental literature pointing out the importance of beliefs about the strategic sophistication of other players (e. g. Agranov et al, 2012). We also highlight the endogeneity of depth of reasoning with respect to incentives (also found in Alaoui and Penta, 2013). Regarding gender, our results show that the combination of incentives and gender priming enhances females' performance. This suggests that the gender differences observed in competitiveness in real-effort tasks may not translate to strategic settings. We find that when gender is made salient, females appear to be more sophisticated than males; males are aware of this and that they adjust their behavior accordingly.

Indirect evidence (e.g. Burnham et al, 2009) seemed to suggest that no gender differences exist in the beauty contest. We observe differences only when we manipulated incentives and gender priming. This might explain why there are so few studies reporting gender differences (or the lack of) in strategic interactions. In incentivized experiments, gender differences might arise only if gender is made salient. Iriberri and Rey-Biel (2013) find that mentioning gender is enough to modify performance in a competition. Nevertheless, we are aware that subjects' characteristics could correlate with gender, e. g. major of study in undergraduate populations, and thus create spurious gender differences.[21] Our subject pool in Study 2 was relatively homogeneous. Our participants were students of Economics or Business, of very similar age and ethnic and cultural background, so we are relatively free from this problem.

Our final remark refers to the portability of our results. The beauty contest is a relatively complex game with a big strategy space. Hence, it is to be expected that players use simpler rules of play. These rules might be different across games. Georganas et al (2010) find that the level of sophistication of subjects is similar across games of the same family but varies across families of games. By definition, simple rules of play are more sensitive to individual characteristics than equilibrium play. In fact, level-k theories can indeed be interpreted as models of rules of thumb grounded on "an instinctive reaction to the game" (Crawford et al, 2013). An open question is whether the gender differences in strategic sophistication that we uncover in the beauty contest remain in games where standard equilibrium predictions are more transparent and where subjects may resort to rules of thumb to a lesser extent. We plan to tackle this in our future research.

---

[21] We thank Colin Camerer for pointing this out.

References

[1] Agranov, M, Potamites, E, Schotter, A, and Tergiman, C. 2012. Beliefs and Endogenous Cognitive Levels: An Experimental Study, *Games and Economic Behavior*, 75(2): 449463.

[2] Alaoui, L, and Penta, A. 2013. Endogenous Depth of Reasoning, unpublished manuscript.

[3] Arad, A, and Rubinstein, A. 2012. The 11-20 Money Request Game: A Level-k Reasoning Study, *American Economic Review*, 102(7): 3561-3573.

[4] Baron-Cohen, S. 1991. Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), *Natural theories of mind: Evolution, development and simulation of everyday mindreading.* Oxford: Basil Blackwell.

[5] Bosch-Domenech, A, Garcia-Montalvo, J, Nagel, R and Satorra, A. 2002. One, Two, (Three), Infinity: Newspaper and Lab Beauty-Contest Experiments, *American Economic Review*, 92(5), 1687-1701.

[6] Brañas-Garza, P, Garcıa-Muñoz, T, and Hernan, R. 2012. Cognitive effort in the beauty contest game. *Journal of Economic Behavior & Organization*, 83(2), 254–260.

[7] Bühren, C, and Björn, F. 2010. Chess players' performance beyond 64 squares: A case study on the limitations of cognitive abilities transfer, unpublished manuscript, Philipps-Universität Marburg.

[8] Burnham, T, Cesarini, D, Johannesson, M, Lichtenstein, P, and Wallace, B. 2009. Higher cognitive ability is associated with lower entries in a p-beauty contest. *Journal of Economic Behavior & Organization*, 72(1): 171–175.

[9] Camerer, C F. 2003. B*ehavioral Game Theory: Experiments on Strategic Interaction*, Princeton, NJ: Princeton University Press.

[10] Camerer, C F, Ho, T-H, and Chong, J K. 2004. A Cognitive Hierarchy Model of Games, *Quarterly Journal of Economics*, 119(3): 861-898.

[11] Camerer, C F, Hogarth, R. 1999. The effects of financial incentives in experiments: A review and capital- labor production theory, *Journal of Risk and Uncertainty,* 19(1–3): 7–42.

[12] Choi, S. 2012. A Cognitive Hierarchy Model of Learning in Networks, *Review of Economic Design*, 16(2), 215-250.

