Institut des Politiques Publiques

A HELPING HAND FOR GIRLS? GENDER BIAS IN MARKS AND ITS EFFECT ON STUDENT PROGRESS

IPP Note

n°14

December 2014

Camille Terrier

www.ipp.eu

Summary

In France, when girls start junior high school, their marks are higher than those of boys in French, but lower in mathematics.This latter gap soon closes, however, and by the last year of junior high school, girls do as well as boys in maths. This IPP Note offers an explanation for these differential success rates and how they change over time, through the consideration of two questions: is there gender bias in how teachers grade pupils? And does any such bias affect the progress of girls compared with boys?

Analysis of marks given both anonymously and not, to students[1] in the first year of junior high school reveals positive discrimination for girls in maths but an absence of gender bias in French: for similar anonymous marks, girls receive higher marks from their maths teachers than boys. The less disruptive behaviour of girls in class does not seem to explain this 'helping hand'.

We then show that in maths, the classes in which teachers show greatest bias in favour of girls are also the classes in which girls progress the most relative to boys. This finding is in line with other research highlighting that grading practices affect pupils' motivation and progress. It is also a possible explanation for the reduction in differential achievements in maths observed between girls and boys during junior high school.

- In this sample, the grades that 6th grade teachers give in maths seem to indicate a bias in favour of girls, who gain on average six per cent more than boys, for similar anonymous grades. No bias is observed in French.
- Girls' better behaviour in class is not behind this nudge forwards.
- In maths, the classes in which teachers exhibit the greatest discrimination in favour of girls are also the classes in which the girls progress the most relative to the boys. In the absence of bias, the girls would neither have progressed more than boys, nor caught up with them.

The Institute for Public Policy (IPP) is a joint scientific project of the PSE and the Centre for Research in Economics and Statistics (CREST). It aims to promote the analysis and quantitative evaluation of public policy using the latest economics research methods.

CREST Centre de Recherche en Économie et Statistique
www.crest.fr

PARIS SCHOOL OF ECONOMICS ÉCOLE D'ÉCONOMIE DE PARIS
www.parisschoolofeconomics.eu

1. Source: Students in this study are from the Créteil academy (an education administrative area)

In most countries, **students' grades have two marked characteristics**: at their entry to junior high school, girls' marks are higher than boys' in French, but lower in mathematics. This observation is confirmed in France, where 89 per cent of girls master the basics in French at the end of primary school, compared with 87 per cent of boys.[2] Conversely, in maths, the difference is in favour of boys: 91.1 per cent master the basics, compared with 88.1 per cent for girls. The second important point concerns the change in these **inequalities over the course of the school years**: in maths, the inequalities in success between girls and boys tend to reduce, even disappear, whereas in French, they remain. The differences between subjects are particularly apparent in international studies such as PIRLS (results for reading) and TIMSS (results in maths).[3] This observation is again confirmed for French students. At the end of junior high school, the proportion of students who master the basics in maths is the same among girls as boys: the girls are no longer behind the boys. In French, however, the gap remains significant.

These findings raise two questions: how do we explain the gap in achievement between girls and boys at the beginning of first year of junior high school? And why does the gap decrease in maths during the course of junior high school, while the gap in French remains? This paper addresses these questions, first by analysing the bias in marks that teachers give, then by investigating the effect of this bias on the relative progress of girls and boys.

Measuring bias in teachers' grades

We had access to grades given in both anonymous and non-anonymous fashion to 4,519 6th grade students in the Créteil academy.

These marks were collected in the context of an experiment on the place of parents in school. The anonymous marks come from standardised exams taken at the beginning and the end of 6th grade. The non-anonymous marks are those given by teachers in each of the three terms of the school year.

To study the existence of gender bias in teachers' grades, **we use the double-difference method**, which consists of comparing the average difference between non-anonymous and anonymous grades for girls with the same difference for boys. The assumption of this method is as follows: if the anonymous and non-anonymous scores measure the same competences, then it is possible that the anonymous exam grading gives rise to a more "severe" correction for both genders, in which case the average grades given by teachers might be higher or lower than the average marks given anonymously. However, the gap between non-anonymous and anonymous scores should be the same for girls and boys. In other words, girls should not systematically gain a higher score than boys when the grades are given by teachers than when the marks are gained in anonymous fashion. If it is the case, then this gap could be interpreted as gender bias in teachers' grading.[5]

