

Desentrañando las diferencias de género en pruebas de opción múltiple: Un análisis detallado en las asignaturas de “Economía Pública” y “Public Economics” del Grado en Economía

Paloma Lanza León ^{1,2}

paloma.lanza@unican.es

Carla Blázquez Fernández ^{1,2}

carla.blazquez@unican.es

David Cantarero Prieto ^{1,2,3}

david.cantarero@unican.es

Natividad Fernández Gómez ^{1,2}

natividad.fernandez@unican.es

¹ *Universidad de Cantabria, Departamento de Economía, Facultad de Ciencias Económicas y Empresariales. Avda. los Castros, s/n, 39006, Santander, Cantabria, España.*

² *Grupo de Investigación en Economía de la Salud, Instituto de Investigación Sanitaria Valdecilla (IDIVAL).*

³ *Santander Financial Institute (SANFI).*

Recibido: 7 de marzo de 2024
Aceptado: 27 de junio de 2024

Resumen

Los resultados académicos de los estudiantes resultan fundamentales tanto para finalizar con éxito sus estudios como para encontrar un puesto de trabajo acorde a su formación. De esta forma, la innovación aquí presentada se basa en analizar si existen diferencias de género en el rendimiento académico al realizar pruebas de evaluación de tipo test/opción múltiple para superar una asignatura obligatoria del tercer curso del Grado en Economía, “Economía Pública” o “Public Economics”. Para contrarrestar el azar, se ha establecido una penalización para que las respuestas incorrectas/erróneas puntúen negativo. De esta manera, es posible analizar la heterogeneidad de las diferencias de género en la tendencia a dejar preguntas en blanco. Para realizar el análisis, hemos aprovechado que la asignatura obligatoria en inglés comienza a impartirse en la Universidad de Cantabria en el curso académico 2021/2022. Esto es, en el presente curso académico 2023/2024, se ha cumplido su tercer año de impartición durante el primer cuatrimestre (septiembre-enero). Así, la idea es comparar, los resultados de los exámenes tipo test realizados el día de la convocatoria ordinaria y ver si existen diferencias entre alumnos y alumnas (durante los dos últimos cursos académicos). Merece la pena mencionar que el examen final, tanto de la convocatoria ordinaria como de la extraordinaria, consta de dos partes diferenciadas. Por un lado, una parte con conocimientos teóricos basada en una serie de preguntas tipo test/repuesta opción múltiple. Por otro lado, una parte con supuestos prácticos y/o preguntas abiertas. Los resultados muestran que no existen grandes diferencias ni en el comportamiento ni en el rendimiento medio de mujeres y hombres. En todo caso, se pone de manifiesto la importancia de la evaluación continua, sistema por el que la mayor parte de los alumnos opta y es el que aquí se analiza.

Palabras clave: Diferencias (Brecha) de género, Prueba de opción múltiple, Resultados académicos, Economía pública.

Códigos JEL: A22, H0

1. INTRODUCCIÓN

Los resultados académicos de los estudiantes resultan fundamentales tanto para finalizar con éxito sus estudios como para encontrar un puesto de trabajo acorde a su formación. Las preguntas de opción múltiple son consideradas medidas objetivas y justas del rendimiento de los estudiantes. De hecho, este método de evaluación se utiliza habitualmente para evaluar los conocimientos de los candidatos en una gran variedad de situaciones. Refleja el conocimiento de la materia por parte del alumnado y, sin embargo, presenta, principalmente, una desventaja. Y es que los examinados pueden intentar adivinar la respuesta sin tener ningún conocimiento sobre ella o no estar seguros de la respuesta (Baldiga, 2014; Akyol et al., 2022). Como solución a este problema, existe la posibilidad de aplicar una puntuación negativa a las respuestas erróneas (penalización por adivinar). En función del conocimiento del estudiante acerca de la pregunta, su decisión de adivinar u omitir la pregunta dependerá tanto de sus propios conocimientos, como de su grado de aversión al riesgo y nivel de confianza. La equidad de las puntuaciones de las pruebas podría verse comprometida por este problema, dado que los distintos grupos pueden presentar diferentes comportamientos.

