

Garcia-Brazales, Javier

**Working Paper**

## Changing Gender Attitudes: The Long-Run Effects of Early Exposure to Female Classmates

*Suggested Citation:* Garcia-Brazales, Javier (2021) : Changing Gender Attitudes: The Long-Run Effects of Early Exposure to Female Classmates, ZBW – Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at:

<https://hdl.handle.net/10419/232502>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

# Changing Gender Attitudes: The Long-Run Effects of Early Exposure to Female Classmates

Javier Garcia-Brazales (CEMFI)\*

First Version: January 2020; This Version: February 2021

## Abstract

Identity norms are an important cause of inequalities and talent misallocation. I leverage a unique opportunity to observe students exogenously allocated to classes across a close-to-nationally-representative set of Vietnamese schools to show that more exposure to female peers during childhood causally decreases the extent of agreement with traditional gender roles in the long-run. This shift in attitudes is accompanied by changes in actual behavior: employing friendship nominations I find that male children have more female friends and spend more time with them outside school. Moreover, both their intensive and extensive margin contributions to home production increase in the short- and the long-run. These results are novel in the attitudes formation and in the long-term effects of peers literature and are important in informing optimal class allocation. Academic spillovers from female classmates are much weaker.

**JEL Codes:** I24; I25; J16.

---

\*The author gratefully acknowledges the funding received from the Fundación Ramón Areces. Address: Calle Casado del Alisal 5, 28014, Madrid, Spain. Mail: javier.garcia@cemfi.edu.es

The data used in this publication come from Young Lives, a 15-year study of the changing nature of childhood poverty in Ethiopia, India, Peru and Vietnam ([www.younglives.org.uk](http://www.younglives.org.uk)). Young Lives is funded by UK aid from the Department for International Development (DFID). The views expressed here are those of the author. They are not necessarily those of Young Lives, the University of Oxford, DFID or other funders.

# 1 Introduction

An increasing body of literature has shown that traditional gender norms have important consequences on individual and family decisions such as labor force participation (e.g. [Bertrand et al., 2015](#); [Bursztyn et al., 2017](#)) and that this may translate into intra-household inequalities (for instance, in the division of home production, [Hwang et al., 2019](#)). At the same time, much attention has been placed on the role of school peers in academic achievement and, to a lesser extent, in other behavioral or late-life outcomes (e.g. [Hoxby, 2000](#); [Kremer and Levy, 2008](#); [Carrell et al., 2018](#)). The aim of this paper is to connect these two strands of the literature. In particular, I address the following, novel, question: does stronger exposure to female classmates at an early age causally affect *attitudes towards gender norms* that persist in the *long-run*?

To answer this question I focus on the Vietnamese case. This is a country where, despite a recent narrowing in gender gaps, women have consistently lagged behind men in many respects. For instance, in 2015, which is around the time when the data employed in my analysis was obtained, only 9% of the 200 members of the Central Committee of the Communist Party were females. What is more, this severe female under-representation in leadership positions extended to gender discrimination in access to occupations, there being a list of 77 jobs legally forbidden to women ([World Bank, 2015](#)).

While the literature on peer effects is subject to important identification issues ([Manski, 1993](#)), I exploit a unique opportunity to answer my question of interest in a context where peer composition is plausibly exogenous. More specifically, I make use of Oxford University’s Young Lives project, which has tracked two cohorts of Vietnamese students over a period of fifteen years and collected information on their academic trajectories, cognitive and non-cognitive skills, and life attitudes. I combine this longitudinal information with an additional release, named School Survey, that aimed at studying the academic context and progress of a subset of the main Young Lives survey respondents, together with their classmates, over one school year (2011-2012, when these children were enrolled in Grade 5). I link peers’ information from the School Survey with own attitudes elicited five years later in the 2016 round of the Young Lives survey. I then leverage a series of features of these datasets that are ideal to address my question of interest and that enable me to make several contributions to the existing literature.

First, teachers were asked to report the method followed to form classes. This allows me to work with the subset of sections where student assignment was said to be “random”

(around 80% of the original sample) or possibly exogenous (an additional 13%). Hence, conditional on accounting for students' selection into schools through school fixed effects, the pressing concerns in the literature about endogenous peer formation are largely mitigated. I do not take the exogenous allocation of students to classes as given and I provide various tests to verify it. For instance, I show that the proportion of female students in a class is not correlated with a wide range of individual observable characteristics. Another exercise is that I isolate the within-school variation in female peers exploited in my analysis. This yields a distribution consistent with the one arising from Monte Carlo simulations where assignment to peers within schools is random by construction.

By exploiting exogenous assignment of students to classes I am able to define peers at the class level and to obtain a clean identification of the causal effects of female peers on views on gender norms. This is in contrast with studies that do not benefit from exogenous peer formation and instead take advantage of idiosyncratic variation in peers' characteristics across cohorts within schools, whose identification is arguably less clear-cut and are also constrained to define peers at a more aggregated level: the Grade level ([Hoxby, 2000](#); [Angrist, 2014](#)). Moreover, because the School Survey is a large-scale project, featuring 92 school sites and being close to nationally-representative (although it has a pro-poor nature), I am able to provide estimates that are potentially more externally valid than other studies that also benefit from plausibly exogenous assignment of peers but that are typically limited to one, or few, specific schools (e.g. [Sacerdote, 2001](#); [Feld and Zölitz, 2017](#)).

Second, while most research on peer effects at school has focused on academic outcomes, I have access to a rather uncommon set of questions about gender norms eliciting views on aspects such as the leadership abilities of females relative to males or the importance that females adhere to traditional life goals emphasizing being a housewife over a successful professional career. This allows me both to provide the first evidence about an unexplored, and long-reaching, dimension that a priori seems likely to be subject to peers' influence as well as to dig into the specific attitudes that were modified. My study complements two recent pieces of work exploring how being around racial minority individuals (black students in the US) and poor classmates in India can foster acceptance towards these groups in academic settings ([Carrell et al., 2019](#); [Rao, 2019](#)).<sup>1</sup> It also adds a new dimension to the effects of peers on the broader realm of non-cognitive skills, in this case non-academic ones, that has

---

<sup>1</sup>Two other novel research has looked at the effects of experimentally-manipulated teams outside the academic context ([Lowe, forthcoming](#); [Dahl et al., forthcoming](#)). Moreover, [Anelli and Peri \(2019\)](#) focus, as I do, on academic gender peer effects, and find that males exposed to more females during secondary education are not more likely to choose college majors traditionally favored by females nor to have their labor market outcomes impacted.

notably been understudied (Gong et al., 2018), and I do so at an stage in the children’s development — early adolescence — that has been recently emphasized as crucial for the development of socio-emotional skills (Choudhury et al., 2006; Ashraf et al., 2020).

Third, another unique feature of the School Survey is that, at the end of the academic year, students were asked to categorize their degree of friendship and the frequency of interactions outside school with each and every surveyed classmate. I am therefore able to, inter alia, verify that the uncovered changes in gender norms reported five years later are in line with *actual* behavior experimented throughout the academic year, as measured by increased interactions with female peers. This finding also provides evidence that exogenously manipulated groups can influence actual network formation, something ex ante not certain (Carrell et al., 2013). I complement this behavioral result by showing that there is also an increase along the intensive and extensive margins in the time devoted to home chores and to caring for relatives — which are closely linked to traditional gender views<sup>2</sup> — among males who shared the classroom with a larger fraction of female peers.

Fourth, Young Lives succeeded at tracking individuals over a period of fifteen years with very low attrition rates. This is important in minimizing concerns of selection out of the sample given my focus on the long-term effects of exposure to female classmates during childhood and the fact that the gender-related questions of interest were elicited five years after the School Survey. Moreover, in as much as the study of the long-run effects of peers is scarce and limited to academic and labor market outcomes in developed countries (Gould et al., 2009; Black et al., 2013; Carrell et al., 2018; Bietenbeck, 2020), my work contributes the first evidence on the persistent effects of peer effects on attitudes, and I do so in the much less understood context of developing countries.

Fifth, Young Lives and the School Survey, unlike the often-used administrative datasets in the peer effect literature (e.g. Carrell et al., 2018; Anelli and Peri, 2019), provide detailed longitudinal information on a wide range of cognitive, non-cognitive and family decisions dimensions of the students, as well as rich sociodemographics on the teachers. This allows me to: (i) explore other interesting outcomes that could also be affected by female peers; (ii) evaluate in detail the key identification assumption of exogenous assignment of peers to classes within schools, and (iii) provide suggestive evidence on the mechanisms behind my results. This last point is particularly important as it contributes novel evidence on attitudes formation and their malleability, especially throughout childhood and adolescence (e.g. Hanaoka et al., 2018; Shigeoka, 2019).

---

<sup>2</sup>See, for instance, Hwang et al. (2019) or Hyun (2020).

Overall, my long-term approach and its wide coverage of outcomes provides important insights that short-term and narrower-scoped studies cannot address. In particular, my findings not only contribute to our understanding of the non-pecuniary returns to education (see, for instance, [Oreopoulos and Salvanes, 2011](#)), but also provide a more complete picture, both in terms of the outcomes explored and their temporal persistence, to inform the long-standing debate on school segregation by gender and optimal class allocation.

My main result is that a ten percentage points increase in exposure to female peers in Grade 5 causally reduces the degree of agreement with a composite measure of traditional gender norms by a tenth of a standard deviation five years later. While these effects are of similar size for male and female students, they are driven by children whose mother had less weight in the household’s decision making processes — as measured a few years before the peer group formation observed in the 2011 School Survey. This points in the direction of exposure to female peers being a force for *convergence* towards more equal views on gender norms. Quantile regressions provide a more nuanced picture by uncovering a significantly larger impact around the median of the distribution of attitudes, but not at its right tail i.e. among the most traditional ones. This suggests that the converging forces are found among children who were somewhat predisposed to having egalitarian views to begin with and that attitudes are less malleable among the very traditional children. Additionally, in order to better understand the margins of this shift in attitudes I then explore the effects across various dimensions of gender norms. For instance, returning to my introductory example, I find that larger exposure to females yields an economically large and statistically significant fall in the beliefs that females are less able to occupy leadership positions than their male counterparts.

These effects on views on gender roles, which are long-lived, are in contrast with those in academic performance in mathematics and Vietnamese: more exposure to female peers improves cognitive abilities contemporaneously but fade away five years later. School dropout at age fifteen was not affected either. My results therefore highlight the importance of going beyond the exploration of purely academic outcomes and provide valuable evidence on which outcomes are more prone to fade away over time and which ones are not. This has implications both for our understanding of sorting as a source of inequalities and for optimal policy on class composition that is less likely to underestimate its social returns. For instance, the fact that under increased gender-mixing, which is a relatively easily-implementable policy, females are persistently regarded as more able to undertake jobs with large responsibilities could significantly reduce the misallocation of talent to professional careers ([Hsieh et al.](#),

2019). The uncovered shifts in attitudes and socio-emotional skills therefore bring in an additional dimension to the cost-benefit analysis of such potential policy that would otherwise be obviated.

The rest of the paper is organized as follows. Section 2 introduces in detail the data employed. Section 3 outlines the empirical approach and provides evidence on the validity of the identification assumptions. Section 4 reports my main results on views on gender norms. Section 5 explores other outcomes of interest, such as cognition and school dropout, and provides evidence on potential mechanisms behind the main results. Section 6 probes the robustness of my results to different specifications. Section 7 concludes.

## 2 Data

### 2.1 Young Lives

Young Lives (YL) is a longitudinal study led by Oxford University aiming at providing researchers with rich data to study the living conditions, environment, and decisions of two cohorts of children (aged 1 and 7 at the first round, in 2002) from childhood to early adulthood (the fifth, and last, round was collected in 2016) across four developing countries.

As an important component of any child’s context, YL additionally undertook two “school surveys”.<sup>3</sup> In the case of Vietnam, the 2011 School Survey consisted in surveying a subset of YL’s *young cohort* and their classmates (up to 20, randomly selected) for each section in the school in Grade 5 at the beginning and at the end of the 2011-2012 academic year. Rich sociodemographic characteristics of all the surveyed students and their teachers were collected. The original 2011 School Survey dataset contains 3,284 students distributed across 176 classes in 56 schools. Among these students, 1,138 belonged to the longitudinal component of the YL survey, meaning that they continued being tracked after the end of the academic year.

In this paper I exploit both the School Survey (“SS”) data and YL’s longitudinal information (“longitudinal YL”, or YL for short) to be able to study the effects of class composition on outcomes measured years later. I now provide a description of the information available in each of these datasets.

**Sociodemographics (SS).** Although the unbiasedness of our results will rely on the exogenous assignment of students to classes, the rich sociodemographic information from the

---

<sup>3</sup>One in the academic year 2011-2012, which is central in the present paper, and another one in the 2016-2017 academic year, which I make use of only at one specific point, as indicated in due time.

SS is valuable for checking the balance in student observables, for exploring heterogeneous effects, and for increasing precision. Standard questions on the gender, age, parental education, ethnicity, and household size were asked. Moreover, an interesting piece of information pertains to the households' ownership of various relevant goods, which allows me to compute a wealth index<sup>4</sup>. This is useful because wealth has been shown to be important for addressing omitted variable biases in academic contexts (Lovenheim and Reynolds, 2013) and for predicting intergenerational transmission of outcomes (Blundell and Risa, 2019). Relevant to our context, there is also information on the number of books at home as well as on physical resources available at home that are likely to be an input in academic production (owning: a calculator; a study desk; a study chair, and a own space at home to study). I construct another index, this time for academic resources at home, as the proportion of affirmative responses.

**Cognitive Information (SS).** In order to evaluate the students' academic progress during one school year, YL designed, distributed, and supervised the undertaking of mathematics and Vietnamese language tests at each of the two visits (i.e. at the very beginning of the school year and right before its end). These tests consisted of 30 multiple-answer questions aiming at testing official curriculum knowledge and containing a subset of items common to both rounds. Two attractive features are: (a) given the multiple-choice nature of the tests and their implementation by YL fieldworkers, the scores do not suffer from differential grading neither within nor across schools, which maximizes their comparability, and (b) since the second round is conducted right before the end of the school year, differential knowledge losses over the holidays are avoided (Fruehwirth and Gagete-Miranda, 2019). I work both with the total raw score and, to ease interpretation, also with their standardized version (mean of 500 and a standard deviation of 100) and explore whether class composition affected academic performance in the short- and the long-run.

**Peers' Information and Friendship Nominations (SS).** The School Survey collects rich information of all surveyed students. This allows me to compute leave-out-means (the standard measure of peer characteristics used in the literature) for the various dimensions of interest *at the class level*, which is a finer level of aggregation than what it is possible in the absence of random assignment. Importantly, the combination of a survey designed *ex profeso* to study potential channels of academic spillovers (e.g. non-cognitive aspects) sets this study

---

<sup>4</sup>I compute such an index as the proportion of positive answers to the questions of whether the household owns each of the following eleven items: phone; television; electric fan; computer; mobile phone; bicycle; air conditioning; internet; radio; motorbike; car.



apart from the rather limited information typically available on peers' characteristics, which is usually limited to those readily available in administrative data (such as gender and age).

Moreover, a unique feature of the school survey is that, at the end of academic year, every child was asked to answer the following two questions with respect to each and every surveyed classmate: (1) how would you describe your friendship with this classmate?, and (2) how much do you do things with this classmate outside of school?<sup>5</sup> This allows me to undertake two important analyses. First, I explore whether the less traditional views on gender roles that I uncover are in line with actual behavior, as proxied by the inclusion of more females in the friend circle of male children. Second, although friendship formation within classes during the 2011-2012 academic year is endogenous, it will allow me to provide suggestive evidence on up to which degree of friendship peer effects may be operational.

**Class Assignment (SS).** A crucial aspect of the SS is that it provides information, as reported by the class teacher, on the way that students were allocated to the section. The responses were as follows: (i) randomly (77.73% of the students); (ii) there was only one Grade 5 class (6.68%); (iii) by location of residence (6.62%); (iv) by general ability (4.75%); (v) by ability in mathematics (2.76%), and (vi) by age (1.47%). These figures show a very attractive characteristic of my context of interest: almost 80% of the students were randomly allocated to classes.

