Gender Differences in Risk Aversion: Do Single-Sex Environments Affect their Development?

Alison Booth
Lina Cardona Sosa
Patrick Nolen

DISCUSSION PAPER NO. 654

December 2011
Gender Differences in Risk Aversion:
Do Single-Sex Environments Affect their Development?  
Alison Booth\textsuperscript{a,b}, Lina Cardona Sosa\textsuperscript{a} and Patrick Nolen\textsuperscript{a}

\textsuperscript{a}University of Essex, \textsuperscript{b}Australian National University

November 2011

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 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 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.

JEL Classification: C9, C91, C92, J16, D01, D80, J16, J24

Keywords: gender, risk preferences, single-sex groups, cognitive ability

Corresponding author:
Professor AL Booth
Research School of Economics,
Australian National University,
Arndt Bldg, Acton 0200 ACT Australia
Email: Alison.booth@anu.edu.au
Tel: +61 (0)2 6125 3285
Fax: +61 (0)2 6125 5124

1 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 PhD students who assisted in conducting the experiment.
I. Introduction

Women are under-represented in high-paying jobs and in high-level occupations. Recent studies in experimental economics have examined to what degree this under-representation may be due to innate differences between men and women. 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. For instance, boys are pushed to take risks when participating in competitive sports while girls are often encouraged to remain cautious. Thus the riskier choices made by males could be due to the nurturing received from parents or peers. Likewise, the disinclination of women to take risks could be the result of parental or peer pressure not to do so.

Gneezy, Leonard and List (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


3 For a survey of new perspectives on gender in economics, see Bertrand (2010). See Charness and Kuhn (2010) for a survey on laboratory experiments in labour economics.
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 two distinct environments or ‘cultures’, Booth and Nolen (2011b) examined the effect on preferences over risk of two types of environmental influences -- educational environment (single-sex or coeducational) and randomly assigned experimental peer-groups. The former represents longer-run nurturing experiences, while the latter -- the experimental group -- captures short-run environmental effects. The experimental subjects in that study were UK students in years 10 and 11 who were 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.

This present paper extends previous experimental work by designing an experiment to look at the effects of nurture over time. 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. While ‘environments’ have been found to affect risk-taking behaviour, it is unclear how long these effects last. Do single-sex environments change risk-taking behaviour temporarily, or do they have a lasting effect? As shown by Gneezy and List (2006), experimental results such as gift-exchange might have only an initial effect and disappear subsequently. If environmental influences are only temporary, then changing educational pre-market factors such as educational or training environments may not affect gender pay gaps or under-representation of women in top jobs. For this reason, it is desirable to explore how the environment or ‘nurture’ affect longer term risk preferences or attitudes. That is the goal of our experiment reported herein.

---

4 In a companion paper, Booth and Nolen (2011a) investigated how competitive behaviour (including the choice between piece-rates and tournaments), is affected by single-sex schooling.
Our subjects are first year college (university) students who 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. 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 Yestrumskas (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 summarized in Eckel and Grossman (2008) and Croson and Gneezy (2009).\(^5\) 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 weeks in a single-sex 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 results are robust to the inclusion or exclusion of controls for personality type, as well as to a number of sensitivity tests reported in Section V.

\(^5\) Harbaugh, Krause and Vesterlund (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.
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.

II. Subjects 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. The students were randomly assigned to small classes 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 student then took part in another paid experiment by filling
out a second risk questionnaire during their class that week. The results from the risk
questionnaire 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. 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; and Dohmen, Falk, Huffman and Sunde, 2010).

There were a total of 37 classes.\textsuperscript{6} Classes were taught by Graduate Teaching Assistants
(GTAs), who were PhD students hired in a competitive hiring process from the pool of PhD
applicants. GTAs were told not to discuss with students any of the details – or objectives - of
the experiments. (We control for the gender of GTAs in our regressions.) Appendix D, at the end
of the paper, gives the full set of instructions provided to GTAs and the set of slides they were
each required to use in the experiment.

\textsuperscript{6} 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.
II.A. Experimental Measures of Risk Aversion

We used paid experiments to measure subjects’ willingness to take risks at two points in time. In both sessions, students were asked to answer a questionnaire designed to assess subject’s 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 £20 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 and that 10% of all the participants (also selected at random) would receive the corresponding payment.7

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. 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.8 Thus our measure of risk aversion is the

---

7 A coin was flipped to determine the outcome of the lottery if the student chose column B in the randomly selected row.
8 This format is also similar to that of 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.
number of risky choices made by the subject.\textsuperscript{9} 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.

Figures 1a and 1b show the distribution of the number of risky choices made by each student in sessions 1 and 2 for the subsample of those attending both sessions. The mean number of risky choices in Session 1 was 11.44 and the median number 10.5, while in Session 2 the mean and median were 12 and 12.40 respectively.\textsuperscript{10}

\textbf{II.B. Our Measure of Cognitive Ability}

Burks, Carpenter, Goette and Rustichini (2009) and Dohmen, Falk, Huffman and Sunde (2010) find that cognitive skills are likely to affect individuals’ economic preferences. Given this, higher

\textsuperscript{9} The row where individual changed and the number of rows where the lottery was chosen were perfectly correlated. Nevertheless we also 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.

\textsuperscript{10} Assuming a constant relative risk aversion utility function of the type $u(x)=x^{1-\sigma}/(1-\sigma)$, where $\sigma$ is the degree of relative risk aversion, we calculate that the value of $\sigma$ making an individual just indifferent between choosing the lottery and the certain outcome lies between 0.4 and 0.3. This is similar to the range of 0.3 to 0.5 found by Holt and Laury (2002), and the range of 0.43 and 0.48 found by Dohmen et al. (2010).
cognitive skills may affect individual’s economic outcomes through greater risk-taking behaviour as well as through the more conventional channels of influence.

How might cognitive skills affect individuals’ economic preferences? 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 1,000 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.11 Research by Dohmen et al. (2010) used a more representative sample and a different IQ measure. Using data from around 1,000 adults representative of the German population, they found that higher cognitive ability is associated with lower risk aversion and less pronounced impatience.12

In our experiment, we use an abbreviated version of the Raven’s Advanced Progressive Matrices to measure IQ. This test, designed to assess eductive ability, is a widely accepted measure of higher order general mental ability that does not rely on cultural context or prior experience (see for example Bors and Stokes, 1998, and Arthur, Tubre, Paul and Sanchez-Ku, 1999, and references therein).13 Indeed, it was originally designed to provide information about a subjects’ ability using a non-verbal setting uncontaminated by linguistic background. The results have been shown to be consistent across cultures and over time (Raven, 1998).

Thus our first measure of cognitive ability is the raw score from a 20 minute version of this test, conducted in the first session of our experiment. The distribution of the scores is shown in Figure 2. It is a measure of ability based on the number of correct answers. The figure

11 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.

12 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 seconds according to a given correspondence. The other, the word fluency test, gives subjects a timed vocabulary test.

13 Edective ability is defined by Raven, et. al (1998) as “the ability to make meaning out of confusion, the ability to forge largely non verbal constructs which make it easy to handle complexity”.
shows that the test is skewed to the left. We reject the null hypothesis that the skewness and kurtosis are zero (i.e. the distribution of the values is normal) at a 99% confidence level. 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.

**Figure 2:** Distribution of the scores for the ability test

II.C. *Personality Traits*

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. The importance of personality traits is that they can be related to individual characteristics in what is called the Five Factor Model (FFM). Under this scheme, each of the personality dimensions is related to one of the following five characteristics: openness to experience, conscientiousness, extraversion, agreeableness and neuroticism. The structure of the questionnaire used in this experiment is in the same as that found in the British Household Panel (BHPS). The questions are answered on a scale from 1 to 7 and the final score results are
obtained by adding the scores within each specific dimension (Heineck, 2007). The components of each dimension are summarized in Table 1.

Insert Table 1 near here

Hence, in some of our regressions, to be reported below, we include the score of each of the five dimensions to control for personality traits. To avoid excluding individuals leaving these questions unanswered, we include dummies for missing responses for each personality dimension.

II.D Other Controls

In addition to being affected by gender, cognitive ability and personality traits, individual’s preferences may also be influenced by variables such as language, subject of study, the gender of the class teacher and the gender environment in high school.\textsuperscript{14} We control for the family background characteristics using information obtained from the demographic questionnaire.

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.

