



2020

¿Mejora o retroceso? El impacto de la reforma a la PSU en la brecha de puntajes según colegio de egreso

Roberto Cases V.



**PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA
MAGISTER EN ECONOMIA**

**TESIS DE GRADO
MAGISTER EN ECONOMIA**

Cases, Vásquez, Roberto Alejandro

Julio, 2020



**PONTIFICIA UNIVERSIDAD CATOLICA DE CHILE
INSTITUTO DE ECONOMIA
MAGISTER EN ECONOMIA**

¿Mejora o retroceso?

**El impacto de la reforma a la PSU en la brecha de puntajes según colegio
de egreso**

Roberto Alejandro Cases Vásquez

Comisión

Nicolás Figueroa
Jeanne Lafortune

Santiago, Julio de 2020

¿Mejora o retroceso?

El impacto de la reforma a la PSU en la brecha de puntajes según colegio de egreso *

Roberto Alejandro Cases Vásquez **

Julio, 2020

Abstract

Resumen

Las pruebas estandarizadas de selección múltiple han jugado un rol importante en la admisión universitaria de muchos países del mundo. Una práctica común de este tipo de pruebas es la regla de penalizar las respuestas incorrectas con el fin de evitar que quienes rindan el test recurran a la adivinación. En este estudio, se demuestra teóricamente que la norma de penalización crea brechas en el rendimiento esperado entre dos grupos de distinto grado de aversión al riesgo, a pesar que no existan diferencias en el nivel de conocimiento entre los individuos. Adicionalmente, esta investigación aprovecha la variación causada por la reforma de la PSU del año 2015 en Chile, que elimina la regla de penalización por respuestas erróneas, para identificar el efecto causal de esta norma sobre la brecha de rendimiento entre colegios privados y no privados. Los resultados muestran una reducción significativa de dicha brecha correspondiente al 9% del promedio históricos de la brecha de rendimiento según el tipo del establecimiento educacional de egreso. Esta disminución se traduce en una caída de quince puntos porcentuales en la brecha referente a la probabilidad de estar matriculado en algún programa de estudios superiores y un aumento significativo en la probabilidad de cumplir los requisitos académicos en la postulación de becas de arancel. Los resultados sugieren que si se quiere fomentar la igualdad de oportunidades, la norma de penalización no debería ser considerada en futuros cambios en los mecanismos de selección universitaria.

*Se agradece al Departamento de Evaluación, Medición y Registro Educacional (DEMRE), de la Universidad de Chile, por facilitar las bases de datos del Proceso de Admisión a la Educación Superior Universitaria vía Prueba de Selección Universitaria para el desarrollo de esta investigación

**Esta tesis fue realizada bajo la tutela de la profesora Jeanne Lafortune y el profesor Nicolás Figueroa, a quienes agradezco profundamente sus comentarios, recomendaciones y por sobre todo su constante dedicación. Sin embargo, cualquier error es de exclusiva responsabilidad del autor. Correo de contacto racases@uc.cl

1. Introducción

En el mundo las pruebas estandarizadas de selección múltiple han jugado un papel importante en la educación terciaria. En Chile, la Prueba de Selección Universitaria (en adelante, PSU) es la principal determinante de quienes participan en la educación superior, a qué universidad o institución ingresan y a qué carrera se matriculan. Adicionalmente, la mayoría de las becas y créditos otorgadas por el gobierno de Chile tiene como requisito obtener un determinado puntaje mínimo en la PSU para poder postular, por lo que esta prueba de selección también tiene un rol en la elegibilidad en estos medios de financiamiento para muchos postulantes a la educación superior.

En general los tests de selección múltiple han sido utilizados para medir las habilidades de las personas y una práctica común en este tipo de pruebas es la regla de penalizar las respuestas incorrectas con el fin de evitar que quienes rindan el test recurran a la adivinación. En el contexto particular de Chile, antes del año 2015 los procesos de selección universitaria contaban con una regla para la PSU que otorgaba 1 punto por respuesta correcta, $-0,25$ puntos por respuesta incorrecta y 0 puntos por pregunta omitida. No obstante, en el año 2015 las autoridades encargadas de este test en Chile decidieron remover la penalización por respuesta incorrecta en la PSU tras seguir la recomendación de una auditoría externa.

En países con altos índices de desigualdad, como Chile, es relevante discutir si los instrumentos que determinan el ingreso de las personas a la educación superior contribuyen a la igualdad de oportunidades. Por ejemplo, resulta interesante analizar si al mismo nivel de conocimiento entre los alumnos, las reglas del instrumento de medición favorecen a personas con mejor *background* socioeconómico (o a estudiantes que egresan de colegios privados) y/o perjudica a personas con peores antecedentes socioeconómico (o a alumnos que pertenecieron a escuelas no privadas). Sin embargo, para el caso chileno y algunos países de América Latina, esta arista de los instrumentos de medición para la selección universitaria ha sido en gran medida olvidada dentro la literatura.

Un modelo sustentado por la teoría de utilidad esperada con un enfoque de aversión al riesgo, en un contexto donde un individuo se enfrenta a la decisión de responder una pregunta empleando la adivinación u omitirla en un test que penaliza las respuestas erróneas nos entrega una conclusión reveladora: individuos con mayor aversión al riesgo obtienen puntajes esperados menores a los individuos que toman posiciones más riesgosas, independiente de que ambos grupos tengan el mismo nivel de conocimiento. De esta forma, se predice que la eliminación de la penalización debería disminuir la brecha de rendimiento entre individuos con distinto grado de aversión al riesgo, particularmente en aquellos que pertenecen a la cola inferior de la distribución de habilidad.

En base a este modelo, la presente investigación aprovecha el cambio de política en el proceso

de selección universitaria de Chile, a fin de examinar el efecto de la eliminación del descuento por respuestas erradas de la PSU en las brechas de rendimiento según la dependencia del establecimiento secundario de origen. Se postula que alumnos con peores situaciones socioeconómicas tienen un mayor grado de aversión al riesgo, por lo que se ven perjudicados ante la existencia de la norma de penalización. El análisis utiliza datos administrativos de gran escala proporcionados por el Departamento de Evaluación, Medición y Registro Educativo (DEMRE). Adicionalmente, se pretende evaluar el efecto de esta política en resultados de largo plazo, analizando cómo la eliminación de la penalización por respuestas incorrectas impacta a la brecha en la probabilidad de cumplir los requisitos de puntajes promedio PSU¹ para recibir becas y/o créditos del estado y el impacto en las brechas de los resultados de matrícula reflejado en la probabilidad de ser admitido en algún programa de estudios universitarios y en la calidad de los programas de educación superior al que se matricula un postulante según la dependencia de su colegio de origen.

La pregunta de esta investigación se sustenta en la evidencia presente en la literatura de aversión al riesgo, la cual señala que las personas con peor *background* socioeconómico tienden a ser más aversas que las personas con mejores ingresos. Si un test posee penalización por respuesta incorrecta, ambos grupos de la población utilizarán la adivinación con distinta probabilidad a la hora de responder, lo que provoca una brecha entre grupos socioeconómicos en la tasa de respuestas omitidas en la PSU. Esta diferencia en la omisión de preguntas repercute en el rendimiento obtenido en la prueba, generando una brecha en los resultados de la PSU entre ambos subconjuntos de personas, la cual desaparecería en el caso que se elimine la penalización por respuestas erróneas.

