Gender Composition in Teams: Differences between Exogenously and Endogenously Formed Teams

Ainoa Aparicio Fenoll, Flavia Coda-Moscarola, Sarah Zaccagni*

PRELIMINARY DRAFT

Abstract

In light of increasing female labor force participation, women's influence in working teams has increasingly received attention. Papers studying endogenously-formed and exogenously-formed teams find different results. We study how the effect of gender composition of teams on team performance, working style, and individual's satisfaction differs between endogenous and exogenous teams. We divide a sample of high school students into two groups: we assign students in one group to teams of varying gender composition and we allow the students in the other group to form teams freely. We find that students who choose their team are significantly more satisfied with the gender composition of their team and most students prefer gender-balanced teams. We also find that female-predominant teams underperform other teams but there are no differences between endogenous and exogenous teams. In contrast, there are differences in the effect of gender composition on team working style and individual's satisfaction for endogenous and exogenous teams.

*Aparicio Fenoll: University of Turin (ESOMAS Department), Corso Unione Sovietica 218 Bis, 10134 Torino, Italy (ainoa.apariciofenoll@unito.it); Coda-Moscarola: CeRP-Collegio Carlo Alberto and Ufficio Studi Compagnia di San Paolo (flavia.codamoscarola@carloalberto.org); Zaccagni: University of Copenhagen (Department of Economics), Øster Farimagsgade 5, Building 35, DK-1353 Copenhagen (Sarah.Zaccagni@econ.ku.dk). We are grateful to Ghazala Azmat, Giovanni Mastrobuoni, Ignacio Monzón, and Juan Morales for their useful comments.

JEL Classification: I21 ; I24 ; D04

Keywords: Team Composition; Gender; Team Formation; Team Dynamics; Team Performance

1 Introduction

Increasing female labor force participation has changed workplace demographics and created more gender-diverse teams. In several countries, these changes have been eased by the introduction of explicit gender quotas in professions such as politics or the corporate sector. With the growing representation of women on political committees, company boards and other group settings, understanding the effects of gender diversity is important for academics and policy makers. Differences in the gender composition of teams may have consequences for corporate and political outcomes.

Azmat and Petrongolo [2014] state that "One of the main problems with studying gender and groups is that groups are typically formed in an endogenous way. While experiments can go some way to solve this issue through random assignment into groups, they tend to create an artificial environment, in which it becomes difficult to distinguish group diversity and group dynamics". Papers that study the impact of the gender ratio on team performance using endogenously-formed and exogenously-formed teams find different results. In this paper, we explore how the gender composition of teams affects team outcomes differently for endogenous and exogenous groups.

We design an experiment in which we randomly assign students to two equally sized groups with the same proportion of females. In the first group, we assign students to teams with different gender compositions. We call this the group of exogenous teams. We let the second group form their groups endogenously. We then analyze the effects of the proportion of female students on satisfaction with the gender composition of the team, team performance, working dynamics, and individual's satisfaction with special emphasis on how these effects differ for endogenous and exogenous teams.

The experimental literature has largely documented gender differences in skills, psychological traits, individual attitudes and preferences (Kagel and Roth [2016] provides a comprehensive review of the studies on the topic). If different psychological traits lead men and women to make different choices in similar contexts, the gender composition of teams becomes a relevant factor in collective decision-making. Moreover, various studies have shown that individual behavior changes in the presence of people from the same or opposite sex (Gneezy and Rustichini [2004]; Antonovics, Arcidiacono, and Walsh [2009]; Ivanova-Stenzel and Kübler [2011]). According to the theory of Becker [1957], taste-based discrimination arises because some individuals receive a dis-utility when they work with women. This may affect individual satisfaction with the team, working dynamics, and also performance. By allowing individuals to endogenously form groups, individuals can self-select according to their gender preferences. As a result, the impact of gender composition on team outcomes may differ greatly across endogenous and exogenous groups.

The mathematics camp is addressed to high school students in grades one to four (ages fourteen to eighteen). Mathematics teachers select the top performing students in each class (the median is two students per class) to attend an intensive three-days mathematics camp at the end of the academic year. During the camp, students work in groups of six members (some teams are formed by five or seven for organizational purposes). Teams work on mathematics problems that are unrelated to the high school curricula using math manipulatives. Each team submits one solution to each of the problems and is evaluated. Hence, the camp is characterized by *peer-to-peer* learning, *"inquiry-oriented"* activities, and a *"hands-on"* learning style.

In an accompanying paper, we find that students participating in the mathematics camp improve their problem-solving skills. The improvement is higher in problems that require logic skills rather than problems that require formal mathematics knowledge (formulas, standard solving methods, etc.). Regarding personality traits, the camp leads to improvements in self-concept: it reduces the incidence of self-declared neuroticism and fosters extroversion among participants.

Recent papers that use field data find mixed evidence regarding how the gender composition of a team influences performance. Apesteguia, Azmat, and Iriberri [2012] use a large online business game to study how female presence on a team affects collective choices and performance. The game is played by teams of three, where each group takes the role of a general manager of a beauty-industry company, competing in a market composed of four other simulated companies. The analysis shows that teams composed of three women are significantly outperformed by any other gender combination. Differences in performance are explained by differences in decision-making: all-women teams are less aggressive in their pricing strategies, invest less in R&D, and invest more in social sustainability initiatives, than any other gender combination. In contrast, Hoogendoorn, Oosterbeek, and Van Praag [2013a] use a similar business game setting where Dutch university students are exogenously allocated to teams of different sizes (the average number of students is 12). They detect an inverse u-shaped relationship between the share of women and a team's business performance, such that teams with an equal gender mix perform best.

We find that students who choose their team are significantly more satisfied with the gender composition of their team and most students prefer gender-balanced teams. We also find that female predominant teams underperform other teams but there are no differences between endogenous and exogenous teams. In contrast, there are differences in the effect of gender composition on team working style and work and team satisfaction for endogenous and exogenous teams.

