



# Early Gendered Performance Gaps in Math: An Investigation on French Data

Thomas Breda, Joyce Sultan Parraud, Lola Touitou

## ► To cite this version:

Thomas Breda, Joyce Sultan Parraud, Lola Touitou. Early Gendered Performance Gaps in Math: An Investigation on French Data. 2024. halshs-04414597

**HAL Id: halshs-04414597**

**<https://shs.hal.science/halshs-04414597>**

Preprint submitted on 24 Jan 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**WORKING PAPER N° 2024 – 01**

**Early Gendered Performance Gaps in Math:  
An Investigation on French Data**

Thomas Breda  
Joyce Sultan Parraud  
Lola Toutilou

JEL Codes: I21 ; I24 ; J16.

Keywords: Gender gaps; Math performance; Education.



# Early Gendered Performance Gaps in Math: An Investigation on French Data

Thomas Breda\*  
Joyce Sultan Parraud<sup>†</sup>  
Lola Touitou<sup>‡</sup>

January 24, 2024

## Abstract

When and where do gender gaps in school performance appear, and what is the role of school in their early evolution? While there is no gap in mathematics performance at the start of first grade, a gap in favor of boys appears and widens during the first year of primary school. Using standardized national assessments administered during first grade (CP) to more than 2.5 million pupils in France between 2018 and 2022, we show that this relative drop in girls' performance is observed for all the cohorts and most of the exercises assessed. The greatest drop-off occurs among the best-performing girls at the start of first grade (those in the top 1 % initially). These girls lose an average of nearly 7 percentile ranks at the start of second grade compared with boys in the same initial percentile. The aggregate dynamic is observed across all social categories and family configurations, and throughout the country. Girls lose slightly less ground compared to boys in classes where the top student in math is a girl, and in priority public schools (REP or REP+). However, characteristics of the school environment explain only a small part of the overall dynamics, suggesting that the girls are losing ground compared to boys in every strata of society.

**Keywords:** Gender gaps; Math performance; Education.

**JEL Codes:** I21 ; I24 ; J16.

---

\*Paris School of Economics, CNRS, Institut des politiques publiques. Email: thomas.breda@psemail.eu

<sup>†</sup>Institut des politiques publiques. Email: joyce.sultan@ipp.eu

<sup>‡</sup>Institut des politiques publiques (IPP).

We would like to thank the Women and Science Chair at Université Paris Dauphine-PSL for funding this project, and the DEPP for providing the data. We also thank participants at the IPP internal seminar are also thanked for their helpful comments. All remaining errors are our own.

# 1 Introduction

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 labour outcomes. Women are under-represented in the occupations considered to be the most prestigious or the most lucrative and this is particularly striking in the field of science, technology, engineering and mathematics (STEM), where women account for only 32% of French graduates in 2021<sup>1</sup>. Further upstream, [Card and Payne \[2021\]](#) show that in Canada and the US, the discrepancy between men and women in high school course selection can actually explain most of the observed difference later in majors' choice in college. This under-representation is an obstacle to gender equality on the labour market: 20 to 30% of the pay gap among higher education graduates can actually be explained by differences in occupation [[Breda et al., 2023](#)].

The under-representation of women also has perverse effects for new generations – with little opportunity to interact with women engineers or scientists, young women are turning away from these occupations thinking they are not for them – while maths-related anxiety is said to be passed on from female teachers to girls ([Beilock et al. \[2010\]](#)). According to the latest results of the PISA 2022 survey, 72% of French girls aged 15 who attend school agree with the statement “I’m afraid of getting bad marks in maths”, compared with 56% of boys.

In a cumulative process, inequality and difference in academic performances very early-on may potentially have lasting impacts on women's career choices and labour outcomes. The beginning of mandatory schooling has furthermore been shown to be a critical period for children internalising gender stereotypes [[Duru-Bellat, 2008, 2016](#)] and associating men more than women with high-level and presumed innate intellectual abilities (gender-brilliance stereotype). [Bian et al. \[2017\]](#) are interested in the acquisition of this stereotype by children: when it comes to imagining a particularly intelligent person, children make no difference according to gender at the age of 5, whereas from the age of 6 girls begin to imagine a man, and they also tend to refuse to take part in games identified as being for very intelligent people. Performance at school can ultimately reinforce the influence of such gender stereotypes through a confirmation effect. They can also directly affect pupils' self-confidence and lead to disinvestment in certain subjects in which they consider themselves to be less talented, thus contributing to the under-representation of women in science [[Napp and Breda, 2022](#)].

For mathematics, using PISA data, boys have been shown to outperform girls at the age of 15 in most OECD countries. Longitudinal studies such as [Fryer Jr and Levitt \[2010\]](#) in the US have also shown a widening of girls' disadvantage over time. Furthermore, recent works have demonstrated that those gender gaps can be highly heterogeneous, and depend for instance on family socio-economic

---

<sup>1</sup>The breakdown of new graduates by gender and by field can be consulted [on the OECD website](#)..



status or race (Penner and Paret [2008]). Using a sample of Australian primary school students, Cobb-Clark and Moschion [2017] 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.

Although gender gaps in performance are well documented, factors explaining their level and evolution do not foster consensus, especially as to their relative importance. Drawing from gender socialisation theories, subfields of the literature have looked into the influence of the family, peers, and institutions. In particular, schools and other educational institutions are critical environments for internalising gender stereotypes. Interactions between teachers and pupils or between pupils and pupils (peer effects) can influence the relative performance of girls compared with boys. Girls whose teachers (either female or male) have significant gender biases conform to their teachers' beliefs and perform less well ("self-fulfilling prophecy"). Terrier [2020] shows that having a teacher who gives better marks to girls than to boys (for equal performance) durably increases their performance and their probability of going into science. In Italy, for secondary school pupils, Carlana [2019] shows that having a teacher with implicit biases associating mathematics more with boys degrades girls' self-confidence and contributes to the gender gap in mathematics, while Lavy and Sand [2018] document the persistence of the effect of having been exposed in primary school to a biased teacher over time in Israel. Dee 2005, 2007 show that having a teacher of the same gender improves middle school performance for both girls and boys in the United States (role model effects). Finally, there is less consensus on the influence of the gendered composition of classes or schools: some studies conclude that there is a positive effect on the mathematical results of girls in non-mixed environments (see Paredes [2022] in Chile), while others point to negative effects on the choice of scientific subjects (see Brenøe and Zölitz [2020] for Denmark).

The aim of this paper is to document the gender gap in mathematics that is likely to emerge from the start of primary school – i.e. at the age of 6 for the majority of pupils – and its possible causes. We use data coming from compulsory standardised national assessments administered to more than 2.5 million pupils in first grade ("*cours préparatoire*") and second grade ("*cours élémentaire première année*") between 2018 and 2022 (see Section 2). The absence of any notable change in French (figure 1) has led us to focus mainly on the mathematics assessments, but we present some results related to French in appendix (Section 6). The ground lost by girls relative to boys in mathematics has already been shown in France by the *Conseil scientifique de l'éducation nationale* Martinot [2021] and the DEPP [2023]. Fischer and Thierry [2022] also show on survey data that boys start outperforming girls in mathematics around age 6 in France. We are continuing this body of work and attempting to understand the factors likely to push girls falling behind boys in mathematics.

The remainder of the paper is structured as follows. Section 2 describes the data as well as details the construction of scores and variables used in the analysis. Section 4 describes the evolution of

the performance of boys and girls in mathematics and reading in first grade (age 6) and how the gender gap in this evolution varies according to initial performance and family background. Section ?? presents the main results from multivariate analyses and section 5 concludes.

## 2 Data

### 2.1 Data sources

#### *Evaluations Repères CP-CE1*

Created by the French Ministry of Education in 2018, the *Évaluations Repères CP-CE1* are exhaustive standardised national assessments. These assessments are designed to detect learning difficulties and concern all first and second grade pupils enrolled in state and private schools under contract. They are carried out in class, under the guidance of the teacher, and take place in three sessions: the first at the start of first grade (September), then in January during first grade and finally in September at the start of second grade. Since 2023, primary school pupils have also been assessed at a fourth session on entering fourth grade (September). After correction, the results of the assessments are entered by the teachers and summarised.

The exercises (items) assessed do not cover the whole curriculum, but are selected to test certain skills in mathematics (*adding, solving problems...*) and French (*writing syllables, reading a text aloud...*), and for each item difficulty thresholds are defined in order to identify struggling pupils. The exercises may vary between different assessment sessions (the exercise *understanding spoken words* is only assessed at the beginning of first grade) and between cohorts (the exercise *reading invented words* only concerns pupils in first grade in 2018-2019), and grading scales are not consistent over cohorts so direct comparisons between cohorts and assessment sessions are limited.

#### *Constats de rentrée en sixième*

Information on pupils' socioeconomic background is merged for the first cohort (in first grade during 2018-2019) using another data source filled by families when enrolling children in middle schools. Note that the information is therefore retropolated as we use family information collected when pupils enter sixth grade (*Sixième*).

Two peices of information are available : the relation of each legal representative (up to three) to the pupil, and their occupation (*Profession et Catégorie Socio-professionnelle*) coded in 29 categories. From the parents' occupation categories, we reconstruct the individual SES index (*Indice de Position Sociale*<sup>2</sup>) of the legal representatives and the family SES index, using the Ministry of Education cross-tabulations<sup>3</sup>. We also reconstruct the family types shown in figure 4, based on the child's legal

---

<sup>2</sup>The IPS is a continuous index created by the Ministry of Education to describe pupils' socioeconomic background.

<sup>3</sup>Available here : <https://www.education.gouv.fr/media/158757/download>

representatives. When the child has only one declared legal representative, he or she is part of a single-parent family. Family other refers to children raised by someone other than their parents (a relative other than their parents, an educator, a family assistant, child welfare, etc.).

## APAE 1D

Administrative information on the characteristics of schools is available in the *Aide au Pilotage et à l'Auto-Évaluation du premier degré* (APAE 1D) application. In addition to the type of school (public, private) and whether or not the school belongs to a priority education network (*Réseau d'Education Prioritaire*, defined based on schools' social composition and share of underprivileged pupils), information on pupil numbers (average number of pupils per class, proportion of girls, school SES index) and teachers (proportion of women, average age, length of service in the school) is provided.

## 2.2 Construction of scores and variables

### Pupils' outcomes and gender gaps

In the following, results are standardised for each cohort and evaluation session.

To give a similar weight to every item<sup>4</sup>, items are first transformed into Z-scores. For item  $j$  and student  $i$ :

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

Provided pupils are evaluated on all items, we then aggregate grades in mathematics and French averages, and standardise once more.

$$score_i = \frac{x_i - \bar{x}}{\sigma(x)} \quad \text{with } x_i = \frac{\sum_{j=1}^n \tilde{item}_{ij}}{n} \text{ and } n \text{ the number of items evaluated}$$

Finally, we rescale average scores between 0 – 100 by computing percentile ranks.

$$rank_i = 100 \times \frac{N_i}{N - 1} \quad \text{with } N_i = \sum_{s \neq i}^N \mathbb{1}(score_s < score_i) \text{ and } N \text{ the sample size}$$

Progress outcomes are the difference between the ranking at the beginning of second grade (CE1) and first grade (CP).

$$prog_i = rank_i^{begCE1} - rank_i^{begCP}$$

Gender gaps in scores and progress are defined as the difference between boys and girls' averages. When positive, they indicate an advantage to boys and *vice versa*.