To control for the differences in language and culture, we include in our analysis a set of dummy variables grouping different nationalities by region-of-origin. These are Asia (excluding China), Africa, China, Europe, Eastern Europe, and the Asia-Europe boundary, with the English-speaking countries forming the base.\textsuperscript{15}

\textsuperscript{14} To control for whether or not our experimental subjects were exposed to single-sex schooling before entering university, we use a dummy variable for attending a coed school. 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 (2001b), whose subject pool comprised only students from government-financed secondary schools.

\textsuperscript{15} 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.
III. 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 usable responses for 486 students turning up for the first lecture and class and signing a consent form. Of these, some left the course and some failed to participate in both the first and second experiment. Consequently we have usable responses for 231 individuals who participated in both experimental sessions.\(^{16}\)

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 and found that attrition is not a problem. 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 the Appendix. The third column in each panel of Table B.1 shows 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. The only exception is found in the final panel, for coeducational classes, where the proportion of female teachers differs across the initial sample and the subsample participating in both experiments.

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. The results of this probit are presented in Appendix Table B.2. The means of the explanatory variables are presented in Table 2a below. The estimates in Appendix Table B.1 show that

\(^{16}\) We restricted our sample to those individuals without missing answers from the risk questionnaires and valid information for the other variables of interest. Out of the 260 students who attended both experiments, 231 completed the whole questionnaire.
neither risk preferences observed in session 1 nor the IQ test score had a statistically significant effect on the attendance probability in session 2.\textsuperscript{17} We also found that female attendance in session 2 was not significantly affected by being in an all-female group.

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. 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. None of these variables had a statistically significant impact. This it seems that attrition was not specific to particular class types.

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

IV. The Results
In this section we first discuss the means of our variables, and then present the estimates of the determinants of the number of risky choices, and discuss how they change across two points in time.

IV. A. 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 2a reports the means of the variables for two different subsamples, first, all those with usable responses who took part in session 1, and second, those 231 individuals who took part in both sessions.

\textsuperscript{17} 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. Our probit estimates indicate that neither risk attitudes nor ability as proxied by IQ score determine attendance.
Our main interest here is in the second column of means in Table 2a. Inspection reveals that around 34% of subjects are female and 75% of our subjects are younger than twenty years. The raw IQ test score was 11.75 and the standardised value was 0.01. Just under half the subsample was enrolled in a straight economics degree. Around 46% of classes were single-gender (27% of all students were in all-male classes while 19% of all students were in all-female classes.) The means show considerable heterogeneity across area-of-origin: 32% of our subjects were from English-speaking regions, 20% were from Eastern Europe, 18% from the rest of continental Europe, 16% from China, 6% from other Asia, 4% from Africa and 3% from Asia-Europe. Finally, it is interesting to note that students were distributed roughly equally across personality types.

Next we compare means across the single-sex and coeducational groups for our estimating subsample. For expositional simplicity, these are presented in Appendix Table B.3.. From this table, we see that the number of risky choices does not differ across the class-types. The variables with statistically significant differences in means across the two groups are age, attendance at a coeducational school prior to entering college, gender by construction,\textsuperscript{18} the gender of the class teacher, student of Asian-European origin, and the global gender gap index (to be used in the sensitivity analysis later in the paper). This indicates that the random assignment to single sex classes was on the whole very effective, but does also suggest that it is important to include these controls in our regressions.

**IV. B Cross-sectional and Fixed Effects Estimates of Number of Risky Choices**

Table 2b reports the means and standard deviations of the dependent variable used in the regression analysis. It shows that the mean number of risky choices was 11.44 in session 1 and 12.40 in session 2.

\textsuperscript{18} Since the majority of students are male, fewer females were assigned to coeducational classes. In fact, 72% of men were allocated to mixed classes while 28% of women were. We explore this variation in more detail later in the paper.
Table 3 presents the ordinary least squares (OLS) and the fixed effects (FE) estimates of the determinants of the number of risky choices (which can range between 0 and 20). The estimated coefficients from the initial session are given in the first pair of columns, headed *Session 1*. Robust standard errors (SEs) are reported in parenthesis. The specification in Column [1] displays the estimates of our variables of interest – gender, cognitive ability, and the gender composition of the group to which the individual was randomly assigned prior to the start of the academic year. Excluded from this specification are controls for the ‘Big Five’ personality types. The other controls are listed in the note under the table. The next column, [2], provides the estimated coefficients from the specification in which personality type is also controlled for.

Insert Table 3 near here.

The results from column [1] show that females are less likely than males to make risky choices, and that an increase in cognitive ability is associated with a slight increase in the number of risky choices made. Both estimated effects are statistically significant at the 10 percent and 5% level respectively. Notice that at this stage, the first class of the academic year, there are no differences between students assigned to the various group types in the number of risky choices made. From column [2], it can be seen that once personality type has been controlled for, the magnitude of the estimated coefficient to the female dummy variable is slightly reduced and it is less precisely determined. The impact of the IQ test score remains statistically significant at the 5 percent level. This confirms the results of Burks et al. (2009) for their sample of trainee US truckers using a very similar IQ score measure, as well as the findings of Dohmen at al (2010) for their sample representative of the German population across a much wider age-range (although our marginal effect of a one standard deviation increase in IQ is smaller than theirs). This may reflect the fact that we have a more homogeneous population (our subjects are all smart enough to be admitted into a degree course, and are all much the same age). We are also using a different measure of IQ, as discussed in Section II.B.

The estimated coefficients for the other controls are reported in full in Appendix C, Table C.1. They show that, in the first session, Asian students and those majoring in economics were less likely to make risky choices, while students with high scores in ‘openness to new experiences’ made a greater number of risky choices. The latter result accords with our priors,
since it suggests that curious people are more likely to take risks. This result was also found by Dohmen et. al (2010).

Next we consider the results 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 columns headed Session 2. In the intervening weeks between Sessions 1 and 2, students have been attending the same classes, as no switching was permitted, and so they have had a considerable period of 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 estimated coefficient to being in an all-female group is large and statistically significant. Thus (using the estimates reported in column (4)), we see that relative to the base of a male student, a woman randomly assigned to an all-female group will choose slightly fewer risky rows (-2.72 +2.48=-0.24). In contrast, if she had been assigned to a coeducational group, she would choose nearly 3 fewer rows of risky choices. This is a large effect, and one that dwarfs the impact of IQ-test scores, as we shall see.

The effect of cognitive ability is now slightly smaller in magnitude and is statistically significant at the 10% level, suggesting that a one SD increase in cognitive ability is associated with a very small increase (by under half an extra risky choice) in the number of risky choices made. From the additional controls included in the analysis (see Appendix C, Table C.1), we found that enrolment in an economics degree is now uncorrelated with the number of risky choices, while being from the Asia-Europe boundary decreases them. As was also observed in session 1, openness to experience is associated with a significantly greater number of risky choices.

Finally, consider column (5). The gender-composition of our classes was, by construction, random. Moreover, we have controls in our analysis to pick up some effects that are typically unobserved in surveys – in particular, cognitive ability and personality type. Nonetheless readers may worry about unobservable factors which may be correlated with some of our exogenous variables and which may also affect risky choices. To allay these concerns, we therefore estimated a fixed effects model based on the balanced panel of 231 individuals (and 462 person-wave observations). In this specification we assigned the group
composition variables to take the value zero at wave 1 and the value unity at wave 2. The estimates from this exercise are presented in the final column of Table 3. Inspection of these estimates reveals that, for men in men-only groups, there is no statistically significant difference in the number of risky choices compared to the base of individuals in coeducational groups. However, for women in women-only groups, the number of risky choices increases by 2.09, an effect that is statistically significant at the 1 percent level.

In summary our results in Table 3 suggest that women are less likely to make risky choices than men at both dates. However, after eight weeks in a single-sex environment, women were significantly more likely to choose the lottery than their counterparts in coeducational groups. These results were robust to the inclusion or exclusion of controls for personality type.

V. Sensitivity Analysis

We next report the results from a number of robustness checks on the specifications reported above. These checks involved experimenting with different estimation methods; different ways of modelling the dependent variable; different ways of transforming the score for cognitive ability; and replacing the country-of-origin dummy variables by the World Economic Forum’s Gender Gap Index (GGI). This last exercise aims to see if risky choices are affected by the “cultural” factors encapsulated in that index. Finally, we investigated if an individual’s IQ test-taking strategy is correlated with risk-aversion and with gender. The estimates from these sensitivity tests are presented in Appendix C. (This appendix is for the referees’ information and not intended for publication unless advised otherwise.)

