Journal of Economic Behavior & Organization 99 (2014) 126–154

Contents lists available at ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

Gender differences in risk aversion: Do single-sex environments affect their development?☆

Alison Booth[b,*], Lina Cardona-Sosa[c], Patrick Nolen[a]

[a] University of Essex, United Kingdom
[b] Australian National University, Australia
[c] Central Bank of Colombia and Universidad EAFIT, Colombia

ARTICLE INFO

Article history:
Received 12 September 2012
Received in revised form 11 October 2013
Accepted 23 December 2013
Available online 2 January 2014

JEL classification:
C9
C91
C92
J16
D01
D80
J16
J24

Keywords:
Gender
Risk preferences
Single-sex groups
Cognitive ability
coeducation
gender identity

ABSTRACT

Single-sex classes within coeducational environments are likely to modify students' risk-taking attitudes in economically important ways. To test this, we designed a controlled experiment using first year college students who made choices over real-stakes lotteries at two distinct dates. Students were randomly assigned to weekly classes of three types: all female, all male, and coeducational. They were not allowed to change group subsequently. We found that women are less likely to make risky choices than men at both dates. However, after eight weeks in a single-sex class environment, women were significantly *more* likely to choose the lottery than their counterparts in coeducational groups. These results are robust to the inclusion of controls for IQ and for personality type, as well as to a number of sensitivity tests. Our findings suggest that observed gender differences in behaviour under uncertainty found in previous studies might partly reflect social learning rather than inherent gender traits.

1. Introduction

Women are under-represented in high-paying jobs and in high-level occupations.[1] Recent studies in experimental economics have examined to what degree this under-representation may be due to innate differences between

☆ For helpful comments on an earlier draft, we thank the editor Ragan Petrie, two anonymous referees, Arthur van Soest, Elaine Liu, and seminar and conference participants. The experiment described in this paper took place at the University of Essex, and received approval from the University's Ethics Committee. We are grateful to the timetabling office for undertaking the random assignment by gender, and to the Ph.D. students who assisted in conducting the experiment.

* Corresponding author at: Research School of Economics, Australian National University, Arndt Building, Acton 0200, ACT, Australia. Tel.: +61 02 6125 3285; fax: +61 02 6125 5124.

E-mail address: Alison.booth@anu.edu.au (A. Booth).

[1] Bertrand and Hallock (2001) look at top corporate jobs in the USA. Arulampalam et al. (2007) summarise the pay gap across the wages distribution in Europe and conjecture as to causes. Blau and Kahn (2006) discuss the evolution of the gender pay gap in the USA over time.

0167-2681/$ – see front matter
http://dx.doi.org/10.1016/j.jebo.2013.12.017

men and women.[2] It has been suggested that gender differences in risk aversion, feedback preferences or fondness for competition may help explain observed gender disparities. For example, if women are more risk averse than men, and if much of the remuneration in high-paying jobs consists of bonuses linked to a company's performance, relatively fewer women will choose high-paying jobs because of the uncertainty.

It is important for policy purposes to understand the extent to which risk preferences are innate or are shaped by the environment. If risk preferences are innate, under-representation of women in certain areas may be solved only by changing the remuneration method. But if risk attitudes are primarily shaped by the environment, changing the educational or training context could help address under-representation. Thus the policy prescription for dealing with under-representation of women in high-paying jobs will depend upon whether or not the reason for the absence is innate to one's gender.

Only recently have economists begun to explore why women and men might have different risk preferences. Broadly speaking, those differences may be due to either nurture, nature, or some combination of the two. Gneezy et al. (2009) explore the role that culture plays in determining gender differences in competitive behaviour. They investigate two distinct societies – the patriarchal Maasai tribe of Tanzania and the matrilineal Khasi tribe in India. While they find that, in the patriarchal society, women are less competitive than men – which is consistent with experimental data from Western cultures – in the matrilineal society, women are more competitive than men. Indeed, the Khasi women were found to be as competitive as Maasai men. The authors interpret this as evidence that culture has an influence. Interestingly, however, they find no evidence that, on average, there are gender differences in risk attitudes within either society.

Also using subjects from distinct environments or 'cultures', Booth and Nolen (2012b) examined the effect on preferences over risk of two types of environmental influences – randomly assigned experimental peer-groups and educational environment (single-sex or coeducational). The latter represents longer-run nurturing experiences, while the former captures short-run environmental effects.[3] The experimental subjects in that study were UK students who were just under 15 years old, and attending either single-sex or coeducational state-funded high schools. The authors found that the gender composition of the experimental group to which a student was randomly assigned, as well as the gender mix of the school the student attended, affected decisions on whether or not to enter a real-stakes lottery with an imposed coefficient of relative risk aversion.

This present paper extends our previous experimental work on risk preferences by designing an experiment that looks at the effects of nurture over time and that also endogenizes the coefficient of relative risk aversion. Our goals are to see if random assignment to a single-sex group produces effects on individuals' behaviour immediately upon assignment and also after some weeks' exposure to that particular environment. Our subjects are first year college (university) students (and thus a different subject pool to that of Booth and Nolen, 2012b). These students were required to make choices over real-stakes lotteries at two distinct dates. Our 'nurturing' environment is the experimental peer-group or class to which students were randomly assigned by the timetabling office. The class groups were of three different types: all-female; all-male; or mixed gender.

Given that the class group was randomly assigned, there are no issues of endogeneity. We test if the experimental environment influences the behaviour under uncertainty of men and women, and we are particularly interested in seeing if individuals who are placed in a same-sex group for the experiment make different choices to otherwise identical individuals placed in a mixed group. Our measure of risk aversion involved students making choices over real-stakes lotteries.

While this group effect has been explored in previous work by Gneezy et al. (2003), Niederle and Vesterlund (2007) and Datta Gupta et al. (2005), those studies focused on competitive tasks. They did not investigate risk attitudes nor did they explore how risk preferences may change over time – the main focus of our investigation.

In our experiment, we repeated the risky-choices rounds at two different time periods: the initial week of term, and again in the eighth week. Our results show that, at both dates, women are significantly *less* likely to make risky choices than men. This finding of gender differences in choices under uncertainty is in line with the majority of experimental studies investigating risk choices at a single point in time, as summarised in Eckel and Grossman (2008) and Croson and Gneezy (2009).[4] We also found that, in the initial week, the sex composition of the classes into which individuals had been randomly assigned had no impact on the choices over real stakes lotteries. However, after eight weekly sessions in the single-sex class environment, women were significantly *more* likely to choose the lottery than their counterparts in coeducational groups, and the magnitude of the effect was quite large. No such result was found for men in the single-sex groups. Moreover, our

[2] For a survey of new perspectives on gender in economics, see Bertrand (2011). See Charness and Kuhn (2011) for a survey on laboratory experiments in labour economics.

[3] In a companion paper, Booth and Nolen (2012a) investigated how competitive behaviour (including the choice between piece-rates and tournaments) is affected by single-sex experimental peer-groups and single-sex schooling.

[4] Harbaugh et al. (2002) find no significant sex differences in risk aversion, although they do find that risk aversion varies with age. Our subjects, however, are all about the same age. Schubert et al. (1999), using as subjects undergraduates from the University of Zurich, show that the context makes a difference. While women do not generally make less risky financial choices than men, they are less likely to engage in an abstract gamble.

results are robust to the inclusion or exclusion of controls for personality type, as well as to a battery of sensitivity tests reported in Section 5.

Our findings are important because they suggest that observed gender differences in behaviour under uncertainty found in previous studies might actually reflect social learning rather than inherent gender traits. Of course this is not to say that inherent gender traits do not exist. Rather it suggests that they can be modified by the environment in which a woman is placed. In particular, single-sex classes within a coeducational environment were found to significantly alter young women's choices over time.

The crucial differences between the new experiment reported in this paper and Booth and Nolen (2012b) are as follows. First, Booth and Nolen (2012b) used a sample of school children (average age just under 15 years) from government-funded schools in two adjacent UK counties, while in the current paper we use first year college students entering an Economics or Business degree, whose mean age was 19, and who came from a number of different countries and school types (public *and* private). Second, Booth and Nolen (2012b) used a simple lottery with an *imposed* coefficient of relative risk aversion, instead of eliciting a subject's coefficient of relative risk aversion using a series of questions as we do in the present paper.[5]

The remainder of the paper is set out as follows. In Section 2 we describe our subjects and the dataset. Section 3 explores attrition and shows that this is not a problem with our data. Section 4 reports the means and discusses class attendance, before presenting results of the models estimating the correlates of the number of risky choices made. Sensitivity checks are reported in Section 5, while Section 6 concludes.

2. Subjects, protocol and data

Our subject pool consisted of first year undergraduate students registered for the course, *Introduction to Economics*, at the University of Essex at the start of the 2010–2011 academic year. Prior to arrival, the students were randomly assigned to small weekly classes which, during term-time, run in tandem with the lecture course. During their first lecture, students filled in a demographic questionnaire and, as part of a paid experiment during their first class, completed a cognitive ability test and risk questionnaire. Eight weeks later students then took part in another paid experiment by filling in a second risk questionnaire during their class that week. Students were thus in the economics class environment for 1 h per week over eight weeks. At no stage were the students told the purpose of the experiment, what the experiment would involve, nor that it was to be repeated in eight weeks time. All students enrolled in the course are supposed to attend the classes and do the compulsory exercises. Lectures and classes begin immediately after student arrival at campus.

The results from the risk questionnaires form the dependent variable in our empirical analysis, with other information used as controls. Our main interest is whether or not women assigned to all-female classes within a coeducational environment take more risks than those in coeducational classes.[6] We are also interested in controlling for the impact of cognitive ability, given the recent evidence suggesting that individuals of higher cognitive ability are more likely to take risks, ceteris paribus (see Burks et al., 2009; Dohmen et al., 2010).

Classes were taught by graduate teaching assistants (GTAs), who were Ph.D. students hired in a competitive hiring process from the pool of Ph.D. applicants.[7] GTAs had two types of training: first, 3 h of general training provided by the university and afterwards augmented by a 1-h session with the Economics Department deputy director of graduate studies. Second, one of us provided specific training for our experiment. This involved explaining to GTAs that tutorial classes would comprise three types (all-male, all-female, and coed classes), and taking the GTAs through the procedures on PowerPoint slides for the first session. This part of the training focused on learning protocol and how to run the first session. The GTAs were not told why we were conducting the experiment or our intended outcome variables. It is therefore highly unlikely that the GTAs would influence students' choices or outcomes. Moreover, GTAs were told not to discuss the experiments or the class arrangements with students. (An appendix F can be found in supplementary material related to this article, as detailed at the end of the paper, gives instructions provided to GTAs and the set of slides they were required to use in the experiment.) In Section 4 we discuss in detail the distribution of GTAs across classes.

[5] In Booth and Nolen (2012b), girls and boys chose between Option 1 (£5 for certain) and Option 2 (flip a coin and get £11 if the coin came up heads or £2 if the coin came up tails). The expected monetary value exceeds the certain outcome. The dependent variable in that paper took the value one if the individual chose to enter the lottery and zero otherwise. The implied coefficient of relative risk aversion (CRRA) was 0.8. We *imposed* the coefficient of relative risk aversion because we had limited resources and a limited number of rounds.

[6] One referee asked if the same effect might be found for extracurricular activities such as all-female undergraduate clubs. Societies and clubs at the University of Essex are not allowed to be single-sex, with the exception of sports clubs. Not only do few of these restrict entry based on gender (two examples are 'Rugby Women' and 'Rugby Men'), but those that do comprise individuals who have self-selected on the basis of sporting prowess and on willingness to compete and to take calculated risks. This is in stark contrast to our randomly assigned class-types – randomised at the student level so our experiment is internally valid.

[7] In classes the instructors discuss with students problem-sets that relate directly to the material taught in that week's *Introduction to Economics* lectures. Students are assigned to specific classes and attend *only* that class in each week for a period of 20 weeks over the academic year.

Table 1
Difference in risky choices between sessions by gender and class composition.

	Women				Men			
	First Session (1)	Second Session (2)	Difference (2) – (1): (3)	Std. Dev. of (3): (4)	First Session (5)	Second Session (6)	Difference (6) – (5): (7)	Std. Dev. of (7): (8)
Single-Sex Classes								
Number of risky choices made	10.48	12.50	2.02**	0.99	12.15	13.11	0.96	0.62
Observations	44	44	44	44	81	81	81	81
Coeducational Classes								
Number of risky choices made	10.23	10.32	0.10	0.94	11.44	12.60	1.16*	0.65
Observations	31	31	31	31	63	63	63	63

Notes: *, ** and ***, corresponds to 10%, 5% and 1% levels of significance, respectively.

2.1. Experimental measures of risk aversion

We used paid experiments to measure subjects' willingness to take risks at two points in time.[8] In both sessions, students were asked to answer a questionnaire designed to assess subjects' risk preferences. This risk questionnaire – reported in full in Appendix A and following the format of Dohmen et al. (2010) – consisted of 20 rows. Each row had two columns, A and B. Column A was a 'safe' option: subjects would receive the amount of money stated if they chose the option in column A. The option presented in each row of column B was a lottery where individuals had a 50:50 chance of receiving £30. (At the time of writing, £30 was worth around US$48.80.) The safe option in column A started with 0 pounds (first row) and increased until £19 in the final (20th) row. Students were asked to make a choice between column A and B in each row (i.e. 20 choices). Subjects were told that a single row would be selected at random if that round were chosen for payment, and that 10% of all the participants (also selected at random) would receive the corresponding payment.