---

<sup>4</sup>Grading scales can vary quite a lot between items

## Other variables

**Teachers' gender** is inferred from a text field related to class name in Repères, that sometimes contain the name and gender titles of teachers ("*Classe de CP-A de Mme X*"). We code three modalities: male, female, unknown and use female as baseline.

**Classes and characteristics** are reconstructed at the cohort  $\times$  school level using a classroom number when available, or the class name field when not. Class characteristics depend on observed pupils, for instance the number of pupils in a class.

**Pedagogy and religion** are variables created at the school level respectively using school names and school names and type (only private schools can be religious in France) and looking for explicit mentions of alternative pedagogies ("*Montessori*", "*Freinet*" etc) or specific usual words ("*Saint*" for catholic schools).

## 3 Descriptive Analysis

### 3.1 Sample selection

For a given evaluation within a cohort (beginning and middle of first grade, beginning of second grade) between 10 and 15% of students' results are missing. We check that missingness does not appear to be related to gender or schools. When computing statistics we use all available observations for a given assignment and do not restrict the sample to pupils with no missing evaluation. In practice, this means that the students observed at the beginning, middle of first grade and beginning of second grade are not exactly the same in the gender gap figures. When defining initial percentile ranks in regressions or figures it is relative to the pupils observed at the beginning of first grade (but not necessarily afterwards). Note that our results are robust to restricting the sample to the pupils evaluated in the three assessments.

Some pupils appear to switch school during the first grade year. Because we use school fixed effects and controls in the regressions, we exclude those that switch between the beginning and middle of first grade evaluations from the sample regressed.

Descriptive statistics on students and schools attended broken down by cohort can be found in table [A1](#). There is 51 % of boys and 49 % of girls in the sample. Considering all cohorts, the average age of pupils upon entering first grade is 6 years and a little over three months. When it comes to schools attended, in the sample 13 % of pupils are schooled in private schools, 69 % in public non priority schools, and 18 % in public schools in priority education networks.

### 3.2 Gender gaps in Mathematics and French

Looking at all the cohorts from 2018 to 2022 (figure 1), there is almost no difference in the average rank between boys and girls in Mathematics when pupils enter first grade. Girls, on the other hand, are better in French (their average rank per hundred pupils is 5.5 higher on average than that of boys). The gap in Mathematics appears at the second assessment in the middle of the first grade year, in January (2.7 ranks), and widens at the last assessment. At the end of second grade, boys are on average 5.8 ranks ahead of girls, which corresponds to an average mark difference of just over 1 point out of 20. In French, the dynamic appears to be different, as the initial advantage for girls is reduced in the middle of first grade (from 5.5 to 2 ranks), before increasing again at the start of second grade (4 ranks).

Identical results were observed for each of the four cohorts studied (see figure A4 in appendix): every year, including during the years 2019-2020 and 2020-2021, which were severely affected by the Covid-19 pandemic, girls' began to lose ground in Mathematics compared to boys as early as five months after starting first grade, falling behind even further at the beginning of second grade.

Statistics on the distributions of Mathematics and French scores for both gender separately are provided in table A2 and figure A1 in appendix. Looking at the Mathematics densities, boys and girls densities quite clearly shift away from one another over time, while there is no such pattern in French. An additional finding in line with the related literature is that the distribution for boys is more spread out than for girls (more extreme performers).

### 3.3 Gender gaps by item

A potential caveat to the previous finding is that because the content of the evaluations is not the same at the beginning of first grade, middle of first grade and beginning of second grade, girls falling behind boys over time when considering the aggregate Mathematics' score could be due to a change in content rather than actual evolution of boys and girls relative ranking. To account for this concern, figure 2 shows the evolution of the rank differences between boys and girls for the eight Mathematics exercises included in at least two separate assessments. For seven of the eight exercises, the advantage for boys is very small at the start of first grade and then increases over time. This shows that the overall drop observed for girls in Mathematics is not linked to a change in the content of the assessments, but reflects a real decline in their performance compared with boys for the same type of test.

The gaps widen more or less over time in favour of boys, depending on the exercises assessed. Thus, at the beginning of second grade, the advantage for boys is of 7 ranks for the exercises *Adding* or *Placing a number on a graduated line*, but only 0.3 rank for the exercise *Reproducing an assembly*. On the other hand, the exercise *Comparing numbers* is the only one for which the trend towards an in-

crease in the gap at aggregate level is contradicted, since the advantage of 2 ranks in favour of boys already present at the beginning of first grade decreases slightly to reach 1.2 ranks in the middle of first grade (this exercise is not available in the second grade assessment).

In French, the partial catching up of boys in the middle of first grade seem to be partly linked to the introduction of new exercises measuring reading skills and cannot therefore be directly interpreted as boys gaining ground over girls. This observation, and the absence of any clear trend in French, leads us to concentrate mostly on the Mathematics assessments. We show the same type of figure for recurring French items in appendix figure A2.

### 3.4 More ground lost for girls among the best pupils

Are the differences between girls and boys greater or lesser for pupils who are having difficulty or for those who are doing better at school? To answer this question, we first compute the proportion of girls at different levels of performance and examine how this changes from one assessment to the next (panel (a) of figure 3). At the start of first grade, girls' performances are more 'centered' than boys' and girls are therefore slightly under-represented among both the worst and best pupils. The proportion of girls among the worst performers is thus lower than the proportion of girls assessed (orange line) up to the 20th percentile and from the 84th percentile upwards. From the middle of first grade, i.e. five months after the first assessment, the under-representation of girls among the best occurs earlier in the distribution, from the 70th percentile, while girls remain under-represented among the worst performers until the 4th percentile. The under-representation of girls among the best pupils becomes more pronounced at the end of the first year of elementary school. At the start of second grade, girls are under-represented in the first two percentiles of the performance distribution (the lowest-performing 2 % of pupils) but are then over-represented up to the 63rd percentile. Their representation finally decreases sharply with the level of performance at the top of the distribution, so that among the very best pupils at the start of second grade, fewer than 25 % are girls.

The phenomenon therefore seems to occur at the top of the distribution, but panel (a) of figure 3 above does not make it possible to determine precisely which girls, on the basis of their performance at the start of first grade, are losing ground. Panel (b) presents the average change in rank in Mathematics for girls minus the same change for boys as a function of the initial level of each pupil measured at the start of first grade. It shows that, whatever their starting level, girls make less progress than boys with the same starting level. This difference in progress between girls and boys increases sharply with the initial level among the lowest-performing pupils and stabilises at around 6 ranks among the half of the highest-performing pupils at the start of first grade.

We perform the same type of analysis for French in appendix figure A3. In terms of over/under-representation in the French distribution girls tend to be over-represented at the top and under-represented at the bottom of the distribution, but the dynamic between assessments is not as clear

as is the case for Mathematics. The difference in French progress between boys and girls depending on starting French level is lower (between -1 and +1 rank) than for Mathematics. While girls initially at the bottom of the distribution appear to progress less than similar boys, at the top of the initial distribution they make more progress.

### 3.5 An evolution common to all social, family and territorial contexts

Figure 4 shows the gender gaps in Mathematics according to different dimensions of heterogeneity relating to the social and family context in which the children live. This analysis is only possible for pupils entering first grade in 2018, for whom we were able to retrieve this information from the *Constat de rentrée en sixième* (see section 2).

The upper part of the figure uses families SES index (IPS), to classify pupils into four categories (very high SES, high SES, medium SES and low SES). The emergence of a rank gap in Mathematics in favour of boys does not depend on the children's social background. The magnitude of the gap is even slightly more pronounced among pupils with a more privileged social status, which is partly linked to their better initial level in Mathematics (see figure 3). In families with a very high SES, boys are on average 7.6 ranks ahead of girls, whereas this gap is just under 5.9 ranks in families with a low SES. Note that this finding is robust when considering each mothers and fathers' SES index separately as in appendix figure A5.

We then examine the same gaps according to the child's family type, since the transmission of gender norms may be influenced by family composition. Our main finding remains that the change in the gap in ranking between boys and girls is to the disadvantage of girls, whatever their family type. The gap between boys and girls widens the most among children whose two legal representatives are of the same sex (same-sex families) (+8 ranks for boys at the start of second grade). Children brought up by two parents of different sexes (heteroparental families) come next. The gap is least pronounced among single-parent families (+5.3 ranks at the start of second grade) or when the legal guardians are not the parents (+3.3 ranks).

The third part of the figure presents results according to a final dimension likely to affect differences in children's performance, namely the occupation of the child's legal representatives. Differences are measured within families where both parents have the same occupation (both parents teachers, engineers, unskilled workers, or employees). Here again, the trend in the gaps between boys and girls can be seen in families where the representatives have very different occupations, some of which are directly linked to mastery of the foundations of Mathematics: the gap at the start of second grade is +6.8 ranks in families where both representatives are engineers and +7.1 ranks when both parents are teachers). The analysis for all the occupation categories of the mother or father of children is presented in appendix figures A7 and A6: whatever the category considered, the evolution of the Mathematics gap is always to the disadvantage of girls.



Furthermore, we do not observe any major territorial difference depending on the child's department of schooling (figure 5): girls lose on average between 2 and 7 ranks compared to boys. Nonetheless, girls appear to lose more ground on average in mainland France than in the overseas départements.

## 4 Main Results

These characteristics studied in the previous section can play together on girls' lower relative progress in math relative to boys. In order to better understand which factors influence this relative progress, we present three separate Tables focusing on the influence of school, class and individual characteristics, respectively. In each Table, we progressively add controls for the student, class or school characteristics that are not the primary focus of the Table. This approach helps us isolate the main drivers (if any) that might be behind girls' lower relative progress. The analysis remains however purely descriptive and no causal statement should be made from the results.

### 4.1 School characteristics

We start with a study of school characteristics: does the school context attenuate or, on the contrary, reinforce the dynamic of girls decline in relative ranking compared with boys in Mathematics? Are there characteristics that help to limit the drop?

Table 1 presents the results. The reference group is a public school, not in a "priority education area" (*Réseau d'éducation prioritaire*), with no indication of extra religious activities, with a standard pedagogy and an average size and with students that are at the average in terms of socioeconomic background. The Table also systematically includes an indicator variable for students' gender. It therefore examines how varies the relative progress of girls and boys in math when schools depart from this reference group in one dimension. In column 1, we see that girls lose less grounds relative to boys in priority education areas and in larger schools. Column (2) adds School\*Cohort fixed effects to control for "school quality", ie. the average progress of students in math within each school (independent of their gender). Such specification allows us to isolate the specific effect of school characteristics on girls relative to boys, holding fixed their effects average students' progress. Results remain broadly unchanged. In column (3) and (4), we further control for individual characteristics (age and initial level in math and reading as well as their interaction with students' gender) and classroom characteristics also interacted with gender. Finally, column (5) controls more finely for students' socioeconomic background and type of family (and their interaction with gender) using only the first cohort for which these pieces of information are available. The attenuated loss for girls in priority schools becomes smaller in magnitude but remains significant. The effect of being a larger schools also remains.