**Teacher Characteristics (SS).** Given the crucial role that teachers have been found to have as a key school-based input ([Jackson et al., 2014](#)), the school survey also collected detailed background characteristics of the teachers. These included, among others, gender, ethnicity, highest educational level, years of experience, and wealth. While this information is useful in reinforcing the evidence that random allocation of students also extended to their assigned teachers, it is unfortunate that the survey did not collect longitudinal information on teacher practices that could have provided direct evidence on whether teacher behavior is adaptive to class composition.

**Non-cognitive Information (SS).** As alluded above, one strength of the school survey is that it collected non-cognitive information for all students including, among other, self-reported level of effort and of perceived performance at school. This is important for the study of the mechanisms in place. Indeed, although I do not have self-reported behavioral information from the teachers, I do observe students' perceptions on the treatment by their teachers both at the beginning and at the end of the year. This is relevant for the study

---

<sup>5</sup>The options available for question (1) were: not close friends; a little/sometimes friends; close friends; very close friends. For (2) they were: none; not very much; quite a lot; a lot.

of peer effects since students are likely to be affected by and act upon such *perceptions*, regardless of actual behavioral changes on the part of the teachers. One relevant aspect to highlight is that the set of elicited non-cognitive information in the SS does not contain views on gender norms, hence precluding me from estimating the short-run effect (i.e. at the end of the academic year) of peers on them.

**Views on Gender Norms (longitudinal YL).** Instead, I obtain the information on the long-term views on gender norms from the fifth wave of the YL longitudinal survey (the only wave where they are elicited). As such, it is available for all YL students, but not for their classmates in the School Survey. This is, however, not problematic for my causal inference of interest, as YL students represent a random and representative sample of the population (I discuss this in Section C.1).

To be more specific, the 2016 round of YL, when the young cohort was aged 15, asked the following battery of questions, to which the students had to answer in a four point scale (strongly disagree, disagree, agree, strongly agree) and that I classify into four categories according to the broad gender-related aspect that was elicited. **“(1) Life purpose”** covers the following statements: (a) girls should be more concerned with becoming good wives and mothers than desiring a professional or business career; (b) girls should have the same freedoms as boys; (c) more encouragement in a family should be given to sons than to daughters to go to college; (d) it is more important for boys than for girls to do well in school. **“(2) Abilities”** contains: (e) boys are better leaders than girls; (f) on the average, girls are as smart as boys. **“(3) Cross-gender interactions”** features: (g) it is all right for a girl to ask a boy out on a date; (h) on a date, the boy should be expected to pay all expenses; (i) it is all right for a girl to want to play rough sports like football; (j) swearing is worse for a girl than for a boy. Finally, **“(4) intra-household decisions”** addresses: (k) in general, the father should have greater authority than the mother in making family decisions; (l) if both husband and wife have jobs, the husband should do a share of the housework such as washing dishes and doing the laundry.

To gauge and overall effect, to increase statistical power, and to facilitate interpretation, in my main analysis I compute the average level of agreement towards traditional norms — higher values indicate a more traditional view on the position of females; I flip the ordering of the responses when relevant, e.g. in question (3).<sup>6</sup> Moreover, in order to explore which specific dimensions of views are affected by class composition I also make use of each

---

<sup>6</sup>I also check robustness to alternative constructions of the attitudinal variables, e.g. using indicators signaling that the child agrees or strongly agrees with a given statement.

dimension separately.

**Sociodemographics and Cognition (longitudinal YL).** As explained, the longitudinal tracking of the two YL cohorts collected rich information about the realities of the children. Moreover, language and mathematics tests were consistently implemented in every YL round. I take advantage of this in two main ways. First, it allows me to provide supportive evidence in favor both of the random allocation of students to classes and of the non-selected choice of the subset of YL children that were sampled for the school survey. Second, it provides valuable information to explore the mechanisms behind my main effects as well as alternative outcomes of interest. For instance, I am able to investigate whether dropout rates systematically differ across individuals who faced different degrees of exposure to female peers, and whether cognitive effects are persistent or fade away over time.

**Sample Selection.** As mentioned, my main outcome of interest, long-run attitudes towards gender norms, is only observed in the fifth wave of the YL longitudinal survey. This means that in my baseline analysis I need to restrict my sample to those children in the SS that belonged to the YL survey (1,138). Additionally, my preferred estimation strategy relies on exogenous assignment of students to classes. Excluding those individuals assigned by ability or by age (which may reflect ability for repeater and grade skippers)<sup>7</sup> and those with small peer groups (classes with less than nine pupils) leaves me with 937 student-level observations distributed across 74 school sites (i.e. including satellite locations<sup>8</sup>) and 152 classes.<sup>9</sup> While this is my main estimating sample, the number of observations is slightly reduced in certain specifications with additional controls due to missing covariates.

## 2.2 Descriptive Statistics

Table 1 provides descriptive statistics for standard sociodemographic characteristics, for the raw scores in the mathematics and Vietnamese tests in the first round of the school survey, and for my main explanatory variable, the proportion of female peers in the class. Briefly, the sample is evenly split in terms of gender.<sup>10</sup> The average age is 10.29, as expected given

---

<sup>7</sup>Feld and Zölitz (2017) show how non-random selection based on ability largely biases the estimation of peer effects upwards.

<sup>8</sup>These are blocks of classrooms located at a distance from the main school — 50 out of the original 56 main schools in my sample — in order to facilitate schooling access to remote areas (Rolleston et al., 2013). Indeed, 96% of the main school principals report that the goal of the satellite sites is to enable children in remote areas to attend school. Therefore, given their geographical separation from the main school in my main specifications I include school site fixed effects (i.e. I treat satellite sites as separate from the main school site).

<sup>9</sup>In Section C.4 I verify that the main results hold under more stringent sample selection criteria.

<sup>10</sup>This is consistent with the fact that the sex ratios at birth in 2001, the year of birth of my cohort of interest, were very close to the natural rate of 105 boys per 100 girls (Guilmoto, 2009) and that primary education attendance is compulsory in Vietnam.

Table 1: Descriptive Statistics of Selected Variables in the School Survey

	(1)	(2)	(3)	(4)	(5)
	Mean	Standard Deviation	Min	Max	Count
Female	0.49	0.50	0	1	932
Age (years)	10.29	0.25	8.75	12.17	922
Father Can Read	0.96	0.20	0	1	915
Mother Can Read	0.95	0.21	0	1	929
Ethnic Minority (non-Kinh)	0.08	0.027	0	1	932
Home Educational Resources Index	0.72	0.25	0	1	937
Wealth Index	0.51	0.15	0	1	937
No Books at Home	0.19	0.39	0	1	929
1-5 Books at Home	0.25	0.43	0	1	929
6-10 Books at Home	0.15	0.35	0	1	929
More than 10 Books at Home	0.42	0.49	0	1	929
No Health Problem	0.71	0.45	0	1	937
Mathematics Raw Score First Test	18.01	5.57	1	30	930
Vietnamese Raw Score First Test	19.88	5.30	2	30	933
Proportion Female Peers	0.47	0.10	0.21	0.86	925

*Notes.* Descriptive statistics computed from the estimating sample for the long-run effects on attitudes (74 school sites and 152 classes). All variables are indicators, unless stated otherwise.

that it pertains to Grade 5, although there are a few students who are either ahead or behind for their age. Around 96% of the parents are able to read<sup>11</sup>. The average values for my two indices are 0.72 for the one measuring educational resources at home, and 0.51 for the overall household wealth index. The raw scores in the first round of the school survey were 18.01 and 19.88 for mathematics and Vietnamese, respectively. Finally, the proportion of female peers has an average of 0.47, a standard deviation of 0.1, and is close to normally distributed (Figure B.3.)<sup>12</sup>

In Table 2 I turn to the descriptives of the attitudinal questions elicited in the fifth round

<sup>11</sup>I choose to report this coarse measure of parental human capital because the alternative one of the highest academic level reached by parents was frequently missed (it was reported by the children).

<sup>12</sup>There is also substantial *within-school* variation in the proportion of female peers. The mean standard deviation of that variable among all schools is 0.07 with a standard deviation of 0.04. Moreover, it is worth remarking that mixed-gender classes are universal in the sample.

of Young Lives. To ease the reporting, I compress responses into a binary variable taking the value of 1 when the child reported that he/she agreed or strongly agreed with the respective traditional view indicated by the row variable — the separate distributions for male and female children using the four possible responses are provided in Online Appendix Figures B.1 and B.2.

Overall, traditional gender norms are well-accepted among the youth. For instance, 62% support the idea that females should aim at being good housewives to the detriment of their professional careers and 42% believe that females lack leadership abilities relative to males. At the same time, one should also highlight that for other dimensions there is significantly less support for traditional norms. For example, only 8% of the respondents consider that males and females differ in intelligence.

In Table A.1 I report the same moments of the data for the *old cohort* of the Young Lives data. These children were around 21 years of age at the time of the 2016 interview. One can appreciate that the figures are very similar to those in Table 1, which provides external validity to the attitudes reported by the young cohort and also suggests that the views (and the estimated changes that I obtain in the rest of the paper) of the fifteen-year-old individuals are likely to persist later in life.

### 3 Empirical Approach

#### 3.1 Baseline Specification

My main empirical strategy is based on the widely used linear-in-means specification to estimate Equation 1:

$$y_{ics2} = \alpha + \beta_0 PF_{-ics1} + \beta_1 Female_{ics} + \beta_2 X_{ics1} + \beta_3 \overline{X}_{-ics1} + \lambda_s + \epsilon_{ics}, \quad (1)$$

where  $y$  refers to our outcome of interest (primarily views on gender norms at the YL’s 2016 round),  $ics$  denotes person  $i$ <sup>13</sup> in class  $c$  in school  $s$ , 1 indicates that the variables were measured in the baseline round of the School Survey, and 2 indicates a moment in time posterior to the baseline round (i.e. either the end-of-year re-interview or our long-term outcomes measured years later through the YL data).

$PF$  measures the proportion (i.e. support is the closed interval 0-1) of female students in the class, excluding the person of reference. Hence,  $\beta_0$ , our coefficient of interest, estimates the average effect of exposure to a higher proportion of female classmates *during the*

---

<sup>13</sup>Adding a “-” sign in front is short-hand notation for indicating that a variable is computed as the leave-out-mean among the peers of individual  $i$ , located in class  $c$  at school  $s$ .

Table 2: Descriptive Statistics of Attitudes towards Gender Norms

	(1)	(2)	(3)	(4)	(5)
	Mean	Standard Deviation	Min	Max	Count
<b>Life Purpose</b>					
Life Goals	0.62	0.48	0	1	936
Freedom	0.08	0.28	0	1	937
Family Encouragement for College	0.17	0.38	0	1	935
Importance of Good Academic Performance	0.20	0.40	0	1	936
<b>Abilities</b>					
Leadership Ability	0.42	0.49	0	1	937
Intelligence	0.08	0.27	0	1	936
<b>Cross-Gender Interactions</b>					
Asking for a Date	0.22	0.42	0	1	926
Pay for a Date	0.56	0.50	0	1	931
Male-dominated Sports	0.09	0.28	0	1	937
Swearing	0.58	0.49	0	1	936
<b>Intra-household Decisions</b>					
Authority in Household Decisions	0.45	0.50	0	1	935
Chores	0.05	0.22	0	1	937
<b>Mean Score</b>	2.18	0.32	1.17	3.17	937

*Notes.* All variables except Mean Score are indicators taking the value one if the child agrees or strongly agrees with a traditional view on gender norms for each dimension of interest. The exact questions were reported in the “Views on Gender Norms (longitudinal YL)” subsection. Mean Score is computed as the average score (on a 1-4 scale) across the twelve dimensions at the individual level.

*academic year when the School Survey took place* on our outcome of interest. It captures an “exogenous” peer effect, as it arises from background characteristics of the students, and not from malleable dimensions such as their achievement or their behavior (i.e. “endogenous effect”). While existing work has explored exogenous effects stemming from gender and race, I add to this literature by focusing on an outcome previously unexplored (views on gender norms), by observing them five years later, which allows me to evaluate the persistence of the effects, and by investigating multiple relevant outcomes and potential mechanisms.

Although it is well-known that the identification of exogenous peer effects is complicated by several aspects inherent to social network formation, my context of interest is particularly well-suited to deal with these. A first difficulty is that direct comparisons of students across schools is likely to yield biased estimates due to selection into schools — these are Manski (1993)’s “correlated effects”. If this was the case, unobserved determinants of a student’s gender views would likely also be correlated with her classmates’ average characteristics, including the proportion of female peers. For instance, families with more traditional views might sort into locations where other families share such views. The standard way of dealing with this problem is through the inclusion of unit-level fixed effects at a higher level of aggregation than the one at which peer effects are measured. In my case, since I define peers at the classroom level, I include school site fixed effects ( $\lambda_s$ ).<sup>14</sup>

The addition of school fixed effects cannot deal, however, with potential sorting of students into classes. For example, it could be the case that schools assign students to classes based on certain student characteristics that correlate with unobserved determinants of our outcome of interest. Another case would be if parents who pay particular care to their child’s academic performance exerted some pressure to the school headmaster to place their children with high-ability students.

Because random assignment of students to classes is rare, particularly when aiming at expanding the sample to a representative one<sup>15</sup>, the existing literature has often exploited plausibly exogenous changes in peers’ background characteristics across cohorts within schools. This requires tracking schools over years, assuming that compositional changes within schools are not correlated with the long-run outcomes of interest, and also forces researchers to de-

---

<sup>14</sup>The inclusion of these fixed effects also accounts for the fact that, while there is annual random assignment of students across classes within schools, there is still persistence in the proportion of females across all classes in a given cohort. This persistence arises because there is little grade repetition and mobility across schools so the proportion of female peers for a given cohort remains similar as these students progress across grades.

<sup>15</sup>I use the term “representative” here in a broad sense (e.g. could be nationally representative, provincially representative, etc.) What I emphasize is the attempt of going beyond simply focusing on a particular school whose idiosyncrasies provided exogenous assignment to peers.

fine peers at a rather broad level: same-grade students. In my case, given that I have direct information on the way sections were formed, and that random assignment is widespread in primary education in Vietnam, this allows me to focus on the schools that exogenously assign students to classes. This not only overcomes the concern of within-school selection but also allows me to consider peers at a finer level (i.e. the classroom). Moreover, one central concern in long-run studies is selection out of the sample, or selection out of the assigned experimental group as years progress, as it is the case in, for instance, project STAR (Krueger, 1999). Because my exogenous variation comes from a given year, and the tracking of the random subset of Young Lives children was done with minimal attrition, these concerns are minimized.<sup>16</sup>

One important aspect to discuss is that, because the School Survey purposefully sampled 20 randomly-selected students per class, there is some measurement error in my peers variables. In particular, although I can observe the actual size of the class, the total number of males and females was not recorded, and I can only compute this proportion within the twenty students observed. If the resulting measurement error — which is present only for classes with more than 20 students<sup>17</sup> — is classical this will bias the estimates towards zero. Under random assignment to classes within schools and if students are missing at random, which should both hold given the design of the school survey, Sojourner (2013) shows that it is possible to account for the partial unobservability of peers by controlling for school fixed effects and their interactions with the proportion of total classmates that are observed (which I can compute since I know the size of each class).

With these aspects in mind, identification in my empirical approach is achieved from variation in the percentage of female classmates within schools across classes. The stability of the results will be verified by introducing additional covariates (both for the individual and the peers — i.e. in  $X$  and in  $\bar{X}$ ) such as wealth or parental education. These inclusions, precisely stated as they become relevant, are not strictly necessary to ensure unbiasedness of my coefficient of interest, but they help in increasing precision.  $\epsilon$  is the error term. I cluster the errors at the class level (the level of randomization) in order to account for possible correlation of the outcomes among classmates.