[13] Coricelli, G, and Nagel, R. 2009. Neural Correlates of Depth of Strategic Reasoning inMedial Prefrontal Cortex, *Proceedings of the National Academy of Sciences*, 106(23): 9163-9168.

[14] Costa-Gomes, M A, Crawford, V P, and Broseta, B. 2001. Cognition and Behavior in Normal-Form Games: An Experimental Study, *Econometrica,* 69, 1193-1235.

[15] Costa-Gomes, M A, and Weizsäcker, G. 2008. Stated Beliefs and Play in Normal-Form Games, *Review of Economic Studies*, 75(3): 729–62.

[16] Cox, J C, Smith V L, and Walker, J M. 1992. Theory and Misbehavior of First-Price Auctions: Comment, *American Economic Review*, 82, 1392-1412.

[17] Crawford, V P, Costa-Gomes, M A, and Iriberri, N. 2013. Structural Models of Nonequilibrium Strategic Thinking: Theory, Evidence, and Applications, *Journal of Economic Literature*, 51, 5-62.

[18] Croson, R. and Gneezy, U. 2009. Gender differences in preferences, *Journal of Economic Literature*, 47(2): 1-27.

[19] Davidson, R, and Duclos, J-Y. 2000. Statistical inference for stochastic dominance and for the measurement of poverty and inequality, *Econometrica*, 68(6): 1435-1464.

[20] Dreber, A, von Essen, E, and Ranehill, E. 2011. Outrunning the Gender Gap – Boys and Girls Compete Equally, *Experimental Economics*, 14(4), 567-582.

[21] Frick, B. 2011. Gender differences in competitiveness: Empirical evidence from professional distance running, *Labour Economics*, 18(3):389-398.

[22] Fryer, R, and Levitt, S. 2010. An Empirical Analysis of the Gender Gap in Mathematics, *American Economic Journal: Applied Economics*, 2(2).

[23] Fryer, R, Levitt, S, and List, J A. 2009. Exploring the Impact of Financial Incentives on Stereotype Threat: Evidence from a Pilot Study, *American Economic Review*, 98(2):370-375.

[24] Georganas, S, Healy, P, and Weber, R. 2011. On the Persistence of Strategic Sophistication, unpublished manuscript.

[25] Gneezy U, Niederle M, and Rustichini A. 2003. Performance in competitive environments: gender differences, *Quarterly Journal of Economics*, 118:1049–74.

[26] Gneezy U and Rustichini A. 2004. Gender and competition at a young age, *American Econonomic Review*, 94:377–81.

[27] Ho, T-H, Camerer, C F, and Weigelt, K. 1998. Iterated Dominance and Iterated Best Response in Experimental 'p-Beauty Contests', *American Economic Review*, 88(4): 947-969.

[28] Iriberri, N, and Rey-Biel, P. 2013. Let's (Not) Talk about Sex: Gender Awareness and Stereotype-Threat on Performance under Competition, unpublished manuscript, Barcelona GSE.

[29] Kocher, M, and Sutter, M. 2005. The decision maker matters. Individual versus team behavior in experimental beauty-contest games, *Economic Journal*, 115: 200-223.

[30] Kovalchik, S, Camerer, C F, Grether, D M, Plott, C R, and Allman, J M. 2005. Aging and decision making: A comparison between neurologically healthy elderly and young individuals, *Journal of Economic Behavior and Organization*, 58, 79–94.

[31] Krendl, A C, Richeson, J A, Kelley, W M, and Heatherton, T F. 2008. The negative consequences of threat: An fMRI investigation of the neural mechanisms underlying women's underperformance in math, *Psychological Science*, 19(2), 168-175.

[32] Lindberg, S M, Hyde, J S, Petersen, J L, and Linn, M C. 2010. New trends in gender and mathematics performance: A meta-analysis, *Psychological Bulletin*, 136, 1123-1135.

[33] Nagel, R. 1995. Unraveling in Guessing Games: An Experimental Study, *American Economic Review*, 85(5): 1313-1326.

[34] O'Brien, L T, and Crandall, C S. 2003. Stereotype threat and arousal: Effects on women's math performance, *Personality and Social Psychology Bulletin,* 29, 782–789.

[35] Östling, R., Wang, J. T., Chou, E. Y., and Camerer, C.F. 2011. Testing Game Theory in the Field: Swedish LUPI Lottery Games. *American Economic Journal: Microeconomics*, 3(3): 1-33.