The main take away, however, from results in Table 1 is that school characteristics have little influence to explain girls' lower progress in math. Indeed, even those that have a significant influence—in part due to our very large sample size—are only able to “cancel” a small fraction of the average girls' dropout of around 7 ranks. The more striking result is perhaps that girls lose grounds relative to boys to a similar extent in schools that employ alternative pedagogy or teaching methods (such as Freinet or Montessori) that are usually considered more progressive and inclusive.

## 4.2 Classroom characteristics

Table 2 examines in a similar fashion the effect of classroom characteristics. Column (1), which does not include any control variables, shows for example that girls lose more ground in classes taught by a man, a result that is consistent with the findings of the literature on teachers' role models. In column (2), we include classroom\*cohort fixed effects and school\*cohort\*girls fixed effects. The former control for the average progress in math within each classroom, while the latter hold constant the relative progress of girls versus boys within each school and cohort. This implies that the specifications in column (2) and following columns examine the relative progress of girls (versus boys) within each class relative to their relative progress within the school. Column (3) further control for individual characteristics, and column (4) for students' socioeconomic background using the first cohort only.

The only effects that remain significant in the most demanding specifications are those related to peer effects: girls lose less ground (i) when the average level in math is initially larger, (ii) when the top performer in the class is a girl, or (iii) when other girls in the class perform better. As for school characteristics, these effects remain however limited and are far from being able to cancel girls lower progress in math in Grade 1.

## 4.3 Students characteristics

Table 3 finally examines if girls lower relative progress in math varies according to other students' characteristics. All models include Classroom\*cohort\*sex fixed effects so that these effects are examined beyond those of school and classroom characteristics on girls' relative progress.

In the three first columns, we use all four cohorts available and examine the influence of students' initial level in math and reading and progress in reading. The objective is to check whether girls' lower progress could be driven by gender differences observed in reading. For example, girls may progress less in math because they progress more in French. We first see that students who are initially better in French progress and/or progress more in French also progress more in math. This is true no matter their gender, and the positive effect of the initial level in French is even slightly larger for girls, meaning that being good in French slightly protects girls against falling behind in math. This result goes against the idea of a strong specialization, all the more that when they progress in

French, girls are not more likely than boys to fall behind in math. A more direct examination reveals that the average relative loss of girls in math performance during Grade 1 is larger (reaching 9 ranks) when controlling for the initial French performance in reading (Table A3). Hence girls better average initial level in reading cannot explain that they fall back in math.

In columns (4) to (7) of the Table, we further examine the role of socioeconomic background and family characteristics using the first cohort only. We see that once controls for other individual characteristics are included, parents' socioeconomic status (using an average of both parents or considering the mother and the father separately) has no or a very small positive effect on girls' relative progress in math. Hence socioeconomic background seems to play no role in examining the general ground lost by girls in math in Grade 1.

We try to challenge this conclusion in Table A4 by isolating parental occupations that might be the most likely to "protect girls" against falling behind in math. We include in these occupations teachers, professors, science-related jobs and engineers and consider 3 cases: only the mother holds one of such jobs, only the father does, both parents do. None of these three configurations are directly beneficial to girls relative progress in math (column (2)).

A last take-away from Table 3 is that being older slightly protects girls against falling behind in math, although, once again, the effect is small in magnitude. These insights are overall confirmed by a specific analysis of students in the top half of the math initial performance distribution (columns (8) to (14)).

## 5 Conclusion

In this paper, we analyse the emergence of the gender gap in mathematics using French exhaustive administrative data covering four cohorts of pupils attending first and second grades from 2018 to 2022. While there are no differences between boys and girls upon entry to school, girls lose 3 ranks compared to boys after 5 months of schooling. This gender gap persists through the first year of schooling, reaching 6 ranks at the beginning of second grade. The overall drop observed for girls is not linked to change in the content of the assessments, but reflects a real decline in their performance compared with boys for the same type of test.

During the Covid-19 pandemic years, pupils less attended school but this did not modify the dynamics observed for years outside the pandemic: the dropout of girls in math as early as five months after starting CP, and its continuation until the start of CE1, are of a similar scale across all four cohorts studied since 2018.

We document the fact the girls are losing ground among the best pupils in maths in first grade. However this dynamics is not correlated with a specialisation of girls in French.

This increasing gender gap in maths is visible across every strata of society. Thanks to the richness of data we used, we were able to add a number of variables to control for potentially important drivers of school achievement. We thus find the influence of factors already identified in the literature as potential mediators (the size of the class, the gender of the teacher, the proportion of girls in the class, for example). However, this influence remains quantitatively too limited to be able to significantly counteract the fact that girls are losing ground in maths relative to boys during the first year of primary school.

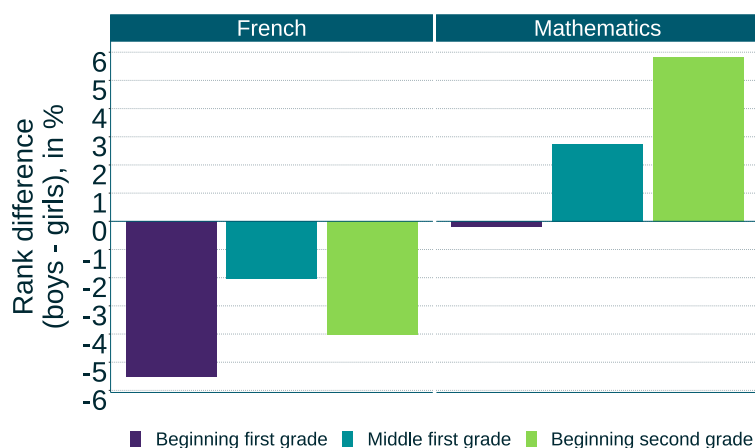
## References

- BEILOCK, S. L., E. A. GUNDERSON, G. RAMIREZ, AND S. C. LEVINE (2010): "Female teachers' math anxiety affects girls' math achievement," *Proceedings of the National Academy of Sciences*, 107, 1860–1863.
- BIAN, L., S.-J. LESLIE, AND A. CIMPIAN (2017): "Gender stereotypes about intellectual ability emerge early and influence children's interests," *Science*, 355, 389–391.
- BREDA, T., J. GRENET, M. MONNET, AND C. VAN EFFENTERRE (2023): "How Effective are Female Role Models in Steering Girls Towards STEM? Evidence from French High Schools," *The Economic Journal*, 133, 1773–1809.
- BREDA, T., E. JOUINI, C. NAPP, AND G. THEBAULT (2020): "Gender stereotypes can explain the gender-equality paradox," *Proceedings of the National Academy of Sciences*, 117.
- BREDA, T. AND C. NAPP (2019): "Girls' comparative advantage in reading can largely explain the gender gap in math-related fields," *Proceedings of the National Academy of Sciences*, 116, 15435–15440.
- BRENØE, A. A. AND U. ZÖLITZ (2020): "Exposure to more female peers widens the gender gap in STEM participation," *Journal of Labor Economics*, 38, 1009–1054.
- CARD, D. AND A. A. PAYNE (2021): "High school choices and the gender gap in STEM," *Economic Inquiry*, 59, 9–28.
- CARLANA, M. (2019): "Implicit stereotypes: Evidence from teachers' gender bias," *The Quarterly Journal of Economics*, 134, 1163–1224.
- COBB-CLARK, D. A. AND J. MOSCHION (2017): "Gender gaps in early educational achievement," *Journal of Population Economics*, 30, 1093–1134.
- DEE, T. S. (2005): "A teacher like me: Does race, ethnicity, or gender matter?" *American Economic Review*, 95, 158–165.
- (2007): "Teachers and the gender gaps in student achievement," *Journal of Human resources*, 42, 528–554.
- DEPP (2023): "Filles et garçons sur le chemin de l'égalité. De l'école à l'enseignement supérieur," Tech. rep., DEPP.
- DURU-BELLAT, M. (2008): "La (re) production des rapports sociaux de sexe: quelle place pour l'institution scolaire?" *Travail! genre et sociétés*, 131–149.
- (2016): "À l'école du genre," *Enfances & Psy*, 90–100.
- FISCHER, J.-P. AND X. THIERRY (2022): "Boy's math performance, compared to girls', jumps at age 6 (in the ELFE's data at least)," *British Journal of Developmental Psychology*, 40, 504–519.
- FRYER JR, R. G. AND S. D. LEVITT (2010): "An empirical analysis of the gender gap in mathematics," *American Economic Journal: Applied Economics*, 2, 210–240.
- LAVY, V. AND E. SAND (2018): "On the origins of gender gaps in human capital: Short-and long-term consequences of teachers' biases," *Journal of Public Economics*, 167, 263–279.

- MARTINOT, PAULINE ET DEHAENE, S. E. B. P. E. H. P. E. P.-W. C. E. S.-C. L. E. Z. J. (2021): "Qu'apprend-on des évaluations de CP-CE1?" Tech. rep., CSEN.
- NAPP, C. AND T. BRED A (2022): "The stereotype that girls lack talent: A worldwide investigation," *Science Advances*, 8, eabm3689.
- PAREDES, V. (2022): "Mixed but not scrambled: Gender gaps in coed schools with single-sex classrooms," *Journal of Research on Educational Effectiveness*, 15, 330–366.
- PENNER, A. M. AND M. PARET (2008): "Gender differences in mathematics achievement: Exploring the early grades and the extremes," *Social Science Research*, 37, 239–253.
- TERRIER, C. (2020): "Boys lag behind: How teachers' gender biases affect student achievement," *Economics of Education Review*, 77, 101981.

## 6 Figures and Tables

Figure 1: Gender gaps over time

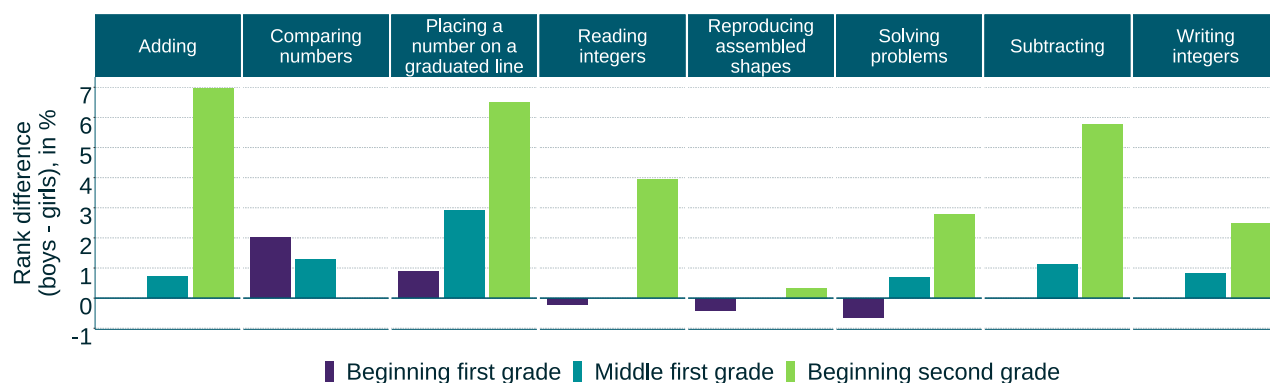


Notes: A positive gap means that boys outperform girls and *vice versa*.