Por lo general, estas pruebas penalizan a los estudiantes más aversos al riesgo, soliendo ser estos las mujeres en comparación con los hombres (Anderson, 1989; Ben-Shakhar y Sinai, 1991; Walstad y Robson, 1997; Eckel y Grossman, 2008; Iriberry y Rey-Biel, 2021 o Riukula, 2023). La explicación radica en que cuando se penaliza por respuestas erróneas, es menos probable que las mujeres se arriesguen a responder de manera incorrecta. Es decir, ellas son menos propensas a adivinar y, por tanto, es posible que obtengan puntuaciones esperadas más bajas. De hecho, Anderson (1989) concluye que las mujeres obtendrían peores resultados que los hombres. Lo cual se justificaría, en parte, por un mayor nivel de abstencionismo. Esto es, las mujeres serían menos propensas a apostar cuando realmente no están seguras de la respuesta que los hombres. De manera intuitiva, la probabilidad de omitir una pregunta se incrementa a medida que la aversión al riesgo del estudiante aumenta y se reduce a medida que considera que tiene mayores posibilidades de responder correctamente (Coffman y Klinowski, 2020). Así, en caso de que las mujeres tengan una menor confianza en su capacidad de responder correctamente o muestren mayor aversión al riesgo, es posible que omitan más preguntas en comparación con sus homólogos masculinos, cuando existe una regla de puntuación diferencial para respuestas incorrectas y preguntas omitidas. Este comportamiento relacionado con la omisión de responder a un mayor número de preguntas puede perjudicar su rendimiento, obteniendo peores puntuaciones en las pruebas (y quizás, en las calificaciones finales) que los hombres, a pesar de poseer el mismo conocimiento sobre la materia (Conde-Ruiz et al., 2020). Es decir, este formato de examen podría generar disparidades de género. De hecho, durante las últimas décadas, el nivel educativo de las mujeres ha aumentado de modo que se ha aceptado y demostrado que ellas suelen estar más formadas que los hombres (Pascual-Saez y Lanza-Leon, 2023).

En este sentido, existe evidencia empírica que analiza la brecha de género en el rendimiento educativo. Hay estudios que sugieren que, en situaciones de presión competitiva, las mujeres rinden peor que los hombres en ciertas tareas (Pekkarinen, 2015). Diversos trabajos han analizado los posibles factores que pueden explicar este hecho. Entre los determinantes, destaca la asunción de riesgos, el exceso de confianza o las actitudes y rendimiento en competencia, entre otros (Baldiga, 2014). Además, otro factor que puede condicionar esta diferencia de género puede ser el nivel de dificultad de las preguntas del examen (Riener y Wagner, 2017). En concreto, los trabajos empíricos en este campo sugieren, tal y como ya se ha resaltado, que las mujeres son más propensas a saltarse preguntas que los

hombres en los exámenes tipo test/opción múltiple (Montolio y Taberner, 2021; Saygin y Atwater, 2021).

Uno de los primeros autores que analiza este tema es Swineford (1941), quien afirma que los hombres tienen una mayor tendencia a la hora de asumir riesgos al responder preguntas en exámenes en comparación con las mujeres. Existe cierta literatura que analiza las puntuaciones de las pruebas de acceso a determinadas universidades, como son los exámenes estandarizados SAT, que consiste en una prueba utilizada en Estados Unidos para la colocación y admisión en casi todos los niveles de escolarización. Sin embargo, los resultados no son del todo concluyentes en lo que a perjudicar a las estudiantes mujeres se refiere. Por su parte, Pekkarinen (2015) analiza las diferencias de género en el rendimiento en los exámenes de elección múltiple de acceso a universidades finlandesas en el ámbito de economía y administración de empresas. Además, Baldiga (2014) y Coffman y Klinowski (2020) muestran que el hecho de que las preguntas omitidas y las respuestas incorrectas no puntúen lo mismo tiene un impacto negativo en la brecha de género relativa al rendimiento. Más recientemente, Akyol et al. (2022) examinan las respuestas de los estudiantes en exámenes de acceso a la Universidad de Turquía, concluyendo que las mujeres son más aversas al riesgo, aunque los efectos negativos para ellas son pequeños.