En línea con lo anterior, a lo largo de este estudio se demuestra que existe una clara correlación positiva entre tener más ingresos y pertenecer a un colegio privado, y una correlación positiva entre tener menos ingresos y estar vinculado a un colegio público o de copago, por lo que si la aversión al riesgo depende de las características socioeconómicas de las personas, esta también debería variar por el tipo de colegio al que pertenece el alumno. Estas relaciones y mecanismos conducen la hipótesis que el cambio de política presente en el proceso de selección universitaria en Chile en el año 2015 impacta de manera positiva en la reducción de la brecha de rendimiento en la PSU entre colegios privados y no privados. A su vez, es intuitivo pensar que esta disminución en la brecha en los resultados en la PSU (a favor de los alumnos de escuelas no privadas) se traduzca en una mayor matrícula en programas de estudios universitarios y un incremento en el porcentaje de inscritos pertenecientes a este tipo de establecimientos que cumplen los requisitos académicos para becas y créditos estatales.

En efecto, la eliminación de la penalización por respuestas erróneas se traduce en una reducción

¹En el contexto de postulación a becas y créditos para la educación superior, el puntaje promedio PSU se define como el promedio simple de los puntajes obtenidos en las pruebas de Lenguaje y Matemáticas.

estadísticamente significativa de 13.51 preguntas en la brecha de preguntas omitidas entre colegios privados y no privados, la cual conlleva una disminución del 9% en la brecha de rendimientos según el tipo del establecimiento educacional de egreso, respecto a su promedio histórico en el período previo al cambio de política. Adicionalmente, las estimaciones sugieren que las mujeres presentan un impacto de mayor magnitud en la reducción del *gap* de rendimiento entre los tipos de establecimientos educacionales que en el caso de los hombres. A su vez, los resultados empíricos demuestran que la magnitud del impacto del cambio de política no necesariamente es creciente en el nivel de habilidad o conocimiento de los alumnos.

De igual manera, la reforma en la Prueba de Selección Universitaria trae consigo una reducción en la brecha general de puntajes PSU entre colegios privados y no privados de 5.24 puntos. Esta disminución es mayor en los tests de índole humanistas (Historia y Lenguaje) que en los relacionados con las ciencias exactas (Matemáticas y Ciencias Naturales). Esta mejora relativa en el rendimiento (favorable para los establecimientos no privados) se manifiesta en los resultados de las matriculas en programas de estudios universitarios y en el cumplimiento de los requisitos académicos en el proceso de postulación de becas. En efecto, la política tiene como consecuencia una reducción de quince puntos porcentuales en la brecha referente a la probabilidad de estar matriculado en alguna carrera universitaria de alguna institución perteneciente al Consejo de Rectores. Asimismo, se estima una reducción significativa de 1,3% en la brecha sobre la probabilidad de cumplir los requisitos académicos para la postulación general de becas. Dicha disminución alcanza un valor estimado de 3.2% en el caso particular del requisito académico de la Beca Vocación de Profesor (BVP).

El resto de la investigación se estructura de la siguiente forma: la sección 2 describe la literatura relacionada a este estudio. En la sección 3 se desarrolla el modelo teórico que motiva la pregunta de investigación y se establece el mecanismo por el cual la eliminación de la regla de penalización por respuestas incorrectas impacta a las brechas de interés. En la sección 4 se describen los datos y la construcción de las muestras a utilizar. En la sección 5 se presenta la estrategia empírica, su justificación y el supuesto de identificación en el cual se sustenta. La sección 6 muestra las principales estimaciones de dicha estrategia, mientras que la sección 7 caracteriza las consecuencias de la política. Finalmente, en la sección 8 se discuten los resultados de la investigación.

2. Antecedentes y Literatura Relacionada

Con el transcurso del tiempo, los mecanismos de selección universitaria en Chile han sufrido diversas modificaciones. A finales de la década de 1960, se implementa por primera vez, un mecanismo único de selección y admisión para todas las instituciones de educación superior a nivel nacional, denominado Prueba de Aptitud Académica (PAA). Este instrumento de medición se mantuvo vigente durante 35 años; sin embargo, tras un intenso debate público, fue reemplazado por

un nuevo test estandarizado designado como Prueba de Selección Universitaria (PSU), el cual es administrado por el Departamento de Evaluación, Medición y Registro Educativo (DEMRE).

Desde entonces, la PSU se realiza de manera anual como parte del proceso de selección universitaria y consta de cuatro instrumentos de medición independientes: dos tests obligatorios (uno de Matemáticas y otro de Lenguaje y Comunicación) y dos optativos (uno de Ciencias y otro de Historia y Ciencias Sociales). Los contenidos de estas pruebas se focalizan en los Contenidos Mínimos Obligatorios (CMO) y se encuentran alineados al nuevo Marco Curricular de la Enseñanza Media establecido por el Ministerio de Educación. En principio, se consideraron una serie de contenidos reducidos que se han ido ampliando gradualmente conforme a nuevas exigencias curriculares.

Los cuatro tests que componen la PSU son administrados durante dos días consecutivos entre la última semana de Noviembre y las primeras semanas de Diciembre y en cada día se toman dos pruebas. Durante estas jornadas los postulante deben rendir de manera presencial las pruebas obligatorias (Lenguaje y Matemática) y al menos una prueba opcional (Historia o Ciencias). Estos tests estandarizados son de selección múltiple y abarcan entre 70 y 80 preguntas, cada una de ellas con cinco posibles respuestas donde sólo una es correcta.

La publicación de los resultados de la PSU se demora alrededor de un mes y posterior a ello, los inscritos comienzan el proceso de postulación a la universidades del Consejo de Rectores (CRUCH) a través del Sistema Único de Admisión. En esta etapa los postulantes deben indicar sus preferencias acerca de las carreras a las que desean ingresar y en que universidad pretende desarrollar sus estudios de educación superior. Posteriormente, las universidades seleccionan a los postulantes tomando como insumo los puntajes obtenidos en los tests obligatorios y optativos, el puntaje NEM calculado a partir del promedio de notas de enseñanza media y desde el 2013, el puntaje ranking del postulante². Estos puntajes se unifican obteniendo un puntaje ponderado de acuerdo a ponderadores que varían según el programa de estudios y la universidad a los cuales una persona postula. Las universidades ordenan a los postulantes de cada programa de estudios según sus puntajes ponderados y seleccionan a sus futuros estudiantes considerando el número de cupos de cada programa.

Un aspecto fundamental de los instrumentos de medición utilizados para los procesos de selección universitaria es la validez predictiva de las pruebas en términos de conocimientos, habilidades y rendimiento futuro en la educación superior. La relevancia de estos factores radica en que son utilizados como mecanismos de ordenamiento de los postulantes, por lo que determinan en gran medida la posterior selección e ingreso a las instituciones de educación superior. En ese marco, los

²Según el DEMRE, “el puntaje ranking expresa la posición relativa del estudiante en cada contexto educativo en el cual estuvo durante su Enseñanza Media, tomando como referencia el desempeño de los estudiantes en las últimas tres generaciones de dicho contexto”.

autores Koljatic and Silva (2010) elaboraron un estudio empírico para analizar la implementación de la PSU en reemplazo de la PAA. Su investigación concluye que el aumento de los contenidos evaluados en los test de medición no garantiza una mejoría sustantiva en la capacidad predictiva de las pruebas de selección universitaria en el país.

En esa misma línea, existe literatura que examina si las características presentes en el diseño de instrumentos de medición utilizados en la selección universitaria perpetúan desigualdades económicas. A modo de ejemplo, se ha estudiado si las particularidades de las pruebas de selección, benefician o perjudican a un grupo determinado de la población en base a características sociodemográficas, tales como el historial socioeconómico del postulante o su colegio de origen. En ese sentido, el estudio correlacional de Koljatic et al. (2013), compara los test de aptitudes (como la PAA) con las pruebas de rendimiento (como la PSU) en base a los cambios de los instrumentos de selección observados en Chile. Los investigadores comprueban que los grupos menos aventajados socioeconómicamente se ven beneficiados por el cambio de pruebas de aptitudes a pruebas de rendimiento; sin embargo, en sistemas de acceso de oportunidades educacionales desiguales, los beneficios adicionales de los tests de rendimiento no son independientes del número de contenidos evaluados en las pruebas de selección. En otras palabras, los autores señalan que cuando los tests de rendimiento evalúan un gran número de contenidos los beneficios marginales desaparecen, por lo que es fundamental asegurar que los contenidos curriculares sean predictores confiables del éxito en el primer año de universidad para garantizar que las pruebas de rendimiento contribuyan a fortalecer la equidad de oportunidades educacionales.