---

<sup>16</sup>The second traditional identification issue in the peer effects literature is that of “reflection”. Intuitively, this problem arises when it is hard to disentangle the effects of peers on an individual, and those from that student on her peers. In my particular exercise, in which the focus is on exogenous peer effects arising from a predetermined variable (gender), and given random classroom allocation, the concern that one’s gender causes peers’ gender is irrelevant.

<sup>17</sup>The median class size among the students in my estimating sample is 28.5 and the 75<sup>th</sup> percentile is 33.



### 3.2 Validity of Exogenous Class Formation

The key identification assumption is random assignment to peers within schools. If this holds, the OLS-estimated  $\beta_0$  in Equation 1 will have a causal interpretation. I perform various tests that consistently point at the validity of this assumption.

First, I explore whether, conditional on school fixed effects, the proportion of female classmates is correlated with any of a wide range of individual observables like ethnicity, parental education, child’s health conditions, and educational resources at home.<sup>18</sup> Importantly, non-cognitive characteristics measured at the beginning of the academic year and that are typically unobserved in similar studies such as whether the child feels pressure to do well at school are also included to explore baseline differences in ex ante relevant dimensions that could hint at potential selection based on attitudes.<sup>19</sup> Table A.4 shows that this is not the case, where only one of the thirteen dimensions investigated is significant at the 10% — as would be expected by pure chance.

As additional pieces of supportive evidence, I take advantage of the longitudinal information from the YL and explore whether female peer composition correlates with children’s observable characteristics in round 3. This wave took place in 2009 and therefore predates the school survey. Once again, in Table A.5 we find no evidence in favor of systematic differences prior to the formation of peer groups across children assigned to classes with exogenously higher proportions of female classmates along a rich set of socio-emotional dimensions including, for instance, self-perceived academic performance, strength of friendships, the value of effort, and the degree of satisfaction with multiple aspects of a child’s life. Moreover, Table A.6 shows that female peer composition is not related to a rich set of observable teacher characteristics such as their gender, tenure, the type of contract, and their beliefs that even the worst students can be helped by the teacher.

Second, I explore whether the observed distribution of the proportion of female peers is consistent with the one that, given the actual class sizes and the total number of female students in each school, would arise under randomization. For this I obtain the residuals from a regression of the proportion of females in a class on school fixed effects. This is the variation that is exploited in my main analysis. I replicate this exercise 1,000 times, in each of them

---

<sup>18</sup>I also check for differences in gender. For this, I follow Guryan et al. (2009) in controlling for the proportion of female peers at the school level to account for the fact that there is a bias arising from the sampling of peers without replacement. No relationship is found at convention levels of significance (p-value of 0.111).

<sup>19</sup>As a comparison, one of the most studied settings benefiting from randomized class allocation, Project STAR, only allows for the exploration of baseline balance for age, gender, race, and free-lunch status. This leaves the door open to potential differences in other dimensions for students across classes and treatment arms (Hanushek, 2003; Chetty et al., 2011).

reshuffling students across classes within their original school, while respecting the number of section and the total of female students. Figure B.4 shows that the actual distribution of the residuals closely matches the one obtained in the Monte Carlo exercise, which strengthens the confidence that the variation in female peers exploited is indeed random. Additionally, in Section C.2 I show by means of permutation tests with resampling that the baseline average characteristics at the class level are consistent with the ones that would arise from randomization.

Third, I follow Chetty et al. (2011) in regressing students' gender on school and class fixed effects. Given random allocation of female students — a predetermined variable — across classes within schools, an F-test of the joint significance of the class dummies should not show a significant relationship after the inclusion of school fixed effects. This is indeed the case, with the test yielding a p-value of 0.97.

## 4 Main Results

**Visual Analysis.** I start by showing raw correlations between the proportion of females and views on gender roles in round 5 when only controlling for school fixed effects in Figure B.6. There are two main readings. First, there is a clear negative correlation between the number of female peers in the classroom in the 2011-2012 academic year and the degree of support towards traditional gender norms in 2016. Second, there do not seem to be significant differences in the correlation for the subsets of males and female students.

### 4.1 Baseline Results

In order to formally explore the statistical significance of this relationship, I present in Table 3 the results from estimating Equation 1, both without correcting for the partial observability of peers (Panel A) and correcting for it (Panel B). Column 1 simply controls for the student's gender and school fixed effects. The negative correlation depicted before remains and is strongly statistically significant. In terms of economic magnitude, a ten percentage points increase in the proportion of female classmates decreases the degree of agreement with traditional norms in almost a tenth (8.6%) of a standard deviation (which is 0.32).

While the successful random assignment of students to classes guarantees that these results are causal and unbiased, in column 2 I add a rich set of controls at the individual level — as listed in the Table's notes — in order to increase precision and to verify that the

point estimate in column 1 is not sensitive to this inclusion, as should be the case under exogenous assignment. Both of these theoretical predictions hold. I proceed similarly in column 3, this time adding the same controls for peers' attributes. The fact that the point estimate is virtually unchanged suggests that we are indeed capturing the effects of gender, and not of other characteristics correlated with it.

In column 4 I explore whether the effects are heterogeneous for males and females. As expected from Figure B.4 this is not the case: while our measure of the proportion of female peers continues being significant and of about the same magnitude, its interaction with the female dummy is neither significant nor of a large economic size. Reassuringly, the results from Panel B are qualitatively the same as in Panel A but larger in size, as expected – the effect is now 15.2% of a standard deviation.

## 4.2 Exploring the Effects on the Different Aspects of Gender Norms

In order to better understand the nature of the changes in views uncovered above, I explore the effects separately for each of the dimensions used to construct the aggregate outcome measure employed in Table 3.

Figure 1 provides the point estimates for the proportion of female peers and the corresponding 90% confidence intervals from separate regressions for each of the twelve outcomes (standardized as z-scores) both with and without correction for partial observability of peers. Several patterns arise. First, estimates are negative over the board. In particular, without correction, there is a significant shift towards more equality in acceptable cross-gender interaction as measured by: females should be able to ask males for dates (g), to play male-dominated sports (i), and to swear as much as males (j). Importantly, females are no longer expected to have life goals that prioritize their professional careers (a). Consistently with the prediction that correcting for partial observability alleviates the existing downward bias, the point estimates in orange are more negative and in some cases become significant, as is the case for who should pay for a date (h). Second, there is convergence towards the view that males and females do not differ in abilities. More specifically, while the descriptive statistics already pointed towards widespread acceptance of the view that males and females do not differ in academic abilities, the one dimension where they were largely believed to differ was in terms of leadership abilities. We see that individuals more exposed to females are less likely to continue holding this belief (e). Also, girls are now regarded as deserving as much freedom as their male counterparts (b).<sup>20</sup> Finally, there tends to be no noticeable change

---

<sup>20</sup>Given that I am testing multiple outcomes one may want to account for this by correcting the threshold p-value

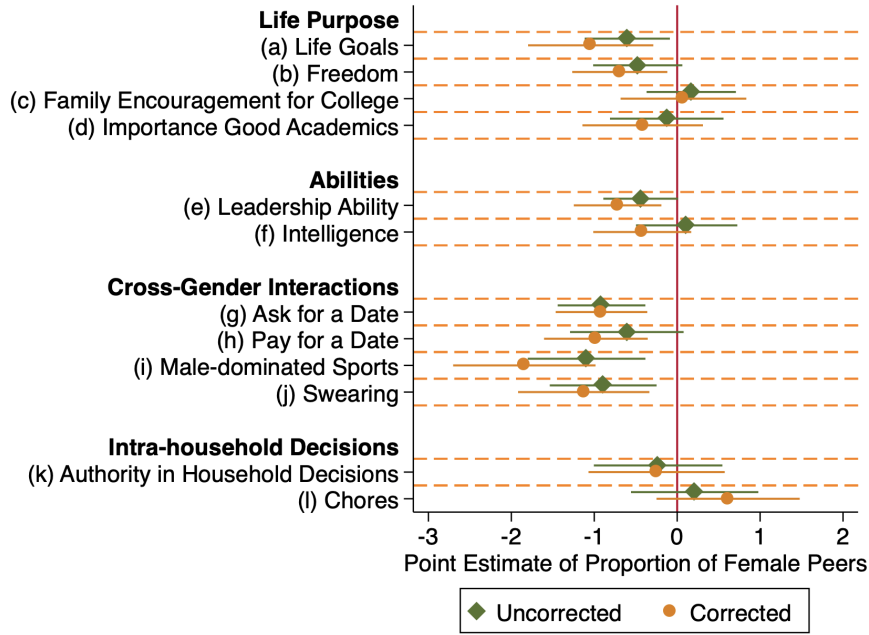
Table 3: Effects of the Proportion of Female Peers on Long-term Views on Gender Norms

<b>Panel A: No Correction</b>				
	(1)	(2)	(3)	(4)
	Mean Agreement with Traditional Gender Norms			
Proportion Female Peers	-0.274** (0.135)	-0.277** (0.128)	-0.267** (0.128)	-0.292* (0.170)
Female	-0.174*** (0.021)	-0.177*** (0.021)	-0.174*** (0.021)	-0.191* (0.107)
Female*Prop. Female Peers				0.030 (0.215)
Additional Controls	None	Indiv.	Indiv.+Peers	Indiv.
Observations	921	886	886	886
R-squared	0.230	0.248	0.258	0.248
<b>Panel B: Correction</b>				
	(5)	(6)	(7)	(8)
	Mean Agreement with Traditional Gender Norms			
Proportion Female Peers	-0.488*** (0.106)	-0.486*** (0.103)	-0.424*** (0.148)	-0.505*** (0.151)
Female	-0.184*** (0.021)	-0.186*** (0.021)	-0.181*** (0.022)	-0.204* (0.114)
Female*Prop. Female Peers				0.038 (0.228)
Additional Controls	None	Indiv.	Indiv.+Peers	Indiv.
Observations	915	880	880	880
R-squared	0.269	0.290	0.299	0.290

*Notes.* All regressions control for gender and school fixed effects. Column 2 adds the following individual controls: age, an indicator for paternal literacy, an indicator for being an ethnic minority, the educational resources and household wealth indices, and the categorical measure of number of books in the household. Column 3 additionally adds controls for the average peer characteristics for the same dimensions included as individual controls. Column 4 estimates the same specification as in column 2, but adding the interaction between a female indicator and the proportion of female classmates. All controls are measured in the baseline round of the 2011 School Survey. The mean of the dependent variable is 2.18 and its standard deviation is 0.32. Panel B proceeds similarly but accounting for the partial observability of peers by including the addition of the interaction between school fixed effects and the proportion of peers observed in the class. The sample contains all YL students with non-missing values present in the School Survey whose class-assignment satisfied the sample selection criteria outlined in Section 2.1 and who are also observed in round 5 of the YL longitudinal dataset. Standard errors clustered at the class level (151 clusters) in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

in those dimensions where there was already relatively weak support towards traditional gender norms, such as whether parents should encourage male children more to go to college (c), or in aspects related to intra-household dimensions (k and l). Indeed, it is only for this already-very-egalitarian views where we find (non-significant) positive point estimates.<sup>21</sup>

Figure 1: Breakdown of the Effects of the Proportion of Female Peers on Long-term Views on Gender Norms



### 4.3 Are the Results in Line with Actual Behavior?

While the above analyses already provide novel and meaningful results, one important aspect to consider is whether stated preferences translate into actual changes in behavior. Young Lives and the School Survey feature two key pieces of information that allow me to explore this avenue: (i) friendship nominations and (ii) time use allocation to home chores and care of relatives.<sup>22</sup>

under which we may claim statistically significant effects while maintaining a false discovery rate of 10%. Following [Benjamini and Hochberg \(1995\)](#)'s correction I find that the significance under partial observability is preserved for all outcomes.

<sup>21</sup>In Online Appendix Figure B.7 I repeat the same analysis separately for male and female children. While members of both groups display an overall shift towards less traditional views, some notable patterns appear. For instance, the impact on males is widely present across the various dimensions contained within the abilities and cross-gender interactions categories, but females' views on own abilities are not impacted. Another important remark is that the overall effects found for life goals were driven by females.

<sup>22</sup>Other interesting behaviors that have a clear mapping with the dimensions elicited in the gender norms questions are generally unobservable at the age of 15. For example, these adolescents are still too young to have formed a household so that we have no information on their chosen household arrangements or actions taken with respect to

**Friendship Formation.** The school survey has a unique feature by which friend nominations for every surveyed classmate are available. I have access to information on how close of a friendship children consider to have developed with each classmate, as well as on how often they do things together outside school. I focus on the male subsample and ask the following question: are children who share the classroom with a higher proportion of female classmates disproportionately more likely to form friendship groups with a higher fraction of females?

In Table 4 I provide the results of estimating Equation 1 for different measures of female friends. Columns 1-3 make use of the question “How would you describe your friendship with this classmate?”, which allows for the following answers: (a) Not close friends; (b) A little/sometimes friends; (c) Close friends; (d) Very close friends, and construct three indicators that are increasing in the strength of the friendship.<sup>23</sup> Columns 4-6 proceed similarly with respect to the following question: “How much do you do things with this classmate outside of school?” where the options were: (a’) None; (b’) Not very much; (c’) Quite a lot; (d’) A lot. Therefore, while columns 1-3 consider self-perceived friendship, columns 4-6 acknowledge that: (i) there may still be interactions outside school with individuals that are not considered to be a friends, and (ii) that there is room for peers to have varying degrees of influence based on the total amount of interactions. Finally, columns 7-9 require both dimensions to hold simultaneously for the outcome variable to take value 1. For instance, the dependent variable in column 7 will be 1 if the answer to the first question is (b), (c) or (d) and, at the same time, the answer for the second one is (b’), (c’) or (d’).

The results are consistent across all these definitions. There is no significant relationship between the proportion of female classmates and the proportion of “close” friends or strong outside-school interactions (i.e. columns 2, 3, 5, 6, 8, and 9) whereas the effect is large and significant in the “somewhat” friends and “above-not very much” contact cases (columns 1, 4, and 7). In order to interpret the magnitude, a useful counterfactual to consider is that, if friendship formation was purely random, the estimated coefficient would be 1. Therefore, given the mean values of the dependent variables reported in the table, it becomes clear that a higher proportion of female peers significantly increases the fraction, and the total number, of nominated female friends. Indeed, in Section C.3<sup>24</sup> I argue that, under reasonable assumptions, these estimates suggest that increasing the proportion of females classmates

---

their children.

<sup>23</sup>For instance, column 1’s “at least a little close” is an indicator taking the value 1 if the child responded option (b), (c) or (d) and 0 otherwise. In a similar vein, “at least close” takes the value 1 if the answer is either (c) or (d).

<sup>24</sup>This section also provides a parallel analysis for females.

raises the proportion of female friends *more than proportionally* when using the current friendship formation patterns as counterfactual. Importantly, the nature of such interactions is relatively limited: male children are more prone to interact with females but not to the point where there is an increase in the proportion of them that become very close friends.

Overall, this unique opportunity of observing friendship nominations provides evidence that there is an actual behavioral change on the part of male children and also highlights that increased interaction of male children with female peers, even if not to the point of becoming close friends, potentially reshaped their views towards females' abilities and identity roles through fostered familiarity and friendship.

**Time Use.** I turn now to information on the children's time use collected in rounds 4 and 5 of the YL survey. Children were asked to indicate the number of hours devoted to domestic tasks and caring for others on a normal week day. This information is an attractive complement to the friendship nominations in that it pertains to observed decisions on an important dimension (time allocation) over the long-run.