Por otro lado, se ha comprobado que el efecto de penalización de las respuestas erróneas en los resultados de los exámenes depende de la especialidad (Coffman y Klinowski, 2020). Especialmente, esta brecha de género se ha analizado en el campo de las matemáticas (Atkins et al., 1991; Ramos y Lambating, 1996). Centrándose en dicho ámbito, Tannenbaum (2012) demuestra que las mujeres evitan responder más preguntas que los hombres debido a, entre otros motivos, la aversión al riesgo. Más allá de ese campo, y considerando un examen de microeconomía, Funk y Perrone (2016) estudian la brecha de género y las medidas de aversión al riesgo, sin obtener resultados claros sobre la existencia en cuanto a diferencias de género que resulten perjudiciales para las estudiantes mujeres. A pesar de que la mayoría de la literatura se centra en exámenes de acceso a la universidad o estudios universitarios, existe evidencia que analiza las diferencias de género en omitir respuestas a ciertas preguntas en pruebas de opción múltiple que se centran en estudiantes de secundaria o edades comprendidas entre los 8 y 18 años (Riener y Wagner, 2017; 2018). Los autores sugieren la existencia de diferencias de género para preguntas difíciles. Además, señalan que esta brecha podría no existir en caso de proporcionar recompensas externas por el rendimiento (Riener y Wagner, 2017). Si distinguen por grupos de edad, encuentran que las niñas, independientemente de su edad, responden a menos preguntas de opción múltiple que los niños (Riener y Wagner, 2018).

En base a todo lo anteriormente expuesto, la innovación aquí presentada se basa en analizar si existen diferencias de género en la forma en que los examinados responden a la incertidumbre sobre las respuestas correctas al realizar pruebas de evaluación de tipo test para superar una asignatura obligatoria del tercer curso del Grado en Economía, “Economía Pública” o “*Public Economics*” (impartida, ésta última, completamente en inglés), en el área de Ciencias Sociales. En concreto, se analiza si las mujeres son más propensas que los hombres a saltarse preguntas en lugar de “seleccionarlas” (“mayor nivel de abstencionismo”). Asimismo, esta investigación se centra en analizar el rendimiento académico en base a este comportamiento.

2. MATERIALES Y MÉTODO: DESCRIPCIÓN DE LA EXPERIENCIA DOCENTE EN EL CONTEXTO DE LAS ASIGNATURAS “ECONOMÍA PÚBLICA” Y “PUBLIC ECONOMICS”

Las asignaturas objeto de estudio en este trabajo, “Economía Pública” y “*Public Economics*”, se imparten en la Facultad de Ciencias Económicas y Empresariales de la Universidad de Cantabria. Dichas asignaturas son obligatorias con una carga lectiva de 6 créditos ECTS cada una y se desarrollan durante el primer cuatrimestre del tercer curso del Grado en Economía y Doble Grado en Administración y Dirección de Empresas y Economía (desde el curso académico 2021/2022, comienza a impartirse la asignatura en inglés). Ambas asignaturas son homólogas, presentan de manera muy similar el temario, los contenidos y el sistema de evaluación, con la única distinción del idioma de impartición. Para llevar a cabo este estudio, nos centraremos en los dos últimos cursos académicos impartidos. Esto es, 2022/2023 y 2023/2024.

Entre los objetivos establecidos para estas asignaturas, se pueden encontrar los siguientes: (1) razonar y fundamentar los motivos económicos y factores determinantes de la economía pública, (2) analizar la participación del sector público (tanto en gastos como en financiamiento) en la economía, (3) evaluar las consecuencias económicas de la tributación, (4) examinar las categorías de financiamiento público, y (5) comprender las aportaciones más actuales en el campo de la Economía Pública y los retos que se presentan en el futuro. Al finalizar la asignatura correspondiente, se espera que los estudiantes adquieran conocimientos teóricos en el ámbito de la economía pública, desarrollen habilidades para abordar y resolver problemas relacionados con el gasto y la financiación del sector público, y sean capaces de argumentar y comprender la relevancia de la intervención del sector público en una economía de mercado.

La organización docente de las asignaturas se divide en tres bloques, que aparecen sintetizados en la Tabla 1.

Tabla 1. Contenidos principales de las asignaturas “Economía Pública” y “*Public Economics*”

Bloque	Tema
1. Fundamentos teóricos del Sector Público	1.1. La economía del Sector Público y su objeto de estudio
2. Análisis económico de la intervención del Sector Público	2.1. La Hacienda Pública normativa 2.2. Teoría de bienes públicos y externalidades 2.3. La teoría de la elección colectiva
3. Ingresos públicos: Criterios normativos y efectos económicos	3.1. Incidencia de la imposición: Eficiencia y Equidad distributiva 3.2. Otros ingresos públicos

Fuente: Elaboración propia.