De igual forma, los autores Koljatic and Silva (2010) observan que al controlar por la variable “nivel socioeconómico”, el aumento de contenidos en la PSU ha sido perjudicial para los estudiantes de grupos vulnerables, especialmente para aquellos que provienen de la educación técnico-profesional. En esa misma línea investigativa, Beyer (2009) analiza la evolución de la igualdad de oportunidades en términos de la selección universitaria en Chile comparando la prueba PSU con la PAA, demostrando que la nueva prueba de selección universitaria parece haber aumentado el peso del capital económico y cultural en la determinación de los desempeños de los estudiantes en la misma prueba, por lo que concluye que el cambio de las pruebas de selección universitaria ha retrocedido marginalmente la igualdad de oportunidades educacionales en Chile.

Otro aspecto fundamental que se debe considerar sobre el diseño de los instrumentos de medición es el tipo de prueba que se va a desarrollar para la selección universitaria. Dentro de las categorías de estas pruebas, destacan los test de selección múltiple de alternativas, los cuales han sido ampliamente utilizados para medir las diversas habilidades de las personas. En ese marco, una regla frecuente en este tipo de pruebas es penalizar las respuestas incorrectas, con el fin de evitar que los individuos que rinden el test recurran a la adivinación de las respuestas. De esta manera,

diversas investigaciones se han centrado en el impacto de esta penalización en algunas variables de resultados, poniendo especial énfasis en las brechas de género de los puntajes obtenidos en este tipo de pruebas. A modo de ejemplo, Ben-Shakhar and Sinai (1991) examinan las diferencias presentes en los resultados de hombres y mujeres en test de selección múltiple desarrolladas por dos muestras diferentes en Israel: una correspondiente a pruebas de conocimiento en escuelas secundarias y otra en el contexto de selección universitaria. En ese sentido, los autores evidencian que los hombres tienen una ventaja comparativa en el desarrollo de las pruebas en ambas muestras de estudio, y a su vez, señalan que las mujeres manifiestan una mayor tasa de omisión en comparación a los hombres.

Siguiendo con esa línea, el estudio experimental de Baldiga (2014), revela que a igualdad de conocimiento de los sujetos de estudio, las mujeres tienden a omitir más que los hombres en los tests de selección múltiple, y señala que remover la penalización por respuestas incorrectas genera que todas las personas contesten la totalidad de las preguntas, eliminando la brecha de género en los puntajes obtenidos. En otra investigación, Coffman and Klinowski (2020) analizan el impacto en el gap de puntajes entre hombres y mujeres utilizando el cambio en las reglas de la PSU en el año 2015, los hallazgos del estudio indican que la nueva regla reduce la brecha de respuestas omitidas y disminuye la brecha en los puntajes obtenidos favoreciendo a las mujeres. De esta forma, los autores proponen a las diferencias de género en el grado de aversión al riesgo como el principal mecanismo por el cual la política genera un impacto. Dado lo anterior, en el análisis del impacto de la reforma de la PSU sobre las brechas de rendimiento según el tipo de colegio de egreso, se evaluará cómo correlaciona el cambio en dicha brecha con las diferencias de género.

De igual forma, el mecanismo de aversión al riesgo también es considerado en la investigación de Akyol et al. (2016). Este estudio modela y estima la decisión de contestar preguntas en test de selección múltiple con regla de penalización, concluyendo que la penalización disminuye la adivinación de las respuestas, y en consecuencia, aumenta la precisión de la medición considerablemente. Aun así, la regla reduce el puntaje esperado de las personas que son más aversas al riesgo, es decir, existe una discriminación hacia los grupos de personas que presentan un mayor grado de aversión. Adicionalmente, el estudio utiliza data de Turquía para documentar diferencias de aversión al riesgo por género y habilidad demostrando que las mujeres y las personas con mayor habilidad tienden a ser más aversos al riesgo.

Para justificar mi acercamiento empírico, se debe argumentar que existe un vínculo entre la aversión al riesgo y algunos aspectos socioeconómicos de las personas. Hartog et al. (2002) utiliza una muestra de tres encuestas diferentes para caracterizar a la personas según su grado de aversión, encontrando que la aversión al riesgo según la medida de Arrow-Pratt cae con el ingreso y la riqueza. Por otro lado, Guiso and Paiella (2008) analizan una encuesta realizada por el Banco de Italia en el año 1995 la cual incluye preguntas respecto a las posiciones de riesgo que toman los individuos.

Los autores encuentran que el grado de aversión al riesgo absoluta es una función decreciente de las dotaciones de los consumidores. En un estudio de naturaleza experimental, Jerrim et al. (2019) analizan la propensión a declarar experiencia en tres construcciones matemáticas que realmente no existen, utilizando el Programa para la Evaluación Internacional de Estudiantes (PISA) para caracterizar a las personas *bullshitters*. Los autores encuentran que las personas con ventajas en el *background* socioeconómico tienden a mentir más sobre el conocimiento de estos constructos matemáticos falsos y por lo tanto, estarían más dispuestos a utilizar la adivinación a la hora de responder un test. Por otra parte, en el estudio de Dohmen et al. (2010) se analiza la existencia de una relación entre las habilidades cognitivas y la aversión al riesgo en el contexto de Alemania. Los resultados de los autores revelan que la aversión al riesgo varía sistemáticamente con las habilidades cognitivas; en efecto, los individuos con mejores habilidades están significativamente más dispuestos a tomar riesgos. En resumen, la literatura ha mostrado evidencia consistente que las personas con mejores antecedentes socioeconómicos y mayores habilidades cognitivas tienden a ser menos aversas al riesgo.

Sumado a lo anterior, cabe destacar que la presente investigación se posiciona como una de las primeras en examinar el cambio de la norma de descuento por respuestas erradas en la PSU en términos del análisis de las brechas de los rendimientos desde una perspectiva socioeconómica y su efecto en relación a los resultados de matriculas y acceso a medios de financiamiento de los programas de educación superior. La evidencia presentada en este estudio es un aporte relevante para las instituciones encargadas del diseño e implementación de futuros instrumentos utilizados en el proceso de selección universitario chileno y si bien la validez externa de los resultados de esta investigación es cuestionable, la experiencia de Chile puede ser un punto de partida importante para posibles reformas en los procesos de selección de otros países similares. Por último, este estudio se presenta como un ejercicio empírico relacionado a la literatura de diseño de mecanismos, y entrega evidencia sustantiva sobre cómo los cambios en las reglas de una prueba de selección pueden modificar el ordenamiento de las personas en términos de la posición en la distribución de rendimiento en cada prueba y en el posterior ingreso a la educación universitaria.

3. Modelo Teórico

En esta sección se desarrolla un modelo teórico basado en la teoría de utilidad esperada con un enfoque de aversión al riesgo para analizar la manera en la que un agente contesta racionalmente una pregunta de un test estandarizado de selección múltiple como la Prueba de Selección Universitaria. Con ello, se establece el principal mecanismo mediante el cual la regla de penalización de incorrectas puede abrir una brecha entre los puntajes esperados de dos grupos de una población, la cual no necesariamente se debe a las diferencias en el nivel de conocimiento entre estos grupos.