V.A Are the results sensitive to estimation method?

We 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.2 reveals.\textsuperscript{19} The same set of variables remain statistical significant.

\textbf{V.B Are the results sensitive to different ways of modelling the dependent variable?}

To account for multiple switching,\textsuperscript{20} 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.\textsuperscript{21} 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, we found that the estimates reported in Table C.3 are very similar to the findings presented in Table 3. Moreover these results do not differ from the estimation using the number of risky choices (baseline) but restricting the analysis to multiple switchers.\textsuperscript{22}

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. The estimates are reported in 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 3. Once again, the results support our main conclusions.\textsuperscript{23}

\textbf{V.C Are the results sensitive to different ways of transforming cognitive ability?}

The results using different scales for the IQ-test are shown in Tables C.5 and C.6 in Appendix C. When using the raw score of the IQ test, we found that the statistical significance of the coefficients of interest do not change and our main conclusions remain the same (see Table C.5, and note that the coefficients for IQ should be interpreted differently under this scale.)

\textsuperscript{19} 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).

\textsuperscript{20} For the subsample used in the analysis, 5% of the individuals never switched between columns, 70% switched once and 25% switched more than once.

\textsuperscript{21} Our analysis is restricted to those students starting with consistent choices, i.e. those starting moving from the lottery option to the safety one.

\textsuperscript{22} The main specification uses the number of risky choices as our dependent variable without restricting for the number of changes.

\textsuperscript{23} Estimates available for the different dependent variables are available from authors on request.
When using the population norms for the IQ test (standardizing the IQ score following the performance of a similar population and reported in Table C.6), the cognitive ability measure is not statistically significant during the second session for any specification (in contrast with column (4) when using the other scales). Nevertheless the estimated effects of all-female classes and of gender are almost identical to the previous analysis.

V.D  Do culturally-driven norms and beliefs affect women’s risky choices?

Our subjects come from a variety of different regions, as Table 2a shows. These countries are characterised by different cultural norms and levels of gender emancipation. Although the region-of-origin dummy variables typically did not have a statistically significant effect on the number of risky choices (see Appendix Table C.1), these dummies are rather crude proxies for cultural differences.\(^{24}\) We therefore experimented with replacing them by the World Economic Forum’s Gender Gap Index (GGI). 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.\(^{25}\) In summary, we found that the GGI was statistically insignificant and that the coefficients to our other variables of interest were robust to inclusion of the GGI. Thus our results reported in Table 3 remain little changed. While the index itself does not have a statistically significant effect, once we interact it with gender we find that women from countries from more ‘gender equality’ are associated with a greater number of risky choices.

V.E  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

---

\(^{24}\) In the second session, students from the Asia-Europe boundary were less likely to make risky choices.

\(^{25}\) 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.
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 choices, as well as all the other controls used in Table 3. We found that neither gender nor the number of risky choices in session 1 have a statistically significant association with the accuracy rate. The estimated coefficient (SE) to female was -0.021 (0.038) and the coefficient to the number of risky choices was 0.005 (0.003).26 (These estimates are available from the authors on request.) Thus 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.

VI. 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 (that is, 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.

Our results are interesting because 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 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 the following. Women, even those endowed with an intrinsic propensity to make riskier choices, may be discouraged from doing so because they are inhibited by culturally-driven norms and beliefs about the appropriate mode of female behaviour – avoiding risk. But once they are placed in an all-female environment, this inhibition is reduced. No longer reminded of their own gender identity and society’s norms, they may find it easier to make riskier choices than do women who are placed in a co-ed class. We hope in future research to investigate these hypotheses further.
References


Table 1: Personality Measures: The Big Five (Five Factor Model)

<table>
<thead>
<tr>
<th>Individual ...</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>is rude to others</td>
<td>Agreeableness</td>
</tr>
<tr>
<td>has a forgive nature</td>
<td></td>
</tr>
<tr>
<td>is considerate and kind</td>
<td></td>
</tr>
<tr>
<td>is talkative</td>
<td>Extraversion</td>
</tr>
<tr>
<td>is outgoing sociable</td>
<td></td>
</tr>
<tr>
<td>is reserved</td>
<td></td>
</tr>
<tr>
<td>worries a lot</td>
<td>Neuroticism</td>
</tr>
<tr>
<td>get nervous easily</td>
<td></td>
</tr>
<tr>
<td>is relaxed</td>
<td></td>
</tr>
<tr>
<td>does a thorough job</td>
<td>Conscientiousness</td>
</tr>
<tr>
<td>tends to be lazy</td>
<td></td>
</tr>
<tr>
<td>does things efficiently</td>
<td></td>
</tr>
<tr>
<td>is original, with new ideas</td>
<td>Openness</td>
</tr>
<tr>
<td>values the art</td>
<td></td>
</tr>
<tr>
<td>has an active imagination</td>
<td></td>
</tr>
</tbody>
</table>
Table 2a: Means of the Two Sample Groups
Difference in means between the whole sample and the sample used in the analysis

<table>
<thead>
<tr>
<th></th>
<th>Initial Sample N=486</th>
<th>Sample used N=231</th>
<th>Difference in Means (1)-(2)=(3)</th>
<th>St. error of the difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.35</td>
<td>0.34</td>
<td>0.01</td>
<td>0.04</td>
</tr>
<tr>
<td>Raw score IQ test</td>
<td>11.76</td>
<td>11.75</td>
<td>0.00</td>
<td>0.25</td>
</tr>
<tr>
<td>Z-score IQ test</td>
<td>0.01</td>
<td>0.01</td>
<td>0.00</td>
<td>0.08</td>
</tr>
<tr>
<td>Below age 20</td>
<td>0.72</td>
<td>0.75</td>
<td>-0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>Economics Degree</td>
<td>0.45</td>
<td>0.48</td>
<td>-0.03</td>
<td>0.04</td>
</tr>
<tr>
<td>Coeducational School</td>
<td>0.63</td>
<td>0.64</td>
<td>-0.01</td>
<td>0.04</td>
</tr>
</tbody>
</table>

**Composition of the class**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Only boys</td>
<td>0.22</td>
<td>0.27</td>
<td>-0.05</td>
<td>0.03</td>
</tr>
<tr>
<td>Only girls</td>
<td>0.21</td>
<td>0.19</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>Female Teacher</td>
<td>0.37</td>
<td>0.41</td>
<td>-0.04</td>
<td>0.04</td>
</tr>
</tbody>
</table>

**Region of origin**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td>0.05</td>
<td>0.04</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>English speaking</td>
<td>0.36</td>
<td>0.32</td>
<td>0.04</td>
<td>0.04</td>
</tr>
<tr>
<td>Asia</td>
<td>0.06</td>
<td>0.06</td>
<td>0.00</td>
<td>0.02</td>
</tr>
<tr>
<td>Asia-Europe</td>
<td>0.02</td>
<td>0.03</td>
<td>-0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>China</td>
<td>0.14</td>
<td>0.16</td>
<td>-0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>East-Europe</td>
<td>0.19</td>
<td>0.20</td>
<td>-0.02</td>
<td>0.03</td>
</tr>
<tr>
<td>Europe</td>
<td>0.17</td>
<td>0.18</td>
<td>-0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>Region Missing</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**Personality Dimensions**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Agreeableness</td>
<td>11.88</td>
<td>11.84</td>
<td>0.04</td>
<td>0.30</td>
</tr>
<tr>
<td>2. Conscientiousness</td>
<td>12.68</td>
<td>12.72</td>
<td>-0.04</td>
<td>0.34</td>
</tr>
<tr>
<td>3. Extraversion</td>
<td>12.60</td>
<td>12.72</td>
<td>-0.12</td>
<td>0.32</td>
</tr>
<tr>
<td>4. Neuroticism</td>
<td>11.36</td>
<td>11.46</td>
<td>-0.11</td>
<td>0.31</td>
</tr>
<tr>
<td>5. Openness</td>
<td>13.52</td>
<td>13.32</td>
<td>0.19</td>
<td>0.38</td>
</tr>
<tr>
<td>Dimension 1 missing</td>
<td>0.07</td>
<td>0.06</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Dimension 2 missing</td>
<td>0.09</td>
<td>0.07</td>
<td>0.01</td>
<td>0.02</td>
</tr>
<tr>
<td>Dimension 3 missing</td>
<td>0.08</td>
<td>0.06</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Dimension 4 missing</td>
<td>0.07</td>
<td>0.05</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>Dimension 5 missing</td>
<td>0.07</td>
<td>0.06</td>
<td>0.02</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Observations 486 231