En lo relativo a los métodos de evaluación, las asignaturas de Grado se dividen en pruebas/actividades de evaluación continua (durante el cuatrimestre) y examen final. Cabe destacar que el sistema de evaluación se ve ligeramente modificado cuando hablamos de los dos cursos académicos considerados (2022/2023 y 2023/2024). En el primero de los cursos analizados (2022/2023), la calificación de la evaluación continua correspondía a un 50% de la calificación final, mientras este porcentaje se redujo a un 40% en el segundo de los cursos

analizados (2023/2024) para fomentar la asistencia y participación en clase de los estudiantes se han modificado bien el porcentaje de la participación activa (“*Public Economics*”) bien el peso de las pruebas de informática (“*Economía Pública*”). La Tabla 2 sintetiza los porcentajes establecidos para cada actividad, según el tipo de prueba.

Señalar en este punto que, si un estudiante no desea seguir la evaluación continua, en la convocatoria bien ordinaria bien extraordinaria (ambas convocatorias con la misma estructura de examen), se examinará sobre un total de 10 puntos (100% de la calificación total de las actividades recuperables de la asignatura).

Tabla 2. Porcentajes establecidos en los métodos de evaluación continua

Pruebas durante el cuatrimestre	Curso académico 2022/23		Curso académico 2023/24	
	Economía Pública	<i>Public Economics</i>	Economía Pública	<i>Public Economics</i>
<i>Prueba de informática</i>	15%	15%	30%	15%
<i>Trabajo en equipo</i>	25%	25%	0%	0%
<i>Participación activa (seminarios, flipped classroom, tareas y actividades en el aula...)</i>	10%	10%	10%	25%
Examen final	50%	50%	60%	60%

Fuente: Elaboración propia.

Las guías docentes completas (para los cursos académicos 2022/2023 y 2023/2024, en los cuales se realiza la experiencia docente) están disponibles en la página web de la Universidad de Cantabria.

Por lo tanto, el objetivo concreto de este estudio se centra en el análisis de una parte del examen final (en convocatoria ordinaria). Resulta, por tanto, fundamental entender cómo se configura el examen final (mismo formato en la convocatoria ordinaria que en la convocatoria extraordinaria). El examen combina una parte de conocimientos teóricos basada en una serie de preguntas tipo test/repuesta opción múltiple y otra parte de ejercicios y supuestos prácticos y/o preguntas abiertas (la parte de contenidos teóricos puntúa 3 puntos en continua y 6 en no continua, en ambas asignaturas). En concreto, tal y como hemos venido comentando, es la prueba de conocimientos teóricos la que se utiliza en este trabajo para mostrar la experiencia docente. Para esta prueba, los estudiantes disponen de los 30 primeros minutos que dura la prueba (1 hora y 30 minutos en total en 2022/2023 y 2 horas en total en 2023/2024 ya que se ampliaron los ejercicios establecidos en este último curso).

En cualquier caso, y para contrarrestar el azar, se ha establecido una penalización para que las respuestas incorrectas/erróneas puntúen negativo. De esta manera, es posible analizar la heterogeneidad de las diferencias de género en la tendencia a dejar preguntas en blanco. De esta forma, se establecen un total de 20 preguntas de opción múltiple en el examen final. Cada pregunta tiene cuatro posibles respuestas, de las cuales sólo una es correcta. El alumnado recibe/suma 0,15 puntos por cada pregunta correctamente respondida, resta 0,05 puntos por cada respuesta errónea, y las respuestas en blanco no suman ni restan puntos (para alumnos que siguen la evaluación continua y constituyen nuestra muestra principal - de hecho, sólo tres

alumnos de la asignatura en español, y para el curso 2023-2024, seguirían el sistema de “no continua”).

Obtenemos así datos a nivel individuo de todos los estudiantes que se examinaron del examen final de la convocatoria ordinaria de las asignaturas “Economía Pública” y “*Public Economics*” de la Universidad de Cantabria en los cursos académicos 2022/2023 y 2023/2024. Los datos que serán analizados en el siguiente apartado incluyen preguntas no respondidas, preguntas respondidas, y resultados para cada estudiante examinado, en cada curso analizado, así como datos sociodemográficos administrativos de cada individuo, en concreto, el sexo (no disponemos de más información).

