PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

MASTER THESIS Analysis and Policy in Economics

Schooling and Early Gendered Performance Gaps: An Investigation on French Data

---

Author:
Lola TOUITOU

Supervisor:
Thomas BREDA

Year:
2022-2023

Referee:
Gabrielle FACK

Schooling and Early Gendered Performance Gaps: An Investigation on French Data

Lola Touitou
Paris School of Economics
Supervisor: Thomas Breda
September 2023

Abstract

When do gender gaps appear, and what is the role of students' schooling experience in their early evolution? In this master's thesis, I exploit exhaustive national standardized student evaluations from France and show that the maths gender gap appears and widens during first grade. Gaps are larger in privileged schools, and at the top of the distribution. Exploring the effect of different factors on maths progress, I don't find any evidence that gendered individual specialisation effects explain the evolution of the gap. However, girls seem to benefit more than boys in terms of rank progress from more experienced teachers. Furthermore, according to both an OLS and an instrumental variable strategy based on a plausibly exogenous variation in the share of female teachers across schools, being assigned to a male teacher has a small negative impact on girls. Finally, the cohort affected by school closure during the COVID-19 lockdown experienced a comparatively smaller increase in the maths gender gap.

Keywords Gender gaps · Educational achievement · Education · France
JEL classifications I21 · I24 · J16

Acknowledgements

I would like to thank Prof. Thomas Breda for his encouragement, guidance and advice on the topic. I am also thankful with the ministry of Education and DEPP for allowing me to access the data, and availability in answering questions. This work benefitted immensely from useful comments and discussions with Frederic Leon Denker, Simon Drapala, Guilhem Lameiras, Rafael Schütz and Violetta Van Veen, as well as participants in the public and labor economics PSE masters' seminar. All remaining errors are my own.

Contents

List of Figures

List of Tables

1 Introduction

Progress has been made in addressing gender disparities, but women still face under-representation and lower pay in various sectors of the labor market. Identifying the extent and causes of these differences, including gender discrimination, has proven challenging. Scholars have found that inequalities may indeed result from discrimination, but also from distinct choices made by men and women and cumulative effects over time. Education, recognized as a prime determinant of labor market outcomes, plays a crucial role. Analyzing gender gaps in education, including academic performance and career choices, is therefore of prime importance for understanding and addressing gender disparities in society. Examining the gender achievement gaps has revealed a paradox that emerged in the last decades: despite women outperforming men on average in school and having higher rates of high school completion, college enrollment, and completion in most OECD countries, they still experience worse labor outcomes than men.

The existing research has highlighted multiple factors that influence educational and occupational choices at different stages of the pipeline and can explain differences in outcomes. For instance, the underrepresentation of women in STEM[1] majors in college has been identified as an important factor for the gender pay gap. Research by Card and Payne (2021) suggests that college majors' choice still accounts for 20% of the pay gender gap of recent US college graduate. In turn, studies on occupational choices show that women tend to be interested and believe more in their ability to succeed in jobs where there are more women (Ji et al., 2004). Further upstream, Card and Payne (2021) demonstrate that in Canada and in the US, although women are more likely than men to be able[2] to enroll in STEM-related courses before the last year of high school, fewer women actually enroll in enough courses to be STEM-ready. This discrepancy in high school course selection can actually explain most of the observed difference in majors' choice in college.

In a cumulative process, inequality and difference in academic performances very early-on may potentially have lasting impacts on women's career choice and labor outcomes. The beginning of mandatory schooling is a critical period for internalising gender stereotypes (Duru-Bellat 2008, 2016), therefore of utmost importance for researchers and policymakers. Understanding the factors that influence early educational performance and gender gaps could have a profound impact on gender disparities in both education and the labor market, ultimately leading to a more equitable and inclusive society.

Using several cohorts of comprehensive national evaluations of first grade students from France (*Evaluations Repères)*, this thesis first aims at providing a clear picture of gender gaps both in French and maths. Taking stock of the findings of a note produced by the scientific council of education (CSEN, 21.03 *Qu'apprend-on des évaluations de CP-CE1?*) on the first cohort surveyed, I first confirm on my data a number of usual findings on gender performance

[1] Science, Technology, Engineering, and Mathematics

[2] In terms of previous results.

in French and maths. As in most of the related literature in developed countries (see OECD reports on PISA data), French male students are on average better in maths (except at the beginning of first grade), while female outperform them in French. Distributions for boys are also more spread out, in both subjects, meaning there are more extreme performers at both ends of the distribution among boys than among girls, whose results are more centered aroung the mean. Consequently, in terms of timing as Fischer and Thierry (2022) I show that a gender gap in French can be observed as early as the beginning of first grade, with girls outperforming boys. For maths however, I find that the gender gap in favor of boys is non-existent at the beginning of first grade, and it appears afterward. Similarly to other longitudinal studies (Fryer and Levitt, 2010), although spanning a shorter amount of time I observe the maths gender gap widening in favor of boys during first grade.

More recent works (Penner and Paret, 2008; Cobb-Clark and Moschion, 2017) have also shown that those gaps can be highly heterogeneous, and depend for instance on family socio-economic status or race. The heterogeneity analysis I conduct to see where the gender gap is biggest in the performance distribution echoes results in Cobb-Clark and Moschion. Using a sample of Australian primary school students, they find that the unconditional raw gender gap is highest at the top of the distribution, and that patterns differ when considering students of different socio-economic background separately. I indeed find that girls are the most underrepresented at the top of the distribution, accounting for only 30% of students scoring int the top maths percentile at the beginning of second grade. I also show that in my sample, gender gaps are lower in schools with students coming from more disadvantaged backgrounds (public REP and REP+ schools, and schools with low IPS), but patterns in the distribution seem to be the same.

Although gender gaps in performance are well documented, factors explaining their level and evolution do not foster consensus, especially as to their relative importance. The reversible nature of gender gaps over time and the variability of their magnitude from one country to another (OECD reports on PISA or TIMMS data) nevertheless make it possible to rule out biological arguments, while highlighting the role of gender (in)equality and cultural determinants. For example, Scandinavian countries display a lower gender gap than does France in the latest PISA survey in maths. Furthermore, the gap seems to decrease over the years, which is not the case in France. Using measures such as the World Economic Forum's Gender Gap Index, Guiso et al (2008) show that the gender gap in maths of 15 years-old students disappear in more gender equal society, and is even more in favour of girls in reading[3].

To understand whether gendered specialization processes can account for the significant increase in the maths gender gap, I begin my analysis by looking at the influence of initial

[3]Fryer and Levitt (2010) show that it is only true for PISA countries, and not for TIMMS countries that also include middle-eastern countries that are highly unequal but show very low level of gender gap in maths. They however also note that those are countries for which the majority of schools have sex-segregated classrooms, and are Muslim countries.

scores in both subjects on maths progress. Building upon the research conducted by Breda and Napp (2019), which demonstrates that girls' individual comparative advantage in reading over maths influences their educational choice, I show that initial scores affect boys and girls' maths progress differently with initial French and maths scores being respectively more positively and more negatively associated with maths progress for girls. However, the positive effect of the initial French scores on progress for girls is inconsistent with the hypothesis that girls progress less in maths due to a French specialization process explained by individual comparative advantage.

Drawing from gender socialization theories, subfields of the literature have looked into the influence of the family, peers, and institutions such as schools and religion on gender gaps. In particular, schools and other educational institutions are critical environments in which gender biases can either be reinforced or decreased, and in which students interact with their peers and teachers. The second part of my work can be framed in the literature related to gender and students' schooling experience.

Autor et al (2016) look at the effect of the quality of schools depending on gender, and show that boys actually benefit more than girls from schools of better quality[4]. The evidence on the effect of the gender composition of students' environment is more mixed. Considering single-sex or coeducational environments, single-sex colleges have been found to reduce gender stereotyping for women (Dasgupta and Asgari, 2004). Paredes (2022) shows that single-sex classes in coeducational high schools in Chile reduces the maths gender gap by half. However, in Denmark, Brenoe et Zölitz (2020) argue that having a larger proportion of female peers in high school actually reduces women's probability of enrolling in and graduating from STEM programs[5]. I look at the influence of the gender composition of peers and teachers and other characteristics at the school level on girls' maths progress using school and cohort fixed effects. While the gender composition (students or teachers) at school does not appear to have a significantly different effect between genders, my analysis reveals that certain school characteristics do affect girls differently. Notably, the average age and seniority of teachers within the school benefit girls' maths progress more than boys'.

Using a classroom name text field in the data which allows me to identify teacher's gender, I also contribute to the literature studying the effect of the gender of teachers. At the college level, using random assignment to mandatory maths and science courses professors, Carrell, Page and West (2009) show that being assigned to a female professor has a significant impact on women's maths and science performances and likelihood to graduate with a STEM degree. In middle school, in the US, Dee (2005, 2007) shows that assignment to a same-gender teacher improves performance for both girls and boys and that it also improves the teacher's perception of the student's performance. To establish the effect of teachers' gender on students' maths progress, I first estimate an OLS specification with cohort and school fixed effect, the treatment being the interaction of the variables female student and male teacher. I find a small negative

[4]Note that they consider outcomes for which boys have a disadvantage: suspensions, and an aggregate maths and reading performance score. School quality is measured as school value added.

[5]One of the proposed mechanism is that women's orientation choices are more influenced by same sex peers in a high school with a high share of women. Interestingly for women with a mother working in STEM, the gender composition of the school does not have an effect.

effect (-0.03 SD). To account for potential confounding factors in the allocation of students to teachers depending on gender, I then instrument teachers' gender by the proportion of female teachers within the school. The results of the IV specifications are of the same order of magnitude as the OLS, providing evidence that French girls progress less in maths compared to boys when assigned to a male teacher. In any case, back-of-the-enveloppe calculations reveal that the effect of having only female teachers on the maths gender gap would be small (-2% at most).

Finally, comparing students with varying length of schooling and controlling for age, the CSEN note shows that the maths gender gap is increasing in schooling, and argues that its evolution is linked to schooling. I propose to test this general hypothesis in the last part of my work, in which I exploit the exogeneity of the COVID-19 lockdown. To compare cohorts, I estimate a triple diff-in-diff specification, the coefficient of interest identifying the difference in evolution of the gender gap in maths between the treated, in lockdown during first grade, and control cohorts. However, identification is not without challenges, and the effect I find is small.

This thesis is organised as follow. In a first section, I present some background information on the main results of other standardized evaluations on French students as well as the *Evaluations Repères*. In a second section, I introduce the data and sample of students and schools. In a third section, I examine the main descriptive results relating to performance and gender, and explore the heterogeneity of gender gaps. The fourth section explores potential mechanisms explaining the evolution of the maths gender gap, from individuals specialization to different schooling aspects. The fifth section discusses the results. I finally provide concluding remarks.

2 Standardized evaluations in France and gender gaps

2.1 International standardized evaluations

French students participate to international standardized evaluations in maths and litteracy, the goal being to compare performances across countries. As gender is a specific focus in most of them, it also allows to compare countries in terms of achievement gaps.

TIMSS assesses the maths and science knowledge of fourth grade (CM1) students (TIMSS 4) or eighth grade (*Quatrième*) students (TIMSS 8) every four years. French students have only been part of the evaluations for at most the last two waves (TIMMS 4), although these evaluations exist since 1995. The last results (2019) show that both for fourth grade and eigth grade France is among the few countries for which there is a significant gap in maths in favour of boys with respectively 13 and 9 points of difference.

PISA evaluates both maths, science and reading skills of 15 year-old students every three years since 2000. The last results available (2018) show that in reading French girls significantly outperform boys by 25 points – while the OECD average is 30 points, but the gap is decreasing.

In maths boys are significantly ahead (7 points, respectively 499 for boys and 492 for girls), a gap that is higher than the OECD average (2.8 points).

PIRLS focuses on the reading comprehension of students in fourth grade (CM1) in 50 countries and is conducted every five years since 2001. the latest available results (2021) show that French girls also have an advantage in reading achievement compared to boys (521 and 507 points respectively).

2.2 National exhaustive standardized evaluations and *Evaluations repères CP-CE1*

Since 2017[6], the DEPP (*Direction de l'Evaluation, de la Prospective et de la Performance*), the DGESCO (*Direction Générale de l'Enseignement SCOlaire*) of the ministry of education and the scientific council of education (CSEN) design and conduct national exhaustive standardized evaluations to assess litteracy and numeracy skills of French students at different schooling points between age 3 and 16 (compulsory education years). A first round of evaluations takes place during first grade (CP) and at the beginning of second grade (CE1). A second one is organised during sixth grade (*Sixième*), and the last one at the beginning of tenth grade (*Seconde*). Unlike international assessments, those evaluations, which are exhaustive both for public and private schools that are overseen by the State, take place every year and are meant to evaluate students' performances on a set of nationally-defined skills. For all three evaluations, the DEPP publishes notes on students' performances and provides statistics broken down by gender.

The *Evaluations Repères CP-CE1* take place in three distinct sessions in a little over a year: first during the last-half of September at the beginning of first grade, then *point d'étape* at the middle of first grade usually during the second half of January, and finally in the second half of September at the beginning of second grade. While teachers are given some flexibility on the exact date at which they conduct the evaluation, they are encouraged to split each session in two so as not to overwhelm students. Each session is composed of series of exercises to test numeracy and literacy skills alternatively, to be completed in a limited amount of time. Following an official protocole, teachers conduct the evaluations by giving the instructions to students in the classroom, or individually for certain exercises (reading). Students are provided exercise notebooks on which they fill out answers. Teachers then correct answers and report the results using a dedicated software.

Exercises (*item*) do not cover every skill in the official curriculum, but are created to identify precisely sources of difficulty for students. Their nature, modality and questions vary over the years and evaluation sessions, so direct comparison between cohorts or over time is not

[6]The first one to be launched was the sixth grade evaluations in 2017. Then came *Evaluations repères CP-CE1*, starting 2018. The last one, the tenth grade evaluation, was first conducted in 2020.

straightforward.[7]. For each item, performance thresholds are defined by the DEPP depending on the item difficulty. The first threshold identifies students that have severe difficulties (*en difficulté*). A second one points out those for which results are fragile (*acquis fragiles*).

Results are shared with school principals, teachers and students' families after each session of evaluation. For each student, diagrams with performance for each skill evaluated in numeracy or literacy are created.

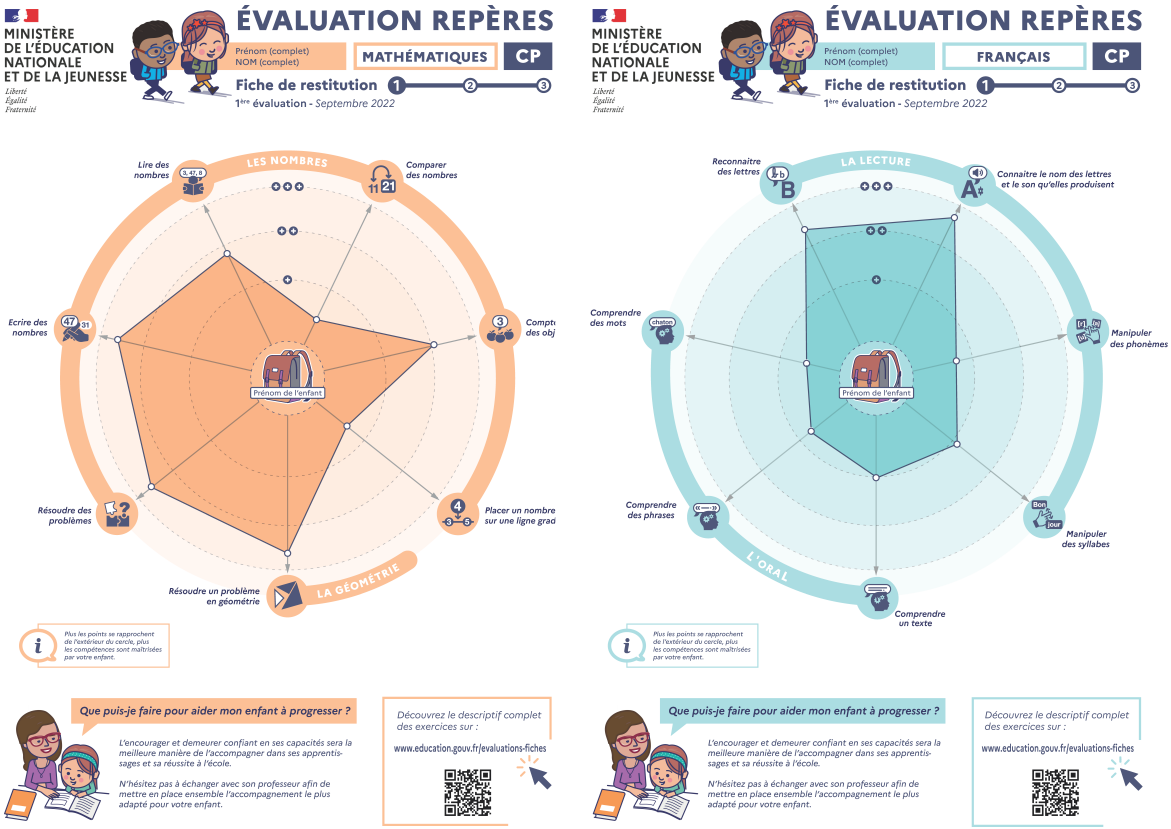


Figure 1: Students' performance diagrams

3 Data

I use two databases from the DEPP. For students' outcomes, I use the results of the *Evaluations Repères CP-CE1*. I complement the results by information on students' environment and teachers at the school level using the APAE 1D database.

[7]For example the item *Comprendre des mots à l'oral* is only evaluated during the first session of the evaluation, while the item *lire des mots inventés* was only assessed for the first cohort.

3.1 Result outcomes (*Evaluations Repères CP-CE1*)

The data contains item results at the individual level. I restrict the sample to observations that are identified by crypted unique student number (*Identifiant National Elève*), sex, birthday, and school identifier.

It should be noted that students can't be linked to their class and that the results data is incomplete, with many students missing item results, or evaluations[8]. I compute aggregate averages (details below), provided I have at least around half the items for a given subject, and it could be a problem if missing items scores were related to gender, as some type of items may advantage one gender more than the other. I check for several items that there does not seem to be more missing values for one gender over the other. More results available upon request.

3.1.1 Item results

Items evaluated and grading scales are not consistent over cohorts and evaluation sessions. I therefore standardize item results for each cohort and evaluation session.

For item j and student i, the result is transformed in the following way:

$$\tilde{item}_{ij} = \frac{item_{ij} - \overline{item_j}}{sd(item_j)}$$

This allows to give a similar weight to every item in the following transformations. Not doing so would implicitly result in considering items that are graded on a bigger scale as more important than others. Note that for the *Evaluations repères CP-CE1* scales can differ quite a lot depending on the item. Reading items for instance can be graded out of 100 while writing words items are graded out of 12.

3.1.2 Aggregate scores

I then construct an aggregate measure of performance for maths and French separately. I average item results previously standardized by subject and compute Z-scores. For every cohort and test session, student i score is:

$$score_i = \frac{x_i - \bar{x}}{sd(x)} \qquad \text{with } x_i = \frac{\sum_{j=1}^{n} \tilde{item}_{ij}}{n}$$

By definition, the created scores have a mean of 0 and standard deviation of 1.

[8] Between 2 and 5% depending on the evaluation in the obtained sample

3.1.3 Progress outcomes

To investigate the evolution of gender gaps, I look at the difference in rank for a given student between two evaluations, that I name "progress" for simplicity[9] and define for student i and evaluation at time t (beginning of first grade, middle of first grade and beginning of second grade) as

$$prog_i^t = score_i^t - score_i^{t-1}$$

I present the results of my main specification between the beginning of first grade and the beginning of second grade. Results using the progress in maths between the beginning and the middle of first grade can also be found in the Appendix.[10] Results using other progress outcomes are available upon request.

3.2 School environment controls (APAE 1D)

3.2.1 Type, enrollment and classes

For the schools I know whether it is a **private** (overseen by the State) **or public school**. For the latter, I also know whether they have **priority status** and are REP or REP+ or whether they are not concerned. I have data on **enrollment numbers** with both total enrollment and enrollment broken down by grades (including first grade). I also have the **number of classes** and type (multigrade or not) for every grade. I therefore construct a dummy variable to distinguish first grade students who are in a **multigrade class or not** with certainty (respectively 1 or 0) or possibly (2) when there are multigrade and regular first grade classes in a school, for a given school year. While it is not possible to know class size, I also have the **average number of students per class** in the school. I know the **proportion of female students** in the school and I have the **proportion of first grade students that are older or younger** than the norm. Filtering by the name of schools, I also create factor variables for schools that use **alternative pedagogies**[11] or are **religious private schools**[12] to see whether the size and dynamic of gender gaps are similar. See the Appendix (Data - Case studies) for more details.

[9] Note that the name of the outcome can actually be misleading as it does not give any indication on students' real progress, but rather gives an indication of the evolution of their rank within their cohort. For instance when considering a student that has negative progress, it could be that although he has improved between sessions in maths, others have on average improved more.

[10] Progress between the beginning and middle evaluations can actually be seen as a robustness check as it does not include the Summer holidays.

[11] I look for mentions of "Freinet", or "Montessori".

[12] For Catholic schools I look for names containing "Saint(e)" or "Notre dame", for Jewish and Muslim schools I research schools online.

3.2.2 Socio-demographic characteristics of students

Concerning students' family background, I have the **IPS of the school**[13], and further I have the **proportion of students** whose head of household is classified within each of the **four main socio-professional categories**. [14] Finally, I have the **median income in the school neighbourhood or town**. As there are more missing values for socio-economic values and school neighbourhood median income, I mainly use the IPS in the analysis part.

3.2.3 Teachers' characteristics

First I have the total **number of teachers** working in the school and **full time equivalent number of teachers**. Combining it with enrollment data I can compute a **teacher per pupil ratio** at the school level. Second I have the **proportion of female teachers**, and **average age and seniority** within the school. More granular information on seniority and age is also available, for instance the proportion of teachers over 50 years old or between 5 and 8 years of seniority.

3.3 Sample

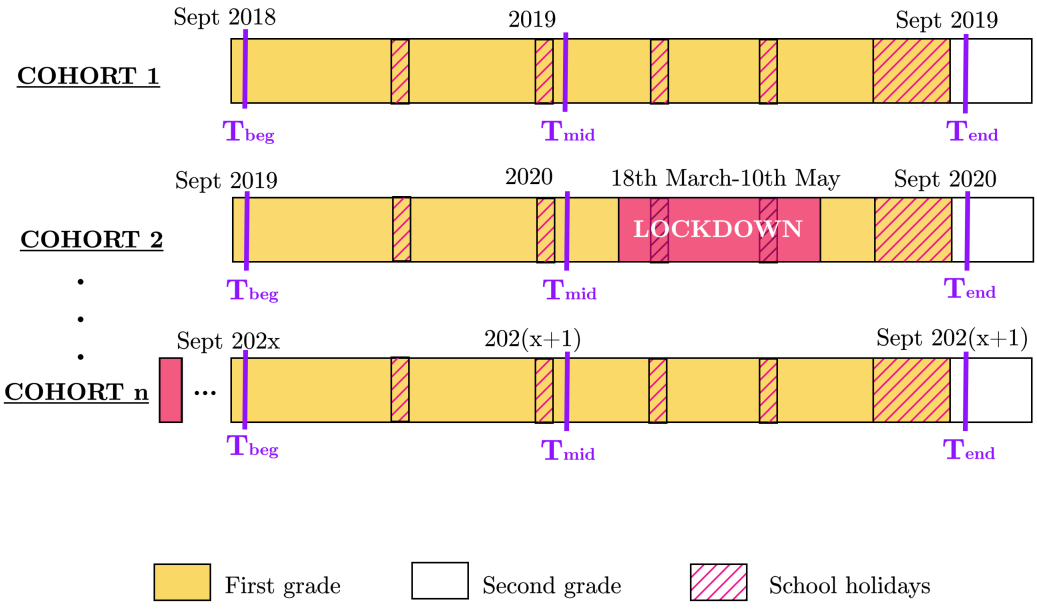


Figure 2: Timeline of evaluations for cohorts

[13] *Indice de Position Sociale*, index created by the DEPP for each student depending on parents' combination of jobs, and aggregated at the school level.

[14] More accurately the data both for IPS and CSP is actually a proxy as it relates to fifth grade students (CM2) that are surveyed in sixth grade (*Sixième*), the year after they have left primary school.

I have data for four cohorts who have completed the three sessions of evaluations. Those are the students registered in public or private (overseen by the State) schools, that were in first grade in 2018-2019, 2019-2020, 2020-2021 and 2021-2022. Note that the second cohort is affected by COVID-19 during first grade, with the general lockdown taking place between the middle and end session of evaluations (middle of first grade and beginning of second grade).

3.3.1 Restrictions

In order to build a more relevant sample for analysis, I restrict it to students that have the most common educational path and are evaluated for the relevant sessions when considering progress outcomes. In particular, I discard observations of students that are significantly younger (less than 5 years old) or older (more than 9 years old) than their peers or repeat a grade. See Appendix (Data - Processing) for more details.

3.3.2 Sample statistics

Students

The main student sample obtained is made up of a little above 2.8 million students for the four cohorts[15] (see Table 10). Girls represent 49% of the students present. 87.5% of students are enrolled in public schools, including 6.5% in REP+ and 10.3% in REP. On average, students are enrolled in schools that also have 49% of female students and have 86% of female teachers. They are registered in schools that have 2.3 first grade classes, and 22.6 students per class. Teachers are 42 years old and have 6.4 years of seniority within the schools on average. The teacher to pupil ratio is 0.05.

I check whether the characteristics of first grade students are balanced between both genders (Table 11). Although significant, differences are small in magnitude. Girls nonetheless appear slightly younger than boys (-0.1 month), are enrolled in schools with a little more first graders (+0.08 student), and in less privileged schools in terms of IPS (-0.09).

Schools

86% of schools in my sample are public, including 4.1% of REP+ and 7.4% of REP. Note that the average school enrollment is lower than the one of the school the average student is enrolled in (157.2 vs 222.1 students), but the proportion of female students (49%) and teachers (87%), average age (41.8 years) and seniority within school (6.5 years) of teachers are similar, as is the ratio of teacher per pupil (0.05). For each school, I observe on average 23 students, among which 49% are girls. See Table 15 in Data Appendix.

[15]Note that observations with missing values of control variables (such as IPS) or outcome variables are dropped when needed in regressions. Sample sizes therefore vary across specifications because of data completeness.

Looking at school characteristics depending on the type of school (public or private and priority status in Table 16), the biggest schools in terms of enrollment are public and REP+ (232 students), followed by public and REP and private schools. The most privileged schools in terms of IPS are private. Classes are the biggest in private schools (23.8 students) compared to public schools without priority status (22.2 students) or with priority (17 for REP, 16.7 for REP+). Consequently the ratio of teacher per students is also highest in REP and REP+ compared to private schools and non-REP public schools. Teachers are oldest and have the most seniority within school in private schools (43.3 years old; 7.5 years of seniority), compared to public schools with priority status (39.3 years old; 5.3 years in REP/REP+). While it is quite common for private and public schools without priority status to have multigrade first grade (65% of private schools and 71% of public schools have at least one multigrade first grade), it is much less common in REP and REP+ schools (respectively 24% and 18%). On average there is also less female teachers in REP and REP+ schools (83 and 80%) compared to private (92%) and public schools without priority status (87%).

4 Gender and performance in first grade

4.1 Gender distributions

4.1.1 Maths and French

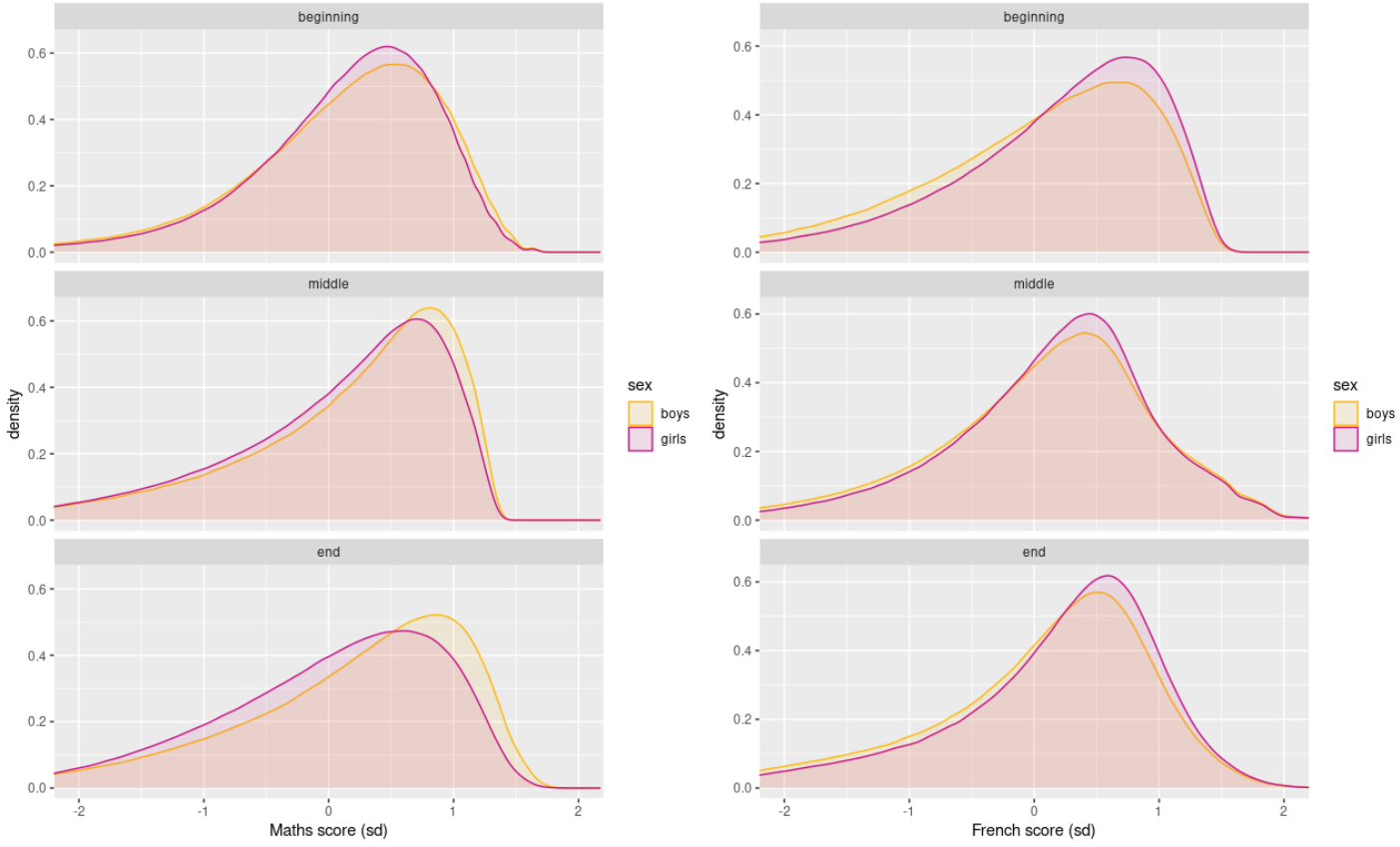
The distribution of first-grade students' results in French and maths by gender reveals several consistent patterns commonly found in standardized test score analysis in developed countries. Boys tend to perform better than girls in maths, while girls perform better than boys in French. Additionally, the distribution of scores is less concentrated for boys compared to girls, indicating greater variance in boys' performance.