Table 2b: Number of risky choices per session

<table>
<thead>
<tr>
<th></th>
<th>Session 1 Mean/St.Dev</th>
<th>Session 2 Mean/St.Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of risky choices</td>
<td>11.44</td>
<td>12.40</td>
</tr>
<tr>
<td></td>
<td>(3.92)</td>
<td>(4.25)</td>
</tr>
</tbody>
</table>

Observations 231 231

Notes: Sample size restricted to individuals answering the risk questions in both sessions and providing valid responses to control variables.
Table 3: Individuals present in both sessions, number of risky choices

<table>
<thead>
<tr>
<th></th>
<th>Session 1 (1)</th>
<th>Session 2 (2)</th>
<th>Session 2 (3)</th>
<th>Session 2 (4)</th>
<th>Fixed Effects (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>-1.43**</td>
<td>-1.20*</td>
<td>-2.83***</td>
<td>-2.72***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(0.65)</td>
<td>(0.78)</td>
<td>(0.81)</td>
<td></td>
</tr>
<tr>
<td>Z-score IQ test</td>
<td>0.48**</td>
<td>0.50**</td>
<td>0.44*</td>
<td>0.44*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.24)</td>
<td>(0.22)</td>
<td>(0.24)</td>
<td></td>
</tr>
<tr>
<td>Class comprises only men</td>
<td>0.54</td>
<td>0.58</td>
<td>-0.25</td>
<td>-0.10</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.71)</td>
<td>(0.60)</td>
<td>(0.74)</td>
<td>(0.56)</td>
</tr>
<tr>
<td>Class comprises only women</td>
<td>0.16</td>
<td>0.15</td>
<td>2.43**</td>
<td>2.48**</td>
<td>2.09***</td>
</tr>
<tr>
<td></td>
<td>(1.04)</td>
<td>(0.99)</td>
<td>(0.93)</td>
<td>(0.91)</td>
<td>(0.77)</td>
</tr>
<tr>
<td>Female Teacher</td>
<td>-0.19</td>
<td>-0.06</td>
<td>0.70</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.66)</td>
<td>(0.45)</td>
<td>(0.46)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>231</td>
<td>231</td>
<td>231</td>
<td>231</td>
<td>462</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.029</td>
<td>0.032</td>
<td>0.038</td>
<td>0.038</td>
<td>0.037</td>
</tr>
</tbody>
</table>

Notes:
(i) Other controls included in specification (1) to (4) include dummies for region-of-origin, age less than 20 years, studying for a degree in economics, and educated at a coeducational high school.
(ii) Robust standard errors, clustered at a module and class level, are in parentheses.
(iii) Columns (2) and (4) include the big five personality dimension. The Big Five personality dimensions: openness to experience, conscientiousness, extraversion, agreeableness and neuroticism.
(iv) *, ** and *** denote 10%, 5% and 1% levels of significance respectively.
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 £3 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.

<table>
<thead>
<tr>
<th>Column A</th>
<th>Column B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) £0.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>2) £1.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>3) £2.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>4) £3.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>5) £4.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>6) £5.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>7) £6.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>8) £7.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>9) £8.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>10) £9.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>11) £10.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>12) £11.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>13) £12.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>14) £13.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>15) £14.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>16) £15.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>17) £16.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>18) £17.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>19) £18.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
<tr>
<td>20) £19.00 for sure or 50% chance of winning £30 and a 50% chance of getting £0</td>
<td></td>
</tr>
</tbody>
</table>
### Table B.2: Difference in Means across Class-types

<table>
<thead>
<tr>
<th>Class composition</th>
<th>All-girls’ classes</th>
<th>All-boys’ classes</th>
<th>Coeducational classes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Initial Sample (1)</td>
<td>Sample used (2)</td>
<td>Difference (1) - (2) = (3)</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw score IQ test</td>
<td>11.48</td>
<td>11.96</td>
<td>-0.48</td>
</tr>
<tr>
<td>Z-score IQ test</td>
<td>-0.07</td>
<td>0.08</td>
<td>-0.15</td>
</tr>
<tr>
<td>Risky choices (1st)</td>
<td>10.01</td>
<td>10.42</td>
<td>-0.41</td>
</tr>
<tr>
<td>Below age 20</td>
<td>0.81</td>
<td>0.84</td>
<td>-0.03</td>
</tr>
<tr>
<td>Economics Degree</td>
<td>0.47</td>
<td>0.44</td>
<td>0.02</td>
</tr>
<tr>
<td>Coeducational School</td>
<td>0.62</td>
<td>0.73</td>
<td>-0.11</td>
</tr>
<tr>
<td>Class composition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only boys</td>
<td>1.00</td>
<td>1.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Only girls</td>
<td>0.00</td>
<td>0.24</td>
<td>0.00</td>
</tr>
<tr>
<td>Female Teacher</td>
<td>0.25</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>Region of origin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td>0.08</td>
<td>0.07</td>
<td>0.01</td>
</tr>
<tr>
<td>English speaking</td>
<td>0.32</td>
<td>0.20</td>
<td>0.12</td>
</tr>
<tr>
<td>Asia</td>
<td>0.04</td>
<td>0.07</td>
<td>-0.03</td>
</tr>
<tr>
<td>Asia-Europe</td>
<td>0.01</td>
<td>0.00</td>
<td>0.01</td>
</tr>
<tr>
<td>China</td>
<td>0.13</td>
<td>0.18</td>
<td>-0.05</td>
</tr>
<tr>
<td>East-Europe</td>
<td>0.23</td>
<td>0.27</td>
<td>-0.04</td>
</tr>
<tr>
<td>Europe</td>
<td>0.20</td>
<td>0.22</td>
<td>-0.02</td>
</tr>
<tr>
<td>Region Missing</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Personality Dimensions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Agreeableness</td>
<td>11.82</td>
<td>11.76</td>
<td>0.07</td>
</tr>
<tr>
<td>2. Conscientiousness</td>
<td>12.35</td>
<td>12.60</td>
<td>-0.25</td>
</tr>
<tr>
<td>3. Extraversion</td>
<td>12.45</td>
<td>12.40</td>
<td>0.05</td>
</tr>
<tr>
<td>4. Neuroticism</td>
<td>11.89</td>
<td>12.02</td>
<td>-0.13</td>
</tr>
<tr>
<td>5. Openness</td>
<td>13.05</td>
<td>13.18</td>
<td>-0.13</td>
</tr>
<tr>
<td>Dimension 1 missing</td>
<td>0.06</td>
<td>0.04</td>
<td>0.01</td>
</tr>
<tr>
<td>Dimension 2 missing</td>
<td>0.11</td>
<td>0.07</td>
<td>0.04</td>
</tr>
<tr>
<td>Dimension 3 missing</td>
<td>0.10</td>
<td>0.04</td>
<td>0.05</td>
</tr>
<tr>
<td>Dimension 4 missing</td>
<td>0.07</td>
<td>0.04</td>
<td>0.02</td>
</tr>
<tr>
<td>Dimension 5 missing</td>
<td>0.08</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>Observations</td>
<td>101</td>
<td>45</td>
<td></td>
</tr>
</tbody>
</table>
Table B.2: Probit Estimates of the probability of participating in session 2