3. RESULTADOS

La muestra final de este análisis consiste en un total de 157 estudiantes (87 correspondientes al curso académico 2022/2023 y 70 pertenecientes al curso 2023/2024), como se puede comprobar en la Tabla 3. Se muestra la información de manera desagregada acerca de la participación en las asignaturas. Además, dado el objetivo del presente trabajo, se diferencia en cuanto al género.

Tabla 3. Muestra final diferenciando por género y asignatura

	Curso académico 2022/2023			Curso académico 2023/2024		
	Economía Pública	<i>Public Economics</i>	Total	Economía Pública	<i>Public Economics</i>	Total
Mujeres	32	14	46	17	7	24
Hombres	32	9	41	36	10	46
Total	64	23	87	53	17	70

Fuente: Elaboración propia.

Tal y como se resumen en la Tabla 4, los resultados muestran que, a priori, y en promedio, las mujeres evitan responder más preguntas que los hombres (salvo para el curso 2023/2024 en “Economía Pública”).

Tabla 4. Promedio de respuestas respondidas y no respondidas por curso académico y género

	Preguntas no respondidas	Preguntas respondidas	Preguntas no respondidas (%)
Mujeres			
Economía Pública			
2022/2023	4.81	15.19	24.06%
2023/2024	2.29	17.71	11.47%
<i>Public Economics</i>			
2022/2023	3.64	16.36	18.21%
2023/2024	3.14	16.86	15.71%
Hombres			
Economía Pública			
2022/2023	3.41	16.59	17.03%
2023/2024	2.78	17.22	13.89%
<i>Public Economics</i>			
2022/2023	3.11	16.89	15.56%
2023/2024	2.60	17.40	13.00%

Fuente: Elaboración propia.

No obstante, al realizar el contraste de Chi-cuadrado para la muestra completa (N=157) se observa que no existen diferencias significativas por género al considerar las dos asignaturas y los dos cursos académicos (Pr = 0.302). Al analizar los resultados por asignatura, serían prácticamente estables para “*Public Economics*” (N= 40), esto es, no habría diferencias significativas por género al obtener una Pr = 0.684. Finalmente, al 10%, se podría señalar que existen diferencias por género en el caso de “Economía Pública” (N=117) con una Pr = 0.082. Si realizamos el contraste por curso académico, obtenemos que no hay diferencias significativas por género ni para el curso 2022/2023 (N=87) ni para el 2023/2024 (N=70). Con unos Pr = 0.248 y Pr = 0.901, respectivamente.

La siguiente cuestión a analizar son los resultados alcanzados en las correspondientes pruebas. En concreto, en la Tabla 5 se muestran las calificaciones obtenidas por los estudiantes distinguiendo por curso académico (2022/2023 y 2023/2024), género (mujeres y hombres) y asignatura (“Economía Pública” y “*Public Economics*”).

Tabla 5. Calificaciones por curso académico, género y asignatura

	Calificación test	Calificación examen final	Calificación final
Mujeres			
Economía Pública			
2022/2023	1.67	3.18	7.08
2023/2024	1.91	3.72	6.49
<i>Public Economics</i>			
2022/2023	1.56	2.25	6.85
2023/2024	1.04	2.16	5.16
Hombres			
Economía Pública			
2022/2023	1.80	3.08	6.64
2023/2024	1.87	3.46	6.25
<i>Public Economics</i>			
2022/2023	1.71	2.71	7.35
2023/2024	1.51	3.79	6.87

Fuente: Elaboración propia.

De esta forma, según los resultados de la Tabla 5, si se tiene en consideración la calificación final de la asignatura las mujeres presentarían mejores resultados que los hombres en “Economía Pública”, pero serían los hombres quienes obtienen mejores resultados en “*Public Economics*”. Para la calificación del examen, se observa la tendencia anterior. No obstante, para la calificación del test, los hombres presentarían mejores resultados (salvo para “Economía Pública” en el curso 2020/2024).