[36] Palacios-Huerta, I, and Volij, O. 2009. Field Centipedes, *American Economic Review*, 99(4): 1619-1635.

[37] Quinn, D, and Spencer, S. 2001. The interference of stereotype threat with women s generation of mathematical problem-solving strategies. *Journal of Social Issues*, 57, 55–71.

[38] Rey-Biel, P. 2009. Equilibrium Play and Best Response to (Stated) Beliefs in Normal Form Games, *Games and Economic Behavior* 65 (2): 572–85.

[39] Sheremeta, R. 2010. Experimental Comparison of Multi-Stage and One-Stage Contests, *Games and Economic Behavior*, 68, 731-747.

[40] Shurchkov, O. 2012. Under pressure: gender differences in output quality and quantity under competition and time constraints, *Journal of the European Economic Association*, 10(5): 1189–1213.

[41] Spencer, S J, Steele, C M, and Quinn, D M. 1999. Stereotype threat and women's math performance, *Journal of Experimental Social Psychology*, 35:4–28.

[42] Stahl, D O, and Wilson, P W. 1994. Experimental evidence on players' models of other players, *Journal of Economic Behavior & Organization*, 25, 309–327

[43] Stahl, D O, and Wilson, P R. 1995. On Players' Models of Other Players: Theory and Experimental Evidence, *Games and Economic Behavior*, 10(1): 218-254.

[44] Steele, C M. 1997. A threat in the air: how stereotypes shape intellectual identity and performance, *American Psychologist*, 52:613–29.

[45] Walton, G M, and Cohen, G L. 2003. Stereotype lift, *Journal of Experimental Social Psychology,* 39, 456–467.

[46] Weber, R A. 2003. 'Learning' with no feedback in a competitive guessing game, *Games and Economic Behavior,* 44(1): 134–144.

Appendix: Instructions (translated from Spanish)

GENERAL INSTRUCTIONS

Hello. Many thanks for taking part in this session.

The purpose of this session is to study how people make decisions in strategic settings.

The session is organized in two parts:

In the first part, you should answer a series of independent questions with the objective of becoming familiar with the rules of the experiment.

In the second part, you should answer another series of independent questions. You will compete with the rest of participants in your room for a monetary prize. The participant with the most correct answer will be the winner.

After reading these instructions you will find the first set of questions. We will read each question aloud. You will have time to answer each question before moving to the next one.

Read carefully each question and take the time you need to answer it.

It is very important that you remain silent during the whole session. Otherwise, the data collected will be useless.

Please do not go to the next question until we tell you to.

Before starting the experiment please write in the box below your participant number.

GENERIC ROUND QUESTION (PHASE 1)

Each one of you should choose a number between 0 and 100 with the objective of guessing (p fraction of) the average of the numbers chosen by all the participants in this room.

The winner will be the participant(s) whose answer is the closest to the (p fraction of the) average of all numbers chosen.

Which number do you choose?

Do not go to the next question until being instructed to do so.

INSTRUCTIONS PHASE 2

Now the second phase of the experiment begins.

In this phase, you will participate in tow independent rounds. The structure and rules are similar to those of phase 1 but there are two main differences:

1. The identity of the participants you will compete with will change in each round.

2. There will be two monetary prizes of 40 euros each.

At the end of the second phase, one of the two rounds will be chosen randomly. The winner of this round will obtain the prize. If there is more than one winner in the chosen round, the prize will be split among the winners.

Again questions will be read aloud.

Read carefully each question and take the time you need to answer it.

Recall that it is very important that you remain silent during the whole session. Otherwise, the data collected will be useless.

Please do not go to the next question until we tell you to.

Before continuing please write in the box below your participant number.

GENERIC ROUND QUESTION (PHASE 2)

Each one of you should choose a number between 0 and 100 with the objective of guessing the "2/3 of the average" of the numbers chosen in this question by all the participants in this room.

The winner will be the participant(s) whose answer is the closest to the 2/3 of the average of all numbers chosen in this question by all the participants in this room.

Which number do you choose?

Now close the booklet and remain silent.