Sample: Students in first grade between 2018 and 2022.

Sources: MENJ-DEPP, Repères first grade-CE1; Authors' computations.

Figure 2: Gender gaps for recurring Mathematics items

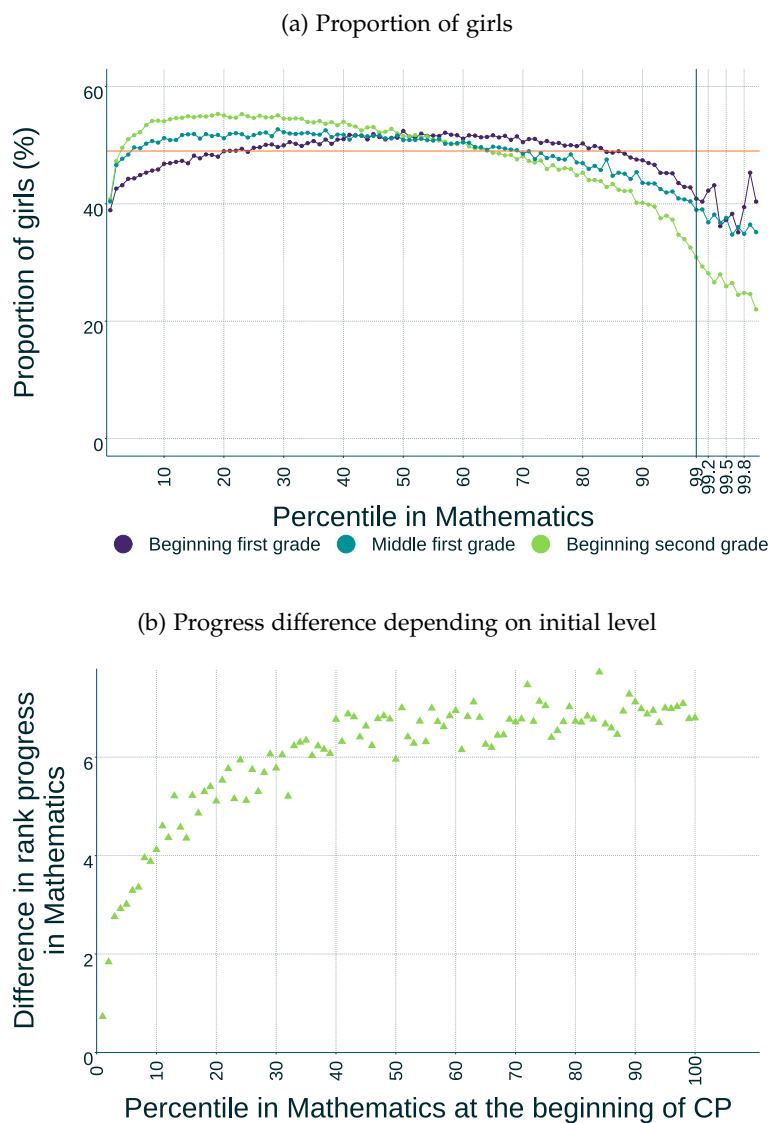


Notes: A positive gap means that boys outperform girls and *vice versa*. Items presented are those that are evaluated in more than one assessment (per cohort). The item *Placing a number on a graduated line* actually pools together the results of two similar items (*Placing a number on a graduated line* or *numerical line*).

Sample: Students in first grade between 2018 and 2022.

Sources: MENJ-DEPP, Repères first grade-CE1; Authors' computations.

Figure 3: Gender gap dynamics in the Mathematics distribution



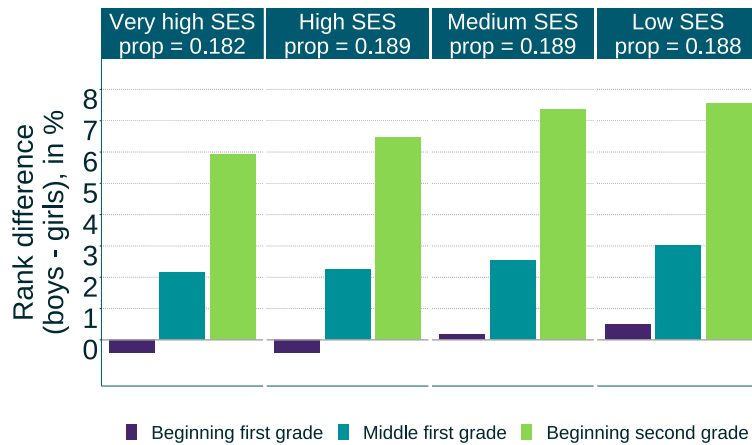
*Notes:* We rank student by score percentile from the smallest (the 1 % with the worst results) to the highest (the 1 % with the best results). Figure a) also includes a focus on the top 1 % of students. In figure b) the initial Mathematics percentile (x-axis) is defined relatively to the beginning of first grade evaluation.

*Sample:* Students in first grade between 2018 and 2022.

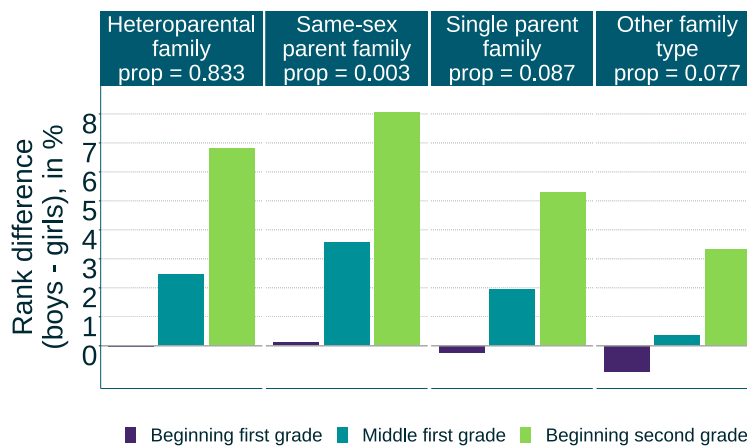
*Sources:* MENJ-DEPP, Repères first grade-CE1; Authors' computations.

Figure 4: Gender gaps depending on family background

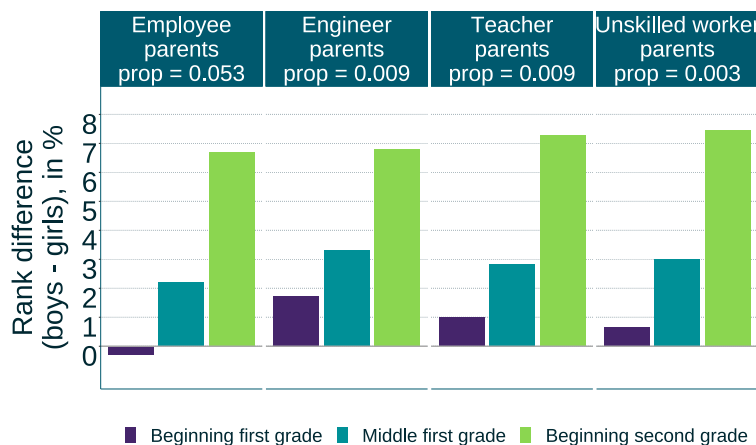
(a) Gender gap by family socioeconomic background



(b) Gender gap by family type



(c) Gender gap by parents' profession



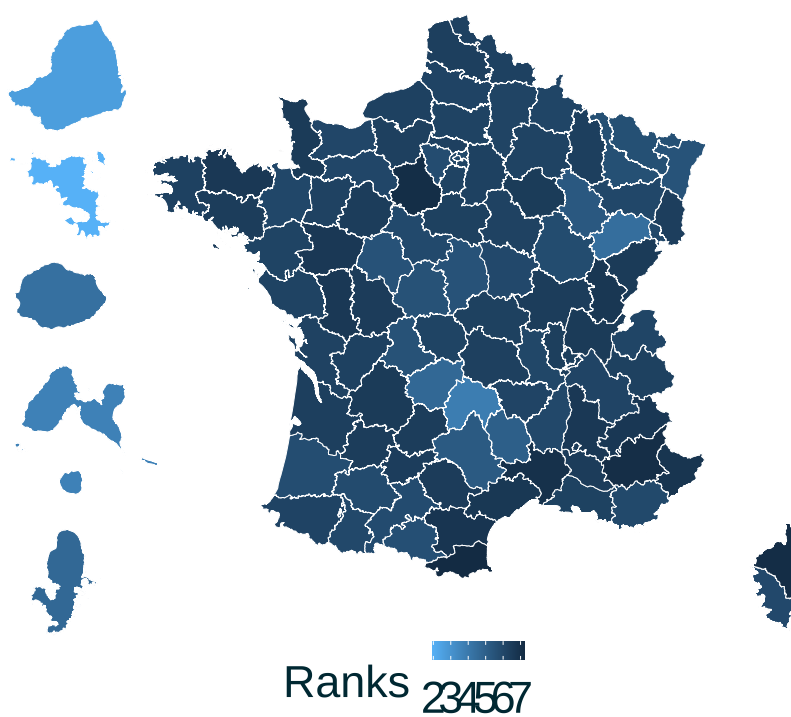
Notes: A positive gap means that boys outperform girls and *vice versa*. The proportion of the sample within each category is indicated. When it does not add up to 1, it means that some categories (including missing observations) are not shown.

Sample: Students in first grade between 2018 and 2022.

Sources: MENJ-DEPP, Repères first grade-CE1 & Constats de rentrée en sixième; Authors' computations.



Figure 5: Ranks lost by girls depending on departement



*Notes:* Overseas departments (not true to scale) are placed on the left of mainland France.

*Sample:* Students in first grade between 2018 and 2022.

*Sources:* MENJ-DEPP, Repères first grade-CE1 & APAE1D; Authors' computations.