1.1 Related Literature

Recent reforms that mandate certain levels of female representation on boards of directors offer a valuable, quasi-experimental setting to study gender composition of teams and performance. One of the first countries to implement gender quota laws was Norway in 2003, followed by Spain, Finland, Iceland and France. To study the causal impact of female presence on boards on firm performance, Ahern and Dittmar [2012] exploit the gender quota reform in Norway, which requires listed companies to achieve 40 percent female board representation within two years. The passage of the law change required that approximately 30 percent of the members of an average board needed to change to be in compliance with the 40-percent quota. However, some firms had a greater proportion of female directors before the quota was imposed and, therefore, faced a smaller constraint than those with fewer women. Given variation in the share of female directors at baseline, the Norwegian setting offers a natural experiment for the study of the impact of gender quotas. Ahern and Dittmar [2012] document important effects of female board representation. In particular, they find that the constraints imposed by the quota implied a decline in stock prices and operating profits.

Understanding the channels through which group composition affects performance is clearly important to design appropriate policy intervention. As for overall gender inequalities, such channels broadly boil down to gender differences in human capital and preferences. Ahern and Dittmar [2012] find that the quota setting in Norway led to less experienced boards, as new female directors had substantially less CEO experience and were younger than existing male directors. However, other channels such as gender differences in preferences or group strategic interactions may not be a priori discarded in Ahern and Dittmarâs study. Matsa and Miller [2013] explicitly study the effects of the Norwegian gender quota on corporate decision-making and conclude that the quota has changed the style of corporate leadership. They compare publicly-listed firms in Norway with a matched sample of unlisted firms in Norway, and listed and unlisted firms in other Nordic countries, and find that most corporate decisions were unaffected after womenâs board representation increased, but note sizeable differences in firmsâ employment policies. In particular, firms affected by the quota undertook fewer employee layoffs, causing an increase in relative labor costs. The results suggest that female directors consider labor hoarding a more profitable long-run strategy, or that they have a greater concern for workersâ vulnerability to unemployment risk. As the gender quota did not lead to less profitable decisions overall, but only changes in human resource management, Matsa and Miller conclude that changes in corporate strategy may not be explained by less experienced boards.

Bagues and Esteve-Volart [2010] provide further quasi-experimental evidence on the relevance of gender in teams, by exploiting the random allocation of applicants to selection committees in the Spanish civil service. They analyze the chances of success of male and female candidates to the Spanish judiciary system and find that female candidates are less likely to be hired whenever they are randomly assigned to a committee with a stronger female presence. Further evidence suggests two potential driving mechanisms. First, female evaluators tend to overestimate the quality of male candidates. Second, the presence of women on committees induces male members to increasingly favor male candidates.

Delfgaauw, Dur, Sol, and Verbeke [2013] provide novel field evidence on gender interactions and firm performance. They run an experiment in a Dutch retail chain to study whether firm performance under competition is affected by the gender of the manager and/or by the gender composition of employees. The experiment introduces short-term sales competition in a random subset of stores, and tournament compensation is offered within sets of treated stores. On average, the introduction of the tournament increases sales growth by five percentage points, and this effect is independent of the managerâs gender or the share of female employees. However, the effect of the tournament on sales growth is highest in stores in which the manager and a sufficiently high fraction of employees are of the same gender.

The interaction between gender composition and the altruistic behavior of groups is analyzed in a dictatorâs game setting by Dufwenberg and Muren [2006]. In their game, groups of three people with varying female presence divide a sum of money among themselves and a fourth person. The authors find that female-majority groups give more to the individual recipients and choose the egalitarian division more often than male-majority groups do. They also find that the most generous groups are those with two women and one man, suggesting that the presence of a man in the group triggers an enhanced generosity in the group.

Recent papers that use field data have also found evidence that the gender composition of a team does influence performance. Apesteguia et al. [2012] use a large online business game to study how female presence on a team affects collective choices and performance. The game is played by teams of three, where each team takes the role of a general manager of a beauty-industry company, competing in a market composed of four other simulated companies. The analysis shows that teams composed of three women are significantly outperformed by any other gender combination. Differences in performance are explained by differences in decision-making: all-women teams are less aggressive in their pricing strategies, invest less in R&D, and invest more in social sustainability initiatives, than any other gender combination. Hoogendoorn et al. [2013a] use a similar business game setting with Dutch university students, in which students are exogenously allocated to teams, and detect an inverse u-shaped relationship between the share of women and a teamâs business performance, such that teams with an equal gender mix perform best. In this paper, we explore how the gender composition of teams involved in a mathematics camp affects team dynamics and performance differently for endogenous and exogenous groups.

The remainder of this paper is organized as follows. In Section 2, we describe the mathematics camp. We provide details on the design of our randomized control trial in Section 3. We describe our data in Section 4. In section 5, we present the results of the randomization and describe the empirical strategy. We discuss the results of our estimates in Section 6. Finally section 7 concludes. The Appendix includes the questionnaire.

2 Background

The Mathesis Mathematics Camp involves students from more than forty different high-schools in Turin (a province in the North of Italy). It is organized by Mathesis, the association of math high-school teachers in the Italian region of Piedmont [1]. The summer camp has been organized yearly since 1995. It takes place some weeks before the end of the academic year (last days of May and early June). The camp lasts three days, in which participants work away from classrooms, in the Olympic Village in Bardonecchia, a touristic resort in a mountain location near Turin. The mathematics camp aims to enhance excellence in mathematics. In each edition around one thousand students from the first to the fourth grades of high-schools participate. They are followed by 120 high-school professors, 6 professors from the Department of Mathematics of the University of Turin, 20 undergraduate students in Mathematics (with didactic specialization), and 8 recent graduates in Mathematics. Due to location capacity constraints, the students are divided into 4 waves. In each wave a different set of schools participate and students from each grade are equally represented.

The camp is characterized by its learning mode and mathematics contents. Students are introduced to traditional mathematics concepts (infinite, series, etc.) which are illus-

[1] Further information on the Mathesis association can be found at http://www.associazionesubalpinamathesis.it/en/

trated through applications to topics of great actuality (QR codes, crypto currencies, etc.). The learning mode is based on “inquiry-oriented” activities and a “hands-on” problem solving methodology: support for the reasoning is given by manipulatives (for instance, geometrical shapes built with recycled materials) which allow for a more effective understanding of the theoretical concepts. Complex problems and mathematical games are proposed in order to urge the students to present original solutions and strategies in a climate of playful competition.