2011

2011/1, Oppedisano, V; Turati, G.: "What are the causes of educational inequalities and of their evolution over time in Europe? Evidence from PISA"

2011/2, Dahlberg, M; Edmark, K; Lundqvist, H.: "Ethnic diversity and preferences for redistribution "

2011/3, Canova, L.; Vaglio, A.: "Why do educated mothers matter? A model of parental help"

2011/4, Delgado, F.J.; Lago-Peñas, S.; Mayor, M.: "On the determinants of local tax rates: new evidence from Spain"

2011/5, Piolatto, A.; Schuett, F.: "A model of music piracy with popularity-dependent copying costs"

2011/6, Duch, N.; García-Estévez, J.; Parellada, M.: "Universities and regional economic growth in Spanish regions"

2011/7, Duch, N.; García-Estévez, J.: "Do universities affect firms' location decisions? Evidence from Spain"

2011/8, Dahlberg, M.; Mörk, E.: "Is there an election cycle in public employment? Separating time effects from election year effects"

2011/9, Costas-Pérez, E.; Solé-Ollé, A.; Sorribas-Navarro, P.: "Corruption scandals, press reporting, and accountability. Evidence from Spanish mayors"

2011/10, Choi, A.; Calero, J.; Escardíbul, J.O.: "Hell to touch the sky? private tutoring and academic achievement in Korea"

2011/11, Mira Godinho, M.; Cartaxo, R.: "University patenting, licensing and technology transfer: how organizational context and available resources determine performance"

2011/12, Duch-Brown, N.; García-Quevedo, J.; Montolio, D.: "The link between public support and private R&D effort: What is the optimal subsidy?"

2011/13, Breuillé, M.L.; Duran-Vigneron, P.; Samson, A.L.: "To assemble to resemble? A study of tax disparities among French municipalities"

2011/14, McCann, P.; Ortega-Argilés, R.: "Smart specialisation, regional growth and applications to EU cohesion policy"

2011/15, Montolio, D.; Trillas, F.: "Regulatory federalism and industrial policy in broadband telecommunications"

2011/16, Pelegrín, A.; Bolancé, C.: "Offshoring and company characteristics: some evidence from the analysis of Spanish firm data"

2011/17, Lin, C.: "Give me your wired and your highly skilled: measuring the impact of immigration policy on employers and shareholders"

2011/18, Bianchini, L.; Revelli, F.: "Green polities: urban environmental performance and government popularity"

2011/19, López Real, J.: "Family reunification or point-based immigration system? The case of the U.S. and Mexico"

2011/20, Bogliacino, F.; Piva, M.; Vivarelli, M.: "The impact of R&D on employment in Europe: a firm-level analysis"

2011/21, Tonello, M.: "Mechanisms of peer interactions between native and non-native students: rejection or integration?"

2011/22, García-Quevedo, J.; Mas-Verdú, F.; Montolio, D.: "What type of innovative firms acquire knowledge intensive services and from which suppliers?"

2011/23, Banal-Estañol, A.; Macho-Stadler, I.; Pérez-Castrillo, D.: "Research output from university-industry collaborative projects"

2011/24, Ligthart, J.E.; Van Oudheusden, P.: "In government we trust: the role of fiscal decentralization"

2011/25, Mongrain, S.; Wilson, J.D.: "Tax competition with heterogeneous capital mobility"

2011/26, Caruso, R.; Costa, J.; Ricciuti, R.: "The probability of military rule in Africa, 1970-2007"

2011/27, Solé-Ollé, A.; Viladecans-Marsal, E.: "Local spending and the housing boom"

2011/28, Simón, H.; Ramos, R.; Sanromá, E.: "Occupational mobility of immigrants in a low skilled economy. The Spanish case"

2011/29, Piolatto, A.; Trotin, G.: "Optimal tax enforcement under prospect theory"

2011/30, Montolio, D; Piolatto, A.: "Financing public education when altruistic agents have retirement concerns"

2011/31, García-Quevedo, J.; Pellegrino, G.; Vivarelli, M.: "The determinants of YICs' R&D activity"

2011/32, Goodspeed, T.J.: "Corruption, accountability, and decentralization: theory and evidence from Mexico"

2011/33, Pedraja, F.; Cordero, J.M.: "Analysis of alternative proposals to reform the Spanish intergovernmental transfer system for municipalities"

2011/34, Jofre-Monseny, J.; Sorribas-Navarro, P.; Vázquez-Grenno, J.: "Welfare spending and ethnic heterogeneity: evidence from a massive immigration wave"

2011/35, Lyytikäinen, T.: "Tax competition among local governments: evidence from a property tax reform in Finland"

2011/36, Brülhart, M.; Schmidheiny, K.: "Estimating the Rivalness of State-Level Inward FDI"

2011/37, García-Pérez, J.I.; Hidalgo-Hidalgo, M.; Robles-Zurita, J.A.: "Does grade retention affect achievement? Some evidence from Pisa"

2011/38, Boffa, f.; Panzar. J.: "Bottleneck co-ownership as a regulatory alternative"

2011/39, González-Val, R.; Olmo, J.: "Growth in a cross-section of cities: location, increasing returns or random growth?"