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th></th>
<th>(2)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Session 1</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of risky choices</td>
<td>0.01 (0.02)</td>
<td>0.01 (0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Individual Variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-0.14 (0.13)</td>
<td>0.10 (0.19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Z-score IQ test</td>
<td>-0.02 (0.06)</td>
<td>-0.02 (0.06)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Below age 20</td>
<td>0.16 (0.15)</td>
<td>0.15 (0.15)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Economics Degree</td>
<td>0.12 (0.12)</td>
<td>0.13 (0.12)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coeducational School</td>
<td>0.02 (0.13)</td>
<td>0.02 (0.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single Sex Classes</td>
<td>0.22* (0.13)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Composition of the class</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Only boys</td>
<td></td>
<td></td>
<td>0.40** (0.16)</td>
<td></td>
</tr>
<tr>
<td>Only girls</td>
<td></td>
<td></td>
<td>-0.08 (0.21)</td>
<td></td>
</tr>
<tr>
<td>Female Teacher</td>
<td>0.15 (0.13)</td>
<td>0.14 (0.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Region of origin</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Africa</td>
<td>0.09 (0.29)</td>
<td>0.13 (0.29)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia</td>
<td>0.26 (0.27)</td>
<td>0.24 (0.27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asia-Europe</td>
<td>1.09*** (0.50)</td>
<td>1.05** (0.50)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>0.41* (0.20)</td>
<td>0.38* (0.20)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>East-Europe</td>
<td>0.31* (0.18)</td>
<td>0.29 (0.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>0.36** (0.18)</td>
<td>0.35** (0.18)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Region Missing</td>
<td>-0.29 (0.70)</td>
<td>-0.38 (0.71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Personality Dimensions</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Agreeableness</td>
<td>-0.03 (0.03)</td>
<td>-0.03 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Conscientiousness</td>
<td>-0.00 (0.03)</td>
<td>-0.00 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Extraversion</td>
<td>0.02 (0.03)</td>
<td>0.02 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Neuroticism</td>
<td>0.00 (0.02)</td>
<td>0.00 (0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Openness</td>
<td>-0.03 (0.02)</td>
<td>-0.03 (0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimension 1 missing</td>
<td>6.09*** (0.89)</td>
<td>6.11 (.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimension 2 missing</td>
<td>-0.18 (0.45)</td>
<td>-0.16 (0.45)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimension 3 missing</td>
<td>-0.92 (0.57)</td>
<td>-0.94 (0.58)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimension 4 missing</td>
<td>-5.37 (.)</td>
<td>-5.41*** (0.90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dimension 5 missing</td>
<td>-0.48 (0.74)</td>
<td>-0.52 (0.75)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-0.05 (0.51)</td>
<td>0.01 (0.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>486</td>
<td></td>
<td>486</td>
<td></td>
</tr>
</tbody>
</table>

Notes: a. Column (1) includes the main effect of being in a single sex class. b. Column (2) uses two different variables to differentiate the effect of male-only classes from female-only classes. Coefficients of the probit specification are reported. The category excluded for region of origin English-speaking countries. Robust standard errors in parenthesis. *, ** and *** denote 10%, 5% and 1% levels of significance respectively.
Table B.3: Difference in means between students attending single sex and coeducational classes

<table>
<thead>
<tr>
<th></th>
<th>Single Sex Class (1)</th>
<th>Coed Class (2)</th>
<th>Difference in Means (1)-(2)=(3)</th>
<th>St. error of the Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of risky choices</td>
<td>11.40</td>
<td>11.48</td>
<td>-0.07</td>
<td>0.52</td>
</tr>
</tbody>
</table>

**Individual Variables**

<table>
<thead>
<tr>
<th></th>
<th>0.42</th>
<th>0.27</th>
<th>0.15**</th>
<th>0.06</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Raw score IQ test</td>
<td>11.35</td>
<td>12.10</td>
<td>-0.76*</td>
<td>0.42</td>
</tr>
<tr>
<td>Below age 20</td>
<td>0.81</td>
<td>0.70</td>
<td>0.11*</td>
<td>0.06</td>
</tr>
<tr>
<td>Economics Degree</td>
<td>0.49</td>
<td>0.47</td>
<td>0.02</td>
<td>0.07</td>
</tr>
<tr>
<td>Coeducational School</td>
<td>0.70</td>
<td>0.58</td>
<td>0.12*</td>
<td>0.06</td>
</tr>
<tr>
<td>Female Teacher</td>
<td>0.28</td>
<td>0.52</td>
<td>-0.24***</td>
<td>0.06</td>
</tr>
</tbody>
</table>

**Region of origin**

<table>
<thead>
<tr>
<th></th>
<th>0.06</th>
<th>0.03</th>
<th>0.02</th>
<th>0.03</th>
</tr>
</thead>
<tbody>
<tr>
<td>Africa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English speaking</td>
<td>0.30</td>
<td>0.35</td>
<td>-0.05</td>
<td>0.06</td>
</tr>
<tr>
<td>Asia</td>
<td>0.06</td>
<td>0.06</td>
<td>-0.00</td>
<td>0.03</td>
</tr>
<tr>
<td>Asia-Europe</td>
<td>0.00</td>
<td>0.06</td>
<td>-0.06**</td>
<td>0.02</td>
</tr>
<tr>
<td>China</td>
<td>0.21</td>
<td>0.13</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td>East-Europe</td>
<td>0.21</td>
<td>0.19</td>
<td>0.02</td>
<td>0.05</td>
</tr>
<tr>
<td>Europe</td>
<td>0.17</td>
<td>0.19</td>
<td>-0.02</td>
<td>0.05</td>
</tr>
<tr>
<td>Region Missing</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Global Gender Gap Index 2010</td>
<td>0.70</td>
<td>0.71</td>
<td>-0.01*</td>
<td>0.01</td>
</tr>
</tbody>
</table>

**Personality Dimensions**

<table>
<thead>
<tr>
<th></th>
<th>11.88</th>
<th>11.80</th>
<th>0.08</th>
<th>0.48</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.Agreeableness</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.Conscientiousness</td>
<td>12.78</td>
<td>12.67</td>
<td>0.11</td>
<td>0.54</td>
</tr>
<tr>
<td>3.Extraversion</td>
<td>12.61</td>
<td>12.82</td>
<td>-0.22</td>
<td>0.51</td>
</tr>
<tr>
<td>4.Neuroticism</td>
<td>11.63</td>
<td>11.32</td>
<td>0.30</td>
<td>0.48</td>
</tr>
<tr>
<td>5.Openness</td>
<td>13.18</td>
<td>13.45</td>
<td>-0.27</td>
<td>0.61</td>
</tr>
<tr>
<td>Dimension 1 missing</td>
<td>0.06</td>
<td>0.06</td>
<td>-0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>Dimension 2 missing</td>
<td>0.09</td>
<td>0.06</td>
<td>0.04</td>
<td>0.03</td>
</tr>
<tr>
<td>Dimension 3 missing</td>
<td>0.07</td>
<td>0.05</td>
<td>0.03</td>
<td>0.03</td>
</tr>
<tr>
<td>Dimension 4 missing</td>
<td>0.06</td>
<td>0.05</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>Dimension 5 missing</td>
<td>0.06</td>
<td>0.06</td>
<td>-0.00</td>
<td>0.03</td>
</tr>
</tbody>
</table>

Observations: 107 124
### Table C.1: Individuals present in both sessions, number of risky choices (full set of estimates for specifications presented in Table 3 in the text)