During the mathematics camp students work in open spaces, one for each of the four grades. Students work in teams of 6 individuals (in the cases in which the total number of students is not multiple of 6, some teams have 5 or 7 students). Each team must submit one common solution to the proposed mathematical problems. The composition of the team remains invariant along the entire duration of the camp. In a regular year, students are casually placed, avoiding students from the same school in the same team. We change this approach in the 2019 edition for the purposes of our study.

At the end of the mathematics camp students are involved in a ”Treasure hunt” on the topics covered in the camp with the dual purpose of verifying the activities carried out and concluding in a joyful environment. Prices are awarded to teams who have performed particularly well during the camp and in the treasure hunt.

The Camp is sponsored by Compagnia di San Paolo, the largest bank foundation in Italy. Students pay a small contribution (around ninety euros) which covers only a small part of the total cost. Some schools pay the student contribution. On several occasions, the initiative was supported by the financial contribution of local Institutions.

3 Experiment Design

We split each group of students in the same grade who participate in the camp in each of the four waves in half. We form teams for the first subgroup. The composition of these teams follows two criteria: obtaining teams with different gender compositions and avoiding teams with students belonging to the same school. Each team was composed of six students (for logistical reasons some teams had five or seven students range). For the second subgroup, we let students form teams freely, asking them to avoid students belonging to the same school.

In practice, we provide the teachers in charge of organizing the teams with one sheet of paper per team. Each sheet includes a heading with the number of the team and seven spaces to be filled with the names of the team components. For the exogenous teams, we filled the spaces with the names of the components in random order. Teachers asked these students to take a seat first. The teachers allowed the remaining half of students to form teams endogenously. After all teams are built, students in endogenous teams report their names in the corresponding sheet of paper.

The precise instructions read by teachers upon students arrival to the open-space are as follows: "Dear students, we are carrying out an experiment in the field to optimize the functioning of the stage, and to this end, we kindly ask for your cooperation. We propose you a game. We now assign a table to some of you, we ask the others to wait for a moment. We now read the composition of the tables that have already been assigned, and when you hear your name, please go and sit at the established table." After all students have taken a seat, teachers kept reading: "Those who have remained standing are going to sit now at the remaining tables forming X groups of 6 students and Z groups of five (or seven) students (on the tables there is a sheet with how many components must sit there). We will start a timer that will last six minutes. Before time runs out, you all need to be seated down. The only condition that you must respect is the following: You *cannot* sit down with students from your school. Ready, set, go!"

The students' activity during the mathematics camp is evaluated in at least five distinct stages. The number of moments of intermediate evaluation varies according to the grade. Teachers are responsible for the evaluation: they report it on a board that students can consult at any time during the three days. For each intermediate test, the evaluator ranks the groups by assigning the first classified a score equal to the number of teams, and a decreasing score for the other teams (eg. If there are 20 teams, on the first come a score of 20 will be given, per second ranked 19, third 18 ... etc). The score in the intermediate evaluation is assigned based on the appropriateness of the answer formulated, the originality in the execution of the work, the speed and any other dimension the teachers consider interesting in order to return an exhaustive overview of performance. The last test consists of a treasure hunt lasting two hours at the end of the last day of the mathematics camp. Students are asked to answer some questions, only by giving the right answer is it possible to move to the next question. The winning team is the first to provide right answers to all questions. At the end of the two hours, the teachers draw up the ranking of the teams by referring to the number of correct answers provided. The final score deciding the winning team is determined by calculating a weighted average where twenty percent is the sum of scores obtained during the intermediate stages and eighty percent is the treasure hunt score.

At the end of the activities, students were given an electronic questionnaire to be filled in on their respective devices (often mobile phones). In the rare event in which one student does not have a mobile phone, he/she can borrow it from other student or the teacher. The specific questions asked in the questionnaire are reported in A. Teachers have the responsibility to control that students send the questionnaires, carrying out this check on the bus that takes students back home.

4 Data and Descriptive Statistics

Our sample is composed of students who participated in the camp and answered the questionnaire administered at the end of the camp. They are 890 students, among which 461 in the group of endogenous teams and 429 in the group of exogenous teams.

Table 1 reports mean, standard deviation, minimum and maximum values of the variables that capture team composition. As explained above, slightly more than half of students belong to endogenous teams. We define as male-predominant teams those teams where more than two thirds of components are male. They are 32% of teams. Similarly, more than two thirds of components are female in female-predominant teams. They are slightly more than 27% of all teams. Among endogenous teams, male-predominant are 27% while female-predominant are 23%. Regarding students' satisfaction with the gender composition of their team, 10% of students would have preferred more males in their teams while 19% of students would have preferred more females in their team.

Table 1: Team Composition

Variable	Mean	Std. Dev.	Min.	Max.	N
Endogenous teams	0.518	0.5	0	1	890
Male-predominant	0.317	0.466	0	1	890
Female-predominant	0.272	0.445	0	1	890
Endogenous male-predominant	0.134	0.341	0	1	890
Endogenous female-predominant	0.115	0.319	0	1	890
More male	0.096	0.294	0	1	890
As it was	0.716	0.451	0	1	890
More female	0.189	0.392	0	1	890

Table 2 shows descriptive statistics for the different outcomes used in the analysis. The standard score is the result of standardizing the raw test score to have mean zero and standard deviation equal one. The average team is in position number nine in the ranking of its grade and wave. As different grades and waves may have different number of teams, we standardized the ranking dividing it by the total number of teams. Students are overall satisfied with the camp: the average score assigned to the camp is more than eight over ten. A similar opinion is declared on team performance: the average student believes the performance of his team deserves a score of eight out of ten. Also the working environment within the team deserves high scores. The average score is slightly below nine. On average, students consider that their individual contribution to the team can be ranked eight out of ten. The average student considers that his/her knowledge, brightness and effort with respect to the rest of members of the team can be evaluated with seven out of ten.