En primer lugar se considerará la decisión marginal en el desarrollo de una prueba de dos tipos de personas con igual nivel de conocimiento pero con distinto grado de aversión al riesgo para examinar si en este escenario la regla de penalización crea una brecha en los resultados obtenidos en un test de selección múltiple, a pesar de que no existan diferencia en el conocimiento. Posteriormente, se realizará un estudio de estática comparativa con el fin de analizar cómo varía la magnitud de esta brecha ante cambios en el nivel de conocimiento de la población y/o la dificultad relativa de diferentes pruebas de selección.

Con el fin de simplificar la modelación, supongamos que una persona de tipo i debe responder marginalmente una pregunta de una prueba estandarizada. Primero, asumamos que la regla de evaluación consiste en la penalización de respuestas incorrectas igual a la PSU antes de la política, es decir, una persona tipo i obtiene 1 punto por respuesta correcta, $-0,25$ puntos por respuesta incorrecta y 0 puntos por pregunta omitida. Además, supongamos que el test consiste en preguntas de selección múltiple con cinco alternativas (A, B, C, D y E) de las cuales sólo una de ellas es correcta y sin pérdida de generalidad, se asume que la alternativa correcta es la A. Sumado a lo anterior, se supone que, condicional a las creencias de cada individuo, una persona tipo i le asigna una probabilidad p a que la alternativa correcta es la alternativa A. En otras palabras, si el individuo decide contestar la pregunta en cuestión, con probabilidad p contestará de manera correcta y con probabilidad $1 - p$ su respuesta será errónea.

Supongamos que las preferencias de la persona tipo i cumplen con los axiomas de von Neumann and Morgenstern (1944), por lo que estas pueden ser representadas mediante utilidad esperada. La función bernoulli de una persona tipo i esta dada por $u_i = u(x; \sigma_i)$ donde x representa un pago asociado a el status de la pregunta que responde el individuo (correcta, incorrecta u omitida) y σ_i es un parámetro de aversión al riesgo, o bien, σ_i determina el nivel de cóncavidad de la función u_i ³. Adicionalmente, supongamos que u es creciente en x y que el parámetro σ_i proviene de una distribución de probabilidad F_i cuyo soporte está definido por el dominio para σ_i en la función u .

Imaginemos que en la población que rinde la prueba existen dos tipos de personas (esto es, $i \in \{1, 2\}$), por ejemplo, personas con bajo y alto nivel de ingresos, alumnos que provienen de colegios públicos y privados, etc. Supongamos que estos grupos de personas se diferencian por su grado de aversión al riesgo y que el nivel de aversión es creciente en el parámetro σ_i ⁴. Lo anterior, también se puede expresar cómo que el coeficiente de Arrow-Pratt de aversión al riesgo absoluta es creciente en σ_i (esto es, $\frac{\partial}{\partial \sigma_i} \left[\frac{-u''(x, \sigma_i)}{u'(x, \sigma_i)} \right] \geq 0$). Sin pérdida de generalidad se asume que el individuo tipo 1 es más averso al riesgo que el individuo tipo 2, es decir, $\sigma_1 > \sigma_2$. Finalmente, se supone que el individuo más averso al riesgo tiene una distribución de probabilidad de σ que domina estocásticamente a la

³En la sección A.1 del Apéndice se discute un ejemplo de la función u_i .

⁴Una manera alternativa de plantear lo anterior, sería definir la relación "más averso al riesgo". Una persona i se dice que es más averso a al riesgo que una persona j ($u(x, \sigma_i) \succeq u(x, \sigma_j)$) si $\sigma_i > \sigma_j$.

distribución de probabilidad del individuo menos averso, es decir, $F_1(\sigma) \leq F_2(\sigma)$ para todo σ .

Una persona i que enfrenta la decisión de responder (recurriendo a la adivinación) u omitir una pregunta, se enfrenta a la elección de dos loterías:

- *Lotería R:*
Ganar 1 punto con probabilidad p .
Perder 0.25 puntos con probabilidad p .
- *Lotería O:*
Obtener 0 puntos con probabilidad 1.

En consecuencia, una persona i decide responder recurriendo a la adivinación si y solo si la utilidad esperada de responder es mayor o igual a la utilidad esperada de omitir la pregunta. Lo anterior se expresa en la siguiente desigualdad:

$$pu(1, \sigma_i) + (1 - p)u(-0.25, \sigma_i) \geq u(0, \sigma_i) \quad (1)$$

Es fácil notar que esta condición se cumple si y sólo si σ_i es menor que un determinado valor $\bar{\sigma}$ ($\sigma_i \leq \bar{\sigma}$) donde $\bar{\sigma}$ es tal que ambas loterías reportan la misma utilidad esperada, es decir:

$$pu(1, \bar{\sigma}) + (1 - p)u(-0.25, \bar{\sigma}) = u(0, \bar{\sigma}) \quad (2)$$

En otras palabras, si una persona es muy aversa al riesgo y supera un determinado valor límite para el parámetro de aversión, no contestará la pregunta y omitirá. Dado lo anterior, las personas más aversas al riesgo contestarán con probabilidad $F_1(\bar{\sigma})$ mientras que las personas menos aversas al riesgo contestarán con probabilidad $F_2(\bar{\sigma})$.

Cabe destacar que de la ecuación (2) se puede inferir que el valor de $\bar{\sigma}$ dependerá de p , por lo que se puede escribir como función de las creencias de los individuos. En esta línea, una pregunta importante es cómo se comporta la función $\bar{\sigma}(p)$ ante cambios en la probabilidad de responder correctamente. Para responder esta interrogante, es útil notar que $\bar{\sigma}(p)$ es el grado de aversión al riesgo asociado a un equivalente cierto para la lotería R tal que el pago certero es cero, por lo que la ecuación (2) se puede reescribir de la siguiente forma:

$$pu(1, \sigma) + (1 - p)u(-0.25, \sigma) = u(c(\sigma, p), \sigma) \quad (3)$$

donde $c(\sigma, p)$ es el equivalente cierto de la lotería R que depende del grado de aversión al riesgo σ y la probabilidad de responder correctamente. En este caso, $\bar{\sigma}(p)$ es tal que $c(\bar{\sigma}(p), p) = 0$. Si diferenciamos ambos lados de esta última igualdad respecto a p se obtiene lo siguiente:

$$\begin{aligned} c(\sigma, p) &= 0 \\ \frac{\partial c(\sigma, p)}{\partial p} &= 0 \\ \frac{\partial c}{\partial \sigma} \frac{\partial \bar{\sigma}}{\partial p} + \frac{\partial c}{\partial p} &= 0 \end{aligned}$$

Al despejar la derivada de $\bar{\sigma}$ respecto a p se obtiene como resultado la siguiente expresión:

$$\frac{\partial \bar{\sigma}}{\partial p} = - \frac{\frac{\partial c}{\partial p}}{\frac{\partial c}{\partial \sigma}} \quad (4)$$

El numerador de la expresión del lado derecho de la ecuación (4) es la derivada parcial del equivalente cierto respecto a la probabilidad de éxito en la pregunta contestada por el alumno. Intuitivamente, ante un aumento en p , el valor de la utilidad esperada obtenida en la lotería R es más grande y en consecuencia, el pago certero que deja indiferente al individuo entre la lotería R y dicho pago debe aumentar. Dicho de otra forma, el aumento de p se traduce en un incremento del equivalente cierto y por ende, el término $\frac{\partial c}{\partial p}$ tiene signo positivo. Por su lado, el denominador de la ecuación (4) es la derivada parcial del equivalente cierto respecto al parámetro de aversión al riesgo. En Mas-Colell et al. (1995), los autores estipulan la equivalencia entre que el coeficiente de Arrow-Pratt de aversión al riesgo absoluto sea creciente en σ y que el equivalente cierto de una lotería sea decreciente en el grado de aversión al riesgo, por lo que a raíz de los supuestos descritos anteriormente se infiere que $\frac{\partial c(\sigma, p)}{\partial \sigma} < 0$ ⁵. En consecuencia, el signo para la derivada parcial de $\bar{\sigma}(p)$ respecto a la probabilidad de contestar de manera correcta es positivo, lo cual nos indica que ante un incremento de p el valor de $\bar{\sigma}(p)$ aumenta y la condición para que un individuo conteste la pregunta en cuestión se cumple para un mayor rango de valores de σ_i ⁶.