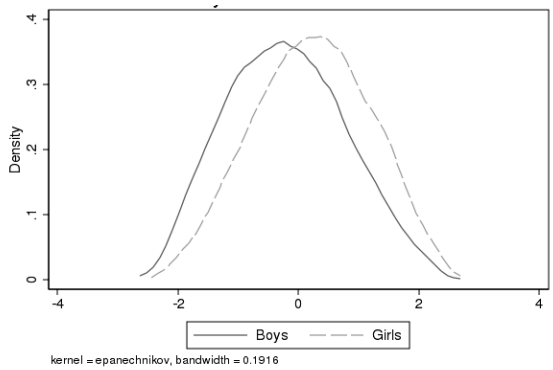
Girls are favoured in maths when marks are given by teachers

The difference between grades can be seen graphically as we trace the distributions of girls' and boys' scores in maths and in French. It is interesting to note that there is a significant difference between the two subjects. Figure 1 suggests that **in French**, girls have better results than boys and that this gap is the same regardless of how the marks were obtained (through either anonymous tests or through teachers' grades).

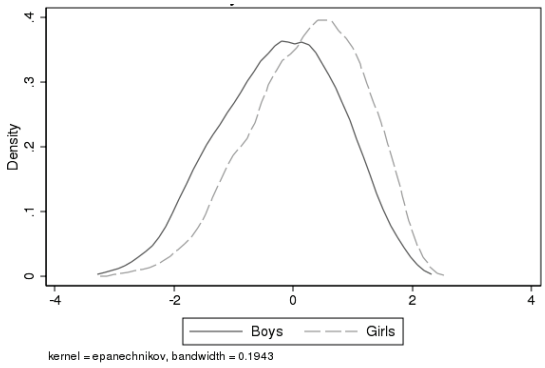
2. Students in this study are from the Créteil academy (an education administrative area).
3. « Méthodologie de l'évaluation des compétences de base en français et en mathématiques en fin d'école et de collège », Note d'Information n°08-37 (2008), Direction de l'Evaluation, de la Prospective et de la Performance.
4. "PIRLS 2011 International Results in Reading", Ina V.S. Mullis, Michael O. Martin, Pierre Foy, and Kathleen T. Drucker, 2012, Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. " TIMSS 2011 International Results in Mathematics", Ina V.S. Mullis, Michael O. Martin, Pierre Foy, and Alka Arora, 2012, Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
5. Formally, this gap (that is, the discrimination) is written: E(NA-A/Girls) -E(NA-A/Boys), with NA= Non-anonymous marks and A=Anonymous marks.

Figure 1: Distribution of girls' and boys' marks **in French**

Anonymous grading (figure 1.A)



Non-anonymous grading (figure 1.B)



Sources: Data collected in an experiment on the place of parents in school.
Note: The marks are standardized such that their mean is equal to 0 and their variance is equal to 1. Girls have an average score higher than that of boys for both types of marks (anonymous and non-anonymous). The average difference between girls' and boys' scores represents 0.434 points of a standard deviation when the marks are anonymous, and 0.460 points when they are non-anonymous..

In maths, however, Figure 2 gives a different picture: girls have lower marks than boys when they are gained anonymously, but their results are better than boys' in the grades given by teachers. This average gap suggests that girls benefit from positive bias in their favour in mathematics. **For equal anonymous marks, grades given by teachers are on average 6.2 per cent higher for girls than for boys in maths.** No such gender bias is observed in French. These results tend to contradict the popular idea that girls suffer from negative discrimination in scientific subjects. We find the reverse and our results confirm the conclusion of similar studies carried out in other countries (Lavy, 2008; Robinson and Lubiensky, 2011).

In order to be certain of the robustness of these results, several alternative specifications were tested to take into account the fact that:

(i) teachers' grades and standardised exams marks were not attributed at exactly the same time,
(ii) girls and boys might be unequally "stressed" or attach more or less importance to the exams, and
(iii) the two evaluations perhaps do not measure exactly the same competences.

All the tests confirm a significant bias in favour of girls in mathematics and the absence of bias in French.

Less disruptive behaviour among girls is not behind the gender bias…

To take these results further, we test whether the positive discrimination in favour of girls in maths reflects their less disruptive behaviour in class. If the teachers' grades take into account – consciously or not – the behaviour of students, then this might explain why the difference between the grades they give and the marks from anonymous tests are higher for girls than boys. We have detailed data on student behaviour, including school suspensions and "conduct warnings" issued by the class councils. As expected, these two sanctions are very unequally distributed between girls and boys: 84.4 per cent of those punished are boys. The inclusion of these variables in the econometric analysis allows us to measure the discrimination that is not attributable to student behaviour. We find that **a negligible part (four per cent) of the bias in favour of girls is explained by their less disruptive behaviour**. Student behaviour does not, therefore, explain the gender discrimination in grades.