A risk-neutral student would choose the lottery up to the point where the safe option offers the same expected value as the lottery (£15). This is where the individual is indifferent between the safe option and the lottery; in our experiment this occurs in row 16. Risk averse students would choose the safe option before row 16 and risk loving students would choose the lottery even after the safe bet of £15 has been offered. Hence the row where students switched serves as an indicator of risk aversion.

If a student had monotonic preferences then she would have a unique switching point. However, nothing prevents them from switching more than once.[9] To account for switching back and forth, we follow Holt and Laury (2002) and use as our dependent variable the number of rows where the student chose the lottery option. Thus our measure of risk aversion is the number of risky choices made by the subject.[10] The advantage of this measure is that we account for the fact that some of the individuals never switched column, i.e. we account for those who always chose the risky option or never did so. During session 1, 70% of the students in our sample switched row only once; 4% never switched and 25% switched more than once. During the second session, students switching only once account for 73% of the sample, 10% never switched during the second session and 16% switched more than once. (We will return to switching issues in Section 5.3 of the sensitivity analysis.)

Table 1 shows differences in risky choices, across the two experiment sessions and class-types, for the sample used in the analysis. The first four columns provide information for women, and the last four columns provide information for men, while the top panels give figures for single-sex classes and the bottom panels give figures for coed classes. There are 44 women in all-female classes and 31 in coeducational classes, while there are 81 men in all-male classes and 63 in coeducational classes. For the all-female classes, the raw data illustrate a significant difference at the 5% level across sessions in the mean number of risky choices. Women in all-female classes chose 10.48 risky choices in the first session and 12.50 in the second

[8] Session 1 not only involved a 20-min IQ test and the risk matrix, but also two rounds of two-digit addition (see the experiment protocol in the last Appendix to this paper, Appendix F). The only payments involved in Session 1 were for risk or addition, as also explained in the appendix. To determine which round was paid, a number was randomly chosen by one of the authors from the set {1, 2, 3} – and where 1 corresponded to the risk questionnaire; 2 to the first round of two-digit addition; 3 to the second round of two-digit addition. If the number 1 were chosen, the student would be paid in accordance with a randomly selected row from the risk matrix. If, for that row, the student had chosen the safe outcome she would be paid that, and if she had chosen the risky outcome in column B, one of us would flip a coin and the person would be paid based on the outcome. If the number 2 were chosen, the student would be paid 20 pence for each correctly solved problem in that round. If the number 3 were chosen, the student would get paid £1 for each problem solved correctly in that round if that student was in the top 20% of performers in this class. Note that if 1 (risk) were randomly chosen, then a random 10% of people were paid, but if one of the two digit rounds was chosen we paid everyone.

[9] This is also similar to Holt and Laury (2002) and Harrison et al. (2007). Unlike Dohmen et al. (2010), we did not ask subjects if they wished to stop completing the table once they had made a switch between safe and risky choices.

[10] The row where individual changed and the number of rows where the lottery was chosen were perfectly correlated. Nevertheless we experimented with using, as a measure for risk, the row where the student changed and the average row (in the case of multiple switching). Our main conclusions remain unchanged by this. This will be discussed under robustness checks later in the paper.

session conducted eight weeks later.[11] This compares with women in coeducational classes, who chose 10.23 risky choices in the first session and 10.32 in the second.[12]

From the last four columns of Table 1, it can be seen that young men in all-male classes chose 12.15 risky choices in the first session and 13.11 in the second, a difference that is not statistically significant. However, men in the coeducational class chose 11.44 risky choices in the first session and 12.60 in the second: this difference is statistically significant, albeit only at the at the 10% level. On average, the risk preferences of women in single-sex classes changed the most, followed by men in coeducational classes. Women in the coeducational classes changed the least. (The coeducational classes will form the base in our regressions presented in Section 4.)

Fig. 1a and b shows, for women and men, respectively, the difference between the number of risky choices made in the second session and the number made in the first session, disaggregated by class-type and for the sample used in the analysis. Inspection of Fig. 1a shows that, for women attending all-female classes, there is a greater increase than for women attending coeducational classes (the histogram in the bottom panel has a longer positive tail than that in the top panel). This implies that women in the treatment group increased the number of risky choices made during the second session in comparison to those women in the coeducational classes. For men, however, the difference in the number of risky choices across sessions for each class-type is less clear. Men attending all-male classes do not show a significant increase in the risky choices made in the second session in comparison with men attending coeducational classes, as Table 1 also reveals.[13]

2.2. Other controls

As well as by class environment, an individual's preferences may be influenced by variables such as subject of study, the gender of the class teacher, cultural factors encapsulated in country-of-origin, the gender environment in high school, and cognitive ability.[14] (In our robustness checks reported in Section 5, we control for the family background characteristics and personality type using information obtained from the demographic questionnaire.)

Burks et al. (2009) argue that differences in perception of risky options due to cognitive ability may systematically affect individuals' choices. The more complex is an option, the larger the noise. If people of high cognitive skills perceive a complex option more precisely than people with low cognitive skills, they will be more likely to choose riskier options. Burks et al. (2009), using a sample of 1000 trainee truckers in the US, found that lower cognitive ability (as measured by a nonverbal IQ test, Raven's matrices) is associated with greater risk aversion and more pronounced impatience.[15] Dohmen et al. (2010) used a more representative sample and a different IQ measure. Utilising data from around 1000 adults representative of the German population, they found that higher cognitive ability is associated with lower risk aversion and less pronounced impatience.[16]

In the first session of our experiment we measured IQ using the number of correct answers from a 20-min version of the Raven's Advanced Progressive Matrices, and the mean for this variable is given in Table 2.[17] The test is a widely accepted indicator of higher order general mental ability that does not rely on cultural context or prior experience (see for example Bors and Stokes, 1998; Arthur et al., 1999, and references therein). Originally designed to provide information about a subjects' ability using a non-verbal setting uncontaminated by linguistic background, its results have been shown to be consistent across cultures and over time (Raven et al., 1998).

Our subjects come from a variety of different regions, as Table 2 shows. These countries are characterised by different cultural norms and levels of gender emancipation. We therefore included among the controls the World Economic Forum's

---

[11] Assuming a constant relative risk aversion utility function of the type $u(x) = x^{1-\sigma}/(1-\sigma)$, where σ is the degree of relative risk aversion, we calculate that the value of σ making an individual just indifferent between choosing the lottery and the certain outcome lies between 0.3 and 0.4. This is similar to the range of 0.3–0.5 found by Holt and Laury (2002), and the range of 0.43 and 0.48 found by Dohmen et al. (2010).

[12] A referee suggested that payouts from the first session might affect risk behaviour in the second session. We checked our payments across class-types, and no significant differences between payments across single-sex classes and coed classes. While men in single-sex classes are slightly more likely to choose risky options than are women in single-sex classes, this is because men in session 1 are more risk taking in general. However, our randomisation payment process (described in footnote 10) meant this did not transfer into significantly higher payments.

[13] The Kolmogorov–Smirnov test for the difference in the distributions between single-sex and coed classes suggests that we cannot reject the hypothesis that both distributions are equal (the p-value is above 0.05 – i.e., 0.27 for girls and 0.93 for boys).

[14] To control for whether or not our experimental subjects were exposed to single-sex high-schooling before entering university, we experiment with using in some specifications a dummy variable for attending a coed secondary school (where the base was single-sex schooling). These estimates will be discussed later in the paper. Since our experiment includes students from a number of different countries in which single-sex schools were predominantly private, our single-sex secondary schooling will be picking up the influence of both private education *and* single-sex schooling. Therefore our results are not directly comparable to those of Booth and Nolen (2012b), whose subject pool comprised *only* students from government-financed secondary schools.

[15] These authors collected three measures of cognitive skills: a nonverbal IQ test (Raven's matrices), a test of the ability to plan (the 'Hit 15' task), and a quantitative literacy (or numeracy) test. In their analysis they reported only the impact of IQ (Raven's matrices) but they note that their results were robust to using the common factor obtained from a factor analysis of all three measures.

[16] Dohmen et al. (2010) used two tests of cognitive ability that were similar to certain modules of the Wechsler Adult Intelligence Scale (WAIS). One was the symbol correspondence test, which asks subjects to match as many numbers and symbols as possible in 90 s according to a given correspondence. The other, the word fluency test, gives subjects a timed vocabulary test.

[17] Our distribution of score is skewed to the left. At the 99% confidence level we reject the null hypothesis that skewness and kurtosis are zero (i.e. the distribution of the values is normal). In fact, for within-age scores of this type of test, the distribution of the results have been found not to be Gaussian (Raven, 2000). Since normality is not a requirement to standardise the results, we use the z-score of the test as our preferred transformation for the cognitive ability test to facilitate the analysis.

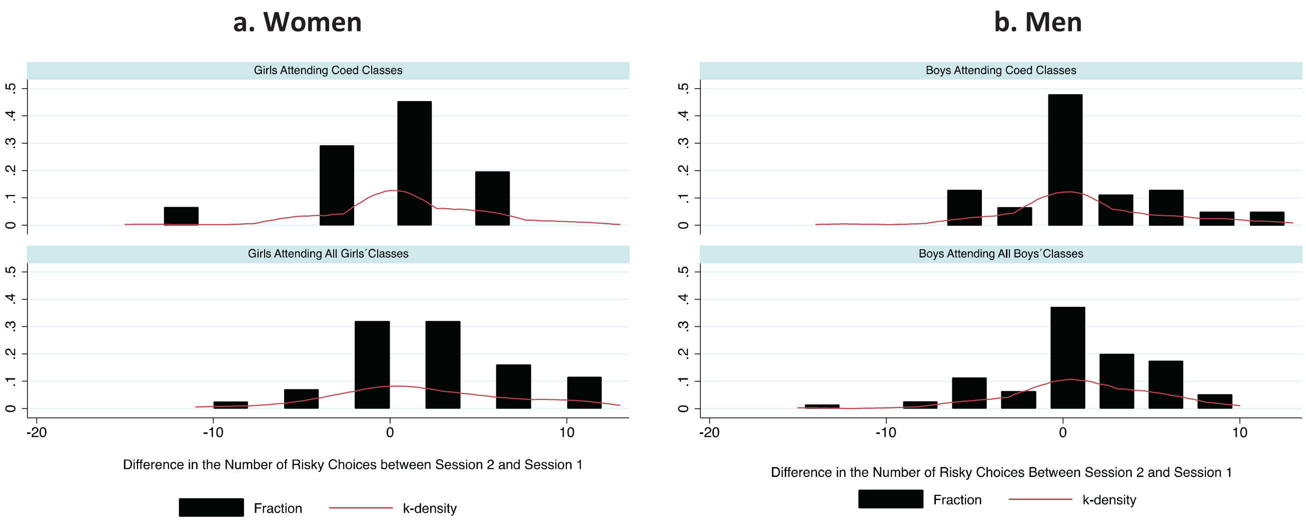

Fig. 1. Difference in the number of risky choices across sessions by class-type and gender. *Notes*: The vertical axis in the figures shows the frequency of students while the horizontal axis refers to the difference in the number of risky choices made between the sessions. The differences in means suggests a higher variation in favour of women attending all women' classes.

Table 2
Difference in Means between students attending only the first session and students attending both sessions.

	Females				Males			
	In Session 1 ONLY (N = 75)[a]	In Both Sessions (N = 75)	Difference in Means (2) – (1) :	Std. Dev. of (3)	In Session 1 ONLY (N = 135)	In Both Sessions (N = 144)	Difference in Means (6) – (5):	Std. Dev. of (7)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Risky choices Session 1	10.37	10.13	–0.24	(0.68)	11.84	11.57	–0.27	(0.47)
Female	1.00	1.00	0.00	(0.00)	0.00	0.00	0.00	(0.00)
Raw score IQ test	12.00	11.60	–0.40	(0.48)	11.63	11.81	0.18	(0.40)
Z-score IQ test	0.08	–0.04	–0.13	(0.15)	–0.03	0.02	0.06	(0.13)
Below the age of 20	0.35	0.33	–0.01	(0.08)	0.38	0.32	–0.06	(0.06)
Economics degree	0.41	0.47	0.05	(0.08)	0.51	0.40	–0.11*	(0.06)
Coeducational school	0.64	0.55	–0.09	(0.08)	0.65	0.66	0.01	(0.06)
Female teacher	0.32	0.28	–0.04	(0.08)	0.33	0.32	–0.01	(0.06)
Single sex class	0.59	0.68	0.09	(0.08)	0.56	0.43	–0.13**	(0.06)
Africa	0.04	0.07	0.03	(0.04)	0.04	0.07	0.03	(0.03)
English speaking	0.16	0.37	0.21***	(0.07)	0.40	0.41	0.01	(0.06)
Asia	0.08	0.03	–0.05	(0.04)	0.05	0.07	0.03	(0.03)
Asia-Europe	0.04	0.01	–0.03	(0.03)	0.02	0.01	–0.01	(0.01)
China	0.21	0.08	–0.13**	(0.06)	0.15	0.15	–0.00	(0.04)
East-Europe	0.27	0.23	–0.04	(0.07)	0.17	0.15	–0.03	(0.04)
Europe	0.20	0.21	0.01	(0.07)	0.17	0.14	–0.03	(0.04)
Global Gender Gap Index 2010	0.69	0.71	0.02**	(0.01)	0.71	0.71	–0.00	(0.01)
1. Agreeableness	11.89	12.20	0.31	(0.59)	11.77	11.81	0.04	(0.48)
2. Conscientiousness	12.59	12.51	–0.08	(0.69)	12.78	12.67	–0.12	(0.52)
3. Extraversion	13.03	12.84	–0.19	(0.64)	12.61	12.24	–0.37	(0.49)
4. Neuroticism	12.25	12.08	–0.17	(0.67)	11.09	10.74	–0.35	(0.45)
5. Openness	13.40	13.51	0.11	(0.77)	13.33	13.61	0.29	(0.59)