The individual characteristics of our sample of students are summarized in Table 3. There are relatively more students from lower grades although the distribution is fairly balanced across grades. There is almost the same number of males and females. The average student has one sibling while the maximum number of siblings is five. Students' fathers are relativey well educated: 41% of students have a high-school graduated father and 43% of students have a university graduated father. The level of education of mothers is even higher: 43% of mothers are high-school grads while as much as 48% of mothers have a university degree. As expected, given the selection process to participate in the

Table 2: Outcomes

Variable	Mean	Std. Dev.	Min.	Max.	N
Standard score	0.004	1.006	-2.653	2.19	890
Ranking	8.801	5.253	1	20	890
Standard ranking	0.527	0.31	0.019	1	890
Specialization	5.867	2.488	1	10	890
Leader	0.135	0.342	0	1	890
Follower	0.521	0.5	0	1	890
Opinion on camp	8.283	1.277	1	10	890
Opinion on team performance	7.907	1.507	2	10	890
Team environment	8.664	1.482	1	10	890
Team influence on individual performance	8.08	1.821	1	10	890
Individual contribution to team	7.331	1.38	1	10	890
Relative knowledge	7.199	1.403	1	10	890
Relative brightness	7.207	1.414	1	10	890
Relative effort	7.392	1.391	1	10	890

camp, the average school test score received by students in our sample during the first quarter at school is higher than eight over ten. Somewhat surprisingly, the minimum score is four.

Table 3: Individual Characteristics

Variable	Mean	Std. Dev.	Min.	Max.	N
Grade	2.267	1.077	1	4	890
Male	0.513	0.5	0	1	890
Number of siblings	1.075	0.778	0	5	890
High-school grad father	0.407	0.492	0	1	890
University grad father	0.43	0.495	0	1	890
High-school grad mother	0.431	0.496	0	1	890
University grad mother	0.478	0.5	0	1	890
Mathematics school score	8.167	0.972	4	10	890

5 Randomization and Econometric Strategy

In Table 4 we compare the average individual characteristics of students in endogenous and exogenous teams. P-values in the third column show that there are no significant differences between the two groups except for a small difference in mathematics test scores at school. These scores are higher on average for students in exogenous teams than for students in endogenous teams. For this reason, we control for test scores at school in all our regressions.

Table 4: Randomization Test

Variable	**Mean Endogenous**	**Mean Exogenous**	**P-Value**
Male	0.513	0.514	0.970
Number of siblings	1.079	1.080	0.843
High-school grad father	.413	0.407	0.732
Father university	0.408	0.430	0.193
Mother high-school	0.436	0.427	0.797
Mother university	0.466	0.488	0.444
Mathematics school score	2.193	2.336	0.048

We test differences in the impact of team's gender composition between endogenous and exogenous teams by means of the following equation:

$$\begin{aligned} Y_{ig} = \beta_0 + \beta_1 EN_{ig} + \beta_2 EN_g * MG_g + \beta_3 EN_g * FG_g + ... \\ ... + \beta_4 MG_g + \beta_5 FG_g + \beta_6 C_{ig} + \beta_7 Grade_{ig} * Wave_{ig} + \epsilon_{ig} \end{aligned} \tag{1}$$

where Y_{ig} is one of the outcomes of the study (tastes for gender mix in the team, team performance, team working style, and opinions about the camp and the team) for individual i working in team g. EN is a dummy equal to 1 if student i belongs to an endogenous team g, MG is a binary indicator for male-predominant team (a team where more than two thirds of the team are males), and FG is an binary variable for female-predominant team (a team where more than two thirds of the team are females). Therefore, the reference category remains gender-balanced exogenous teams. The vector C contains the set of individual characteristics composed of a male dummy, indicators for number of siblings, mother high-school and mother university education binary variables, father high-school and father university education dummies, and the mathematics test score obtained in school in the first quarter. *Grade* by *Wave* are indicators for each combination of grade (one to four) and each of the four waves of the camp. Finally, ϵ represents the error term.

6 Results

We first observe that endogenous teams differ from exogenous teams in terms of gender mix (see Figure 1). Gender-balanced teams are over-represented in the set of endogenous teams. This indicates that the average student has a preference for this type of teams. We then study how students' satisfaction with their team's gender mix changes between endogenous and exogenous formed teams. We estimate Equation (1) using the answer to the following questions as dependent variables: (i) would you have preferred to have more males in your group?, (ii) would you have preferred to have more females in your group?, and (iii) were you satisfied with the gender mix in your team?. We show the result of these estimations in Table 5. For exogenous teams, students in male-predominant teams are unsatisfied with the gender mix. They would have preferred more females in their teams. Similarly, students in female-predominant teams would have preferred more males in their teams. In contrast, students in endogenous teams are overall satisfied with the gender mix in their teams. Some students in male-predominant teams would have preferred more females and the same is true for some students in female-predominant teams who would have preferred more males, but this happens significantly less often than in exogenous teams.

We are mainly interested in the impact of team's gender composition on team performance for endogeneous and exogenous teams. We address this by estimating Equation (1) using different versions of the team's final score as dependent variable. Table 9 shows the results of these estimations. In the first column, we use team's final test score standardized by grade and wave to have mean zero and standard deviation equal to one. Second column uses as dependent variable the position of the team in the ranking of the corresponding grade and wave. Finally, we standardized the ranking position in column 3 by dividing it by the total number of teams to account for the differ-

Table 5: Satisfaction with Team's Gender Mix

	More Male	As It Was	More Female
	(1)	(2)	(3)
Male-predominant	-.084	-.320	0.404
	(0.03)***	(0.104)***	(0.099)***
Female-predominant	0.19	-.186	-.004
	(0.076)**	(0.105)*	(0.053)
Endogenous male-predominant	0.073	0.116	-.189
	(0.039)*	(0.101)	(0.101)*
Endogenous female-predominant	-.134	0.168	-.034
	(0.082)*	(0.105)	(0.06)
Endogenous team	-.044	0.03	0.013
	(0.039)	(0.06)	(0.053)
Male	-.042	-.093	0.134
	(0.024)*	(0.034)***	(0.031)***
Observations	890	890	890
R^2	0.169	0.147	0.3