In Table 5 I focus again on the subsample of male YL children. Column 1 shows that being exposed to a higher fraction of female peers in Grade 5 causes a larger amount of hours devoted to home production five years later (in round 5, column 2), while the effects are not sufficiently precisely estimated for round 4 (column 1). Because of this, in columns 3 and 4 I consider instead as outcome an indicator taking the value of 1 if the child spent more than 2 hours doing these activities, and zero otherwise. The statistical difference is now present both in the short and in the long-run. Finally, columns 5 and 6 suggest that the impact of female peers may extend also to the extensive margin of home production, at last in the long-run (column 6).<sup>25</sup>

## 5 Other Outcomes and Mechanisms

**Academic Outcomes.** Apart from the effects on gender norms, there are several other outcomes whose long-run effects are of clear interest. I first exploit another strength of the YL dataset by which participants' mathematics and verbal skills were tested in every round. This is an attractive feature since it provides actual cognitive ability, rather than having to resort to proxying for the effects on human capital through grade completion or school

---

<sup>25</sup>A regression of time allocation in round 3, i.e. prior to the School Survey using the same specification as in the Table 5 does not yield significant estimates of female peer composition, which provides further evidence that class assignment in my sample was exogenous and that the divergence in outcomes occurs after class formation. In particular the point estimate is -0.06 (p-value of 0.80) when the outcome is an indicator for providing a positive number of home or care hours and it is 0.09 (p-value 0.46) when the outcome is an indicator for contributing over 2 hours per day.

Table 4: Effects of the Proportion of Female Peers on the Female Composition of Male Children's Friendship Networks

<b>Panel A: Degree of Friendship</b>			
	(1) At Least a Little Close	(2) At Least Close	(3) Very Close
Proportion Female Peers	0.657*** (0.208)	0.116 (0.158)	-0.020 (0.187)
Mean (sd) Dep. Var.	0.35 (0.19)	0.21 (0.21)	0.12 (0.23)
Observations	441	436	382
R-squared	0.447	0.364	0.203
<b>Panel B: Contact Outside School</b>			
	(4) At Least not Many	(5) At Least Quite a Lot	(6) A Lot
Proportion Female Peers	0.573*** (0.153)	0.165 (0.188)	0.089 (0.164)
Mean (sd) Dep. Var.	0.33 (0.21)	0.20 (0.24)	0.15 (0.26)
Observations	432	425	355
R-squared	0.399	0.338	0.243
<b>Panel C: Combined Measures</b>			
	(7) 1+4	(8) 2+5	(9) 3+6
Proportion Female Peers	0.588*** (0.178)	0.077 (0.152)	0.030 (0.170)
Mean (sd) Dep. Var.	0.30 (0.22)	0.16 (0.23)	0.10 (0.22)
Observations	432	419	312
R-squared	0.417	0.299	0.201

*Notes.* All regressions replicate the specification in Table 3's column 2 for the subsample of male children. The outcomes are measured in the second round of the 2011 School Survey. The dependent variables in columns 1-3 refer to friendship nomination. Columns 4-6 refer to the frequency of doing things together. Columns 7-9 require both definitions to be satisfied at the same time. The number of observations varies as the dependent variable is the proportion of female friends *among total friends*. Therefore, if a child does not nominate any friend, regardless of the gender, for a given category this situation will lead to a missing value. Standard errors clustered at the class level (132 clusters) in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1



Table 5: Effects of the Proportion of Female Peers on the Contribution to Home Production

	(1)	(2)	(3)	(4)	(5)	(6)
	R4 Total	R5 Total	R4 Over 2	R5 Over 2	R4 Any	R5 Any
	Home Hours	Home Hours	Home Hours	Home Hours	Home Hours	Home Hours
Proportion Female Peers	0.606	2.126**	0.569**	0.489**	0.158	0.427*
	(0.799)	(0.992)	(0.272)	(0.213)	(0.276)	(0.216)
Mean (sd) Dep. Var.	1.36 (1.24)	1.64 (1.42)	0.15 (0.36)	0.18 (0.38)	0.78 (0.42)	0.86 (0.35)
Observations	433	433	433	433	433	433
R-squared	0.301	0.352	0.257	0.260	0.334	0.283

*Notes.* All regressions replicate the specification in Table 3’s column 6 for the subsample of male children. Home hours are computed as the number of hours devoted on a normal weekday to either home chores or caring for relatives in rounds 4 and 5 of the longitudinal YL survey (collected in 2013 and 2016, respectively). Columns 1 and 2 use the levels of home hours as outcome while columns 3-6 use the relevant indicator according to the column’s header. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

enrollment, which are noisy measures, particularly in developing countries (Singh, 2016). Crucially, these cognitive scores are available for every YL student, *irrespective* of her school enrollment status.

Columns 1 and 2 in Table 6 show the results for Vietnamese and mathematics at the end of the 2011-2012 academic year (i.e. the retake exam of the school survey). A higher proportion of female classmates is causally related with higher academic performance. This suggests that there are endogenous peer effects. The rest of the table asks whether these gains persist over the years. In particular, columns 3 and 4 use as outcome variables the cognitive scores from the fourth round of the YL survey (collected in 2013-2014) while columns 5 and 6 do so for the fifth round (collected in 2016-2017). While the positive effects are still significant for Vietnamese in the fourth round, these fade away over time (although they remain of a sizable positive economic size and close to being significant at conventional levels). This pattern of initially sustained cognitive spillovers that slowly turn non-significant has been commonly found for the United States (e.g. Carrell et al., 2018; Bietenbeck, 2020).

At this point, one may wonder whether the effects on academic scores are indeed driven by gender or they could simply be capturing the fact that, in Vietnamese primary schools, females perform better than males. Two pieces of evidence are against this hypothesis. First,

Table 6: Effects of the Proportion of Female Peers on Academic Outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
	Vietnamese	Math SS	Vietnamese	Vietnamese	Math	Math
	SS Round 2	Round 2	Round 4	Round 5	Round 4	Round 5
Proportion Female Peers	7.710*	7.259**	62.316**	25.574	41.685	35.887
	(4.244)	(3.275)	(26.713)	(41.065)	(29.327)	(37.199)
Observations	876	875	870	880	869	877
R-squared	0.425	0.521	0.323	0.299	0.412	0.374

*Notes.* All regressions replicate the specification in Table 3's column 2. The outcome variables in columns 1 and 2 are the raw Vietnamese and mathematics scores (out of 30 points) in the second round of the school survey. The one in columns 3-6 come from the fourth and fifth rounds of the YL data (collected in 2013 and 2016, respectively), and are standardized to have a mean of 500 and a standard deviation of 100. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

in Table A.10 I show that the estimate of the effects is robust to the inclusion of a wide range of peer characteristics, including peers' baseline test scores. Therefore, differences in peers' ability do not explain the effects of increased exposure to female classmates. Second, the size of the point estimate for mathematics and Vietnamese are about the same, while the baseline gender gap in tests scores is significantly larger in Vietnamese (19.38 for females vs. 16.27 for males) than in mathematics (16.27 vs. 16.19). Therefore, it is unlikely that the academic effects are indeed driven by females' differential academic ability.

I complement the above analysis by also looking at the effects on dropout at age 15, a time where that rate is about 20% in the YL sample. Table A.11 shows that such effects are not present. Overall, the results on academic outcomes do not seem to support the idea that the effects observed on gender norms are purely driven by more human capital accumulation.<sup>26</sup>

**Non-cognitive Skills.** Table 7 explores the effects on other school-related, non-cognitive, outcomes. These outcomes are interesting both in themselves, and also to better understand potential mechanisms. Having said this, it is clear that it is an almost impossible endeavour to attempt to explore all possible non-cognitive channels. Indeed, an important candidate,

<sup>26</sup>I explore, but do not find, evidence that either the average cognitive or enrollment results may be masking heterogeneity by gender.

Table 7: Effects of the Proportion of Female Peers on Non-cognitive Academic Outcomes

	(1)	(2)	(3)	(4)
	Low School Effort	Positive View on Teachers	Thinks Does not Do Better than Others at School	Working Hard will Not Help me to Attend College
Proportion Female Peers	-0.729* (0.437)	0.496 (0.436)	0.768 (0.630)	-0.317 (0.300)
Observations	822	842	840	865
R-squared	0.254	0.245	0.198	0.184

*Notes.* All regressions replicate the specification in Table 3's column 2 with outcome variables collected at the second round of the 2011 School Survey. A description of the outcome variables used is provided in Section D. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

namely improved class atmosphere, cannot be studied as measures for this dimension are unavailable to me. Still, valuable lessons can be learnt. Column 1 show that children exposed to more females are significantly less likely to put low academic effort. Such a peer effect arising from female classmates has previously been found in the literature and can partly explain the contemporaneous improvements in cognition uncovered above. Column 2 show that behavioral teacher responses to classroom composition may not have taken place, as students do not experience a change in their perceptions of their instructors. Similarly, there were no significant effects on the self-perception of own performance nor on the views that effort can help students manage to attend college. These results are consistent with the lack of effects on dropout rates and the eventual fade away of cognitive spillovers.<sup>27</sup>

**Fertility Preferences.** In Table 8 I explore whether the *ideal* self-reported age of first childbearing, as well as the total number of children and their gender composition, were affected by female exposure. Panel A and B report the results for males and females, respectively. Unlike for views on gender norms, here there is significant heterogeneity across genders. Males who experienced more contact with female peers wish to have significantly less children, and this is true equally for sons and daughters. On the other hand, females did not experiment any change. This leads to a convergence in the stated preferences across

<sup>27</sup>One important related aspect, however, is whether academic *aspirations* are affected by the composition of the peer group. In Table A.7 in the Online Appendix I construct an indicator variable taking the value of 1 if the child claims, in round 5 of the YL, that his/her desired level of education is at least a college degree, and 0 otherwise. While I do not find any significant impact on males, there is a strong positive effect for females. This is consistent with the finding in Figure B.7 in the Online Appendix that females' agreement with traditional life goals was largely decreased.

genders (females started from higher average levels of desired children).

**Heterogeneous Effects.** Ex ante, one could expect the effects on gender norms to be heterogeneous along multiple dimensions. For example, perhaps individuals from less wealthy families have more deeply-rooted views that are harder to change. However, the converse could very much be true. It therefore remains an empirical question whether the effects uncovered are heterogeneous and, if so, in which directions.

In Table 9's column 1 I interact the proportion of female peers with an indicator, computed using information from the second round of YL (collected in 2006), that takes the value of 1 if the child's mother was reported to have a saying in at least 50% of a set of important family decisions (enumerated in Section D), and zero otherwise. By focusing on the subset of male children, this analysis can provide important insights on how the effects of female exposure may differ across children with varying levels of initial acceptance of traditional gender roles, as proxied by the maternal decision power within the household experienced by the child prior to class formation. The findings are the following.

First, as expected, male children who grew up in households where the mother had more power in the family decisions display less traditional gender norms (-0.30, on average). Second, being exposed to more female peers continues to strongly reduce agreement with traditional norms (-0.69). These effects differ, however, based on the children's mother decision power: the effects of female exposure are significantly smaller for those children whose mother had more saying in household decisions, as shown by the positive and significant coefficient on the interaction term. These results suggest that higher exposure to female peers operated mainly through individuals who would otherwise have been likely to display more traditional views on gender norms — in a sense, affecting the “extensive margin” of agreement with more equal gender roles — rather than through increasing the equality views of children who were more prone to have them to begin with based on their family upbringing. Indeed, while both children whose mothers were and were not important decision makers benefit from female exposure, the absolute value of the effect of the increase in the proportion of female peers (from 0 to 1) is much larger for those whose mother was not (-0.694) than for those whose mother was (-0.694 + 0.643). This points to a converge towards more equal gender views across children with varying levels of baseline views, rather than to a polarization.<sup>28</sup>

---

<sup>28</sup>One may note that, in the extreme case of undertaking the thought experiment of shifting a male student whose mother was not the main decision maker from a class with 0% female peers to another one with 100%, his predicted agreement with traditional gender norms would be lower than for someone whose mother was an important decision maker (-0.694 vs. -0.694 - 0.304 + 0.643). However, for more plausible scenarios that do not require extreme extrapo-

Table 8: Effects of the Proportion of Female Peers on Fertility Preferences

<b>Panel A: Males</b>				
	(1) Age First Child	(2) Total Number of Children	(3) Number of Sons	(4) Number of Daughters
Proportion Female Peers	0.616 (1.437)	-0.861** (0.367)	-0.337* (0.203)	-0.321** (0.141)
Mean (sd) Dep. Var.	27.28 (3.02)	1.99 (0.39)	1.03 (0.25)	0.97 (0.28)
Observations	435	446	446	446
R-squared	0.328	0.235	0.162	0.257
<b>Panel B: Females</b>				
	(5) Age First Child	(6) Total Number of Children	(7) Number of Sons	(8) Number of Daughters
Proportion Female Peers	1.556 (1.394)	0.024 (0.262)	0.136 (0.138)	0.022 (0.188)
Mean (sd) Dep. Var.	26.72 (2.96)	1.95 (0.40)	0.97 (0.25)	1.00 (0.29)
Observations	418	440	439	440
R-squared	0.386	0.202	0.198	0.147

*Notes.* All regressions replicate the specification in Table 3's column 2, separately for male respondents (Panel A) and female respondents (Panel B). The outcomes are self-reported preferences for ideal quantities stated in the fifth round of the YL survey. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

In the remaining of the table I consider an extensive set of other potential sources of heterogeneity, pooling male and female children, including: (1) an indicator of any of the parents being illiterate; (2) an indicator for belonging to an ethnic minority; (3) an indicator for being located above the median in terms of the wealth index; (4) a continuous measure of self-perception of not being exceptionally good at academics; (5) a continuous measure of low effort at school, and (6) a continuous measure of having high interest about school. The results show that heterogeneous effects along these dimensions are not present: the level effects of the proportion of female peers variable continue being negative, strongly significant and of about the same initial magnitude, and its interaction terms are never significant.<sup>29</sup>

## 6 Robustness Checks

While one of the key strengths of my empirical approach (and whose validity I discussed in Section 3.2) is that it overcomes the main identification challenges in the peer effects literature, there are still certain aspects to be considered.

**No Selection out of Sample.** Section C.1 has shown that the sample employed does not show systematic differences from the original one. It could still be, however, that there is selective attrition in between the school survey and the time when I measure gender-related views (the fifth round of the YL survey). I start by noting that this is unlikely: in the fifteen-year period between the first and last waves of the YL survey, the attrition rate was remarkably low, standing at 2.5% (Young Lives, 2017). In my particular case of interest, out of the 1,138 YL children interviewed in the school survey, only 17 (1.5%) did not participate in the fifth round, and only 1 out of the 17 was part of my estimating sample for the main analyses.

Table A.8 formally shows that there are no systematic differences in predetermined characteristics by successful-round-5-matching status. More specifically, none of the dimensions considered is predictive of matching status when introduced separately in columns 1-6 nor jointly in column 7 (the p-value of the F-test of joint significance being 0.631).

**No Selection into Subsequent Peer Groups.** The key benefit of observing my outcome of interest years later is that I am able to explore the long-run effects, but it also allows for the possibility that students and families endogenously react to the initial female

---

lation — e.g. moving from the 10<sup>th</sup> (0.333) to the 90<sup>th</sup> (0.611) percentile of the distribution of female peers, the fact that children with mothers who make more decision have lower values of agreement with traditional norms (-0.304) would prevail over the marginal effects of increases in the proportion of female peers, hence making children born to mothers with more decision power still more likely to report less traditional gender views.

<sup>29</sup>I do not report the levels of the interacting variables for the sake of parsimony.