Notes: (a) The number of females who did not participate in session 2 and the number of females who participate in both sessions are 75 and 76, respectively. However, the number of observations without missing information in other controls comprises the sample of females attending both sessions to 75. Hence the number coincides in the table as shown. *, ** and ***, corresponds to 10%, 5% and 1% levels of significance, respectively. The sample compared here corresponds to those students without missing answers in the risk experiment during the respective sessions analyzed.

gender gap index (GGI).[18] Our goal here was not only to see if risky choices are affected by the 'cultural' factors encapsulated in that index, but also to see if the coefficients to our variables of interest alter when we control for gender-relevant cultural factors.[19] As part of the sensitivity analysis we use a set of dummy variables grouping different nationalities by region-of-origin to control for language and other culture differences. These are Asia (excluding China), Africa, China, Europe, Eastern Europe, and the Asia-Europe boundary, with the English-speaking countries forming the base.[20]

3. Are there attrition issues?

As noted above, students filled in a risk questionnaire as part of a paid experiment during their first class, and eight weeks later they completed a second risk questionnaire as part of a paid experiment during their class that week. To measure the initial effect of the class assignment on risk attitudes and its effect over time, we focus on the subset of students taking part in both experiments. We have responses for 476 students turning up for the first lecture and class, signing a consent form and filling the demographic questionnaire as well as the IQ test. Of these, there were 424 with usable responses for crucial variables, and we use these when estimating the attrition probability (where the dependent variable takes the value one if the student attended sessions 1 and 2, and zero if s/he attended only session 1). (These estimates are presented in Appendix Table B.2, and will be discussed below.) We have usable responses for 219 individuals who participated in *both* experimental sessions.[21]

Clearly attrition would matter to our econometric analysis if individuals participating in the second session differed in their risk preferences from those who did not. Attrition would also matter if the two samples differed in terms of mean characteristics of the other crucial control variables. For instance, if women assigned to all-female groups were systematically more likely to attend the second session, this might affect our results of interest.

We conducted a number of analyses to explore the attrition issue. Since these show that attrition is not a problem, in the interests of space we confine details to Appendix B although we briefly summarise our strategy here. First, we compared the means of all the variables for each of our three class-types: all-female, all-male, and coeducational. This comparison is reported in Table B.1 of appendix. Columns (3), (7) and (11) of Table B.1 show the difference in means, for each class-type, between the initial sample and the subsample of students turning up for both experiments. The difference between the means is not statistically significant for any class type.

To further address the attrition issue, we then estimated a model of the probability of attending the second session conditional on characteristics observed in the first session and with controls for GTA fixed effects.[22] The estimates are presented in Appendix Table B.2. The first specification disaggregates single-sex classes into all-female and all-male, and the second specification combines the two into a 'single-sex class' dummy variable. The estimates show that risk preferences observed in session 1 did not affect the session 2 participation probability for either specification: the coefficient to the number of risky choices in session 1 is insignificantly different from zero. We also found that the attendance probability in session 2 was significantly increasing in IQ score and decreasing in the Global Gender Gap Index (at the 10 percent level in both cases). While in the first specification these were the only statistically significant variables, in the second specification, the coefficient to female gender was positive and significant at the 5 percent level. (This reflected the fact that the single-sex class-types were not disaggregated into all-female and all-male classes.) We shall be including controls for these variables in our subsequent regressions.

The important point to note here is that individuals who had made a lot of risky choices in session 1 were as likely to turn up to session 2 as individuals who had made very few risky choices. In other words, people participating in the second session do not differ in their risk preferences from those who did not participate.

Third, we ran an OLS regression at the *class-level* of the determinants of the proportion of all students failing to attend the second experiment (the results are shown in appendix in Table B.3). Our controls included teacher gender, mean IQ score in that class, whether the class was all-female or all-male (the base was coeducational), an indicator if the course was for an economics rather than business degree, day of the week (only four days were assigned to classes), and indicators for early or late in the day. There was no statistically significant correlation between single-sex class (either all-female or all-male) and attendance at the second experiment. Thus it seems that attrition was not specific to particular class types.

---

[18] The GGI ranks national gender gaps in economic, political, educational and health aspects. The values for the index go from 0 to 1. Hence a country with a high index is a country with smaller gender gaps in the access to the country' resources regardless the level of resources.

[19] Using this index, Guiso et al. (2008) found, in a cross-country analysis using PISA data, that girls' underperformance in math relative to boys is eliminated in more gender-equal societies. In contrast, Fryer and Levitt (2010) show, using the bigger sample of countries from the TIMMS data that the effect of the GGI is not statistically significant, and that in countries like Iran with high gender inequality girls actually out-performed boys.

[20] English-speaking countries correspond to the UK and North America. Asia (excluding China) also includes countries from the Middle East. Eastern Europe includes Kosovo, Latvia, Lithuania, Poland, Slovakia, Ukraine, Albania, Bulgaria and Romania. The Asia-Europe boundary includes Azerbaijan, Kazakhstan and Russia.

[21] We restricted our sample to those individuals without missing answers from the risk questionnaires in both sessions and valid information for the other variables of interest. Out of the 240 students who attended both experiments and attended the first lecture, 219 completed the whole questionnaire.

[22] High absence rates are common in all undergraduate classes in the UK, probably because students are able to download lecture and class materials from the course materials website. However, there is always a core majority of students who attend lectures and classes regularly.

An additional way of exploring whether or not attrition is likely to affect our risk results is to compare means across the two mutually exclusive sample groups – those showing up *only* in session 1, compared with those turning up at *both* sessions. This comparison, to be discussed in Section 4.1 below, confirms our conclusion that selection is not a problem for our analysis.

4. The results

In this section we first discuss the means of our variables, then estimate the determinants of individuals' class-attendance in the first eight weeks. This is followed by the cross-sectional estimates of the determinants of the number of risky choices, and finally by the estimates exploiting the panel nature of the data.

4.1. Descriptive statistics

As noted above, our subjects are first year students enrolling in Economics 100 or 111 at the University of Essex. All students were randomly assigned to a single sex or coeducational class by the central timetabling office at the university. Table 2 reports the means of the variables for two different subsamples, each disaggregated by gender. The two *mutually exclusive* subsamples are individuals (i) *only* participating in the first session, and (ii) participating in both the 1st *and* 2nd session. That the same number (75) of women appear in both subsamples is a coincidence, as the note at the bottom of the table explains. The number is the same but their identities differ. Summing across men and women participating in *both* sessions produces the 219 observations used in our estimation, and 34% of this number is female.

For the sample of 219 students participating in both sessions, we had 16 GTAs teaching 32 classes. Of those 16 GTAs, 5 were female. By design no GTA was perfectly correlated with the treatment. In our GTA allocation we also aimed to ensure, as much as possible, a teacher gender balance across class-types. The five female GTAs for our sample together taught a total of ten classes. Only one female GTA taught all-female classes (and she also taught coeducational classes), two were teaching all-male classes, and four were teaching coeducational classes. This compares with a total of eleven male GTAs teaching a total of 22 classes. Four male GTAs taught all-female classes, six taught all-male classes and six taught coed classes. (In our sensitivity analysis in Section 5.6, we control for potential GTA effects in a number of different ways, to see if the results still hold.)

Inspection of the means in Table 2 reveals that there are very few statistically significant differences in means for the subsamples of concern. First, we compare women who turned up only to session 1, and those turning up to both sessions. English-speaking women were more likely to be found in the participating subsample, and those from China were less likely to. The difference in means in the Global Gender Index also reflects these differences (recall this index is based on country-level data). There are also differences in the variable 'missing data for Extraversion'. Now compare the means for men. Here there are statistically significant differences in means only for single-sex class and for degree type. Around 56% of men in single-sex classes and 51% of men doing an economics degree attended only session 1, and 43% of men in single-sex classes and 40% of men doing an economics degree attended both sessions.[23] This suggests that it is important to include these controls in our regressions.

4.2. Estimates of class attendance across the eight weeks

Our data contain information about individuals' attendance at the intermediate classes between the first and eighth sessions. If class-type is significantly correlated with attendance through the term, then any estimated impact of class-type on risky-choices may be picking up an effect of differential class-attendance. In order to investigate this, we estimated an OLS regression of the number of weeks that an individual attended the weekly class in economics, with controls including gender, class-type and degree (economics or business) (see Table B.5). The sample size is slightly smaller (214 observations rather than 219), since attendance information between sessions 1 and 8 was unavailable for a handful of students. The first column reports results *without* GTA fixed effects, while the second column includes these. (The base for all our estimates in which we include GTA fixed effects was the GTA who had the most classes. This GTA was male and he taught coeducational and all-male classes only.)

In specification [1], the estimated coefficient to the constant was 5.34. However, in specification [2] with GTA fixed effects it was 4.85, indicating that on average students in our sample attended roughly 5 out of the eight classes, a respectable number. At the University of Essex, the gender balance within economics and business undergraduate degrees is roughly two-thirds men to one-third women. It is possible that the higher attendance rates for women in single-sex classes arose because women felt more comfortable in this environment within the predominately male faculty, or relatedly, because

[23] Readers might be interested to know that our students are from a number of different countries: 31% of our subjects were from English-speaking regions, 20% were from Eastern Europe, 18% from the rest of continental Europe, 18% from China, 6% from other Asia, 4% from Africa and 3% from Asia-Europe. It is interesting to note that all students were distributed roughly equally across personality types, and that the mean for the global gender gap index – which we will use as a proxy for cultural gender background factors – is 0.70.

Table 3
OLS estimates of the determinants of the number of risky choices.

	Number of risky choices made in session 1					Number of risky choices made in session 2				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	−1.71**	−1.69**	−1.48	−1.52	−1.38	−1.17*	−1.18*	−2.02**	−2.06**	−1.90*
	(0.68)	(0.69)	(0.92)	(0.96)	(0.89)	(0.59)	(0.59)	(0.91)	(0.90)	(0.95)
Economics degree		1.28***	0.34	0.60	0.31		−0.74	−1.27	-0.99	−1.31
		(0.37)	(0.69)	(0.67)	(0.69)		(0.89)	(0.85)	(0.82)	(0.84)
All-Male Class			1.21	1.31	1.27			0.60	0.70	0.67
			(0.85)	(0.89)	(0.87)			(0.71)	(0.76)	(0.72)
All-Female Class			1.05	1.14	1.03			2.95***	3.04***	2.93***
			(1.15)	(1.22)	(1.19)			(0.96)	(1.02)	(1.01)
Below the age of 20			−0.12	−0.13	−0.17			−0.69	−0.71	−0.74
			(0.58)	(0.58)	(0.61)			(0.60)	(0.61)	(0.63)
Z-score IQ test				0.48**					0.50*	
				(0.20)					(0.25)	
Global Gender Gap Index 2010					5.15					5.87
					(7.47)					(8.65)
Constant	10.80***	9.74***	9.74***	9.63***	6.09	12.42***	13.04***	13.48***	13.36***	9.31
	(0.36)	(0.14)	(0.36)	(0.38)	(5.17)	(0.78)	(0.12)	(0.36)	(0.37)	(6.15)
GTA's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	*219*	*219*	*219*	*219*	*219*	*219*	*219*	*219*	*219*	*219*
R^2	*0.140*	*0.142*	*0.150*	*0.164*	*0.152*	*0.081*	*0.081*	*0.103*	*0.116*	*0.106*

Notes: Dummies for each GTA were included in the analysis. We left out the GTA with the most students in our estimating sample. Robust standard errors clustered by module and class are reported in parenthesis.
* Corresponds to 10% levels of significance.
** Corresponds to 5% levels of significance.
*** Corresponds to 1% levels of significance.

that class-type facilitated the formation of same-sex friendships that encouraged attendance. We will return to attendance issues when discussing robustness in Section 5.

4.3. Cross-sectional estimates of number of risky choices in sessions 1 and 2

Table 3 presents the ordinary least squares (OLS) estimates of the determinants of the number of risky choices (which can range between 0 and 20). These are estimated separately for sessions 1 and 2 and for five different specifications in each case. Columns [1]–[5] report the session 1 estimates, while columns [6]–[10] show the results for session 2. All specifications include 15 dummy variables, one for each GTA. (The base was the GTA who had the most classes. This GTA was male and he taught only coeducational and all-male classes.) We term these *GTA fixed effects* hereafter.[24] Specification (1) includes dummy variables for gender while Specification (2) adds degree type. Even though all our students were taking the introductory economics course, two different formats of the course, EC111 and EC100, were available depending on the degree subject studied (either economics or business-related). The course for students majoring in economics has a higher math component than the course for students majoring in business and related areas (e.g. accounting, finance). We therefore control for degree subject in our estimation.

Specification (3) augments Specification (2) by adding dummy variables for class-type (the base is coeducational class) and an indicator variable for being younger than 20 years old. Specification (4) augments specification (3) by including a measure for cognitive ability and finally, specification (5) replaces this by the global gender gap index. (We also experimented with a number of expanded specifications not reported here, but which we return to in the sensitivity analyses of Section 5.)