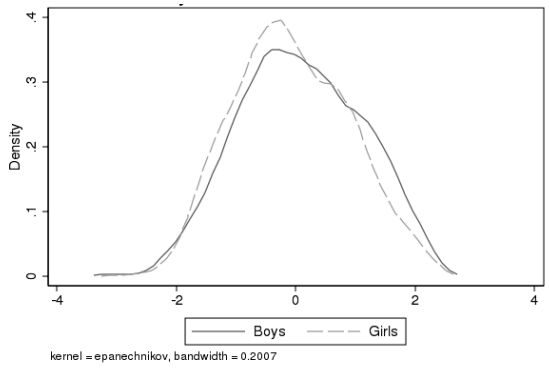
… but their weaker results in maths partially explains it.

It is also possible that the bias in favour of girls partly represents an "over-marking" of the weaker students – the aim of which is to encourage them. Figures 1.A and 2.A show that girls are weaker than boys in mathematics, but better in French. Independently of all gender bias, if teachers have a tendency to mark higher the weakest students, then this could increase girls' marks in maths. To test this hypothesis, we introduced into the regression information about whether students belonged to the first or last decile of the anonymous marks distribution. The results suggest that **in mathematics, 7.5 per cent of the bias in favour of girls is explained by a positive bias in favour of weaker students.**

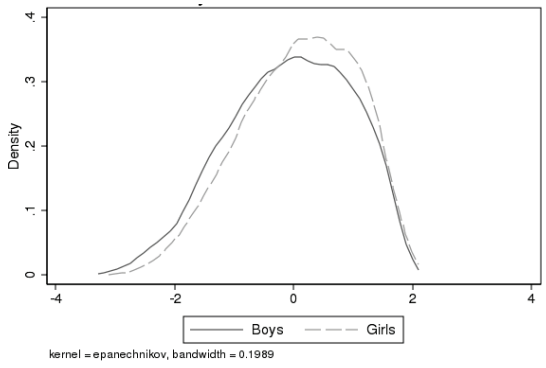
Thus far, our results offer an explanation of the disparity in success rates between girls and boys in maths. Next, we investigate the change in these inequalities in the course of schooling, to find out why the differences disappear in mathematics, but remain in French. To do so, we posed the following question: what is the effect of the observed gender bias on girls' progress over the year, compared with that of boys?

Figure 2: Distribution of girls' and boys' marks **In Mathematics**

Anonymous grading (figure 2.A)



Non-anonymous grading (figure 2.B)



Source: Data collected in an experiment on the role of parents in school.
Note: The marks are standardized such that their mean is equal to 0 and their variance is equal to 1. Girls have an average score lower than that of boys when the scores are anonymous, but higher when they are given by teachers. The average difference between girls' and boys' scores is negative and represents 0.147 points of a standard deviation when the marks are anonymous; it is positive and represents 0.169 points when they are non-anonymous.

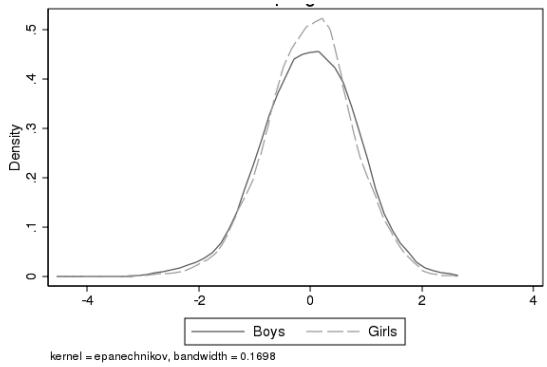
The impact of gender discrimination on girls' progress

We define student progress as the gap between the results from standardised exams at the beginning of the year and at the end of it. Figure 3 suggests that girls and boys progress in similar fashion in French, while in maths, girls make more progress than boys.