Una vez analizada la forma en que los individuos deciden si contestan una pregunta o la omiten ante una regla de penalización para las respuestas incorrectas, cabe preguntarse cuáles son las consecuencias sobre el puntaje esperado de quienes rinden la prueba. Por un lado, la subpoblación que es más aversa al riesgo obtiene como puntaje esperado $[1 - F_1(\bar{\sigma}(p))] * 0 + F_1(\bar{\sigma}(p)) * [p * 1 - (1 - p) * \frac{1}{4}] = F_1(\bar{\sigma}(p)) \frac{5p-1}{4}$, mientras que la subpoblación de individuos que son menos aversos al riesgo obtienen

⁵Mas-Colell et al. (1995) Capítulo VI.

⁶En efecto, la Figura A.1 del Apéndice muestra la función $\bar{\sigma}(p)$ utilizando una función de utilidad que cumple con los supuestos presentes en el modelo y muestra una clara relación monótona creciente entre la probabilidad de éxito para un alumno y $\bar{\sigma}$.

$F_2(\bar{\sigma}(p))^{\frac{5p-1}{4}}$. Debido a la condición de dominancia estocástica de primer orden es posible notar que $F_1(\bar{\sigma})^{\frac{5p-1}{4}} < F_2(\bar{\sigma})^{\frac{5p-1}{4}}$, por lo que individuos con más aversión al riesgo obtienen en promedio menos puntaje que aquellos que tienen un coeficiente de aversión menor. Dicho de otra forma, la existencia de la regla de penalización establece una brecha de puntajes promedios que no depende del nivel de conocimiento de los individuos. Esta brecha está dada por la diferencia de puntaje esperado entre los dos grupos:

$$Brecha = \frac{5p-1}{4} [F_2(\bar{\sigma}(p)) - F_1(\bar{\sigma}(p))] \quad (5)$$

Ahora supongamos que la prueba no posee una regla que penaliza las respuestas incorrectas, es decir, una persona obtiene 1 punto por respuesta correcta, 0 puntos por respuesta incorrecta y 0 puntos por pregunta omitida. Es fácil notar que tras la modificación de este supuesto todas las personas (independiente de su grado de aversión al riesgo) deciden responder y por lo tanto, ambos grupos de personas obtienen el mismo puntaje esperado igual a p . En consecuencia, remover la regla de penalización por respuestas incorrectas disminuye completamente la brecha de puntajes entre grupos que, a pesar de tener igual nivel de conocimiento, tienen un nivel de aversión al riesgo distintos.

En este modelo, el parámetro p puede ser interpretado como una métrica del conocimiento del alumno ya que si un individuo se familiariza con los contenidos de una pregunta y/o prueba es más probable que conteste de manera correcta, mientras que una persona que no conoce los tópicos relacionados con la pregunta tendrá menos éxito en la contestación de esta. En esa misma línea, el parámetro p también está directamente relacionado con la dificultad de una prueba en particular. Por ejemplo, si en un tipo de test es más fácil descartar alternativas al momento de contestar, la probabilidad de éxito es mayor que en un tipo de prueba que no es posible descartar alternativas o que tiene opciones que son distractores y operan como “falsas correctas”. Por consiguiente, es plausible pensar que la reducción de la brecha de puntajes entre dos grupos (que poseen distinto grado de aversión al riesgo) producto de la eliminación de la penalización por respuesta incorrectas varíe según el valor del parámetro p . No obstante, la dirección del efecto de un movimiento en p sobre la reducción de la brecha no es trivial y es necesario realizar un análisis de estática comparativa.

Diferenciando respecto a p la ecuación (5) se obtiene:

$$\frac{\partial Brecha}{\partial p} = \frac{5}{4} [F_2(\bar{\sigma}(p)) - F_1(\bar{\sigma}(p))] + \frac{5p-1}{4} \frac{\partial \bar{\sigma}}{\partial p} [f_2(\bar{\sigma}(p)) - f_1(\bar{\sigma}(p))] \quad (6)$$

Supongamos que p aumenta. Por la condición de dominancia estocástica de primer orden expuesta anteriormente, el primer término de la ecuación (6) es positivo y representa el incremento en la brecha proveniente de aquellas personas que aún deciden contestar y obtienen un puntaje

esperado mayor en el caso que contestan correctamente ya que aciertan con una probabilidad superior. El segundo término corresponde al efecto marginal en la brecha esperada provocada por el cambio en la decisión de contestar/omitir por parte de los individuos.

En cuanto al primer componente del segundo término de la ecuación (6) es plausible asumir que $p \geq 0,2$ dado que si un individuo contesta completamente al azar, cada alternativa tiene una probabilidad $\frac{1}{5}$ de estar correcta, por lo que podemos suponer que ninguna persona contesta de peor forma que la completa adivinación. Por otra parte, anteriormente se mencionó que a partir de la ecuación (2) es posible deducir que $\bar{\sigma}$ es una función de p y se demostró que esta función depende positivamente de la probabilidad de contestar correctamente. Dado lo anterior, el signo del segundo término y de la derivada parcial de la magnitud del impacto en la brecha de puntajes respecto a p dependen de la diferencia entre $f_2(\bar{\sigma}(p))$ y $f_1(\bar{\sigma}(p))$.

Dado que la dirección del efecto del cambio de política sobre la brecha de rendimiento ante cambios en la probabilidad de éxito no es monótonico, en la sección A.2 del Apéndice se muestran los resultados de dos simulaciones que utilizan la función de utilidad descrita en el ejemplo del Apéndice y dos tipos de distribuciones para el parámetro de aversión al riesgo. Las Figuras A.2 y A.3 del Apéndice muestran las funciones de probabilidad acumulada de ambos grupos de la población (F_1 y F_2), para dos familias de distribuciones para el parámetro de aversión al riesgo que cumplen con la condición de dominancia estocástica de primer orden. La Figura A.2 utiliza distribuciones uniformes mientras que la Figura A.3 considera distribuciones exponenciales. Adicionalmente, las Figuras A.4 y A.5 de la sección del Apéndice presentan simulaciones para el efecto marginal de la magnitud del impacto de la política ante cambios en la probabilidad de éxito en la contestación de una pregunta. Ambas figuras desagregan el efecto marginal total en el “margen intensivo”, el cual corresponde al primer término de la ecuación (6) y el “margen extensivo” correspondiente al segundo término de dicha ecuación.

Los dos ejercicios de simulación respecto a la estática comparativa del impacto de política ante cambios en p presentan un patrón similar. El signo de la derivada parcial de la magnitud del impacto de la eliminación de la penalización respecto a la probabilidad de éxito en la contestación es positivo para la mayor parte del rango de valores de p . En otras palabras, a medida que la probabilidad de éxito aumenta, el cambio de política tiene un efecto mayor en la disminución de las brechas de rendimiento entre los dos grupos de la población. Sin embargo, para valores muy grandes de p (aproximadamente mayores a 0.9) el margen extensivo negativo domina en magnitud al margen intensivo y por lo tanto, la derivada parcial es menor a cero. De esta forma, no siempre una mayor probabilidad de éxito implica que la política tendrá un mayor efecto, más bien, la magnitud del impacto es menor para individuos que contestan con seguridad la pregunta en cuestión.