The estimated coefficients from the initial session are given in the first five columns, headed *Session 1*. Robust standard errors (SEs), clustered at the module and class level, are reported in parentheses. The specification in column [2] displays the estimates for being female and for being enrolled in an economics degree. Here the base male, enrolled in a business degree and with the base GTA, chooses 9.74 risky rows, while a women enrolled in a business degree and with the base GTA chooses 9.74 − 1.69 = 8.05 rows. Thus women are significantly less likely than men to take risks.

Now consider the session 1 estimates from the expanded specifications in columns [3] and [4]. The randomly assigned class-types have no statistically significant correlation with the number of risky choices. This is as expected, given the

[24] Our results are also robust to *exclusion* of GTA fixed effects. Full estimates are available in our CEPR Discussion Paper, and also in Appendix C of this paper, where in Tables C.1 and C.2 we present a sensitivity analysis with and without GTA fixed effects for the reader's convenience.

random allocation and the fact that the students have not yet spent any time in their treatment group. The specification that includes our measure of cognitive ability is given in column [4]. While higher ability students do make more risky choices, the magnitude is quite small. The estimated coefficient, statistically significant at the 5% level, suggests that a one SD increase in cognitive ability is associated with an increase of less than half an extra risky choice in the number of risky choices made. Notice also that in specifications [3]–[5], the impact of the female dummy variable is statistically insignificant.

We now turn to the results from session 2, when the risky choices experiment was repeated for the same set of individuals eight weeks later. These estimates are presented in Table 3 in the last five columns, headed *Session 2*. In the intervening weeks between Sessions 1 and 2, students have been attending the same classes for the economics course, as no switching was permitted, and so they have had a number of weeks' exposure to the sex-composition of the group to which they were randomly assigned. Here we see that females are significantly less likely to make risky choices, but the last three columns show that the estimated coefficient to being in an all-female group is large and statistically significant at the five and ten percent level, respectively.

To illustrate how men and women compare in their *session 2 risky choices*, using the results from column [8], we first consider the base-case person, a man doing a business degree and randomly assigned to a co-ed class. For such an individual only the constant applies, and his number of risky choices is 13.48. How does this compare with the outcome for an otherwise identical woman? Subtraction of the gender coefficient from the constant reveals that she will pick 13.48 − 2.02 = 11.46 risky rows. However, women in all-female classes are making more risky choices than this, as the estimated coefficient to all-female classes reveals. Moreover women in all-female classes are not only making more risky choices in session 2 than their coed counterparts, but they are also making slightly more risky choices than men – regardless of whether the men are in coed or all-male classes.

Some of the other variables are also statistically significant. Higher ability students continue to make more risky choices in session 2, as shown in column [9]. Again the magnitude is quite small, just as it was in session 1 (see column [4]).

In summary our results in Table 3 suggest that on average women are less likely to make risky choices than men at both dates. However, after eight weeks in a single-sex environment, women are significantly *more* likely to choose the lottery than their counterparts in coeducational groups. These results are robust to the inclusion or exclusion of controls for personality type, as well as a number of other robustness checks, as we shall see in Section 5.

Booth and Nolen (2012b) showed that single-sex schooling for girls had a significantly positive impact on willingness to take risks of 14 year-olds in publicly funded schools. Given this earlier finding and the focus of our present paper, we also experimented with including a dummy variable 'coeducational secondary schooling' as a control in all the regressions reported in Tables 3 and 4, where the base is single-sex schooling. (A separate appendix, available from the authors on request, reports its estimated coefficient and standard error.) In all cases, this variable was statistically insignificant. Since our experiment in the present paper includes students from a number of different countries in which single-sex schools were predominantly private, the base for secondary schooling will be picking up the influence of both private education *and* single-sex schooling. Therefore our results are not directly comparable to those of Booth and Nolen (2012b), whose subject pool comprised *only* students from government-financed secondary schools in Britain. In other words, we are not here able to disentangle the private schooling effect from the single-sex schooling effect using our data; moreover, one may be positively associated with risk attitudes, the other negatively, and this may be why we find no effect of the combined variable in the current experiment.

4.4. *Fixed effects estimates of number of risky choices*

We estimated two additional models exploiting the panel nature of the data, and these are reported in Table 4. The first estimates the determinants of the number of risky choices in session 2 including a variable that measures the number of session 1 risky choices. The second additional model is where we difference the number of risky choices made in each session.

Columns [1]–[5] of Table 4 report the results of models that explain the number of risky choices made in session 2 by the number of risky choices chosen in session 1 and other variables. They represent our overall preferred regressions. Column [1] includes only gender and the number of risky choices made in session 1, plus the GTA dummy variables, while column [2] augments this with economics degree and class type. Column [3] then adds the student's age (with the base 20 years of age or above). Column [4] adds normalised IQ test results while column [5] replaces this by the Gender Gap Index. It is interesting that the estimates for class type alter little across these different specifications. This suggests that the main result – that young women attending single-sex classes increased their number of risky choices in session 2 – is robust to the inclusion of a further control, namely the number of risky choices made in session 1.

It is also interesting to compare the magnitude of the estimated coefficient to the variable measuring the session 1 number of risky choices. This varies from 0.36 in column [4] to 0.38 in columns [1]–[3]. If the estimated coefficient to this variable were close to one, we might conclude that the model should be better estimated in first-differenced form. While this is clearly not the case, for the reader's interest we nonetheless undertake this additional exercise, as explained below. However, we do not view that as the preferred model.

Table 4
OLS estimates of the determinants of the number of risky choices.

	Number of risky choices made in session 2 explained by those made in session 1					Difference in the number of risky choices made between sessions				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	−0.52 (0.71)	−1.29 (1.01)	−1.46 (1.06)	−1.51 (1.05)	−1.39 (1.09)	0.55 (0.87)	0.51 (0.88)	−0.54 (1.36)	−0.54 (1.37)	−0.52 (1.36)
Risky choices made in Session 1	0.38*** (0.08)	0.38*** (0.08)	0.38*** (0.08)	0.36*** (0.08)	0.37*** (0.08)					
Economics degree		−1.49** (0.73)	−1.39* (0.74)	−1.21 (0.73)	−1.42* (0.74)		−2.02*** (0.57)	−1.61** (0.72)	−1.59** (0.72)	−1.61** (0.73)
All-Male class		0.22 (0.63)	0.14 (0.66)	0.22 (0.68)	0.19 (0.66)			−0.61 (0.88)	−0.61 (0.88)	−0.61 (0.88)
All-Female class		2.35*** (0.74)	2.55*** (0.81)	2.62*** (0.85)	2.54*** (0.84)			1.89* (0.94)	1.90* (0.96)	1.89* (0.94)
Below the age of 20			−0.64 (0.60)	−0.66 (0.61)	−0.68 (0.62)			−0.57 (0.76)	−0.57 (0.76)	−0.58 (0.77)
Z-score IQ test				0.33 (0.22)					0.03 (0.24)	
Global Gender Gap Index 2010					3.95 (6.51)					0.72 (5.10)
Constant	8.31*** (1.22)	9.51*** (0.91)	9.82*** (1.01)	9.85*** (1.01)	7.04 (4.91)	1.61** (0.66)	3.30*** (0.18)	3.74*** (0.52)	3.73*** (0.52)	3.23 (3.61)
GTA's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	*219*	*219*	*219*	*219*	*219*	*219*	*219*	*219*	*219*	*219*
R^2	*0.188*	*0.200*	*0.205*	*0.210*	*0.206*	*0.073*	*0.073*	*0.085*	*0.085*	*0.085*

Notes: Dummies for each GTA were included in the analysis. We left out the GTA with more students in our estimating sample. Robust standard errors clustered by module and class are reported in parenthesis.
* Corresponds to 10% levels of significance.
** Corresponds to 5% levels of significance.
*** Corresponds to 1% levels of significance.

Columns [6]–[10] of Table 4 present the estimates from a simple differenced model, following the approach outlined below. Let y_{it} denote the number of risky choices made by individual i at time t, and suppose for notational convenience we have just 2 explanatory variables, both measured at time $t = 1$ (as in our dataset). Then at period 2 we have:

$$Y_{i2} = a_2 + b_2 X_1 + c_2 Z_1 + v + e_2 \tag{1}$$

while at period 1 we have:

$$Y_{i1} = a_1 + b_1 X_1 + c_1 Z_1 + v + e_1 \tag{2}$$

When we difference these two equations we obtain:

$$(Y_{i2} - Y_{i1}) = (a_2 - a_1) + (b_2 - b_1)X_1 + (c_2 - c_1)Z_1 + (e_2 - e_1) \tag{3}$$

As an illustration, we can see that an estimated positive coefficient to $(b_2 - b_1)$ indicates that the impact of X_1 is growing over time, while an estimated negative coefficient suggests that the impact is declining over time.

We now turn to the estimates obtained from variants of this model, which are provided in columns [6]–[10] of Table 4. The negative (and statistically insignificant) estimated coefficient to female gender of around −0.5 in columns [8]–[10] indicates that the female effect is declining over time. However, the coefficient is not precisely estimated. Now consider the main variables in which we are interested, the class-types. Notice that the effect of being in an all-female class is increasing over time and is, in all cases, statistically significant. Across columns, the estimated coefficient is around 1.9, indicating that women allocated to all-female classes increase the number of risky choices by over two, relative to the base. The coefficient to being a man in an all-male class is decreasing, albeit the magnitude is small at −0.61, and is very imprecisely estimated.

5. Sensitivity analysis

We next report the results from a number of robustness checks. These checks involved adding in personality measures; experimenting with different estimation methods; experimenting with different ways of modelling the dependent variable; experimenting with different transformations of cognitive-ability score; and replacing the World Economic Forum's gender gap index (GGI) by the country of origin dummy variables. This last exercise aims to see if risky choices are affected by

Table 5
Personality measures: the big five (five factor model).

Individual	Dimension
Is rude to others Has a forgive nature Is considerate and kind	Agreeableness
Is talkative Is outgoing sociable Is reserved	Extraversion
Worries a lot Get nervous easily Is relaxed	Neuroticism
Does a thorough job Tends to be lazy Does things efficiently	Conscientiousness
Is original, with new ideas Values the art Has an active imagination	Openness

language and other cultural factors. We also investigated if an individual's IQ test-taking strategy (accuracy) is correlated with risk-aversion and with gender. (The estimates from these sensitivity tests are presented in Appendix C.) Finally, we experimented with different ways of controlling for GTA fixed effects.

5.1. Are the results sensitive to the inclusion of class attendance and personality?

Tables 3 and 4 are extended to include an additional control variable – class attendance and personality type – in Tables C.1 and C.2, respectively. Here we present results with and without GTA fixed effects, for comparison. Thus columns [3] to [5] contain the specifications *with* GTA fixed effects, as do columns [8] to [10]. Note that the sample size here is slightly smaller (214 observations rather than 219), since attendance information between sessions 1 and 8 was unavailable for a handful of students. The estimates show that higher class attendance is associated with fewer risky choices, and this effect is statistically significant for all specifications of the number of session 2 risky choices. Importantly, the coefficient to all-female classes in session 2 remains statistically significant at the one percent level, and its magnitude remains large. This suggests that the slightly higher attendances for all female classes found in Table B.5 is not explaining the large and statistically significant effect of all-female classes on the number of risky choices in session 2.

Following the psychological literature (see for example Heineck, 2007, and references therein), we included in the demographic questionnaire a set of 15 questions regarding personality traits, using the same format found in the British household panel survey (BHPS). Personality traits can be linked to individual characteristics in what is called the five factor model (FFM). Under this scheme, each personality dimension is related to one of the following five characteristics: openness to experience, conscientiousness, extraversion, agreeableness and neuroticism. The questions are answered on a scale from 1 to 7 and the final scores obtained by adding the scores within each specific dimension (Heineck, 2007). The components of each dimension are summarised in Table 5.

Estimates in the last four columns of each panel of Tables C.1 and C.2 include the score of each of the five dimensions to control for personality traits. To avoid excluding individuals who did not answer these questions, we also control for missing responses. The literature suggests that personality could affect the way individuals take decisions. People who are more open to different experiences may be less conservative and consequently might take more risks, while individuals who pay more attention to detail or who are more precise could make fewer risky choices. Results from Tables C.1 and C.2 suggest that in fact individuals with a high score in the trait "openness" are associated with a higher number or risky choices. Similarly, individuals with higher score in the "conscientiousness" component reduce the number of risky choices made in each session. This was the case when using as dependent variable the number of risky choices made in session 1 and session 2, and is what we would expect. For the specification that uses the difference in the number of risky choices between both sessions, the personality traits were not found to be statistically significant (see last panel of Table C.2).

Regarding the coefficient of being in all-female classes and all-male classes when the personality traits are included, we find our main conclusions observed in Tables 3 and 4 are reinforced. Indeed, the GTA fixed effects estimation in columns [5] and [10] shows that when attendance and personality type is included the magnitude and precision of the coefficient to all-female class increase.

5.2. Are the results sensitive to estimation method?

We next conducted an interval regression analysis to take into account the presence of salient numbers in the stakes (£5 pounds in row 6, £10 pounds in row 11, £15 pounds in row 16) that could make the students switch in that row. Use of an interval as the dependent variable instead of the number of risky choices does not overturn our results, as inspection of

Table C.3 reveals.[25] The same set of variables remain statistical significant. (Again we include estimates without and without GTA fixed effects, for the referees' comparison. Since our results in Tables 3 and 4 are robust to estimation method, we do not refer to this in any further detail here.)