Table 9: Heterogeneous Effects of the Proportion of Female Peers on Long-term Views on Gender Norms

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Mean Agreement with Traditional Gender Norms						
Proportion Female Peers	-0.694*** (0.254)	-0.297** (0.124)	-0.267** (0.130)	-0.232* (0.128)	-0.289** (0.125)	-0.306** (0.140)	-0.250** (0.122)
Mother is Decision Maker	-0.304* (0.173)						
Mother is Decision Maker*Proportion Female Peers	0.643* (0.365)						
Any Parent Illiterate*Proportion Female Peers		0.535 (0.420)					
Ethnic Minority*Proportion Female Peers			-0.293 (0.343)				
High Wealth Index*Proportion Female Peers				-0.080 (0.219)			
Not Better Student than Other*Proportion Female Peers					-0.042 (0.100)		
Low School Effort*Proportion Female Peers						-0.160 (0.097)	
High Interest School*Proportion Female Peers							0.024 (0.091)
Sample	Males	All	All	All	All	All	All
Observations	445	886	886	886	855	850	857
R-squared	0.274	0.254	0.249	0.249	0.254	0.261	0.256

*Notes.* All regressions replicate the specification in Table 3's column 2 (although column 1 restricts the sample to male children). A description of the interacting variables used is provided in Section D. Standard errors clustered at the class level in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

peer composition by selecting into future classrooms within or across schools. Although I am not able to observe the class composition for each student after Grade 5, I provide two pieces of suggestive evidence that subsequent selective classroom formation is unlikely to explain the main findings.

First, I make use of YL’s academic history module to obtain information on whether the child continued being enrolled in the same school or transferred to a different one between Grade 5 and Grade 9<sup>30</sup>. Column 1 in Table A.12 shows no correlation between the female peer composition and whether the child changes schools at some point until Grade 9, inclusive. This is not surprising as only 3% of the children in my sample changed schools.

Second, Young Lives conducted a second School Survey in 2017 when students were already in Grade 10. While the subset of YL children that were sampled in both school surveys is relatively small, and subject to the caveat that the progression to Grade 10 is accompanied by a change in the school attended as children leave their middle schools, it is still informative to correlate the proportion of female peers that each child was exposed to in each of the two academic years surveyed in the school surveys to see if there’s persistence in the classroom gender composition faced by students. Column 2 in Table A.12 does not show an statistically significant relationship.

**Effects not Driven by Persistent Peer Matching.** One potential concern about interpreting the effects on attitudes as “long-term” ones is that female exposure could be persistent over the years in-between the school survey and the fifth round of YL. While the randomized setting and the findings in the previous point are useful in reducing these concerns, I am not able to observe class formation throughout these years. However, one particular aspect of the Vietnamese context can help us dismiss this possibility. In particular, as children transition into high school at Grade 10, there is an almost universal change in school attended, as mentioned before. This implies that classroom composition certainly takes place one year before the fifth round of the YL survey is fielded. Hence, the effects uncovered cannot be purely driven by sustained exposure, since these children were no longer exposed to the same set of peers.

**Randomization Based Inference: Placebo Peer Allocation.** In order to verify that the effects uncovered are not mechanical nor driven by unobserved factors, I provide a falsification test based on simulating random allocations of students to classes within schools, without replacement, and maintaining the original size of the classes and number of females in each class. I then run the specification in Table 3’s column 1 and replicate this exercise

---

<sup>30</sup>There is a universal school transfer in Grade 10, since students finish middle school and promote to high school.



1,000 times. If we were to find that our baseline effect is consistent with those ones obtained in this placebo analysis, where students are assigned to classmates who were not their actual peers, we may conclude that the main effects were not actually driven by interacting with those peers. Reassuringly, the baseline estimate of -0.274 lies outside the 90% confidence interval obtained from the simulations, which is [-0.211, 0.218].<sup>31</sup>

**No Changes in Educational Investments.** It could be the case that parental investments and views on their child’s education react to the composition of the class. While, unfortunately, the second round of the school survey does not feature information that could be useful in this sense, I can exploit related information from the fourth wave of the YL survey.

Table A.14 shows that the proportion of female peers is not correlated with the extensive margin of participation in extra tutorials outside school (column 1), nor on the total number of hours or money spent on such activity (columns 2 and 3, respectively). Moreover, there was no change in the ideal education level that the parents wish their child achieves nor in whether they claim to know the names of their child’s friends (which I take to capture parental involvement and time investment on the child). These findings, coupled with the lack of changes in classroom composition years later mentioned above, suggest that parental endogenous reactions to class composition are not present.

**Quantile Regressions.** The study of heterogeneous effects suggests that there are larger shifts away from traditional norms for those children whose mother had a lower weight in family decisions. A complementary view is to undertake a quantile analysis, which will be informative of the parts of the distribution of views on gender norms that drive the baseline, average effects, reported in Table 3. The results in Table A.15 confirm a tendency for convergence towards less traditional views, but these effects are primarily located around the median of the distribution of attitudes. This suggests that, while attitudes are relatively malleable for children with some initial predisposition to more egalitarian views, extremely traditional views are harder to change, and some divergence from the rest of the — more modern — distribution took place.

**Alternative Summary Index for Views on Gender Norms.** The baseline analysis has made use of an averaged measure of raw views on gender norms. In this section I follow the widely used approach in Kling et al. (2007) where I first standardize each of

---

<sup>31</sup>An alternative approach to this exercise is to compute an empirical p-value as the proportion of total cases in which the absolute value of the point estimate is bigger than the one obtained in my estimation using actual data (Athey and Imbens, 2017). There are 42 such cases, hence leading to a p-value of 0.042.

the twelve dimensions of gender norms considered so that they have a mean of zero and a standard deviation of one (as in Table 1), then I obtain their average value for each individual, and I standardize this value again across all children. This approach was conceived to increase statistical power for same-directional effects. In order to explore the reliability of this measure I compute the coefficient alpha (Cronbach's alpha) of the twelve items used in the computation of the summary score, which yields a scale coefficient of 0.71 — a value widely accepted as reliable in the personality psychology literature (Kline, 2000). Table A.16 shows that my main results are robust: a ten percentage points increase in female peers decreases the index measure of agreement with traditional gender norms by 8.6% of a standard deviation when I do not account for partial observability of peers, and by 15.2% when I do — the same sizes as in Table 3.

**Results Using the Full YL Sample.** While an attractive contextual feature is that almost 80% of the students are randomly assigned to classes — and an additional 15% of the original sample are exogenously assigned, which I also include in my main analysis<sup>32</sup> — it is relevant to explore whether the main results hold for the full sample, where some selection into classes could be present and bias the results. Table A.17 shows that the main qualitative results hold, which further increases the external validity of the baseline results.<sup>33</sup>

**Inclusion of Further Controls.** I complement the multiple exercises undertaken with the aim of showing the validity of my identification assumption by re-estimating my baseline regression with the inclusion of a rich set of additional controls — at the individual and peer levels — both for cognitive abilities (the baseline mathematics score in the SS survey) and for non-cognitive and attitudinal dimensions, e.g. effort exerted at school or the perception that by working hard the child can achieve college. The results in Table A.18 show that the point estimate, once again, remains virtually unaltered. This not only reinforces the credibility of my empirical exercise but also provides strong evidence that it is indeed exposure to female students, and not gender-correlated attributes of the peers, that is driving the changes in views towards gender roles.

---

<sup>32</sup>The figures, in levels, are the following. Out of the total 3,264 students in the School Survey, 3,071 satisfy my sample selection criteria.

<sup>33</sup>One may note that the point estimate using the full sample is bigger in absolute value than the baseline one (-0.378 vs. -0.277). This suggests that, among the non-exogenously allocated classes, there may be a positive correlation between a determinant of less traditional gender norms and the proportion of female peers. This is consistent with the fact that the excluded classes are formed on the basis of academic achievement and that Vietnamese girls perform significantly better than males, particularly in mathematics.

## 7 Conclusion

Identity norms play a key role throughout a person’s life. In the aggregate, they have important consequences on welfare as they can, for instance, lead to severe misallocation of talent across occupations. In this paper I aim at answering a novel question within the peer effect literature: does exposure to more female classmates during childhood shift attitudes towards gender norms and, if so, how persistent are these changes?

My focus is on Vietnam, a country with marked gender gaps and strong traditional gender roles, even among the current youth. I take advantage of an ideal setting that allows me to estimate the causal effects of female peers on the long-run views on gender norms in the absence of selection into peers and other commonplace threats to identification in the peer effects literature ([Manski, 1993](#)). A rich battery of tests lends strong support to the identification assumption based on exogenous assignment of students to classes within schools. The main result is that a ten percentage points increase in the proportion of female classmates leads to a reduction of a tenth of a standard deviation in the agreement with more traditional views.

While a concern could be that survey responses may not translate into actual behavior, I exploit information on friendship nominations to show that male children spend more time and develop more ties with female peers. This finding also suggests that exposure and familiarity with females is likely an important channel explaining the shifts in views uncovered. This claim is supported by the fact that the spillovers from female peers into academic outcomes — both cognitive and noncognitive — are relatively minor and short-lived, which is at odds with a story of pure improvements in education leading to more inclusive views. I complement this behavioral exploration by showing that home hours, a fundamental allocation of a person’s available time previously found to be closely related with views on gender roles, are also significantly larger for male children exposed to larger fractions of female peers.

Overall, my findings provide a rich picture of gender peer effects, both in the scope of outcomes explored and in the extended time-frame along which they are observed. This approach, which is novel in the literature and highlights the limitations of exclusively accounting for academic spillovers, is achieved by means of a clean identification strategy with an arguably large external validity, and is crucial to not obviate the socio-emotional gains of certain class configurations and to better inform policy on class formation. For instance, identifying a pathway towards less women facing less traditional gender roles in the society

is prone to help increase female labor force participation and career development opportunities, particularly in the numerous sectors where they are under-represented, which is a long-standing concern in a country where one in every five job postings makes explicit gender requirements ([International Labour Organization, 2015](#)). This, in turn, would have implications for the female spouse’s bargaining power within the household, as well as intergenerational spillovers arising from the transmission of attitudes from parents to children and the fact that child outcomes have been shown to be enhanced when the mother has a heavier weight in family decisions (e.g. [Hoddinott and Haddad, 1995](#); [Duflo, 2003](#)).

## References

- Anelli, Massimo and Giovanni Peri**, “The Effects of High School Peers’ Gender on College Major, College Performance and Income,” *The Economic Journal*, 2019.
- Angrist, Joshua D**, “The perils of peer effects,” *Labour Economics*, 2014, *30*, 98–108.
- Ashraf, Nava, Natalie Bau, Corinne Low, and Kathleen McGinn**, “Negotiating a better future: How interpersonal skills facilitate intergenerational investment,” *The Quarterly Journal of Economics*, 2020, *135* (2), 1095–1151.
- Athey, Susan and Guido W Imbens**, “The state of applied econometrics: Causality and policy evaluation,” *Journal of Economic Perspectives*, 2017, *31* (2), 3–32.
- Benjamini, Yoav and Yosef Hochberg**, “Controlling the false discovery rate: a practical and powerful approach to multiple testing,” *Journal of the Royal Statistical Society: Series B (Methodological)*, 1995, *57* (1), 289–300.
- Bertrand, Marianne, Emir Kamenica, and Jessica Pan**, “Gender identity and relative income within households,” *The Quarterly Journal of Economics*, 2015, *130* (2), 571–614.
- Bietenbeck, Jan**, “The long-term impacts of low-achieving childhood peers: Evidence from project STAR,” *Journal of the European Economic Association*, 2020, *18* (1), 392–426.
- Black, Sandra E, Paul J Devereux, and Kjell G Salvanes**, “Under pressure? The effect of peers on outcomes of young adults,” *Journal of Labor Economics*, 2013, *31* (1), 119–153.
- Blundell, Jack and Erling Risa**, “Income and Family Background: Are we Using the Right Models?,” *Available at SSRN 3269576*, 2019.

- Bursztyn, Leonardo, Thomas Fujiwara, and Amanda Pallais**, “‘Acting Wife’: Marriage Market Incentives and Labor Market Investments,” *American Economic Review*, 2017, *107* (11), 3288–3319.
- Carrell, Scott E, Bruce I Sacerdote, and James E West**, “From natural variation to optimal policy? The importance of endogenous peer group formation,” *Econometrica*, 2013, *81* (3), 855–882.
- , **Mark Hoekstra, and Elira Kuka**, “The long-run effects of disruptive peers,” *American Economic Review*, 2018, *108* (11), 3377–3415.
- , – , and **James E West**, “The impact of college diversity on behavior toward minorities,” *American Economic Journal: Economic Policy*, 2019, *11* (4), 159–82.
- Chetty, Raj, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan**, “How does your kindergarten classroom affect your earnings? Evidence from Project STAR,” *The Quarterly Journal of Economics*, 2011, *126* (4), 1593–1660.
- Choudhury, Suparna, Sarah-Jayne Blakemore, and Tony Charman**, “Social cognitive development during adolescence,” *Social cognitive and affective neuroscience*, 2006, *1* (3), 165–174.
- Dahl, Gordon, Andreas Kotsadam, and Dan-Olof Rooth**, “Does integration change gender attitudes? The effect of randomly assigning women to traditionally male teams,” *Quarterly Journal of Economics Forthcoming*, forthcoming.
- Duflo, Esther**, “Grandmothers and granddaughters: old-age pensions and intrahousehold allocation in South Africa,” *The World Bank Economic Review*, 2003, *17* (1), 1–25.
- Feld, Jan and Ulf Zölitz**, “Understanding peer effects: On the nature, estimation, and channels of peer effects,” *Journal of Labor Economics*, 2017, *35* (2), 387–428.
- Fruehwirth, Jane Cooley and Jessica Gagete-Miranda**, “Your peers’ parents: Spillovers from parental education,” *Economics of Education Review*, 2019, *73*, 101910.
- Gong, Jie, Yi Lu, and Hong Song**, “The effect of teacher gender on students’ academic and noncognitive outcomes,” *Journal of Labor Economics*, 2018, *36* (3), 743–778.

- Gould, Eric D, Victor Lavy, and M Daniele Paserman**, “Does immigration affect the long-term educational outcomes of natives? Quasi-experimental evidence,” *The Economic Journal*, 2009, *119* (540), 1243–1269.
- Guilmoto, Christophe**, “Recent change in the sex ratio at birth in Viet Nam: a review of evidence,” 2009.
- Guryan, Jonathan, Kory Kroft, and Matthew J Notowidigdo**, “Peer effects in the workplace: Evidence from random groupings in professional golf tournaments,” *American Economic Journal: Applied Economics*, 2009, *1* (4), 34–68.
- Hanaoka, Chie, Hitoshi Shigeoka, and Yasutora Watanabe**, “Do risk preferences change? evidence from the great east japan earthquake,” *American Economic Journal: Applied Economics*, 2018, *10* (2), 298–330.
- Hanushek, Eric A**, “The failure of input-based schooling policies,” *The Economic Journal*, 2003, *113* (485), F64–F98.
- Hoddinott, John and Lawrence Haddad**, “Does female income share influence household expenditures? Evidence from Côte d’Ivoire,” *Oxford Bulletin of Economics and Statistics*, 1995, *57* (1), 77–96.
- Hoxby, Caroline**, “Peer effects in the classroom: Learning from gender and race variation,” Technical Report, National Bureau of Economic Research 2000.
- Hsieh, Chang-Tai, Erik Hurst, Charles I Jones, and Peter J Klenow**, “The allocation of talent and us economic growth,” *Econometrica*, 2019, *87* (5), 1439–1474.
- Hwang, Jisoo, Chulhee Lee, and Esther Lee**, “Gender norms and housework time allocation among dual-earner couples,” *Labour Economics*, 2019, *57*, 102–116.
- Hyun, Yeseul**, “Does Intrafamily Time Allocation Respond to Economic Shocks? The Role of Gender Norms,” 2020.
- International Labour Organization**, “Gender Equality in Recruitment and Promotion Practices in Viet Nam,” ILO’s Policy Brief 2015.
- Jackson, C Kirabo, Jonah E Rockoff, and Douglas O Staiger**, “Teacher effects and teacher-related policies,” *Annu. Rev. Econ.*, 2014, *6* (1), 801–825.
- Kline, Paul**, *The handbook of psychological testing*, Psychology Press, 2000.

- Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz**, “Experimental analysis of neighborhood effects,” *Econometrica*, 2007, 75 (1), 83–119.
- Kremer, Michael and Dan Levy**, “Peer effects and alcohol use among college students,” *Journal of Economic Perspectives*, 2008, 22 (3), 189–206.
- Krueger, Alan B**, “Experimental estimates of education production functions,” *The Quarterly Journal of Economics*, 1999, 114 (2), 497–532.
- Lovenheim, Michael F and C Lockwood Reynolds**, “The effect of housing wealth on college choice: Evidence from the housing boom,” *Journal of Human Resources*, 2013, 48 (1), 1–35.
- Lowe, Matt**, “Types of contact: A field experiment on collaborative and adversarial caste integration,” *American Economic Review*, forthcoming.
- Manski, Charles F**, “Identification of endogenous social effects: The reflection problem,” *The Review of Economic Studies*, 1993, 60 (3), 531–542.
- Oreopoulos, Philip and Kjell G Salvanes**, “Priceless: The nonpecuniary benefits of schooling,” *Journal of Economic Perspectives*, 2011, 25 (1), 159–84.
- Rao, Gautam**, “Familiarity does not breed contempt: Generosity, discrimination, and diversity in Delhi schools,” *American Economic Review*, 2019, 109 (3), 774–809.
- Rolleston, Caine, Zoe James, Laure Pasquier-Doumer, and Tran Ngo Thi Minh Tam**, *Making progress: Report of the Young Lives school survey in Vietnam*, Young Lives, 2013.
- Sacerdote, Bruce**, “Peer effects with random assignment: Results for Dartmouth roommates,” *The Quarterly Journal of Economics*, 2001, 116 (2), 681–704.
- Shigeoka, Hitoshi**, “Long-term consequences of growing up in a recession on risk preferences,” Technical Report, National Bureau of Economic Research 2019.
- Singh, Abhijeet**, “Learning more with every year: School year productivity and international learning divergence,” *Journal of the European Economic Association*, 2016.
- Sojourner, Aaron**, “Identification of peer effects with missing peer data: Evidence from Project STAR,” *The Economic Journal*, 2013, 123 (569), 574–605.

**World Bank**, “Vietnam: Breaking gender stereotypes that hinder women’s empowerment,”  
World Bank’s East Asia Pacific on the Rise Series 2015.

**Young Lives**, “Young Lives Survey Design and Sampling in Vietnam,” 2014.

— , “Young Lives Survey Design and Sampling (Round 5): Viet Nam,” 2017.



## A Additional Tables (For Online Publication)

### Life Purpose

Table A.1: Descriptive Statistics of Attitudes towards Gender Norms for the Old Cohort

	(1)	(2)	(3)	(4)	(5)
	Mean	Standard Deviation	Min	Max	Count
<b>Life Purpose</b>					
Life Goals	0.56	0.50	0	1	910
Freedom	0.05	0.22	0	1	909
Family Encouragement for College	0.16	0.37	0	1	908
Importance of Good Academic Performance	0.15	0.36	0	1	910
<b>Abilities</b>					
Leadership Ability	0.45	0.50	0	1	908
Intelligence	0.06	0.23	0	1	908
<b>Cross-Gender Interactions</b>					
Asking for a Date	0.11	0.32	0	1	909
Pay for a Date	0.55	0.50	0	1	908
Male-dominated Sports	0.07	0.26	0	1	909
Swearing	0.58	0.49	0	1	908
<b>Intra-household Decisions</b>					
Authority in Household Decisions	0.49	0.50	0	1	909
Chores	0.05	0.21	0	1	909

*Notes.* All variables are indicators taking the value one if the child agrees or strongly agrees with a traditional view on gender norms for each dimension of interest. The exact questions were reported in the “Views on Gender Norms (longitudinal YL)” subsection. Respondents are from the Old Cohort and are interviewed in the fifth round of the YL survey (2016).

Table A.2: Balance of Characteristics at Round 3 by Status for School Survey Inclusion

<b>Panel A</b>						
	(1) Rural	(2) Female	(3) BMI-for-age z-score	(4) Health	(5) Enrolled	(6) Father's Education
In School Survey	-0.000 (0.003)	0.010 (0.026)	-0.029 (0.067)	0.003 (0.035)	0.036*** (0.007)	0.075 (0.162)
Observations	1,961	1,961	1,926	1,954	1,911	1,869
R-squared	0.915	0.005	0.138	0.075	0.082	0.291
<b>Panel B</b>						
	(7) Mother's Education	(8) Father's age	(9) Mother's age	(10) Household Size	(11) Atheist	(12) Wealth Index
In School Survey	0.016 (0.162)	0.054 (0.294)	0.400 (0.285)	-0.043 (0.068)	-0.009 (0.017)	0.015** (0.007)
Observations	1,929	1,883	1,940	1,961	1,960	1,935
R-squared	0.335	0.094	0.086	0.099	0.124	0.503
<b>Panel C</b>						
	(13) Plans the Future	(14) Studying Hard Will be Rewarded	(15) Friends Look Up to You as a Leader?	(16) Included in Other Children Games?	(17) Hard to Talk to Other Children?	(18) Position in Ladder of Best Possible Life
In School Survey	0.052 (0.036)	0.040 (0.030)	-0.057 (0.037)	-0.032 (0.020)	0.005 (0.024)	0.189* (0.108)
Observations	1,893	1,914	1,907	1,947	1,942	1,928
R-squared	0.057	0.063	0.050	0.054	0.024	0.135

*Notes.* Regressions of the indicated outcome variables (measured in 2009 for Round 3 of the YL survey) on an indicator taking the value 1 if the Young Lives child was sampled for the School Survey and 0 otherwise and sentinel site fixed effects. The sample are all Young Lives students present in round 3. Robust standard errors in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A.3: Balance of Characteristics between YL Children and their Classmates at the 2011 School Survey

<b>Panel A</b>						
	(1) Female	(2) Any Parent Illiterate	(3) Ethnic Minority	(4) No Health Problems	(5) Home Educational Resources	(6) Wealth Index
YL Child	0.026 (0.020)	-0.011 (0.010)	-0.012 (0.009)	0.012 (0.019)	-0.006 (0.010)	-0.003 (0.006)
Observations	2,956	2,965	2,962	2,971	2,971	2,971
R-squared	0.020	0.439	0.680	0.096	0.329	0.385
<b>Panel B</b>						
	(7) Number Books	(8) If I Work Hard I can go to College	(9) No Rein- forcement Class	(10) Enjoy School	(11) Worry about Exams	
YL Child	-0.046 (0.042)	-0.024 (0.032)	-0.002 (0.017)	-0.004 (0.011)	-0.011 (0.028)	
Observations	2,953	2,942	2,971	2,945	2,942	
R-squared	0.271	0.084	0.350	0.051	0.094	

*Notes.* Regressions of the indicated outcome variables (measured in the first round of the school survey) on an indicator taking the value 1 if the child belongs to the longitudinal Young Lives study and 0 otherwise and school fixed effects. The sample contains all students with non-missing values present in the School Survey whose class-assignment satisfied the sample selection criteria outlined in Section 2.1. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.4: Does Peers' Female Composition Predict Observable Individual Characteristics?

<b>Panel A</b>					
	(1) Ethnic Minority	(2) Repeater	(3) Any Health Issue	(4) Any Parent Illiterate	(5) Number of Books at Home
Proportion Female Peers	-0.017 (0.039)	-0.021 (0.034)	-0.110 (0.108)	-0.054 (0.043)	0.192 (0.374)
Observations	2,886	2,883	2,894	2,889	2,877
R-squared	0.674	0.048	0.099	0.427	0.267
<b>Panel B</b>					
	(6) Wealth Index	(7) Home Educational Resources	(8) Takes Private Classes	(9) No Rein- forcement Class	
Proportion Female Peers	0.086* (0.050)	0.047 (0.055)	-0.097 (0.211)	0.070 (0.209)	
Observations	2,894	2,894	2,887	2,894	
R-squared	0.387	0.316	0.347	0.348	
<b>Panel C: Child's Baseline Attitudes</b>					
	(10) Working Hard I Can Go to College	(11) Feel Pressure at School?	(12) Difficult to Complete Homework?	(13) Hit by Parents?	
Proportion Female Peers	0.093 (0.211)	-0.219 (0.237)	-0.050 (0.207)	0.056 (0.172)	
Observations	2,867	2,856	2,846	2,868	
R-squared	0.071	0.092	0.082	0.085	

*Notes.* All outcome variables in Panel A are indicators (except the 4-category variable for number of books at home). A description of the variables in Panel C is provided in Section D. All outcomes were measured in the first round of the 2011 School Survey. Regressions control for school fixed effects. The sample contains all students with non-missing values present in the School Survey whose class-assignment satisfied the sample selection criteria outlined in Section 2.1. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.5: Does Peers' Female Composition Correlate with YL Children's Round 3 Characteristics?

<b>Panel A</b>								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Perceived School Performance	Teacher Punished Classmate	Teacher Punished Self	Safe in Neighborhood	Can Count on Friends	Try Hard Can Improve Life	Study Hard For Better Job	Plans Ahead
Proportion Female Peers	-0.135 (0.197)	-0.026 (0.177)	0.041 (0.229)	-0.068 (0.178)	-0.044 (0.080)	-0.054 (0.097)	0.047 (0.058)	0.087 (0.110)
Observations	943	943	916	945	944	936	941	933
R-squared	0.107	0.091	0.117	0.121	0.096	0.128	0.068	0.136

<b>Panel B</b>								
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
	Included in Others' Games	Happy About: Owned Things	Happy About: Health	Happy About: Things Wish Be Good At	Happy About: Friends	Happy About: Safety	Happy About: Doing Things Away from Home	Happy About: Future
Proportion Female Peers	0.062 (0.109)	-0.054 (0.123)	-0.237 (0.207)	-0.099 (0.124)	-0.335** (0.166)	0.155 (0.156)	0.038 (0.075)	-0.073 (0.132)
Observations	944	945	945	941	941	943	943	902
R-squared	0.131	0.106	0.115	0.151	0.100	0.101	0.137	0.152

*Notes.* All outcome variables are measured in Round 3 of the Young Lives survey, which was conducted in 2009. The sample includes all YL children who were also surveyed in the School Survey and whose class-assignment satisfied the sample selection criteria outlined in Section 2.1. A description of the variables used is provided in Section D. Regressions control for school fixed effects. Standard errors clustered at the class level in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Table A.6: Are Teacher Characteristics Correlated with Peers' Female Composition?

<b>Panel A</b>					
	(1) Female	(2) Highest Training	(3) Temporary Contract	(4) Has Multiple Jobs	(5) Current Years of Tenure
Proportion Female Peers	0.397 (0.253)	0.108 (0.538)	-0.066 (0.146)	-0.005 (0.255)	-4.498 (7.233)
Observations	2,894	2,875	2,874	2,894	2,894
R-squared	0.569	0.597	0.261	0.582	0.445

<b>Panel B</b>				
	(6) Total Years of Tenure	(7) Wealth Index	(8) Winner of Excellent Teacher Award	(9) Teacher Believes Possible to Help Worst Students
Proportion Female Peers	2.519 (6.254)	0.125 (0.081)	0.246 (0.218)	-0.106 (0.762)
Observations	2,924	2,924	2,924	2,904
R-squared	0.519	0.688	0.379	0.576

*Notes.* Regressions of the indicated teacher characteristics (measured in the first round of the 2011 School Survey) on the proportion of female peers and school fixed effects. The sample contains all students with non-missing values present in the School Survey whose class-assignment satisfied the sample selection criteria outlined in Section 2.1. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.7: Effects of the Proportion of Female Peers on Academic Aspirations

	(1)	(2)
	Aspires to Reach University	
Proportion Female Peers	-0.100 (0.303)	0.300** (0.133)
Sample	Males	Females
Observations	428	431
R-squared	0.384	0.365

*Notes.* These regressions replicate the specification in Table 3's column 2 separated for the subsamples of male and female children. The dependent variable is an indicator taking the value of 1 if the child stated that his/her desired level of academic achievement would be at least a college degree. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.8: Non-selective Attrition

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Matched with YL's Round 5						
Female	-0.001 (0.008)						-0.002 (0.007)
Ethnic Minority		0.004 (0.009)					-0.001 (0.012)
Home Educational Resources			-0.018 (0.016)				-0.013 (0.017)
Literate Mother				-0.026 (0.021)			-0.026 (0.024)
Low School Effort					-0.001 (0.003)		-0.002 (0.003)
Wealth Index						-0.042 (0.037)	-0.035 (0.040)
In Estimating Sample							0.000 (0.003)
P-value F-test Joint Significance							0.631
Observations	1,129	1,132	1,138	1,128	1,090	1,138	1,069
R-squared	0.183	0.183	0.184	0.184	0.183	0.185	0.186

*Notes.* Regressions of an indicator taking the value 1 if the YL child surveyed in the 2011 School Survey was also present (“matched”) with the attitudinal information from YL’s round 5 (collected in 2016) and 0 otherwise on the baseline characteristics indicated in the rows and measured in the first round of the school survey and school fixed effects. I do not report a separate regression for “In Estimating Sample” as only 1 out of the 17 children who are not matched was not part of my estimating sample. Standard errors clustered at the class level in parentheses. \*\*\* p<0.01, \*\* p<0.05, \* p<0.1



Table A.9: Effects of the Proportion of Female Peers on the Female Composition of Female Children's Friendship Networks

<b>Panel A: Degree of Friendship</b>			
	(1) At Least a Little Close	(2) At Least Close	(3) Very Close
Proportion Female Peers	1.049*** (0.154)	0.241 (0.147)	-0.004 (0.106)
Mean (sd) Dep. Var.	0.63 (0.22)	0.81 (0.23)	0.89 (0.23)
Observations	438	432	368
R-squared	0.428	0.358	0.330
<b>Panel B: Contact Outside School</b>			
	(4) At Least not Many	(5) At Least Quite a Lot	(6) A Lot
Proportion Female Peers	0.882*** (0.159)	0.421*** (0.121)	0.158 (0.169)
Mean (sd) Dep. Var.	0.67 (0.23)	0.80 (0.25)	0.87 (0.26)
Observations	429	405	325
R-squared	0.379	0.351	0.268
<b>Panel C: Combined Measures</b>			
	(7) 1+4	(8) 2+5	(9) 3+6
Proportion Female Peers	0.947*** (0.166)	0.315** (0.129)	0.128 (0.142)
Mean (sd) Dep. Var.	0.71 (0.24)	0.85 (0.23)	0.91 (0.22)
Observations	427	391	282
R-squared	0.397	0.377	0.417

*Notes.* All regressions replicate the specification in Table 4 for the subsample of female children. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.10: Effects on Cognition: Robustness to Including Peers Baseline Cognition

	(1)	(2)
	Vietnamese Raw	Mathematics Raw
	Score Round 2	Score Round 2
Proportion Female Peers	7.298** (3.304)	7.653** (2.964)
Observations	876	875
R-squared	0.445	0.544

*Notes.* Replications of the first two columns in Table 6 with the inclusion of controls for the following peer characteristics: age, parental literacy, ethnic minority, wealth, the number of books in the household, grade repeaters, school interest, and baseline raw test score, together with school fixed effects. Standard errors clustered at the class level in parentheses.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.11: Is Long-run School Enrollment Affected by Peers' Female Composition?

	(1)	(2)
	Currently Enrolled	Currently Enrolled
	Round 4	Round 5
Proportion Female Peers	0.002 (0.013)	-0.157 (0.168)
Observations	883	883
R-squared	0.111	0.230

*Notes.* Replication of the specification in Table 3's column 2 for enrollment rates of YL children in rounds 4 (column 1) and 5 (column 5). Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.12: Is There Selection into Subsequent Academic Peer Groups?