Notes: Data is from the questionnaire administered at the end of the camp. All regressions include a male dummy, indicators for number of siblings, mother high-school and mother university education binary variables, father high-school and father university education dummies, the mathematics test score obtained in school prior to the camp, and indicators for each combination of grade and wave. Standard errors are clustered at team level.

ent number of teams in each grade and wave. Results are consistent across the different measures of performance. We find that female-predominant teams underperform gender-balanced and male-predominant teams. The magnitude of the penalty associated to female-predominant teams equals 0.36 standard deviations of scores, almost 2 positions in the ranking, and 0.12 standard deviations of ranking position. These results are in line with findings in Apesteguia et al. [2012] and contradicts the conclusions in Hoogendoorn et al. [2013a]. Moreover, it rules out that the difference between the findings in the two papers relies on the endogeneity of team formation in Apesteguia et al. [2012] and the exogeneity of team formation in Hoogendoorn et al. [2013a].

Apesteguia et al. [2012] show that their estimated differences in performance are explained by differences in decision-making. We explore whether teams with different gender compositions and different formation rules differ in their working styles. To test this, we use three different proxies of working style which correspond to the three columns in Table 7. In the first column, we use the question "did the members of your team specialize in different aspects of the work?" which is coded from one to ten where one is not at all and ten is totally. The second and third columns have dummies for "the student identifies himself or herself as leader of his/her team" and "the student does not identify himself

Table 6: Performance by Gender-Mix and Endogeneity of Team Formation

	Std Score	Ranking	Std Ranking
	(1)	(2)	(3)
Male-predominant	0.08	-.750	-.045
	(0.234)	(1.040)	(0.063)
Female-predominant	-.361	1.801	0.117
	(0.183)**	(0.852)**	(0.052)**
Endogenous male-predominant	-.142	1.828	0.09
	(0.322)	(1.478)	(0.094)
Endogenous female-predominant	0.228	-.441	-.055
	(0.33)	(1.488)	(0.09)
Endogenous team	-.028	-.015	0.013
	(0.236)	(1.149)	(0.068)
Male	0.009	-.168	-.005
	(0.066)	(0.323)	(0.02)
Observations	890	890	890
R^2	0.03	0.198	0.157

Notes: Data is from the questionnaire administered at the end of the camp and teachers' reports. All regressions include a male dummy, indicators for number of siblings, mother high-school and mother university education binary variables, father high-school and father university education dummies, the mathematics test score obtained in school prior to the camp, and indicators for each combination of grade and wave. Standard errors are clustered at team level.

or herself as leader of his/her team". We study the two latter aspects separated because 34% of respondents declared that they do not know whether they are the leader of their teams. We only find significant coefficients for *follower*, i.e., those who declare not identify themselves as leaders. Students in endogenous teams are less likely to be followers. This results can be explained if as stated by Ahn, Isaac, and Salmon [2008], individuals exert more effort in endogenous teams and acting as a follower requires less effort. Similarly, students in female-predominant teams are also less likely to be followers. However, these negative effects are significantly weaker in gender-unbalanced (both male-predominant and female-predominant) endogenously-formed teams.

In Table 5 we showed that the team's gender mix influences students satisfaction and that the level of satisfaction with the team's gender team is higher for endogenous teams. Thus, it is possible that students' perception of the entire camp experience and of their team changes according to the gender mix and the team formation rule. To study this possibility, we re-estimate Equation (1) using three different outcomes: We first use a one to ten score on student overall opinion about the camp, with one meaning extremely bad and ten standing for extremely good. Second, we employ a one to ten score reflecting student's opinion on his or her team's performance. Finally, we use student's assessment

Table 7: Team's Working Style

	Specialization	Leader	Follower
	(1)	(2)	(3)
Male-predominant	-.033	-.016	-.066
	(0.197)	(0.05)	(0.055)
Female-predominant	-.356	-.007	-.174
	(0.293)	(0.042)	(0.053)***
Endogenous male-predominant	0.043	0.013	0.175
	(0.396)	(0.062)	(0.063)***
Endogenous female-predominant	0.212	-.023	0.209
	(0.524)	(0.048)	(0.055)***
Endogenous team	0.164	-.024	-.153
	(0.29)	(0.037)	(0.034)***
Male	0.132	0.09	-.142
	(0.138)	(0.031)***	(0.052)***
Observations	890	890	890
R^2	0.043	0.05	0.051

Notes: Data is from the questionnaire administered at the end of the camp. All regressions include a male dummy, indicators for number of siblings, mother high-school and mother university education binary variables, father high-school and father university education dummies, the mathematics test score obtained in school prior to the camp, and indicators for each combination of grade and wave. Standard errors are clustered at team level.

about the team's working environment which is also reported in a one (extremely bad) to ten (extremely good) scale. As shown in Table 8, we find that female-predominant teams have a more negative opinion about the overall camp experience with no differences between endogenous and exogenous teams. This can be explained if the lower performance of those teams reduced their level of satisfaction with the experience. In contrast, male-predominant teams have a significantly more positive judgment of their team's performance. We could not detect significant differences in the perceived working environment across the different team compositions.