5.3. Are the results sensitive to different ways of modelling the dependent variable?

To account for multiple switching,[26] we now average the rows corresponding to the first and last changes between the column for the lottery option and the column for the safe option.[27] Since the range where the individual switches back and forth correspond to an indifference situation (Harrison et al., 2007), the average row refers to the certain equivalent for those individuals who are "multiple switchers". When using the average row as the new dependent variable, the main conclusion still holds for the estimates reported in Tables C.4 and C.5. Moreover these results do not differ from the estimation using the number of risky choices (baseline) but restricting the analysis to multiple switchers.[28]

We also experimented with using the row in which the individual changed as a measure of risk for those individuals who switched from column B to column A only once ("single switchers"). The estimates are reported in columns [1] and [2] of Table C.4. As before, the results are comparable when using the number of risky choices for this "one switch" subsample. The estimates from this procedure are very similar to those presented in Table C.3. Once again, the results support our main conclusions.

5.4. Are the results sensitive to different transformations of cognitive ability?

How are our results affected by inclusion of different transformations of IQ score? The estimates using different scales for the IQ-test are shown in Tables C.6 (raw IQ score) and C.7 (a population norm standardisation) in Appendix C. (Note that we now revert to presenting all our results as GTA fixed effects, in the interests of space.) The statistical significance of the coefficients of interest does not change and our main conclusions remain the same. Importantly, the estimated effects of all-female classes and of gender are very similar to the previous analysis.

5.5. Do culturally driven norms and beliefs affect women's risky choices?

In Tables C.1, C.2, 3 and 4, we included the global gender gap index in the preferred specifications considered for the analysis (the number of risky choices in sessions 1 and 2 separately, and the difference in the number of risky choices made between both sessions). The estimates suggest that the GGI was statistically insignificant and that the coefficients to our other variables of interest were robust to the inclusion of the GGI. While the index itself does not have a statistically significant effect, once we interact it with gender (columns [3] and [6] of Tables C.8 and C.9) we find that women from countries with more 'gender equality' are associated with a greater number of risky choices.

Now we replace the global gender gap index in our preferred specifications (Tables 3 and 4) by the region-of-origin dummy variables as shown in Tables C.8 and C.9 (columns [1] and [4]) in appendix. It can be seen that in many cases the region of origin variables have a statistically significant effect on the number of risky choices. However, the impact of the all-female classes remains robust to this different way of controlling for cultural background. In our main specifications we retain the global gender gap index in our preferred regressions, in the interests of maximising degrees of freedom (recall we already have 15 GTA dummies).[29]

5.6. Is IQ test-taking strategy correlated with risk-aversion and with gender?

As Dohmen et al. (2010) note, risk-averse individuals might take longer to complete the test of cognitive ability because their risk aversion translates into a desire to avoid mistakes and achieve greater accuracy. The conjecture is thus that accuracy is increasing in risk aversion. Therefore there might be a negative correlation between accuracy of responses in the test of cognitive ability and the number of rows of risky choices.

We investigated if individuals' accuracy in completing the Raven's matrices in wave 1 is correlated with Session 1 risk-aversion and with gender. We defined accuracy as the number of correct answers divided by the total number of answers completed by each individual. The mean of the sample's accuracy rate is 0.67 and its standard deviation is 0.18. Using the accuracy rate as the dependent variable, we estimated a model in which controls included gender and the number of risky

[25] Five intervals for the number of risky choices were constructed. They were built around the following rows: (1,4), (5,8), (9,13), (14,17), (18,20).

[26] For the subsample used in the analysis, in session 1, 5% of individuals never switched between columns, 70% switched once and 25% switched more than once. In session 2 the percentage of multiple switchers dropped to 16%. This compares with multiple switchers in the literature of 18% (Maier and Rüger, 2010), 25% (Bruner, 2007) and around 6% in Andersen et al. (2006).

[27] Our analysis is restricted to those students starting with consistent choices, i.e. those starting moving from the lottery option to the safety one.

[28] The main specification uses the number of risky choices as our dependent variable without restricting for the number of changes.

[29] We also experimented with interacting the GGI with the all-female class variable. We divided the GGI into two variables (above the median and below it). Of course cell sizes reduced considerably when we use the interaction of GGI* above the median and all-female class. While the interaction is not significant, the point estimate of the treatment remains statistically significant and positive.

choices, as well as all the other controls used in Table 3 (see Table C.10 of Appendix C). We found that female gender has a statistically significant association with the accuracy rate only before we control for additional variables. In specification [3], the estimated coefficient (SE) to female was 0.04 (0.05) and the coefficient to the number of risky choices was 0.00 (0.00). In specifications [3] and [6], individuals making a greater number of risky choices do not achieve a lower accuracy rate in the IQ-test. Moreover, being female, controlling for risk aversion, does not translate into greater accuracy. We also experimented with including instead the number of risky choices made in session 2, and find similar results. What is also interesting is that men in all-male classes were slightly less accurate in the IQ test.

5.7. Are the results sensitive to GTAs?

Are our findings robust to controlling in a different way for GTAs? In this subsection we first report the results from re-estimation dropping from the estimating sample those students who were affected by a change in GTA, leaving a subsample of 207 students. The second approach was to re-estimate everything dropping one GTA at a time. Below we briefly discuss the results of these procedures.

With the exception of one GTA, the same GTAs were employed throughout the year. To see if changing this one GTA between experimental sessions 1 and 2 had any effect on our results, we reran all the regressions in Tables 3 and 4 on a subsample in which we dropped the students from this particular GTA's classes, leaving 207 students. (These results are reported in an appendix available on request. The results are robust to this re-estimation.) Indeed, the level of statistical significance remains the same for the coefficient to all-female class, and its magnitude increases slightly.

Our next approach was to re-estimate the regressions reported in Tables 3 and 4 leaving out the observations of each GTA separately. Of course, by dropping a GTA from each estimation, we are reducing the number of observations. More importantly, some of the GTAs were teaching single-sex classes (as well as coed), and by taking out their observations our cell sizes for single-sex classes become small. Indeed, we lose between 13% and 25% of the treatment in some cases. In almost all cases, the estimated coefficient to all-female class remains similar in magnitude and statistical significance to the estimates of Tables 3 and 4. The chief exception is that two of the differenced estimates become statistically insignificant. Perhaps this is unsurprising given that the smaller number of cases means they are less precisely estimated. However, the point estimates for the variable of interest in each estimation remain well within one standard deviation of the result using all observations in Tables 3 and 4.

6. Conclusions

To test if single-sex classes within a coeducational environment modify students' risk-taking behaviour, we designed a controlled experiment using first year economics and business students from a British university. The subjects were asked to make choices over real-stakes lotteries at two distinct dates. Prior to the start of the academic year, students were randomly assigned to classes (i.e., experimental groups) of three types: all-female, all-male, and coeducational, and they were not allowed to change group subsequently.

We found that women are less likely to make risky choices than men at both dates. However, after eight weeks in the single-sex class environment – within the larger coeducational milieu – women were significantly *more* likely to choose the lottery than their counterparts in coeducational groups. No such result was found for men in the single-sex groups. Moreover, our results were robust to a number of sensitivity checks. This finding is relevant to the policy debate on the impact of single-sex classes within coed schools or colleges on individuals' behaviour. Whether or not this outcome carries over into other subject areas apart from economics and business remains a topic for future research.

How do our results differ from Booth and Nolen (2012b)?[30] That study was, to our knowledge, the first to examine the issues also studied in the present paper. Clearly it is important to investigate the external validity of those findings. It is also to be expected that subsequent studies using different methods of eliciting risk preferences, different subject pools, and different treatment-group sizes will find slightly differing results. In fact, the difference in results between Booth and Nolen (2012b) and the present paper lies in the finding that, in the former paper, the 4-person experimental group to which students were randomly assigned had an immediate effect for all-female groups, whereas in the current paper it does not have an effect until the (larger) all-female class has been going for some time. We cannot tell if this is due to the different method used to elicit preferences, the different age-group, the different treatment-group sizes, or the different samples. However, we hope that subsequent work will use the methodology in the present paper to check the external validity of our findings.

How might our results be interpreted? They suggest that a part of the observed gender difference in behaviour under uncertainty found in previous studies might actually reflect social learning rather than inherent gender traits. Of course this is not to say that inherent gender traits do not exist. Rather it suggests that they can be

[30] In a different context using a different risk measure, Booth and Nolen (2012c) showed that salience affects young men and women differently. Though we did not report single-sex class effects in the published version of that paper, in preliminary estimates we found that they had no effect on switching behaviour, nor did we expect them to.

modified by the environment in which a woman is placed. Our experiment does not allow us to tease out *why* these behavioural changes were observed for young women in all-female groups. Conjectures as to the reasons for the changes might include a reduction in stereotype effects. Women who are placed in an all-female environment are no longer reminded of their own gender identity and may find it easier to make riskier choices than do women who are placed in a co-ed class. A related hypothesis is that being placed, shortly after arrival at university, in an all-female group – even though it is only for one class a week – facilitates the formation of friendships within a faculty environment that is disproportionately male. These friendships can assist in building the comfort and confidence of the women so placed. They may also facilitate the formation of work-groups or networks, leading these women to feel more comfortable in making risky choices than women in coed classes. We hope that future research will investigate these hypotheses further.

Appendix A. Risk questionnaire

Instructions: Below is a table with 20 rows. In each row there are two choices, one in column A and one in column B. In each row consider your two choices and, when you have decided which option you prefer, you should circle that choice. Pick only one choice for each row; do not discuss your choices with anyone else.

After everyone has chosen their preferred option, the instructor will randomly choose a number 1–20. One out of every 10 people will then get paid for their choice in that row.

For example, if you chose "50% chance of winning £30 and a 50% chance of getting £0" in row 1 and "1" gets randomly chosen then, if you are one of the 10% who get paid, the instructor will flip a coin and you will have a 50% chance of winning £30. If you chose "£0.00 for sure" in row 1, though, you will receive £0.00 if "1" gets randomly chosen and you are one of the 10% who get paid.

Please raise your hand if you have any questions.

	Column A		Column B
(1)	£0.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(2)	£1.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(3)	£2.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(4)	£3.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(5)	£4.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(6)	£5.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(7)	£6.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(8)	£7.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(9)	£8.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(10)	£9.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(11)	£10.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(12)	£11.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(13)	£12.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(14)	£13.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(15)	£14.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(16)	£15.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(17)	£16.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(18)	£17.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(19)	£18.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0
(20)	£19.00 for sure	or	50% chance of winning £30 and a 50% chance of getting £0

Appendix B. Attrition issues

See Tables B.1–B.5.

Table B.1
Difference in Means between the whole sample and the sample under analysis.

	All-male classes				All-female classes				Coeducational classes			
	Whole sample (N = 150) (1)	Sub-Sample (N = 81) (2)	Diff. in Means (2) – (1) (3)	Std. Dev. of (3) (4)	Whole sample (N = 107) (5)	Sub-sample (N = 44) (6)	Diff. in Means (6) – (5) (7)	Std. Dev. of (7) (8)	Whole sample (N = 219) (9)	Sub-sample (N = 94) (10)	Diff. in Means (10) – (9) (11)	St. error of (11) (12)
Risky choices made in session 1	11.93	12.15	−0.22	(0.53)	9.91	10.48	−0.57	(0.81)	11.01	11.04	−0.03	(0.47)
Female	0.00	0.00	0.00	(0.00)	1.00	1.00	0.00	(0.00)	0.29	0.33	−0.04	(0.06)
Raw score IQ test	11.43	11.23	0.19	(0.48)	11.47	11.93	−0.46	(0.51)	12.08	12.13	−0.05	(0.38)
Z-score IQ test	−0.10	−0.16	0.06	(0.15)	−0.08	0.06	−0.15	(0.16)	0.11	0.13	−0.02	(0.12)
Below the age of 20	0.36	0.37	−0.01	(0.07)	0.48	0.50	−0.02	(0.09)	0.26	0.31	−0.04	(0.06)
Economics degree	0.37	0.44	−0.07	(0.07)	0.48	0.45	0.02	(0.09)	0.48	0.51	−0.03	(0.06)
Coeducational school	0.65	0.64	0.01	(0.07)	0.61	0.73	−0.12	(0.09)	0.62	0.61	0.01	(0.06)
Region of origin												
Africa	0.06	0.05	0.01	(0.03)	0.08	0.07	0.02	(0.05)	0.05	0.02	0.03	(0.03)
English speaking	0.37	0.35	0.02	(0.07)	0.32	0.18	0.14*	(0.08)	0.35	0.35	−0.00	(0.06)
Asia	0.07	0.06	0.00	(0.03)	0.04	0.07	−0.03	(0.04)	0.06	0.05	0.01	(0.03)
Asia-Europe	0.01	0.00	0.01	(0.01)	0.01	0.00	0.01	(0.01)	0.03	0.06	−0.04	(0.02)
China	0.19	0.23	−0.04	(0.06)	0.13	0.18	−0.05	(0.06)	0.14	0.12	0.02	(0.04)
East-Europe	0.17	0.17	−0.01	(0.05)	0.21	0.27	−0.06	(0.08)	0.18	0.20	−0.02	(0.05)
Europe	0.14	0.14	0.00	(0.05)	0.21	0.23	−0.02	(0.07)	0.19	0.19	0.00	(0.05)
Global Gender Gap Index 2010	0.71	0.71	0.00	(0.01)	0.70	0.69	0.01	(0.01)	0.71	0.71	−0.00	(0.01)
Personality Dimensions												
1. Agreeableness	11.93	11.81	0.12	(0.51)	11.71	11.73	−0.02	(0.68)	11.83	11.85	−0.02	(0.48)
2. Conscientiousness	12.55	12.68	−0.13	(0.54)	12.26	12.55	−0.28	(0.81)	12.75	12.83	−0.08	(0.53)
3. Extraversion	12.43	12.58	−0.15	(0.53)	12.36	12.43	−0.08	(0.75)	12.77	13.05	−0.28	(0.50)
4. Neuroticism	11.01	11.02	−0.02	(0.49)	11.90	12.00	−0.10	(0.77)	11.26	11.65	−0.39	(0.48)
5. Openness	13.54	13.22	0.32	(0.67)	12.93	13.25	−0.32	(0.87)	13.73	13.51	0.22	(0.59)

Table B.2
Probit estimates of the probability of participating in session 2.