	(1)	(2)
	Ever Transfers	Proportion Female
	Schools	Peers 2017
Proportion Female Peers 2011	0.007 (0.052)	0.225 (0.214)
Observations	884	221
R-squared	0.134	0.447

*Notes.* Replication of the specification in Table 3's column 2 for: (1) an indicator taking the value 1 if the child switches schools between Grade 6 and Grade 9 as retrieved from round 5 of the YL survey; (2) the proportion of female classmates in Grade 10, as measured in the 2017 School Survey (hence the observations are restricted to YL children who participated both in the 2011 and in the 2017 school surveys). Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.13: Robustness to Accounting for Partial Peer Unobservability and to Different Sample Selections

	(1)	(2)	(3)	(4)	(5)
	Mean	Friendship	Vietnamese	Math	Low School
	Norms	1+4	Round 2	Round 2	Effort
Proportion Female Peers	-0.273*	0.602***	5.397	3.906	-0.420
	(0.142)	(0.155)	(3.888)	(3.569)	(0.455)
Observations	730	353	876	875	822
R-squared	0.298	0.441	0.447	0.519	0.251

*Notes.* Replication of the specification in Table 3's column 2 with the addition of the interaction between school fixed effects and the proportion of peers observed in the class. Columns 1 and 2 use only individuals who were randomly assigned to classes. Columns 3-5 use the same sample as in the main text. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.14: Are Parental Investments Affected by Peers' Female Composition?

	(1)	(2)	(3)	(4)	(5)
	Any	Hours Extra	Money	Ideal Education	Knows Names of
	Extra Class	Classes	Extra Classes	Level Reached	Child's Friends
Proportion Female Peers	0.068	0.575	-95.251	0.114	0.006
	(0.135)	(1.978)	(342.070)	(0.353)	(0.143)
Observations	882	881	876	863	884
R-squared	0.376	0.359	0.220	0.178	0.169

*Notes.* Replication of the specification in Table 3's column 2 for parental investment information collected in round 4 of the YL survey. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.15: Quantile Regressions of the Effects of Peers' Female Composition on Views on Gender Norms

	(1)	(2)	(3)	(4)	(5)
Mean Agreement with Traditional Gender Norms					
	Q. 0.2	Q. 0.4	Q. 0.5	Q. 0.6	Q. 0.8
Proportion Female Peers	-0.216 (0.247)	-0.356** (0.177)	-0.394* (0.215)	-0.464** (0.202)	-0.018 (0.266)
Observations	446	446	446	446	446
R-squared	0.134	0.204	0.224	0.209	0.161

*Notes.* Quantile regressions for the male subsample using the same controls as in Table 3's column 2. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.16: Robustness of Effects on Gender Norms to Standardizing the Outcome

	(1)	(2)
Standardized Norms		
Proportion Female Peers	-0.856** (0.403)	-1.524*** (0.324)
Peers Correction	No	Yes
Observations	886	880
R-squared	0.252	0.292

*Notes.* Replication of the specification in Table 3's columns 2 and 6 where the outcome variable is constructed first standardizing each of its components, then obtaining their mean, and standardizing it again. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table A.17: Robustness of Effects on Gender Norms to Using the Full Sample

	(1)
	Mean Agreement
Proportion Female Peers	-0.378*** (0.105)
Observations	1,032
R-squared	0.251

*Notes.* Replication of the specification in Table 3's column 2 using the full sample (i.e. no sample selection undertaken based on classroom formation rules). Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

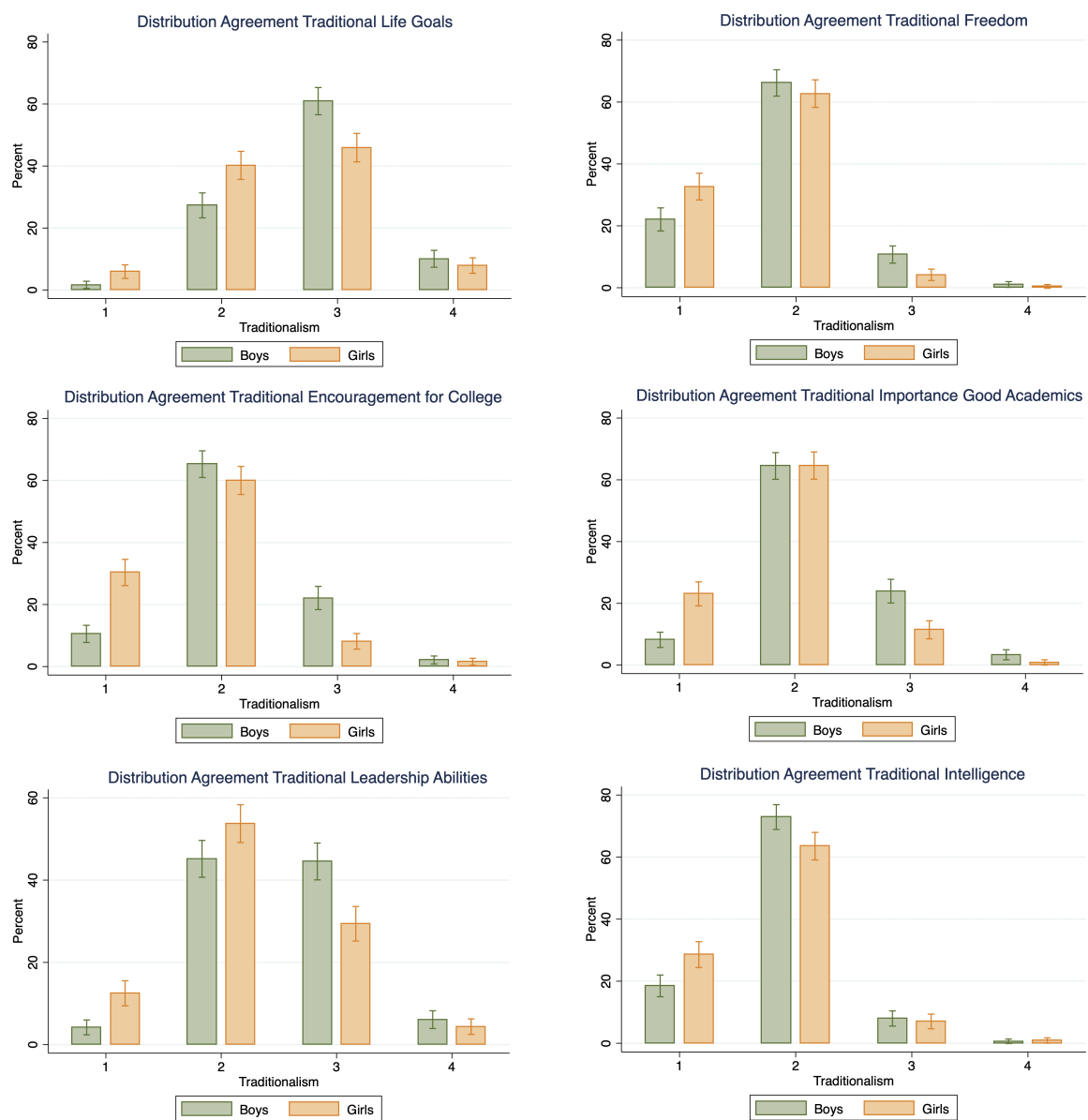
Table A.18: Robustness of Effects on Gender Norms to the Inclusion of Further Controls

	(1)
	Mean Agreement
Proportion Female Peers	-0.281* (0.148)
Observations	798
R-squared	0.303

*Notes.* Replication of the specification in Table 3's column 2 with the inclusion of the following additional controls both at the individual and peer levels: worry about exams, baseline mathematics score, low school effort, high interest in school, difficult to complete homework, feel pressure at school, if I work hard I can go to college as defined in Section C.1. Standard errors clustered at the class level in parentheses. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

# B Additional Figures (For Online Publication)

Figure B.1: Distribution Views on the Different Dimensions of Traditionalism by Gender<sup>34</sup>



<sup>34</sup>The horizontal axis labels are as follows: 1: strongly disagree; 2: disagree; 3: agree; 4: strongly agree. 95% confidence intervals are also reported.



Figure B.2: Distribution Views on the Different Dimensions of Traditionalism by Gender (Continued)

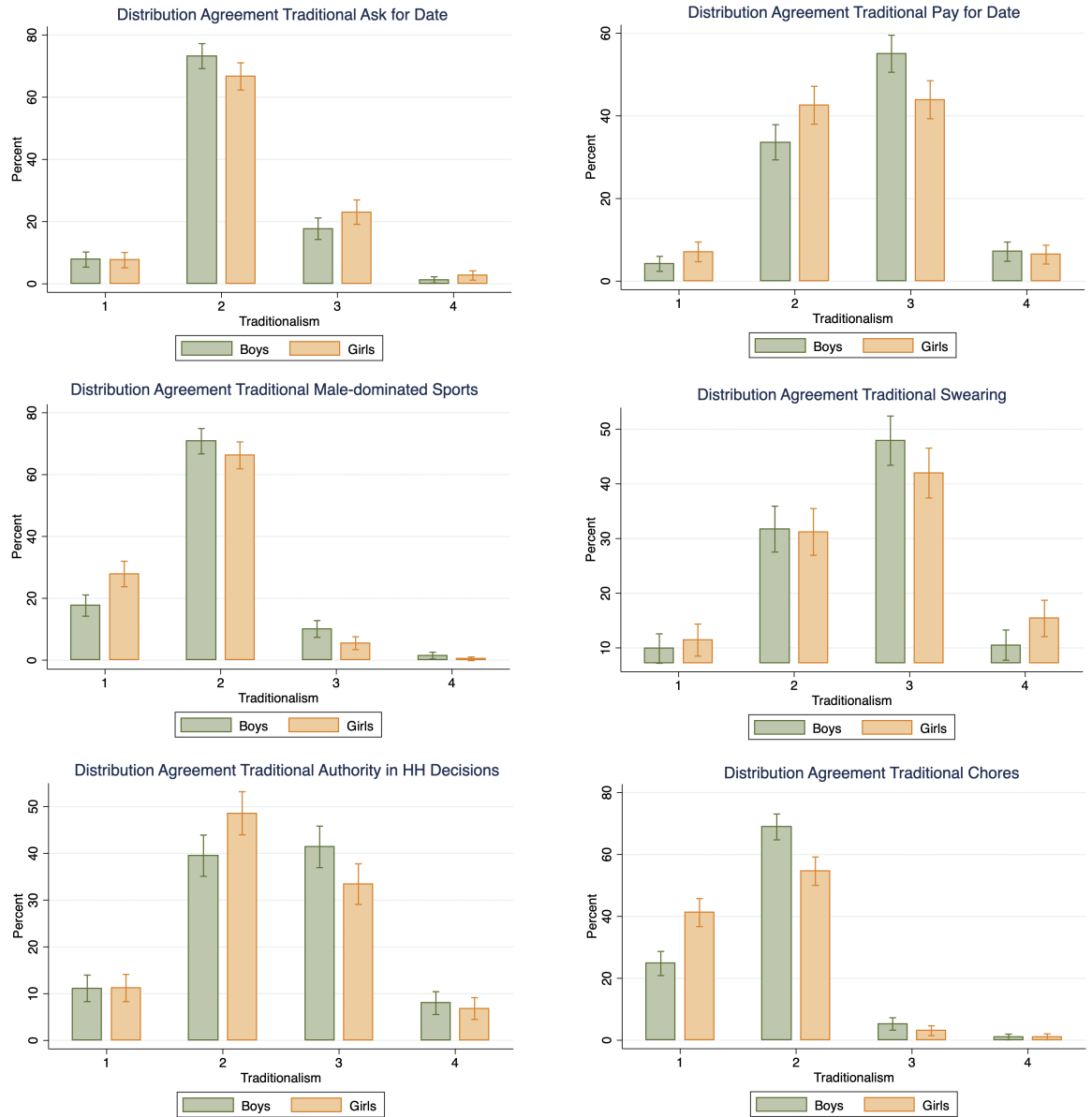


Figure B.3: Distribution of the Proportion of Female Peers

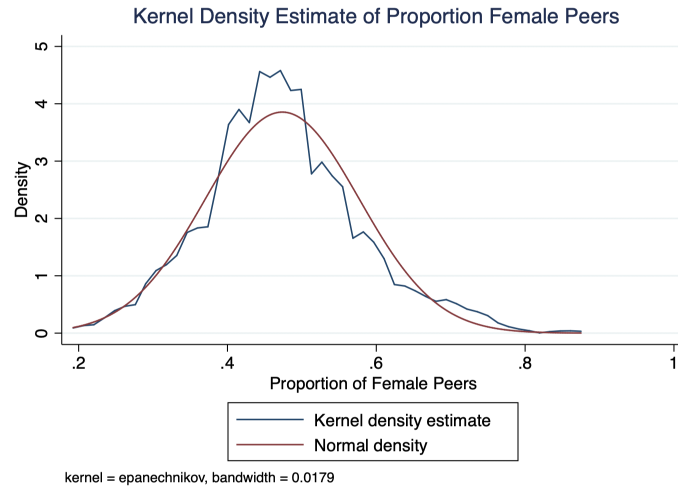


Figure B.4: Distribution of Actual and Simulated Variation in Female Peers

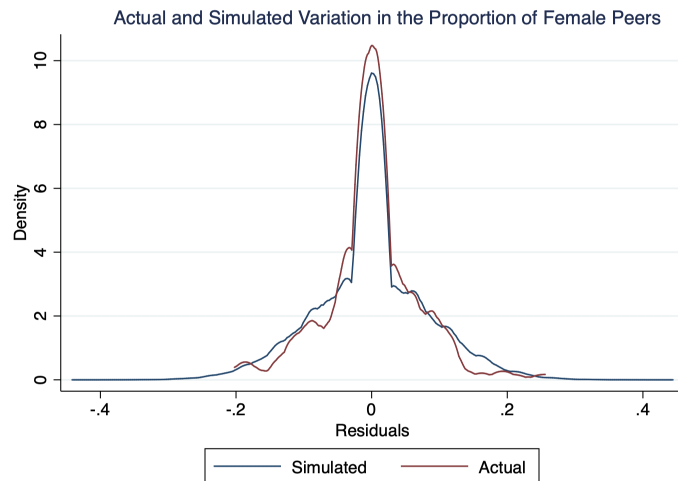


Figure B.5: Distribution of the Difference Between Actual and Simulated Proportion of Female Peers for One Randomly-selected Class per School

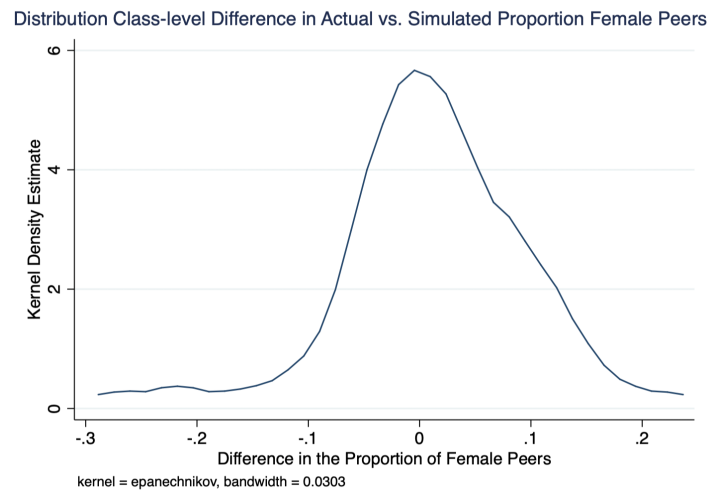


Figure B.6: Correlations of Proportion of Female Peers and Long-term Views on Gender Norms by Student Gender