	Maths						French					
	Beginning		Middle		End		Beginning		Middle		End	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
mean	0.07	0.09	0.12	0.05	0.14	-0.02	-0.01	0.16	0.05	0.12	0.01	0.14
min	-8.07	-7.99	-4.94	-4.91	-4.44	-4.57	-5.34	-5.34	-5.71	-5.53	-6.02	-5.61
max	1.95	2.18	1.68	1.55	1.78	1.85	1.59	1.61	3.51	3.74	3.36	3.16
Q10	-1.02	-0.90	-1.19	-1.23	-1.23	-1.32	-1.34	-1.03	-1.17	-0.99	-1.34	-1.11
Q25	-0.34	-0.28	-0.34	-0.42	-0.39	-0.57	-0.54	-0.29	-0.43	-0.32	-0.46	-0.29
Q50	0.25	0.25	0.37	0.27	0.36	0.14	0.18	0.35	0.19	0.24	0.22	0.33
Q75	0.69	0.66	0.80	0.72	0.88	0.69	0.71	0.81	0.66	0.67	0.67	0.75
Q90	1.00	0.95	1.05	0.99	1.18	1.03	1.04	1.10	1.11	1.08	1.02	1.07
SD	0.92	0.84	0.92	0.89	0.97	0.92	0.96	0.88	0.93	0.85	0.96	0.90
Missing (%)	1.18	1.16	4.23	4.14	0.36	0.33	0.32	0.31	4.21	4.11	0.36	0.33

Table 1: Descriptive statistics: Performance by gender
Note: Statistics are expressed in proportion of a standard deviation

In French, girls consistently have a higher average score than boys throughout first grade. Conversely, in maths, boys have a higher average score than girls, except for the first test session where girls outperform boys by a small margin. Those differences are statistically significant when performing an equality of means test. Over time, boys' average maths scores and position in the distribution consistently increase (0.07, then 0.12 and finally 0.14 SD), whereas girls' position in the distribution decreases (0.09, then 0.05 and finally -0.02 SD). In French, the average scores fluctuate more for both genders, gaining ground for boys at the second test session and then falling behind, while the pattern is the opposite for girls.

The dispersion of the distributions further supports the observation that boys' performance in both subjects is more spread out compared to girls, as evidenced by their higher standard deviation. Additionally, boys' distributions have more extreme values, as indicated by higher minimum and maximum scores, standard deviation, and more extreme first and last deciles.



(a) Maths

(b) French

Figure 3: Score density by gender
Note: Scores are expressed in proportion of a standard deviation

Plotting the distribution of results for girls and boys separately for each subject and test session (Figure 3), a few additional differences can be noted. In French, the distribution of boys' results is slightly more skewed to the left compared to girls', indicating that a higher

proportion of boys are poor performers, especially during the first test session. For maths, the distributions are quite similar at the beginning, with the girls' distribution being more centered around the mean. However, over time, the curves visibly shift away from one another, with boys outperforming girls. Interestingly, it appears that both for girls and boys, the distribution of results in maths becomes more spread out with time, indicating greater variance in performance. On the other hand, in French, the distribution tightens, suggesting that performance differences between students in this subject become less pronounced over time.

4.1.2 Maths *vs* French

The analysis of individual test scores reveals a strong positive correlation between performance in maths and French for both boys and girls. The correlation coefficient ranges between 0.7 and 0.75 for both genders across different test sessions (Table 22 and 23).

To further explore the distribution of performance in each subject, I compare the percentage of students who perform better in maths than in French and *vice versa* (Table 19). Over time, a higher proportion of boys are found to be better at maths than at French (in terms of position in the distribution), with the percentage increasing from 53.4% at the beginning to 58.9% at the end of the evaluation period. A similar trend is observed for girls, with the majority performing better in French than in maths at the beginning and even more at the end of the evaluation period (57% and 59.3%, respectively). Interestingly, there is a slight rebound in the middle evaluation, where only 52.6% of girls perform better in French compared to maths.

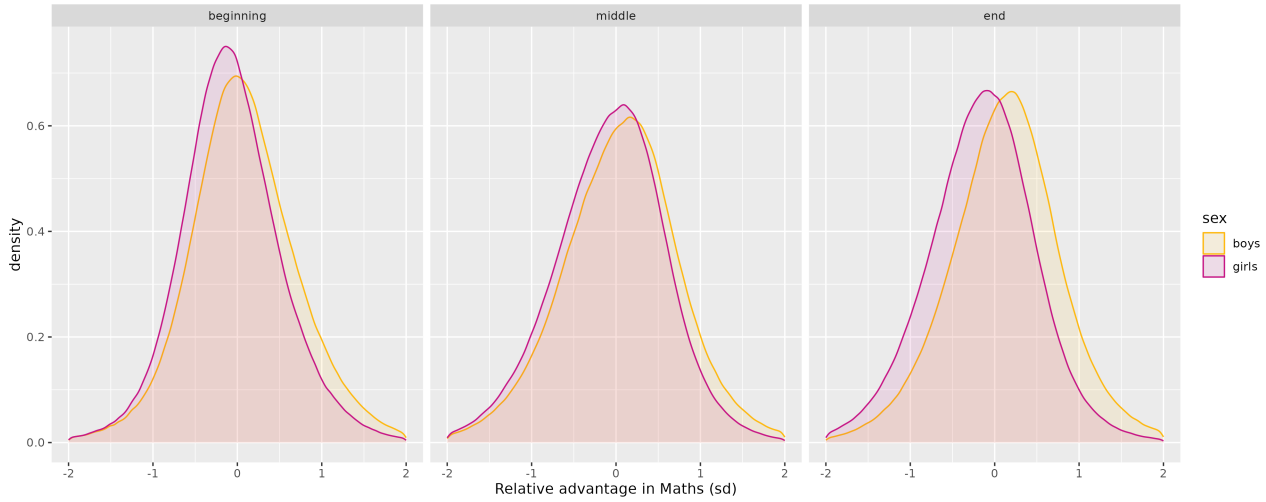


Figure 4: Relative advantage in Maths by gender
Note: the relative advantage is the difference between Maths and French scores for a pupil at a given session. It is expressed in proportion of a standard deviation

Investigating the magnitude of the individual difference in performance in both subject, I examine the distribution of the relative advantage in maths (i.e., the difference between maths and French test scores) for both boys and girls (Figure 4). Boys start first grade with a density curve centered around 0, indicating similar performance in maths and French on

average. However, over time, the density curve shifts to the right, indicating an increasing relative advantage in maths for boys. On the other hand, the density curve for girls shifts to the left, indicating a relative disadvantage in maths compared to French. By the end of the evaluation period, the two density curves are distinct, with girls experiencing a relative disadvantage in maths compared to French, and boys having a relative advantage in maths.

4.2 Gender gaps

In the subsequent sections of this thesis, gender gaps will be computed as the difference between the average scores of boys and girls. A negative gender gap indicates that girls have an advantage over boys and *vice versa*. Similarly, an increasing gender gap implies that boys are gaining ground relative to girls, and a decreasing gender gap indicates that girls are making progress relative to boys.

4.2.1 General results

The analysis of aggregate score performances for the four cohorts combined reveals the presence of gender gaps in French from the beginning of first grade. On average, girls score 0.18 SD higher than boys in French. However, this advantage reduces to 0.07 standard deviation at the second test session. Surprisingly, the gender gap in French widens again between the second and third test sessions, reaching 0.13 SD. In contrast, gender gaps in maths are almost non-existent at the aggregate level at the beginning of first grade or even in favor of girls (0.02 SD). However, the gender gap reverses in the following test session, favoring boys with a gap of 0.07 SD. This gap continues to widen in favor of boys, reaching 0.16 SD at the beginning of second grade.

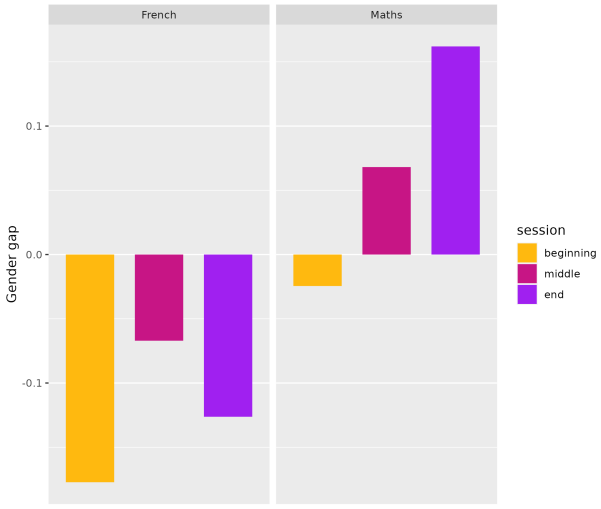


Figure 5: Gender gaps (pooled cohorts)
Note: Gender gaps are expressed in proportion of a standard deviation, and computed as $\overline{score}(boys)^t - \overline{score}(girls)^t$

Considering the evolution of gender gaps in both maths and French, there are distinct dynamics observed. In maths, the gender gap consistently increases over time, indicating that boys are gaining ground relative to girls. On the other hand, in French, the pattern is more nuanced and confirms a finding already mentioned in the CSEN note. Boys seem to catch up during the school year, narrowing the gender gap, before falling behind again at the beginning of second grade. These dynamics hold true when looking at cohorts individually, considering the month of birth, type of schools, and school IPS (Figures 10, 11, 12, and 13).

However, the content of the tests also varies widely between the sessions. For instance, reading items in French are introduced during the second session and make up to half the content of the test. This feature of the evaluations therefore makes it necessary to see whether the content of the test can explain the evolution of gender gaps. See the next section for an analysis of the robustness of gender gaps in maths and French depending on the items considered.

4.2.2 Robustness check

To ensure that the observed pattern in gender gaps is not merely a result of data artifacts, I look at gender gaps for items that are present in multiple evaluation sessions. By focusing on these items, any changes in the gender gaps are less likely to be confounded by variations in the skills evaluated, which could potentially favor one gender over the other and not be related to the evolution of students' ranking in the distribution. It may however be at the expense of precision in evaluating the general level of students in each subject.

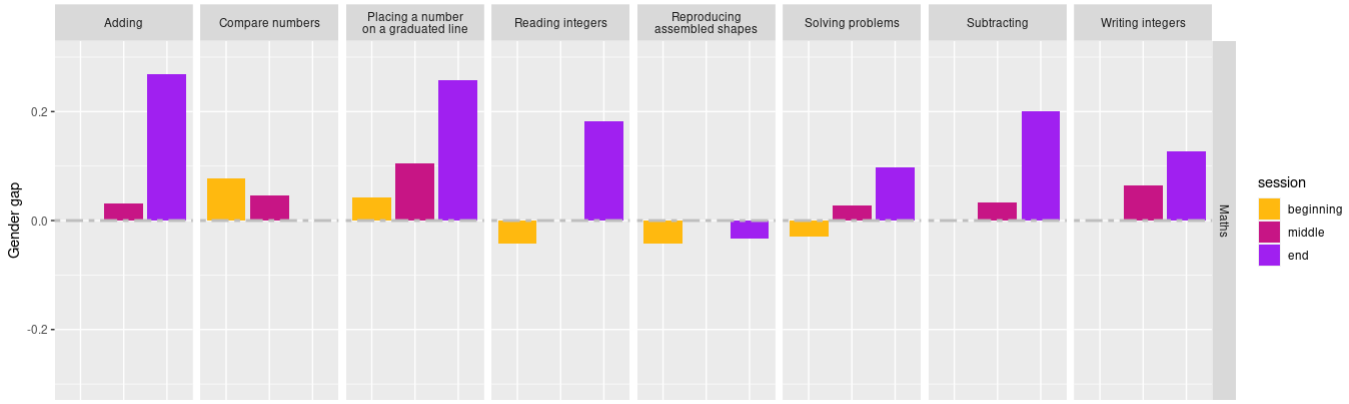
It is striking that for almost every returning Maths items[16], the evolution of the gender gap is exactly the same as for the aggregate Maths score, *ie* it increases over time. Gender gaps in favour of boys seem to increase the most in the *Adding* and *Subtracting*, and *Placing a number on a graduated line* or *Reading integers* items, with the gender gap reaching around 0.2 SD in the last evaluation. It increases much less for the item *Solving problems*. Interestingly, for the *Reproducing assembled shapes* item, first grade girls retain an advantage[17]. Considering the progress made on average by each gender on these returning items (Figure 16) between sessions, it appears that in Maths, the gender gap increases most during the middle and end exam sessions (bottom right panel). However, one also has to take into account that about eight months separate the two, while there are only four months between the first and second sessions.

In French, although the only item that is present in the three evaluation sessions (*Understanding spoken sentences*) and others display gender gaps that are consistent with the aggregate pattern, it appears less clear compared to maths when considering the magnitude of

[16]The exception is the *Compare numbers* item that is present at the beginning and middle evaluation, with boys' advantage decreasing in the second evaluation.

[17]Studying French middle school girls, Huguet and Regnet (2007) have shown, using a similar type of task, that girls that were told the task was a geometry exercise performed worse than those who had been told it was a drawing, providing a prime example of stereotype threat.

the increase. The content of the evaluations changes more in French than in maths (see Table 8 and 9). For instance, reading items are introduced during the second evaluation and together make up to half the content of the test in French. Excluding them and using a restricted French score leads to a doubling of the computed gender gap in favour of girls during the second session, although there is still a small recovery for boys (from 0.17 to 0.14 SD) compared to the first test (Figure 18).



(a) Maths

(b) French

Figure 6: Gender gaps for returning items
Note: Gender gaps are expressed in proportion of a standard deviation, and computed as $\overline{score}(boys)^t - \overline{score}(girls)^t$

The dynamics of the French gender gap does not appear as robust and the magnitude of the phenomenon is not as striking as for maths. Additionally, because maths is said to act as a "filter" (Sells, 1980; Watt et al 2017) for STEM careers, the literature interested in understanding the origin of gender gaps later-on traditionally emphasizes the maths gender gap more. In the French education system, maths remain particularly discriminatory and determine access to higher education and most prestigious schools. I therefore choose to focus most on the maths gender gap in the rest of this master's thesis.

4.2.3 Heterogeneity of the maths gender gap

Score distribution

The analysis of the whole distribution (including both girls and boys) reveals interesting patterns in the under and over-representation of girls across different maths score percentiles (Figure 7). At the beginning of first grade they are underrepresented at the bottom of the distribution and at the top (below the 25^{th} percentile and above the 80^{th} percentile) in equal measure (around 40%). However, by the beginning of second grade, girls become overrepresented at the bottom of the distribution (below the median) and, at the same time, even more underrepresented at the top of the distribution. The most striking change occurs at the top end, where the proportion of girls in the top-performing percentile drops from 40% at the beginning of first grade to only 30% by the beginning of second grade.

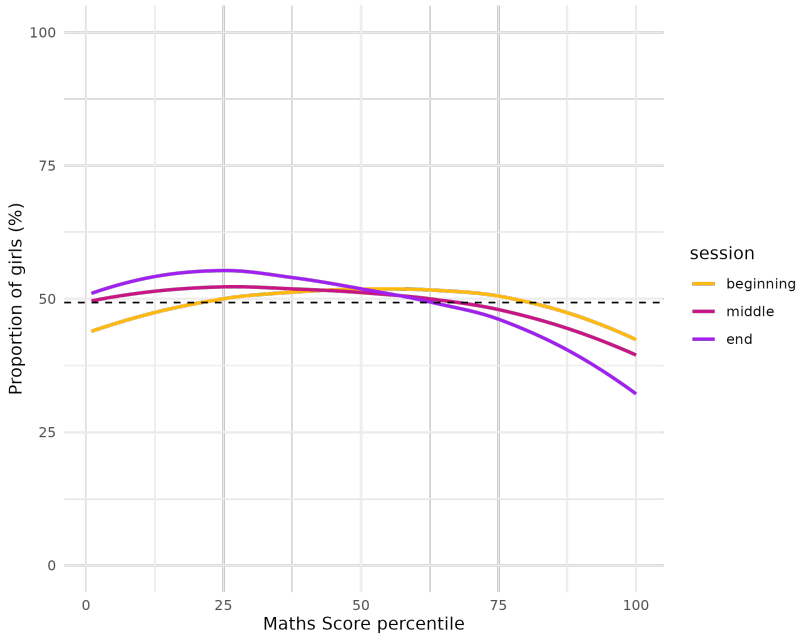


Figure 7: Proportion of girls in each maths percentile

Notes: The dotted line represents the proportion of girls in the sample considered. All curves are smoothed using LOESS.

School type

The maths gender gap displays the same evolution over time, and same heterogeneity pattern in every type of school (public, REP and REP+ and private, see Figure 12 and 22), and whatever the school IPS decile (Figure 13 and 21). What can however be noticed is that although the increase in the maths gender gap looks similar in magnitude across the different types of schools, in levels the gender gap is highest in private and public schools without priority status (0.16 and 0.18 SD). Indeed, in REP and REP+ public schools, girls actually begin first grade with an advantage over boys (respectively -0.06 and -0.08 SD). When considering schools IPS decile it also appears that the gender gap in maths at the beginning of second grade is increasing

in schools' IPS decile, *ie* most privileged schools are also those for which the difference in maths performance between boys and girls is the largest. This finding is consistent with the heterogeneity analysis conducted by Cobb-Clark and Moschion (2017) on Australian primary school students, showing that the gender gap in maths is the highest for high SES students.

Case studies

When examining schools with alternative pedagogies or private and religious schools, it is important to keep in mind that school characteristics and student characteristics are not strictly comparable across different types of schools, and the sample sizes for some specific school types may be small[18]. Nevertheless, in Freinet and Montessori schools, the gender gap in maths is higher (respectively 0.25 and 0.2 SD) compared to schools with no specific pedagogy. In Freinet schools, the gender gap follows a similar evolution as the aggregate gender gaps, increasing over time. In Montessori schools, the gender gap peaks during the second evaluation session before slightly decreasing. In Catholic religious schools, the gender gap in maths is similar in size and evolution compared to other types of schools. In Jewish schools, girls begin first grade with an advantage (0.08 SD), which then reverses, and the gender gap peaks at the beginning of second grade (around 0.055 SD). Both the level and increase of the gap in Jewish schools are lower compared to other school types. Muslim schools display the largest gender gap by the beginning of second grade. However, the sample size for Muslim schools may be too small to draw robust conclusions.

5 Investigating the evolution of the maths gender gap

While the previous section has tried to shed light on the magnitude, dynamics and heterogeneity of gender gaps, it remains mainly descriptive. In the following section I explore several hypothesis as to the reasons of their evolution. I first examine individual factors, focusing on initial scores as a baseline to control for initial ability differences between boys and girls. Next, I investigate the influence of schools and schooling on students' progress, considering school characteristics and educational experiences. Finally, I leverage the COVID lockdown as an exogenous shock to assess whether remote schooling had any impact on gender gaps.

5.1 Individual performance and specialization effects

Do initial performance levels influence the way students' rankings change over time? It should be kept in mind that results and information on individual performances are available to students and families as well as teachers after each evaluation, potentially affecting how students perform over time. Furthermore, it is possible that it is not only performance in a given subject that explains progress, but performances in other subjects with complementarity or substituability phenomena at play. If these mechanisms operate differently for boys and girls, they could

[18]In particular for Montessori schools, or Muslim schools who are in France very few to be overseen by the State

contribute to the dynamics of gender gaps, particularly in the context of the maths gender gap. As girls are initially more numerous to be individually better at French than maths (see section 4.1.2 above), their good performance in French could for instance be associated to their decline in maths. Gendered specialization processes could explain the maths gender gap.

I first look at the influence of initial scores on maths progress by estimating the following baseline specification for student i in cohort c, with t referring to the timing of evaluations:

$$prog_{ic}^{t} = \beta_0 + \beta_1 female_i + \beta_2 score\_maths_{ic}^{t-1} + \beta_3 score\_french_{ic}^{t-1} + \beta_4 female_i \times score\_maths_{ic}^{t-1} + +\beta_5 female_i \times score\_french_{ic}^{t-1} + \varepsilon_{ic}^{t}$$

I cluster standard errors by cohort × school. In alternative specifications I also add school characteristics and age controls X_{ic}^{t}, as well as introduce fixed effects in an additive manner: first only cohort fixed effects, then cohort and school fixed effects and finally I also add cohort × school fixed effects.

Table 2: Maths progress (beginning - end) and initial scores

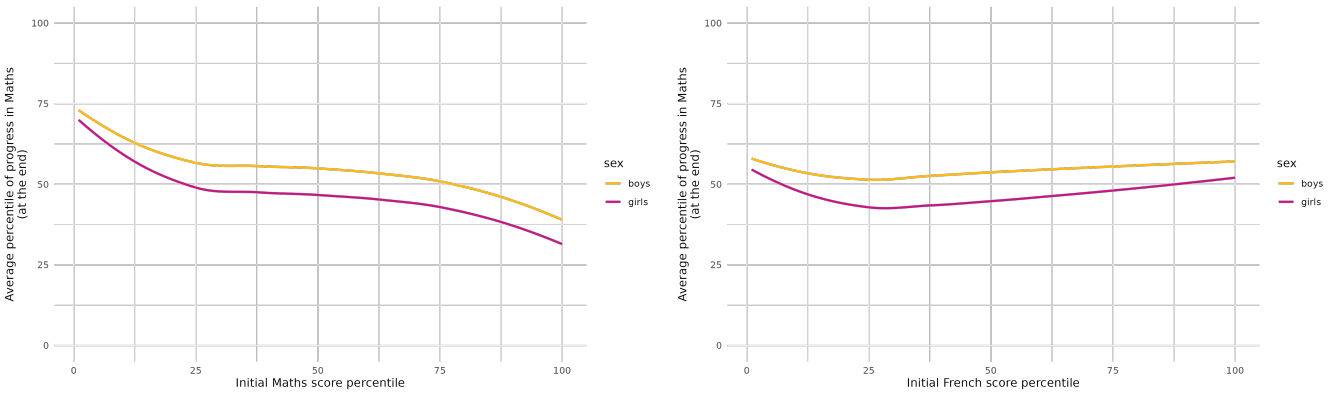
Dependent Variable:	Maths progress									
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Constant	0.1202***	-9.610***								
	(0.0010)	(0.1918)								
Female	-0.2428***	-0.2430***	-0.2428***	-0.2430***	-0.2443***	-0.2464***	-0.2443***	-0.2464***	-0.2462***	-0.2481***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Initial maths score	-0.5829***	-0.5853***	-0.5829***	-0.5853***	-0.5901***	-0.5916***	-0.5901***	-0.5916***	-0.5840***	-0.5858***
	(0.0014)	(0.0015)	(0.0014)	(0.0015)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)
Initial French score	0.3659***	0.3529***	0.3659***	0.3529***	0.3747***	0.3704***	0.3747***	0.3704***	0.3827***	0.3782***
	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0011)	(0.0012)	(0.0011)	(0.0012)	(0.0011)	(0.0012)
Female × Initial maths score	-0.0283***	-0.0270***	-0.0283***	-0.0270***	-0.0251***	-0.0245***	-0.0251***	-0.0245***	-0.0239***	-0.0231***
	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0017)	(0.0018)	(0.0017)	(0.0018)	(0.0018)	(0.0018)
Female × Initial French score	0.0501***	0.0498***	0.0500***	0.0497***	0.0493***	0.0502***	0.0493***	0.0502***	0.0500***	0.0510***
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0015)	(0.0016)	(0.0015)	(0.0016)	(0.0015)	(0.0016)
Controls										
School characteristics and age		Yes		Yes		Yes		Yes		Yes
Fixed-effects										
cohort			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
school_id					Yes	Yes	Yes	Yes	Yes	Yes
cohort × school_id									Yes	Yes
Fit statistics										
Observations	2,789,246	2,634,286	2,789,246	2,634,286	2,789,246	2,634,286	2,789,246	2,634,286	2,789,246	2,634,286
R^2	0.23037	0.24221	0.23039	0.24229	0.30022	0.30221	0.30022	0.30221	0.35812	0.35811
Within R^2			0.23034	0.24222	0.24184	0.24600	0.24184	0.24600	0.24200	0.24608

Notes: Clustered (school_id × cohort) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

The coefficients of interest (interactions between the female dummy and initial maths and French score) maintain consistent sign and magnitude across the various specifications and are all statistically significant at the 1% level. Boys with initial maths and French scores at 0 SD exhibit a 0.12 SD higher progress in maths during first grade compared to girls. The negative coefficient associated with the initial maths score for both boys and girls can be understood in light of mean reversion — it is easier for students initially below average to make progress than for those who were already at the top. In contrast, for the initial French score, the coefficient is positive for both genders, implying that a higher starting point in French leads to greater

progress in maths during first grade.
Regarding the interaction terms, the coefficient associated to being a girl and the initial maths score is negative (-0.0231 SD in the specification with controls and all fixed effects). This indicates that, on average, given the same initial maths score, girls make less progress in maths than boys. Conversely, for the interaction with the initial French score, the effect is positive (0.051 SD). While the sign of the first interaction term coefficient doesn't contradict the hypothesis of a gendered specialization process in first grade according to individual comparative advantages that could contribute to the maths gender gap, the second term (female $\times$ initial French score) however suggests that increasing initial French performance benefits girls' maths progress more than boys'.
It's noteworthy that the corresponding estimation with French progress as the outcome (Table 7) reveals that both interaction term coefficients are positive (although small), indicating that higher initial scores in both maths and French result in greater progress in French for girls. Thus there appears to be a specific effect of the initial maths score of girls' maths progress.

To see whether the average effect estimated by OLS is the same throughout the distribution, I plot the average progress in maths (and French in Appendix Figure 23) by initial maths and French (and maths minus French in Appendix Figure 23) score percentile for both gender.



(a) Initial maths percentile

(b) Initial French percentile

Figure 8: Average progress in maths depending on initial scores by gender

For an identical initial maths and French score percentile, boys tend to belong to higher progress percentiles compared to girls. The gender gap in progress between boys and girls is most pronounced and almost uniform between the 25^{th} and last percentile of initial maths scores, indicating that it is primarily among initial average and high performers that the gap widens the most (8 percentiles). Considering French initial score percentile, it is in the middle of the distribution that the difference in maths progress is the highest (8 percentiles). Note that while progress in maths is monotically decreasing in initial maths score (mean reversion), it is not so clear considering initial French scores. When considering the corresponding graphs for French progress (Figure 23), the curves for both gender are almost identical.

Finally, to see whether those that progress in maths also progress similarly in French (and *vice versa*) and whether there are gender differences, I relate progress in maths and progress in French by looking at the joint distribution of progress quartiles separately for girls and boys (Table 21). At first glance, distributions are very similar. For instance both for boys and girls, about 35% progress as much in maths and French (summing the percentages on diagonal cells), meaning that most do not progress at the same pace in both subjects. However, subtle disparities emerge upon closer inspection. Girls are slightly less in the top-right and bottom-left corners of the table, where the greatest differences in progress between the two subjects are found. Consequently, it can be inferred that girls tend to demonstrate a more coordinated advancement in both maths and French, suggesting a higher level of complementarity between progress in these subjects. Conversely, boys are more frequently observed to progress significantly in one subject while regressing in the other, indicating a less synchronized pattern of progress between maths and French.

In summary, the analyses I conduct do not provide conclusive evidence to establish the presence of a gendered specialization effect. Notably, there is a more positive association between initial French scores and subsequent maths progress for girls, implying that girls with higher initial proficiency in French also tend to make greater progress in maths than boys. Examining the average progress contingent on the initial maths score percentile reveals that the most substantial disparities occur above the 25^{th} initial percentile. This finding resonates with the observed heterogeneity of the gender gap, as previously demonstrated. Looking at the joint distribution of progress in maths and French, disparities between boys and girls do not appear significant enough to explain the pronounced surge in the maths gender gap favoring boys.

5.2 First grade schooling

Having ruled out that initial individual performances can account for the evolution of the maths gender gap, I investigate a second hypothesis mentioned in the CSEN note: that the maths gender gap widens in relation with schooling.

5.2.1 Learning environment and schools

Unfortunately, I do not have variables at the class level, nor can I link students to classes within schools. I therefore use school fixed effects and characteristics at the school level as proxies to see whether schooling characteristics have an influence over the maths gender gap and progress. Note that in a given school, characteristics can change over time[19], therefore when referring to schools, the unit of observation is actually school $\times$ cohort, hence I have variation not only between schools (and cohorts) but also within schools over cohorts. I observe students from approximately 120,000 schools (Table 15).

[19]With the exception of pedagogy and private religious variables.

First, to see whether differences in schooling can explain differences in progress and gender gaps I look at the R^2 of models at the individual level with school fixed effects, but no regressor, for student i, in cohort c and in school s

$$prog_{ics} = \beta_0 + school\_id_{cs} + \varepsilon_{ics}$$

Indeed, a significant finding is that school fixed effects account for more of the variance in the progress made in maths during the first grade for girls compared to boys. As displayed in Table 7, the proportion of explained variance in a model solely incorporating school fixed effects is 8.8% for boys and 10.0% for girls. The R^2 increases to 19.1% for boys and 21.1% for girls when introducing school × cohort fixed effects. Schooling and characteristics associated to schools therefore explain a lot of how students progress over time, and because their importance differ between gender, understanding the mechanisms at play could be key in understanding the evolution of the maths gender gap.

I then explore school-level variables that can be categorized into three primary aspects of students' educational experience: peers-related variables (IPS, proportion of female students), teachers-related variables (average age, seniority within school, proportion of female teachers) and school variables (enrollment, classroom size, teacher per student ratio).
While I arguably can't claim to identify a causal effect for many school characteristics as students are not randomly allocated between schools[20], and school characteristics may play a role in the allocation process[21], for some variables like the proportion of female students, or female teachers within a school, student selection seems less of a threat. In any case, even if only as descriptive evidence to be further investigated, correlations can be informative. I conduct the analysis at two levels: individual and aggregated at the school × cohort level.

First I use individual students' maths progress, and I estimate the following baseline specification with for student i, in cohort c, school s and X_{ics} a vector of school characteristics:

$$prog_{ics} = \beta_0 + \beta_1 female_i + X_{ics}\ \beta_2 + female_i \times X_{ics}\ \beta_3 + \varepsilon_{ics}$$

I cluster standard errors at the school × cohort level.

At the aggregated level I use as an outcome the difference in progress between boys and girls[22], and estimate the baseline specification:

[20] At the primary level, for public schools, the allocation mainly depends on geographical criteria with the existence of catchment areas at the municipality level. Location strategies can however not totally be ruled out, as well as avoidance strategies from families to avoid enrolling their children in the designated public school or in private schools. Public schools can't reject students if they have availability. For private schools, selection is the rule as families can choose where to enroll their children, and schools can reject students.