	Specification 1 (1)		Specification 2 (2)	
Risky choices in session 1	−0.00	(0.02)	−0.00	(0.02)
Female	0.37	(0.25)	0.40**	(0.18)
Z-score IQ test	0.13*	(0.07)	0.13*	(0.07)
Below the age of 20	0.04	(0.14)	0.05	(0.14)
Economics degree	−5.38	(194.04)	−5.52	(252.95)
Coeducational school	0.20	(0.15)	0.19	(0.15)
All-Male class	0.15	(0.29)		
All-Female class	0.23	(0.37)		
Single sex class			0.18	(0.27)
Global Gender Gap Index 2010	−2.77*	(1.55)	−2.77*	(1.55)
1. Agreeableness	−0.04	(0.03)	−0.04	(0.03)
2. Conscientiousness	0.00	(0.03)	0.00	(0.03)
3. Extraversion	0.02	(0.03)	0.02	(0.03)
4. Neuroticism	0.04	(0.03)	0.04	(0.03)
5. Openness	−0.04	(0.02)	−0.04	(0.02)
GTA's fixed effects	*Yes*		*Yes*	
Observations	*424*		*424*	

Notes: Column (1) includes the main effect of being in a single sex class. Column (2) use two different variables to differentiate the effect of all-male class from all-female class. Coefficients of the probit specification are reported. Standard errors in parenthesis. We also included dummy variables to control for whether any aspect of the Big Five was missing.
* Corresponds to 10% levels of significance.
** Corresponds to 5% levels of significance.
*** Corresponds to 1% levels of significance.

Table B.3
OLS regression: % student lost in the second session.

	(1) Specification 1	(2) Specification 2
Mean IQ	0.07**	0.06
	(0.03)	(0.05)
Class characteristics		
All-Male Class	−0.06	−0.04
	(0.06)	(0.14)
All-Female Class	0.10	0.03
	(0.09)	(0.18)
Class on Fridays	−0.23**	−0.07
	(0.11)	(0.26)
Class on Thursdays	−0.16*	0.07
	(0.09)	(0.27)
Class on Tuesday	−0.20	−0.14
	(0.12)	(0.26)
Class at 9 am	−0.18*	−0.11
	(0.09)	(0.27)
Class at 10 am	−0.15	−0.12
	(0.15)	(0.34)
Class at 11 am	−0.12	−0.14
	(0.21)	(0.39)

Table B.3 (*Continued*)

	(1) Specification 1	(2) Specification 2
Class at 2 pm	−0.34***	−0.34*
	(0.04)	(0.16)
Class at 3 pm	−0.13	−0.25
	(0.13)	(0.25)
Class at 4 or 5 pm	−0.19*	−0.32
	(0.10)	(0.25)
GTA was changed	0.10	0.12
	(0.12)	(0.21)
Constant	−0.08	0.05
	(0.36)	(0.56)
GTA's fixed effects	*No*	*Yes*
Observations	*37*	*37*

Notes: The category omitted from day's class correspond to Wednesday (There was no classes scheduled on Monday). Classes at 12 pm were in the excluded category from class' time. Robust standard errors are shown in parenthesis.

* Corresponds to 10% levels of significance.

** Corresponds to 5% levels of significance.

*** Corresponds to 1% levels of significance.

Table B.4

Difference in Means between students attending coed classes and single-sex classes.

	Coed classes ($N=94$) (1)	Single-sex classes ($N=125$) (2)	Difference in Means (1) − (2)	Std. Dev. (4)
Risky choices in session 1	11.04	11.56	−0.52	0.53
Risky choices in session 2	11.85	12.90	−1.04*	0.57
Female	0.33	0.35	−0.02	0.07
Raw score IQ test	12.13	11.48	0.65	0.44
Z-score IQ test	0.13	−0.08	0.20	0.14
Below the age of 20	0.31	0.42	−0.11	0.07
Economics degree	0.51	0.45	0.06	0.07
Coeducational school	0.61	0.67	−0.07	0.07
Region of origin				
Africa	0.02	0.06	−0.03	0.03
English speaking	0.35	0.29	0.06	0.06
Asia	0.05	0.06	−0.01	0.03
Asia-Europe	0.06	0.00	0.06***	0.02
China	0.12	0.22	−0.10*	0.05
East-Europe	0.20	0.21	−0.01	0.06
Europe	0.19	0.17	0.02	0.05
Global Gender Gap Index 2010	0.71	0.70	0.01	0.01
Pct. classes attended	0.83	0.84	−0.00	0.03
Personality dimensions				
1. Agreeableness	11.85	11.78	0.07	0.50
2. Conscientiousness	12.83	12.63	0.20	0.56
3. Extraversion	13.05	12.53	0.53	0.53
4. Neuroticism	11.65	11.37	0.28	0.51
5. Openness	13.51	13.23	0.28	0.64
Observations	94	125		

Table B.5
OLS estimates of the determinants of class' attendance.

	Baseline model (1)	Adding GTAs' FE (2)
Female	−0.05	−0.26
	(0.25)	(0.31)
All-Male class	−0.02	−0.39
	(0.25)	(0.29)
All-Female class	0.50**	0.73**
	(0.23)	(0.28)
Economics degree	0.73***	1.80***
	(0.19)	(0.25)
Constant	5.34***	4.85***
	(0.18)	(0.06)
GTA's fixed effects	*No*	*Yes*
Observations	*214*	*214*
R^2	*0.097*	*0.202*

Notes: Robust standard errors clustered by module and class, are reported in parenthesis.
* Corresponds to 10% levels of significance.
** Corresponds to 5% levels of significance.
*** Corresponds to 1% levels of significance.

Appendix C. Sensitivity analysis

See Tables C.1–C.10.

Table C.1
OLS estimates of the determinants of the number of risky choices (including additional controls).

	Number of risky choices made in session 1					Number of risky choices made in session 2				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	−1.26	−1.25	−1.23	−1.37	−0.97	−2.54***	−2.63***	−2.05***	−2.21**	−1.80**
	(0.91)	(0.98)	(1.04)	(1.11)	(1.00)	(0.90)	(0.97)	(0.94)	(0.92)	(0.99)
Economics degree	−0.70	−0.74	0.20	0.76	−0.01	−0.02	−0.09	−1.07	−0.45	−1.28
	(0.59)	(0.60)	(0.76)	(0.75)	(0.75)	(0.55)	(0.50)	(0.99)	(0.91)	(1.01)
Below the age of 20	−0.22	−0.25	−0.18	−0.24	−0.30	−0.79	−0.77	−0.70	−0.76	−0.82
	(0.58)	(0.62)	(0.65)	(0.66)	(0.70)	(0.54)	(0.52)	(0.57)	(0.60)	(0.62)
All-Male class	0.80	0.68	1.42	1.42	1.59	0.48	0.47	0.89	0.89	1.05
	(0.60)	(0.66)	(0.96)	(1.01)	(0.97)	(0.58)	(0.55)	(0.80)	(0.82)	(0.78)
All-Female class	0.53	0.48	1.62	1.79	1.65	2.65**	2.75***	3.86***	4.05***	3.89***
	(1.12)	(1.09)	(1.33)	(1.42)	(1.42)	(1.02)	(1.00)	(0.93)	(0.98)	(1.02)
Class attendance	−0.23	−0.32	−0.29	−0.37*	−0.30	−0.39**	−0.53***	−0.46**	−0.56**	−0.47**
	(0.18)	(0.19)	(0.22)	(0.22)	(0.22)	(0.18)	(0.17)	(0.21)	(0.21)	(0.22)
1. Agreeableness		−0.04	−0.05	−0.05	−0.03		−0.21	−0.21	−0.21	−0.19
		(0.10)	(0.12)	(0.12)	(0.12)		(0.12)	(0.14)	(0.15)	(0.14)
2. Conscientiousness		−0.23**	−0.17	−0.18	−0.18		−0.21	−0.21***	−0.21***	−0.19**
		(0.11)	(0.12)	(0.13)	(0.12)		(0.10)	(0.12)	(0.12)	(0.11)
3. Extraversion		−0.09	−0.11	−0.09	−0.12		−0.09	−0.09	−0.06	−0.10
		(0.11)	(0.12)	(0.12)	(0.13)		(0.12)	(0.15)	(0.15)	(0.15)
4. Neuroticism		−0.01	−0.05	−0.05	−0.08		0.03	0.00	0.01	−0.02
		(0.11)	(0.10)	(0.10)	(0.10)		(0.13)	(0.13)	(0.13)	(0.14)
5. Openness		0.16*	0.18*	0.17*	0.20*		0.25***	0.25**	0.23**	0.26**
		(0.08)	(0.09)	(0.10)	(0.10)		(0.09)	(0.10)	(0.10)	(0.10)
Z-score IQ test				0.52**					0.59*	
				(0.23)					(0.30)	

Table C.1 (*Continued*)

	Number of risky choices made in session 1					Number of risky choices made in session 2				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Global Gender Gap Index 2010					10.85 (8.65)					10.31 (10.11)
Constant	13.04*** (1.16)	16.20*** (2.60)	13.66*** (2.69)	13.95*** (2.62)	6.17 (6.59)	15.14*** (1.09)	19.27*** (2.62)	19.05*** (3.04)	19.39*** (3.07)	11.94 (7.76)
GTA's fixed effects	*No*	*No*	*No*	*Yes*	*Yes*	*No*	*No*	*No*	*Yes*	*Yes*
Observations	*214*	*214*	*214*	*214*	*214*	*214*	*214*	*214*	*214*	*214*
R^2	*0.056*	*0.111*	*0.195*	*0.210*	*0.206*	*0.068*	*0.152*	*0.187*	*0.203*	*0.195*

Notes: Robust standard errors clustered by module and class are reported in parenthesis. The sample used here is dropped to 214 observations due to missing information in the attendance variable that was not included before. Columns (2)–(5) and (7)–(10) include missing categories for personality traits.

* Corresponds to 10% levels of significance.

** Corresponds to 5% levels of significance.

*** Corresponds to 1% levels of significance.

Table C.2

OLS estimates of the determinants of the number of risky choices (including additional controls).

	Number of risky choices made in session 2 explained by those made in session 1					Difference in the number of risky choices between sessions				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Risky choices	0.37*** (0.08)	0.32*** (0.08)	0.32*** (0.09)	0.31*** (0.09)	0.32*** (0.08)					
Female	−2.08** (0.99)	−2.23** (1.00)	−1.65 (1.02)	−1.79* (0.99)	−1.50 (1.05)	−1.28 (1.26)	−1.37 (1.33)	−0.82 (1.37)	−0.84 (1.37)	−0.83 (1.35)
Economics degree	0.24 (0.48)	0.14 (0.44)	−1.14 (0.84)	−0.68 (0.77)	−1.27 (0.87)	0.68 (0.60)	0.64 (0.60)	−1.28 (0.73)	−1.20 (0.73)	−0.27 (0.76)
Below the age of 20	−0.71 (0.54)	−0.69 (0.53)	−0.64 (0.59)	−0.69 (0.60)	−0.72 (0.62)	−0.57 (0.70)	−0.52 (0.75)	−0.52 (0.82)	−0.53 (0.82)	−0.51 (0.84)
All-Male class	0.18 (0.56)	0.25 (0.53)	0.43 (0.70)	0.45 (0.70)	0.55 (0.66)	−0.33 (0.69)	−0.21 (0.72)	−0.53 (0.82)	−0.53 (0.82)	−0.54 (0.81)
All-Female class	2.46** (1.02)	2.60** (1.02)	3.34*** (0.79)	3.50*** (0.82)	3.37*** (0.84)	2.12 (1.31)	2.27 (1.35)	2.24** (1.01)	2.26** (1.04)	2.24** (1.02)
Class attendance	−0.31* (0.17)	−0.43** (0.16)	−0.37* (0.20)	−0.44** (0.20)	−0.37** (0.21)	−0.16 (0.21)	−0.21 (0.21)	−0.17 (0.25)	−0.18 (0.25)	−0.17 (0.25)
1. Agreeableness		−0.20 (0.13)	−0.19 (0.14)	−0.20 (0.14)	−0.18 (0.14)		−0.17 (0.15)	−0.16 (0.17)	−0.16 (0.17)	−0.16 (0.16)
2. Conscientiousness		−0.19* (0.11)	−0.17 (0.13)	−0.18 (0.13)	−0.18 (0.13)		−0.03 (0.15)	−0.06 (0.17)	−0.06 (0.17)	−0.06 (0.18)
3. Extraversion		−0.06 (0.11)	−0.05 (0.14)	−0.03 (0.14)	−0.06 (0.14)		0.00 (0.13)	0.02 (0.14)	0.03 (0.14)	0.02 (0.14)
4. Neuroticism		0.04 (0.1)	0.02 (0.13)	0.02 (0.13)	0.00 (0.14)		0.04 (0.13)	0.06 (0.14)	0.06 (0.14)	0.06 (0.15)
5. Openness		0.20** (0.07)	0.19** (0.08)	0.18** (0.09)	0.20** (0.09)		0.09 (0.08)	0.06 (0.09)	0.06 (0.10)	0.06 (0.10)
Z-score IQ test				0.43 (0.27)					0.07 (0.28)	
Global Gender Gap Index 2010					6.89 (8.01)					−0.55 (6.53)
Constant	10.36*** *(1.30)*	14.14*** *(3.22)*	14.64*** *(3.56)*	15.09*** *(3.64)*	10.00 *(6.33)*	2.10* *(1.23)*	3.80 *(3.45)*	5.40 *(3.74)*	5.44 *(3.75)*	5.77 *(4.04)*
GTA's fixed effects	*No*	*No*	*Yes*	*Yes*	*Yes*	*No*	*No*	*Yes*	*Yes*	*Yes*
Observations	*214*	*214*	*214*	*214*	*214*	*214*	*214*	*214*	*214*	*214*
R^2	*0.175*	*0.228*	*0.258*	*0.267*	*0.262*	*0.024*	*0.042*	*0.102*	*0.102*	*0.102*

Notes: Robust standard errors clustered by module and class are reported in parenthesis. The sample used here is dropped to 214 observations due to missing information in the attendance variable that was not included before. Columns (2)–(5) and (7)–(10) include missing categories for personality traits.