En todo caso, al realizar el contraste de Chi-cuadrado, dichas diferencias no serían significativas, ni para “Economía Pública” ($Pr = 0,424$, $Pr = 0,569$, $Pr = 0,585$; para la calificación final, el examen final y la prueba tipo test, respectivamente) ni para “*Public Economics*” ($Pr = 0,382$, $Pr = 0,422$, $Pr = 0,322$; para la calificación final, el examen final y la prueba tipo test, respectivamente). De igual forma, para la muestra completa ($N = 157$) los resultados indican que no existen diferencias significativas por género en la calificación final al considerar las dos asignaturas y los dos cursos académicos ($Pr = 0,150$). Tampoco aparecerían diferencias en el resultado del test ($Pr = 0,441$) ni en el examen final ($Pr = 0,633$).

En síntesis, no se puede establecer una correspondencia entre responder un mayor/menor número de preguntas y resultados. Este resultado puede deberse bien a que los estudiantes que estudian más estas asignaturas están más preparados en el examen final, bien a que el tipo test perjudica puede perjudicar a los más aversos al riesgo (no siendo necesariamente éstos las mujeres).

4. DISCUSIÓN Y CONCLUSIONES

Los resultados académicos de los estudiantes universitarios pueden medirse con distintas pruebas/actividades. En todo caso, si bien es cierto que, tras la creación del Espacio

Europeo de Educación Superior, las actividades de evaluación continua han ido adquiriendo un mayor peso en las calificaciones, los exámenes finales (ordinarios y extraordinarios) continúan siendo fundamentales para superar asignaturas obligatorias como la que se analiza en este trabajo. En este sentido, las preguntas de tipo test/opción múltiple han sido consideradas medidas objetivas del rendimiento de los estudiantes y por ello, parte de la evaluación de la asignatura a analizar en este estudio se basan en ellas. Así, este análisis es fundamental para comprender si se dan estas diferencias de género en la puntuación del examen de opción múltiple y en el rendimiento académico de los distintos grupos.

En concreto, en este trabajo se ha analizado la perspectiva de género en pruebas tipo test/de opción múltiple. Para ello, se han revisado los resultados de los exámenes ordinarios de las asignaturas de “Economía Pública” y “*Public Economics*” de la Universidad de Cantabria durante los dos últimos cursos académicos. Los resultados alcanzados con una muestra de 157 estudiantes (70 mujeres y 87 varones) muestran que no existen claras diferencias por género y curso académico.

Con todo, podríamos resumir que las alumnas no omiten un mayor número de preguntas que los estudiantes varones en las pruebas de opción múltiple, tal y como sugería la literatura (Ben-Shakhar y Sinai, 1991; Pekkarinen, 2015; Riukula, 2023; entre otros). Tampoco se evidencian diferencias en los resultados alcanzados por razón de género. La mayor parte de la evidencia empírica indicaba que las mujeres tienden a obtener peores resultados en los exámenes de opción múltiple cuando se dan reglas de puntuación diferencial entre preguntas omitidas y respuestas incorrectas (Baldiga, 2014; Iriberry y Rey-Biel, 2021; Akyol et al., 2022). Una de las razones indicadas se basa en la mayor aversión al riesgo o la falta de confianza por parte de las mujeres, lo que las hace menos dispuestas a adivinar y reduce su rendimiento. Es decir, las mujeres obtienen peores resultados que los hombres (Anderson, 1989). No obstante, en línea a nuestros resultados, Pekkarinen (2015) muestra que no existen grandes diferencias en el rendimiento medio de mujeres y hombres en las pruebas de opción múltiple de acceso a la universidad.

Adicionalmente, los resultados alcanzados ponen de manifiesto la importancia de la evaluación continua, sistema por el que la mayor parte de los alumnos opta y según el cual los estudiantes van acumulando puntos a lo largo del cuatrimestre (nuestra muestra), a la vez que minimizan los riesgos de tener que superar la asignatura con una única prueba al final del cuatrimestre.

REFERENCIAS

- Akyol, P., Key, J., & Krishna, K. (2022). Hit or miss? test taking behavior in multiple choice exams. *Annals of Economics and Statistics*, (147), 3-50.
- Anderson, J. (1989). Sex-related differences on objective tests among undergraduates. *Educational Studies in Mathematics*, 20(2), 165-177.
- Atkins, W. J., Leder, G. C., O'Halloran, P. J., Pollard, G. H., & Taylor, P. (1991). Measuring risk taking. *Educational Studies in Mathematics*, 22, 297-308.
- Baldiga, K. (2014). Gender differences in willingness to guess. *Management Science*, 60(2), 434-448.