2011/40, Anesi, V.; De Donder, P.: "Voting under the threat of secession: accommodation vs. repression"

2011/41, Di Pietro, G.; Mora, T.: "The effect of the l'Aquila earthquake on labour market outcomes"

2011/42, Brueckner, J.K.; Neumark, D.: "Beaches, sunshine, and public-sector pay: theory and evidence on amenities and rent extraction by government workers"

2011/43, Cortés, D.: "Decentralization of government and contracting with the private sector"

2011/44, Turati, G.; Montolio, D.; Piacenza, M.: "Fiscal decentralisation, private school funding, and students' achievements. A tale from two Roman catholic countries"

2012

2012/1, Montolio, D.; Trujillo, E.: "What drives investment in telecommunications? The role of regulation, firms' internationalization and market knowledge"

2012/2, Giesen, K.; Suedekum, J.: "The size distribution across all "cities": a unifying approach"

2012/3, Foremny, D.; Riedel, N.: "Business taxes and the electoral cycle"

2012/4, García-Estévez, J.; Duch-Brown, N.: "Student graduation: to what extent does university expenditure matter?"

2012/5, Durán-Cabré, J.M.; Esteller-Moré, A.; Salvadori, L.: "Empirical evidence on horizontal competition in tax enforcement"

2012/6, Pickering, A.C.; Rockey, J.: "Ideology and the growth of US state government"

2012/7, Vergolini, L.; Zanini, N.: "How does aid matter? The effect of financial aid on university enrolment decisions"

2012/8, Backus, P.: "Gibrat's law and legacy for non-profit organisations: a non-parametric analysis"

2012/9, Jofre-Monseny, J.; Marín-López, R.; Viladecans-Marsal, E.: "What underlies localization and urbanization economies? Evidence from the location of new firms"

2012/10, Mantovani, A.; Vandekerckhove, J.: "The strategic interplay between bundling and merging in complementary markets"

2012/11, Garcia-López, M.A.: "Urban spatial structure, suburbanization and transportation in Barcelona"

2012/12, Revelli, F.: "Business taxation and economic performance in hierarchical government structures"

2012/13, Arqué-Castells, P.; Mohnen, P.: "Sunk costs, extensive R&D subsidies and permanent inducement effects"

2012/14, Boffa, F.; Piolatto, A.; Ponzetto, G.: "Centralization and accountability: theory and evidence from the Clean Air Act"

2012/15, Cheshire, P.C.; Hilber, C.A.L.; Kaplanis, I.: "Land use regulation and productivity – land matters: evidence from a UK supermarket chain"

2012/16, Choi, A.; Calero, J.: "The contribution of the disabled to the attainment of the Europe 2020 strategy headline targets"

2012/17, Silva, J.I.; Vázquez-Grenno, J.: "The ins and outs of unemployment in a two-tier labor market"

2012/18, González-Val, R.; Lanaspa, L.; Sanz, F.: "New evidence on Gibrat's law for cities"

2012/19, Vázquez-Grenno, J.: "Job search methods in times of crisis: native and immigrant strategies in Spain"

2012/20, Lessmann, C.: "Regional inequality and decentralization – an empirical analysis"

2012/21, Nuevo-Chiquero, A.: "Trends in shotgun marriages: the pill, the will or the cost?"

2012/22, Piil Damm, A.: "Neighborhood quality and labor market outcomes: evidence from quasi-random neighborhood assignment of immigrants"

2012/23, Ploeckl, F.: "Space, settlements, towns: the influence of geography and market access on settlement distribution and urbanization"

2012/24, Algan, Y.; Hémet, C.; Laitin, D.: "Diversity and local public goods: a natural experiment with exogenous residential allocation"

2012/25, Martinez, D.; Sjögren, T.: "Vertical externalities with lump-sum taxes: how much difference does unemployment make?"