[21] For instance IPS that is also strongly correlated to students' family background which I don't observe, or teacher per pupil ratio that may also be an indicator of school resources that come in part from tuition fees in private schools

[22] Defined as, for school s and cohort c: $gendergap\_prog_{cs} = prog(boys)_{cs} - prog(girls)_{cs}$

$$gendergap\_prog_{cs} = \beta_0 + X_{cs}\ \beta_1 + \varepsilon_{cs}$$

For both, I add cohort and school fixed effects in alternative specifications.

As I am interested in the effect of multiple characteristics interacted with gender, note that there is an increased risk of type I error or false positive (multiple hypothesis testing). Although it is not informative as to individual variable significance, when performing a joint significance test (the null hypothesis being that all the coefficients of the interacted terms are 0) the F-statistic is high and allows me to reject the null at the 5% risk level.

Individual results In public schools compared to private schools, girls show a slightly greater regression in maths relative to boys (-0.007 SD in the model with all controls and fixed effects). In addition, in REP and REP+ schools, girls exhibit less regression compared to schools without priority status (0.014 SD and 0.017 SD, respectively). This implies that in REP and REP+ schools, not only do girls begin first grade with a stronger maths advantage over boys (as noted earlier), but they also experience less regression compared to boys. This finding suggests that the lower gender gap observed at the start of second grade in these schools is not solely attributed to its initial lower level. Although not always significant, girls seem to regress more in multilevel classes. The ratio teacher per pupil has a positive and significant effect in all specifications (+0.43 SD). Looking at the effect of peers, the school IPS has a slightly more positive influence on girls than on boys (+0.0008 SD), but the effect of the proportion of female students, positive, is not significant[23]. For the effect of teachers' variables, the proportion of women does not have a significant effect (although the coefficients are positive in the different specification), but it appears that girls regress a bit less than boys with older teachers that have higher seniority within the school (respectively +0.0006 and +0.0019 SD).

School-level results Although most coefficients are not significant, in the preferred specification with both schools and cohort fixed effects, their signs are mostly consistent with the coefficients of the interaction terms at the individual level[24], and in the different fixed effects specifications. The only weakly significant effect (10% level) I estimate is that an increase of one year of the average teachers' age decrease the gender gap in progress of 0.011 SD. See Table 31 in Appendix for more details.

[23]Note that public and private schools overlooked by the State are almost all coeducational in France, and that variation may therefore be low.

[24]One may expect that a positive effect on maths progress for girls translates to a negative effect on the gender gap in progress at the school level, and *vice versa*.

Table 3: Progress in maths (beginning - end), school characteristics and gender - individual level

Dependent Variable:	Maths progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-0.2561***	-4.623***				
	(0.0247)	(0.2196)				
Individual characteristics						
Female	-0.3391***	-0.3411***	-0.3390***	-0.3410***	-0.3342***	-0.3362***
	(0.0234)	(0.0233)	(0.0234)	(0.0233)	(0.0221)	(0.0221)
School characteristics						
Type‡ Public *(Private)*	0.0346***	0.0339***	0.0345***	0.0338***		
	(0.0034)	(0.0034)	(0.0034)	(0.0033)		
Priority‡ REP *(None)*	0.0934***	0.0926***	0.0926***	0.0917***		
	(0.0054)	(0.0054)	(0.0054)	(0.0054)		
Priority‡ REP+ *(None)*	0.1197***	0.1184***	0.1187***	0.1171***		
	(0.0070)	(0.0070)	(0.0070)	(0.0070)		
Enrollment & Resources						
Enrollment	0.0002***	0.0002***	0.0002***	0.0002***	0.0001*	0.0001*
	(1.31×10^{-5})	(1.31×10^{-5})	(1.31×10^{-5})	(1.31×10^{-5})	(6.98×10^{-5})	(6.98×10^{-5})
Teacher/student	-0.2609	-0.2082	-0.1713	-0.0938	1.003***	0.9772***
	(0.1996)	(0.1996)	(0.2002)	(0.2002)	(0.2828)	(0.2826)
First grade multilevel‡ Yes *(No)*	-0.0709***	-0.0716***	-0.0707***	-0.0714***	-0.0275***	-0.0278***
	(0.0032)	(0.0032)	(0.0032)	(0.0032)	(0.0038)	(0.0038)
First grade multilevel‡ Possibly *(No)*	-0.0085***	-0.0087***	-0.0082***	-0.0083***	-0.0010	-0.0013
	(0.0026)	(0.0026)	(0.0026)	(0.0026)	(0.0026)	(0.0026)
Student composition						
Female†	-0.0021	-0.0016	-0.0026	-0.0020	-0.0071	-0.0066
	(0.0253)	(0.0253)	(0.0253)	(0.0253)	(0.0354)	(0.0354)
IPS	0.0025***	0.0024***	0.0025***	0.0024***	4.56×10^{-5}	4.86×10^{-5}
	(8.83×10^{-5})	(8.83×10^{-5})	(8.83×10^{-5})	(8.82×10^{-5})	(0.0004)	(0.0004)
Teacher compostition						
Female†	0.0073	0.0076	0.0078	0.0083	0.0146	0.0154
	(0.0099)	(0.0099)	(0.0099)	(0.0099)	(0.0148)	(0.0148)
Age (years)	0.0006*	0.0006*	0.0006*	0.0007**	0.0010*	0.0010*
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0006)	(0.0006)
Seniority within school (years)	-0.0006	-0.0006	-0.0005	-0.0006	-0.0019*	-0.0019*
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0010)	(0.0010)
Gender interactions						
Female × Type Public	-0.0090***	-0.0091***	-0.0091***	-0.0091***	-0.0072**	-0.0072**
	(0.0031)	(0.0031)	(0.0031)	(0.0031)	(0.0030)	(0.0030)
Female × Priority REP	0.0146***	0.0148***	0.0146***	0.0148***	0.0142***	0.0144***
	(0.0049)	(0.0049)	(0.0049)	(0.0049)	(0.0047)	(0.0047)
Female × Priority REP+	0.0137**	0.0144**	0.0135**	0.0142**	0.0166***	0.0174***
	(0.0062)	(0.0062)	(0.0062)	(0.0062)	(0.0061)	(0.0060)
Female × Enrollment	-9.71×10^{-6}	-9.92×10^{-6}	-9.68×10^{-6}	-9.89×10^{-6}	-9.57×10^{-6}	-9.77×10^{-6}
	(1.11×10^{-5})	(1.11×10^{-5})	(1.11×10^{-5})	(1.11×10^{-5})	(1.07×10^{-5})	(1.07×10^{-5})
Female × Teacher/student	0.5114***	0.4825**	0.5124***	0.4832**	0.4596**	0.4294**
	(0.1885)	(0.1882)	(0.1885)	(0.1882)	(0.1810)	(0.1808)
Female × First grade multilevel Yes	-0.0035	-0.0032	-0.0035	-0.0032	-0.0031	-0.0028
	(0.0031)	(0.0031)	(0.0031)	(0.0031)	(0.0029)	(0.0029)
Female × First grade multilevel Possibly	-0.0080***	-0.0079***	-0.0080***	-0.0079***	-0.0083***	-0.0083***
	(0.0024)	(0.0023)	(0.0024)	(0.0023)	(0.0023)	(0.0023)
Female × (Students) Female	0.0389	0.0376	0.0389	0.0375	0.0334	0.0323
	(0.0266)	(0.0266)	(0.0266)	(0.0266)	(0.0236)	(0.0235)
Female × (Students) IPS	0.0007***	0.0008***	0.0007***	0.0008***	0.0008***	0.0008***
	(7.97×10^{-5})	(7.97×10^{-5})	(7.97×10^{-5})	(7.96×10^{-5})	(7.69×10^{-5})	(7.68×10^{-5})
Female × (Teachers) Female	0.0016	0.0009	0.0016	0.0009	0.0036	0.0029
	(0.0088)	(0.0087)	(0.0088)	(0.0087)	(0.0084)	(0.0084)
Female × (Teachers) Age	0.0007**	0.0007**	0.0007**	0.0007**	0.0006**	0.0006**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Female × (Teachers) Seniority	0.0018***	0.0018***	0.0018***	0.0018***	0.0018***	0.0019***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Controls						
Individual		Yes		Yes		Yes
Fixed-effects						
cohort			Yes	Yes	Yes	Yes
school_id					Yes	Yes
Fit statistics						
Observations	2,373,120	2,373,120	2,373,120	2,373,120	2,373,120	2,373,120
R^2	0.02246	0.02469	0.02257	0.02483	0.09399	0.09625
Within R^2			0.02249	0.02474	0.01650	0.01894

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared, var† indicates a proportion; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Case studies: alternative modes of schooling

I also explore the influence of alternative pedagogies and schooling by using the variables I created on pedagogy (Montessori, Freinet) and religion in private schools (Catholic, Jewish and Muslim) on maths progress. Controlling for school characteristics X_{ics}, I estimate the following specifications at the individual level, for student i of cohort c and school s:

$$prog_{ics} = \beta_0 + \beta_1 female_i + Z_s\ \beta_2 + female_i \times Z_s\ \beta_3 + X_{ics}\ \beta_4 + \varepsilon_{ics}$$

with Z either the *pedagogy* or *religion* variable. I cluster by school × cohort and add fixed effects in alternative specifications. Omitted categories are no alternative pedagogies, and public school.

Looking at the results of the specifications without fixed effects, in schools with Montessori or Freinet pedagogy students progress less in maths than in schools for which I don't identify any particular pedagogy. Now turning to the interaction terms, girls progress even less than boys (enrolled in regular schools) in Freinet schools (-0.078 SD, significant at the 5% level). For Montessori, the coefficient is positive, but not significant (+0.074 SD).

Compared to public school, boys progress significantly less in Catholic private schools and other private schools, when controlling for school characteristics (respectively -0.036 and -0.040 SD in the specification without fixed effects), but there is no significant effect for Jewish schools nor Muslim schools. Girls regress significantly (1% level) less than boys in private Catholic and other private schools (respectively +0.017 and +0.024 SD). Jewish schools seem to have a positive effect for girls, bigger (+0.05 SD) than Catholic schools but it is not significant, and neither is the negative estimated effect of Muslim schools(-0.111 SD). It should however be reminded that sample size for Muslim schools is extremely small, estimates should therefore be taken with caution.

5.2.2 Teachers' gender

In the French primary school system, a significant gender disparity exists among teachers, with a substantial majority being women. As of 2020, around 84% of teachers in public schools and up to 92% in state-overlooked private schools were women. This gender imbalance among teachers gradually diminishes in middle schools, secondary education, and higher education. While the prevalence of female teachers is the norm for first-grade students, it remains pertinent to investigate whether teacher gender has an impact on gender gaps at this early stage, as studies traditionally focus most on older students or later outcomes.

I use a 'class name' text field in the *Evaluations Repères* data. This field isn't consistently filled, but it allows me to infer teacher gender from descriptions containing gender-specific titles (*Monsieur* or *Madame*). Additionally, I use schools where the proportion of female teachers from the APAE data is either 0 or 1 to determine teacher gender for students. This approach

has allowed me to establish a link between approximately 1,200,000 student observations and the gender of their first-grade teachers. Looking at the descriptive statistics of students (see Table 12 in the Appendix), I observe that about 6% of observations in the obtained sample are associated with male teachers. This percentage is in contrast to the average proportion of male teachers across the observed schools (13%, as shown in Table 15 in the Appendix). In terms of gender gaps, there does not seem to be a significant difference in performance between students who have male teachers and those with female teachers by the end of the first grade.

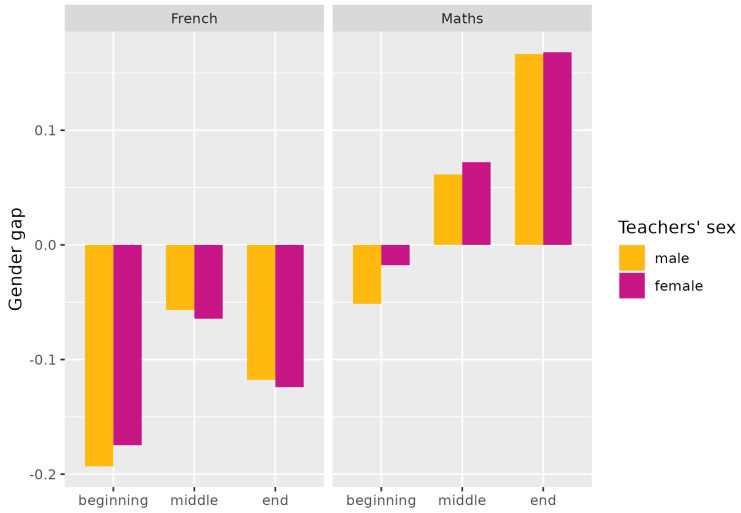


Figure 9: Gender gaps depending on teacher's gender
Note: Gender gaps are expressed in proportion of a standard deviation, and computed as $\overline{score}(boys)^t - \overline{score}(girls)^t$

However, despite not observing a statistical difference in the proportion of boys and girls based on teacher gender, there exist notable differences between students who have male or female teachers. For instance, students with a female teacher tend to be enrolled in larger schools with more students per class, and these schools generally have slightly older teachers on average. Moreover, the school's IPS is significantly higher for students with female teachers. Importantly, students with a female first-grade teacher also begin their academic journey with better results in both maths and French. It is therefore difficult to conclude anything on the level or increase in gender gaps for students with male or female teachers with simple statistics.

OLS strategy

To look at the influence of teachers' gender on maths progress I first estimate the following specification, for student i in cohort c and school s:

$$prog_{ics} = \beta_0 + \beta_1 female_i + \beta_2 teacher\_male_{ics} + \beta_3 female_i \times teacher\_male_{ics} + \varepsilon_{ics}$$

I cluster standard errors by school × cohort, and add a vector of school controls X_{ics} and fixed effects. Note that because having a female teacher is the norm, I use a dummy for having a

male teacher instead.

Table 4: Progress in maths (beginning - end) and pupil and teacher gender

Dependent Variable:	Maths progress											
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	0.0784***	-4.867***	-5.067***									
	(0.0015)	(0.2981)	(0.3056)									
Female student	-0.1854***	-0.1870***	-0.1872***	-0.1853***	-0.1869***	-0.1872***	-0.1847***	-0.1863***	-0.1867***	-0.1849***	-0.1865***	-0.1868***
	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0013)	(0.0013)	(0.0014)	(0.0014)	(0.0014)	(0.0014)
Male Teacher	0.0115*	0.0125**	-0.0011	0.0114*	0.0123**	-0.0011	-0.0004	0.0003	-0.0009	-0.0021	-0.0013	-0.0014
	(0.0061)	(0.0061)	(0.0064)	(0.0061)	(0.0061)	(0.0064)	(0.0070)	(0.0070)	(0.0071)	(0.0079)	(0.0079)	(0.0080)
Female student × Male teacher	-0.0308***	-0.0315***	-0.0320***	-0.0308***	-0.0315***	-0.0319***	-0.0306***	-0.0313***	-0.0319***	-0.0286***	-0.0293***	-0.0300***
	(0.0059)	(0.0059)	(0.0061)	(0.0059)	(0.0059)	(0.0061)	(0.0058)	(0.0058)	(0.0059)	(0.0058)	(0.0058)	(0.0060)
Controls												
Individual		Yes	Yes		Yes	Yes		Yes	Yes		Yes	Yes
School level			Yes			Yes			Yes			Yes
Fixed-effects												
cohort				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
school_id							Yes	Yes	Yes	Yes	Yes	Yes
cohort × school_id										Yes	Yes	Yes
Fit statistics												
Observations	1,254,663	1,254,663	1,188,520	1,254,663	1,254,663	1,188,520	1,254,663	1,254,663	1,188,520	1,254,663	1,254,663	1,188,520
R^2	0.01580	0.01826	0.02604	0.01584	0.01833	0.02614	0.10728	0.10965	0.10832	0.18321	0.18545	0.18214
Within R^2				0.01580	0.01829	0.02609	0.01695	0.01956	0.01972	0.01785	0.02055	0.02058

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions.

In every specification, having a male teacher has a small but significant effect for girls compared to boys, as it decreases girls' maths progress of 0.03 SD.

Instrumental variable strategy

However, it is unclear whether the allocation process of students to teachers, and teachers to classroom, made at the school level, is dependent on teachers and pupils' gender. For instance, one could think of school principals allocating students that have difficulties with female or male teachers more, in case there are several first grade classes. Within the pool of teachers in the school, they could also allocate teacher of one gender to first grade classes more frequently. If so, the previous result would be biased. To more credibly establish the causality of the effect of having a male teacher, I therefore instrument teacher's gender. The instrument I use is the proportion of female teachers at the school from the APAE dataset, the rationale being that having a higher proportion of female teachers within the school makes it less likely to have a male first grade teacher. For this strategy to be valid, the instrument has to meet two requirements.

Relevance is not an issue: I show in Table 7 in Appendix the F statistics in the different specifications of the two first stage regressions. In both cases, the F-statistic is always high and well above 10 (Stock & Yogo 2002), which intuitively makes sense: having a high proportion of female teachers within a school decreases the probability of having a male teacher.

Exogeneity is more challenging, as it necessitates that the impact of the $teachers\_prop\_female$ variable on students' progress occurs exclusively through the heightened likelihood of having a female teacher in the first grade. This assumption appears plausible, if one considers that the

bulk of student-teacher interactions occur within the classroom. However, situations in which students have more interactions with other grade teachers, or scenarios where the proportion of female teachers at the school level influences students' educational experience in alternative manners, would violate the principle of exogeneity. Moreover, the strategy could be compromised if the gender of teachers allocated to schools is not random but depend on variables that also have an influence on students' outcomes, for instance IPS. In such instances, the estimator would become inconsistent.

I then estimate the effect of interest using two-stage least squares. I have two endogenous variables[25], therefore two first stage equations, and adding the second stage equation:

$$\begin{aligned}
teacher_male_{ics} &= \beta_0 + \beta_1 female_i + \beta_2 teachers_prop_female_{ics} \\
&\quad + \beta_3 female_i \times teachers_prop_female_{ics} + e_{ics} \\
female_i \times teacher_male_{ics} &= \beta_0 + \beta_1 female_i + \beta_2 teachers_prop_female_{ics} \\
&\quad + \beta_3 female_i \times teachers_prop_female_{ics} + \epsilon_{ics} \\
prog_{ics} &= \beta_0 + \beta_1 female_i + \beta_2 \widehat{teacher_male}_{ics} \\
&\quad + \beta_3 female_i \times \widehat{teacher_male}_{ics} + \varepsilon_{ics}
\end{aligned}$$

In alternative specifications I add a vector of controls, first at the individual level (age and age squared) and then school characteristics. I cluster standard errors at the school × cohort level.

The results obtained with 2SLS are very close to the ones estimated by OLS, although the effect is slightly bigger. With fixed effects and controls, having a male teacher decreases girls' maths progress by 0.039 SD on average. Results are significant at the 1% level. However, because the proportion of male teachers is small, back-of-the-enveloppe calculations reveal that switching to only female teachers in first grade would only reduce the end maths gender gap by 2% at most (see supplementary results in Appendix for more details). When considering French progress, note that both using OLS and 2SLS I also estimate a negative effect for female students to be assigned to a male first grade teacher. The effect even seems to be larger than in maths (-0.027 SD in OLS; -0.07 SD in 2SLS).

[25]Teacher's gender and the interaction between student's gender and teacher's gender.

Table 5: IV teacher gender: Second stage, maths progress beginning - end

Dependent Variable:	Maths progress								
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.0718***	-4.870***	-5.089***						
	(0.0017)	(0.2982)	(0.3054)						
Female student	-0.1853***	-0.1868***	-0.1871***	-0.1853***	-0.1868***	-0.1871***	-0.1845***	-0.1860***	-0.1862***
	(0.0016)	(0.0016)	(0.0017)	(0.0016)	(0.0016)	(0.0017)	(0.0016)	(0.0016)	(0.0017)
Male teacher	0.1217***	0.1264***	0.0388**	0.1205***	0.1248***	0.0361*	0.0351	0.0341	0.0148
	(0.0172)	(0.0173)	(0.0197)	(0.0173)	(0.0173)	(0.0197)	(0.0518)	(0.0518)	(0.0593)
Female student × Male teacher	-0.0311*	-0.0323**	-0.0340*	-0.0311*	-0.0324**	-0.0340*	-0.0338**	-0.0353**	-0.0387**
	(0.0160)	(0.0160)	(0.0177)	(0.0160)	(0.0160)	(0.0177)	(0.0156)	(0.0156)	(0.0173)
Controls									
Individual		Yes	Yes		Yes	Yes		Yes	Yes
School-level			Yes			Yes			Yes
Fixed-effects									
cohort				Yes	Yes	Yes	Yes	Yes	Yes
school_id							Yes	Yes	Yes
Fit statistics									
Observations	1,254,449	1,254,449	1,188,520	1,254,449	1,254,449	1,188,520	1,254,449	1,254,449	1,188,520
R^2	0.01458	0.01697	0.02591	0.01465	0.01707	0.02603	0.10721	0.10958	0.10831
Within R^2				0.01460	0.01702	0.02598	0.01690	0.01951	0.01971

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions.

5.2.3 COVID-19 lockdown and cohort comparison

An additional avenue for research could involve examining the impact of the COVID-19 pandemic, which coincided with the second cohort's first-grade year and entailed an extended lockdown period (from March 18^{th} to May 10^{th}). During this period schools closed, and schooling switched to remote classroom. Students likely were exposed to an increased family influence, and potentially shaped by distinct gender expectations and stereotypes compared to those encountered in school and among peers. In fact this question can be framed in the broader debate mentioned in the CSEN note and whether schooling actually has an effect on the maths gender gap. While it is not clear beforehand in which direction one would expect the effect to go, and whether school mitigates gender gaps or aggravates them, investigating whether the second cohort's experience during lockdown led to a divergent evolution of the maths gender gap compared to other cohorts could yield insightful findings for policymakers.

Upon initial examination, a comparison of the maths gender gap for the second cohort (2019-2020) with adjacent cohorts does not immediately reveal a discernible impact of the COVID-19 lockdown. While the gender gap at the onset of second grade appears higher for the first cohort than for the COVID-19 cohort, it closely resembles the third cohort's gender gap.

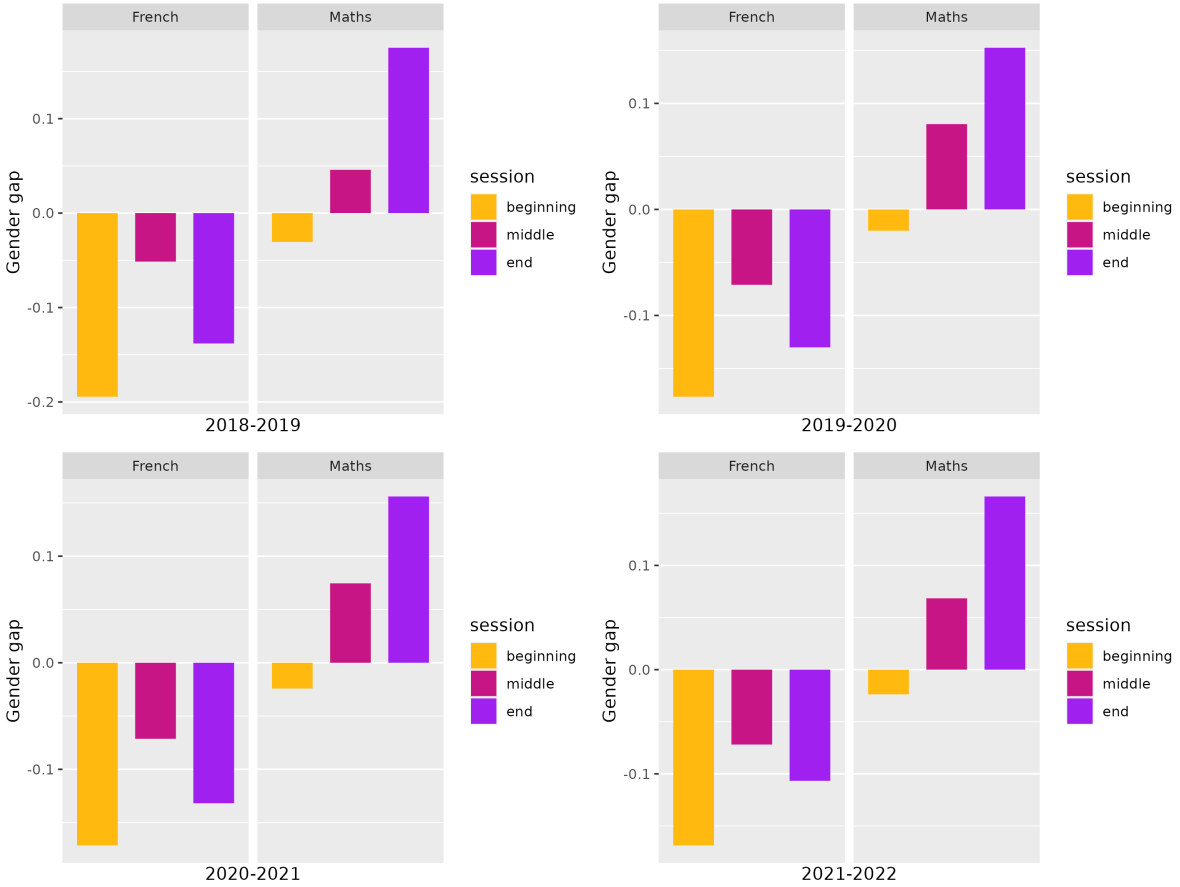


Figure 10: Gender gaps (by cohort)
Note: Gender gaps are expressed in proportion of a standard deviation, and computed as $\overline{score}(boys)^t - \overline{score}(girls)^t$

Triple difference-in-difference

A potential setting to investigate this question is to use a (Triple) difference-in-difference strategy to estimate the effect of two months of remote schooling on the maths gender gap, taking the first cohort as control group. Formally, using the individual maths score (not progress) $score_{ict}$ as outcome, and considering the effect of at-home schooling during the lockdown ($treated$):

$$\begin{aligned} score_{ict} = {} & \beta_0 + \beta_1 female_i + \beta_2 treated_c + \beta_3 after_t \\ & + \beta_4 female_i \times treated_c + \beta_5 female_i \times after_t + \beta_6 treated_c \times after_t \\ & + \beta_7 female_i \times treated_c \times after_t + \varepsilon_{ict} \end{aligned}$$

I add a vector of school characteristics and age controls X_{ict} in alternative specifications, and cluster standard errors by school $\times$ cohort.

β_7 identifies the effect of interest (school closure and remote schooling) under the assumption that without two months of remote at-home schooling, the maths gender gap would have trended similarly in the control and treated groups (parallel trend assumption), and that no other shock

affects the treated cohort's outcome. Olden and Møen (2022) also show that the estimate of the triple difference-in-difference estimator can be rewritten as the difference in the gender gap before and after the lockdown for the cohort that experiences it (treated) minus the difference in the gender gap (before minus after) for the control cohort.

$$\begin{aligned}\widehat{\beta_7} &= \Big[\big(\overline{score}^{treated}_{before}(boys) - \overline{score}^{treated}_{before}(girls)\big) - \big(\overline{score}^{treated}_{after}(boys) - \overline{score}^{treated}_{after}(girls)\big)\Big] \\ &- \Big[\big(\overline{score}^{control}_{before}(boys) - \overline{score}^{control}_{before}(girls)\big) - \big(\overline{score}^{control}_{after}(boys) - \overline{score}^{control}_{after}(girls)\big)\Big] \\ &= \Big(\overline{gendergap}^{treated}_{before} - \overline{gendergap}^{treated}_{after}\Big) - \Big(\overline{gendergap}^{control}_{before} - \overline{gendergap}^{control}_{after}\Big)\end{aligned}$$

However, comparing cohorts to establish a causal link between the COVID-19 lockdown and its potential impact on the maths gender gap presents several challenges. First, variations in the content of evaluations over the years, especially during the first and second cohorts, can make it challenging to select an appropriate control cohort. While the first cohort might differ more in terms of evaluated content, the third cohort also experienced the COVID-19 pandemic and lockdown, albeit not during their first-grade year. Secondly, distinguishing the specific effects of remote schooling from broader factors related to the pandemic's circumstances, such as anxiety and movement restrictions, complicates the estimation of the treatment's impact. Notably, research indicates that women are more susceptible to anxiety, and early studies have identified higher anxiety and depression symptoms among girls compared to boys during the pandemic (Panchal et al. 2021, DEPP Note d'information 19 2021, Mendolia et al. 2022).

Finally, while there is *a priori* no reason for the parallel trend assumption not to hold, it seems less clear when examining pre-treatment trends, particularly for the first cohort. Visual inspection of gender gap trajectories among cohorts highlights potential non-parallel trends before the lockdown, accompanied by significant interaction term in pre-trend regression analysis for the first cohort (Figure 24 and Table 7). For the third cohort however, both visually and estimating whether there is an effect before treatment, trends seem to be parallel before treatment.

For all these reasons, the following results need to be taken with caution. Using the third cohort as control, the maths gender gap increase 0.01 SD less after the lockdown for the cohort that experienced remote schooling. This result is consistent with the hypothesis that schooling does play a role in the increase of the maths gender gap. Although significant (1%), the estimated effect nonetheless appears to be very small.

Table 6: COVID-19 effect on the maths gender gap

Dependent Variable:	Maths score			
Control Group:	First Cohort		Third Cohort	
Model:	(1)	(2)	(3)	(4)
female	-0.0470***	-0.0456***	-0.0750***	-0.0726***
	(0.0025)	(0.0025)	(0.0023)	(0.0023)
treated	0.0287***	0.0290***	-0.0034	0.0022
	(0.0044)	(0.0037)	(0.0042)	(0.0035)
after	0.0595***	0.0606***	0.0302***	0.0312***
	(0.0022)	(0.0022)	(0.0019)	(0.0019)
female × treated	-0.0350***	-0.0344***	-0.0069**	-0.0065**
	(0.0034)	(0.0034)	(0.0033)	(0.0033)
female × after	-0.1304***	-0.1303***	-0.0827***	-0.0830***
	(0.0019)	(0.0020)	(0.0016)	(0.0017)
treated × after	-0.0282***	-0.0285***	0.0011	0.0009
	(0.0030)	(0.0031)	(0.0028)	(0.0028)
female × treated × after	0.0582***	0.0579***	0.0106***	0.0105***
	(0.0026)	(0.0026)	(0.0024)	(0.0024)
Controls		Yes		Yes
Fit statistics				
Observations	2,337,952	2,240,682	2,546,774	2,444,864
R^2	0.00477	0.06545	0.00460	0.06853
Adjusted R^2	0.00477	0.06544	0.00459	0.06852

Notes: Clustered (school_id × cohort) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Controls include average students' age and school-level variables used in schooling environment regressions.

6 Discussion

6.1 Limits and measurement issues

While this Master's thesis has tried to highlight descriptive aspects of gender gaps for French pupils during first grade, and present potential factors explaining their evolutions, several caveats should be kept in mind.