* Corresponds to 10% levels of significance.

** Corresponds to 5% levels of significance.

*** Corresponds to 1% levels of significance.

Table C.3
Interval regression estimates of the determinants of the number of risky choices.

	Interval for the number of risky choices made in session 1		Interval for the number of risky choices made in session 2	
	(1)	(2)	(3)	(4)
Female	−1.61*	−1.64**	−2.42***	−1.95**
	(0.83)	(0.81)	(0.81)	(0.85)
Economics degree	−0.59	0.46	−0.42	−0.75
	(0.50)	(0.34)	(0.57)	(1.06)
Below the age of 20	0.20	0.25	−0.76	−0.73
	(0.55)	(0.56)	(0.47)	(0.51)
All-Male class	0.48	1.06	0.53	0.80
	(0.55)	(0.70)	(0.62)	(0.62)
All-Female class	0.24	0.70	2.35**	2.45**
	(1.00)	(1.05)	(0.96)	(1.24)
Constant	11.98***	9.94***	13.12***	12.56***
	(0.67)	(1.01)	(0.56)	(1.04)
lnsigma				
Constant	1.25***	1.19***	1.32***	1.31***
	(0.06)	(0.36)	(0.05)	(0.05)
GTA's fixed effects	*No*	*Yes*	*No*	*Yes*
Observations	*219*	*219*	*218*	*218*
R^2				

Notes: The intervals have been built around the prominent numbers: 5 pounds (i.e., around row 6) 10 pounds (i.e., around row 11), 15 pounds (i.e., around row 16). Robust standard errors clustered by module and class are reported in parenthesis.
* Corresponds to 10% levels of significance.
** Corresponds to 5% levels of significance.
*** Corresponds to 1% levels of significance.

Table C.4
OLS estimates of the determinants of the number of risky choices modeling differently the dependent variable.

	Single switchers		*Multiple switchers*	
	Number of risky choices (1)	Row where the student switched (2)	Number of risky choices (3)	Average row (4)
Session 1				
Female	−2.61**	−2.61**	−2.12***	−2.37***
	(0.99)	(0.99)	(0.71)	(0.75)
All-Male class	−0.38	−0.38	0.35	−0.27
	(0.95)	(0.95)	(0.60)	(0.65)
All-Female class	−0.37	−0.37	0.68	0.55
	(1.73)	(1.73)	(0.93)	(0.80)
Economics degree	2.15**	2.15**	0.59	0.93*
	(0.87)	(0.87)	(0.56)	(0.60)
Below the age of 20	−1.16*	−1.16*	−0.36	−0.74
	(0.66)	(0.66)	(0.54)	(0.52)
Constant	9.44***	10.44***	9.97***	11.47***
	(0.35)	(0.35)	(0.31)	(0.31)
GTA's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	*155*	*155*	*206*	*206*
R^2	*0.195*	*0.195*	*0.161*	*0.177*

Table C.4 (*Continued*)

	Single switchers		*Multiple switchers*	
	Number of risky choices (1)	Row where the student switched (2)	Number of risky choices (3)	Average row (4)
Session 2				
Female	−1.80**	−2.00**	−1.36**	−1.38*
	(0.72)	(0.76)	(0.64)	(0.68)
All-Male class	−0.25	−0.17	0.19	0.13
	(0.98)	(1.05)	(0.75)	(0.80)
All-Female class	3.68**	3.16*	2.37*	2.21**
	(1.53)	(1.79)	(1.25)	(1.03)
Economics degree	−0.01	−0.06	−0.23	−0.30
	(1.09)	(1.16)	(0.99)	(0.99)
Below the age of 20	−1.04**	−1.22**	−1.15**	−0.83*
	(0.51)	(0.51)	(0.49)	(0.49)
Constant	11.71***	12.81***	11.63***	12.55***
	(0.25)	(0.26)	(0.23)	(0.24)
GTA's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	*160*	*160*	*192*	*192*
R^2	*0.189*	*0.183*	*0.146*	*0.158*

Notes: The estimations used information from all the students without missing answers during both sessions. During session 1, 70% of the students in our sample switched row only once; 4% never switched and 25% switched more than once. During the second session, students switching only once account for the 73% of the sample, 10% never switched during the second session and 16% switched more than once. The estimations for single switchers (columns 1 and 2) use 155 and 160 observations with usable information during the first and second session, respectively. The estimations for all the respondents (columns 3 and 4) compare the total number of risky choices with the average row where the student switched from B to A; i.e., we exclude those students with inconsistent preferences. In such case, we ended up with 206 and 198 observations (out of 219) during the first and second session, respectively. Robust standard errors clustered by module and class are reported in parenthesis.

* Corresponds to 10% levels of significance.
** Corresponds to 5% levels of significance.
*** Corresponds to 1% levels of significance.

Table C.5
OLS estimates of the determinants of the number of risky choices modeling differently the dependent variable.

	Single switchers		*Multiple switchers*	
	Number of risky choices in session 2 (1)	Row where the student switched (2)	Number of risky choices in session 2 (3)	Average row (4)
Risky choices in session 1	0.36***		0.37***	
	(0.10)		(0.09)	
Female	−2.42**	−1.68**	−1.44	−0.42
	(1.03)	(0.75)	(0.93)	(0.73)
All-Male class	−0.13	−0.40	0.22	0.32
	(0.94)	(0.61)	(0.73)	(0.76)
All-Female class	3.82***	2.21**	2.65***	1.24
	(1.00)	(0.87)	(0.83)	(0.88)
Economics degree	−0.13	−0.95	−1.64*	−1.05
	(0.93)	(0.60)	(0.83)	(0.91)
Below the age of 20	−0.97	−0.78	−0.54	−0.44
	(0.74)	(0.56)	(0.63)	(0.52)
Row where those changing once changed		0.44***		
		(0.09)		
Average row where the student switched				0.36***
				(0.08)
Constant	8.71***	8.46***	9.85***	8.84***
	(1.07)	(0.89)	(1.03)	(0.87)
GTA's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	*155*	*130*	*206*	*184*
R^2	*0.249*	*0.410*	*0.192*	*0.253*

Table C.5 (*Continued*)

Alternative dependentVariable	*Single switchers*		*Multiple switchers*	
	Difference in the number of risky choices	Row where the student switched	Difference in the number of risky choices	Average row (multiple switchers)
Female	−0.27	−0.68	0.60	−0.78
	(1.08)	(1.06)	(1.10)	(1.11)
All-Male class	−0.57	0.11	−0.04	0.08
	(0.49)	(0.56)	(0.64)	(0.56)
All-Female class	1.15	2.62**	0.71	2.68**
	(0.71)	(1.05)	(0.73)	(1.07)
Economics degree	−1.33***	−2.28***	−1.49**	−2.28***
	(0.40)	(0.44)	(0.55)	(0.43)
Below the age of 20	−0.94	−0.30	−1.00	−0.29
	(0.60)	(0.65)	(0.68)	(0.65)
Constant	2.80***	2.74***	2.60***	2.77***
	(0.37)	(0.36)	(0.38)	(0.37)
GTA's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	*160*	*130*	*192*	*129*
R^2	*0.188*	*0.165*	*0.143*	*0.168*

Notes: The estimations used information from all the students without missing answers during both sessions. During session 1, 70% of the students in our sample switched row only once; 4% never switched and 25% switched more than once. During the second session, students switching only once account for the 73% of the sample, 10% never switched during the second session and 16% switched more than once. The estimations for single switchers (columns 1 and 2) use 155 and 160 observations with usable information during the first and second session, respectively. The estimations for all the respondents (columns 3 and 4) compare the total number of risky choices with the average row where the student switched from B to A; i.e., we exclude those students with inconsistent preferences. In such case, we ended up with 206 and 198 observations (out of 219) during the first and second session respectively. Robust standard errors clustered by module and class are reported in parenthesis.

* Corresponds to 10% levels of significance.

** Corresponds to 5% levels of significance.

*** Corresponds to 1% levels of significance.

Table C.6

OLS estimates of the determinants of the number of risky choices modeling IQ with the raw score of the test.

	Number of risky choices made in session 1 (1)	Number of risky choices made in session 2 (2)	Number of risky choices made in session 2 (3)	Difference in risky choices between sessions (4)
Female	−1.52	−2.06**	−1.51	−0.54
	(0.96)	(0.90)	(1.05)	(1.37)
All-Male class	1.31	0.70	0.22	−0.61
	(0.89)	(0.76)	(0.68)	(0.88)
All-Female class	1.14	3.04***	2.62***	1.90*
	(1.22)	(1.02)	(0.85)	(0.96)
Economics degree	0.60	−0.99	−1.21	−1.59**
	(0.67)	(0.82)	(0.73)	(0.72)
Below the age of 20	−0.13	−0.71	−0.66	−0.57
	(0.58)	(0.61)	(0.61)	(0.76)
Raw score IQ test	0.15**	0.16*	0.10	0.01
	(0.06)	(0.08)	(0.07)	(0.08)
Risky choices			0.36***	
			(0.08)	
Constant	7.86***	11.48***	8.62***	3.63***
	(0.95)	(1.05)	(1.33)	(1.06)
GTA's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	*219*	*219*	*219*	*219*
R^2	*0.164*	*0.116*	*0.210*	*0.085*

Notes: The *Z*-score of the IQ corresponds to the standardized measure (mean zero and standard deviation equal to one) of the IQ test among the students who responds the test. The raw score of the IQ corresponds to the number of correct answers responded by the student. The IQ-population norms correspond to the percentile where the student belongs following the international standards. Robust standard errors clustered by module and class are reported in parenthesis.

* Corresponds to 10% levels of significance.

** Corresponds to 5% levels of significance.

*** Corresponds to 1% levels of significance.

Table C.7
OLS estimates of the determinants of the number of risky choices modeling IQ with the population norms standardization of the test.

	Number of risky choices made in session 1 (1)	Number of risky choices made in session 2 (2)	Number of risky choices made in session 2 (3)	Difference in risky choices between sessions (4)
Female	−1.52	−2.06**	−1.51	−0.54
	(0.96)	(0.90)	(1.05)	(1.37)
All-Male Class	1.31	0.70	0.22	−0.61
	(0.89)	(0.76)	(0.68)	(0.88)
All-Female Class	1.14	3.04***	2.62***	1.90*
	(1.22)	(1.02)	(0.85)	(0.96)
Economics degree	0.60	−0.99	−1.21	−1.59**
	(0.67)	(0.82)	(0.73)	(0.72)
Below the age of 20	−0.13	−0.71	−0.66	−0.57
	(0.58)	(0.61)	(0.61)	(0.76)
IQ score (population norms)	0.08**	0.08*	0.05	0.00
	(0.03)	(0.04)	(0.03)	(0.04)
Risky choices			0.36***	
			(0.08)	
Constant	7.86***	11.48***	8.62***	3.63***
	(0.95)	(1.05)	(1.33)	(1.06)
GTA's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*
Observations	*219*	*219*	*219*	*219*
R^2	*0.164*	*0.116*	*0.210*	*0.085*

Notes: The *Z*-score of the IQ corresponds to the standardized measure of the IQ test (mean zero and standard deviation equal to one) among the students who responds the test. The raw score of the IQ corresponds to the number of correct answers responded by the student. The IQ-population norms correspond to the percentile where the student belongs following the international standards. Robust standard errors clustered by module and class are reported in parenthesis.
* Corresponds to 10% levels of significance.
** Corresponds to 5% levels of significance.
*** Corresponds to 1% levels of significance.

Table C.8
OLS estimates of the determinants of the number of risky choices adding the Global Gender Gap Index and its interaction with the gender.