2012/26, Cubel, M.; Sanchez-Pages, S.: "The effect of within-group inequality in a conflict against a unitary threat"

2012/27, Andini, M.; De Blasio, G.; Duranton, G.; Strange, W.C.: "Marshallian labor market pooling: evidence from Italy"

2012/28, Solé-Ollé, A.; Viladecans-Marsal, E.: "Do political parties matter for local land use policies?"

2012/29, Buonanno, P.; Durante, R.; Prarolo, G.; Vanin, P.: "Poor institutions, rich mines: resource curse and the origins of the Sicilian mafia"

2012/30, Anghel, B.; Cabrales, A.; Carro, J.M.: "Evaluating a bilingual education program in Spain: the impact beyond foreign language learning"

2012/31, Curto-Grau, M.; Solé-Ollé, A.; Sorribas-Navarro, P.: "Partisan targeting of inter-governmental transfers & state interference in local elections: evidence from Spain"

2012/32, Kappeler, A.; Solé-Ollé, A.; Stephan, A.; Välilä, T.: "Does fiscal decentralization foster regional investment in productive infrastructure?"

2012/33, Rizzo, L.; Zanardi, A.: "Single vs double ballot and party coalitions: the impact on fiscal policy. Evidence from Italy"

2012/34, Ramachandran, R.: "Language use in education and primary schooling attainment: evidence from a natural experiment in Ethiopia"

2012/35, Rothstein, J.: "Teacher quality policy when supply matters"

2012/36, Ahlfeldt, G.M.: "The hidden dimensions of urbanity"

2012/37, Mora, T.; Gil, J.; Sicras-Mainar, A.: "The influence of BMI, obesity and overweight on medical costs: a panel data approach"

2012/38, Pelegrín, A.; García-Quevedo, J.: "Which firms are involved in foreign vertical integration?"

2012/39, Agasisti, T.; Longobardi, S.: "Inequality in education: can Italian disadvantaged students close the gap? A focus on resilience in the Italian school system"

2013

2013/1, Sánchez-Vidal, M.; González-Val, R.; Viladecans-Marsal, E.: "Sequential city growth in the US: does age matter?"

2013/2, Hortas Rico, M.: "Sprawl, blight and the role of urban containment policies. Evidence from US cities"

2013/3, Lampón, J.F.; Cabanelas-Lorenzo, P-; Lago-Peñas, S.: "Why firms relocate their production overseas? The answer lies inside: corporate, logistic and technological determinants"

2013/4, Montolio, D.; Planells, S.: "Does tourism boost criminal activity? Evidence from a top touristic country"

2013/5, Garcia-López, M.A.; Holl, A.; Viladecans-Marsal, E.: "Suburbanization and highways: when the Romans, the Bourbons and the first cars still shape Spanish cities"

2013/6, Bosch, N.; Espasa, M.; Montolio, D.: "Should large Spanish municipalities be financially compensated? Costs and benefits of being a capital/central municipality"

2013/7, Escardíbul, J.O.; Mora, T.: "Teacher gender and student performance in mathematics. Evidence from Catalonia"

2013/8, Arqué-Castells, P.; Viladecans-Marsal, E.: "Banking towards development: evidence from the Spanish banking expansion plan"

2013/9, Asensio, J.; Gómez-Lobo, A.; Matas, A.: "How effective are policies to reduce gasoline consumption? Evaluating a quasi-natural experiment in Spain"

2013/10, Jofre-Monseny, J.: "The effects of unemployment benefits on migration in lagging regions"

2013/11, Segarra, A.; García-Quevedo, J.; Teruel, M.: "Financial constraints and the failure of innovation projects"

2013/12, Jerrim, J.; Choi, A.: "The mathematics skills of school children: How does England compare to the high performing East Asian jurisdictions?"

2013/13, González-Val, R.; Tirado-Fabregat, D.A.; Viladecans-Marsal, E.: "Market potential and city growth: Spain 1860-1960"

2013/14, Lundqvist, H.: "Is it worth it? On the returns to holding political office"

2013/15, Ahlfeldt, G.M.; Maennig, W.: "Homevoters vs. leasevoters: a spatial analysis of airport effects"

2013/16, Lampón, J.F.; Lago-Peñas, S.: "Factors behind international relocation and changes in production geography in the European automobile components industry"

2013/17, Guío, J.M.; Choi, A.: "Evolution of the school failure risk during the 2000 decade in Spain: analysis of Pisa results with a two-level logistic mode"

2013/18, Dahlby, B.; Rodden, J.: "A political economy model of the vertical fiscal gap and vertical fiscal imbalances in a federation"

2013/19, Acacia, F.; Cubel, M.: "Strategic voting and happiness"

2013/20, Hellerstein, J.K.; Kutzbach, M.J.; Neumark, D.: "Do labor market networks have an important spatial dimension?"

