

Diferencias de género en pruebas test bajo presión y su impacto en el rendimiento académico: un diseño cuasi-experimental

El rendimiento de los estudiantes, especialmente en la universidad, es un fuerte condicionante de las decisiones individuales y de las oportunidades futuras en el mercado laboral. La más que demostrada desigualdad de género en los resultados laborales (ver Blau y Kahn, 2010 para una revisión reciente) pone de relieve la necesidad de entender los potenciales determinantes de las diferencias de género en el rendimiento académico para mitigarlos y promover la igualdad de género en la educación y, como resultado, también en las oportunidades laborales.

En los últimos años, los investigadores económicos se han mostrado cada vez más interesados en las diferencias de género en el rendimiento en situaciones de presión (ver, entre otros, Shurchkov, 2012). Así mismo, pocos estudios han analizado el rol de la presión en el rendimiento académico sobre el género utilizando datos reales. Sabemos poco sobre cómo la presión de los exámenes afecta al rendimiento de los estudiantes según su género a nivel universitario. Durante la etapa de educación superior, los estudiantes realizan un importante número de exámenes que les exponen a diferentes niveles de presión, dependiendo del total de créditos que hay en juego, del peso del examen sobre la nota final, de las normas de evaluación específicas que pueden incentivar un mejor rendimiento de los estudiantes, de la dificultad del contenido del examen, del tipo de test o del simple hecho de tener que hacer la prueba.

Además, los test de respuesta múltiple son uno de los formatos más frecuentes en la universidad a la hora de evaluar los conocimientos de los estudiantes. También en este caso, los estudios publicados han documentado diferencias de género en los resultados obtenidos y en el comportamiento a la hora de dar respuesta. Aún así, se sabe poco sobre cómo la brecha de género cambia en función del nivel de presión al que se enfrentan los estudiantes, evidenciando la necesidad de más estudios que puedan dar luz sobre las diferencias de género atribuibles a la presión relacionada con este formato de examen.

Los chicos rinden mejor que las chicas cuando realizan exámenes bajo gran presión, pero esta brecha de género se hace más pequeña a medida que la presión se reduce, hasta el punto que desaparece.

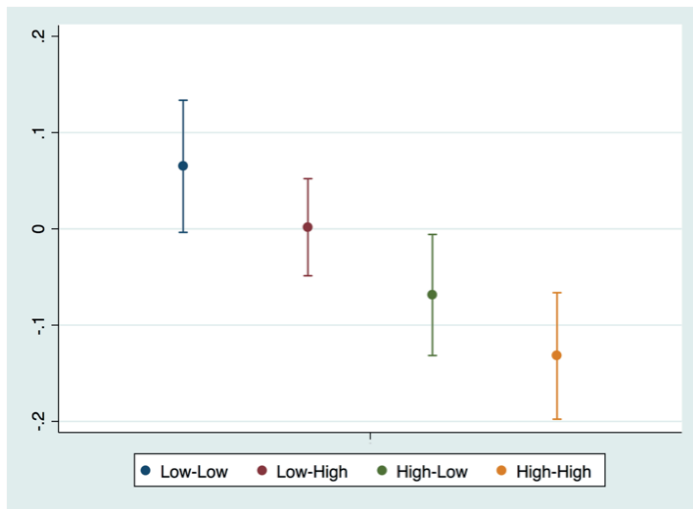
En nuestro reciente artículo ("Gender differences under test pressure and their impact on academic performance: a quasi-experimental design", IEB Working Paper 2018/21) estudiamos las diferencias de género en el rendimiento académico bajo diferentes niveles de presión cuando se hacen exámenes tipo test de respuesta múltiple. La introducción de la evaluación continua como método de evaluación en un curso universitario nos permite explotar un entorno cuasi-experimental en el que los mismos estudiantes hacen exámenes tipo test similares, pero bajo condiciones de presión diferentes. Esta configuración nos permite estructurar el conjunto de datos disponibles de dos formas y explotar cada una al máximo: secciones transversales independientes y datos de panel. En este estudio utilizamos datos de dos fuentes: datos administrativos y datos del curso. En el caso de los datos administrativos, fueron facilitados por la Facultad de Economía y Empresa de la Universidad de Barcelona. Aportan información demográfica y académica completa sobre todos los estudiantes que forman parte del curso Fundamentos de la Fiscalidad durante siete cursos académicos. Por su parte, los datos del curso fueron facilitados por el Grupo de Innovación Docente en Economía de los Impuestos (GIDEI). Contienen toda la información sobre los grupos y las notas obtenidas. La utilización de datos administrativos únicos nos proporciona un conjunto rico de información individual que nos permite controlar las características personales y del grupo de los estudiantes; características que también pueden determinar su rendimiento y, por tanto, afectar las diferencias de género.

Tabla 1: Diferencias de género: secciones transversales agrupadas vs datos de panel de dos periodos.

	Pooled Cross-Section		Datos de panel (dos periodos)	
	Nota EC (1)	Nota test final (2)	Pooled OLS (3)	RE (4)
Diferencia de género en la EC (presión baja)	0.032	-	0.0427	0.0427
Diferencia de género en el test final (presión alta)	-	-0.086***	-0.091***	-0.091***

Nota: EC se refiere a la evaluación continua.