Este análisis entrega conclusiones valiosas en términos de un potencial mecanismo detrás del impacto del cambio de política en la PSU y los potenciales efectos heterogéneos de la eliminación de la penalización por respuestas incorrectas. En primer lugar, esta sección del modelo establece a las diferencias en el grado de aversión al riesgo entre distintos grupos de población como un mecanismo plausible mediante el cual la regla de penalización actúa generando una brecha de puntajes entre ambas subpoblaciones. En segundo lugar, la estática comparativa respecto a la probabilidad de contestar correctamente predice que, para pruebas que por su naturaleza y diseño poseen diferentes niveles de dificultad, el cambio de política tendrá un impacto de distinta magnitud para cada test, por lo que es interesante evaluar efectos heterogéneos a nivel de tipo de prueba. Finalmente, el mismo ejercicio de estática comparativa sugiere que a mayor nivel de conocimiento y/o habilidad de los individuos (y por ende, mayor probabilidad de éxito en la contestación de una pregunta) el impacto de la eliminación de la penalización puede variar dependiendo de los dos efectos marginales descritos en la ecuación (6), por lo que es importante indagar sobre posibles efectos heterogéneos por nivel de conocimiento o habilidad de los inscritos.

4. Datos

Para lograr el desarrollo del ejercicio descrito anteriormente, es necesario obtener datos precisos sobre el rendimiento de alumnos pertenecientes a colegios de distinta dependencia administrativa antes y después del cambio de política. El Departamento de Evaluación, Medición y Registro Educativo (DEMRE) posee bases de datos de uso restringido para académicos e investigadores para los procesos de selección universitario desde el año 2004 hasta el 2019. Estas bases contienen los puntajes de cada prueba por estudiante, el número de respuestas según su status de contestación (correctas, incorrectas y omitidas) en cada test, variables sociodemográficas autoreportadas que fueron recolectadas durante el proceso de inscripción de la PSU e información relacionada al proceso de postulación y matrícula a programas de estudios del postulante.

Los datos poseen variables de los postulantes como el colegio de egreso, la dependencia del colegio⁷, su género, edad, comuna, status laboral, status domiciliario en caso de comenzar sus estudios y su cobertura de salud. Adicionalmente, existen variables sobre características y antecedentes socioeconómicos del grupo familiar como quien es el jefe de familia, el número de personas que pertenecen al grupo familiar, cuantos trabajan, cuantos estudian, el rango de ingreso familiar bruto al que pertenece el grupo familiar, la educación de los padres, la situación laboral de los padres y el domicilio donde reside el grupo familiar.

⁷La dependencia del colegio hace referencia a su fuente de financiamiento y la administración de los establecimientos educacionales. En Chile, existen colegios municipales, particulares subvencionados, particulares pagados, dependiente de corporaciones municipales y de corporaciones de administración delegada.

Cabe destacar que la muestra considerada para los análisis de esta investigación corresponde a los inscritos que rinden cada uno de los cuatro instrumentos de evaluación de la PSU entre los años 2010 y 2017. La elección de este período de tiempo radica en que antes del año 2010 no se cuenta con identificadores para cada uno de los establecimientos educacionales y desde el año 2018 hubo un cambio de codificación para gran parte de las variables sociodemográficas, las cuales son utilizadas como controles individuales en las especificaciones empíricas. Adicionalmente, se excluyen de la muestra todas las observaciones que no tengan información para alguna de las variables usadas en las estimaciones econométricas, así como también las observaciones que presentan valores atípicos alguna determinada variable⁸.

Dado que el análisis principal de esta investigación es el impacto del cambio de política en la regla de penalización por respuestas incorrectas en las brechas de rendimiento por colegio de egreso, resulta necesario definir los grupos de comparación respecto a los establecimientos educacionales. Para ello, se elige segmentar los colegios en establecimientos privados y no privados (donde este último incluye todos los colegios de dependencia distinta a la escuelas particulares pagadas) debido a que la calidad de la educación en Chile presenta una diferencia importante entre estos dos grupos de colegios. Adicionalmente, los alumnos egresados representan dos grupos de la población que se diferencian por su grado de aversión al riesgo.

En efecto, la literatura sobre el vínculo entre la aversión al riesgo y las características individuales de las personas presenta evidencia sobre la correlación negativa entre mejores características socioeconómicas y el grado de aversión. En el contexto de Chile, es posible utilizar los datos de la Encuesta de Protección Social (EPS) desarrollada por la Subsecretaría de Previsión Social para determinar si existe una relación negativa entre ingresos y el grado de aversión al riesgo. La EPS posee un conjunto de preguntas sobre áreas como educación, salud, seguridad social, capacitación laboral, patrimonio y activos, historia familiar e información sobre el hogar. A su vez, la encuesta del año 2012 posee tres preguntas sobre el grado de aversión al riesgo de los individuos, las cuales permiten clasificar a los individuos cuatro categorías de aversión (baja, media baja, media alta, alta)⁹. Mediante un modelo de regresión simple entre las categorías de aversión al riesgo y los tramos de salario mensual de los individuos proporcionados por la EPS del año 2012, es posible computar una relación negativa y estadísticamente distinta de cero entre los salarios y el grado de aversión. Estos resultados confirman los hallazgos de la literatura en el caso particular de Chile e indican que las personas menos aventajadas en términos socioeconómicos en promedio son más aversas al riesgo. Por otro lado, los datos proporcionado por el DEMRE revelan que pertenecer a un tramo de ingreso bruto familiar más alto está asociado a una reducción de la probabilidad de asistir a un colegio

⁸Estas exclusiones corresponde a individuos que rinden la PSU con una edad inferior a los 17 años y superior a los 60 años, además de los datos para el promedio de notas de enseñanza media con valores menores a 4 y mayores a 7.

⁹Tapia Griñen (2012).

no privado en un 5,8%¹⁰. De esta forma, es plausible pensar que aquellas personas que asisten a instituciones educacionales no privadas poseen menores ventajas socioeconómicas que los individuos que egresan de colegios particulares pagados y por esta razón, los alumnos de grupo de escuelas son más aversos al riesgo que los estudiantes de los establecimientos privados.

En esa misma línea, es indispensable definir una medida adecuada para medir el rendimiento de los alumnos en la prueba de selección universitaria. Una primera opción es utilizar los puntajes PSU determinados por el DEMRE, los cuales se construyen a través de un proceso de estandarización que considera el número de preguntas correctas (corregido por las respuestas erróneas en el caso que exista una regla de penalización). En su estudio, Coffman and Klinowski (2020) mencionan que esta medida no es del todo transparente y clara, ya que no es posible conocer por completo todas las configuraciones involucradas en el proceso de estandarización. En contraste, los autores proponen el “*raw score*” (número de correctas, corregidas por incorrectas si es que existe penalización) como la métrica más transparente sobre el rendimiento de los alumnos. Adicionalmente, Coffman and Klinowski mencionan que para cuantificar de manera precisa las ganancias en término de rendimiento por parte de los alumnos, se necesita una medida estandarizada que tome en cuenta las diferencias anuales en la varianza del rendimiento.

