

Are the Gender Gaps in PISA Influenced by Its Methodology?

Laura Zieger & John Jerrim

PISA measures 15-year-olds achievement scores in mathematics, reading and science and has evolved into a powerful tool in politics, as the scores can be compared across countries and over time. Apart from the scores themselves, there is also significant interest in the achievement differences between boys and girls.

What most people do not know about PISA is that not every child is directly assessed in all three subjects. Moreover, roughly 60% of the students only answer questions in two of the three subjects. Yet, everyone is assigned an achievement score in mathematics, reading and science.

In order to allocate test scores to all students, predictions are made from the subjects where children actually answered questions and their background characteristics – including gender. This is known as ‘conditioning’ in the academic literature and is considered vital to correctly estimating gender gaps in achievement. Indeed, psychometricians argue that gender differences will tend to be underestimated unless this ‘conditioning’ takes place.

How does conditioning affect the PISA gender gaps in reality?

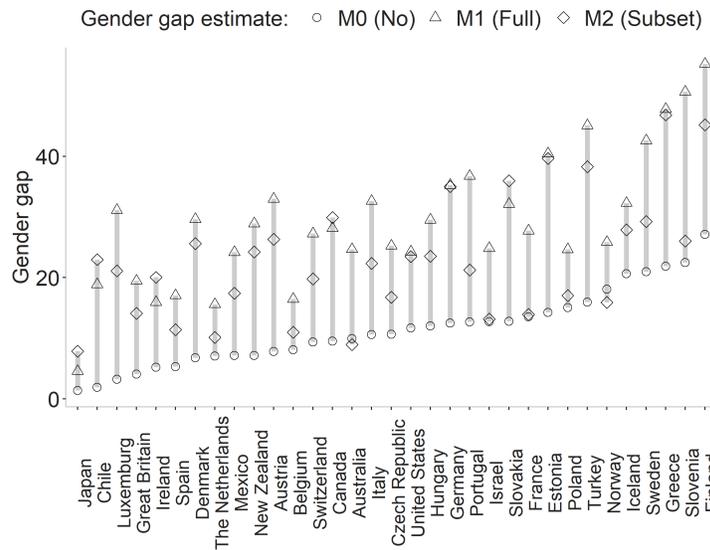
In theory, as soon as gender is included in the conditioning model, we should correctly estimate the difference in achievement between boys and girls. But theory and reality can be different things. For this reason, we have investigated the [conditioning](#)

Laura Zieger
University College London, e-mail: l.zieger@ucl.ac.uk

John Jerrim
University College London, e-mail: j.jerrim@ucl.ac.uk

model¹ used in PISA 2012. As part of this project, we computed three alternative versions of the students' scores in each of the PISA subjects. In the first model, the scores for each student are inferred just from the test responses in the different domains (no conditioning; M0). In the second model, we used the responses to the PISA test questions and all background variables (full conditioning; M1). In the last model, test responses were combined with just a subset of the background variables (gender, grade, mothers and fathers socio-economic index and booklet IDs) to test the sensitivity of the results (conditioning on subset; M2).

The figure below shows the gender gap in reading using model M0 (no conditioning - circle), M1 (full conditioning - triangle) and M2 (conditioning on subset - diamond). We expect a big difference between no conditioning and both versions of conditioning, while M1 and M2 should be very similar. And indeed, for most countries, the triangle (M1) and the diamond (M2) are pointing in the same direction, and for about a third of countries, they even sit on top of each other. This suggests that, in most countries, the gender gap is not sensitive to the exact specification of the conditioning model (once gender has been included) with a potential small increase or decrease when more variables are included. There are, nevertheless, some important changes to the results for some individual countries (that are somewhat difficult to explain). For instance, in Australia, Israel, France, Poland, Slovenia, and Norway the estimated gender gap from M0 and M2 are similar. Yet there is a large gap in the magnitude of the gender gap in M1.



¹ <https://repec-cepeo.ucl.ac.uk/cepeow/cepeowp20-09.pdf>

What does this mean?

Our research led to two conclusions for gender gaps: First, the gender gaps amplify to a less biased estimate as soon as gender is included in the conditioning model. Second, theoretically, the exact specification of the included background variables should not matter as long as gender is included. While this holds true for the majority of countries, our research shows that some countries experience delicate impacts on their gender gap which are also reflected in changes in the ranking of those countries. Summarising we can say that regarding gender gaps, it matters if and which background variables are used in the computation of the students' scores.

This text has been posted on the blog international-education.blog and it is available in different languages on international-education.blog

Zieger, L. & Jerrim, J. (2020). Are the gender gaps in PISA influenced by its methodology? Retrieved from <https://international-education.blog/en/are-the-gender-gaps-in-pisa-influenced-by-its-methodology>

The blog International-Education.Blog is an outcome of the project OCCAM. This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 76500.