Table 1: School characteristics and girls' mathematics progress

Mathematics initial level	Dependent Variable: Mathematics progress									
	all students					top 50%				
	all cohorts				first cohort	all cohorts				first cohort
Cohort	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
girl × school type: Public-priority	0.8489*** (0.1109)	0.7819*** (0.1061)	0.7443*** (0.0861)	0.4015*** (0.1043)	0.4420* (0.2359)	1.057*** (0.1706)	1.030*** (0.1679)	0.9510*** (0.1412)	0.6330*** (0.1624)	0.3282 (0.3856)
girl × school type: Public-priority+	1.064*** (0.1429)	1.134*** (0.1368)	1.194*** (0.1136)	0.9019*** (0.1281)	0.7897*** (0.2941)	0.9659*** (0.2409)	1.323*** (0.2333)	1.433*** (0.1997)	1.176*** (0.2158)	1.206** (0.5136)
girl × school type: Private	0.2405 (0.1493)	0.0738 (0.1413)	0.1322 (0.1099)	0.0847 (0.1122)	0.2968 (0.2711)	0.0726 (0.1761)	0.0193 (0.1690)	0.1482 (0.1441)	0.1277 (0.1480)	0.3926 (0.3662)
girl × school size	0.0748*** (0.0280)	0.0906*** (0.0270)	0.1615*** (0.0220)	0.1906*** (0.0246)	0.2170*** (0.0559)	0.0465 (0.0362)	0.0642* (0.0353)	0.1058*** (0.0298)	0.1206*** (0.0330)	0.2076** (0.0817)
girl × school SES index	0.0043* (0.0023)	0.0045** (0.0022)	-0.0077*** (0.0018)	-0.0148*** (0.0019)	-0.0267*** (0.0044)	0.0246*** (0.0030)	0.0232*** (0.0029)	-0.0025 (0.0025)	-0.0068** (0.0027)	-0.0251*** (0.0063)
girl × school pedagogy: alternative	-0.6918 (1.086)	-0.9306 (1.096)	-1.104 (0.8036)	-1.070 (0.8099)	0.5895 (1.973)	-2.009 (1.616)	-1.995 (1.630)	-1.792 (1.169)	-1.839 (1.164)	-1.204 (3.304)
girl × school religion: yes	-0.2290 (0.1777)	-0.0856 (0.1676)	-0.1617 (0.1308)	-0.1551 (0.1314)	-0.3614 (0.3195)	0.0101 (0.2071)	0.0402 (0.1990)	-0.1309 (0.1696)	-0.1007 (0.1710)	-0.4941 (0.4231)
<i>Controls</i>										
individual			Yes	Yes	Yes			Yes	Yes	Yes
classroom				Yes	Yes				Yes	Yes
family					Yes					Yes
<i>Fixed-effects</i>										
cohort-school_id		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
cohort-sex		Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
<i>Fit statistics</i>										
Observations	2,584,130	2,584,130	2,502,106	2,455,155	528,114	1,321,419	1,321,419	1,290,005	1,264,630	264,719
R <sup>2</sup>	0.02025	0.17910	0.51319	0.51485	0.52633	0.02792	0.22853	0.47227	0.47387	0.49791
Within R <sup>2</sup>		5.76 × 10 <sup>-5</sup>	0.40585	0.40883	0.40732		0.00013	0.31446	0.31820	0.32023
Adjusted R <sup>2</sup>	0.02025	0.14208	0.49052	0.49263	0.50393	0.02791	0.15855	0.42325	0.42564	0.44905
Within Adjusted R <sup>2</sup>		5.39 × 10 <sup>-5</sup>	0.40585	0.40882	0.40708		0.00012	0.31445	0.31818	0.31966

*Notes:* Clustered (cohort × school\_id) standard-errors in parentheses. Signif. Codes: \*\*\*, 0.01, \*\*, 0.05, \*, 0.1. Controls are included both individually and interacted with gender. Individual controls include: pupils' age (in months) their initial Mathematics and French score (beginning of first grade) and progress in French between the beginning of first grade and CE1; classroom controls include: the gender of teacher (female as baseline), class size, whether the class is simple or multilevel (simple level as baseline), the proportion of girls, the class initial (beginning of first grade) level, difference between boys and girls and gender of top performer in Mathematics; family controls include: the family SES index (IPS), occupations (PCS) of parents and type of family (heteroparental, single sex, single parent, other). All continuous explanatory variables are centered to the mean.

*Sample:* Students in first grade between 2018 and 2022.

*Sources:* MENJ-DEPP, Repères first grade-CE1 & APAE1D; Authors' computations.

Table 2: Class characteristics and girls' mathematics progress

Mathematics initial level	Dependent Variable: Mathematics progress							
	all students				top 50%			
	all cohorts		first cohort		all cohorts		first cohort	
Cohort	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
girl $\times$ teacher gender: male	-0.7631*** (0.1776)	-0.4968** (0.2361)	-0.0839 (0.2021)	0.1528 (0.4308)	-0.4349* (0.2393)	-0.4278 (0.3693)	-0.0541 (0.3236)	0.1033 (0.7265)
girl $\times$ teacher gender: uncertain	0.0565 (0.0663)	-0.2486 (0.2877)	-0.2295 (0.2528)	-0.0922 (0.5509)	0.0230 (0.0875)	-0.1789 (0.4382)	-0.2024 (0.3918)	-0.4247 (0.8783)
girl $\times$ class size	-0.0710*** (0.0062)	0.0069 (0.0162)	-0.0182 (0.0135)	0.0005 (0.0293)	-0.0291*** (0.0083)	0.0085 (0.0223)	-0.0134 (0.0198)	-0.0129 (0.0443)
girl $\times$ class multilevel: yes	-0.8161*** (0.0881)	4.956*** (1.845)	3.389** (1.623)	-1.699 (3.133)	-0.5199*** (0.1195)	5.701** (2.580)	3.702 (2.677)	0.0601 (4.311)
girl $\times$ class multilevel: uncertain	-0.1880** (0.0760)	5.546*** (1.853)	3.976** (1.628)	-1.558 (3.146)	-0.2095** (0.1014)	5.980** (2.590)	3.981 (2.686)	0.4684 (4.329)
girl $\times$ class girls proportion	4.139*** (0.3011)	0.1773 (0.5301)	0.2800 (0.4562)	-0.7823 (1.049)	2.686*** (0.3887)	1.299 (0.8212)	0.6515 (0.7236)	0.8707 (1.704)
girl $\times$ class initial maths level†	0.0081*** (0.0022)	0.0042 (0.0040)	0.0321*** (0.0036)	0.0326*** (0.0079)	0.0247*** (0.0030)	0.0089 (0.0066)	0.0111* (0.0064)	0.0127 (0.0145)
girl $\times$ class initial maths top performer: girl	-4.582*** (0.0670)	0.1169 (0.0924)	0.3299*** (0.0790)	0.5803*** (0.1816)	-3.267*** (0.0856)	0.0676 (0.1363)	0.2775** (0.1191)	0.4979* (0.2750)
girl $\times$ class initial maths gender gap†	-0.0653*** (0.0023)	0.0168*** (0.0031)	0.0085*** (0.0026)	0.0197*** (0.0059)	-0.0475*** (0.0029)	0.0079* (0.0047)	-0.0014 (0.0041)	0.0062 (0.0098)
<i>Controls</i>								
individual			Yes	Yes			Yes	Yes
family				Yes				Yes
<i>Fixed-effects</i>								
cohort $\times$ classroom_id		Yes	Yes	Yes		Yes	Yes	Yes
cohort $\times$ sex		Yes	Yes	Yes		Yes	Yes	Yes
cohort $\times$ school_id $\times$ sex		Yes	Yes	Yes		Yes	Yes	Yes
<i>Fit statistics</i>								
Observations	2,675,192	2,675,192	2,591,822	557,941	1,372,895	1,372,895	1,340,828	280,958
R <sup>2</sup>	0.04868	0.36909	0.56593	0.57635	0.03030	0.40679	0.56852	0.59012
Within R <sup>2</sup>		0.12008	0.39159	0.38686		0.04458	0.30068	0.30057
Adjusted R <sup>2</sup>	0.04868	0.24819	0.47967	0.49169	0.03029	0.14825	0.37483	0.39870
Within Adjusted R <sup>2</sup>		0.12007	0.39158	0.38662		0.04457	0.30067	0.29989

Notes: Clustered (cohort  $\times$  school\_id) standard-errors in parentheses. Signif. Codes: \*\*\*, 0.01, \*\*, 0.05, \*, 0.1. Class variables are reconstructed using a class number. † variables are computed for each student by excluding their own observation. Controls are included both individually and interacted with gender. Individual controls include: pupils' age (in months) their initial Mathematics and French score (beginning of first grade) and progress in French between the beginning of first grade and CE1; family controls include: the family SES index (IPS), occupations (PCS) of parents and type of family (heteroparental, single sex, single parent, other). All continuous explanatory variables are centered to the mean.

Sample: Students in first grade between 2018 and 2022.

Sources: MENJ-DEPP, Repères first grade-CE1 & APAE1D; Authors' computations.

Table 3: Specialisation, family characteristics and girls' mathematics progress

Mathematics initial level	Dependent Variable: Mathematics progress													
	all students							top 50%						
	all cohorts			first cohort				all cohorts			first cohort			
Cohort	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
age	-0.4410*** (0.0064)	-0.3544*** (0.0065)	-0.2316*** (0.0054)			-0.2283*** (0.0119)	-0.2265*** (0.0119)	-0.3422*** (0.0097)	-0.2582*** (0.0099)	-0.1750*** (0.0085)			-0.1961*** (0.0190)	-0.1937*** (0.0190)
initial maths level	-0.5664*** (0.0010)	-0.2984*** (0.0007)	-0.6286*** (0.0009)			-0.6341*** (0.0020)	-0.6342*** (0.0020)	-0.5648*** (0.0020)	-0.3453*** (0.0020)	-0.6213*** (0.0019)			-0.6144*** (0.0044)	-0.6146*** (0.0044)
initial french level	0.3895*** (0.0010)		0.5445*** (0.0010)			0.5803*** (0.0022)	0.5794*** (0.0022)	0.3543*** (0.0014)		0.4960*** (0.0015)			0.5301*** (0.0033)	0.5290*** (0.0033)
girl × age	0.1714*** (0.0089)	0.2219*** (0.0090)	0.1528*** (0.0076)			0.1858*** (0.0166)	0.1855*** (0.0166)	0.1674*** (0.0137)	0.2728*** (0.0142)	0.1580*** (0.0120)			0.2044*** (0.0271)	0.2037*** (0.0271)
girl × initial maths level	-0.0598*** (0.0013)	-0.0355*** (0.0010)	-0.0595*** (0.0013)			-0.0679*** (0.0028)	-0.0679*** (0.0028)	-0.0209*** (0.0029)	-0.0034 (0.0029)	-0.0248*** (0.0028)			-0.0350*** (0.0063)	-0.0349*** (0.0063)
girl × initial french level	0.0304*** (0.0014)		0.0305*** (0.0014)			0.0085*** (0.0030)	0.0087*** (0.0030)	0.0798*** (0.0021)		0.0982*** (0.0021)			0.0820*** (0.0046)	0.0823*** (0.0047)
progress in french		0.1996*** (0.0010)	0.3858*** (0.0010)			0.3728*** (0.0022)	0.3723*** (0.0022)		0.1587*** (0.0014)	0.3398*** (0.0014)			0.3354*** (0.0032)	0.3349*** (0.0032)
girl × progress in french		-0.0019 (0.0014)	-0.0006 (0.0013)			-0.0103*** (0.0029)	-0.0102*** (0.0029)		0.0285*** (0.0020)	0.0431*** (0.0020)			0.0359*** (0.0045)	0.0361*** (0.0045)
SES index family				0.0260*** (0.0016)		0.0052*** (0.0013)					0.0640*** (0.0021)		0.0156*** (0.0019)	
girl × SES index family				-1.94 × 10 <sup>-5</sup> (0.0022)		0.0082*** (0.0018)					0.0083*** (0.0031)		0.0053** (0.0027)	
SES index mother					0.0173*** (0.0019)		0.0013 (0.0016)					0.0450*** (0.0025)		0.0076*** (0.0022)
SES index father					0.0239*** (0.0019)		0.0102*** (0.0016)					0.0491*** (0.0025)		0.0173*** (0.0022)
girl × SES index mother					-0.0054** (0.0026)		0.0037* (0.0021)					0.0002 (0.0036)		0.0015 (0.0031)
girl × SES index father					0.0021 (0.0028)		0.0060*** (0.0022)					0.0090** (0.0036)		0.0037 (0.0032)
<i>Fixed-effects</i>														
cohort × classroom_id × sex	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>														
Observations	2,653,582	2,620,211	2,620,211	584,195	584,195	564,159	564,159	1,370,208	1,356,969	1,356,969	292,484	292,484	284,674	284,674
R <sup>2</sup>	0.50598	0.45048	0.58366	0.33272	0.33293	0.59174	0.59182	0.53134	0.47608	0.59721	0.45155	0.45231	0.61557	0.61567
Within R <sup>2</sup>	0.28047	0.19758	0.39206	0.00111	0.00142	0.38494	0.38505	0.18963	0.09153	0.30157	0.00904	0.01041	0.30016	0.30035
Adjusted R <sup>2</sup>	0.41710	0.35022	0.50771	0.21756	0.21780	0.51869	0.51878	0.35870	0.28115	0.44735	0.25163	0.25266	0.47180	0.47194
Within Adjusted R <sup>2</sup>	0.28047	0.19758	0.39206	0.00110	0.00141	0.38492	0.38504	0.18962	0.09153	0.30157	0.00903	0.01039	0.30013	0.30031

Notes: Clustered (cohort × school\_id) standard-errors in parentheses. Signif. Codes: \*\*\*, 0.01, \*\*, 0.05, \*, 0.1. All continuous explanatory variables are centered to the mean.