The gender composition of the team and the team formation rules could also affect how students perceive themselves with respect to their group. We test this possibility using five different outcomes: (1) a one to ten score measuring whether the student considers that other team members had a positive influence on his or her individual performance, (2) a one to ten score indicating the self-assessed contribution to the team's results, (3) a one to ten score measuring how knowledgeable the student believes he or she is with respect to other team members, (4) a one to ten score illustrating student's opinion on how his/her brightness compares to other team members', and (5) a one to ten score measuring how much effort the student's believe to have exerted as compares

Table 8: Self-Perceived Experience

	Opinion Camp	Opinion Performance	Team Environment
	(1)	(2)	(3)
Male-predominant	-.258	0.381	0.191
	(0.173)	(0.215)*	(0.295)
Female-predominant	-.294	-.064	0.381
	(0.124)**	(0.219)	(0.253)
Endogenous male-predominant	0.102	-.334	-.231
	(0.222)	(0.271)	(0.319)
Endogenous female-predominant	0.157	0.318	-.042
	(0.18)	(0.322)	(0.274)
Endogenous team	-.045	0.174	0.224
	(0.132)	(0.271)	(0.245)
Male	-.252	-.057	0.129
	(0.112)**	(0.068)	(0.129)
Obs.	890	890	890
R^2	0.109	0.05	0.049

Notes: Data is from the questionnaire administered at the end of the camp. All regressions include a male dummy, indicators for number of siblings, mother high-school and mother university education binary variables, father high-school and father university education dummies, the mathematics test score obtained in school prior to the camp, and indicators for each combination of grade and wave. Standard errors are clustered at team level.

to other team members. Our findings indicate that members of endogenous teams are more likely to think that the team has had a positive influence on their individual performance. This may simply be explained by ex-post rationalization. We could not detect significant differences across team types for outcomes (2)-(4). In male-predominant teams students believe that they exerted more effort than their team-mates if their team was formed exogenously but they are more likely to declare to have exerted less effort than their team-mates in endogenous teams.

Therefore, we find differences between endogenous and exogenous teams only for the probability of being a follower (which is lower for endogenous teams) and team's influence on positive performance (which is higher for endogenous teams). We also find that students in male-predominant endogenous teams are less likely to consider that they have exerted more effort than their team-mates. Even though there are no differences in the estimated impact of team gender composition on performance between endogenous and exogenous teams, these differences may arise in other contexts where the probability of being a follower, team's influence on performance, and/or the relationship between individual's and team-mates' effort matters more for performance.

Table 9: Self-Perceived Performance

	Team Good Influence	Contribution	Knowledge	Talent	Effort
	(1)	(2)	(3)	(4)	(5)
Male-predominant	0.131	0.038	0.077	0.087	0.413
	(0.16)	(0.152)	(0.235)	(0.148)	(0.177)**
Female-predominant	-.082	0.098	0.083	-.005	-.029
	(0.203)	(0.146)	(0.186)	(0.104)	(0.123)
Endogenous male-predominant	-.244	0.004	-.151	-.231	-.698
	(0.212)	(0.22)	(0.288)	(0.246)	(0.23)***
Endogenous female-predominant	0.088	-.228	-.040	0.111	-.110
	(0.247)	(0.21)	(0.195)	(0.131)	(0.241)
Endogenous team	0.266	-.012	-.042	-.088	0.17
	(0.103)***	(0.12)	(0.166)	(0.083)	(0.15)
Male	-.359	0.275	0.407	0.443	-.245
	(0.169)**	(0.094)***	(0.096)***	(0.108)***	(0.122)**
Obs.	890	890	890	890	890
R^2	0.045	0.051	0.076	0.071	0.059

Notes: Data is from the questionnaire administered at the end of the camp. All regressions include a male dummy, indicators for number of siblings, mother high-school and mother university education binary variables, father high-school and father university education dummies, the mathematics test score obtained in school prior to the camp, and indicators for each combination of grade and wave. Standard errors are clustered at team level.

7 Conclusion

We are interested in whether the impact of team gender composition on performance, team working style and individual satisfaction change between endogenous and exogenous teams. We implement an experiment in the context of a mathematics camp where students work in teams of six members and compete in solving mathematical problems. We divide students in two equally sized groups where one group works in teams designed by us (exogenous teams) while the other group is free to form teams (endogenous teams). We then estimate how teams' outcomes change according to the gender composition of the team and whether the team was endogenously or exogenously formed.

We find that students who choose their team are significantly more satisfied with the gender composition of their team. Most students prefer gender-balanced teams. We also find that female-predominant teams underperform other teams but there are no differences in performance between endogenous and exogenous teams. In contrast, there are differences in the effect of gender composition on team working style and work and team satisfaction for endogenous and exogenous teams.

We conclude that, although in our setup there are no differences in the estimated impacts of gender composition on team performance between endogenous and exogenous teams, these differences may arise in contexts where team working style and work and team satisfaction have a higher effect on performance.

References

A. Abdulkadiroglu. School choice. *The handbook of market design*, pages 138–169, 2013.

K. R. Ahern and A. K. Dittmar. The changing of the boards: The impact on firm valuation of mandated female board representation. *The Quarterly Journal of Economics*, 127(1): 137–197, 2012.

T.-K. Ahn, R. M. Isaac, and T. C. Salmon. Endogenous group formation. *Journal of Public Economic Theory*, 10(2):171–194, 2008.

K. Antonovics, P. Arcidiacono, and R. Walsh. The effects of gender interactions in the lab and in the field. *The Review of Economics and Statistics*, 91(1):152–162, 2009.

J. Apesteguia, G. Azmat, and N. Iriberri. The impact of gender composition on team performance and decision making: Evidence from the field. *Management Science*, 58(1): 78–93, 2012.

S. Athey and G. W. Imbens. The econometrics of randomized experiments. In *Handbook of Economic Field Experiments*, volume 1, pages 73–140. Elsevier, 2017.

A. Aughinbaugh. The effects of high school math curriculum on college attendance: Evidence from the nlsy97. *Economics of Education Review*, 31(6):861–870, 2012.

G. Azmat and B. Petrongolo. Gender and the labor market: What have we learned from field and lab experiments? *Labour Economics*, 30:32–40, 2014.

M. F. Bagues and B. Esteve-Volart. Can gender parity break the glass ceiling? evidence from a repeated randomized experiment. *The Review of Economic Studies*, 77(4):1301–1328, 2010.

G. S. Becker. *The economics of discrimination*. University of Chicago press, 1957.

D. Blazar. Effective teaching in elementary mathematics: Identifying classroom practices that support student achievement. *Economics of Education Review*, 48:16–29, 2015.

L. Borghans, A. L. Duckworth, J. J. Heckman, and B. Ter Weel. The economics and psychology of personality traits. *Journal of human Resources*, 43(4):972–1059, 2008.

L. Borghans, J. J. Heckman, B. H. Golsteyn, and H. Meijers. Gender differences in risk aversion and ambiguity aversion. *Journal of the European Economic Association*, 7(2-3): 649–658, 2009.

