infoIEB Number 30 · December 2018

IEB Institut d'Economia de Barcelona

Universitat de Barcelona

Gender Differences Under Test Pressure and their Impact on Academic Performance: a Quasi-Experimental Design

Student performance, especially at university, is a strong determinant of individual decisions and future labour opportunities. The well-documented gender gap in labour outcomes (see Blau and Kahn, 2010 for a recent review) highlights the need to understand the potential determinants of gender differences in academic performance so as to mitigate them and to promote increased gender equality in education and, hence, in labour opportunities.

In recent years, economic researchers have become increasingly interested in gender differences in performance when under pressure (see, among others, Shurchkov, 2012). However, few studies have analysed the role of pressure in educational performance across gender using real-world data. Little is known about how test pressure affects student performance across gender at the university level. In higher education, students sit a significant number of exams exposing them to different levels of pressure, depending on the number of credits at stake, the weight of the exam in the final grade, the specific rules of assessment which might require students to perform better, the difficulty of the exam content, the type of exam, or the simple fact of having to sit an exam.

Additionally, a multiple choice test is the most frequently adopted format to assess student knowledge at university. Here, also, published studies have documented gender differences in the results obtained on multiple choice tests and in the respective answering behaviour by gender. However, little is known about how this gender gap varies according to the level of pressure faced, highlighting the need for more studies that can shed light on gender differences attributable to pressure on this test format.

Male students perform better than female students when under greater test pressure, but that this gender gap narrows as test pressure decreases until mitigated.

In our recent article ("Gender differences under test pressure and their impact on academic performance: a quasi-experimental design", IEB Working Paper 2018/21), we study gender differences in academic performance under different levels of pressure when sitting multiple choice tests. The introduction of continuous assessment in the evaluation system of a university course allows us to exploit a unique quasi-experimental setting in which the same students take similar tests but exposed to different levels of pressure. This setting allows us to structure the dataset in two ways and to exploit each to the full: independently pooled cross-sections and panel data. This study uses data from two sources: administrative and course data. First, the administrative data was provided by the University of Barcelona's Faculty of Economics and Business. They contain full demographic and academic information for all students enrolled on the course Principles of Taxation over the seven academic years. Second, the course data were provided by the Economics of Taxation Teaching Innovation Group (GIDEI, from the acronym in Catalan). It contains a full set of information about the grades and groups. The use of unique student administrative data provides us with a

Table 1: Gender Gap – Pooled Cross-Section vs Two-Period Panel Data

	Pooled Cross-Section		Two-Period Panel Data	
	CA Grade (1)	Final Test Grade (2)	Pooled OLS (3)	RE (4)
CA Gender Gap (low pressure)	0.032	-	0.0427	0.0427
Final Test Gender Gap (high pressure)	-	–0.086***	–0.091***	–0.091***

Note: CA stands for continuous assessment.