Sample: Students in first grade between 2018 and 2022.

Sources: MENJ-DEPP, Repères first grade-CE1, APAE1D & Constats de rentrée en sixième; Authors' computations.

# A Appendix

## A.1 Descriptive statistics

Table A1: Students and schools attended

Variable	Statistic	First cohort (2018-2019)	Second cohort (2019-2020)	Third cohort (2020-2021)	Fourth cohort (2021-2022)	All cohorts
<b>Individual characteristics</b>						
sex	missing	0	0	0	0	0
	prop. summary	boys: 0.51 girls: 0.49	boys: 0.51 girls: 0.49	boys: 0.51 girls: 0.49	boys: 0.51 girls: 0.49	boys: 0.51 girls: 0.49
age	missing	0	0	0	0	0
	mean	75.9	75.87	75.92	75.91	75.9
	sd	3.9	3.87	3.83	3.79	3.85
	range	[ 14 : 212 ]	[ 26 : 241 ]	[ 16 : 253 ]	[ 12 : 207 ]	[ 12 : 253 ]
<b>School characteristics</b>						
school type	missing	0	0	0	0	0
	prop. summary	Private: 0.13 Public: 0.7 Public - priority: 0.11 Public - priority+: 0.07	Private: 0.13 Public: 0.69 Public - priority: 0.11 Public - priority+: 0.07	Private: 0.13 Public: 0.69 Public - priority: 0.11 Public - priority+: 0.07	Private: 0.13 Public: 0.69 Public - priority: 0.11 Public - priority+: 0.07	Private: 0.13 Public: 0.69 Public - priority: 0.11 Public - priority+: 0.07
school religion	missing	0	0	0	0	0
	prop. summary	0: 0.94 1: 0.06	0: 0.93 1: 0.07	0: 0.92 1: 0.08	0: 0.92 1: 0.08	0: 0.93 1: 0.07
school pedagogy	missing	0	0	0	0	0
	prop. summary	0: 1 1: 0	0: 1 1: 0	0: 1 1: 0	0: 1 1: 0	0: 1 1: 0
school class size	missing	0.01	0.01	0.01	0	0
	mean	23.09	22.68	22.4	22.19	22.59
	sd	3.42	3.75	3.68	3.62	3.64
	range	[ 2 : 143.33 ]	[ 2.86 : 165 ]	[ 2 : 100 ]	[ 2 : 97.14 ]	[ 2 : 165 ]
pupils prop. female	missing	0.01	0.01	0.01	0	0
	mean	0.49	0.49	0.49	0.49	0.49
	sd	0.04	0.04	0.04	0.04	0.04
	range	[ 0 : 1 ]	[ 0 : 1 ]	[ 0.04 : 1 ]	[ 0 : 1 ]	[ 0 : 1 ]
school SES index	missing	0.05	0.06	0.05	0.05	0.05
	mean	102.49	102.45	102.37	102.32	102.41
	sd	18.52	18.54	18.55	18.56	18.54
	range	[ 43.32 : 157.6 ]	[ 47.41 : 156.5 ]	[ 47.94 : 155.09 ]	[ 49.56 : 155.59 ]	[ 43.32 : 157.6 ]
teachers prop. female	missing	0	0.01	0.01	0.01	0.01
	mean	0.86	0.86	0.86	0.86	0.86
	sd	0.13	0.13	0.13	0.13	0.13
	range	[ 0 : 1 ]	[ 0 : 1 ]	[ 0 : 1 ]	[ 0 : 1 ]	[ 0 : 1 ]
teachers' age	missing	0	0.01	0.01	0.01	0.01
	mean	41.61	41.9	42.14	42.3	41.99
	sd	4.41	4.44	4.44	4.45	4.44
	range	[ 22 : 63 ]	[ 23 : 64 ]	[ 23 : 63 ]	[ 22.5 : 64 ]	[ 22 : 64 ]
teachers' seniority	missing	0	0.01	0.01	0.01	0.01
	mean	6.15	6.29	6.5	6.63	6.39
	sd	2.82	2.89	2.93	2.96	2.91
	range	[ 0.1 : 37.1 ]	[ 0 : 37.2 ]	[ 0 : 32.2 ]	[ 0 : 33.2 ]	[ 0 : 37.2 ]

Notes: School, pupils and teacher variables are averages (at the school level).

Sample: Students in CP between 2018 and 2022.

Sources: MENJ-DEPP, Repères CP-CE1; Authors' computations.

Table A2: Score distribution statistics by gender

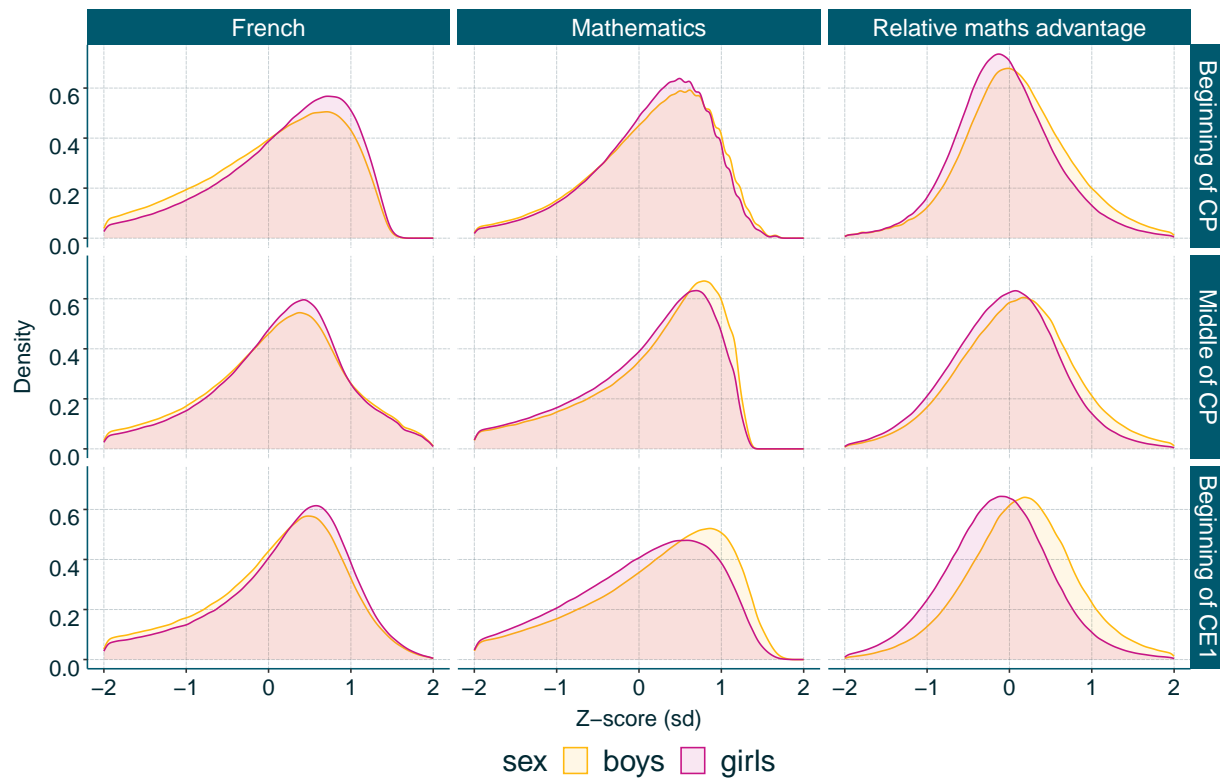
	French						Mathematics					
	Beginning first grade		Middle first grade		Beginning second grade		Beginning first grade		Middle first grade		Beginning second grade	
	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls	Boys	Girls
Mean	47.3	52.81	49	51.04	48.03	52.05	49.9	50.1	51.34	48.59	52.86	47.03
Min	0	0	0	0	0	0	0	0	0	0	0	0
Max	99.99	99.99	99.99	99.99	100	100	99.93	99.93	99.79	99.79	99.99	99.99
SD	29.02	28.44	29.44	28.22	28.91	28.68	29.56	28.12	29.42	28.19	29.57	27.81
Q10	8.3	12.39	8.72	11.54	8.82	11.46	9.1	11.01	9.9	10.1	10.39	9.67
Q25	21.75	28.85	23.11	27.02	22.83	27.49	23.85	26.16	25.71	24.34	27.32	23.09
Q75	72.21	77.52	74.76	75.22	72.73	77.05	75.95	74.06	77.32	72.47	79.1	70.24
Q90	88.61	91.16	90.27	89.71	89.04	90.82	90.94	88.93	91.38	88.21	92.31	86.58
N	1505893	1445475	1452257	1394754	1575740	1519471	1522391	1459800	1466636	1407817	1580979	1522953
Prop.NA	0.13	0.12	0.16	0.15	0.09	0.08	0.12	0.11	0.15	0.15	0.09	0.08

Notes: Statistics are expressed in percentile ranks.

Sample: Students in CP between 2018 and 2022.

Sources: MENJ-DEPP, Repères CP-CE1; Authors' computations.

Figure A1: Densities by gender



Notes: Gender densities shown use Z-scores, and not corresponding percentile ranks. The relative Mathematics advantage is for a given student and assessment the difference between her Mathematics and French Z-score.

Sample: Students in CP between 2018 and 2022.

Sources: MENJ-DEPP, Repères CP-CE1; Authors' computations.