	Number of risky choices made in session 1			Number of risky choices made in session 2		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−1.58*	−1.38	−27.43***	−2.06**	−1.90*	−31.88***
	(0.92)	(0.89)	(8.21)	(0.94)	(0.95)	(10.10)
Economics degree	0.28	0.31	0.14	−1.03	−1.31	−1.50
	(0.68)	(0.69)	(0.69)	(0.83)	(0.84)	(0.89)
Below the age of 20	−0.05	−0.17	−0.13	−0.72	−0.74	−0.70
	(0.69)	(0.61)	(0.61)	(0.66)	(0.63)	(0.64)
All-Male class	1.22	1.27	1.07	0.42	0.67	0.44
	(0.90)	(0.87)	(0.83)	(0.77)	(0.72)	(0.75)
All-Female class	0.95	1.03	1.20	2.76***	2.93***	3.11***
	(1.31)	(1.19)	(1.22)	(0.93)	(1.01)	(1.02)
English speaking	0.36			0.45		
	(1.11)			(2.18)		
Asia	−0.56			0.95		
	(1.06)			(1.90)		
Asia-Europe	1.44			−2.50		
	(1.06)			(2.45)		
China	0.94			1.04		
	(1.24)			(2.29)		
East-Europe	0.44			0.62		
	(1.08)			(2.01)		
Europe	1.25			1.01		
	(1.31)			(2.27)		

Table C.8 (*Continued*)

	Number of risky choices made in session 1			Number of risky choices made in session 2		
	(1)	(2)	(3)	(4)	(5)	(6)
Global Gender Gap Index 2010		5.15 (7.47)	−6.41 (7.74)		5.87 (8.65)	−7.43 (9.84)
GGI × Female			37.04*** (11.53)			42.63*** (13.94)
Constant	9.21*** (1.02)	6.09 (5.17)	14.65** (5.44)	12.73*** (2.05)	9.31 (6.15)	19.17** (7.12)
GTA's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	*219*	*219*	*219*	*219*	*219*	*219*
R^2	*0.164*	*0.152*	*0.180*	*0.121*	*0.106*	*0.137*

Notes: The *Z*-score of the IQ corresponds to the standardized measure of the IQ test (mean zero and standard deviation equal to one) among the students who responds the test. The raw score of the IQ corresponds to the number of correct answers responded by the student. The IQ-population norms correspond to the percentile where the student belongs following the international standards. Robust standard errors clustered by module and class are reported in parenthesis.

* Corresponds to 10% levels of significance.
** Corresponds to 5% levels of significance.
*** Corresponds to 1% levels of significance.

Table C.9
OLS estimates of the determinants of the number of risky choices adding the Global Gender Gap Index and its interaction with the gender.

	Number of risky choices made in session 2			Difference in the number of risky choices made between sessions		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	−1.46 (1.08)	−1.39 (1.09)	−22.36** (10.17)	−1.05 (1.26)	−0.52 (1.36)	−4.45 (9.17)
Risky choices	0.38*** (0.08)	0.37*** (0.08)	0.35*** (0.09)			
All-Male class	−0.04 (0.76)	0.19 (0.66)	0.07 (0.71)	−0.38 (0.73)	−0.61 (0.88)	−0.64 (0.92)
All-Female class	2.40*** (0.77)	2.54*** (0.84)	2.70*** (0.87)	1.75 (1.28)	1.89* (0.94)	1.92* (0.96)
Economics degree	−1.14 (0.79)	−1.42* (0.74)	−1.55* (0.77)	0.66 (0.65)	−1.61** (0.73)	−1.64** (0.71)
Below the age of 20	−0.70 (0.61)	−0.68 (0.62)	−0.65 (0.63)	−0.59 (0.66)	−0.58 (0.77)	−0.57 (0.77)
English speaking	0.31 (2.06)			0.24 (1.78)		
Asia	1.16 (1.91)			2.04 (1.85)		
Asia-Europe	−3.05 (2.42)			−3.89 (2.59)		
China	0.68 (2.29)			0.40 (2.14)		
East-Europe	0.45 (1.97)			0.68 (1.83)		
Europe	0.53 (2.03)			0.29 (1.69)		
Global Gender Gap Index 2010		3.95 (6.51)	−5.21 (8.32)		0.72 (5.10)	−1.02 (6.53)
GGI × Female			29.77** (13.62)			5.58 (11.91)
Constant	9.22*** (2.22)	7.04 (4.91)	14.08** (6.58)	0.80 (1.65)	3.23 (3.61)	4.52 (4.82)

Table C.9 (*Continued*)

	Number of risky choices made in session 2			Difference in the number of risky choices made between sessions		
	(1)	(2)	(3)	(4)	(5)	(6)
GTA's fixed effects	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*	*Yes*
Observations	*219*	*219*	*219*	*219*	*219*	*219*
R^2	*0.224*	*0.206*	*0.221*	*0.053*	*0.085*	*0.086*

Notes: Robust standard errors clustered by module and class are reported in parenthesis.
* Corresponds to 10% levels of significance.
** Corresponds to 5% levels of significance.
*** Corresponds to 1% levels of significance.

Table C.10
OLS estimates for the effect of risk preferences on the accuracy rate of the test.

	Accuracy rate in the IQ test explained by the risky choices made in session 1			Accuracy rate in the IQ test explained by the risky choices made in session 2		
	(1)	(2)	(3)	(4)	(5)	(6)
Risky choices in the respective session	0.01* (0.00)	0.01* (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
Female	−0.02 (0.04)	−0.01 (0.04)	0.04 (0.05)	−0.02 (0.04)	−0.01 (0.04)	0.04 (0.05)
All-Male class	−0.07** (0.03)	−0.06** (0.03)	−0.02 (0.05)	−0.07*** (0.03)	−0.06** (0.03)	−0.02 (0.05)
All-Female class	0.02 (0.04)	0.02 (0.04)	−0.01 (0.06)	0.01 (0.04)	0.01 (0.04)	−0.02 (0.06)
Economics degree	0.00 (0.02)	−0.00 (0.03)	−0.13 (010)	0.00 (0.02)	−0.00 (0.02)	−0.13 (0.10)
Below the age of 20	−0.01 (0.03)	−0.02 (0.03)	−0.01 (0.03)	−0.01 (0.03)	−0.01 (0.03)	−0.01 (0.03)
Global Gender Gap Index 2010		0.38 (0.31)	0.54 (0.33)		0.39 (0.31)	0.54 (0.33)
1. Agreeableness			0.00 (0.01)			0.00 (0.01)
2. Conscientiousness			0.00 (0.01)			0.00 (0.01)
3. Extraversion			−0.01** (0.01)			−0.01** (0.01)
4. Neuroticism			−0.00 (0.01)			−0.00 (0.01)
5. Openness			0.01** (0.00)			0.01** (0.00)
Constant	0.64*** (0.05)	0.38* (0.22)	0.37 (0.26)	0.65*** (0.05)	0.37* (0.22)	0.35 (0.26)
GTA's fixed effects	*No*	*No*	*Yes*	*No*	*No*	*Yes*
Observations	*219*	*219*	*219*	*219*	*219*	*219*
R^2	*0.039*	*0.045*	*0.185*	*0.035*	*0.043*	*0.187*

Notes: The variable of accuracy rate is the ratio of correct answers to the total possible answers. Robust standard errors clustered by module and class are reported in parenthesis. In columns (3) and (6) we have included indicator variables for missing categories in the personality answers.
* Corresponds to 10% levels of significance.
** Corresponds to 5% levels of significance.
*** Corresponds to 1% levels of significance.

Appendix D. Determinants of class' attendance

See Table D.1.

Table D.1
Determinants of class' attendance.

	Attendance during the first 8 weeks	
	Coefficient	Standard error
Female	−0.29	(0.31)
Economics degree	1.62**	(0.67)
All-Male class	−0.29	(0.34)
All-Female class	0.79*	(0.44)
Global Gender Gap Index 2010	0.56	(2.26)
1. Agreeableness	0.00	(0.04)
2. Conscientiousness	−0.07*	(0.04)
3. Extraversion	−0.07*	(0.04)
4. Neuroticism	0.05	(0.04)
5. Openness	0.04	(0.03)
GTA fixed effects	*Yes*	
Observations	*214*	

Notes: Standard errors in parenthesis.
* Corresponds to 10% levels of significance.
** Corresponds to 5% levels of significance.
*** Corresponds to 1% levels of significance.
We also included dummy variables to control for whether any aspect of the Big Five was missing.

Appendices E and F. Supplementary data

Supplementary material related to this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jebo.2013.12.017.

References

Andersen, S., Harrison, G.W., Lau, M.I., Rutstrom, E.E., 2006. Elicitation using multiple price list formats. Experimental Economics 9 (4), 383–405.
Arthur Jr., W., Tubre, T.C., Paul, D.S., Sanchez-Ku, M.L., 1999. College-sample psychometric and normative data on a short form of the Raven Advanced Progressive Matrices Test. Journal of Psychoeducational Assessment 17, 354–361.
Arulampalam, W., Booth, A.L., Bryan, M.L., 2007. Is there a glass ceiling over Europe? Exploring the gender pay gap across the wages distribution. Industrial and Labor Relations Review 60 (2), 163–186.
Bertrand, M., 2011. New perspectives on gender. In: Ashenfelter, O., Card, O.D. (Eds.), Handbook of Labor Economics, vol. 4B. Elsevier, pp. 1543–1590.
Bertrand, M., Hallock, K.F., 2001. The gender gap in top corporate jobs. Industrial and Labor Relations Review 55 (1), 2–21.
Blau, F.D., Kahn, L.M., 2006. The US gender pay gap in the 1990s: slowing convergence. Industrial and Labor Relations Review 60 (1), 45–66.
Booth, A.L., Nolen, P.J.P.J., 2012a. Gender differences in competition: the role of single-sex education. Journal of Economic Behavior and Organisation 81, 542–555.
Booth, A.L., Nolen, P.J., 2012b. Gender differences in risk behaviour: does nurture matter? Economic Journal 122 (February (668)), F56–F78.
Booth, A.L., Nolen, P.J., 2012c. Salience, risky choices and gender. Economics Letters 117, 517–520.
Bors, D.A., Stokes, T.L., 1998. Raven's advanced progressive matrices: norms for first year university students and the development of a short form. Educational and Psychological Measurement 58, 382–398.
Bruner, D., 2007. Multiple Switching Behavior in Multiple Price Lists. Working Paper. Appalachian State University.
Burks, S.V., Carpenter, J.P., Goette, L., Rustichini, A.A., 2009. Cognitive skills affect economic preferences, strategic behavior, and job attachment. Proceedings of the National Academy of Sciences of the United States of America 106 (19), 7745–7750.
Charness, G., Kuhn, P.J., 2011. Lab labor: what can labor economists learn from the lab? In: Ashenfelter, O., Card, O.D. (Eds.), Handbook of Labor Economics, vol. 4A. Elsevier, pp. 229–330.
Croson, R., Gneezy, U., 2009. Gender differences in preferences. Journal of Economic Literature 47 (2), 1–27.
Datta Gupta, N., Poulsen, A., Villeval, M., 2005, November. Male and Female Competitive Behavior: Experimental Evidence, IZA Working Paper 1833.
Dohmen, T.J., Falk, A., Huffman, D., Sunde, U., 2010. Are risk aversion and impatience related to cognitive ability? American Economic Review 100 (3), 1238–1260.
Eckel, C., Grossman, P., 2008. Men, women and risk aversion: experimental evidence. In: Charles, R., Plott, Vernon, L., Smith (Eds.), Handbook of Experimental Economics Results. Elsevier Science B.V./North-Holland, Amsterdam, The Netherlands, pp. 1078–1086 (Chapter 114).
Fryer, R., Levitt, S., 2010. An empirical analysis of the gender gap in mathematics. American Economic Journal: Applied Economics 2 (2), 210–240.
Gneezy, U., Leonard, K.L., List, J.A., 2009. Gender differences in competition: evidence from a matrilineal and a patriarchal society. Econometrica 77 (5), 1637–1664.
Gneezy, U., Niederle, M., Rustichini, A., 2003. Performance in competitive environments: gender differences. Quarterly Journal of Economics 118 (3), 1049–1074.
Guiso, L., Monte, F., Sapienza, P., Zingales, L., 2008. Culture, gender and math. Science 320, 1164–1165.

Harrison, G.W., Morten, I.L., Rutström, E.E., 2007. Estimating risk attitudes in Denmark: a field experiment. Scandinavian Journal of Economics 109 (2), 341–368.
Heineck, G., 2007. Does it pay to be nice? Personality and earnings in the UK. LASER Discussion Papers. Paper No. 3–9.
Holt, C., Laury, A.S.K., 2002. Risk aversion and incentive effects. American Economic Review 92 (5), 1644–1655.
Niederle, M., Vesterlund, L., 2007. Do women shy away from competition? Do men compete too much? Quarterly Journal of Economics 122 (3), 1067–1101.
Raven, J., Raven, J.C., Court, J.H., 1998. Manual for Raven's Progressive Matrices and Vocabulary Scales. Section 4: The Advanced Progressive Matrices. Harcourt Assessment, San Antonio, TX.
Raven, J., 2000. The Raven's progressive matrices: change and stability over culture and time. Cognitive Psychology 41, 1–48.
Schubert, R., Brown, M., Gysler, M., Brachinger, H.W., 1999. Financial decision-making: are women really more risk-averse? American Economic Review 89 (2), 381–385.