<table>
<thead>
<tr>
<th></th>
<th>Session 1</th>
<th>Session 2</th>
<th>FE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Female</td>
<td>-1.43**</td>
<td>-1.20*</td>
<td>-2.83***</td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(0.65)</td>
<td>(0.78)</td>
</tr>
<tr>
<td>Z-score IQ test</td>
<td>0.48**</td>
<td>0.50**</td>
<td>0.44*</td>
</tr>
<tr>
<td></td>
<td>(0.24)</td>
<td>(0.24)</td>
<td>(0.22)</td>
</tr>
<tr>
<td>Below age 20</td>
<td>-0.87</td>
<td>-0.83</td>
<td>-0.73</td>
</tr>
<tr>
<td></td>
<td>(0.63)</td>
<td>(0.62)</td>
<td>(0.59)</td>
</tr>
<tr>
<td>Economics Degree</td>
<td>-1.06*</td>
<td>-1.14*</td>
<td>-0.32</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
<td>(0.61)</td>
<td>(0.55)</td>
</tr>
<tr>
<td>Coeducational School</td>
<td>0.02</td>
<td>0.11</td>
<td>-0.34</td>
</tr>
<tr>
<td></td>
<td>(0.53)</td>
<td>(0.56)</td>
<td>(0.59)</td>
</tr>
<tr>
<td>Class comprises only men</td>
<td>0.54</td>
<td>0.58</td>
<td>-0.25</td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.71)</td>
<td>(0.60)</td>
</tr>
<tr>
<td>Class comprises only women</td>
<td>0.16</td>
<td>0.15</td>
<td>2.43**</td>
</tr>
<tr>
<td></td>
<td>(1.04)</td>
<td>(0.99)</td>
<td>(0.93)</td>
</tr>
<tr>
<td>Female Teacher</td>
<td>-0.19</td>
<td>-0.06</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.66)</td>
<td>(0.45)</td>
</tr>
<tr>
<td>Africa</td>
<td>-0.90</td>
<td>-1.22</td>
<td>-0.41</td>
</tr>
<tr>
<td></td>
<td>(0.97)</td>
<td>(0.86)</td>
<td>(1.47)</td>
</tr>
<tr>
<td>Asia</td>
<td>-2.08**</td>
<td>-2.66***</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(0.81)</td>
<td>(0.89)</td>
<td>(1.18)</td>
</tr>
<tr>
<td>Asia-Europe</td>
<td>-0.09</td>
<td>0.20</td>
<td>-3.36**</td>
</tr>
<tr>
<td></td>
<td>(1.17)</td>
<td>(1.24)</td>
<td>(1.34)</td>
</tr>
<tr>
<td>China</td>
<td>-0.37</td>
<td>-0.68</td>
<td>0.36</td>
</tr>
<tr>
<td></td>
<td>(0.86)</td>
<td>(0.86)</td>
<td>(1.25)</td>
</tr>
<tr>
<td>East-Europe</td>
<td>-0.77</td>
<td>-1.05</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>(0.72)</td>
<td>(0.72)</td>
<td>(0.80)</td>
</tr>
<tr>
<td>Europe</td>
<td>0.25</td>
<td>-0.25</td>
<td>0.52</td>
</tr>
<tr>
<td></td>
<td>(0.79)</td>
<td>(0.83)</td>
<td>(0.82)</td>
</tr>
<tr>
<td>1. Agreeableness</td>
<td>-0.06</td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.13)</td>
<td></td>
</tr>
<tr>
<td>2. Conscientiousness</td>
<td>-0.13</td>
<td>-0.15</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.12)</td>
<td>(0.10)</td>
<td></td>
</tr>
<tr>
<td>3. Extraversion</td>
<td>-0.04</td>
<td>-0.04</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.09)</td>
<td></td>
</tr>
<tr>
<td>4. Neuroticism</td>
<td>-0.09</td>
<td>-0.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.11)</td>
<td>(0.11)</td>
<td></td>
</tr>
<tr>
<td>5. Openness</td>
<td>0.14*</td>
<td>0.18*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
<td>(0.10)</td>
<td></td>
</tr>
<tr>
<td>Dimension 1 missing</td>
<td>-2.75</td>
<td>0.99</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(4.03)</td>
<td>(1.59)</td>
<td></td>
</tr>
<tr>
<td>Dimension 2 missing</td>
<td>-0.50</td>
<td>-0.95</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.85)</td>
<td>(2.02)</td>
<td></td>
</tr>
<tr>
<td>Dimension 3 missing</td>
<td>4.49</td>
<td>2.95</td>
<td></td>
</tr>
<tr>
<td>Dimension 4 missing</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>-----</td>
<td>-----</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(3.41)</td>
<td>(4.30)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-4.06</td>
<td>-5.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.64)</td>
<td>(5.85)</td>
<td></td>
</tr>
<tr>
<td>Dimension 5 missing</td>
<td>2.47*</td>
<td>1.73</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.22)</td>
<td>(1.34)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>13.30***</td>
<td>15.36***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.78)</td>
<td>(2.15)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>13.54***</td>
<td>15.40***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.92)</td>
<td>(2.06)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11.66***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.34)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>231</td>
<td>231</td>
<td></td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.029</td>
<td>0.032</td>
<td></td>
</tr>
</tbody>
</table>

Notes: Other controls included in specification (1) to (4) include dummies for country-of-origin, age less than 20 years, studying a degree in economics, past attendance at a coeducational high school. Robust standard errors, clustered at a module and class level, are in parentheses. Columns (2) and (4) include the big five personality dimension. The Big Five personality dimensions: openness to experiences, conscientiousness, extraversion, agreeableness and neuroticism. The dummy variables “Dimension x missing” are for cases with missing information. *, ** and *** denote 10%, 5% and 1% levels of significance respectively.
## APPENDIX C: Sensitivity analyses estimates (not for publication)

Table C.2: Interval Regression

<table>
<thead>
<tr>
<th></th>
<th>Dep. Variable: number of risky choices</th>
<th>Dep. Variable: row where student changed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Session 1</td>
<td>Session 2</td>
</tr>
<tr>
<td>Female</td>
<td>-1.42***</td>
<td>-2.75***</td>
</tr>
<tr>
<td></td>
<td>(0.55)</td>
<td>(0.81)</td>
</tr>
<tr>
<td>Z-score IQ test</td>
<td>0.49**</td>
<td>0.56**</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>Class comprises only men</td>
<td>0.50</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>(0.62)</td>
<td>(0.66)</td>
</tr>
<tr>
<td>Class comprises only women</td>
<td>0.20</td>
<td>2.45***</td>
</tr>
<tr>
<td></td>
<td>(0.88)</td>
<td>(0.91)</td>
</tr>
<tr>
<td>Female Teacher</td>
<td>-0.22</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
<td>(0.53)</td>
</tr>
<tr>
<td>Observations</td>
<td>231</td>
<td>229</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: (i) The intervals are built around the following rows (1,4), (5,8), (9,13), (14-17), (18,20). (ii) The number of observations between the first and the second session differ due to the presence of individuals choosing zero risky choices and that are not included in the interval analysis. (iii). By using the actual row where the student change for those students switching row only once, the number of observations reduced considerably, with a greater reduction for the first session. (iv) Other controls included in specification (1) to (4) include dummies for country-of-origin, age less than 20 years, studying a degree in economics, past attendance at a coeducational high school. (v) Robust standard errors clustered at a module and class level are in parentheses. (vi) Columns (2) and (4) include the big five personality dimension. The Big Five personality dimensions: openness to experiences, conscientiousness, extraversion, agreeableness and neuroticism. (vii) *, ** and *** denote 10%, 5% and 1% levels of significance respectively.
### Table C.3: Individuals present in both sessions, switching more than once

<table>
<thead>
<tr>
<th></th>
<th>Session 1</th>
<th></th>
<th>Session 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of risky choices</td>
<td>Average row</td>
<td>Number of risky choices</td>
<td>Average row</td>
</tr>
<tr>
<td>Female</td>
<td>-1.96***</td>
<td>-2.21***</td>
<td>-2.09***</td>
<td>-2.16***</td>
</tr>
<tr>
<td></td>
<td>(0.57)</td>
<td>(0.59)</td>
<td>(0.75)</td>
<td>(0.65)</td>
</tr>
<tr>
<td>Z-score IQ test</td>
<td>0.22</td>
<td>0.17</td>
<td>0.40**</td>
<td>0.39**</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.18)</td>
<td>(0.17)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>Class comprises only men</td>
<td>0.28</td>
<td>-0.18</td>
<td>0.27</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(0.56)</td>
<td>(0.64)</td>
<td>(0.64)</td>
</tr>
<tr>
<td>Class comprises only women</td>
<td>0.29</td>
<td>0.24</td>
<td>2.09**</td>
<td>1.88**</td>
</tr>
<tr>
<td></td>
<td>(0.77)</td>
<td>(0.64)</td>
<td>(0.91)</td>
<td>(0.81)</td>
</tr>
<tr>
<td>Female Teacher</td>
<td>0.17</td>
<td>0.19</td>
<td>0.62</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td>(0.62)</td>
<td>(0.59)</td>
<td>(0.54)</td>
<td>(0.49)</td>
</tr>
<tr>
<td>Observations</td>
<td>216</td>
<td>216</td>
<td>202</td>
<td>202</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.035</td>
<td>0.057</td>
<td>0.025</td>
<td>0.026</td>
</tr>
</tbody>
</table>

Notes: Other controls included in specification (1) to (4) include dummies for country-of-origin, age less than 20 years, studying a degree in economics, past attendance at a coeducational high school. Robust standard errors, clustered at a module and class level, are in parentheses. The Big Five personality dimensions are openness to experiences, conscientiousness, extraversion, agreeableness and neuroticism. *, ** and *** denote 10%, 5% and 1% levels of significance respectively.
Table C.4: Individuals present in both sessions, switching only once