## A.2 Supplementary results

### A.2.1 French

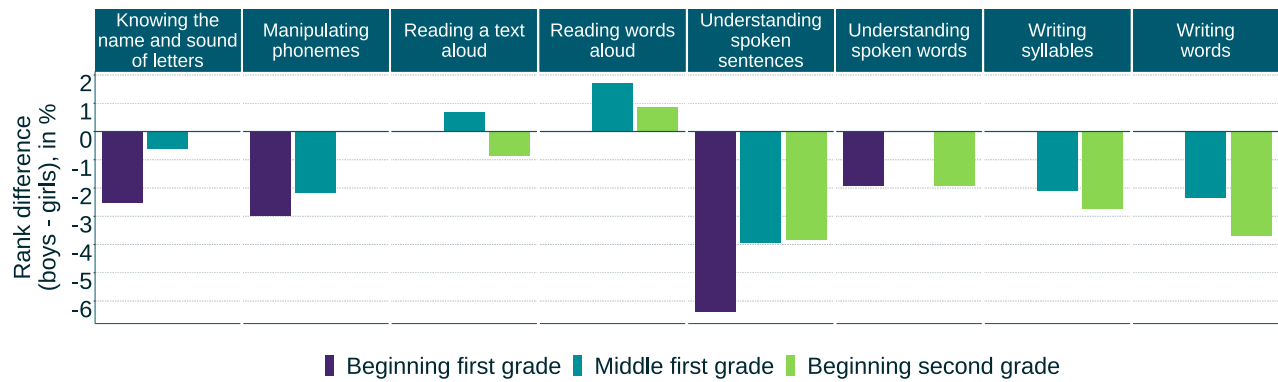


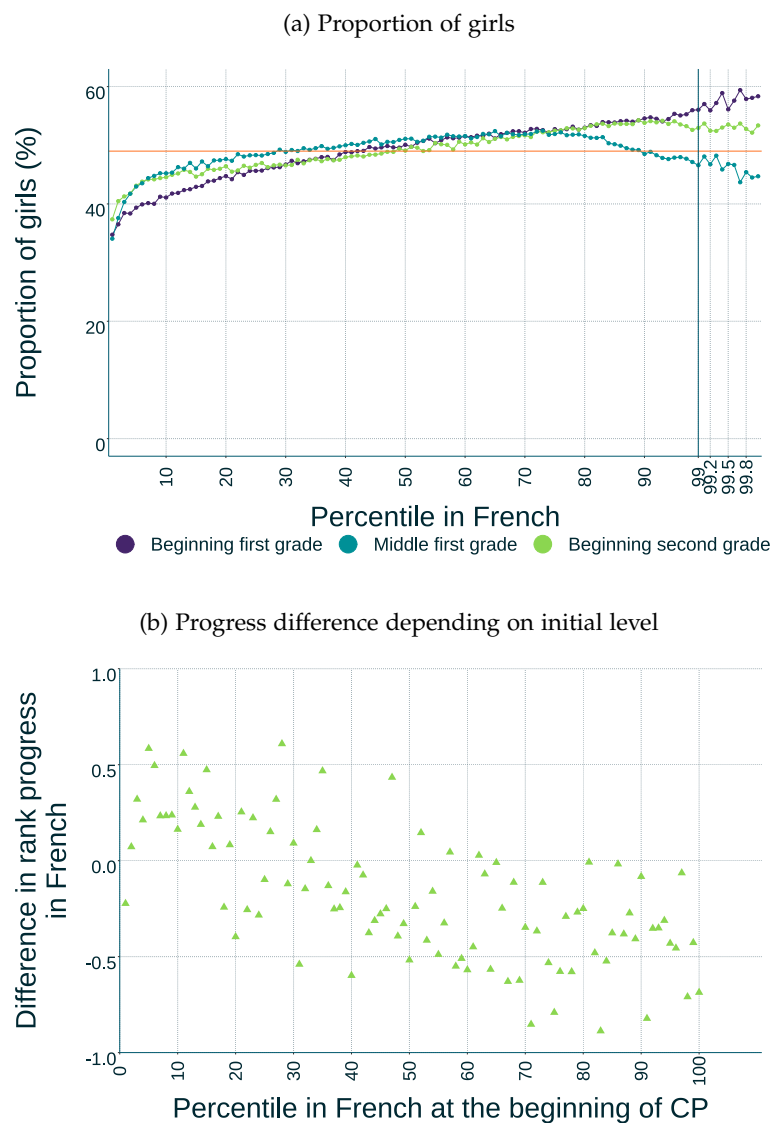
Figure A2: Gender gaps for recurring French items

Notes: A positive gap means that boys outperform girls and *vice versa*.

Sample: Students in CP between 2018 and 2022.

Sources: MENJ-DEPP, Repères CP-CE1; Authors' computations.

Figure A3: Gender gap dynamics in the French distribution



*Notes:* We rank student by score percentile from the smallest (the 1 % with the worst results) to the highest (the 1 % with the best results). Figure a) also includes a focus on the top 1 % of students. In figure b) the initial French percentile (x-axis) is defined relatively to the beginning of CP evaluation.

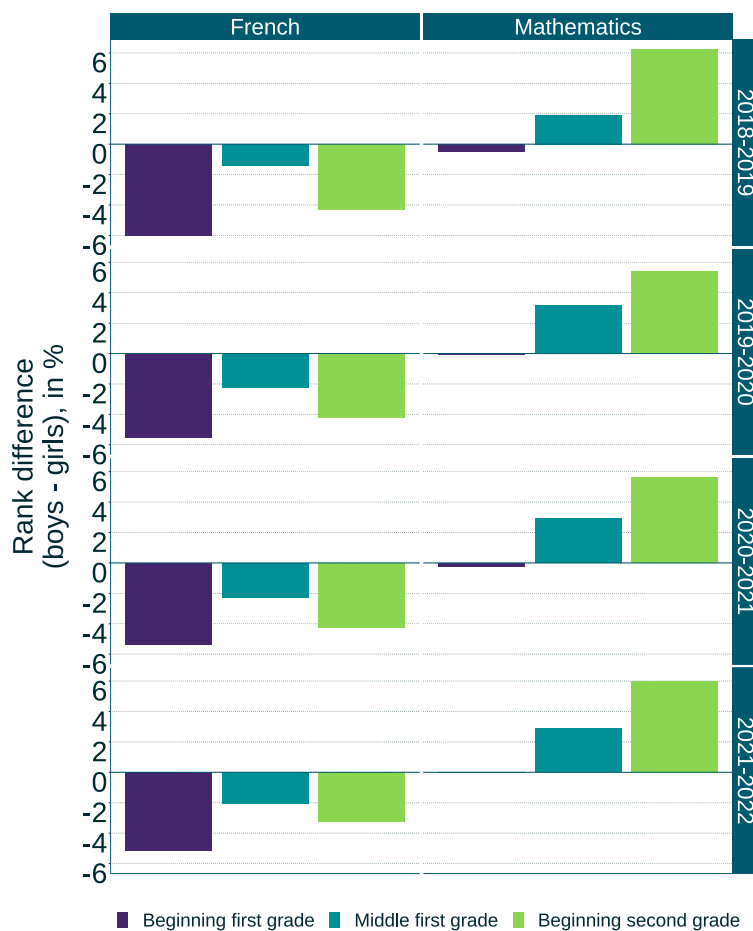
*Sample:* Students in CP between 2018 and 2022.

*Sources:* MENJ-DEPP, Repères CP-CE1; Authors' computations.



## A.2.2 Gender gaps

Figure A4: Gender gaps for each cohort

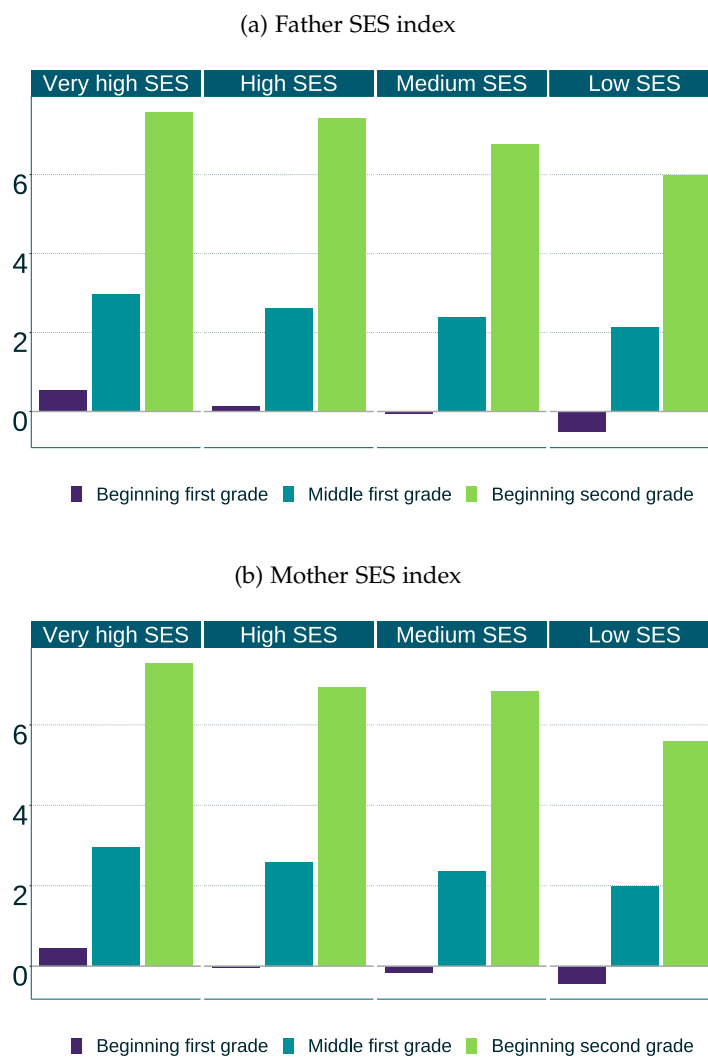


Notes: A positive gap means that boys outperform girls and *vice versa*.

Sample: Students in CP between 2018 and 2022.

Sources: MENJ-DEPP, Repères CP-CE1 & Constats de rentrée en sixième; Authors' computations.

Figure A5: Mathematics gender gaps depending on parents' SES index



Notes: A positive gap means that boys outperform girls and *vice versa*.

Sample: Students in CP between 2018 and 2022.

Sources: MENJ-DEPP, Repères CP-CE1 & Constats de rentrée en sixième; Authors' computations.