6.1.1 Test modalities and gender gaps

While I could not find any skill associated to a type of exercice that has been demontrated to favor more one gender or the other in the Literature, there is some evidence that tests modalities actually matter. Several studies have found greater gender differences in favour of girls on constructed-response items compared to multiple-choice items (Bridgeman and Lewis,

1994; Lindberg, Hyde, Petersen, and Linn, 2010; Griselda 2020). Boys, compared to girls, tend to take more risk in multiple choices that end up to be rewarding. More specifically, when considering reading exercises, the content of the text matters (Lafontaine and Monseur, 2009) as narrative and long texts advantage girls while boys are more influenced by the topic and gender of the narrator. Finally high stake tests are usually found to favor boys compared to tests taking place in more relaxed environments like within a classroom (Arias et al, 2023). This of course has more far-reaching implications. For instance in the US, in the context of college admission tests, the issue is known as the female underprediction effect: SATs and ACTs have been shown to favor boys[26] compared to girls when taking as a counterfactual performances and success in maths after admission (Fischer et al, 2013). The consequences for gender equality are obvious, and failing to predict girls' success in STEM fields not only leads to underadmission of female candidates, but also has consequences in terms of funding and scholarship opportunities.

Coming back to gender gaps, the issue is that including certain type of exercises to be evaluated in French or maths in *Evaluations Repères* is not necessarily neutral and can affect the measured level of gender gaps. Even when considering evolutions, because exercises change between evaluations, the content of the different evaluations could favor one gender or the other in different ways over time. Relating the literature to the *Evaluation*, the format is closer to multiple choice that would favor boys. The texts can be either descriptive or narrative, but when there are protagonnists, they either are neutral (a couple, Nathan and Chloé...), or male (Medor the dog, Dino the dinosaur, or an unnamed young boy). Finally, and while in the French education system grading has an important place, because the evaluation is not graded and does not have any direct consequence on students' life (although results are communicated to parents and teachers), one can also wonder about the motivation of students, and of boys in particular.

6.1.2 Timing of tests and Summer Learning Loss

It also has to be taken into account that the last evaluation happens after the Summer break, which in France lasts two months. If the break affects boys and girls in a different way, my results could either be seen as upper or lower bounds of the true effect of schooling. In the literature, research on a differentiated effect of the Summer break in learning, using growth models, has traditionally put the emphasis more on socio-economic background or ethnicity rather than gender. Nonetheless, Downey et al (2022) hypothesize that school actually has a mitigating effect on gender gaps. Using a sample of 900,000 US children they show that girls learn faster when school is out (over the Summer), both in reading and in maths, but that this is reversed when school is in session. Workman, von Hippel and Merry (2023) performing a meta-review however show that results often fail to replicate. Using three different data source[27], for boths maths and English, they get non-consistent results of the effect of the Summer break depending

[26] Especially in the quantitative section, while scores are comparable on the reading and writing sections
[27] ECLS-K:2011, Renaissance Learning and NWEA

on gender[28]. Finally, reconciling the effect of Summer holidays and of two months of remote schooling during the COVID-19 lockdown, as both involve an increased family influence, note that the results I find go against Downey et al's claim that school has a mitigating effect on disparities, at least when it comes to gender. It therefore seems difficult to conclude anything as to the direction of the potential bias.

In that regard, alternative measures of progress, for instance between the first (September) and second evaluation (January) may constitute an interesting robustness check to account for a possible differentiated gender effect of the Summer break. I present the main maths specifications using this alternative outcome in the Appendix. If not of the same magnitude, most results have a consistent sign. For instance using the second progress outcome, the effect of having a male teacher on maths progress for female students is -0.02 SD when estimated using OLS, and -0.037 SD when estimated using 2SLS.

6.1.3 Other data limitations and measurement error

One important limit of this work is that I do not have other individual variables than gender and age, hence cannot properly control for the influence of family background. Ideally, one would want to control for socio-economic factors such as household income, parents' jobs and education level, or even siblings... This is a problem because students are obviously not randomly enrolled into schools, and failing to account for it diminishes the credibility of causality because of omitted variable bias. Likewise, because I can't link students to classrooms, I cannot adequately control for peer effects within the classroom, the influence of which may be expected to have a role on students' progress. Apart from students' classroom gender composition, it would for instance be interesting to be able to investigate whether the initial ranks of other students has an impact on students' progress in maths, or whether the gender of the maths/French top performers influences their classmate of the same and opposite gender differently.

45,000 teachers are involved in the evaluation sessions and although guidelines are precise, there is necessarily some variation in the test conditions and grading experienced by students. There is no bias as long as the measurement error on scores is not correlated with the truth and explanatory variables (classical measurement error). Considering explanatory variables, note that I do not have variables at the classroom level, that would be the most relevant unit of observation, but instead use variables aggregated at the school level. Hence, with x_{ics} the true value of the classroom variable for individual i in cohort c and school s, and $\tilde{x}_{cs}$ the observed value that is an average of the considered characteristic at the school level, the true classroom regressor can be expressed as $x_{ics} = \tilde{x}_{cs} + u_{ics}$. The resulting measurement error on classroom variables is not of the classical type, but of the optimal predictor type, *ie* the error is uncorrelated with the observed regressor ($u_{ics} \perp \tilde{x}_{cs}$). Hyslop and Imbens (2001) show that in such cases, the OLS estimator remains consistent.

[28] ≪ *In both reading and math, the NWEA and Renaissance tests showed girls gaining on or pulling ahead of boys during most but not all summers, and boys regaining lost ground during most but not all school years. On the ECLS-K:2011 tests, however, no seasonality in gender gaps was evident.*≫

6.2 Mechanisms and leads for further work

While I hope this work has shed light on several phenomena of interest, my results leave room for interpretation in terms of mechanisms and work is still needed to understand the dynamics of gender gaps in first grade and beyond, and the influence of schooling.

For instance, various potential channels through which teacher gender could influence educational outcomes have been explored in the literature[29]. I have shown that being assigned to a male teacher has a negative effect on French female students' progress in maths. It is however not clear what the reason is, and the data does not allow me to investigate further. One could understand it in light of role model theories, female teachers having a positive influence on female students. Further work could try to establish whether male or female teachers hold more or less gender stereotypical views as to students' maths ability[30]. It could also try to open the black box of classroom interactions to see whether teacher-students interactions can explain maths performance. My results could for instance be compatible with the interpretation that female teachers hold less gender stereotypical views on maths ability, and have more gender neutral teaching practices. There is also a literature on the attitudes of teachers to maths depending on gender (Relich et al, 1991; Relich 1996). My results are however not consistent with the hypothesis that female teachers' lower maths self-concept could be more detrimental to girls, or at least this effect does not predominate.

Further work could also try to more robustly identify the effect of remote schooling during the COVID-19 pandemic, taking into account the limits described above. Ideally, one would *a minima* want to control for factors such as anxiety or lockdown conditions to disentangle the actual effect of a change in schooling and the more general COVID-19 situation. During the lockdown, in France, some students that had parents working at the hospital actually remained at school, even though not under normal conditions[31] also because the measure was mainly implemented to provide a daycare option for parents that were also medical workers and allow them to go to work. Of course such students also experienced less severe lockdowns in terms of freedom of movement restrictions. Identifying the children that benefitted from such a measure would nonetheless be interesting and could provide a control group when considering gender gaps.

[29]These include gendered teaching practices or interactions with students (Bassi et al, 2018), the role model effect of having a teacher of the same gender, the Pygmalion effect where teachers' expectations shape students' performance, and the influence of implicit stereotypes held by teachers (Lavy and Sand, 2018; Carlana, 2019).

[30]And whether patterns of heterogeneity exist for instance depending on teachers' age or level of education.

[31]For instance, not necessarily in their usual schools

7 Conclusion

As we continue to strive for equality and inclusivity, the findings from this study may provide valuable insights for policymakers, educators, and researchers aiming to address and mitigate gender-based disparities in early academic experiences.

I document that in France although there is a gender gap already in favour of girls at the beginning of first grade, the maths gender gap in favour of boys is non existent, but widens during first grade. The heterogeneity analysis I conduct reveals that gender gaps are most pronounced in more privileged schools, and that girls stall most at the top, becoming increasingly less represented.

Looking at the individual factors that may explain the evolution of gender gaps, I find a distinct influence of initial maths and French score on maths progress between girls and boys. However my findings do not support the hypothesis of gendered specialisation whereby girls would regress in maths because they initially are (individually) better in French. Further work would be needed to make sense of how students progress in maths in relation to both maths and French initial results, and whether the process is differentiated between gender.

When considering students' schooling experience, and acknwoledging data limitations that prevent me from addressing all endogeneity concerns, my findings notably highlight the importance of teachers. Girls benefit more than boys of having more experienced teachers, both in terms of age and of seniority within the school. Furthermore, instrumenting teachers' gender by the proportion of female teachers within the school, I establish that female students progress less in maths compared to boys when assigned to a male teacher. Contrasting school and family influence in the context of the COVID-19 lockdown, I find a small effect of the lockdown with the gender gap increasing slightly less for the cohort that experiences two months of remote schooling during the lowkdown. While the sign of the effect is consistent with the hypothesis of school as a less gender neutral environment than the family, further work is needed to disentangle treatment from other effects, and more robustly establish causality.

References

[1] Oscar Arias, Catalina Canals, Alejandra Mizala, and Francisco Meneses. Gender gaps in Mathematics and Language: The bias of competitive achievement tests. *Plos one*, 18(3):e0283384, 2023.

[2] David Autor, David Figlio, Krzysztof Karbownik, Jeffrey Roth, and Melanie Wasserman. School quality and the gender gap in educational achievement. *American Economic Review*, 106(5):289–295, 2016.

[3] Marina Bassi, Mercedes Mateo Díaz, Rae Lesser Blumberg, and Ana Reynoso. Failing to notice? Uneven teachers' attention to boys and girls in the classroom. *IZA journal of labor economics*, 7(1):1–22, 2018.

[4] David N Beede, Tiffany A Julian, David Langdon, George McKittrick, Beethika Khan, and Mark E Doms. Women in STEM: A gender gap to innovation. *Economics and Statistics Administration Issue Brief*, (04-11), 2011.

[5] Thomas Breda, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre. Do female role models reduce the gender gap in science? Evidence from french high schools. 2021.

[6] Thomas Breda and Clotilde Napp. Girls' comparative advantage in reading can largely explain the gender gap in math-related fields. *Proceedings of the National Academy of Sciences*, 116(31):15435–15440, 2019.

[7] Anne Ardila Brenøe and Ulf Zölitz. Exposure to more female peers widens the gender gap in STEM participation. *Journal of Labor Economics*, 38(4):1009–1054, 2020.

[8] Brent Bridgeman and Charles Lewis. The relationship of essay and multiple-choice scores with grades in college courses. *Journal of Educational Measurement*, 31(1):37–50, 1994.

[9] David Card and A Abigail Payne. High school choices and the gender gap in STEM. *Economic Inquiry*, 59(1):9–28, 2021.

[10] Michela Carlana. Implicit stereotypes: Evidence from teachers' gender bias. *The Quarterly Journal of Economics*, 134(3):1163–1224, 2019.

[11] Michela Carlana and Lucia Corno. Parents and peers: Gender stereotypes in the field of study. 2021.

[12] Deborah A Cobb-Clark and Julie Moschion. Gender gaps in early educational achievement. *Journal of Population Economics*, 30:1093–1134, 2017.

[13] CSEN. Qu'apprend-on des évaluations de CP-CE1? In *Note 21-03*.

[14] Nilanjana Dasgupta and Shaki Asgari. Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of experimental social psychology*, 40(5):642–658, 2004.

[15] DEPP. Confinement et fermeture des écoles au printemps 2020 : le vécu des familles d'enfants scolarisés en CP et en CE1. In *Note d'information 22.03*.

[16] DEPP. Dispositif d'évaluation des conséquences de la crise sanitaire : Comment les élèves ont-ils vécu le confinement de mars-avril 2020 ? In *Note d'information 21.19*.

[17] DEPP. Évaluations Repères 2021 de début de CP et de CE1 : les effets négatifs de la crise sanitaire de 2020 surmontés en 2021. In *Note d'information 22.01*.

[18] DEPP. PISA 2018 : Culture mathématique, culture scientifique et vie de l'élève. In *Note d'information 19.50*. 2019.

[19] DEPP. PISA 2018 : Stabilité des résultats en compréhension de l'écrit. In *Note d'information 19.49*. 2019.

[20] DEPP. TIMSS 2019 – Évaluation internationale des élèves de CM1 en mathématiques et en sciences : les résultats de la France toujours en retrait. In *Note d'information 20.46*. 2020.

[21] DEPP. TIMSS 2019 – Évaluation internationale des élèves de CM1 en mathématiques et en sciences : les résultats de la France toujours en retrait. In *Note d'information 20.47*. 2020.

[22] DEPP. Filles et garçons sur le chemin de l'égalité. De l'école à l'enseignement supérieur. 2022.

[23] DEPP. PIRLS 2021 : La France stabilise ses resultats contrairement aux autres pays européens majoritairement en baisse. In *Note d'information 23.21*. 2023.

[24] Douglas B Downey, Megan Kuhfeld, and Margriet van Hek. Schools as a relatively standardizing institution: The case of gender gaps in cognitive skills. *Sociology of Education*, 95(2):89–109, 2022.

[25] Marie Duru-Bellat. La (re) production des rapports sociaux de sexe: quelle place pour l'institution scolaire? *Travail! genre et sociétés*, (1):131–149, 2008.

[26] Marie Duru-Bellat. À l'école du genre. *Enfances & Psy*, (1):90–100, 2016.

[27] Franziska T Fischer, Johannes Schult, and Benedikt Hell. Sex-specific differential prediction of college admission tests: A meta-analysis. *Journal of Educational Psychology*, 105(2):478, 2013.

[28] Roland G Fryer Jr and Steven D Levitt. An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, 2(2):210–240, 2010.

[29] Marco Giesselmann and Alexander W Schmidt-Catran. Interactions in fixed effects regression models. *Sociological Methods & Research*, 51(3):1100–1127, 2022.

[30] Silvia Griselda et al. Different Questions, Different Gender Gap: Can the Format of Questions Explain the Gender Gap in Mathematics? Technical report, Job Market Papers, 2020.

[31] Luigi Guiso, Ferdinando Monte, Paola Sapienza, and Luigi Zingales. Culture, gender, and math. *Science*, 320(5880):1164–1165, 2008.

[32] Pascal Huguet and Isabelle Regner. Stereotype threat among schoolgirls in quasi-ordinary classroom circumstances. *Journal of educational psychology*, 99(3):545, 2007.

[33] Dean R Hyslop and Guido W Imbens. Bias from classical and other forms of measurement error. *Journal of Business & Economic Statistics*, 19(4):475–481, 2001.

[34] Peter Y Ji, Richard T Lapan, and Kevin Tate. Vocational interests and career efficacy expectations in relation to occupational sex-typing beliefs for eighth grade students. *Journal of Career Development*, 31:143–154, 2004.

[35] Dominique Lafontaine and Christian Monseur. Gender gap in comparative studies of reading comprehension: To what extent do the test characteristics make a difference? *European Educational Research Journal*, 8(1):69–79, 2009.

[36] Victor Lavy and Edith Sand. On the origins of gender gaps in human capital: Short-and long-term consequences of teachers' biases. *Journal of Public Economics*, 167:263–279, 2018.

[37] Larry L Leslie, Gregory T McClure, and Ronald L Oaxaca. Women and minorities in science and engineering: A life sequence analysis. *The journal of higher education*, 69(3):239–276, 1998.

[38] Sara M Lindberg, Janet Shibley Hyde, Jennifer L Petersen, and Marcia C Linn. New trends in gender and mathematics performance: a meta-analysis. *Psychological bulletin*, 136(6):1123, 2010.

[39] Bertrand Marianne. New perspectives on gender. In *Handbook of labor economics*, volume 4, pages 1543–1590. Elsevier, 2011.

[40] Michael O Martin, Ina VS Mullis, Pierre Foy, and Martin Hooper. International Results in Reading in PIRLS 2016. *International Association for the Evaluation of Educational Achievement*, 2017.

[41] Michael O Martin, Ina VS Mullis, Pierre Foy, Dana L. Kelly, and Bethany Fishbein. TIMSS 2019 International Results in Mathematics and Science. *International Association for the Evaluation of Educational Achievement*, 2020.

[42] Silvia Mendolia, Agne Suziedelyte, and Anna Zhu. Have girls been left behind during the COVID-19 pandemic? Gender differences in pandemic effects on children's mental wellbeing. *Economics Letters*, 214:110458, 2022.

[43] Natalia Nollenberger and Núria Rodríguez-Planas. Let the Girls Learn! It is Not Only About Math... It's About Gender Social Norm. 2017.

[44] OCDE. *PISA 2018 Results (Volume II)*. 2019.

[45] OCDE. *Education at a Glance 2022*. 2022.

[46] Andreas Olden and Jarle Møen. The triple difference estimator. *The Econometrics Journal*, 25(3):531–553, 2022.

[47] Valentina Paredes. Mixed but not scrambled: Gender gaps in coed schools with single-sex classrooms. *Journal of Research on Educational Effectiveness*, 15(2):330–366, 2022.

[48] Andrew M Penner and Marcel Paret. Gender differences in mathematics achievement: Exploring the early grades and the extremes. *Social Science Research*, 37(1):239–253, 2008.

[49] Joe Relich. Gender, self-concept and teachers of mathematics: Effects on attitudes to teaching and learning. *Educational studies in mathematics*, 30(2):179–195, 1996.

[50] Joseph D Relich, John Conroy, and Vicki Webber. *Gender differences in self-concept: Male and female mathematics teachers as role models*. NSW Department of School Education and University of Western Sydney, Nepean, 1991.

[51] Lucy W Sells. Mathematics: The invisible filter. *Engineering Education*, 70(4):340–41, 1980.

[52] James H Stock and Motohiro Yogo. Testing for weak instruments in linear IV regression, 2002.

[53] Helen MG Watt, Janet S Hyde, Jennifer Petersen, Zoe A Morris, Christopher S Rozek, and Judith M Harackiewicz. Mathematics—a critical filter for stem-related career choices? a longitudinal examination among australian and us adolescents. *Sex Roles*, 77:254–271, 2017.

[54] Joseph Workman, Paul T von Hippel, and Joseph Merry. Findings on summer learning loss often fail to replicate, even in recent data. *Sociological Science*, 10:251–285, 2023.

Appendix

Data Appendix

Variables - *Repères* and APAE

Table 7: Variables APAE1D

code_indicateur	code_modalité	variable
F1D_IND_NAT_503	Effd-El-Tot	Total enrollment
	Effd-El-CP	First grade enrollment
F1D_IND_NAT_502	Nb_classes_el	Total number of classes
	Nb_classes_CP	First grade classes
F1D_IND_NAT_493	EC_ele_Etab	Average number of students per class
F1D_IND_NAT_489	Pc_filles_Etab	Proportion of students - girls
F1D_IND_NAT_528	ava_cp_Etab	Proportion of first grader - younger than norm
F1D_IND_NAT_506	ret_cp_Etab	Proportion of first grader - older than norm
F1D_IND_NAT_690	IPS_Etab	School IPS
F1D_IND_NAT_490	PCS_1_Etab	Proportion of students - Socio professional category 1
	PCS_2_Etab	Proportion of students - Socio professional category 2
	PCS_3_Etab	Proportion of students - Socio professional category 3
	PCS_4_Etab	Proportion of students - Socio professional category 4
	PCS_NR_Etab	Proportion of students - Socio professional category missing
F1D_IND_NAT_547	Rev_Med_Etab	Median income near school
F1D_IND_NAT_509	Ens_eff	Total number of teachers
F1D_IND_NAT_510	Ens_etp	Full-time equivalent number of teachers
F1D_IND_NAT_516	Femmes_Etab	Proportion of teachers - women
F1D_IND_NAT_518	Agemoy_Etab	Average age of teachers
F1D_IND_NAT_519	Prop_35_Etab	Proportion of teachers - below 35 y.o.
	Prop_3550_Etab	Proportion of teachers - between 35 and 50 y.o.
	Prop_50_Etab	Proportion of teachers - above 50 y.o.
F1D_IND_NAT_520	Ancmoy_Etab	Average seniority of teachers
F1D_IND_NAT_521	Prop2_Etab	Proportion of teachers - less than 2 years of seniority
	Prop25_Etab	Proportion of teachers - between 2 and 5 years of seniority
	Prop58_Etab	Proportion of teachers - between 5 and 8 years of seniority
	Prop8_Etab	Proportion of teachers - above 8 years of seniority

Table 8: *Évaluations Repères CP–CE1* — Literacy Subskills

	First grade — September					First grade — January					Second grade — September				
	18/19	19/20	20/21	21/22	22/23	18/19	19/20	20/21	21/22	22/23	18/19	19/20	20/21	21/22	22/23
— *From speaking to writing* —															
Writing syllables dictated by the teacher.						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Writing words dictated by the teacher.						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
— *Letter recognition* —															
Knowing the name of letters and the sound they make.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓					
Recognizing the different written forms of a letter.	✓	✓	✓	✓	✓										
Recognizing letters among signs.	✓														
Comparing sequences of letters.	✓	✓	✓	✓							✓				
— *Phonology* —															
Manipulating phonemes.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓					
Manipulating syllables.	✓	✓	✓	✓	✓										
— *Oral comprehension* —															
Understanding spoken words.	✓	✓	✓	✓	✓						✓	✓	✓	✓	✓
Understanding spoken sentences.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Understanding spoken texts.	✓	✓	✓	✓	✓										
— *Reading and writing comprehension* —															
Reading words aloud.						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Reading made-up words aloud.						✓									
Reading a text aloud.						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Understanding sentences read on your own.							✓	✓	✓	✓		✓	✓	✓	✓
Understanding texts read on your own.											✓	✓	✓	✓	✓
Answering questions read on your own.											✓				

Table 9: *Évaluations Repères CP–CE1* — Numeracy Subskills

	First grade — September					First grade — January					Second grade — September				
	18/19	19/20	20/21	21/22	22/23	18/19	19/20	20/21	21/22	22/23	18/19	19/20	20/21	21/22	22/23
— *Reading and writing numbers* —															
Writing integers.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Reading integers.	✓	✓	✓	✓	✓						✓	✓	✓	✓	✓
Representing integers.											✓	✓	✓	✓	
Placing a number on a graduated line.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
— *Problem solving* —															
Solving problems.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
— *Calculation* —															
Adding.						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Subtracting.						✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Calculating mentally.											✓	✓	✓	✓	✓
— *Comparing numbers* —															
Counting collections.	✓	✓	✓	✓	✓										
Comparing numbers.	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓				
— *Geometry and space* —															
Reproducing assembled shapes.		✓	✓	✓	✓							✓	✓	✓	✓
Recognizing and using the notions of alignment, right angle, equality of lengths, midpoint, symmetry.											✓				

Processing

1. Identification of observations
   (a) Get ridd of observations for which INE, sex or dob is missing and cannot be recovered from other observations.
   (b) Remove observations identified by INE for which there are too many different dob or sex. *Allow for one mistake, for instance on dob, and replace by last observation.*
   (c) Remove INE observations that appear in several schools at the same time.
2. Item results
   (a) For INE observations that have several different results for a same item at the same time take the average of the two.
   (b) For reading items, which have no effective maximum, cap the highest value to three times the theoretical maximum score.
   (c) Only compute mean French and Maths scores for pupils for which half or more items (in French and Maths separately) have been evaluated.
3. Merge *Repères* and APAE 1D by school and school year.
4. Further restrictions on sample
   - Cross-section analysis:
     (a) Remove INE observations that are significantly younger or older than the population norm.
     (b) Remove INE observations that repeat grades or go backwards (From Second to First grade).
   - Longitudinal analysis: same as above and:
     (a) Remove INE observations for which there is any missing evaluation session (beginning middle or end) or item of interest.
     (b) Remove INE observations that switch school between the first and second test session (school switchers).

Case studies

Alternative pedagogies

Maria Montessori (1870-1952) was an Italian physician and a pioneer in the field of child development and education. She is known for developing an education method named after her. Montessori's approach emphasized a child-centered learning environment where students are given the freedom to explore and learn at their own pace. She believed that children are naturally curious and capable of learning through hands-on experiences with specially designed

educational materials. She wrote several books and articles describing her methods.

A range of practices nowadays exists under the Montessori name, but schooling generally involves:

- Mixed-age groupings with classrooms of age 3-6 years, 6-9 years and 9-12 years which fosters peer learning;
- Longer blocks of uninterrupted work time (3 hours) compared to regular French primary schools to allow students to concentrate deeply and develop focus;
- More freedom for students to choose what they want to work on, encouraging self-discipline and sense of responsability, individualised learning;
- Hands-on and experience learning and emphasis on practical life skills;
- Grading and testing is discouraged, and progress is assessed by observation.

In France today, there are more than 300 Montessori schools. Almost all are private and not overseen by the State. Schooling fees are high as they do not get any subventions, and they are generally regarded as elitist, far from Montessori's initial focus on poor children.

Célestin (1896-1966) and Elise (1898-1983) Freinet were a couple of French teachers who created a pedagogy aimed at making students more active in the learning process and develop critical thinking and civicness. Célestin Freinet wrote several books describing the *techniques* associated to the Freinet pedagogy.

Some pillars and methods of the Freinet pedagogy include:

- Learning by experimenting, field trips so that students can learn from their own curiosity and make observations about their local environments;
- Autonomy and personalized workplan defined by the students;
- Fostering cooperation among students, group works such as exposes;
- Democratic decision-making in the classroom, in which the teacher is not the sole decision maker;
- Communication and self expression, creation of classroom and school newspapers.

In France, and although I do not have precise numbers, the number of schools that apply the Freinet pedagogy is small, but judging from my sample, several are public.

To identify both type of schools in my data I select schools with names that either contain "Freinet" or "Montessori". Note that most Montessori schools in France are private, but not

overseen by the State which means I unfortunately do not have student observations. The misclassification risk in the case of pedagogies is high, as for instance schools can also be named after the people that developped the methods without applying them.

Religious private schools

In France, private schools are predominantly religious and Catholic, and most overseen by the State. There are some Jewish schools, that are also most of the time in contract with the French ministry of Education. There are however very few Muslim schools, and even less overseen by the State. Teachings generally include some form of non-compulsory religious education, and apart from having to follow the official curriculum, private religious schools have more flexibility in terms of organisation and pedagogy.

To identify Catholic schools, I filter school names that contain "Saint(e)" or "Notre dame". For Jewish schools, I look for relevant schools networks names (for instance "Ozar Hatorah") and other keywords based on a sample of jewish schools I have found online. There are only two Muslim primary schools overseen by the State, therefore I select them directly. Considering the resulting sample, I probably mainly misclassify private schools by failing to classify them as Catholic.

Sample statistics

Students

Table 10: Students - Main sample of analysis

	Mean/N‡	SD/%‡	Min	Median	Max	Missing (%)
Individual characteristics						
Sex	0.49	0.50	0.00	0.00	1.00	0
Age (months)	75.54	2.86	60.00	75.00	104.00	0
Cohort‡						
2018-2019	635393	22.41				
2019-2020	722030	25.47				
2020-2021	741732	26.16				
2021-2022	735835	25.96				
Attended schools						
Type‡						
Private	355480	12.54				
Public	2001960	70.62				
Public - REP	290750	10.26				
Public - REP+	186060	6.56				
Enrollment & Resources						
Enrollment	222.12	117.57	2.00	207.00	1092.00	0
First grade enrollment	38.79	21.86	1.00	35.00	197.00	0
First grade classes	2.30	1.29	1.00	2.00	14.00	0
Students/class	22.62	3.61	2.00	23.50	165.00	0
Teacher/student	0.05	0.01	0.01	0.04	0.25	0
Multilevel first grade‡						
No	1362637	48.06				
Yes	608279	21.46				
Possibly	854384	30.14				
Student composition						
Younger†	0.01	0.02	0.00	0.00	0.69	1
Older†	0.03	0.04	0.00	0.02	1.00	1
Female†	0.49	0.04	0.00	0.49	1.00	0
IPS	102.79	18.36	48.75	103.50	157.60	5
Very High SES†	0.22	0.17	0.00	0.19	1.00	5
High SES†	0.12	0.06	0.00	0.12	0.80	5
Medium SES†	0.26	0.09	0.00	0.25	1.00	5
Low SES†	0.34	0.18	0.00	0.32	1.00	5
Median income	21 887.72	6340.32	1830.00	21 400.00	69 080.00	54
Teacher composition						
N	11.04	5.92	1.00	10.00	41.00	1
FTE	10.20	5.33	0.50	9.33	40.00	0
Female†	0.86	0.13	0.00	0.88	1.00	0
Age (years)	42.06	4.42	22.50	42.30	64.00	0
Seniority within school	6.44	2.90	0.00	6.20	37.10	0

Notes: var† indicates a proportion; var‡ indicates a categorical variable; Multilevel first grade "possibly" indicates that there are both multilevel and regular first grade classes in school; Younger and Older variables refer to students that begin primary school younger/older than the norm.

Table 11: Students by gender - Main sample of analysis

	Boys		Girls			
	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡	Diff. in Means	SE
Individual characteristics						
Age (months)	75.59	2.93	75.49	2.78	-0.10***	0.00
Cohort‡						
2018-2019	323297	22.5	312096	22.3		
2019-2020	366182	25.5	355848	25.5		
2020-2021	374775	26.1	366957	26.3		
2021-2022	373549	26.0	362286	25.9		
Attended schools						
Type‡						
Private	181427	12.6	174053	12.5		
Public	1015332	70.6	986628	70.6		
Public - REP	146910	10.2	143840	10.3		
Public - REP+	93771	6.5	92289	6.6		
Enrollment & Resources						
Enrollment	222.00	117.49	222.24	117.66	0.24+	0.14
First grade enrollment	38.74	21.83	38.83	21.90	0.08**	0.03
First grade classes	2.30	1.29	2.30	1.29	0.00**	0.00
Students/class	22.62	3.60	22.61	3.63	-0.01*	0.00
Teacher/student	0.05	0.01	0.05	0.01	0.00+	0.00
Multilevel first grade‡						
No	689631	48.0	673006	48.2		
Yes	309420	21.5	298859	21.4		
Possibly	433799	30.2	420585	30.1		
Student composition						
Younger†	0.01	0.02	0.01	0.02	0.00***	0.00
Older†	0.03	0.04	0.03	0.04	0.00***	0.00
Female†	0.49	0.04	0.49	0.04	0.01***	0.00
IPS	102.83	18.34	102.74	18.39	-0.09***	0.02
Very High SES†	0.23	0.17	0.22	0.17	0.00**	0.00
High SES†	0.12	0.06	0.12	0.06	0.00*	0.00
Medium SES†	0.26	0.09	0.26	0.09	0.00*	0.00
Low SES†	0.34	0.18	0.34	0.18	0.00**	0.00
Median income	21905.78	6344.76	21869.07	6335.67	-36.71***	11.05
Teacher composition						
N	11.03	5.91	11.04	5.92	0.02*	0.01
FTE	10.19	5.33	10.21	5.34	0.02**	0.01
Female†	0.86	0.13	0.86	0.13	0.00*	0.00
Age (years)	42.06	4.43	42.06	4.42	0.00	0.01
Seniority within school	6.44	2.90	6.44	2.90	-0.01	0.00
Students number	1437803		1397187			

Notes: Signif. code: +=.1, *=.05, **=.01, ***=0.001; var† indicates a proportion; var‡ indicates a categorical variable; Multilevel first grade "possibly" indicates that there are both multilevel and regular first grade classes in school; Younger and Older variables refer to students that begin primary school younger/older than the norm.