C. T. Clotfelter, H. F. Ladd, and J. L. Vigdor. The aftermath of accelerating algebra evidence from district policy initiatives. *Journal of Human Resources*, 50(1):159–188, 2015.

K. E. Cortes, J. S. Goodman, and T. Nomi. Intensive math instruction and educational attainment long-run impacts of double-dose algebra. *Journal of Human Resources*, 50(1): 108–158, 2015.

J. Currie and D. Almond. Human capital development before age five. In *Handbook of labor economics*, volume 4, pages 1315–1486. Elsevier, 2011.

J. Delfgaauw, R. Dur, J. Sol, and W. Verbeke. Tournament incentives in the field: Gender differences in the workplace. *Journal of Labor Economics*, 31(2):305–326, 2013.

W. Dobbie and R. G. Fryer Jr. Getting beneath the veil of effective schools: Evidence from new york city. *American Economic Journal: Applied Economics*, 5(4):28–60, 2013.

M. Dufwenberg and A. Muren. Gender composition in teams. *Journal of Economic Behavior & Organization*, 61(1):50–54, 2006.

G. Ellison and A. Swanson. The gender gap in secondary school mathematics at high achievement levels: Evidence from the american mathematics competitions. *Journal of Economic Perspectives*, 24(2):109–28, 2010.

G. Ellison and A. Swanson. Do schools matter for high math achievement? evidence from the american mathematics competitions. *American Economic Review*, 106(6):1244–77, 2016.

M. Ferkany. The educational importance of self-esteem. *Journal of Philosophy of Education*, 42(1):119–132, 2008.

D. N. Figlio and M. E. Lucas. Do high grading standards affect student performance? *Journal of Public Economics*, 88(9-10):1815–1834, 2004.

D. Fuchs, L. S. Fuchs, P. G. Mathes, and D. C. Simmons. Peer-assisted learning strategies: Making classrooms more responsive to diversity. *American Educational Research Journal*, 34(1):174–206, 1997.

U. Gneezy and A. Rustichini. Gender and competition at a young age. *American Economic Review*, 94(2):377–381, 2004.

A. Griffith and K. Rask. The influence of the us news and world report collegiate rankings on the matriculation decision of high-ability students: 1995–2004. *Economics of Education Review*, 26(2):244–255, 2007.

L. Guiso, F. Monte, P. Sapienza, and L. Zingales. Culture, gender, and math. *Science*, 320 (5880):1164–1165, 2008.

S. Hoogendoorn, H. Oosterbeek, and M. Van Praag. The impact of gender diversity on the performance of business teams: Evidence from a field experiment. *Management Science*, 59(7):1514–1528, 2013a.

S. M. Hoogendoorn et al. *Diversity and team performance: A series of field experiments*. Tinbergen Institute, 2013b.

J. Horstschräer. University rankings in action? the importance of rankings and an excellence competition for university choice of high-ability students. *Economics of Education Review*, 31(6):1162–1176, 2012.

R. Ivanova-Stenzel and D. Kübler. Courtesy and idleness: Gender differences in team work and team competition. 2005.

R. Ivanova-Stenzel and D. Kübler. Gender differences in team work and team competition. *Journal of Economic Psychology*, 32(5):797–808, 2011.

M. Justman and S. J. Méndez. Gendered choices of stem subjects for matriculation are not driven by prior differences in mathematical achievement. *Economics of Education Review*, 64:282–297, 2018.

J. H. Kagel and A. E. Roth. *The handbook of experimental economics*, volume 2. Princeton university press, 2016.

E. I. Knudsen, J. J. Heckman, J. L. Cameron, and J. P. Shonkoff. Economic, neurobiological, and behavioral perspectives on building america?s future workforce. *Proceedings of the National Academy of Sciences*, 103(27):10155–10162, 2006.

V. Maestri. Promoting scientific faculties: Does it work? evidence from italy. *Economics of Education Review*, 32:168–180, 2013.

D. A. Matsa and A. R. Miller. A female style in corporate leadership? evidence from quotas. *American Economic Journal: Applied Economics*, 5(3):136–69, 2013.

M. Niederle and L. Vesterlund. Do women shy away from competition? do men compete too much? *The quarterly journal of economics*, 122(3):1067–1101, 2007.

S. Rothmann and E. P. Coetzer. The big five personality dimensions and job performance. *SA Journal of Industrial Psychology*, 29(1):68–74, 2003.

H.-g. Song, M. Restivo, A. van de Rijt, L. Scarlatos, D. Tonjes, and A. Orlov. The hidden gender effect in online collaboration: An experimental study of team performance under anonymity. *Computers in human Behavior*, 50:274–282, 2015.

Appendices

A Camp questionnaire

QUESTIONNAIRE FOR STUDENTS

Instructions

Dear Student,

at the end of your experience in Bardonecchia, we ask you to answer the short questionnaire which we present below. As for the first questionnaire, all the information you provide us will be treated in full respect of your anonymity.

Thanks a lot for the collaboration!!

BASIC INFORMATION

1. **Name**

   ....

2. **Surname**

   ....

3. **Class**

   ....

4. **Section**

   . . . .

5. **Type of school**

   . . . .

6. **Name of school**

   . . . .

7. **Starting date of the stage at Bardonecchia**

   . . . .

8. **Working table (number)**

   . . . .

EVALUATION OF THE CAMP EXPERIENCE

9. **How do you rate the internship experience? (from 1 - disappointing - to 10 - outstanding):**

. . . .

10. **What is the final score achieved by your group?**

    . . . .

11. **How do you rate your group's performance? (from 1 - disappointing - to 10 - outstanding):**

    . . . .

12. **What do you think was your contribution to solving the exercises proposed during the camp? (from 1 - nothing - to 10 - essential):**

    . . . .

13. **What do you think was your contribution to solving the exercises proposed during the camp? (from 1 - nothing - to 10 - essential):**

    . . . .

14. **How was the atmosphere within your group? (from 1 - conflictive - to 10 - cooperative):**

    . . . .