Figure A6: Mathematics gender gaps and father's occupation

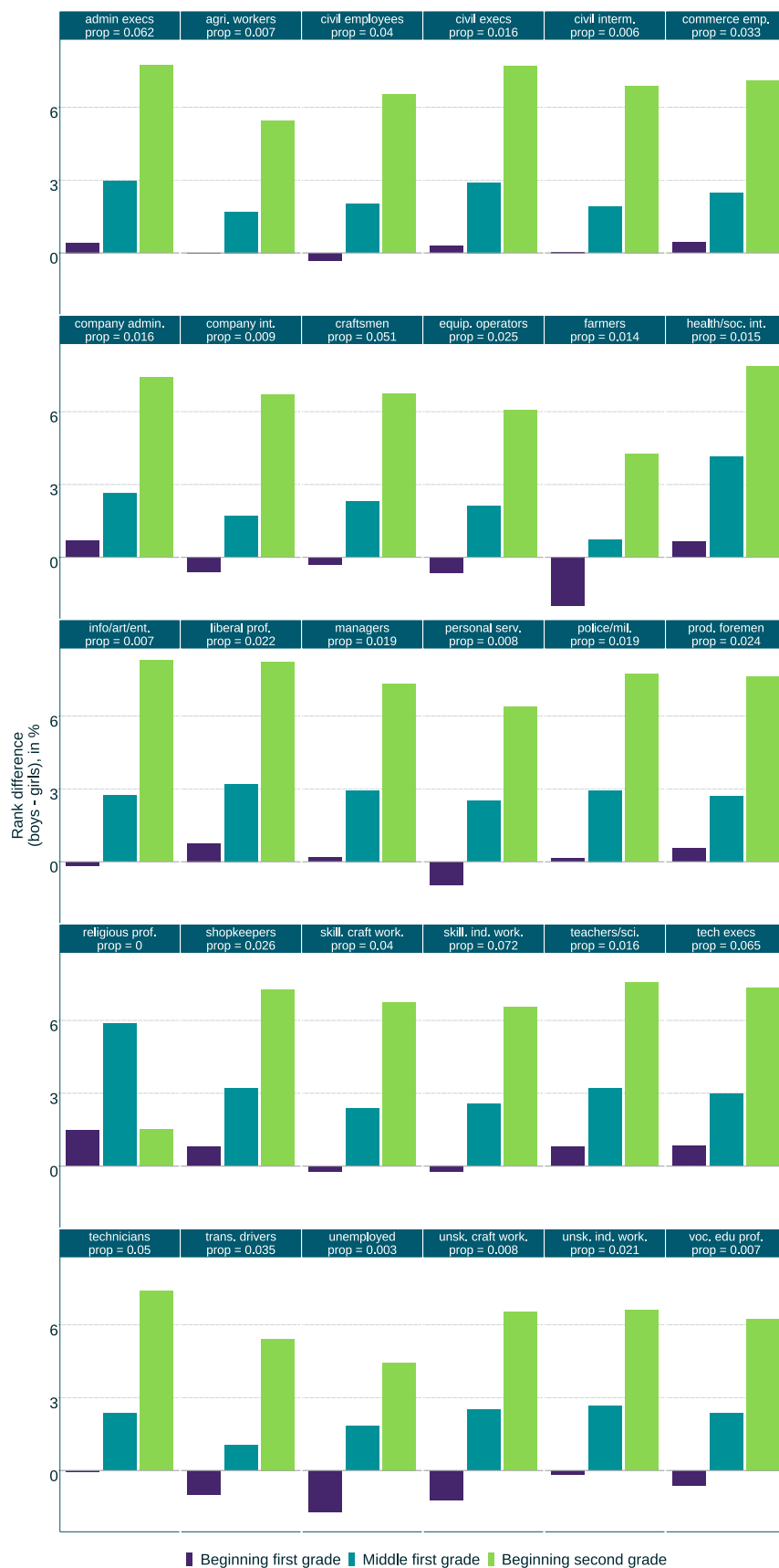


Figure A7: Mathematics gender gaps and mother's occupation



### A.3 Robustness checks

Table A3: Effect of gender and controls on girls' mathematics progress

Mathematics initial level	Dependent Variable: Mathematics progress											
	all students						top 50%					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
girl	-6.725*** (0.0652)	-6.812*** (0.0649)	-6.703*** (0.0593)	-8.912*** (0.0572)	-9.110*** (0.0539)	-9.194*** (0.0547)	-7.408*** (0.0885)	-7.398*** (0.0875)	-8.019*** (0.0844)	-10.21*** (0.0807)	-10.19*** (0.0775)	-10.19*** (0.0866)
age		-0.4979*** (0.0115)	-0.1932*** (0.0099)	-0.2424*** (0.0095)	-0.2968*** (0.0091)	-0.3230*** (0.0093)		-0.1760*** (0.0156)	-0.0035 (0.0153)	-0.0903*** (0.0144)	-0.1964*** (0.0138)	-0.2331*** (0.0155)
SES index family		-0.0224*** (0.0042)	-0.0435*** (0.0038)	-0.0327*** (0.0036)	-0.0174*** (0.0032)	-0.0145*** (0.0033)		-0.0474*** (0.0058)	-0.0526*** (0.0055)	-0.0377*** (0.0052)	-0.0221*** (0.0049)	-0.0171*** (0.0055)
family type: other		-2.183*** (0.2208)	-5.276*** (0.2227)	-4.419*** (0.1969)	-3.513*** (0.1645)	-3.371*** (0.1680)		4.272*** (0.3658)	6.224*** (0.3637)	4.015*** (0.3244)	3.506*** (0.2895)	3.082*** (0.3228)
family type: single parent		-0.6544*** (0.1770)	-0.7075*** (0.1641)	-0.8049*** (0.1526)	-0.4195*** (0.1204)	-0.4177*** (0.1225)		-0.6517*** (0.2505)	-0.6313*** (0.2387)	-0.6419*** (0.2228)	-0.2315 (0.1935)	-0.1665 (0.2177)
family type: single sex		1.119* (0.6328)	3.186*** (0.5554)	1.755*** (0.5235)	1.841*** (0.4858)	1.808*** (0.4991)		1.862** (0.7805)	2.512*** (0.7411)	1.221* (0.7018)	1.177* (0.6554)	1.415* (0.7435)
initial maths level			-0.4087*** (0.0015)	-0.6424*** (0.0018)	-0.6223*** (0.0015)	-0.6093*** (0.0015)			-0.4943*** (0.0034)	-0.6795*** (0.0035)	-0.6052*** (0.0030)	-0.5774*** (0.0034)
initial french level				0.3812*** (0.0018)	0.4172*** (0.0015)	0.4270*** (0.0015)				0.3631*** (0.0023)	0.3994*** (0.0021)	0.4108*** (0.0023)
<i>Controls</i>												
Parents' occupations and gender interactions		Yes	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes	Yes
<i>Fixed-effects</i>												
school_id					Yes	Yes					Yes	Yes
classroom_id						Yes						Yes
<i>Fit statistics</i>												
Observations	616,233	616,233	616,233	600,309	600,309	569,930	308,367	308,367	308,367	302,099	302,099	286,851
R <sup>2</sup>	0.01780	0.02621	0.21714	0.31363	0.44408	0.48594	0.02309	0.04664	0.13442	0.23370	0.42086	0.48354
Within R <sup>2</sup>					0.31369	0.30701					0.22879	0.22636
Adjusted R <sup>2</sup>	0.01779	0.02608	0.21703	0.31353	0.41706	0.41352	0.02309	0.04637	0.13418	0.23347	0.36329	0.32184
Within Adjusted R <sup>2</sup>					0.31358	0.30689					0.22855	0.22605

Notes: Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1. All continuous explanatory variables are centered to the mean.

Sample: Students in first grade between 2018 and 2022.

Sources: MENJ-DEPP, Repères first grade-CE1, APAE1D & Constats de rentrée en sixième; Authors' computations.

Table A4: More detailed effect of family configurations on girls' mathematics progress

Mathematics initial level	Dependent Variable: Mathematics progress							
	all students				top 50%			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
girl	-6.697*** (0.0625)	-6.696*** (0.0672)	-7.006*** (0.0681)	-7.138*** (0.0758)	-7.279*** (0.0906)	-7.370*** (0.0971)	-7.272*** (0.0920)	-7.388*** (0.1026)
SES index mother	0.0429*** (0.0037)			0.0341*** (0.0038)	0.0739*** (0.0052)			0.0692*** (0.0054)
SES index father	0.0510*** (0.0041)			0.0471*** (0.0043)	0.0786*** (0.0056)			0.0757*** (0.0059)
SES index family	-0.0369*** (0.0053)			-0.0309*** (0.0053)	-0.0434*** (0.0075)			-0.0437*** (0.0076)
girl × SES index mother	-0.0305*** (0.0050)			-0.0311*** (0.0052)	-0.0166** (0.0073)			-0.0202*** (0.0075)
girl × SES index father	-0.0240*** (0.0057)			-0.0269*** (0.0060)	-0.0076 (0.0079)			-0.0114 (0.0084)
girl × SES index family	0.0327*** (0.0073)			0.0280*** (0.0074)	0.0243** (0.0106)			0.0244** (0.0107)
scient./prof./engineer mother		2.438*** (0.2127)		0.9872*** (0.2400)		3.676*** (0.2384)		1.141*** (0.2749)
scient./prof./engineer father		2.130*** (0.1781)		0.2432 (0.2107)		3.384*** (0.2040)		0.8451*** (0.2431)
scient./prof./engineer parents		2.857*** (0.2665)		0.2673 (0.2961)		4.925*** (0.2888)		0.5702* (0.3283)
girl × scient./prof./engineer mother		-0.3441 (0.2975)		0.8644** (0.3380)		0.2891 (0.3412)		0.5124 (0.3927)
girl × scient./prof./engineer father		-0.2187 (0.2518)		0.7562** (0.2990)		0.5894** (0.2944)		0.3876 (0.3530)
girl × scient./prof./engineer parents		0.6084 (0.3780)		2.251*** (0.4234)		1.476*** (0.4052)		1.520*** (0.4675)
family type: other			-3.420*** (0.1619)	-3.357*** (0.1643)			1.025*** (0.3080)	1.884*** (0.3114)
family type: single parent			-1.595*** (0.1601)	-1.418*** (0.1628)			-2.001*** (0.2505)	-1.571*** (0.2503)
family type: single sex			1.914** (0.8387)	1.706** (0.8361)			1.802* (1.053)	1.566 (1.046)
girl × family type: other			2.727*** (0.2265)	2.860*** (0.2285)			0.3427 (0.4513)	0.4453 (0.4549)
girl × family type: single parent			1.314*** (0.2194)	1.208*** (0.2239)			0.4257 (0.3393)	0.4761 (0.3406)
girl × family type: single sex			-1.573 (1.191)	-1.299 (1.188)			-2.085 (1.527)	-1.758 (1.513)
<i>Fixed-effects</i>								
cohort × classroom_id	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
<i>Fit statistics</i>								
Observations	584,195	584,195	584,195	584,195	292,484	292,484	292,484	292,484
R <sup>2</sup>	0.27442	0.27418	0.27396	0.27525	0.35672	0.35345	0.35016	0.35736
Within R <sup>2</sup>	0.02397	0.02364	0.02334	0.02508	0.03907	0.03419	0.02928	0.04003
Adjusted R <sup>2</sup>	0.21579	0.21553	0.21529	0.21667	0.24809	0.24428	0.24043	0.24881
Within Adjusted R <sup>2</sup>	0.02395	0.02363	0.02333	0.02505	0.03904	0.03417	0.02925	0.03996

Notes: Clustered (cohort × school\_id) standard-errors in parentheses. Signif. Codes: \*\*\*: 0.01, \*\*: 0.05, \*: 0.1. All continuous explanatory variables are centered to the mean.

Sample: Students in first grade between 2018 and 2022.

Sources: MENJ-DEPP, Repères first grade-CE1, APAE1D & Constats de rentrée en sixième; Authors' computations.