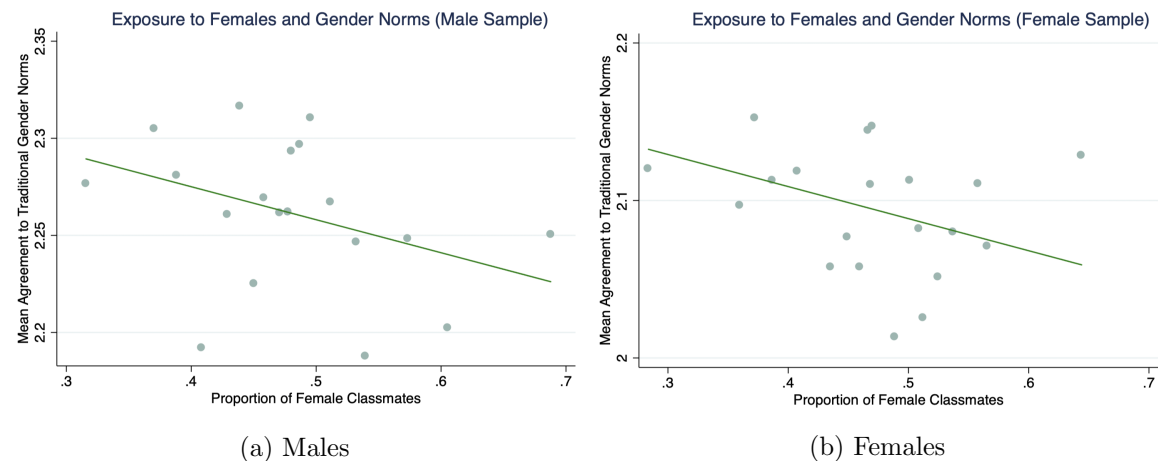
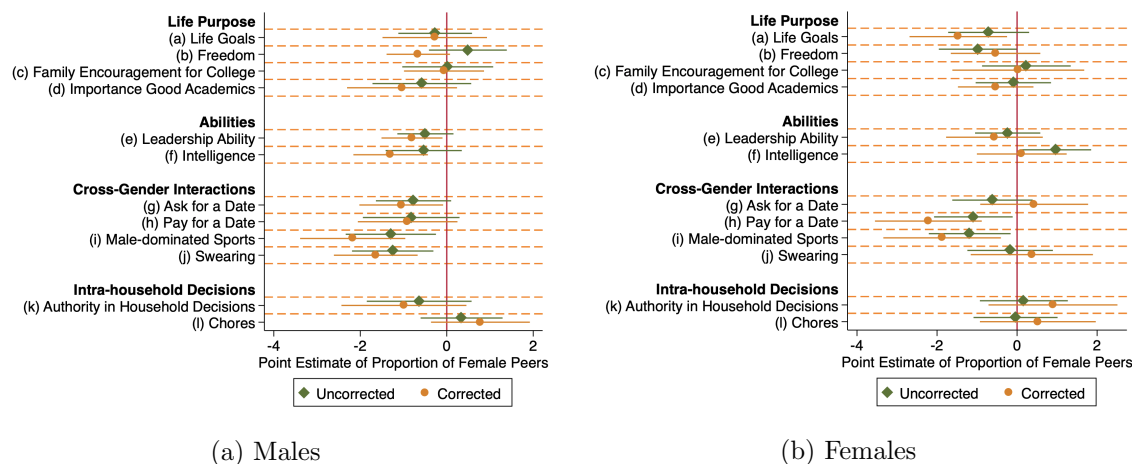


Figure B.7: Breakdown by Gender of the Effects of the Proportion of Female Peers On Views on Gender Norms



## C Extended Analyses (For Online Publication)

### C.1 Representativity of the Sample

As mentioned, the Young Lives survey was designed with a pro-poor approach. Therefore, while it is well-suited to study the dynamics of the children, the final dataset is slightly poorer than nationally representative ones (Young Lives, 2014). In this section I show that: (i) the subsample of children that participate in the school survey is not a selected one within the Young Lives data; (ii) the YL children in the school survey are not different from their non-YL-participating classmates.

For the first point, I use the full YL sample available at the third wave<sup>35</sup> and regress a battery of observable characteristics on an indicator of whether the child was also included in the school survey. Table A.2 shows that the majority of the estimates are small in magnitude and not statistically significant. This suggests that those YL children that were selected for the school survey were indeed a random subsample of the original sample. As expected, those children included in the school survey are slightly more likely to be enrolled in school (4%), since this was a requirement to be eligible to be part of the school survey in the first place.

For the second point, I turn to the school survey and I focus, as in the main analysis, on exogenously formed classes, but I do not impose that the child needs to be part of the YL main survey. Instead, the interest here is on keeping also the classmates and exploring whether an indicator of being a YL child is statistically significant. This would indicate that YL children are systematically different from their peers. Table A.3 shows that this is not the case.<sup>36</sup>

### C.2 Further Exploration of Random Allocation of Children to Classes: Permutation Tests

In order to verify that the average characteristics at the class level are in line with those obtained from randomizing the existing students in a school across the available sections, I proceed as follows.

First, I form synthetic classes of the same size as in the actual data by randomly allocating actual students in a school across sections. I then compute the average characteristics at the class level for a series of important dimensions such as gender, parental education, the wealth index, and non-cognitive skills (e.g. interest in school). I repeat this exercise 1,000

---

<sup>35</sup>The original cross-sectional dimension (i.e. at the first round) of this cohort was 2,000.

<sup>36</sup>See Section D for more details on the variables used.

times. For each class and simulation, I compute the distance between the actual value of each characteristic in the original class configuration and its simulated counterpart. Calculating the average distance between the original and simulated values across the 1,000 repetitions yields absolute values of about 0.0001 for all the observables considered, with very small standard deviations. For example, for gender, the mean distance is -0.0001 and the standard deviation is 0.001. This analysis therefore provides further support to the exogenous allocation of students to classes within schools.<sup>37</sup>

### C.3 Further Discussion on Friendship Formation

As mentioned in the main text, a reasonable counterfactual to benchmark the estimates in Table 4 is the one in which friendships are formed at random. One should note, however, that since male children are significantly more likely to form friendships with other males than with females (among those students classified as friends in column 1 in Table 4, 35% are females and 65% are males), the thought experiment of removing one male peer and substituting it for a female one would mechanically increase the proportion of female peers.<sup>38</sup>

With this in mind, and given that I do not pursue the endeavour of modelling friendship formation, one pathway is to undertake Monte Carlo simulations to compute the average change in the proportion of female friends that would be expected, under the current gender partition of friendship networks, after replacing one male peer by a female one among the actual classroom composition. To be more specific, I proceed as follows. I first compute the proportion of female and male friends in the actual data that satisfy the friendship definition in column 1 of Table 4 — 0.35, as reported in the Table. Second, for every *male* child, I randomly draw a number from a uniform distribution in the 0-1 interval for each of his classmates. If the number allocated is below 0.35 for a female classmate or below 0.65 for a male one, I define that child as a friend of the reference child. Then, for every child, I replace one male classmate by a female classmate. If the random number assigned to that child is below 0.35 I define her as a friend. I then compute the proportion of female friends over the total number of friends. I repeat this exercise 1,000 times, obtaining an average female proportion of total friends of 0.393 — which is  $100 \times (0.393 - 0.35) \approx 4.3$  percentage

---

<sup>37</sup>To check whether these very low mean values of the distance variable are a consequences of averaging out larger differences across classrooms within schools, in Figure B.5 I plot the kernel density estimate of the distribution of the gap between the actual and the 1,000 simulated classroom formations of the proportion of female peers for the first class within each school (such class ordering was randomly selected by the School Survey fieldworkers). The mean of this variable is of similar size as found when averaging across all sections in a school, hence pointing against towards a lack of non-random assignment into classes.

<sup>38</sup>This is because of two forces: (i) there are more females to be friends with, and (ii) the *total* number of friends (which is at the denominator when computing the proportion of female friends among all friends) decreases because males form more friendships with other males than with females.

points higher than the empirical one.<sup>39</sup>

Now, because this figure is obtained by increasing the number of females in one *unit*, I need to compute how much this increase represents in percentage points of female classmates, which is how my independent variable of interest in Table 4 is measured. In a first step I compute the proportion of female peers when, taking the actual data, I substitute one male per class by a female. I find that an increase in one female in the class raises the proportion of female peers in approximately 13.05 percentage points. Since the point estimate in column 1 in Table 4 is 0.657 per every percentage point increase in the proportion of female classmates, the addition of one female, which represents an increase in 13.05 percentage points, is predicted to raise the proportion of female peers in  $13.05 \times 0.657 \approx 8.57$  percentage points. Therefore, under the above assumptions, there is evidence that after the swapping of a male for a female peer not only male children would form friendships with a higher number of female classmates, but also to do so disproportionately more than predicted ( $8.57 \gg 4.3$  percentage points).

Additionally, given that Table 3 does not show heterogeneity in the effects on attitudes by gender, in Table A.9 I proceed similarly to Table 4 for the subsample of female children. Several aspects are worth highlighting. First, the mean of the outcome variables is increasing in the degree of “closeness”. For instance, while 63% of the “at least a little close” friends in column 1 are females, this figure increases to 89% for “very close” in column 3. This, which is the opposite case to Table 4, was expected: it reinforces the idea that children are more likely to establish closer relationships with same-gendered peers. Second, the qualitative results are in line with those for males: more exposure to female peers significantly increases the proportion of “weak” friends and, in the present case, also of close ones (but not of extremely close ones.) Third, the estimated coefficients are larger, as consistent with the fact that females are much more prone to form networks with other females in the first place. Overall, these findings provide further support for increased interactions with females to indeed be a likely channel behind my main effects on attitudes, *both* for males *and* females.

#### C.4 Accounting for Measurement Error in Peers and Different Sample Selection Criteria

In this section I probe the robustness of the results to accounting for measurement error in peers and different samples. In the main analyses I have worked with all students who were not allocated to classes based on ability or age. In Table A.13’s columns 1 and 2 I explore

---

<sup>39</sup>In order to put bounds on the new proportion of 0.393, I report that 5% of the simulations for the proportions of female friends are below 0.387 and 5% are above 0.398.

the effects of female peers on views on gender norms and on friendship with females using exclusively students who were randomly allocated. The qualitative and quantitative results hold.

Additionally, in columns 3-5 I return to using the same subsample as in the main text, and I explore whether the secondary results on cognition and academic effort are unchanged under the inclusion of the interaction between school fixed effects and the proportion of classmates observed. This is not the case. Therefore, this confirms that, while the effects on views on gender norms are large and likely driven by stronger interactions with female peers, other forms of spillovers, particularly academic ones, do not seem to be strongly in place.

## D Variables Description (For Online Publication)

In this appendix I provide further details on the construction of certain variables that were not described in the main text.

- “No health problems”: an indicator taking the value of 1 if the child stated not to have any health problem that affects him or her in school.
- “If I work hard I can go to college”: a four-category (strongly agree; agree; disagree; strongly disagree) variable for the following statement “If I work hard I think I can go to college or university.”
- “No reinforcement class”: an indicator taking the value of 1 if the child stated not to be attending extra classes in any subject.
- “Enjoy school”: a three-category (always; sometimes; rarely/never) variable for the question “Do you enjoy school?”
- “Worry about exams”: a three-category (always; sometimes; rarely/never) variable for the question “Do you worry about exams/tests?”
- “Low school effort” is the principal component of the following five items (each has four categories: strongly agree; agree; disagree; strongly disagree):
  - I pay attention to the teachers during lessons
  - I study hard for my tests
  - I am willing to do my best to pass all the subjects
  - I do not give up easily when I am faced with a difficult question in my schoolwork
  - I am not willing to put in more effort in my school work (I invert it)
- “Positive view on teacher” is the principal component of the following two items (each has four categories: strongly agree; agree; disagree; strongly disagree):
  - My teachers feel that I am poor in my work
  - I get frightened when I am asked a question by the teachers
- “Think does not do better than others” is the principal component of the following two items (each has four categories: strongly agree; agree; disagree; strongly disagree):
  - I often forget what I have learnt



- I always do poorly in tests
- “High interest in school” is the principal component of the following five items (each has four categories: strongly agree; agree; disagree; strongly disagree):
  - I daydream a lot in class
  - I often do my homework without thinking
  - I am usually interested in my schoolwork (I invert it)
  - I often feel like quitting school
  - I am always waiting for the lessons to end
- “Any parent illiterate”: indicator taking the value of 1 if any of the parents were reported to be illiterate, and zero otherwise.
- “Takes private classes”: indicator taking the value of 1 if the child takes extra private (paid) classes and zero otherwise in the last academic year.
- “Feel pressure at school?”: a three-category (always; sometimes; rarely/never) variable for the question “Do you feel under pressure to perform well at school?”
- “Difficult to complete homework?”: a three-category (always; sometimes; rarely/never) variable for the question “Do you find it difficult to complete your homework?”
- “Hit by parents?”: a three-category (always; sometimes; rarely/never) variable for the question “Are you hit by your parents?”
- “Hours extra classes”: number of hours responded to the question “During an average week, how many hours has your child attended extra classes?”
- “Money extra classes”: rupees amount responded to the question “During the school months of the last academic year, approximately how much money has the household paid on average per months for this child’s extra classes?”
- “Ideal education level reached”: categorical level of education answered by the child’s caretaker when asked: “Ideally, what level of formal education would you like your child to complete?”
- “Knows names of child’s friends”: an indicator taking the value of 1 if the caretaker answered yes to the question “Do you know the names of your child’s friends?”

- “Mother is a decision maker”<sup>40</sup>: an indicator taking the value of 1 if the proportion of affirmative responses to whether the child’s mother is the main decision maker (or decisions are taken jointly by the household) is at least 50% across the following dimensions:
  1. Who is the main person responsible for making the key decisions about most of the land?
  2. Who mainly controls the use of the earnings from the sale of goods or rent from most of the land?
  3. Who is the main person that can make decisions about most of the animals?
  4. Who is the main person who controls the use of the earnings from the sale of goods or rent from these animals?
  5. Who is the main person responsible for making the key decisions about the sale of livestock or livestock products?
  6. Who is the main person responsible for controlling the earnings from the sale of livestock or livestock products?
  7. Who is the main person responsible for making the key decisions about work for wages activities?
  8. Who is the main person responsible for controlling the earnings from work for wages activities?
  9. Who is the main person responsible for making the key decisions about business and self-employment activities?
  10. Who is the main person responsible for controlling the earnings from business and self-employment activities?
- “Perceived School Performance:” an indicator taking the value of 1 if the child stated that (s)he thinks that. he is performing worse than their classmates (other options were: “about the same”, and “better”).
- “Teacher Punished Classmate”: indicator taking the value of 1 if the child stated that in the last week that (s)he attended school (s)he never saw a teacher using physical punishment on other students (other options were: “once or twice”, and “more than twice or often”).

---

<sup>40</sup>The mean in my estimating sample is 0.68 with a standard deviation of 0.47.

- “Teacher Punished Self”: indicator taking the value of 1 if the child stated that in the last week that (s)he attended school (s)he was never physically punished by a teacher (other options were: “once or twice”, and “more than twice or often”).
- “Safe in Neighborhood”: indicator taking the value of 1 if the child stated that (s)he agrees or strongly agrees to the statement that (s)he feels safe when he goes out of the house on his/her own.
- “Count on Friends”: indicator taking the value of 1 if the child stated that (s)he agrees or strongly agrees to the statement that his/her friends will stand by him/her during difficult times.
- “Try Hard Can Improve Life”: indicator taking the value of 1 if the child agrees or strongly agrees that if (s)he tries hard (s)he can improve his/her situation in life
- “Study Hard for Better Job”: indicator taking the value of 1 if the child stated that (s)he agrees or strongly agrees to the statement that if (s)he studies hard at school (s)he will be rewarded by a better job in future.
- “Plans Ahead”: indicator taking the value of 1 if the child stated that (s)he agrees or strongly agrees to the statement that (s)he likes to make plans for his/her future studies and work.
- “Included in Others’ Games”: indicator taking the value of 1 if the child stated that other children always include him/her in their games (other options were: “sometimes”, and “never”).
- The set of questions collected in round 3 about the degree of happiness had a common phrasing: “How happy or sad do you feel about...” I create indicator variables taking the value of 1 if the answer was “a bit happy” or “very happy” (other options were: “very sad”, “a bit sad”, and “neither happy nor sad”).
  - The things you have? Like the money you have and the things you own?
  - Your health?
  - The things you want to be good at?
  - Getting on with the people you know?
  - How safe you feel?
  - Doing things away from home?

- What may happen to you later on in your life?