En esta investigación se considera la medida de rendimiento propuesta en Coffman and Klinowski (2020). A fin de construir esta métrica, es importante notar que dependiendo de la asignatura del test y del año, el número de preguntas varía entre 70 y 80 preguntas. Por esta razón, primero es necesario homologar el número de preguntas que posee cada test para que estas sean comparables. Para ello, se calcula el porcentaje de respuestas correctas, omitidas y erróneas para cada individuo que rinde una determinada prueba en un determinado año:

$$x'_{i,a,t} = \frac{x_{i,a,t}}{c_{i,a,t} + o_{i,a,t} + e_{i,a,t}}$$

Donde c , o y e corresponden al número de respuestas correctas, omitidas y erróneas respectivamente. A su vez, $x = \{c, o, e\}$ y x' es el porcentaje para cada uno de los status de contestación en relación al total de preguntas. Además, se denota por i a los individuos, por a a las asignaturas y por t a los años. Para obtener el total de preguntas para cada status de contestación, se reescala el total de preguntas multiplicando los porcentajes anteriores por 80, por lo que estos valores están expresados en términos del mismo nivel para todas las asignaturas y años (80 preguntas). Posteriormente, se construyen lo que Coffman and Klinowski denominan “*raw score*”:

¹⁰Las bases de datos para investigadores proporcionadas por el DEMRE incluye una variable de ingreso familiar bruto que indica la posición del grupo familiar en relación a doce tramos para el ingreso bruto. En base a lo anterior, se estima un modelo de probabilidad lineal donde la variable dependiente es el indicador $NoPrivado_{ic}$

$$R_{i,a,t} = C_{i,a,t} - \mathbb{1}_{\{t < 2015\}} * 0,25 * E_{i,a,t}$$

Donde $R_{i,a,t}$ es el “*raw score*” obtenido por el individuo i en la prueba de asignatura a en el año t , C y E corresponden al total de preguntas correctas y erróneas considerando la reescalación descrita anteriormente. Por su parte, $\mathbb{1}_{\{t < 2015\}}$ es un indicador que toma el valor 1 para el período previo al cambio de política. En otras palabras, esta medida de rendimiento considera directamente el número de respuestas correctas ajustadas por el número de respuestas incorrectas en el caso en que la regla de penalización esté vigente.

Es fácil notar que los “*raw scores*” aumentarán una vez que se elimine la norma de penalización por respuestas erróneas, a pesar que no exista un aumento en el rendimiento de los alumnos. Por esta razón, los autores argumentan que es necesario estandarizar los “*raw scores*” para cada asignatura, restando su media anual y dividiendo por la desviación estándar conjunta anual entre los dos tipos de colegios, con el fin de obtener una medida más precisa del rendimiento. En esa línea, se define la siguiente métrica:

$$Z_{i,a,t} = \frac{R_{i,a,t} - \bar{R}_{a,t}}{\frac{N_{p,a,t}}{N_{a,t}} \sigma_{p,a,t} + \frac{N_{np,a,t}}{N_{a,t}} \sigma_{np,a,t}}$$

Donde $Z_{i,a,t}$ es la estandarización del “*raw score*” (el cual en adelante llamaremos “Z-puntaje”), $\bar{R}_{a,t}$ es el “*raw score*” promedio del año t en el test de la asignatura a . A su vez, $N_{p,a,t}$ y $N_{np,a,t}$ corresponden al número de personas que pertenecen a los colegios privados y no privados (respectivamente) y $N_{a,t}$ es el número total de personas que rinden la prueba de la asignatura a en el año t . Finalmente, $\sigma_{p,a,t}$ y $\sigma_{np,a,t}$ es la desviación estándar de los “*raw score*” para ambos grupos.

5. Estrategia Empírica

Usando los datos descritos en la sección anterior, se seguirá la estrategia empírica presente en el trabajo de Coffman and Klinowski (2020) y su estudio del impacto de la penalización por respuestas incorrectas en la brecha de género de los puntajes obtenidos en la PSU. En este trabajo los autores compararan a las personas que rinden la PSU antes y después del cambio en la política, realizando múltiples chequeos de robustez para lidiar con el problema de confundir el efecto causal con tendencias generales en el tiempo. A lo largo de esta sección se describirá la adaptación de este enfoque econométrico y sus respectivas pruebas de robustez a esta investigación.

Siguiendo la estrategia de Coffman and Klinowski (2020) utilizaré una comparación “*before-after*” entre dos grupos de la población de personas que rinden la PSU. Este enfoque de diferencias en diferencias busca comparar el comportamiento de distintas variables de interés, antes y después de la eliminación de la penalización por respuestas incorrectas entre colegios privados y no privados. La justificación de esta estrategia empírica radica en que el cambio de política nos entrega un escenario cuasi experimental que genera variación entre los dos tipos de colegios y bajo ciertas condiciones (detalladas más adelante) es posible identificar su impacto causal a través de la comparación mencionada anteriormente.

Para examinar el impacto de la eliminación de la regla de penalización por respuestas incorrectas en la PSU sobre las brechas de preguntas omitidas y de rendimientos según la dependencia del establecimiento secundario de origen del postulante, se estimará la siguiente especificación:

$$y_{icta} = \beta Post_t * NoPrivado_{ic} + \delta_t + \mu_c + \mu_c * t + \gamma Z_i + \epsilon_{icta} \quad (7)$$

donde y_{icta} corresponde a la variable de resultado de interés (número de preguntas omitidas o rendimiento) del individuo i que pertenece al colegio c , en el año t para la prueba de la asignatura a . El indicador $Post_t$ toma el valor 1 para todos los años posteriores al 2014 y cero en caso contrario. La variable dummy $NoPrivado_{ic}$ toma el valor 1 si el colegio no pertenece a los colegios de dependencia particular pagada y cero en caso contrario¹¹. Por otro lado, δ_t y μ_c corresponde a efectos fijos a nivel año y colegio respectivamente. Además, el término $\mu_c * t$ indica que la especificación incluye tendencias lineales en el tiempo específicas de cada colegio. Estas tendencias lineales se incorporan en el modelo econométrico con el fin de evitar confundir el efecto causal del cambio de política con diferencias generales en el tiempo entre escuelas. Por su parte, el vector Z_i incluye una serie de variables referentes a características socio-demográficas individuales que se utilizarán como controles en la especificación. Finalmente, se asume que los términos del error puede correlacionar a nivel colegio, por lo que en las estimaciones se considerarán errores estándar clusterizados según el establecimiento educacional al que pertenece el estudiante.

El análisis principal de esta investigación radica en identificar el impacto de la penalización por respuestas incorrectas en las brechas en el número de preguntas omitidas y en el rendimiento obtenido en la PSU entre los colegios privados y no privados, por lo que el coeficiente de interés es β . Para justificar la validez interna de esta estrategia empírica es necesario presentar evidencia a favor del cumplimiento de tendencias paralelas, esto es, que las tendencias de las variables de resultado de interés para grupos de personas de establecimientos educacionales privado y no privados antes de la

¹¹En la base de datos otorgada por el DEMRE, los colegios son clasificados según su dependencia administrativa en alguno de los siguientes grupos de establecimientos educacionales: particulares pagados, particulares subvencionados y municipales.

política han evolucionado conjuntamente de manera paralela antes de la política. En otras palabras, se debe demostrar que ante la ausencia de la eliminación de la regla de penalización, las diferencias en las variables de interés entre los dos tipos de colegios es constante a través del tiempo.

Por otro lado, para evitar confundir el efecto causal de la penalización por respuestas incorrectas con tendencias generales en el tiempo, es necesario realizar chequeos de robustez a favor de la validez interna de los resultados de las estimaciones. En principio, al igual que en el estudio Coffman and Klinowski, se acotarán los datos a un intervalo de años cercanos a la política del cambio de reglas en la PSU, en particular, para dos años antes del cambio y dos años después del 2015, a fin de evitar que los cambios en tendencias temporales entre los grupos a comparar afecten las estimaciones. En segundo lugar, se analizará la muestra de alumnos que repiten la prueba para analizar el diferencial de rendimiento obtenido en ambos procesos de selección durante el periodo anterior y durante el cambio de política, con el fin de aportar evidencia adicional al cumplimiento del supuesto de tendencias paralelas. Finalmente, se realizará un test de tratamientos placebos sobre cuatro variables relacionadas con la situación laboral de los integrantes del grupo familiar y dos variables asociadas al rendimiento escolar, para comprobar que la estimación del impacto causal de la política no es conducida simplemente por un cambio sistemático en la composición de los colegios en el año 2015.