<table>
<thead>
<tr>
<th></th>
<th>Session 1</th>
<th></th>
<th>Session 2</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number of risky choices</td>
<td>Row number</td>
<td>Number of risky choices</td>
<td>Row number</td>
</tr>
<tr>
<td>Female</td>
<td>-2.68***</td>
<td>-2.68***</td>
<td>-2.69***</td>
<td>-2.64***</td>
</tr>
<tr>
<td></td>
<td>(0.79)</td>
<td>(0.79)</td>
<td>(0.75)</td>
<td>(0.80)</td>
</tr>
<tr>
<td>Z-score IQ test</td>
<td>0.13</td>
<td>0.13</td>
<td>0.42**</td>
<td>0.56***</td>
</tr>
<tr>
<td></td>
<td>(0.26)</td>
<td>(0.26)</td>
<td>(0.18)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>Class comprises only men</td>
<td>-0.18</td>
<td>-0.18</td>
<td>0.27</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>(0.72)</td>
<td>(0.72)</td>
<td>(0.65)</td>
<td>(0.61)</td>
</tr>
<tr>
<td>Class comprises only women</td>
<td>0.29</td>
<td>0.29</td>
<td>2.82**</td>
<td>2.33**</td>
</tr>
<tr>
<td></td>
<td>(1.10)</td>
<td>(1.10)</td>
<td>(1.05)</td>
<td>(1.05)</td>
</tr>
<tr>
<td>Female Teacher</td>
<td>0.51</td>
<td>0.51</td>
<td>1.22**</td>
<td>1.17**</td>
</tr>
<tr>
<td></td>
<td>(0.80)</td>
<td>(0.80)</td>
<td>(0.50)</td>
<td>(0.50)</td>
</tr>
<tr>
<td>Observations</td>
<td>162</td>
<td>162</td>
<td>168</td>
<td>168</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.056</td>
<td>0.056</td>
<td>0.036</td>
<td>0.047</td>
</tr>
</tbody>
</table>

Notes: Other controls included in specification (1) to (4) include dummies for country-of-origin, age less than 20 years, studying a degree in economics, past attendance at a coeducational high school. Robust standard errors clustered at a module and class level are in parentheses. The Big Five personality dimensions are openness to experiences, conscientiousness, extraversion, agreeableness and neuroticism. *, ** and *** denote 10%, 5% and 1% levels of significance respectively.
Table C.5: USING DIFFERENT SCALES FOR THE IQ-SCORE

(a) Individuals present in both sessions, number of risky choices (using the raw score of the IQ test)

<table>
<thead>
<tr>
<th></th>
<th>Session 1 (1)</th>
<th>Session 2 (2)</th>
<th>Session 2 (3)</th>
<th>Session 2 (4)</th>
<th>Fixed Effects (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>-1.43***</td>
<td>-1.20*</td>
<td>-2.83***</td>
<td>-2.72***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.70)</td>
<td>(0.65)</td>
<td>(0.78)</td>
<td>(0.81)</td>
<td></td>
</tr>
<tr>
<td>Raw score IQ test</td>
<td>0.15**</td>
<td>0.16**</td>
<td>0.14*</td>
<td>0.14*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.07)</td>
<td>(0.08)</td>
<td></td>
</tr>
<tr>
<td>Class comprises only men</td>
<td>0.54</td>
<td>0.58</td>
<td>-0.25</td>
<td>-0.10</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.71)</td>
<td>(0.60)</td>
<td>(0.74)</td>
<td>(0.56)</td>
</tr>
<tr>
<td>Class comprises only women</td>
<td>0.16</td>
<td>0.15</td>
<td>2.43**</td>
<td>2.48**</td>
<td>2.09***</td>
</tr>
<tr>
<td></td>
<td>(1.04)</td>
<td>(0.99)</td>
<td>(0.93)</td>
<td>(0.91)</td>
<td>(0.77)</td>
</tr>
<tr>
<td>Female Teacher</td>
<td>-0.19</td>
<td>-0.06</td>
<td>0.70</td>
<td>0.77</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.66)</td>
<td>(0.45)</td>
<td>(0.46)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>231</td>
<td>231</td>
<td>231</td>
<td>231</td>
<td>462</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.029</td>
<td>0.032</td>
<td>0.038</td>
<td>0.038</td>
<td>0.037</td>
</tr>
</tbody>
</table>

Notes: Other controls included in specification (1) to (4) include dummies for country-of-origin, age less than 20 years, studying a degree in economics, past attendance at a coeducational high school. Robust standard errors clustered at a module and class level are in parentheses. Columns (2) and (4) include the big five personality dimension. The Big Five personality dimensions: openness to experiences, conscientiousness, extraversion, agreeableness and neuroticism. *, ** and *** denote 10%, 5% and 1% levels of significance respectively.
Table C.6:
(b) Individuals present in both sessions, number of risky choices (using the population norms for the IQ)

<table>
<thead>
<tr>
<th></th>
<th>Session 1 (1)</th>
<th>Session 2 (3)</th>
<th>Session 2 (4)</th>
<th>Fixed Effects (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>-1.41**</td>
<td>-2.82***</td>
<td>-2.70***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.69)</td>
<td>(0.78)</td>
<td>(0.81)</td>
<td></td>
</tr>
<tr>
<td>IQ percentile</td>
<td>0.01*</td>
<td>0.01</td>
<td>0.01*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
<td>(0.01)</td>
<td>(0.01)</td>
<td></td>
</tr>
<tr>
<td>Class comprises only men</td>
<td>0.52</td>
<td>-0.29</td>
<td>-0.13</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.60)</td>
<td>(0.75)</td>
<td>(0.56)</td>
</tr>
<tr>
<td>Class comprises only women</td>
<td>0.18</td>
<td>2.44**</td>
<td>2.49***</td>
<td>2.09***</td>
</tr>
<tr>
<td></td>
<td>(1.05)</td>
<td>(0.93)</td>
<td>(0.91)</td>
<td>(0.77)</td>
</tr>
<tr>
<td>Female Teacher</td>
<td>-0.18</td>
<td>0.69</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.61)</td>
<td>(0.45)</td>
<td>(0.46)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>231</td>
<td>231</td>
<td>231</td>
<td>462</td>
</tr>
<tr>
<td>Adjusted $R^2$</td>
<td>0.029</td>
<td>0.036</td>
<td>0.037</td>
<td>0.037</td>
</tr>
</tbody>
</table>

Notes: Other controls included in specification (1) to (4) include dummies for country-of-origin, age less than 20 years, studying a degree in economics, past attendance at a coeducational high school. Robust standard errors clustered at a module and class level are in parentheses. Columns (2) and (4) include personality dimension. The Big Five personality dimensions are openness to experiences, conscientiousness, extraversion, agreeableness and neuroticism. *, ** and *** denote 10%, 5% and 1% levels of significance respectively.
Appendix D:
Instruction Sheet for the First Class

This sheet will help you use the First Class PowerPoint slides, though, if the computer in your room is not working you should be able to do the class with just this sheet.

When you get to class, turn on the computer and make sure to load the PowerPoint slides that were designed for this class.
When the students arrive take attendance:
  o If a student is not on your attendance sheet have them write their name down CLEARLY.
  Make sure you can read that student’s name!
  Do not let the students talk to each other while they are doing these things

Once it seem like everyone is in the class (at least by 5 min after your class was suppose to begin) you can start the class.

1) Read through Slide Two:

   
   Introduction
   
   Please make sure you signed in on the attendance sheet.

   Today we will ask you to do three things:
   - Choose Matrices
   - Fill out one form
   - Add up two-digit numbers

   You should be done in 35min.

When you are done with Slide One hand out the Matrix Packets AND the Matrix Answer Sheets that you were given. MAKE SURE STUDENTS DO NOT WRITE ON THE MATRIX PACKETS!!!
After handing out the packets and answer sheets read slide three:

Matrices

You will get a packet of “matrices” and an answer sheet. Fill out your name, registration number, class, etc. on the answer sheet.

Do not write on the matrix packet.

We will go through two practice matrices together now. Do not look at the non-practice matrices before being told to do so!

When done with Go to slide four:

Slide four shows “Practice 1” on the handout. If your computer is not working just tell the students to look at “Practice 1” Matrix when you give the next instructions.

Tell the students to “Look at how there are patterns in the rows going down the matrix and patterns going across the matrix. Examine the matrix and try to find which of the 8 options completes the matrix.”

Then give the students 1 minute to figure out which answer is correct. After enough time tell them “OK. The answer was #8. Now to practice again we will look at the next Practice matrix, turn the page.”

Now go onto slide five:

Slide five shows “Practice 2” in the handout. If your computer is not working tell them to look at “Practice 2” Matrix. Then say “Now look at this matrix and the part that is missing. Examining the matrix up and down, and left to right, try to figure out which part of the matrix is missing.”