Figure 1: Gender gap ordered from the lowest to the highest pressure scenario

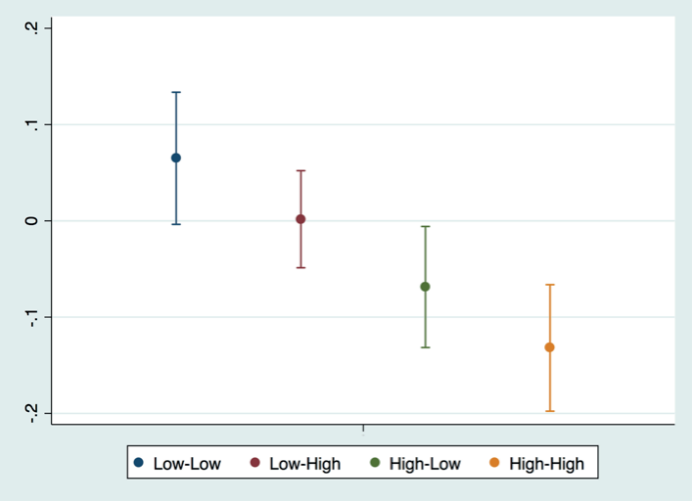
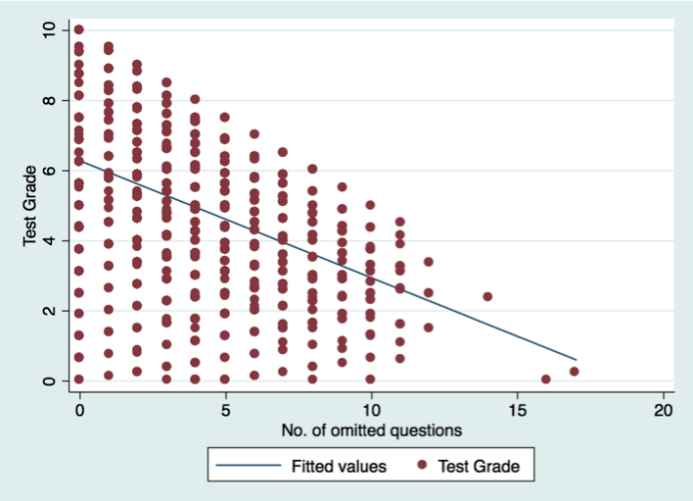


Figure 2: Scatter plot of final test grade and number of omitted questions on the final test



rich individual-based dataset to control for personal and group characteristics that might also determine student performance and affect gender differences.

We define test pressure as the increased need that students feel to perform well on an examination due to the increasing importance attached to that test. The principal source of test pressure on students in our main setting is the greater weight attached to a test in the final grade awarded for the course. Moreover, because in our setting some of the rules of evaluation change over the years, modifying the pressure students are under when sitting these exams, we perform a heterogeneity analysis. In this case, the secondary source of test pressure is identified as the rules of evaluation which require a better performance from the students. The strength of our quasi-experimental set up is based on five specific characteristics: (i) the tests are machine/computer corrected so require no subjective bias correction, (ii) the tests present an almost identical format (multiple choice), employing the same questions, with the same level of difficulty and a very similar structure, (iii) the same cohort of students sit these tests in scenarios characterised by different levels of pressure, (iv) the pressure students are under is analysed in a real world environment i.e. sitting their university exams, and (v) our data and empirical strategy allow us to explore the possible mechanisms responsible for the results obtained, that is, we are able to disentangle the main drivers of the gender differences observed.

The empirical results (table 1) show that, after controlling for individual and group characteristics, male students perform better than female students when under greater test pressure, but that this gender gap narrows as test pressure decreases until mitigated, and even reversed, in favour of female students (figure 1). Our pooled cross-sectional estimates allow us to obtain the main results for each performance separately, to analyse differences along each grade distribution and to assess self-selection bias related to the decision to take or not take each exam. Our panel data enable us to reshape the data set and to analyse the setting as a sequential game of two periods, allowing us to confirm our main results and, moreover, to test for heterogeneity effects in relation to exam pressure.

We also examine the potential mechanisms that might account for these gender gaps and find that they are likely to differ depending on the source of test pressure. When we analyse the pressure attributable to the weight of the test, female students appear more likely to choke under pressure, while male students maintain their level of performance. Thus, the gender gap results from a fall in the performance of female students. However, our analysis of the test pressure attributable to the rules of the evaluation system shows that male students are more likely to excel while the performance of female students remains unchanged. Thus, in this instance, the gender gap results from a hike in the performance of male students. Moreover, there is suggestive evidence that the top female students omit more items than male students on the final test as test pressure rises (figure 2). Consequently, education systems should be designed to assess student knowledge and abilities, and grades should not be influenced by student test taking strategies.

References

Blau, F. D. and Kahn, L. M. (2010). Gender Differences in Pay. Journal of Economic Perspectives, 14(4), 75-99.

Shurchkov, O. (2012). Under Pressure: Gender Differences in Output Quality and Quantity Under Competition and Time Constraints. Journal of the European Economic Association, 10(5), 1189-1213.

Daniel Montolio, IEB Researcher
montolio@ub.edu

Pere A. Taberner, KSNET Researcher
peretaberner@gmail.com