Table 12: Students by teacher gender - Teacher gender sample

	Male teacher		Female teacher			
	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡	Diff. in Means	SE
Individual characteristics						
Sex	0.49	0.50	0.49	0.50	0.00	0.00
Age (months)	75.58	2.86	75.52	2.84	-0.06***	0.01
Cohort‡						
2018-2019	18565	24.5	273921	22.9		
2019-2020	19277	25.4	304848	25.5		
2020-2021	18976	25.0	308571	25.8		
2021-2022	18975	25.0	308780	25.8		
Initial maths score	0.05	0.92	0.12	0.83	0.07***	0.00
Initial French score	0.02	0.95	0.12	0.88	0.11***	0.00
Attended schools						
Type‡						
Private	1619	2.1	150446	12.6		
Public	56634	74.7	892075	74.6		
Public - REP	10221	13.5	97522	8.2		
Public - REP+	7319	9.7	56061	4.7		
Enrollment & Resources						
Enrollment	220.13	102.88	204.78	111.40	-15.35***	0.39
First grade enrollment	40.85	20.42	35.98	20.80	-4.87***	0.08
First grade classes	2.48	1.32	2.11	1.13	-0.37***	0.00
Students/class	22.12	3.61	22.78	3.44	0.66***	0.01
Teacher/student	0.05	0.01	0.05	0.01	0.00***	0.00
Multilevel first grade‡						
No	38973	51.4	536338	44.8		
Yes	11712	15.5	307961	25.7		
Possibly	25108	33.1	351657	29.4		
Student composition						
Younger†	0.01	0.02	0.01	0.02	0.00***	0.00
Older†	0.03	0.04	0.03	0.04	0.00***	0.00
Female†	0.49	0.04	0.49	0.05	0.00***	0.00
IPS	98.50	17.95	104.06	17.78	5.57***	0.07
Very High SES†	0.19	0.14	0.23	0.17	0.04***	0.00
High SES†	0.12	0.05	0.13	0.06	0.01***	0.00
Medium SES†	0.25	0.08	0.26	0.09	0.01***	0.00
Low SES†	0.38	0.18	0.33	0.17	-0.05***	0.00
Median income	20916.45	6084.47	22160.86	6183.73	1244.42***	32.64
Teacher composition						
N	11.39	5.45	9.98	5.52	-1.42***	0.02
FTE	10.44	5.00	9.25	4.90	-1.19***	0.02
Female†	0.73	0.18	0.92	0.11	0.19***	0.00
Age (years)	42.13	4.47	42.01	4.42	-0.12***	0.02
Seniority within school	6.55	2.95	6.57	2.94	0.01	0.01
Students number	75793		1196120			

Notes: Signif. code: +=.1, *=.05, **=.01, ***=0.001; var† indicates a proportion; var‡ indicates a categorical variable; Multilevel first grade "possibly" indicates that there are both multilevel and regular first grade classes in school; Younger and Older variables refer to students that begin primary school younger/older than the norm.

Table 13: Students in schools with alternative pedagogy

	Freinet		Montessori		Rest	
	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡
Individual characteristics						
Sex	0.49	0.5	0.48	0.5	0.49	0.5
Age (months)	75	2.8	76	2.7	76	2.9
Cohort‡						
2018-2019	473	20.8	97	19.5	634823	22.4
2019-2020	565	24.8	130	26.2	721335	25.5
2020-2021	620	27.3	131	26.4	740981	26.2
2021-2022	616	27.1	139	28.0	735080	26.0
Initial maths score	-0.059	0.98	-0.12	1.2	0.08	0.88
Initial French score	-0.13	1	-0.16	1.1	0.073	0.93
Attended schools						
Type‡						
Private	0	0.0	199	40.0	355281	12.5
Public	1923	84.6	138	27.8	1999899	70.6
Public - REP	261	11.5	160	32.2	290329	10.3
Public - REP+	90	4.0	0	0.0	185970	6.6
Enrollment & Resources						
Enrollment	203	79	245	99	222	118
First grade enrollment	39	19	31	11	39	22
Students/class	23	3.1	22	5	23	3.6
Teacher/Student	0.046	0.008	0.048	0.015	0.047	0.009
Multilevel first grade‡						
No	963	42.3	292	58.8	1361382	48.1
Yes	510	22.4	141	28.4	607628	21.5
Possibly	801	35.2	64	12.9	853519	30.1
Student composition						
Younger†	0.0076	0.015	0.0014	0.014	0.0061	0.018
Older†	0.031	0.034	0.031	0.033	0.028	0.038
Female†	0.49	0.04	0.48	0.037	0.49	0.043
IPS	97	16	99	34	103	18
Very High SES†	0.17	0.12	0.29	0.33	0.22	0.17
High SES†	0.13	0.057	0.083	0.041	0.12	0.058
Medium SES†	0.24	0.062	0.16	0.045	0.26	0.086
Low SES†	0.41	0.17	0.37	0.25	0.34	0.18
Teacher composition						
Female†	0.79	0.14	0.81	0.24	0.86	0.13
Age (years)	40	5.4	43	4.4	42	4.4
Seniority within school	5.7	3	7.8	3.5	6.4	2.9
Students number	2274		497		2832219	
Schools number	24		7		33127	

Notes: var† indicates a proportion; var‡ indicates a categorical variable; Multilevel first grade "possibly" indicates that there are both multilevel and regular first grade classes in school; Younger and Older variables refer to students that begin primary school younger/older than the norm.

Table 14: Students in private religious schools

	Catholic		Jewish		Muslim		Other Private		Public	
	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡	Mean	SD/%‡	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡
Individual characteristics										
Sex	0.49	0.5	0.54	0.5	0.51	0.5	0.49	0.5	0.49	0.5
Age (months)	75	2.8	75	3.5	75	2.7	75	2.9	76	2.9
Cohort‡										
2018-2019	44821	21.0	410	14.3	26	20.3	29117	20.9	560941	22.6
2019-2020	53833	25.3	817	28.5	25	19.5	34497	24.8	632730	25.5
2020-2021	58052	27.2	897	31.2	25	19.5	38188	27.4	644246	26.0
2021-2022	56489	26.5	747	26.0	52	40.6	37484	26.9	640853	25.9
Initial maths score	0.26	0.71	0.14	0.89	0.064	0.76	0.26	0.71	0.054	0.9
Initial French score	0.32	0.73	0.092	0.87	0.25	0.69	0.31	0.76	0.038	0.94
Attended schools										
Enrollment & Resources										
Enrollment	276	153	374	139	186	64	298	171	213	107
First grade enrollment	36	21	53	23	27	0.63	39	24	39	22
Students/class	25	2.8	22	3.4	27	0.25	25	4.7	22	3.4
Teacher/student	0.042	0.0054	0.04	0.0089	0.037	0.00098	0.042	0.0066	0.048	0.0091
Multilevel first grade‡										
No	103165	48.4	2747	95.7	128	100.0	75292	54.1	1181305	47.7
Yes	62003	29.1	51	1.8	0	0.0	35085	25.2	511140	20.6
Possibly	47669	22.4	73	2.5	0	0.0	28588	20.5	778054	31.4
Student composition										
Younger†	0.01	0.024	0.067	0.058	0	0	0.012	0.027	0.0053	0.016
Older†	0.024	0.038	0.029	0.036	0.014	0.029	0.024	0.039	0.029	0.037
Female†	0.49	0.05	0.52	0.13	0.49	0.022	0.49	0.046	0.49	0.042
IPS	116	15	125	11	107	15	118	16	101	18
Very High SES†	0.35	0.2	0.46	0.16	0.26	0.17	0.38	0.22	0.2	0.15
High SES†	0.15	0.063	0.081	0.045	0.13	0.035	0.14	0.066	0.12	0.056
Medium SES†	0.28	0.1	0.36	0.13	0.33	0.16	0.27	0.11	0.25	0.082
Low SES†	0.2	0.14	0.06	0.045	0.22	0.028	0.18	0.14	0.36	0.17
Teacher composition										
Female†	0.92	0.087	0.95	0.075	0.82	0.088	0.93	0.091	0.85	0.13
Age (years)	44	3.8	45	3.2	36	1.2	44	3.8	42	4.4
Seniority within school	8	3	8.7	3	4.3	1.6	8.2	3.1	6.2	2.8
Students number	213195		2871		128		139286		2478770	
Schools number	2749		32		2		1773		28573	

Notes: var† indicates a proportion; var‡ indicates a categorical variable; Multilevel first grade "possibly" indicates that there are both multilevel and regular first grade classes in school; Younger and Older variables refer to students that begin primary school younger/older than the norm.

Schools

Table 15: Schools characteristics

	Mean/N‡	SD/%‡	Min	Median	Max	Missing (%)
Year‡						
2018-2019	28019	22.78				
2019-2020	31296	25.45				
2020-2021	31859	25.91				
2021-2022	31803	25.86				
Type‡						
Private	16388	13.33				
Public	106532	86.63				
Priority‡						
None	108830	88.50				
REP	9051	7.36				
REP+	5039	4.10				
Enrollment & Resources						
Enrollment	157.27	103.40	2.00	135.00	1092.00	0
First grade enrollment	25.60	18.51	1.00	21.00	197.00	0
First grade classes	1.71	1.01	1.00	1.00	14.00	0
Students/class	21.78	3.87	2.00	22.50	165.00	0
Teacher/student	0.05	0.01	0.01	0.05	0.25	0
Multilevel first grade‡						
No	42727	34.74				
Yes	55802	45.38				
Possibly	24027	19.54				
Student composition						
Younger†	0.01	0.02	0.00	0.00	0.69	5
Older†	0.03	0.05	0.00	0.00	1.00	5
Female†	0.49	0.06	0.00	0.49	1.00	0
IPS	102.65	16.45	48.75	103.03	157.60	9
Very High SES†	0.20	0.15	0.00	0.17	1.00	9
High SES†	0.13	0.06	0.00	0.12	0.80	9
Medium SES†	0.27	0.10	0.00	0.26	1.00	9
Low SES†	0.34	0.17	0.00	0.33	1.00	9
Median income	21 609.21	5341.55	1830.00	21 110.00	69 080.00	53
Teacher composition						
Number	7.91	5.30	1.00	7.00	41.00	3
FTE	7.34	4.71	0.50	6.00	40.00	0
Female†	0.87	0.15	0.00	0.89	1.00	0
Age (years)	41.84	4.89	22.50	42.00	64.00	0
Seniority within school	6.48	3.35	0.00	6.20	37.10	0
Observed students						
Number	23.05	16.90	1.00	19.00	192.00	0
Female†	0.49	0.15	0.00	0.50	1.00	0
Age (months)	75.53	0.93	62.00	75.50	101.00	0

Notes: apart from last line, one school observation corresponds to a school x year; var† indicates a proportion; var‡ indicates a categorical variable; Multilevel first grade "possibly" indicates that there are both multilevel and regular first grade classes in school; Younger and Older variables refer to students that begin primary school younger/older than the norm.

Table 16: Schools characteristics by type of school

	Private		Public		Public - REP		Public - REP+	
	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡
Year‡								
2018-2019	3430	20.9	21506	23.3	1963	21.7	1107	22.0
2019-2020	4198	25.6	23455	25.4	2334	25.8	1294	25.7
2020-2021	4399	26.8	23749	25.7	2378	26.3	1314	26.1
2021-2022	4361	26.6	23732	25.7	2376	26.3	1324	26.3
Enrollment & Resources								
Enrollment	191.80	132.93	142.92	93.60	199.80	97.11	232.17	95.63
First grade enrollment	24.33	17.80	23.84	17.55	36.21	20.07	43.15	19.52
First grade classes	1.47	0.70	1.53	0.73	3.01	1.51	3.54	1.48
Students/class	23.77	4.68	22.18	3.29	16.95	2.12	16.66	1.70
Teacher/student	0.04	0.01	0.05	0.01	0.06	0.01	0.06	0.01
Multilevel first grade‡								
No	5733	35.0	26004	28.1	6910	76.3	4080	81.0
Yes	8465	51.7	46507	50.3	733	8.1	97	1.9
Possibly	2177	13.3	19642	21.2	1408	15.6	800	15.9
Student composition								
Younger†	0.01	0.03	0.01	0.02	0.01	0.02	0.01	0.02
Older†	0.03	0.05	0.03	0.05	0.04	0.05	0.04	0.04
Female†	0.48	0.06	0.49	0.06	0.49	0.04	0.49	0.04
IPS	111.74	14.66	104.84	13.78	82.19	10.86	72.64	9.03
Very High SES†	0.29	0.19	0.21	0.14	0.08	0.07	0.05	0.04
High SES†	0.14	0.07	0.13	0.06	0.08	0.04	0.06	0.04
Medium SES†	0.30	0.11	0.28	0.10	0.23	0.08	0.18	0.07
Low SES†	0.24	0.14	0.33	0.14	0.52	0.14	0.59	0.14
Median income	21838.88	5518.39	22449.84	4777.30	16451.04	4285.49	12902.23	4532.99
Teacher composition								
Number	8.59	5.75	6.89	4.45	12.73	5.87	15.23	5.80
FTE	8.05	5.05	6.38	3.86	11.90	5.37	14.38	5.34
Female†	0.92	0.11	0.87	0.16	0.83	0.13	0.80	0.13
Age (years)	43.31	4.54	41.97	4.92	39.28	4.12	39.32	4.02
Seniority within school	7.55	3.40	6.48	3.41	5.28	2.47	5.27	2.26
Observed students								
Number	21.69	16.72	21.66	16.09	32.12	18.36	36.92	17.85
Female†	0.49	0.17	0.49	0.15	0.50	0.12	0.50	0.10
Age (months)	75.49	1.01	75.51	0.94	75.65	0.79	75.68	0.75
Schools number	16388		92442		9051		5039	

Notes: apart from last line, one school observation corresponds to a school x year; var† indicates a proportion; var‡ indicates a categorical variable; Multilevel first grade "possibly" indicates that there are both multilevel and regular first grade classes in school; Younger and Older variables refer to students that begin primary school younger/older than the norm.

Table 17: Schools characteristics - alternative pedagogies

	Freinet		Montessori		Rest	
	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡
Year‡						
2018-2019	17	19.3	5	20.8	27997	22.8
2019-2020	23	26.1	6	25.0	31267	25.4
2020-2021	24	27.3	6	25.0	31829	25.9
2021-2022	24	27.3	7	29.2	31772	25.9
Type‡						
Public	88	100	9	37.5	106492	86.7
Private	0	0	15	62.5	16373	13.3
Priority‡						
None	76	86.4	19	79.2	108792	88.5
REP	8	9.1	5	20.8	9038	7.4
REP+	4	4.5	0	0	5035	4.1
Enrollment & Resources						
Enrollment	157	83	193	110	157	103
First grade enrollment	28	18	24	13	26	19
Students/class	22	3.4	24	5.2	22	3.9
Teacher/student	0.047	0.009	0.041	0.016	0.049	0.01
Multilevel first grade‡						
No	27	30.7	10	41.7	42690	34.7
Yes	39	44.3	12	50.0	55751	45.4
Possibly	22	25.0	2	8.3	24003	19.5
Student composition						
Younger†	0.0064	0.016	0.006	0.029	0.0069	0.024
Older†	0.028	0.039	0.031	0.041	0.03	0.046
Female†	0.49	0.047	0.47	0.041	0.49	0.056
IPS	100	17	112	33	103	16
Very High SES†	0.18	0.13	0.4	0.32	0.2	0.15
High SES†	0.14	0.065	0.091	0.048	0.13	0.063
Medium SES†	0.26	0.079	0.16	0.049	0.27	0.099
Low SES†	0.37	0.17	0.27	0.24	0.34	0.17
Teacher composition						
Female†	0.81	0.15	0.77	0.33	0.87	0.15
Age (years)	39	5.9	45	6.2	42	4.9
Seniority within school	6.1	3.4	9	5.2	6.5	3.4
Observed students						
Initial maths score	0.018	0.43	0.12	0.61	0.14	0.39
Initial French score	-0.014	0.53	0.13	0.69	0.14	0.46
Age (months)	75	0.84	76	0.63	76	0.93
Female†	0.48	0.12	0.44	0.13	0.49	0.15
Schools number	24		7		33127	

Notes: apart from last line, one school observation corresponds to a school x year; var† indicates a proportion; var‡ indicates a categorical variable; Multilevel first grade "possibly" indicates that there are both multilevel and regular first grade classes in school; Younger and Older variables refer to students that begin primary school younger/older than the norm.

Table 18: Schools characteristics - private religious schools

	Catholic		Jewish		Muslim		Other Private		Public	
	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡	Mean/N‡	SD/%‡
Year‡										
2018-2019	2113	21.2	14	14.1	1	20.0	1302	20.6	24576	23.1
2019-2020	2550	25.6	28	28.3	1	20.0	1619	25.6	27083	25.4
2020-2021	2660	26.7	31	31.3	1	20.0	1707	27.0	27441	25.8
2021-2022	2632	26.4	26	26.3	2	40.0	1701	26.9	27432	25.8
Enrollment & Resources										
Enrollment	190	127	281	159	185	72	193	141	152	97
First grade enrollment	24	17	37	23	27	0.71	25	19	26	19
Students/class	24	3.5	21	4.5	27	0.29	24	6.1	21	3.6
Teacher/student	0.044	0.0079	0.045	0.017	0.037	0.0011	0.045	0.0096	0.049	0.011
Multilevel first grade‡										
No	3337	33.5	89	89.9	5	100.0	2302	36.4	36994	34.7
Yes	5235	52.6	8	8.1	0	0.0	3222	50.9	47337	44.4
Possibly	1377	13.8	2	2.0	0	0.0	798	12.6	21850	20.5
Student composition										
Younger†	0.011	0.031	0.073	0.075	0	0	0.013	0.035	0.006	0.022
Older†	0.03	0.051	0.037	0.044	0.015	0.033	0.031	0.053	0.029	0.045
Female†	0.48	0.062	0.51	0.14	0.49	0.025	0.49	0.06	0.49	0.055
IPS	111	14	123	13	107	17	112	15	101	16
Very High SES†	0.28	0.19	0.43	0.17	0.27	0.19	0.29	0.2	0.19	0.14
High SES†	0.15	0.065	0.075	0.043	0.13	0.04	0.14	0.069	0.12	0.062
Medium SES†	0.3	0.11	0.38	0.15	0.33	0.19	0.3	0.11	0.27	0.096
Low SES†	0.24	0.14	0.06	0.048	0.22	0.031	0.23	0.14	0.36	0.16
Teacher composition										
Female†	0.92	0.1	0.96	0.064	0.82	0.099	0.92	0.12	0.86	0.15
Age (years)	43	4.4	45	3.8	36	1.3	43	4.7	42	4.9
Seniority within school	7.5	3.3	7.8	3.1	4.2	1.8	7.6	3.5	6.3	3.3
Observed students										
Initial maths score	0.25	0.31	0.11	0.42	0.067	0.35	0.24	0.33	0.12	0.4
Initial French score	0.3	0.35	0.083	0.43	0.25	0.36	0.26	0.45	0.12	0.46
Age (months)	75	0.98	75	1.7	75	0.47	75	1.1	76	0.92
Female†	0.48	0.16	0.54	0.25	0.51	0.083	0.49	0.17	0.49	0.15
Schools number	2749		32		2		1773		28573	

Notes: apart from last line, one school observation corresponds to a school x year; var† indicates a proportion; var‡ indicates a categorical variable; Multilevel first grade "possibly" indicates that there are both multilevel and regular first grade classes in school; Younger and Older variables refer to students that begin primary school younger/older than the norm.

Supplementary results

French and maths

Table 19: Best subject by gender and test session

		Boys		Girls	
		N	%	N	%
Beginning	French	660913	46.6	788080	57.1
	Maths	758166	53.4	591435	42.9
	|Maths - French|	97253	6.7	196645	14.9
Middle	French	622296	45.3	702628	52.6
	Maths	750734	54.7	633115	47.4
	|Maths - French|	128436	9.4	69513	5.2
End	French	588544	41.1	823709	59.3
	Maths	841666	58.9	566664	40.7
	|Maths - French|	253122	17.8	257045	18.6

Table 20: Descriptive statistics - Individual maths advantage

	Beginning		Middle		End	
	Boys	Girls	Boys	Girls	Boys	Girls
mean	0.08	-0.08	0.06	-0.07	0.13	-0.16
min	-7.65	-8.04	-5.89	-5.66	-4.85	-5.85
max	6.00	5.99	5.84	5.63	5.26	5.52
Q10	-0.70	-0.80	-0.83	-0.93	-0.72	-0.98
Q25	-0.34	-0.46	-0.39	-0.49	-0.29	-0.56
Q50	0.05	-0.10	0.08	-0.04	0.14	-0.14
Q75	0.47	0.29	0.51	0.37	0.55	0.26
Q90	0.92	0.69	0.92	0.73	0.95	0.62
SD	0.68	0.62	0.72	0.67	0.68	0.65
Missing (%)	1.30	1.26	4.50	4.40	0.53	0.49

Note: Statistics are expressed in proportion of a standard deviation.

Table 21: Progress joint distribution: beginning - end (quartiles)

(a) Boys

maths - French	Q1	Q2	Q3	Q4
Q1	10.40	5.86	4.74	4.00
Q2	6.42	7.02	6.40	5.16
Q3	4.82	6.86	7.14	6.18
Q4	3.36	5.26	6.72	9.66

(b) Girls

maths - French	Q1	Q2	Q3	Q4
Q1	10.62	5.80	4.64	3.95
Q2	6.49	6.86	6.30	5.36
Q3	4.69	6.87	7.10	6.34
Q4	3.20	5.48	6.96	9.35

Table 22: Correlation matrix scores: Boys

	Maths (1)	Maths (2)	Maths (3)	French (1)	French (2)	French (3)	M.(2)-M.(1)	M.(3)-M.(2)	M.(3)-M.(1)	F.(2)-F.(1)	F.(3)-F.(2)	F.(3)-F.(1)
Maths (1)	1	0.686	0.666	0.742	0.597	0.602	-0.387	0.011	-0.361	-0.202	0.018	-0.185
Maths (2)	0.686	1	0.746	0.661	0.699	0.639	0.404	-0.303	0.118	0.027	-0.078	-0.036
Maths (3)	0.666	0.746	1	0.657	0.667	0.747	0.110	0.408	0.455	-0.004	0.140	0.111
French (1)	0.742	0.661	0.657	1	0.705	0.705	-0.090	0.030	-0.062	-0.410	0.014	-0.391
French (2)	0.597	0.699	0.667	0.705	1	0.805	0.134	-0.003	0.124	0.359	-0.290	0.123
French (3)	0.602	0.639	0.747	0.705	0.805	1	0.056	0.188	0.214	0.110	0.334	0.377
M.(2)-M.(1)	-0.387	0.404	0.110	-0.090	0.134	0.056	1	-0.398	0.597	0.291	-0.123	0.190
M.(3)-M.(2)	0.011	-0.303	0.408	0.030	-0.003	0.188	-0.398	1	0.498	-0.043	0.309	0.203
M.(3)-M.(1)	-0.361	0.118	0.455	-0.062	0.124	0.214	0.597	0.498	1	0.236	0.153	0.360
F.(2)-F.(1)	-0.202	0.027	-0.004	-0.410	0.359	0.110	0.291	-0.043	0.236	1	-0.392	0.676
F.(3)-F.(2)	0.018	-0.078	0.140	0.014	-0.290	0.334	-0.123	0.309	0.153	-0.392	1	0.412
F.(3)-F.(1)	-0.185	-0.036	0.111	-0.391	0.123	0.377	0.190	0.203	0.360	0.676	0.412	1

Notes: Figures indicate evaluation timing: 1 = beginning, 2 = middle, 3 = end; Differences are progress outcomes *eg* F.(3)-F.(1) denotes progress in French between the beginning and end evaluation.

Table 23: Correlation matrix scores: Girls

	Maths (1)	Maths (2)	Maths (3)	French (1)	French (2)	French (3)	M.(2)-M.(1)	M.(3)-M.(2)	M.(3)-M.(1)	F.(2)-F.(1)	F.(3)-F.(2)	F.(3)-F.(1)
Maths (1)	1	0.670	0.641	0.735	0.586	0.588	-0.331	-0.015	-0.329	-0.202	0.037	-0.165
Maths (2)	0.670	1	0.719	0.661	0.702	0.629	0.478	-0.343	0.138	0.033	-0.070	-0.024
Maths (3)	0.641	0.719	1	0.651	0.668	0.744	0.156	0.407	0.513	0.007	0.159	0.138
French (1)	0.735	0.661	0.651	1	0.694	0.692	-0.026	0.009	-0.019	-0.413	0.038	-0.369
French (2)	0.586	0.702	0.668	0.694	1	0.796	0.198	-0.018	0.170	0.369	-0.263	0.149
French (3)	0.588	0.629	0.744	0.692	0.796	1	0.106	0.175	0.257	0.115	0.376	0.416
M.(2)-M.(1)	-0.331	0.478	0.156	-0.026	0.198	0.106	1	-0.418	0.560	0.284	-0.134	0.168
M.(3)-M.(2)	-0.015	-0.343	0.407	0.009	-0.018	0.175	-0.418	1	0.519	-0.034	0.308	0.214
M.(3)-M.(1)	-0.329	0.138	0.513	-0.019	0.170	0.257	0.560	0.519	1	0.235	0.154	0.357
F.(2)-F.(1)	-0.202	0.033	0.007	-0.413	0.369	0.115	0.284	-0.034	0.235	1	-0.382	0.667
F.(3)-F.(2)	0.037	-0.070	0.159	0.038	-0.263	0.376	-0.134	0.308	0.154	-0.381	1	0.434
F.(3)-F.(1)	-0.165	-0.024	0.138	-0.369	0.149	0.416	0.168	0.214	0.357	0.668	0.434	1

Notes: Figures indicate evaluation timing: 1 = beginning, 2 = middle, 3 = end; Differences are progress outcomes *eg* F.(3)-F.(1) denotes progress in French between the beginning and end evaluation.

Gender gaps

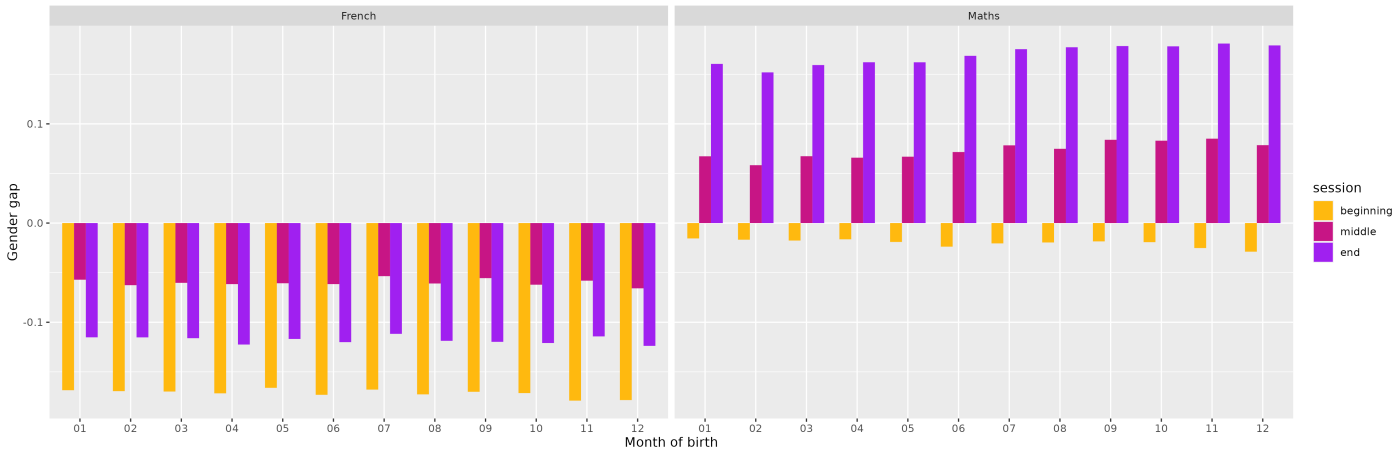


Figure 11: Gender gaps by students' month of birth

Note: Gender gaps are expressed in proportion of a standard deviation, and computed as $\overline{score}^t_{boys} - \overline{score}^t_{girls}$

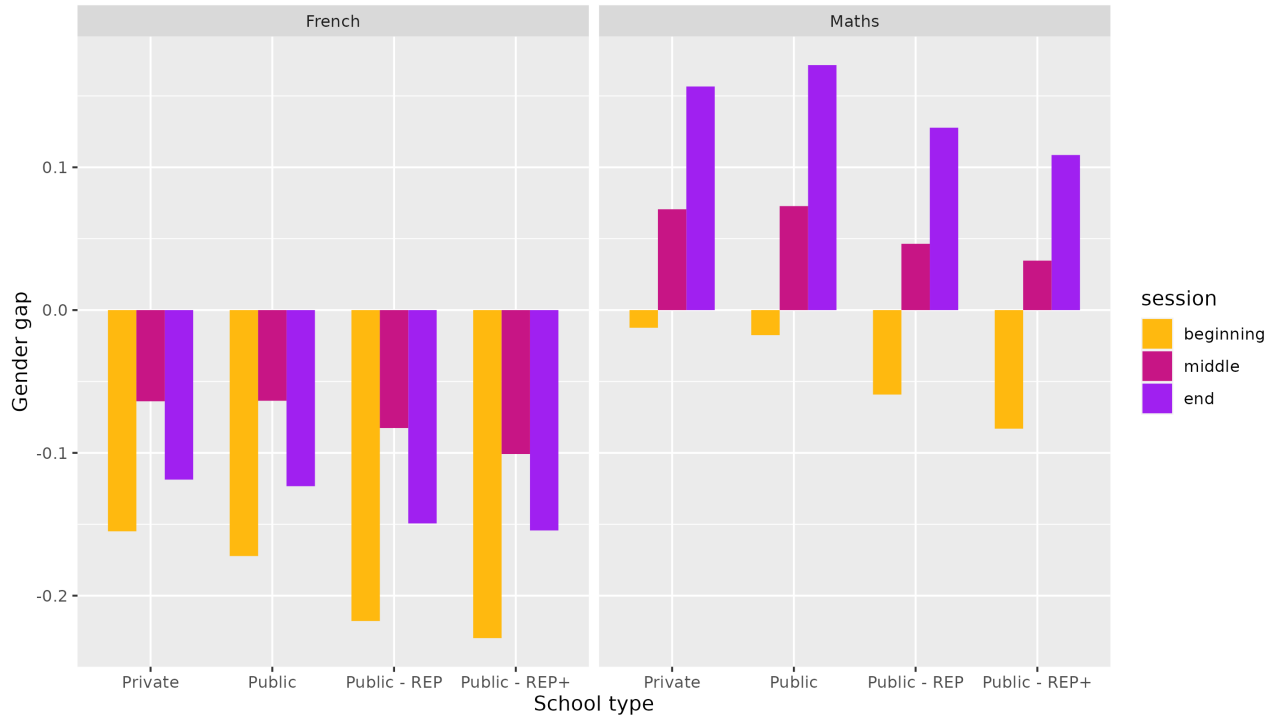


Figure 12: Gender gaps by type of school

Note: Gender gaps are expressed in proportion of a standard deviation, and computed as $\overline{score}^t_{boys} - \overline{score}^t_{girls}$