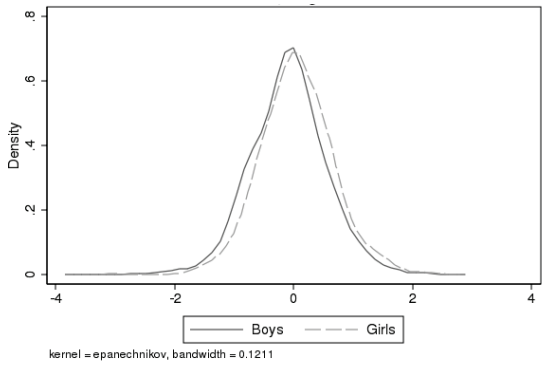
In order to identify the effect on the progress of girls (relative to that of boys) of being assigned to the class of a teacher who exhibits a high degree of discrimination, we use the significant variation in the discriminatory behaviour of teachers. Indeed, not all teachers show gender bias in their marks and among those whose marks do favour girls, some show a strong degree of discrimination while others are much less biased.

Figure 3: Comparison of girls' and boys' progress

In French (figure 3.A)



In maths (figure 3.B)

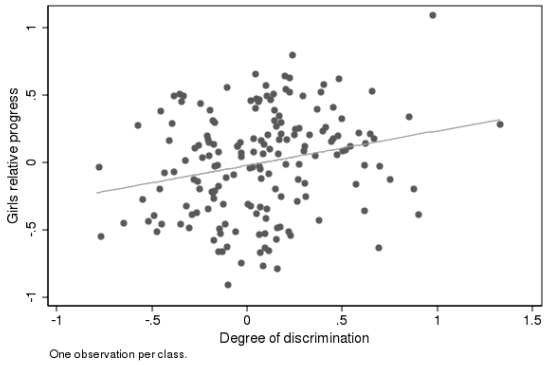


Source: Data collected in an experiment on the role of parents in school.
Note: Progress is defined as the difference between the results from standardised exams taken at the beginning and at the end of the year. The distribution of girls' progress is similar to that of boys in French. In maths, girls make more progress than boys.

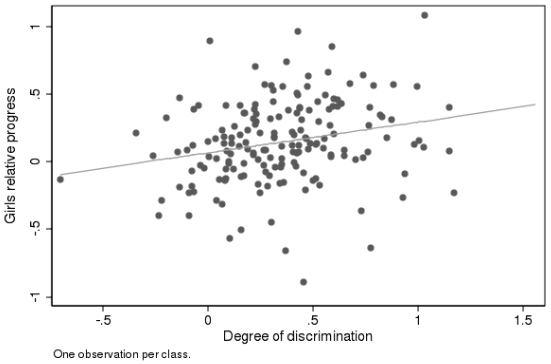
Figure 4 shows the degree of correlation between the level of bias in a class and the progress of girls relative to that of boys. Each point represents one class. We see a positive correlation: **the greater the bias in favour of girls in a class, the more the girls tend to progress relative to the boys.**
Above and beyond these correlations, to understand how a teacher's discriminatory behaviour can causally affect the relative progress of girls, we must distinguish it from other mechanisms that might explain why girls progress more than boys in maths. Indeed, it is possible that girls have a propensity to make more progress than boys independently of any discrimination. In addition, teachers who demonstrate a greater degree of positive discrimination might also be teachers whose classes make more progress, independently of any bias in their grades.

Figure 4: Correlation between the degree of discrimination of a teacher and the progress of girls in that class relative to boys

In French (figure 4.A)



In maths (figure 4.B)



Source: Data collected in an experiment on the role of parents in school.
Note: Each point represents one class. The degree of discrimination (x-axis) represents, for each class, the coefficient of discrimination calculated in the first part of the analysis, that is the difference between the average non-anonymous and anonymous marks of the girls in a class, minus that difference for boys. The relative progress of girls (y-axis) represents the progress of boys, which is subtracted from that of girls. The slope of the line indicates the coefficient of correlation between the two variables.

To take these three separate effects into account, our method uses the following comparison: we consider two classes that initially had the same difference in results between girls and boys[6] and observe how the assignment of a teacher more or less discriminating to each of these classes affects that difference at the end of the academic year. In other words, for two identical classes, if one teacher greatly favours girls, and the girls in that class progress more than the boys, then the difference in results between genders at the end of third trimester will be higher in that class. It is precisely the existence (or not) of this difference that we seek to estimate. [7]

The main result of our analysis suggests that **in mathematics, the classes in which teachers show the greatest degree of positive discrimination in favour of girls are also the classes in which the girls progress the most relative to the boys**. As an indication, we obtain an average coefficient of discrimination of 0.31 in the first part of the analysis. For this level of bias, the average marks of girls will be 1.9 per cent higher than those of boys at the end of the year (starting from equal marks). It is interesting to note that **the catch-up by girls in maths is fully explained by the teachers' positive discrimination: without it, the difference in results between girls and boys would remain constant**. Finally, in French, we observe no effect of discrimination on progress.