2013/21, Pellegrino, G.; Savona, M.: "Is money all? Financing versus knowledge and demand constraints to innovation"

2013/22, Lin, J.: "Regional resilience"

2013/23, Costa-Campi, M.T.; Duch-Brown, N.; García-Quevedo, J.: "R&D drivers and obstacles to innovation in the energy industry"

2013/24, Huisman, R.; Stradnic, V.; Westgaard, S.: "Renewable energy and electricity prices: indirect empirical evidence from hydro power"

2013/25, Dargaud, E.; Mantovani, A.; Reggiani, C.: "The fight against cartels: a transatlantic perspective"

2013/26, Lambertini, L.; Mantovani, A.: "Feedback equilibria in a dynamic renewable resource oligopoly: pre-emption, voracity and exhaustion"

2013/27, Feld, L.P.; Kalb, A.; Moessinger, M.D.; Osterloh, S.: "Sovereign bond market reactions to fiscal rules and no-bailout clauses – the Swiss experience"

2013/28, Hilber, C.A.L.; Vermeulen, W.: "The impact of supply constraints on house prices in England"

2013/29, Revelli, F.: "Tax limits and local democracy"

2013/30, Wang, R.; Wang, W.: "Dress-up contest: a dark side of fiscal decentralization"

2013/31, Dargaud, E.; Mantovani, A.; Reggiani, C.: "The fight against cartels: a transatlantic perspective"

2013/32, Saarimaa, T.; Tukiainen, J.: "Local representation and strategic voting: evidence from electoral boundary reforms"

2013/33, Agasisti, T.; Murtinu, S.: "Are we wasting public money? No! The effects of grants on Italian university students' performances"

2013/34, Flacher, D.; Harari-Kermadec, H.; Moulin, L.: "Financing higher education: a contributory scheme"

2013/35, Carozzi, F.; Repetto, L.: "Sending the pork home: birth town bias in transfers to Italian municipalities"

2013/36, Coad, A.; Frankish, J.S.; Roberts, R.G.; Storey, D.J.: "New venture survival and growth: Does the fog lift?"

2013/37, Giulietti, M.; Grossi, L.; Waterson, M.: "Revenues from storage in a competitive electricity market: Empirical evidence from Great Britain"

2014

2014/1, Montolio, D.; Planells-Struse, S.: "When police patrols matter. The effect of police proximity on citizens' crime risk perception"

2014/2, Garcia-López, M.A.; Solé-Ollé, A.; Viladecans-Marsal, E.: "Do land use policies follow road construction?"

2014/3, Piolatto, A.; Rablen, M.D.: "Prospect theory and tax evasion: a reconsideration of the Yitzhaki puzzle"

2014/4, Cuberes, D.; González-Val, R.: "The effect of the Spanish Reconquest on Iberian Cities"

2014/5, Durán-Cabré, J.M.; Esteller-Moré, E.: "Tax professionals' view of the Spanish tax system: efficiency, equity and tax planning"

2014/6, Cubel, M.; Sanchez-Pages, S.: "Difference-form group contests"

2014/7, Del Rey, E.; Racionero, M.: "Choosing the type of income-contingent loan: risk-sharing versus risk-pooling"

2014/8, Torregrosa Hetland, S.: "A fiscal revolution? Progressivity in the Spanish tax system, 1960-1990"

2014/9, Piolatto, A.: "Itemised deductions: a device to reduce tax evasion"

2014/10, Costa, M.T.; García-Quevedo, J.; Segarra, A.: "Energy efficiency determinants: an empirical analysis of Spanish innovative firms"

2014/11, García-Quevedo, J.; Pellegrino, G.; Savona, M.: "Reviving demand-pull perspectives: the effect of demand uncertainty and stagnancy on R&D strategy"

2014/12, Calero, J.; Escardíbul, J.O.: "Barriers to non-formal professional training in Spain in periods of economic growth and crisis. An analysis with special attention to the effect of the previous human capital of workers"

Document de