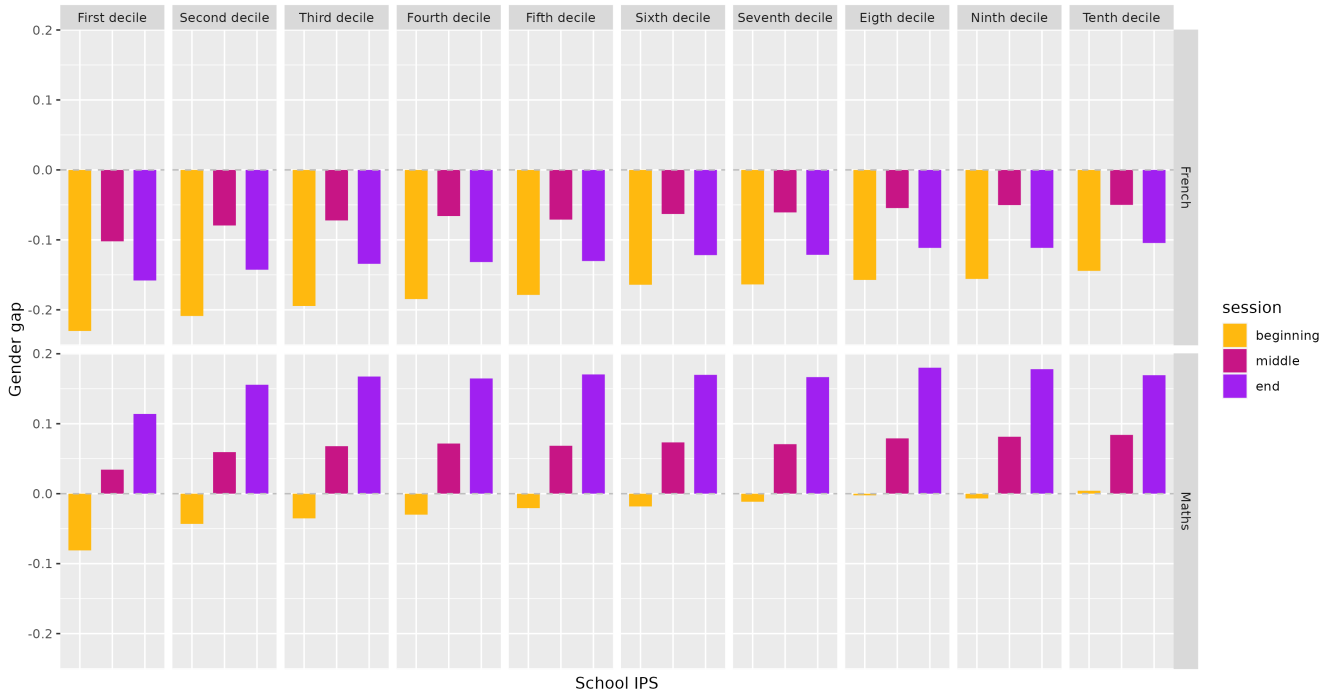
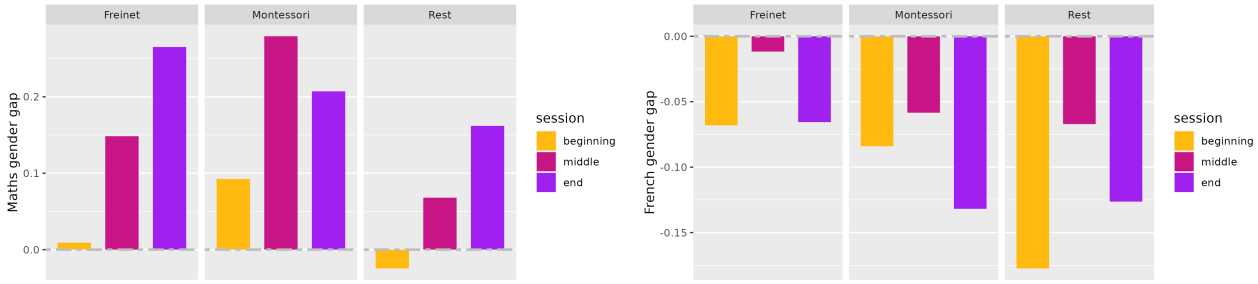


Figure 13: Gender gaps by school IPS decile

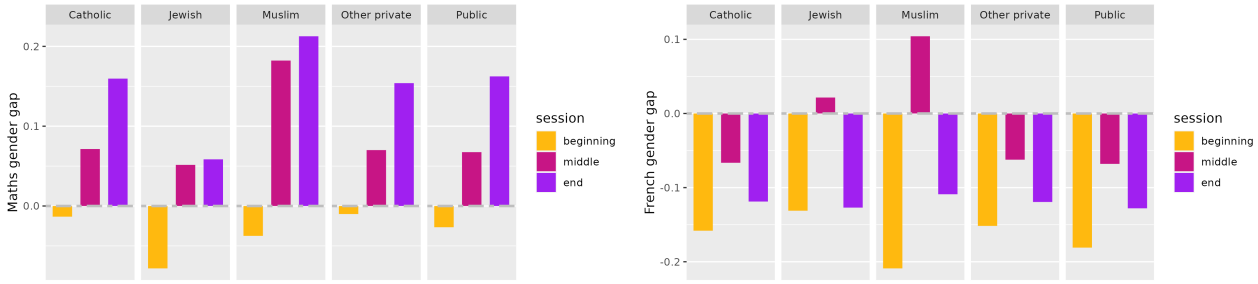
Note: Gender gaps are expressed in proportion of a standard deviation, and computed as $\overline{score}^t_{boys} - \overline{score}^t_{girls}$



(a) Maths (b) French

Figure 14: Gender gaps in schools with alternative pedagogies

Note: Gender gaps are expressed in proportion of a standard deviation, and computed as $\overline{score}^t_{boys} - \overline{score}^t_{girls}$



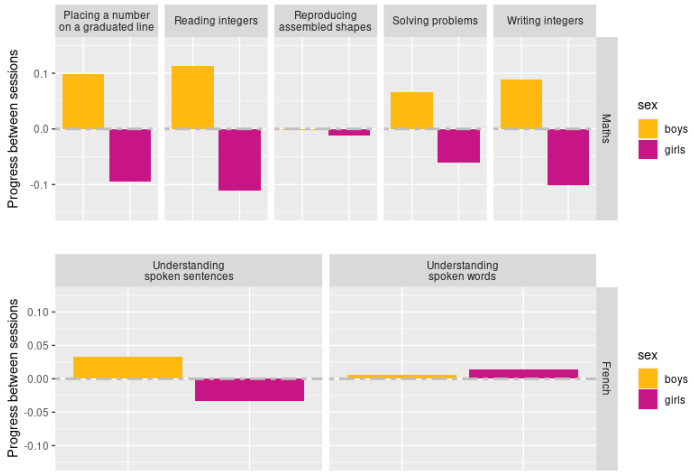
(a) Maths (b) French

Figure 15: Gender gaps in private religious schools

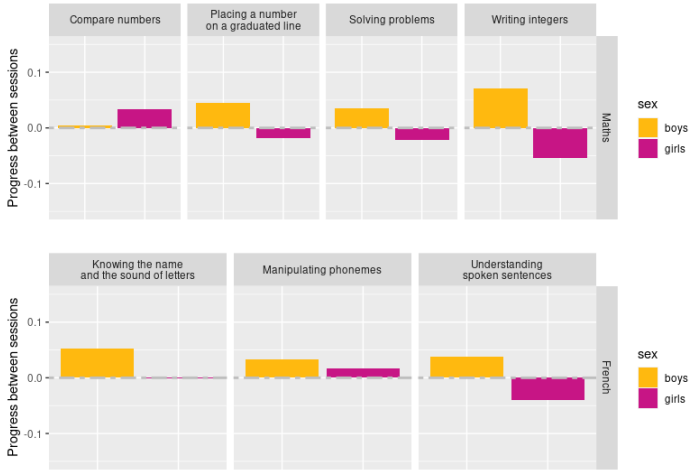
Note: Gender gaps are expressed in proportion of a standard deviation, and computed as $\overline{score}^t_{boys} - \overline{score}^t_{girls}$

Figure 16: Progress for returning items - averaged

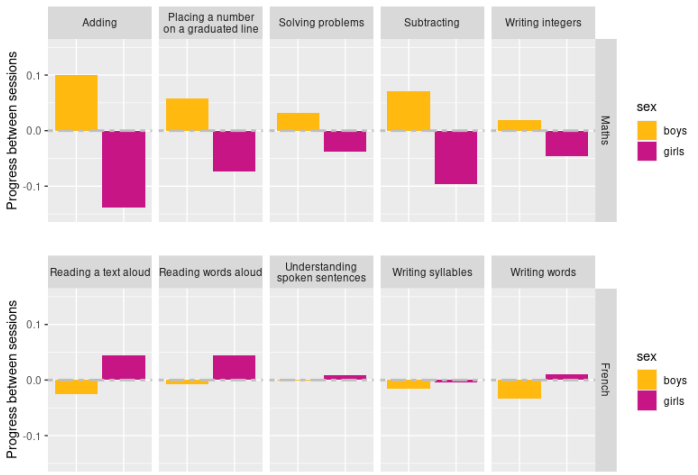
Note: Progress is expressed in proportion of a standard deviation, and computed as $\overline{score}^{t+1} - \overline{score}^t$ using the same items in $t+1$ and t.



(a) beginning - end



(b) beginning - middle



(c) middle - end

Figure 17: Progress for returning items - items

Note: Progress is expressed in proportion of a standard deviation, and computed as $\overline{score}^{t+1} - \overline{score}^{t}$.

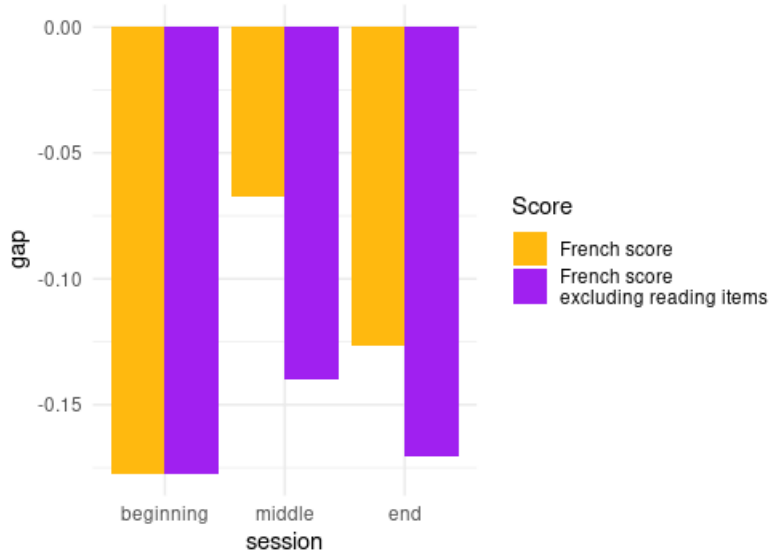
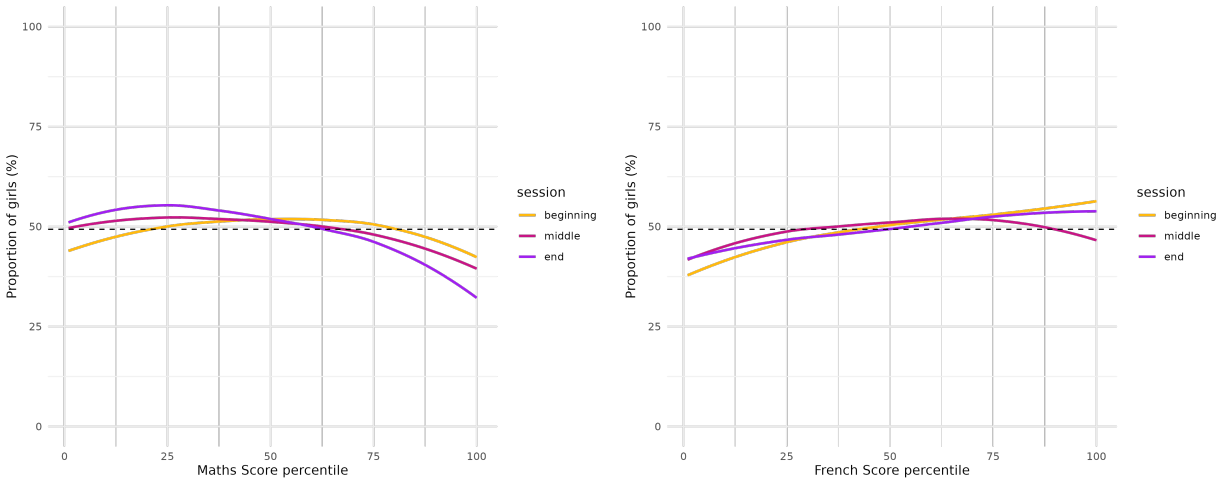


Figure 18: French gender gap - with and without reading items

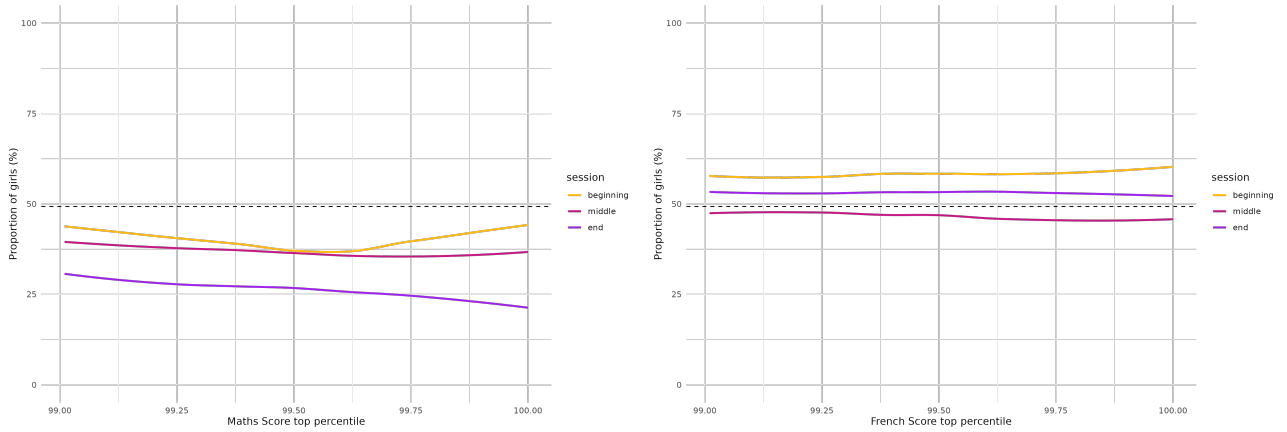
Note: Gender gaps are expressed in proportion of a standard deviation, and computed as $(\overline{score}^{t+1}_{boys} - \overline{score}^{t}_{boys}) - (\overline{score}^{t+1}_{girls} - \overline{score}^{t}_{girls})$



(a) Maths (b) French

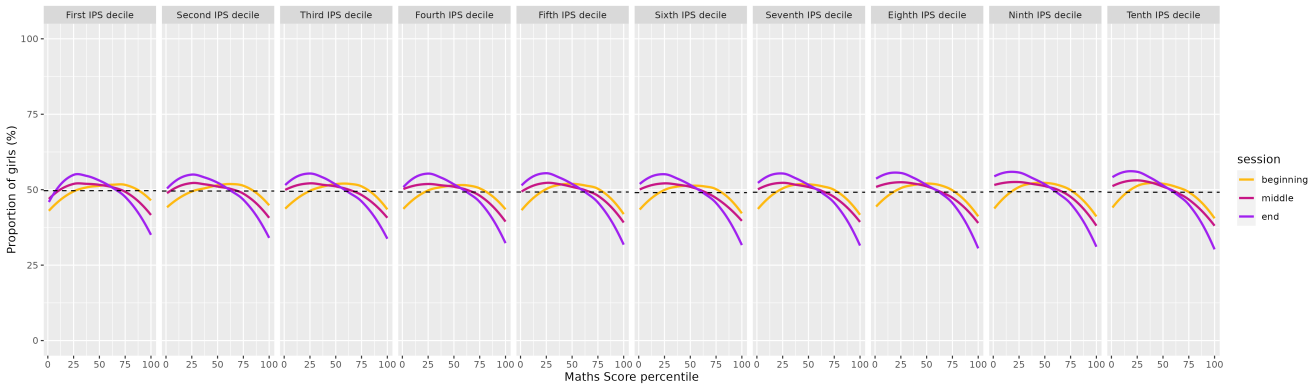
Figure 19: Proportion of girls in the percentile distribution

Note: The dotted line represents the proportion of girls in the sample considered. All curves are smoothed using LOESS.

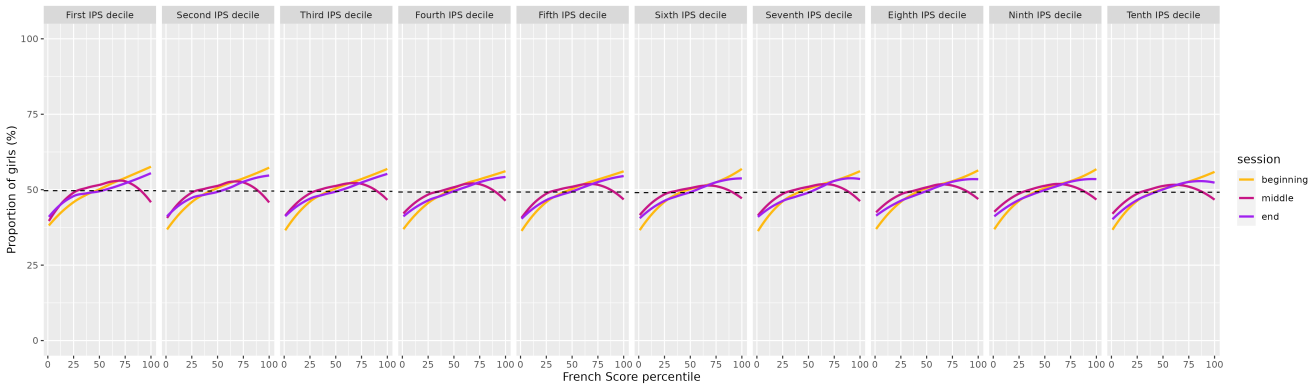


(a) Maths (b) French

Figure 20: Proportion of girls in the top percentile (zoom)
Note: The dotted line represents the proportion of girls in the sample considered. All curves are smoothed using LOESS.

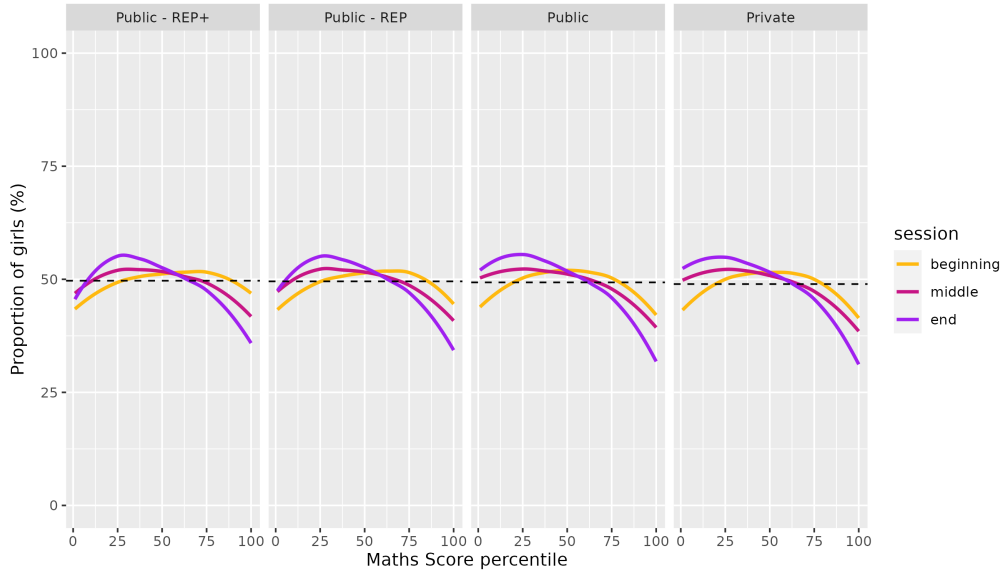


(a) Maths

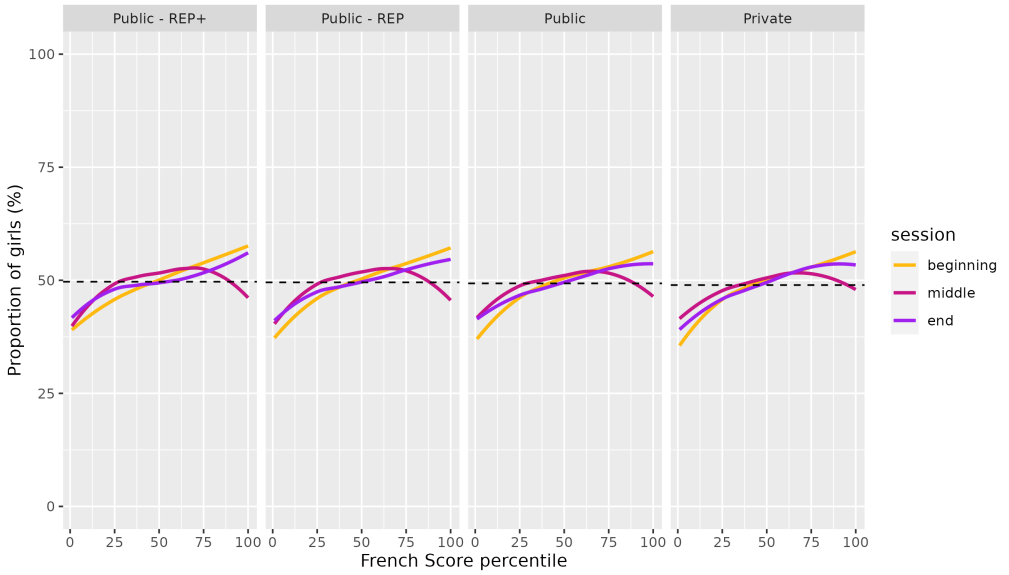


(b) French

Figure 21: Proportion of girls in the percentile distribution by school IPS decile
Note: The dotted line represents the proportion of girls in the sample considered. All curves are smoothed using LOESS.



(a) Maths



(b) French

Figure 22: Proportion of girls in the percentile distribution by school type

Note: The dotted line represents the proportion of girls in the sample considered. All curves are smoothed using LOESS.

Initial scores

Table 24: Maths progress (beginning - middle) and initial scores

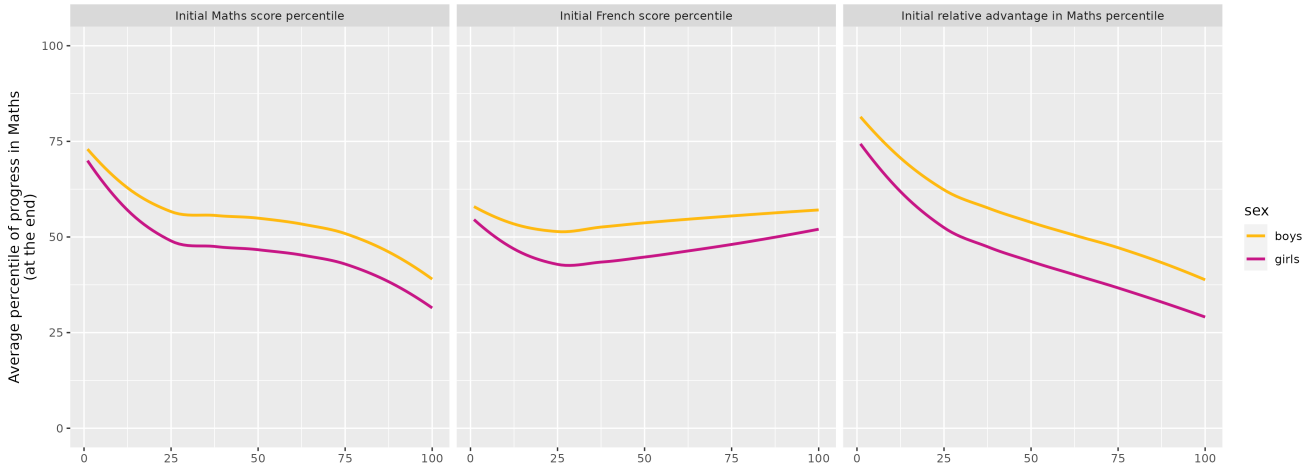
Dependent Variable:	Maths progress									
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Constant	0.0861***	-6.629***								
	(0.0009)	(0.1750)								
Female	-0.1435***	-0.1467***	-0.1435***	-0.1467***	-0.1473***	-0.1487***	-0.1473***	-0.1487***	-0.1492***	-0.1504***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
Initial maths score	-0.5613***	-0.5619***	-0.5613***	-0.5618***	-0.5698***	-0.5712***	-0.5698***	-0.5712***	-0.5659***	-0.5673***
	(0.0014)	(0.0014)	(0.0014)	(0.0014)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)
Initial French score	0.3288***	0.3320***	0.3288***	0.3320***	0.3458***	0.3427***	0.3458***	0.3427***	0.3510***	0.3477***
	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
Female × Initial maths score	-0.0072***	-0.0068***	-0.0072***	-0.0068***	-0.0067***	-0.0065***	-0.0067***	-0.0065***	-0.0059***	-0.0056***
	(0.0018)	(0.0018)	(0.0018)	(0.0018)	(0.0017)	(0.0018)	(0.0017)	(0.0018)	(0.0017)	(0.0018)
Female × Initial French score	0.0502***	0.0515***	0.0503***	0.0515***	0.0523***	0.0525***	0.0523***	0.0525***	0.0527***	0.0532***
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0015)
Controls										
School characteristics and age		Yes		Yes		Yes		Yes		Yes
Fixed-effects										
cohort			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
school_id					Yes	Yes	Yes	Yes	Yes	Yes
cohort × school_id									Yes	Yes
Fit statistics										
Observations	2,683,872	2,534,302	2,683,872	2,534,302	2,683,872	2,534,302	2,683,872	2,534,302	2,683,872	2,534,302
R^2	0.22597	0.23246	0.22600	0.23250	0.28676	0.28788	0.28676	0.28788	0.34750	0.34702
Within R^2			0.22593	0.23242	0.23537	0.23782	0.23537	0.23782	0.23686	0.23915

Notes: Clustered (school_id × cohort) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.

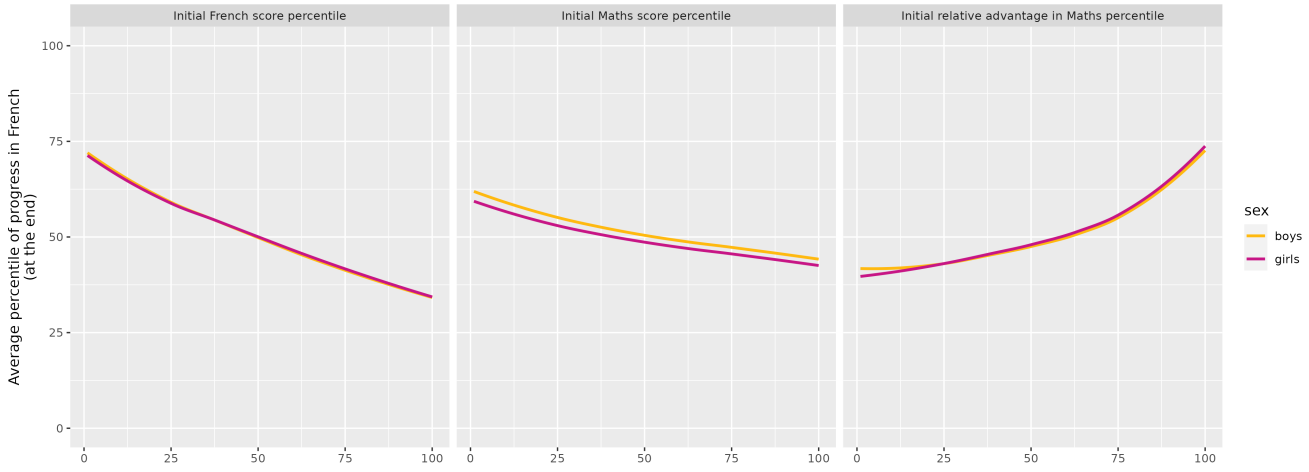
Table 25: French progress (beginning - end) and initial scores

Dependent Variable:	French progress									
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Constant	0.0079***	-10.66***								
	(0.0010)	(0.1824)								
Female	0.0189***	0.0217***	0.0189***	0.0217***	0.0223***	0.0193***	0.0223***	0.0193***	0.0196***	0.0167***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
Initial French score	-0.4273***	-0.4560***	-0.4273***	-0.4560***	-0.4404***	-0.4435***	-0.4404***	-0.4435***	-0.4279***	-0.4313***
	(0.0013)	(0.0012)	(0.0013)	(0.0012)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)	(0.0011)
Initial maths score	0.1821***	0.1784***	0.1821***	0.1785***	0.1695***	0.1697***	0.1695***	0.1697***	0.1705***	0.1706***
	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)	(0.0012)
Female × Initial French score	0.0074***	0.0048***	0.0074***	0.0047***	0.0015	0.0026*	0.0015	0.0026*	0.0031**	0.0042***
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0014)	(0.0015)	(0.0014)	(0.0015)	(0.0014)	(0.0015)
Female × Initial maths score	0.0018	0.0048***	0.0019	0.0049***	0.0080***	0.0096***	0.0080***	0.0096***	0.0082***	0.0098***
	(0.0017)	(0.0017)	(0.0017)	(0.0017)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0016)
Controls										
School characteristics and age		Yes		Yes		Yes		Yes		Yes
Fixed-effects										
cohort			Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
school_id					Yes	Yes	Yes	Yes	Yes	Yes
cohort × school_id									Yes	Yes
Fit statistics										
Observations	2,789,204	2,634,262	2,789,204	2,634,262	2,789,204	2,634,262	2,789,204	2,634,262	2,789,204	2,634,262
R^2	0.16666	0.19260	0.16672	0.19277	0.25733	0.26103	0.25733	0.26103	0.31053	0.31240
Within R^2			0.16663	0.19267	0.17349	0.18098	0.17349	0.18098	0.16424	0.17230

Notes: Clustered (school_id × cohort) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1.