15. **Do you think working within your team has improved your performance compared to work alone? (from 1 - absolutely no - to 10 - absolutely yes):**

   . . . .

16. **Do team members specialize in one activity during the three days? (from 1 - absolutely no - to 10 - absolutely yes):**

   . . . .

17. **Regarding the composition of your group would you have preferred:**

   □ More females

   □ More males

   □ It was ok as it was

NOW WE ASK YOU AN OPINION ABOUT YOU ...

We specify that what you indicate in this section will not be disclosed to classmates or professors; nor will it be used to form any judgment about you. We only need the collected data to explore mechanisms within the working group in which you were inserted.

18. **What do you think is your level of mathematical knowledge compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

   . . . .

19. **What do you think your level of intelligence is compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

20. **What do you think is your level of commitment compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

21. **Would you identify yourself as leader of your work group at the camp?**

☐ Yes

☐ No

☐ I don't know

NOW WE ASK YOU AN OPINION ABOUT THE MEMBERS OF YOUR TEAM

We specify that what you indicate in this section will not be disclosed to classmates or professors; nor will it be used to form any judgment on people. We only need the collected data to exploremechanisms of operation of the working group in which you were inserted.

22. **Name of member 1:**

. . . .

23. **Surname of member 1:**

. . . .

24. **What do you think is her/his level of mathematical knowledge compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

25. **What do you think her/his level of intelligence is compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

26. **What do you think is her/his level of commitment compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

27. **Would you identify her/himself as leader of your work group at the camp?**

    □ Yes

    □ No

    □ I don't know

28. **Name of member 2:**

. . . .

29. **Surname of member 2:**

    . . . .

30. **What do you think is her/his level of mathematical knowledge compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

    . . . .

31. **What do you think her/his level of intelligence is compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

    . . . .

32. **What do you think is her/his level of commitment compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

    . . . .

33. **Would you identify her/himself as leader of your work group at the camp?**

    □ Yes

    □ No

    □ I don’t know

34. **Name of member 3:**

. . . .

35. **Surname of member 3:**

. . . .

36. **What do you think is her/his level of mathematical knowledge compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

37. **What do you think her/his level of intelligence is compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

38. **What do you think is her/his level of commitment compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

39. **Would you identify her/himself as leader of your work group at the camp?**

    □ Yes

    □ No

□ I don't know

40. **Name of member 4:**

. . . .

41. **Surname of member 4:**

. . . .

42. **What do you think is her/his level of mathematical knowledge compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

43. **What do you think her/his level of intelligence is compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

44. **What do you think is her/his level of commitment compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

45. **Would you identify her/himself as leader of your work group at the camp?**

□ Yes

□ No

□ I don’t know

46. **Name of member 5:**

. . . .

47. **Surname of member 5:**

. . . .

48. **What do you think is her/his level of mathematical knowledge compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

49. **What do you think her/his level of intelligence is compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

50. **What do you think is her/his level of commitment compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

51. **Would you identify her/himself as leader of your work group at the camp?**

- ☐ Yes
- ☐ No
- ☐ I don't know

52. **Name of member 6:**

. . . .

53. **Surname of member 6:**

. . . .

54. **What do you think is her/his level of mathematical knowledge compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

55. **What do you think her/his level of intelligence is compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

56. **What do you think is her/his level of commitment compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

57. **Would you identify her/himself as leader of your work group at the camp?**

   □ Yes

   □ No

   □ I don't know

58. **Name of member 7:**

   . . . .

59. **Surname of member 7:**

   . . . .

60. **What do you think is her/his level of mathematical knowledge compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

   . . . .

61. **What do you think her/his level of intelligence is compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

   . . . .

62. **What do you think is her/his level of commitment compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

63. **Would you identify her/himself as leader of your work group at the camp?**

    □ Yes

    □ No

    □ I don't know

64. **Name of member 8:**

    . . . .

65. **Surname of member 8:**

    . . . .

66. **What do you think is her/his level of mathematical knowledge compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

    . . . .

67. **What do you think her/his level of intelligence is compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

    . . . .

68. **What do you think is her/his level of commitment compared to the other members of the group? (from 1 - insufficient - to 10 - excellent):**

. . . .

69. **Would you identify her/himself as leader of your work group at the camp?**

- □ Yes
- □ No
- □ I don't know

BIG FIVE QUESTIONS (scale 0-10)

70. **Do you consider yourself an extrovert and sociable person?**

. . . .

71. **Do you consider yourself a friendly person?**

. . . .

72. **Do you consider yourself a sociable and conscientious person?**

. . . .

73. **Do you consider yourself a neurotic person?**

. . . .

74. **Do you consider yourself an open-mind person?**

. . . .

Let us briefly make the point: to avoid losing pieces and not being able to adequately exploit all the information there you have provided. As part of the evaluation of the Mathematics Stage, this should be the second questionnaire you fill out. Think about it ...

75. **Did you fulfill the Pre-stage questionnaire?**

☐ Yes

☐ No

☐ I son't remember

If you answered no or I don't know, the test continues with some quick ones socio-demographic and attitudinal questions.

76. **Write your date of birth**

. . . .

77. **Are you male or female?**

☐ M

☐ F

□ I don't want to answer

78. **Which is the Postal Code (CAP) of your home?**

. . . .

79. **Indicate your mother's education**

□ Graduate or Post-graduate

□ High-School

□ Compulsory school

□ Nothing

80. **Indicate your father's education**

□ Graduate or Post-graduate

□ High-School

□ Compulsory school

□ Nothing

81. **How many brothers/sisters do you have?**

. . . .

82. **Are your brothers/sisters younger or older than you?**

Brother/sister 1

- ☐ Older
- ☐ Younger

Brother/sister 2

- ☐ Older
- ☐ Younger

Brother/sister 3

- ☐ Older
- ☐ Younger

Brother/sister 4

- ☐ Older
- ☐ Younger

83. **Write name and surname of your math teacher**

. . . .

Information for scientific research (articles 1314 of the EU Reg 2016/679)

The test is finished, thanks for your help! Remember to click ”submit / submit” before closing the page