- Ben-Shakhar, G., & Sinai, Y. (1991). Gender differences in multiple-choice tests: the role of differential guessing tendencies. *Journal of Educational Measurement*, 28(1), 23-35.
- Coffman, K. B., & Klinowski, D. (2020). The impact of penalties for wrong answers on the gender gap in test scores. *Proceedings of the National Academy of Sciences*, 117(16), 8794-8803.
- Conde-Ruiz, J. I., Ganuza, J. J., & García, M. (2020). Gender gap and multiple-choice exams in public selection processes. *Hacienda Pública Española*, (235), 11-28.
- Eckel, C. C., & Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. *Handbook of experimental economics results*, 1, 1061-1073.
- Funk, P., & Perrone, H. (2016). Gender differences in academic performance: The role of negative marking in multiple-choice exams. CEPR Discussion Papers: Technical report.
- Iriberry, N., & Rey-Biel, P. (2021). Brave boys and play-it-safe girls: Gender differences in willingness to guess in a large scale natural field experiment. *European Economic Review*, 131, 103603.
- Montolio, D., & Taberner, P. A. (2021). Gender differences under test pressure and their impact on academic performance: a quasi-experimental design. *Journal of Economic Behavior & Organization*, 191, 1065-1090.
- Pascual-Saez, M., & Lanza-Leon, P. (2023). The mismatch of over-education and earnings in Spain: how big are differences? *Applied Economics*, 55(46), 5369-5383.
- Pekkarinen, T. (2015). Gender differences in behaviour under competitive pressure: Evidence on omission patterns in university entrance examinations. *Journal of Economic Behavior & Organization*, 115, 94-110.
- Ramos, I., & Lambating, J. (1996). Gender differences in risk-taking behavior and their relationship to sat-mathematics performance. *School Science and Mathematics*, 96(4), 202-207.
- Riukula, K. (2023). Gender differences in multiple-choice questions and the risk of losing points. *Journal of the Finnish Economic Association*, 4(1).
- Riener, G., & Wagner, V. (2017). Shying away from demanding tasks? Experimental evidence on gender differences in answering multiple-choice questions. *Economics of Education Review*, 59, 43-62.
- Riener, G., & Wagner, V. (2018). Gender differences in willingness to compete and answering multiple-choice questions—The role of age. *Economics letters*, 164, 86-89.
- Saygin, P. O., & Atwater, A. (2021). Gender differences in leaving questions blank on high-stakes standardized tests. *Economics of Education Review*, 84, 102162.
- Swineford, F. (1941). Analysis of a personality trait. *Journal of Educational Psychology*, 32(6), 438.

Tannenbaum, D. I. (2012). Do gender differences in risk aversion explain the gender gap in SAT scores? Uncovering risk attitudes and the test score gap. *Unpublished paper*, University of Chicago, Chicago.

Walstad, W. B., & Robson, D. (1997). Differential item functioning and male-female differences on multiple-choice tests in economics. *The Journal of Economic Education*, 28(2), 155-171.

Unraveling Gender Differences in Multiple-Choice Tests: A Detailed Analysis in the 'Public Economics' Course of the Economics Degree.

Abstract

The academic outcomes of students are crucial both for successfully completing their studies and for securing a job that aligns with their education. The innovation presented here focuses on analysing whether there are gender differences in academic performance when taking multiple-choice tests to pass a mandatory third-year course in the Economics degree, "Public Economics" (both in Spanish and in English). To mitigate randomness, a penalty has been established so that incorrect answers score negatively. This approach allows for the examination of gender differences in the tendency to leave questions unanswered. To conduct the analysis, we took advantage of the fact that this mandatory course began in English at the University of Cantabria in the academic year 2021/2022. In the current academic year 2023/2024, it has completed its third year of delivery during the first semester (September-January). The idea is to compare the results of multiple-choice exams taken on the day of the regular examination and see if there are differences between male and female students (over the last two academic years). It is worth mentioning that the final exam, both in the regular and extraordinary sessions, consists of two differentiated parts. On one hand, a section with theoretical knowledge based on a series of multiple-choice test questions. On the other hand, a section with practical cases and/or open-ended questions. The results show that there are no major differences in the behaviour or average performance of women and men. In any case, the importance of continuous evaluation is evident, a system that most students opt for, and it is the one analysed here.

Key words: Gender Gap, Multiple-Choice Test, Academic Outcomes, Public Economics.

JEL Codes: A22, H0.