(a) Maths



(b) French

Figure 23: Average progress in maths and French depending on initial scores by gender

Note: Initial relative advantage in maths is computed for student i as $RA_i^{beg} = maths\_score_i^{beg} - french\_score_i^{beg}$

School characteristics

Table 26: Progress in maths (beginning - end): variation explained by school FE by gender

Dependent Variable:	Maths progress			
Sex:	Boys		Girls	
Model:	(1)	(2)	(3)	(4)
Fixed-effects				
school_id	Yes	Yes	Yes	Yes
school_id × cohort		Yes		Yes
Fit statistics				
Dependent variable mean	0.08425	0.08425	-0.10189	-0.10189
R^2	0.08810	0.19081	0.10019	0.21054

Table 27: Progress in maths (beginning - middle): variation explained by school FE by gender

Dependent Variable:	Maths progress			
Sex:	Boys		Girls	
Model:	(1)	(2)	(3)	(4)
Fixed-effects				
school_id	Yes	Yes	Yes	Yes
school_id × cohort		Yes		Yes
Fit statistics				
Dependent variable mean	0.04615	0.04615	-0.04500	-0.04500
R^2	0.07648	0.17869	0.08412	0.19415

Table 28: Progress in French (beginning - end): variation explained by school FE by gender

Dependent Variable:	French progress			
Sex:	Boys		Girls	
Model:	(1)	(2)	(3)	(4)
Fixed-effects				
school_id	Yes	Yes	Yes	Yes
school_id × school_year		Yes		Yes
Fit statistics				
Dependent variable mean	0.03272	0.03272	-0.01885	-0.01885
R^2	0.11504	0.21408	0.11452	0.21649

Table 29: Progress in maths (beginning - middle), school characteristics and gender - individual level

Dependent Variable:	Maths progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-0.0445*	-2.171***				
	(0.0231)	(0.2034)				
Individual characteristics						
Female	-0.1869***	-0.1880***	-0.1869***	-0.1879***	-0.1718***	-0.1730***
	(0.0222)	(0.0221)	(0.0222)	(0.0221)	(0.0208)	(0.0208)
School characteristics						
Type Public‡ *(Private)*	0.0394***	0.0390***	0.0393***	0.0389***		
	(0.0030)	(0.0030)	(0.0030)	(0.0030)		
Priority REP‡ *(None)*	0.0865***	0.0860***	0.0865***	0.0859***		
	(0.0050)	(0.0050)	(0.0050)	(0.0050)		
Priority REP+‡ *(None)*	0.1305***	0.1296***	0.1304***	0.1294***		
	(0.0068)	(0.0068)	(0.0068)	(0.0068)		
Enrollment & Resources						
Enrollment	8.12×10^{-5}***	8.14×10^{-5}***	8.24×10^{-5}***	8.3×10^{-5}***	3.36×10^{-5}	3.37×10^{-5}
	(1.28×10^{-5})	(1.28×10^{-5})	(1.28×10^{-5})	(1.28×10^{-5})	(6.61×10^{-5})	(6.61×10^{-5})
Teacher/student	0.7064***	0.7411***	0.7144***	0.7648***	1.431***	1.416***
	(0.1869)	(0.1868)	(0.1877)	(0.1876)	(0.2690)	(0.2690)
First grade multilevel‡ Yes *(No)*	-0.0619***	-0.0623***	-0.0618***	-0.0622***	-0.0322***	-0.0324***
	(0.0029)	(0.0029)	(0.0029)	(0.0029)	(0.0036)	(0.0036)
First grade multilevel‡ Possibly *(No)*	-0.0123***	-0.0124***	-0.0122***	-0.0123***	-0.0037	-0.0039
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0025)	(0.0025)
Student composition						
Female†	0.0559**	0.0564**	0.0547**	0.0553**	0.0177	0.0180
	(0.0232)	(0.0232)	(0.0232)	(0.0232)	(0.0335)	(0.0335)
IPS	5.75×10^{-5}	2.04×10^{-6}	6.15×10^{-5}	6.89×10^{-6}	0.0003	0.0003
	(8.25×10^{-5})	(8.24×10^{-5})	(8.25×10^{-5})	(8.24×10^{-5})	(0.0004)	(0.0004)
Teacher composition						
Female†	-0.0254***	-0.0253***	-0.0258***	-0.0255***	-0.0209	-0.0204
	(0.0091)	(0.0091)	(0.0091)	(0.0091)	(0.0140)	(0.0140)
Age (years)	-3.05×10^{-5}	-1.65×10^{-5}	-3.24×10^{-5}	-1.23×10^{-5}	0.0014***	0.0014***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0005)	(0.0005)
Seniority witin school (years)	-0.0006	-0.0007	-0.0007	-0.0007	-0.0016*	-0.0016*
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0009)	(0.0009)
Gender interactions						
Female × Type Public	-0.0008	-0.0008	-0.0008	-0.0008	0.0008	0.0008
	(0.0029)	(0.0029)	(0.0029)	(0.0029)	(0.0028)	(0.0028)
Female × Priority REP	0.0017	0.0019	0.0017	0.0018	0.0010	0.0012
	(0.0046)	(0.0045)	(0.0046)	(0.0045)	(0.0044)	(0.0044)
Female × Priority REP+	-0.0021	-0.0017	-0.0022	-0.0018	-0.0014	-0.0009
	(0.0058)	(0.0058)	(0.0058)	(0.0058)	(0.0057)	(0.0057)
Female × Enrollment	-3.51×10^{-5}***	-3.52×10^{-5}***	-3.51×10^{-5}***	-3.53×10^{-5}***	-3.17×10^{-5}***	-3.18×10^{-5}***
	(1.06×10^{-5})	(1.06×10^{-5})	(1.06×10^{-5})	(1.06×10^{-5})	(1.03×10^{-5})	(1.03×10^{-5})
Female × Teacher/student	0.1356	0.1153	0.1371	0.1166	0.1250	0.1037
	(0.1744)	(0.1741)	(0.1743)	(0.1741)	(0.1694)	(0.1692)
Female × First grade multilevel Yes	-0.0005	-0.0003	-0.0005	-0.0003	0.0003	0.0004
	(0.0028)	(0.0028)	(0.0028)	(0.0028)	(0.0027)	(0.0027)
Female × First grade multilevel Possibly	-0.0030	-0.0030	-0.0030	-0.0030	-0.0043**	-0.0043**
	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0022)
Female × (Students) Female	0.0209	0.0198	0.0211	0.0201	-0.0149	-0.0157
	(0.0271)	(0.0270)	(0.0271)	(0.0270)	(0.0225)	(0.0225)
Female × (Students) IPS	0.0005***	0.0005***	0.0005***	0.0005***	0.0005***	0.0005***
	(7.47×10^{-5})	(7.46×10^{-5})	(7.47×10^{-5})	(7.46×10^{-5})	(7.25×10^{-5})	(7.25×10^{-5})
Female × (Teachers) Female	0.0054	0.0049	0.0054	0.0049	0.0084	0.0080
	(0.0082)	(0.0082)	(0.0082)	(0.0082)	(0.0080)	(0.0080)
Female × (Teachers) Age	0.0005	0.0005	0.0005	0.0005	0.0004	0.0004
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Female × (Teachers) Seniority	0.0015***	0.0015***	0.0015***	0.0015***	0.0016***	0.0016***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Controls						
Individual		Yes		Yes		Yes
Fixed-effects						
cohort			Yes	Yes	Yes	Yes
school_id					Yes	Yes
Fit statistics						
Observations	2,368,606	2,368,606	2,368,606	2,368,606	2,368,606	2,368,606
R^2	0.01105	0.01224	0.01117	0.01238	0.07120	0.07239
Within R^2			0.01106	0.01226	0.00468	0.00595

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared, var† indicates a proportion; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Table 30: Progress in French (beginning - end), school characteristics and gender - individual level

Dependent Variable:	French progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Constant	-0.3087***	-2.671***				
	(0.0258)	(0.1926)				
Individual characteristics						
Female	-0.1046***	-0.1065***	-0.1044***	-0.1063***	-0.1004***	-0.1026***
	(0.0215)	(0.0215)	(0.0215)	(0.0215)	(0.0203)	(0.0202)
School characteristics						
Type Public‡ *(Private)*	-0.0112***	-0.0118***	-0.0113***	-0.0120***		
	(0.0033)	(0.0033)	(0.0033)	(0.0033)		
Priority REP‡ *(None)*	0.1145***	0.1136***	0.1137***	0.1126***		
	(0.0057)	(0.0057)	(0.0057)	(0.0057)		
Priority REP+‡ *(None)*	0.1071***	0.1055***	0.1060***	0.1041***		
	(0.0075)	(0.0075)	(0.0075)	(0.0075)		
Enrollment & Resources						
Enrollment	0.0002***	0.0002***	0.0002***	0.0002***	9.46×10^{-5}	9.46×10^{-5}
	(1.39×10^{-5})	(1.39×10^{-5})	(1.39×10^{-5})	(1.39×10^{-5})	(6.55×10^{-5})	(6.54×10^{-5})
Teacher/student	0.7301***	0.7933***	0.8299***	0.9227***	0.6658**	0.6368**
	(0.2124)	(0.2126)	(0.2130)	(0.2131)	(0.2662)	(0.2658)
First grade multilevel‡ Yes *(No)*	-0.0831***	-0.0839***	-0.0829***	-0.0837***	0.0009	0.0008
	(0.0032)	(0.0032)	(0.0032)	(0.0032)	(0.0035)	(0.0035)
First grade multilevel‡ Possibly *(No)*	-0.0003	-0.0005	4.06×10^{-5}	-2.53×10^{-5}	0.0065***	0.0061**
	(0.0027)	(0.0027)	(0.0027)	(0.0027)	(0.0024)	(0.0024)
Student composition						
Female†	-0.0158	-0.0148	-0.0163	-0.0153	0.0106	0.0112
	(0.0260)	(0.0260)	(0.0260)	(0.0260)	(0.0327)	(0.0326)
IPS	0.0011***	0.0010***	0.0012***	0.0010***	6.74×10^{-6}	3.38×10^{-6}
	(9.21×10^{-5})	(9.21×10^{-5})	(9.21×10^{-5})	(9.21×10^{-5})	(0.0003)	(0.0003)
Teacher composition						
Female†	0.1183***	0.1185***	0.1187***	0.1192***	0.0143	0.0154
	(0.0107)	(0.0107)	(0.0107)	(0.0107)	(0.0136)	(0.0135)
Age (years)	0.0016***	0.0016***	0.0016***	0.0017***	0.0011**	0.0012**
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0005)	(0.0005)
Seniority within school (years)	0.0005	0.0005	0.0005	0.0005	-0.0026***	-0.0026***
	(0.0005)	(0.0005)	(0.0005)	(0.0005)	(0.0009)	(0.0009)
Gender interactions						
Female × Type Public	-0.0090***	-0.0091***	-0.0091***	-0.0091***	-0.0080***	-0.0080***
	(0.0029)	(0.0028)	(0.0029)	(0.0028)	(0.0027)	(0.0027)
Female × Priority REP	-0.0104**	-0.0101**	-0.0104**	-0.0101**	-0.0101**	-0.0098**
	(0.0046)	(0.0046)	(0.0046)	(0.0046)	(0.0043)	(0.0043)
Female × Priority REP+	-0.0130**	-0.0121**	-0.0131**	-0.0123**	-0.0080	-0.0071
	(0.0058)	(0.0058)	(0.0058)	(0.0058)	(0.0055)	(0.0055)
Female × Enrollment	-3.51×10^{-5}***	-3.54×10^{-5}***	-3.51×10^{-5}***	-3.54×10^{-5}***	-3.07×10^{-5}***	-3.1×10^{-5}***
	(1.02×10^{-5})	(1.02×10^{-5})	(1.02×10^{-5})	(1.02×10^{-5})	(9.77×10^{-6})	(9.75×10^{-6})
Female × Teacher/student	0.2423	0.2043	0.2432	0.2049	0.1559	0.1135
	(0.1770)	(0.1767)	(0.1770)	(0.1767)	(0.1672)	(0.1667)
Female × First grade multilevel Yes	0.0093***	0.0096***	0.0093***	0.0096***	0.0109***	0.0111***
	(0.0028)	(0.0028)	(0.0028)	(0.0028)	(0.0027)	(0.0027)
Female × First grade multilevel Possibly	-0.0006	-0.0005	-0.0006	-0.0005	-0.0008	-0.0007
	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0021)	(0.0021)
Female × (Students) Female	0.0182	0.0162	0.0181	0.0161	0.0093	0.0081
	(0.0244)	(0.0243)	(0.0244)	(0.0243)	(0.0221)	(0.0220)
Female × (Students) IPS	0.0005***	0.0005***	0.0005***	0.0005***	0.0005***	0.0005***
	(7.37×10^{-5})	(7.35×10^{-5})	(7.37×10^{-5})	(7.35×10^{-5})	(7.02×10^{-5})	(7×10^{-5})
Female × (Teachers) Female	-0.0057	-0.0064	-0.0057	-0.0064	-0.0043	-0.0050
	(0.0081)	(0.0081)	(0.0081)	(0.0081)	(0.0077)	(0.0077)
Female × (Teachers) Age	3.76×10^{-5}	3.72×10^{-5}	4.13×10^{-5}	4.16×10^{-5}	6.92×10^{-5}	6.8×10^{-5}
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
Female × (Teachers) Seniority	0.0004	0.0004	0.0004	0.0004	0.0002	0.0003
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Controls						
Individual		Yes		Yes		Yes
Fixed-effects						
cohort			Yes	Yes	Yes	Yes
school_id					Yes	Yes
Fit statistics						
Observations	2,392,359	2,392,359	2,392,359	2,392,359	2,392,359	2,392,359
R^2	0.00888	0.01374	0.00904	0.01395	0.10253	0.10758
Within R^2			0.00892	0.01383	0.00161	0.00723

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared, var† indicates a proportion; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Table 31: Difference in progress in maths (beginning - end) and school characteristics - School level

Dependent Variable:	Gender gap in maths progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.5384***	0.5396***				
	(0.1237)	(0.1236)				
Type Public‡ *(Private)*	-0.0012	-0.0006	-0.0013	-0.0007		
	(0.0194)	(0.0193)	(0.0194)	(0.0193)		
Priority REP‡ *(None)*	-0.0443*	-0.0483**	-0.0441*	-0.0482**		
	(0.0235)	(0.0235)	(0.0235)	(0.0235)		
Priority REP+‡ *(None)*	-0.0327	-0.0349	-0.0325	-0.0349		
	(0.0285)	(0.0285)	(0.0285)	(0.0285)		
Enrollment & Resources						
Enrollment	-6.26×10^{-5}	-6.14×10^{-5}	-6.32×10^{-5}	-6.17×10^{-5}	-0.0004	-0.0004
	(5.9×10^{-5})	(5.89×10^{-5})	(5.91×10^{-5})	(5.9×10^{-5})	(0.0006)	(0.0006)
Teacher/student	-0.8064	-0.7167	-0.8205	-0.7160	-0.5380	-0.5132
	(0.9312)	(0.9301)	(0.9371)	(0.9359)	(2.997)	(2.989)
First grade multilevel‡ Yes *(No)*	0.0070	0.0069	0.0070	0.0070	-0.0610	-0.0629
	(0.0182)	(0.0181)	(0.0182)	(0.0181)	(0.0502)	(0.0500)
First grade multilevel‡ Possibly *(No)*	0.0086	0.0086	0.0087	0.0087	0.0138	0.0136
	(0.0123)	(0.0123)	(0.0123)	(0.0123)	(0.0231)	(0.0231)
Student composition						
Female†	0.0400	0.0234	0.0408	0.0241	-0.5708	-0.5737
	(0.1349)	(0.1346)	(0.1349)	(0.1346)	(0.3629)	(0.3627)
IPS	-0.0006	-0.0007	-0.0006	-0.0007	0.0022	0.0021
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0043)	(0.0043)
Teacher composition						
Female†	-0.0994**	-0.0958**	-0.0993**	-0.0955**	-0.1124	-0.1112
	(0.0474)	(0.0473)	(0.0474)	(0.0474)	(0.1573)	(0.1571)
Age (years)	-0.0041**	-0.0039**	-0.0041**	-0.0039**	-0.0116*	-0.0114*
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0062)	(0.0062)
Seniority within school (years)	0.0014	0.0015	0.0014	0.0015	-0.0020	-0.0020
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0114)	(0.0114)
Age controls		Yes		Yes		Yes
Fixed-effects						
cohort			Yes	Yes	Yes	Yes
school_id					Yes	Yes
Fit statistics						
Observations	31,179	31,179	31,179	31,179	31,179	31,179
R^2	0.00077	0.00434	0.00084	0.00441	0.56371	0.56489
Within R^2			0.00077	0.00434	0.00104	0.00373

Notes: Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared, var† indicates a proportion; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Table 32: Difference in progress in maths (beginning - middle) and school characteristics - School level

Dependent Variable:	Gender gap in maths progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.2031*	0.2031*				
	(0.1224)	(0.1223)				
Type Public‡ *(Private)*	-0.0109	-0.0109	-0.0103	-0.0104		
	(0.0193)	(0.0193)	(0.0193)	(0.0193)		
Priority REP‡ *(None)*	-0.0272	-0.0301	-0.0253	-0.0283		
	(0.0238)	(0.0238)	(0.0238)	(0.0238)		
Priority REP+‡ *(None)*	-0.0037	-0.0055	-0.0017	-0.0035		
	(0.0286)	(0.0286)	(0.0286)	(0.0286)		
Enrollment & Resources						
Enrollment	-3.4×10^{-5}	-3.36×10^{-5}	-4.1×10^{-5}	-4.03×10^{-5}	-0.0001	-0.0001
	(5.98×10^{-5})	(5.97×10^{-5})	(5.98×10^{-5})	(5.98×10^{-5})	(0.0006)	(0.0006)
Teacher/student	-0.0292	0.0256	-0.2642	-0.2010	4.339	4.394
	(0.9399)	(0.9401)	(0.9439)	(0.9441)	(3.147)	(3.146)
First grade multilevel‡ Yes *(No)*	-0.0099	-0.0099	-0.0104	-0.0104	-0.0339	-0.0342
	(0.0181)	(0.0180)	(0.0181)	(0.0180)	(0.0574)	(0.0573)
First grade multilevel‡ Possibly *(No)*	0.0127	0.0127	0.0117	0.0118	0.0084	0.0087
	(0.0123)	(0.0123)	(0.0123)	(0.0123)	(0.0242)	(0.0242)
Student composition						
Female†	0.1482	0.1420	0.1483	0.1421	0.2554	0.2518
	(0.1326)	(0.1324)	(0.1326)	(0.1324)	(0.3891)	(0.3888)
IPS	-4.26×10^{-5}	-9.25×10^{-5}	-7.58×10^{-5}	-0.0001	0.0044	0.0042
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0045)	(0.0045)
Teacher composition						
Female†	-0.0170	-0.0149	-0.0190	-0.0169	0.1121	0.1143
	(0.0475)	(0.0474)	(0.0475)	(0.0474)	(0.1686)	(0.1686)
Age (years)	-0.0044***	-0.0043***	-0.0045***	-0.0044***	-0.0062	-0.0062
	(0.0016)	(0.0016)	(0.0016)	(0.0016)	(0.0067)	(0.0066)
Seniority within school (years)	0.0053**	0.0053**	0.0051**	0.0052**	0.0094	0.0094
	(0.0024)	(0.0024)	(0.0024)	(0.0024)	(0.0123)	(0.0122)
Age controls		Yes		Yes		Yes
Fixed-effects						
cohort			Yes	Yes	Yes	Yes
school_id					Yes	Yes
Fit statistics						
Observations	27,283	27,283	27,283	27,283	27,283	27,283
R^2	0.00049	0.00215	0.00076	0.00240	0.58591	0.58640
Within R^2			0.00050	0.00214	0.00069	0.00188

Notes: Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared, var† indicates a proportion; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Table 33: Difference in progress in French (beginning - end) and school characteristics - School level

Dependent Variable:	Gender gap in French progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.1175	0.1180				
	(0.1108)	(0.1105)				
Type Public‡ *(Private)*	-0.0119	-0.0113	-0.0119	-0.0114		
	(0.0177)	(0.0177)	(0.0178)	(0.0177)		
Priority REP‡ *(None)*	-0.0106	-0.0169	-0.0102	-0.0166		
	(0.0214)	(0.0213)	(0.0214)	(0.0213)		
Priority REP+‡ *(None)*	-0.0327	-0.0367	-0.0321	-0.0363		
	(0.0261)	(0.0261)	(0.0261)	(0.0261)		
Enrollment & Resources						
Enrollment	5.66×10^{-5}	5.92×10^{-5}	5.5×10^{-5}	5.81×10^{-5}	-0.0001	-0.0001
	(5.36×10^{-5})	(5.34×10^{-5})	(5.36×10^{-5})	(5.34×10^{-5})	(0.0005)	(0.0005)
Teacher/student	0.5580	0.7053	0.5133	0.6791	-1.995	-1.949
	(0.8478)	(0.8458)	(0.8512)	(0.8490)	(2.686)	(2.672)
First grade multilevel‡ Yes *(No)*	-0.0127	-0.0131	-0.0126	-0.0130	-0.0740	-0.0762*
	(0.0166)	(0.0165)	(0.0166)	(0.0165)	(0.0458)	(0.0454)
First grade multilevel‡ Possibly *(No)*	0.0085	0.0083	0.0086	0.0085	0.0022	0.0021
	(0.0113)	(0.0112)	(0.0113)	(0.0112)	(0.0211)	(0.0210)
Student composition						
Female†	0.0597	0.0414	0.0615	0.0431	-0.0098	-0.0162
	(0.1227)	(0.1223)	(0.1227)	(0.1223)	(0.3334)	(0.3325)
IPS	-0.0003	-0.0004	-0.0003	-0.0004	-0.0018	-0.0022
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0039)	(0.0039)
Teacher composition						
Female†	-0.0207	-0.0160	-0.0206	-0.0157	0.0013	0.0035
	(0.0428)	(0.0428)	(0.0429)	(0.0428)	(0.1431)	(0.1428)
Age (years)	-0.0019	-0.0018	-0.0019	-0.0018	-0.0034	-0.0030
	(0.0015)	(0.0015)	(0.0015)	(0.0015)	(0.0057)	(0.0056)
Seniority within school (years)	0.0002	0.0003	0.0001	0.0003	-0.0064	-0.0066
	(0.0022)	(0.0022)	(0.0022)	(0.0022)	(0.0103)	(0.0103)
Age controls		Yes		Yes		Yes
Fixed-effects						
cohort			Yes	Yes	Yes	Yes
school_id					Yes	Yes
Fit statistics						
Observations	31,530	31,530	31,530	31,530	31,530	31,530
R^2	0.00029	0.00786	0.00063	0.00819	0.55394	0.55698
Within R^2			0.00029	0.00785	0.00045	0.00726

Notes: Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared, var† indicates a proportion; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Case studies

Table 34: Progress in maths (beginning - end) and alternative pedagogies - Individual level

Dependent Variable:	Maths progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.0778***	-4.414***	-4.630***			
	(0.0011)	(0.2093)	(0.2151)			
Female	-0.1863***	-0.1880***	-0.1892***	-0.1863***	-0.1880***	-0.1892***
	(0.0009)	(0.0009)	(0.0010)	(0.0009)	(0.0009)	(0.0010)
Pedagogy‡ Freinet *(Rest)*	-0.0195	-0.0208	-0.0015	-0.0195	-0.0208	-0.0011
	(0.0415)	(0.0413)	(0.0410)	(0.0415)	(0.0414)	(0.0411)
Pedagogy‡ Montessori *(Rest)*	-0.1755**	-0.1736**	-0.1757*	-0.1754**	-0.1734**	-0.1754*
	(0.0867)	(0.0868)	(0.0914)	(0.0867)	(0.0867)	(0.0912)
Female × Pedagogy Freinet	-0.0767**	-0.0765**	-0.0786**	-0.0765**	-0.0762**	-0.0782**
	(0.0385)	(0.0384)	(0.0383)	(0.0385)	(0.0383)	(0.0383)
Female × Pedagogy Montessori	0.0918	0.0888	0.0736	0.0923	0.0892	0.0740
	(0.1000)	(0.0994)	(0.1045)	(0.1002)	(0.0997)	(0.1048)
Controls						
Individual		Yes	Yes		Yes	Yes
School-level			Yes			Yes
Fixed-effects						
cohort				Yes	Yes	Yes
Fit statistics						
Observations	2,792,522	2,792,522	2,637,407	2,792,522	2,792,522	2,637,407
R^2	0.01478	0.01718	0.02439	0.01485	0.01728	0.02453
Within R^2				0.01478	0.01721	0.02444

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Table 35: Progress in maths (beginning - middle) and alternative pedagogies - Individual level

Dependent Variable:	Maths progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.0434***	-2.048***	-2.326***			
	(0.0010)	(0.1887)	(0.1946)			
Female	-0.0904***	-0.0914***	-0.0933***	-0.0903***	-0.0913***	-0.0933***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Pedagogy‡ Freinet *(Rest)*	0.0515	0.0505	0.0506	0.0514	0.0504	0.0507
	(0.0378)	(0.0377)	(0.0363)	(0.0379)	(0.0378)	(0.0365)
Pedagogy‡ Montessori *(Rest)*	-0.0131	-0.0120	0.0086	-0.0132	-0.0120	0.0089
	(0.0911)	(0.0911)	(0.0870)	(0.0911)	(0.0910)	(0.0869)
Female × Pedagogy Freinet	-0.0441	-0.0437	-0.0451	-0.0439	-0.0435	-0.0449
	(0.0333)	(0.0332)	(0.0331)	(0.0333)	(0.0331)	(0.0331)
Female × Pedagogy Montessori	-0.0600	-0.0622	-0.0770	-0.0595	-0.0617	-0.0764
	(0.0720)	(0.0711)	(0.0737)	(0.0722)	(0.0714)	(0.0740)
Controls						
Individual		Yes	Yes		Yes	Yes
School-level			Yes			Yes
Fixed-effects						
cohort				Yes	Yes	Yes
Fit statistics						
Observations	2,686,517	2,686,517	2,536,816	2,686,517	2,686,517	2,536,816
R^2	0.00401	0.00513	0.01210	0.00410	0.00522	0.01222
Within R^2				0.00401	0.00514	0.01212

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Table 36: Progress in French (beginning - end) and alternative pedagogies - Individual level

Dependent Variable:	French progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.0270***	-2.515***	-2.862***			
	(0.0011)	(0.1804)	(0.1848)			
Female	-0.0509***	-0.0529***	-0.0539***	-0.0509***	-0.0529***	-0.0539***
	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)	(0.0009)
Pedagogy‡ Freinet *(Rest)*	-0.0009	-0.0024	0.0201	-0.0010	-0.0025	0.0204
	(0.0465)	(0.0465)	(0.0461)	(0.0463)	(0.0464)	(0.0460)
Pedagogy‡ Montessori *(Rest)*	-0.1497***	-0.1466**	-0.1406**	-0.1498***	-0.1466**	-0.1403**
	(0.0575)	(0.0584)	(0.0630)	(0.0575)	(0.0584)	(0.0628)
Female × Pedagogy Freinet	0.0439	0.0438	0.0432	0.0442	0.0442	0.0436
	(0.0327)	(0.0326)	(0.0326)	(0.0327)	(0.0326)	(0.0327)
Female × Pedagogy Montessori	0.0999*	0.0970*	0.0744	0.1005*	0.0976*	0.0748
	(0.0521)	(0.0525)	(0.0516)	(0.0523)	(0.0527)	(0.0518)
Controls						
Individual		Yes	Yes		Yes	Yes
School-level			Yes			Yes
Fixed-effects						
cohort				Yes	Yes	Yes
Fit statistics						
Observations	2,816,468	2,816,468	2,660,012	2,816,468	2,816,468	2,660,012
R^2	0.00126	0.00629	0.01365	0.00136	0.00641	0.01384
Within R^2				0.00126	0.00632	0.01373

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Table 37: Progress in maths (beginning - end) and religious private schools - Individual level

Dependent Variable:	Maths progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6))
Constant	0.0776***	-4.440***	-4.606***			
	(0.0011)	(0.2094)	(0.2151)			
Female	-0.1889***	-0.1906***	-0.1919***	-0.1888***	-0.1906***	-0.1918***
	(0.0010)	(0.0010)	(0.0010)	(0.0010)	(0.0010)	(0.0010)
RPS‡ Other private *(Public)*	0.0049	0.0041	-0.0404***	0.0050	0.0042	-0.0403***
	(0.0046)	(0.0046)	(0.0047)	(0.0046)	(0.0046)	(0.0047)
RPS‡ Catholic *(Public)*	-0.0020	-0.0029	-0.0361***	-0.0018	-0.0027	-0.0359***
	(0.0037)	(0.0037)	(0.0038)	(0.0037)	(0.0037)	(0.0038)
RPS‡ Jewish *(Public)*	0.0797	0.0793	-0.0232	0.0812	0.0812	-0.0213
	(0.0503)	(0.0504)	(0.0512)	(0.0506)	(0.0507)	(0.0515)
RPS‡ Muslim *(Public)*	-0.0004	-0.0042	-0.0113	-0.0008	-0.0045	-0.0103
	(0.0596)	(0.0620)	(0.0667)	(0.0616)	(0.0644)	(0.0696)
Female × RPS Other private	0.0251***	0.0255***	0.0241***	0.0251***	0.0256***	0.0242***
	(0.0041)	(0.0041)	(0.0041)	(0.0041)	(0.0041)	(0.0041)
Female × RPS Catholic	0.0160***	0.0166***	0.0164***	0.0160***	0.0167***	0.0165***
	(0.0033)	(0.0033)	(0.0033)	(0.0033)	(0.0033)	(0.0033)
Female × RPS Jewish	0.0501	0.0489	0.0506	0.0496	0.0483	0.0502
	(0.0529)	(0.0530)	(0.0540)	(0.0528)	(0.0530)	(0.0540)
Female × RPS Muslim	-0.1055	-0.1031	-0.1105	-0.1061	-0.1037	-0.1113
	(0.0916)	(0.0872)	(0.0888)	(0.0918)	(0.0873)	(0.0889)
Controls						
Individual		Yes	Yes		Yes	Yes
School-level			Yes			Yes
Fixed-effects						
cohort				Yes	Yes	Yes
Fit statistics						
Observations	2,791,805	2,791,805	2,637,407	2,791,805	2,791,805	2,637,407
R^2	0.01483	0.01724	0.02440	0.01491	0.01734	0.02454
Within R^2				0.01483	0.01726	0.02445

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Table 38: Progress in maths (beginning - middle) and religious private schools - Individual level

Dependent Variable:	Maths progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Constant	0.0512***	-2.031***	-2.302***			
	(0.0011)	(0.1886)	(0.1944)			
Female	-0.0917***	-0.0928***	-0.0948***	-0.0917***	-0.0928***	-0.0947***
	(0.0009)	(0.0009)	(0.0010)	(0.0009)	(0.0009)	(0.0010)
RPS‡ Other private *(Public)*	-0.0581***	-0.0587***	-0.0419***	-0.0581***	-0.0586***	-0.0418***
	(0.0041)	(0.0041)	(0.0043)	(0.0041)	(0.0041)	(0.0043)
RPS‡ Catholic *(Public)*	-0.0672***	-0.0677***	-0.0472***	-0.0671***	-0.0676***	-0.0470***
	(0.0033)	(0.0033)	(0.0035)	(0.0033)	(0.0033)	(0.0035)
RPS‡ Jewish *(Public)*	0.1334***	0.1324***	0.1252***	0.1345***	0.1337***	0.1267***
	(0.0432)	(0.0432)	(0.0436)	(0.0433)	(0.0433)	(0.0438)
RPS‡ Muslim *(Public)*	-0.2706	-0.2714	-0.2636	-0.2710	-0.2719	-0.2637
	(0.2277)	(0.2294)	(0.2297)	(0.2283)	(0.2304)	(0.2314)
Female × RPS Other private	0.0117***	0.0120***	0.0126***	0.0117***	0.0120***	0.0127***
	(0.0038)	(0.0038)	(0.0039)	(0.0038)	(0.0038)	(0.0039)
Female × RPS Catholic	0.0088***	0.0091***	0.0099***	0.0088***	0.0092***	0.0099***
	(0.0031)	(0.0031)	(0.0032)	(0.0031)	(0.0031)	(0.0032)
Female × RPS Jewish	-0.0500	-0.0508	-0.0561	-0.0504	-0.0513	-0.0565
	(0.0664)	(0.0663)	(0.0671)	(0.0666)	(0.0664)	(0.0674)
Female × RPS Muslim	-0.0705	-0.0737	-0.0720	-0.0705	-0.0738	-0.0724
	(0.0658)	(0.0635)	(0.0634)	(0.0658)	(0.0635)	(0.0633)
Controls						
Individual		Yes	Yes		Yes	Yes
School-level			Yes			Yes
Fixed-effects						
cohort				Yes	Yes	Yes
Fit statistics						
Observations	2,685,807	2,685,807	2,536,816	2,685,807	2,685,807	2,536,816
R^2	0.00479	0.00591	0.01214	0.00487	0.00601	0.01227
Within R^2				0.00478	0.00592	0.01217

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Table 39: Progress in French (beginning - end) and religious private schools - Individual level

Dependent Variable:	French progress					
Model:	(1)	(2)	(3)	(4)	(5)	(6))
Constant	0.0234***	-2.553***	-2.882***			
	(0.0012)	(0.1804)	(0.1847)			
Female	-0.0529***	-0.0550***	-0.0561***	-0.0529***	-0.0550***	-0.0561***
	(0.0009)	(0.0009)	(0.0010)	(0.0009)	(0.0009)	(0.0010)
RPS‡ Other private *(Public)*	0.0255***	0.0242***	0.0014	0.0256***	0.0243***	0.0015
	(0.0047)	(0.0047)	(0.0048)	(0.0047)	(0.0047)	(0.0048)
RPS‡ Catholic *(Public)*	0.0310***	0.0296***	0.0157***	0.0311***	0.0297***	0.0158***
	(0.0036)	(0.0036)	(0.0037)	(0.0036)	(0.0036)	(0.0037)
RPS‡ Jewish *(Public)*	0.0706	0.0666	-0.0025	0.0720	0.0684	-0.0005
	(0.0442)	(0.0445)	(0.0455)	(0.0442)	(0.0446)	(0.0456)
RPS‡ Muslim *(Public)*	0.0867	0.0824	0.1000	0.0859	0.0817	0.1012
	(0.1528)	(0.1545)	(0.1566)	(0.1528)	(0.1543)	(0.1564)
Female × RPS Other private	0.0207***	0.0211***	0.0205***	0.0208***	0.0212***	0.0205***
	(0.0037)	(0.0037)	(0.0037)	(0.0037)	(0.0037)	(0.0037)
Female × RPS Catholic	0.0138***	0.0145***	0.0145***	0.0138***	0.0145***	0.0145***
	(0.0030)	(0.0030)	(0.0030)	(0.0030)	(0.0030)	(0.0030)
Female × RPS Jewish	0.0490	0.0469	0.0495	0.0485	0.0464	0.0491
	(0.0491)	(0.0490)	(0.0491)	(0.0491)	(0.0490)	(0.0491)
Female × RPS Muslim	-0.0576	-0.0562	-0.0601	-0.0583	-0.0569	-0.0609
	(0.0874)	(0.0821)	(0.0828)	(0.0876)	(0.0822)	(0.0828)
Controls						
Individual		Yes	Yes		Yes	Yes
School-level			Yes			Yes
Fixed-effects						
cohort				Yes	Yes	Yes
Fit statistics						
Observations	2,815,733	2,815,733	2,660,012	2,815,733	2,815,733	2,660,012
R^2	0.00158	0.00659	0.01366	0.00168	0.00672	0.01386
Within R^2				0.00158	0.00663	0.01375

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions; var‡ (*baseline*) identifies factor variables, baseline category is indicated on the right.