We do not have enough information to understand what other mechanisms operating simultaneously might explain how the gender bias in teachers' grades affects the progress of girls in maths but not in French. **We can, however, formulate hypotheses about two mechanisms discussed in the literature.** First, some studies show that student interest in a subject depends positively on the marks they obtain in it, and negatively on the marks of their peers (Trautwein et al., 2006; Marsh & Craven, 1997).
Thus, the bias in favour of girls works in two ways: it increases girls' results in absolute fashion, but also in relative fashion in relation to the boys in the class. The cumulative influence of these two effects might be at the origin of girls' increased interest in maths and inversely, at the origin of boys' decreased interest. In addition, it is possible that girls are affected by stereotype threats in mathematics (Spencer et al., 1999). Some studies show that girls' and women's results in maths are lower when they have internalised the stereotypical idea that women are less capable than men in the subject. This threat applies only to subjects about which there is a negative stereotyping of girls and women, so it is likely that girls' performances are more influenced by this effect in mathematics than in French. This could explain why positive discrimination in favour of girls only affects their progress in maths.

Conclusion

This study suggests that in the first year of junior high school, girls benefit from a bias in their favour in mathematics when the grades are given by teachers, while marks in French are not affected by the student's gender. The less disruptive behaviour of girls in class is not the cause of this benefit. On the other hand, girls' initially lower scores than boys' do explain part of it. Beyond this short-term effect, we show that gender bias in marks affects the evolution of the gender achievement gap over time. In maths, the more the teachers' grades are biased in favour of girls, the more girls tend to progress relative to boys in the course of a school year, reducing the initial gap in favour of boys in maths.

These results are interesting for three reasons. First, they highlight that grades given by some teachers in maths do not reflect solely the abilities of the students. Since several important decisions in school life are made on the basis of student grades (choice of stream at the beginning of upper secondary school, whether to repeat a year, choice of subject paths, etcetera), marks obtained through standardised exams have the advantage of being free from bias and constituting a viable alternative, though they are open to other criticisms. Our study also raises the broader question of the role of marks: are they intended only to evaluate the competence of students, or are they also an instrument with which teachers improve student progress? In the latter case, since we note that marks in maths influence the progress of students, they could be a way to reduce the inequalities in achievement between boys and girls. In the same way, in the humanities subjects, they might encourage boys enough to eliminate their lag. In the first case, if we consider that the role of marks is above all to evaluate the competences of students, then the existence of bias can seem problematic. A simple and non-costly policy to remedy this would consist of informing teachers about conscious or unconscious stereotyping and its potential effects on the grades that they give. Some simple training modules on these themes could provoke greater vigilance among teachers, education providers and inspectors. Finally, in certain countries, the age at which students begin to be graded is higher than in France (11 years old, for example, in Sweden), which slows the possibility of bias in the marks given – although teacher behaviour can still be biased.

References:
« Giving a Little Help to Girls ? Evidence on Grade Discrimination and its Effect on Students Achievement. », Camille Terrier, PSE Working Papers n 2014-36. 2014.

Author:
Camille Terrier is a PhD student at the Paris School of Economics.

6. The gap between girls and boys under consideration is the difference in the standardised exam in the first trimester.
7. In order that the estimate of the discrimination effect on progress not be biased, the allocation of a teacher exhibiting a strong degree of bias must be independent of the unobserved characteristics of the girls whom they might potentially help to progress more than boys. This hypothesis seems plausible to the extent that assigning girls with a high progress potential to the most discriminating teachers would require that the directors of the school know, at the moment when classes are being composed, both the discriminating teachers and the girls with the most potential to succeed. The study being carried out on students entering the first year of junior high school (that is, not present in the schools during the preceding years), it is not likely that the school director has this information.
8. Marsh, Herbert. W., and Rhonda G. Craven. "Academic self-concept: Beyond the dustbowl". In G. D. Phye (Ed.), Handbook of classroom assessment. San Diego, CA: Academic Press. (1997): 131-198. Trautwein, Ulrich, Oliver Ludtke, Herbert W. Marsh, Olaf Koller, and Jurgen Baumert. "Tracking, Grading, and Student Motivation: Using Group Composition and Status to Predict Self-Concept and Interest in Ninth-Grade Mathematics." Journal of Educational Psychology 98, no 4 (November 2006): 788-806.