Give the students one minute and then say “The correct answer is #4.”
Now go to slide six and read it:

**Matrices**

Now you will have 20 min to do 16 matrices. 
Do not begin until I say go. 
Read the instructions on the answer sheet 
and make sure to clearly mark the matrix you 
believe to be correct.

Any questions?

OK… Go.

Read the instructions on the answer sheet “Place a single line across the number of the answer you 
think is correct. If you make a mistake, put a cross (X) through your incorrect answer and a signle line 
across the correct one. Please mark your choices firmly with a pencil. NOTE: Do NOT write on the test packet.”

If there are no questions say “Go” and MAKE SURE TO KEEP TRACK OF 20 MINUTES!!!

Once 20 Minutes has gone by say “STOP” and make sure all students finish. Then:

Collect matrix packets AND answer sheets. 
AFTER collecting the sheets back pass out the form entitled “Questionnaire for Class 1”

Then read slide 7:

**Questionnaire**

You will now have a questionnaire in front of you. 
Fill out your name, registration number, class information, etc. and read the instructions.

The Questionnaire has 20 rows. 
In each row there are two options: one in column A and one in Column B. 
In each row consider which option you prefer and circle that option.
When done reading slide 7 go onto slide 8 and read it:

**Questionnaire**

After you are finished choosing an option in each row the GTA will randomly choose a number between 1 and 20 and 1 in every 10 people (10% of the students in EC111/EC100) will be paid for their choice.

For example, say “1” gets randomly chosen:
- IF you chose ‘£0.00 for sure’ in row 1, you will get paid nothing.
- IF you chose ‘50% chance of winning £30 and a 50% chance of winning £0.00’ then the GTA will flip a coin and if it comes up heads you will get £30 if you are one of the 10% getting paid, otherwise you will get £0.00.

Any Questions?!

Go ahead and fill the form out.

If there are no questions then tell them to fill out the form. Once everyone has finished collect the form and then turn over this page and read the random number that I have written on the back of it. That is the number of the row that will be paid, tell that number to the students. Then take out the coin that is with these instruction and flip it. Tell the students if it comes up HEADS or TAILS. Write down the class number you are doing and then the result of the coin flip result Here:

<table>
<thead>
<tr>
<th>Class ____:</th>
<th>Class ____:</th>
<th>Class ____:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coin Flip Outcome</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Now go on to slide 9. Read it and answer any questions:

**Two-Digit Addition Rounds**

You will now be asked to add up two-digit numbers as in the example below:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>45</td>
<td>62</td>
<td>81</td>
<td>35</td>
</tr>
</tbody>
</table>

You will do two rounds and you will receive a payment for your performance based on how you do in each round.

- Payment will be different in each round.
- After you complete both rounds we will randomly choose a number 1 or 2 and you will be paid according to the rules in that round.

You will have 3 minutes to do as many problems as you can.

You cannot use a calculator!!

Any questions?

After you have answered any questions tell them that the answer to the above problem is 246.
Now pass out the sheet entitled “Two-Digit Addition: ROUND ONE” MAKE SURE NO ONE STARTS TO DO THE PROBLEMS!!!

Also hand out the scrap paper. The scrap paper has mazes on one side but tell them to ignore the mazes. Give each student 2 or 3 pieces of scrap paper.

Then read slide 10:

Two-Digit Round One

You now have a sheet of paper two-digit math problems in front of you.
Fill in your name, registration number, class information, etc.

In this round you will get paid £0.20 for each problem you solve correctly.
Example: If you solve 10 problems correctly you will get £2 if this round gets chosen for payment.
You should try to solve as many problems as you can in the 3 minutes you have.

Any questions?....

OK: Go!

If there are no questions tell the students they can begin. KEEP TRACK OF THREE MINUTES! WHEN THREE MINUTES ARE UP YELL “STOP!!” Make sure the students stop writing on the answer sheet.

Collect the ROUND ONE sheets and then give everyone 2 or 3 more sheets of scrap paper if they need it.

After you have collected all the sheets and given out scrap paper had out the sheet entitled “Two-Digit Addition: ROUND TWO” MAKE SURE NO ONE STARTS TO DO THE PROBLEMS!!!!

Now read slide 11:

Two-Digit Round Two

You now have another sheet of two-digit math problems.
Fill in your name, registration number, class information, etc.

In this round you will get paid £1 for each problem you solve correctly if you are in the top 20% of performers in this class.
Example:
Say there are 20 students in this class.
If you add up 10 problems correctly you will get £10 if this round gets chosen for payment AND you are one of the top 4 problem solvers in this class.
You should try to solve as many problems as you can in the 3 minutes you have.

Any questions?....

OK: Go!
After you have read through the slide and example answer any questions that may come up. Then say “GO” and KEEP TRACK OF THREE MINUTES!!! WHEN THREE MINUTES ARE UP YELL “STOP!!” Make sure the students stop writing on the answer sheet.

Once everyone is done collect the ROUND TWO SHEETS and then turn over this sheet. On the back of this sheet is the number 1 or 2 and that is the round that will be paid.

Finally go to the last slide and say THANK YOU.

Thank You!

Thank you for participating.

You will get paid by the end of next week for your performance.

MAKE SURE YOU HAVE ALL THE SHEETS AND INFORMATION BEFORE YOU LEAVE THE CLASS. AND THAT YOU HAVE WRITTEN DOWN THE COIN FLIP RESULT.
Class Meeting
Introduction

Please make sure you signed in on the attendance sheet.

Today we will ask you to do three things:
- Choose Matrices
- Fill out one form
- Add up two-digit numbers

You should be done in 35 min.
Matrices

You will get a packet of “matrices” and an answer sheet. Fill out your name, registration number, class, etc. on the answer sheet.

Do not write on the matrix packet.

We will go through two practice matrices together now. Do not look at the non-practice matrices before being told to do so!
Matrices

Now you will have 20 min to do 16 matrices. Do not begin until I say go. Read the instructions on the answer sheet and make sure to clearly mark the matrix you think is correct.

Any questions?

OK… Go.
You will now have a questionnaire in front of you. Fill out your name, registration number, class information, etc. and read the instructions.

The Questionnaire has 20 rows. In each row there are two options: one in column A and one in Column B. In each row consider which option you prefer and circle that option.
After you are finished choosing an option in each row the GTA will randomly choose a number between 1 and 20 and 1 in every 10 people (10% of the students in EC111/EC100) will be paid for their choice.

For example, say “1” gets randomly chosen:

- IF you chose ‘£0.00 for sure’ in row 1, you will get paid nothing.
- IF you chose ‘50% chance of winning £30 and a 50% chance of winning £0.00’ then the GTA will flip a coin and if it comes up heads you will get £30 if you are one of the 10% getting paid otherwise you will get £0.00.

Any Questions?!

Go ahead and fill the form out.
Two-Digit Addition Rounds

You will now be asked to add up two-digit numbers as in the example below:

<table>
<thead>
<tr>
<th>Numbers</th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>23</td>
<td>45</td>
<td>62</td>
<td>81</td>
</tr>
<tr>
<td>35</td>
<td></td>
<td></td>
<td>35</td>
</tr>
</tbody>
</table>

You will do two rounds and you will receive a payment for your performance based on how you do in each round.
- Payment will be different in each round.
- After you complete both rounds we will randomly choose a number 1 or 2 and you will be paid according to the rules in that round.

You will have 3 minutes to do as many problems as you can.
You cannot use a calculator!!

Any questions?
Two-Digit Round One

You now have a sheet of paper two-digit math problems in front of you. Fill in your name, registration number, class information, etc.

In this round you will get paid £0.20 for each problem you solve correctly. Example: If you solve 10 problems correctly you will get £2 if this round gets chosen for payment. You should try to solve as many problems as you can in the 3 minutes you have.

Any questions?....

OK: Go!
Two-Digit Round Two

You now have another sheet of two-digit math problems. Fill in your name, registration number, class information, etc.

In this round you will get paid £1 for each problem you solve correctly if you are in the top 20% of performers in this class. Example:

- Say there are 20 students in this class.
- If you add up 10 problems correctly you will get £10 if this round gets chosen for payment AND you are one of the top 4 problem solves in this class.

You should try to solve as many problems as you can in the 3 minutes you have.

Any questions?....
OK: Go!
Thank You!

Thank you for participating.

You will get paid by the end of next week for your performance.