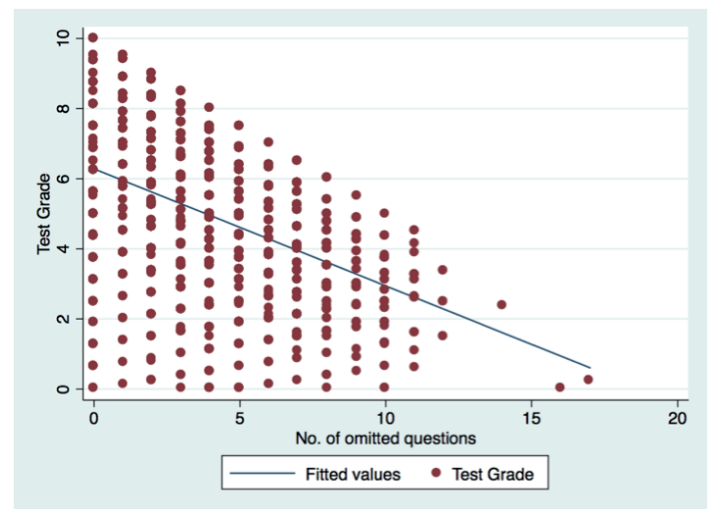
Figura 1: Diferencias de género ordenadas de menor a mayor escenario de presión.



Definimos la presión de un examen como la necesidad que experimentan los estudiantes de rendir mejor en la prueba como causa de la creciente importancia de la misma. Así, la principal fuente de presión sobre los estudiantes en nuestro estudio es el peso del examen sobre la nota final del curso. Además, puesto que durante los cursos académicos que analizamos algunas de las reglas de evaluación también cambiaron y modificaron la presión sobre los estudiantes a la hora de hacer el examen, realizamos un análisis de heterogeneidad. En este caso, la fuente secundaria de presión en los exámenes se identifica como las normas de evaluación que requieren un mejor rendimiento por parte de los estudiantes. La fortaleza de nuestro estudio cuasi-experimental se basa en cinco características específicas: (i) los exámenes tipo test se corrigen con ordenadores, por lo que no se dan correcciones con sesgo subjetivo, (ii) los exámenes tipo test presentan un formato casi idéntico (respuesta múltiple), utilizando las mismas preguntas, con el mismo nivel de dificultad y una estructura muy similar; (iii) la misma cohorte de estudiantes afronta estas pruebas en escenarios que se caracterizan por diferentes niveles de presión, (iv) la presión a la que están sometidos los estudiantes se analizan en un entorno real, es decir, cuando los estudiantes afrontan los exámenes que definen sus calificaciones (v) nuestros datos y la estrategia empírica nos permiten explorar los posibles mecanismos que son responsables de los resultados obtenidos, es decir, podemos identificar las razones principales de las diferencias de género observadas.

Los resultados empíricos (tabla 1) muestran que, después de controlar por las características individuales y grupales, los chicos rinden mejor que las chicas cuando realizan exámenes bajo gran presión, pero esta brecha de género se hace más pequeña a medida que la presión se reduce, hasta el punto que desaparece e, incluso, se da la vuelta en favor de las chicas (figura 1). Nuestras estimaciones con datos por secciones transversales agrupadas nos permiten obtener los principales resultados de cada evaluación por separado, para analizar diferencias entre cada distribución de notas y así poder evaluar el posible sesgo de autoselección relacionado con la decisión de hacer o no cada examen. El hecho de tener también datos de panel nos permite reformular el conjunto de datos y analizar la cuestión como un

Figura 2: Diagrama de dispersión de las notas obtenidas y el número de respuestas omitidas en el examen final tipo test.



juego secuencial de dos períodos, lo que nos permite confirmar nuestros principales resultados y, además, analizar los efectos heterogéneos relacionados con la presión de los exámenes.

También examinamos los potenciales mecanismos que pueden explicar esta brecha de género y constatamos que tienden a diferir según la fuente de presión. Cuando analizamos la presión atribuible al peso del examen sobre la nota final, las chicas parecen más propensas a sufrir la presión, mientras que los chicos mantienen su rendimiento. Por tanto, la brecha de género se produce por una caída en el rendimiento de las chicas. Así mismo, nuestro análisis de la presión de los exámenes atribuible a las normas del sistema de evaluación muestra que para los chicos es más probable que mejoren los resultados, mientras que el rendimiento de las chicas no cambia. Así, en este caso, la brecha de género se da por un aumento en el rendimiento de los chicos. Además, hay evidencia que sugieren que las mejores estudiantes dejan en blanco más preguntas que los chicos en los exámenes finales, cuando la presión aumenta (figura 2). En consecuencia, los sistemas educativos se deberían diseñar para evaluar conocimientos y habilidades de los estudiantes, y las calificaciones no deberían verse influenciadas por las estrategias de toma de decisiones de los estudiantes frente a exámenes tipo test.

Referencias

Blau, F. D. y 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, Investigador del IEB
montolio@ub.edu

Pere A. Taberner, Investigador de KSNET
peretaberner@gmail.com