Teacher gender

Back-of-the-enveloppe computations (maths)

$$\begin{aligned} GG_{end} &= GG_{beg} + (GG_{end} - GG_{beg}) \\ &= GG_{beg} + GG\_prog \end{aligned}$$

With $GG_{beg} = -0.02$ SD and $GG_{end} = 0.16$ SD, $GG\_prog = 0.18$ SD.

On the teacher gender sample, I find (Tables 5 and 6, specifications (1)) that the effect of

having a male teacher accounts for 14% of the negative effect on maths progress of being female ($\frac{0.03}{0.03+0.18}$)).

I consider several alternatives about the proportion of students having a male teacher, and compute the effect of switching to female teachers only on the end maths gender gap.

Hypothesis 1: If every student had a male teacher.

$$\begin{aligned} GG_{end}^{hyp1} &= GG_{beg} + (1 - 0.14) \times GG\_prog \\ &= -0.02 + 0.86 \times 0.18 \\ &= 0.135 \end{aligned}$$

The end maths gender gap (GG_{end}) would be reduced by 16%.

Hypothesis 2: If 13% of students had a male teacher (based on proportion of average proportion of male teachers in school × cohort sample).

$$\begin{aligned} GG_{end}^{hyp1} &= GG_{beg} + 0.13 \times (1 - 0.14) \times GG\_prog + (1 - 0.13) \times GG\_prog \\ &= -0.02 + 0.13 \times 0.86 \times 0.18 + 0.87 \times 0.18 \\ &= 0.156 \end{aligned}$$

The end maths gender gap (GG_{end}) would only be reduced by 2%.

Hypothesis 3: If 6% of students had a male teacher (based on proportion in the teacher gender sample).

$$\begin{aligned} GG_{end}^{hyp1} &= GG_{beg} + 0.06 \times (1 - 0.14) \times GG\_prog + (1 - 0.06) \times GG\_prog \\ &= -0.02 + 0.06 \times 0.86 \times 0.18 + 0.94 \times 0.18 \\ &= 0.16 \end{aligned}$$

The end maths gender gap (GG_{end}) would not be significantly reduced.

Table 40: Progress in maths (beginning - middle) and student and teacher gender

Dependent Variable:	Maths progress											
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	0.0388***	-2.390***	-2.710***									
	(0.0014)	(0.2678)	(0.2775)									
Female student	-0.0885***	-0.0895***	-0.0906***	-0.0885***	-0.0895***	-0.0906***	-0.0892***	-0.0903***	-0.0906***	-0.0899***	-0.0910***	-0.0912***
	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)
Male teacher	0.0515***	0.0522***	0.0265***	0.0514***	0.0520***	0.0264***	0.0255***	0.0260***	0.0273***	0.0213***	0.0219***	0.0246***
	(0.0058)	(0.0058)	(0.0061)	(0.0058)	(0.0058)	(0.0061)	(0.0069)	(0.0069)	(0.0070)	(0.0078)	(0.0078)	(0.0079)
Female student × Male teacher	-0.0213***	-0.0217***	-0.0228***	-0.0213***	-0.0217***	-0.0228***	-0.0219***	-0.0224***	-0.0238***	-0.0180***	-0.0185***	-0.0199***
	(0.0056)	(0.0056)	(0.0057)	(0.0056)	(0.0056)	(0.0057)	(0.0055)	(0.0055)	(0.0056)	(0.0055)	(0.0055)	(0.0056)
Controls												
Individual		Yes	Yes		Yes	Yes		Yes	Yes		Yes	Yes
School-level			Yes			Yes			Yes			Yes
Fixed-effects												
cohort				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
school_id							Yes	Yes	Yes	Yes	Yes	Yes
cohort × school_id										Yes	Yes	Yes
Fit statistics												
Observations	1,223,887	1,223,887	1,159,234	1,223,887	1,223,887	1,159,234	1,223,887	1,223,887	1,159,234	1,223,887	1,223,887	1,159,234
R^2	0.00436	0.00557	0.01221	0.00442	0.00564	0.01226	0.08384	0.08513	0.08395	0.15979	0.16103	0.15821
Within R^2				0.00436	0.00558	0.01221	0.00453	0.00593	0.00611	0.00480	0.00627	0.00630

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions.

Table 41: Progress in French (beginning - end) and pupil and teacher gender

Dependent Variable:	French progress											
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Constant	0.0326***	-2.651***	-2.871***									
	(0.0015)	(0.2637)	(0.2682)									
Female student	-0.0504***	-0.0523***	-0.0530***	-0.0503***	-0.0523***	-0.0530***	-0.0503***	-0.0524***	-0.0529***	-0.0504***	-0.0524***	-0.0528***
	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0013)	(0.0012)	(0.0012)	(0.0013)	(0.0012)	(0.0012)	(0.0013)
Male teacher	-0.0047	-0.0033	-0.0222***	-0.0049	-0.0036	-0.0223***	-0.0159**	-0.0149**	-0.0180***	-0.0140*	-0.0128*	-0.0148*
	(0.0062)	(0.0062)	(0.0065)	(0.0062)	(0.0061)	(0.0065)	(0.0067)	(0.0067)	(0.0069)	(0.0075)	(0.0075)	(0.0077)
Female student × Male teacher	-0.0255***	-0.0262***	-0.0256***	-0.0255***	-0.0262***	-0.0256***	-0.0261***	-0.0269***	-0.0262***	-0.0263***	-0.0271***	-0.0269***
	(0.0055)	(0.0055)	(0.0056)	(0.0055)	(0.0055)	(0.0056)	(0.0053)	(0.0052)	(0.0054)	(0.0053)	(0.0052)	(0.0054)
Controls												
Individual		Yes	Yes		Yes	Yes		Yes	Yes		Yes	Yes
School-level			Yes			Yes			Yes			Yes
Fixed-effects												
cohort				Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
school_id							Yes	Yes	Yes	Yes	Yes	Yes
cohort × school_id										Yes	Yes	Yes
Fit statistics												
Observations	1,264,626	1,264,626	1,197,990	1,264,626	1,264,626	1,197,990	1,264,626	1,264,626	1,197,990	1,264,626	1,264,626	1,197,990
R^2	0.00144	0.00671	0.01571	0.00161	0.00693	0.01597	0.11416	0.11968	0.11804	0.18532	0.19059	0.18742
Within R^2				0.00144	0.00677	0.01582	0.00160	0.00782	0.00796	0.00167	0.00813	0.00822

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions.

Table 42: IV teacher gender: First stage, maths progress beginning - end

Dependent Variables:	Male Teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Variables																		
Constant	0.7046***	-8.88×10^{-16}***	0.5966***	-0.1305***	0.5692***	-0.1333***												
	(0.0091)	(2.25×10^{-17})	(0.0698)	(0.0426)	(0.0737)	(0.0446)												
Female student	-0.0080*	0.6966***	-0.0080*	0.6966***	-0.0062	0.6694***	-0.0080*	0.6966***	-0.0080*	0.6966***	-0.0062	0.6694***	-0.0106***	0.6922***	-0.0106***	0.6921***	-0.0104***	0.6643***
	(0.0044)	(0.0090)	(0.0044)	(0.0090)	(0.0046)	(0.0096)	(0.0044)	(0.0090)	(0.0044)	(0.0090)	(0.0046)	(0.0096)	(0.0025)	(0.0089)	(0.0025)	(0.0089)	(0.0027)	(0.0095)
Proportion of female teachers	-0.7091***	8.88×10^{-16}***	-0.7091***	-7.06×10^{-6}	-0.6860***	-0.0040***	-0.7092***	-5.66×10^{-5}	-0.7092***	-6.13×10^{-5}	-0.6861***	-0.0040***	-0.4605**	0.1215***	-0.4605***	0.1216***	-0.4261***	0.1244***
	(0.0093)	(2.42×10^{-17})	(0.0093)	(1.27×10^{-5})	(0.0108)	(0.0012)	(0.0093)	(9.58×10^{-5})	(0.0093)	(9.62×10^{-5})	(0.0108)	(0.0012)	(0.0156)	(0.0083)	(0.0156)	(0.0083)	(0.0159)	(0.0086)
Female student × Proportion of female teachers	0.0072*	-0.7015***	0.0077*	-0.7015***	0.0058	-0.6724***	0.0077*	-0.7015***	0.0077*	-0.7015***	0.0058	-0.6724***	0.0107***	-0.6966***	0.0107***	-0.6966***	0.0107***	-0.6668***
	(0.0045)	(0.0093)	(0.0045)	(0.0093)	(0.0048)	(0.0099)	(0.0045)	(0.0093)	(0.0045)	(0.0093)	(0.0048)	(0.0099)	(0.0026)	(0.0092)	(0.0026)	(0.0092)	(0.0028)	(0.0098)
Controls																		
Individual			Yes	Yes	Yes	Yes			Yes	Yes	Yes	Yes			Yes	Yes	Yes	Yes
School-level					Yes	Yes					Yes	Yes					Yes	Yes
Fixed-effects																		
cohort							Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
school_id													Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics																		
F-test (1st stage)	97,102.5	96,498.3	97,088.8	96,495.2	67,705.4	72,506.2	97,050.1	96,471.7	97,037.7	96,468.6	67,648.1	72,473.6	7,511.9	62,109.5	7,511.6	62,108.0	5,526.6	49,318.6

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions.

Table 43: IV teacher gender: First stage, maths progress beginning - middle

Dependent Variables:	Male Teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Constant	0.7014***	-1.24×10^{-14}***	0.6006***	-0.1347***	0.5750***	-0.1374***												
	(0.0092)	(2.12×10^{-16})	(0.0714)	(0.0433)	(0.0755)	(0.0454)												
Female student	-0.0070	0.6944***	-0.0070	0.6943***	-0.0050	0.6668***	-0.0070	0.6944***	-0.0070	0.6943***	-0.0050	0.6668***	-0.0099***	0.6896***	-0.0099***	0.6896***	-0.0098***	0.6613***
	(0.0044)	(0.0091)	(0.0044)	(0.0091)	(0.0047)	(0.0097)	(0.0044)	(0.0091)	(0.0044)	(0.0091)	(0.0047)	(0.0097)	(0.0025)	(0.0090)	(0.0025)	(0.0090)	(0.0027)	(0.0096)
Proportion of female teachers	-0.7060***	1.42×10^{-14}***	-0.7060***	-1.01×10^{-5}	-0.6821***	-0.0039***	-0.7061***	-4.76×10^{-5}	-0.7061***	-5.45×10^{-5}	-0.6822***	-0.0040***	-0.4557***	0.1217***	-0.4557***	0.1217***	-0.4226***	0.1235***
	(0.0094)	(2.29×10^{-16})	(0.0094)	(1.35×10^{-5})	(0.0109)	(0.0012)	(0.0094)	(0.0001)	(0.0094)	(0.0001)	(0.0109)	(0.0012)	(0.0157)	(0.0083)	(0.0157)	(0.0083)	(0.0159)	(0.0086)
Female student × Proportion of female teachers	0.0067	-0.6993***	0.0067	-0.6993***	0.0045	-0.6699***	0.0067	-0.6993***	0.0067	-0.6993***	0.0044	-0.6699***	0.0100***	-0.6941***	0.0100***	-0.6941***	0.0100***	-0.6638***
	(0.0046)	(0.0093)	(0.0046)	(0.0093)	(0.0049)	(0.0100)	(0.0046)	(0.0093)	(0.0046)	(0.0093)	(0.0049)	(0.0100)	(0.0025)	(0.0093)	(0.0025)	(0.0093)	(0.0027)	(0.0099)
Controls																		
Individual			Yes	Yes	Yes	Yes			Yes	Yes	Yes	Yes			Yes	Yes	Yes	Yes
School-level					Yes	Yes					Yes	Yes					Yes	Yes
Fixed-effects																		
cohort							Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
school_id													Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics																		
F-test (1st stage)	94,339.4	93,825.5	94,325.1	93,822.8	65,730.2	70,457.9	94,235.7	93,773.7	94,223.2	93,770.8	65,632.0	70,404.4	7,213.4	60,390.8	7,213.3	60,388.8	5,332.3	47,884.4

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions.

Table 44: IV teacher gender: First stage, French progress beginning - end

Dependent Variables:	Male Teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher	Male teacher	Female student × Male teacher
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Constant	0.7042***	-1.78×10^{-15}***	0.5889***	-0.1341***	0.5647***	-0.1368***												
	(0.0091)	(3.91×10^{-17})	(0.0689)	(0.0422)	(0.0729)	(0.0442)												
Female student	-0.0076*	0.6966***	-0.0076*	0.6965***	-0.0059	0.6694***	-0.0076*	0.6966***	-0.0076*	0.6965***	-0.0059	0.6694***	-0.0103***	0.6921***	-0.0103***	0.6921***	-0.0102***	0.6643***
	(0.0044)	(0.0090)	(0.0044)	(0.0090)	(0.0046)	(0.0096)	(0.0044)	(0.0090)	(0.0044)	(0.0090)	(0.0046)	(0.0096)	(0.0025)	(0.0089)	(0.0025)	(0.0089)	(0.0026)	(0.0095)
Proprtion of female teachers	-0.7087***	2.66×10^{-15}***	-0.7087***	-8.63×10^{-6}	-0.6856***	-0.0040***	-0.7088***	-6.13×10^{-5}	-0.7088***	-6.74×10^{-5}	-0.6858***	-0.0040***	-0.4606***	0.1209***	-0.4606***	0.1209***	-0.4259***	0.1240***
	(0.0093)	(4.54×10^{-17})	(0.0093)	(1.29×10^{-5})	(0.0108)	(0.0012)	(0.0093)	(9.65×10^{-5})	(0.0093)	(9.69×10^{-5})	(0.0108)	(0.0012)	(0.0155)	(0.0083)	(0.0155)	(0.0083)	(0.0159)	(0.0086)
Female student × Proportion of female teachers	0.0073	-0.7014***	0.0073	-0.7014***	0.0055	-0.6724***	0.0073	-0.7014***	0.0073	-0.7014***	0.0054	-0.6724***	0.0105***	-0.6966***	0.0105***	-0.6966***	0.0105***	-0.6668***
	(0.0045)	(0.0093)	(0.0045)	(0.0093)	(0.0048)	(0.0099)	(0.0045)	(0.0093)	(0.0045)	(0.0093)	(0.0048)	(0.0099)	(0.0025)	(0.0092)	(0.0025)	(0.0092)	(0.0027)	(0.0098)
Controls																		
Individual			Yes	Yes	Yes	Yes			Yes	Yes	Yes	Yes			Yes	Yes	Yes	Yes
School-level					Yes	Yes					Yes	Yes					Yes	Yes
Fixed-effects																		
cohort							Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
school_id													Yes	Yes	Yes	Yes	Yes	Yes
Fit statistics																		
F-test (1st stage)	97,686.1	97,089.6	97,671.8	97,086.4	68,124.9	72,980.1	97,632.0	97,063.4	97,618.9	97,060.1	68,066.2	72,948.1	7,570.1	62,515.2	7,569.8	62,513.4	5,555.3	49,667.1

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions.

Table 45: IV teacher gender: Second stage, maths progress beginning - middle

Dependent Variable:	Maths progress								
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.0269***	-2.396***	-2.739***						
	(0.0016)	(0.2684)	(0.2771)						
Female student	-0.0880***	-0.0890***	-0.0899***	-0.0880***	-0.0890***	-0.0899***	-0.0886***	-0.0896***	-0.0897***
	(0.0015)	(0.0015)	(0.0016)	(0.0015)	(0.0015)	(0.0016)	(0.0015)	(0.0015)	(0.0016)
Male teacher	0.2506***	0.2540***	0.0820***	0.2507***	0.2537***	0.0817***	0.1466***	0.1464***	0.1387**
	(0.0165)	(0.0165)	(0.0184)	(0.0165)	(0.0165)	(0.0184)	(0.0488)	(0.0488)	(0.0568)
Female student × Male teacher	-0.0276*	-0.0287*	-0.0338**	-0.0277*	-0.0288*	-0.0339**	-0.0306**	-0.0319**	-0.0368**
	(0.0152)	(0.0152)	(0.0168)	(0.0152)	(0.0152)	(0.0168)	(0.0147)	(0.0147)	(0.0165)
Controls									
Individual		Yes	Yes		Yes	Yes		Yes	Yes
School-level			Yes			Yes			Yes
Fixed-effects									
cohort				Yes	Yes	Yes	Yes	Yes	Yes
school_id							Yes	Yes	Yes
Fit statistics									
Observations	1,223,674	1,223,674	1,159,234	1,223,674	1,223,674	1,159,234	1,223,674	1,223,674	1,159,234
R^2	-7.43×10^{-5}	0.00102	0.01194	-1.89×10^{-5}	0.00110	0.01201	0.08308	0.08438	0.08332
Within R^2				-8.14×10^{-5}	0.00104	0.01195	0.00369	0.00511	0.00542

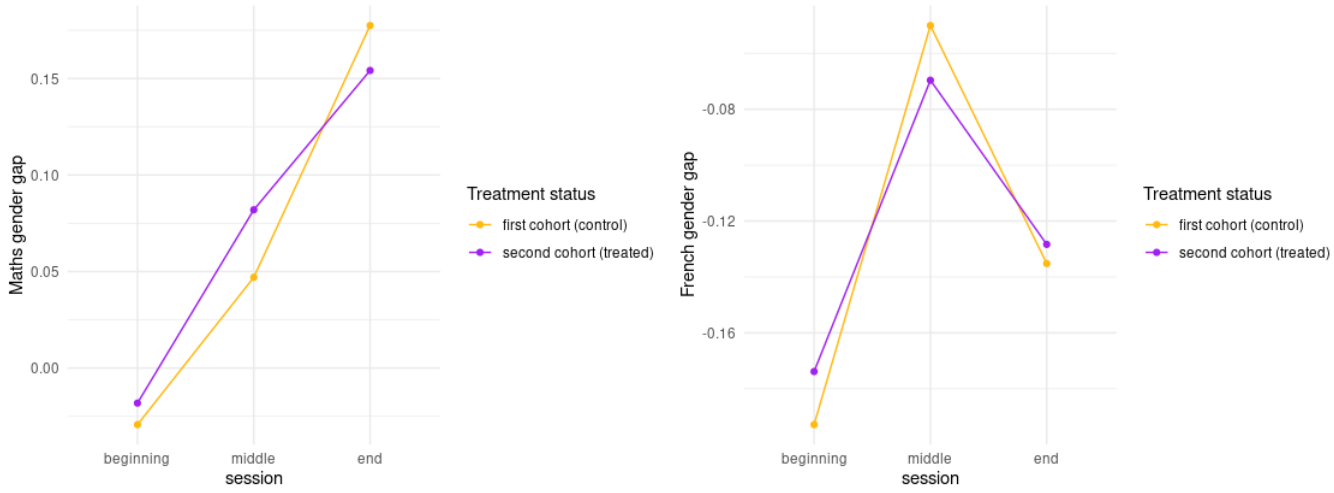
Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions.

Table 46: IV teacher gender: Second stage, French progress beginning - end

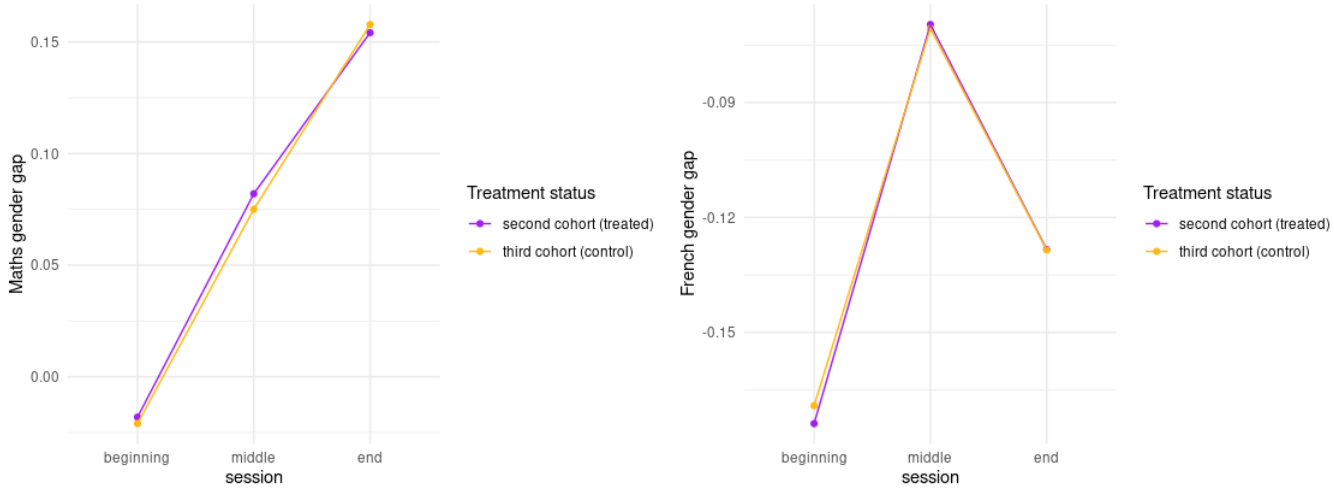
Dependent Variable:	French progress								
Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.0234***	-2.662***	-2.884***						
	(0.0018)	(0.2643)	(0.2679)						
Female student	-0.0478***	-0.0497***	-0.0504***	-0.0478***	-0.0497***	-0.0503***	-0.0475***	-0.0495***	-0.0499***
	(0.0015)	(0.0015)	(0.0016)	(0.0015)	(0.0015)	(0.0016)	(0.0014)	(0.0014)	(0.0015)
Male teacher	0.1501***	0.1565***	0.0121	0.1486***	0.1547***	0.0092	0.0633	0.0635	0.0425
	(0.0182)	(0.0182)	(0.0207)	(0.0182)	(0.0182)	(0.0207)	(0.0475)	(0.0474)	(0.0548)
Female student × Male teacher	-0.0670***	-0.0683***	-0.0698***	-0.0671***	-0.0685***	-0.0700***	-0.0724***	-0.0742***	-0.0745***
	(0.0152)	(0.0152)	(0.0166)	(0.0152)	(0.0152)	(0.0166)	(0.0143)	(0.0143)	(0.0159)
Controls									
Individual		Yes	Yes		Yes	Yes		Yes	Yes
School-level			Yes			Yes			Yes
Fixed-effects									
cohort				Yes	Yes	Yes	Yes	Yes	Yes
school_id							Yes	Yes	Yes
Fit statistics									
Observations	1,264,412	1,264,412	1,197,990	1,264,412	1,264,412	1,197,990	1,264,412	1,264,412	1,197,990
R^2	-0.00070	0.00442	0.01564	-0.00049	0.00470	0.01591	0.11378	0.11931	0.11790
Within R^2				-0.00066	0.00453	0.01575	0.00133	0.00756	0.00779

Notes: Clustered (cohort × school_id) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Individual controls include age and age squared; School-level controls include variables included in schooling environment regressions.

Cohorts comparison



(a) First cohort



(b) Third cohort

Figure 24: Gender gaps: control and treated cohorts

Note: Gender gaps are expressed in proportion of a standard deviation, and computed as $\overline{score}^t_{boys} - \overline{score}^t_{girls}$

Dependent Variable:	Maths score					
Control Group:	First cohort		Third cohort		Fourth cohort	
Model:	(1)	(2)	(3)	(4)	(5)	(6)
Variables						
female	0.0293***	0.0327***	0.0210***	0.0245***	0.0214***	0.0227***
	(0.0024)	(0.0024)	(0.0022)	(0.0022)	(0.0023)	(0.0022)
treated	0.0340***	0.0384***	0.0042	0.0091***	0.0285***	0.0335***
	(0.0031)	(0.0030)	(0.0028)	(0.0027)	(0.0029)	(0.0028)
sessionMiddle	0.0447***	0.0481***	0.0471***	0.0486***	0.0528***	0.0554***
	(0.0024)	(0.0025)	(0.0019)	(0.0019)	(0.0020)	(0.0020)
female × treated	-0.0111***	-0.0119***	-0.0029	-0.0029	-0.0033	-0.0014
	(0.0032)	(0.0032)	(0.0031)	(0.0030)	(0.0031)	(0.0031)
female × sessionMiddle	-0.0763***	-0.0774***	-0.0960***	-0.0967***	-0.0901***	-0.0906***
	(0.0021)	(0.0021)	(0.0017)	(0.0018)	(0.0018)	(0.0018)
treated × sessionMiddle	-0.0052*	-0.0068**	-0.0076***	-0.0072***	-0.0133***	-0.0140***
	(0.0028)	(0.0029)	(0.0023)	(0.0024)	(0.0024)	(0.0025)
female × treated × sessionMiddle	-0.0238***	-0.0230***	-0.0041	-0.0038	-0.0100***	-0.0098***
	(0.0027)	(0.0028)	(0.0025)	(0.0025)	(0.0025)	(0.0026)
Age and school-level controls						
		Yes		Yes		Yes
Fit statistics						
Observations	2,337,952	2,240,682	2,546,774	2,444,864	2,545,644	2,442,558
R^2	0.00099	0.05796	0.00108	0.06686	0.00110	0.06243
Adjusted R^2	0.00099	0.05795	0.00108	0.06685	0.00110	0.06242

Notes: Clustered (school_id × cohort) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Controls include average students' age and school-level variables used in schooling environment regressions.

Table 47: Parallel trends test - Maths

Table 48: COVID-19 effect on the French gender gap

Dependent Variable:	French score			
Control Group:	First Cohort		Third Cohort	
Model:	(1)	(2)	(3)	(4)
female	0.0501***	0.0512***	0.0706***	0.0731***
	(0.0025)	(0.0025)	(0.0023)	(0.0023)
treated	0.0087*	0.0064	-0.0036	0.0024
	(0.0047)	(0.0040)	(0.0046)	(0.0037)
after	-0.0316***	-0.0315***	-0.0359***	-0.0354***
	(0.0021)	(0.0022)	(0.0017)	(0.0017)
female × treated	0.0196***	0.0202***	-0.0010	-0.0010
	(0.0034)	(0.0034)	(0.0032)	(0.0032)
female × after	0.0851***	0.0852***	0.0578***	0.0579***
	(0.0017)	(0.0017)	(0.0014)	(0.0014)
treated × after	-0.0023	-0.0025	0.0020	0.0014
	(0.0029)	(0.0030)	(0.0026)	(0.0027)
female × treated × after	-0.0264***	-0.0264***	0.0009	0.0010
	(0.0022)	(0.0023)	(0.0020)	(0.0021)
Controls		Yes		Yes
Fit statistics				
Observations	2,355,838	2,257,606	2,569,034	2,466,224
R^2	0.00328	0.08094	0.00332	0.09007
Adjusted R^2	0.00328	0.08093	0.00332	0.09006

Notes: Clustered (school_id × cohort) standard-errors in parentheses; Signif. Codes: ***: 0.01, **: 0.05, *: 0.1; Controls include average students' age and school-level variables used in schooling environment regressions.