Dentro de la literatura sobre aversión al riesgo y su relación con características demográficas individuales hay múltiples estudios que reportan una relación clara entre la aversión al riesgo y el género: las mujeres tienden a ser más aversas al riesgo que los hombres (Jianakoplos and Bernasek, 1998; Halek and Eisenhauer, 2001; Eckel and Grossman, 2008). Dado lo anterior, el parámetro de interés de la especificación (7) se estimará manera separada para hombres y mujeres con el fin de estudiar si existen efectos heterogéneos entre individuos que son más aversos al riesgo que otros dentro de cada grupo de colegios.

Finalmente, la estática comparativa desarrollada en la sección 3 sugiere que podrían existir diferencias en el impacto de la política a través de los cuatro tipos de prueba que conforman la PSU, ya que cada una de ellas poseen características propias en términos de diseño y dificultad. Adicionalmente, en esa misma sección se concluyó que el aumento o disminución en el nivel de conocimiento (traducido en una mayor probabilidad de contestar correctamente) produce variaciones en el efecto de la política. Dado lo anterior, se estudian efectos heterogéneos por tipos de prueba y según el nivel de conocimiento o habilidad de los individuos. Para esta última heterogeneidad en el efecto de la eliminación de la penalización se utilizará como *proxy* de conocimiento la posición de un alumno en la distribución del puntaje NEM en un determinado año, considerando que personas con un menor promedio de notas en la enseñanza es un indicador de un menor nivel de conocimiento.

6. Resultados en respuestas omitidas y en Z-puntajes

6.1. Resultados principales

En esta sección se emplea la estrategia empírica descrita anteriormente utilizando como variable dependiente el número de respuestas omitidas y el rendimiento (medido como el Z-puntaje) obtenido por los estudiantes en cada instrumento de la PSU. Dado el modelo de la sección 3, se espera que los alumnos más aversos al riesgo (es decir, aquellos que provienen de colegios no privados) omitan menos y en consecuencia, obtengan un mayor puntaje esperado después de la implementación del cambio de política.

El panel A de la Tabla 1 muestra el número de preguntas omitidas promedio por tipo de colegio para cada uno de los cuatro instrumentos de evaluación que componen la PSU durante el período de estudio. Antes de la política, se observa que a lo largo de todas las asignaturas los alumnos egresados de escuelas no privadas omiten en promedio un mayor número de preguntas que los estudiantes de establecimientos privados. A modo de ejemplo, antes del año 2015, los alumnos de colegios no privados omitían en promedio 9.73 preguntas más que los egresados de escuelas privadas en la prueba de Lenguaje Comunicación. Asimismo, la Tabla 1 muestra que esta brecha es mayor para el resto de las pruebas, alcanzando 12.44 , 14.11 y 16.84 preguntas para los tests de Historia, Ciencias y Matemáticas respectivamente. No obstante, después del cambio de política en el año 2015, el número de preguntas omitidas en ambos grupos se reduce abruptamente y las diferencias entre tipos de colegios alcanzan valores cercanos a cero.

El panel A de la Tabla 2 formalmente mide el impacto de la eliminación en la brecha del número de preguntas omitidas por tipo de colegio a través de la estimación de la especificación (7) para la muestra comprendida entre los años 2010 y 2017. Los resultados de este panel revelan que el cambio de política reduce de manera significativa la brecha de preguntas omitidas entre colegios privados y no privados. En promedio, esta brecha se reduce en 9.51 preguntas en el caso de la prueba de Lenguaje y 16.37 preguntas en el test de Matemáticas. Por otro lado, para las evaluaciones optativas se registra una reducción de 13.96 y 15.16 preguntas para las pruebas de Historia y Ciencias respectivamente. Estas estimaciones sugieren que la modificación en la Prueba de Selección Universitaria reduce completamente las diferencias en la tasa de omisión entre dos grupos que poseen distintos grados de aversión al riesgo, lo cual es un resultado directo del modelo teórico expuesto anteriormente. A su vez, el panel A muestra que estos resultados son robustos a la exclusión de variables de control a nivel individual, en términos de la magnitud y significancia estadística.

Siguiendo la estrategia empírica de Coffman and Klinowski (2020) y con el fin de evitar que la estimación del impacto causal del cambio de política en la PSU sea confundido con tendencias

generales en el tiempo, la muestra se acota a dos años previos a la eliminación de la regla de penalización y dos años posteriores a esta modificación. El panel B de la Tabla 2 presenta las estimaciones de la especificación (7) para la muestra restringida al período que abarca desde el año 2013 hasta el 2016. Los resultados de este panel muestran que la estimación del impacto de la eliminación de la penalización es levemente menor cuando se acota el período de estudio. No obstante, las estimaciones puntuales de la reducción de las brechas del número de preguntas omitidas para cada prueba sigue representando un porcentaje importante de las brechas promedios previas al cambio de política. Asimismo, los coeficientes asociados al impacto de la modificación de la PSU continúan siendo estadísticamente distintos de cero. Por consiguiente, el panel B de la Tabla 2 posee evidencia a favor de la idea que es poco probable que las estimaciones del panel A de la misma tabla confunda de manera significativa el impacto de la política con tendencias generales en el tiempo.

A pesar que la modificación de la regla de penalización afecta directamente a las diferencias en la omisión de preguntas entre colegios privados y no privados, no es tautológico que el cierre de estas brechas se verá reflejado en una mejora relativa del rendimiento de los alumnos de escuelas no privadas. En esta línea, resulta relevante explorar el impacto de la reforma en la PSU sobre las diferencias de rendimiento entre establecimientos educacionales privados y no privados.

Como se mencionó anteriormente, en esta investigación se utiliza la medida de rendimiento propuesta en el estudio de Coffman and Klinowski (2020), ya que esta métrica es más clara y toma en consideración cualquier cambio en la varianza que es inducida por el cambio de política. Para visualizar el comportamiento de esta medida de rendimiento a través del tiempo, el panel B de la Tabla 1 muestra el promedio del Z-puntaje por tipo de colegio para cada uno de los cuatro instrumentos de evaluación que componen la PSU durante el período de estudio. Este panel indica que tras la implementación de la política, los colegios privados tienen en promedio un menor rendimiento (en comparación al período previo a la reforma) en todas las pruebas y las escuelas no privadas exhiben un aumento promedio de los Z-puntajes, lo cual se traduce en una caída de la brecha promedio de rendimiento entre ambos tipos de establecimientos educacionales.

La Tabla 3 presenta las estimaciones del impacto de la eliminación de la penalización por respuestas incorrectas en las brechas de rendimiento entre colegios privados y no privados. Esta tabla utiliza la misma estructura que la Tabla 2: en el panel A se muestran los resultados para todo el período de estudio, mientras que en el panel B se utiliza una muestra restringida a los años 2013-2016. En este punto, cabe destacar que los valores de los Z-puntajes son interpretados como fracciones de una desviación estándar (en adelante SD, por su sigla en inglés). Las estimaciones sugieren que el cambio de política reduce de manera general la brecha en el rendimiento entre colegios privados y no privados en 0.12 SD en promedio. Esta reducción es estadísticamente significativa y corresponde a un nueve por ciento de la brecha promedio del rendimiento considerando el período

previo a la modificación de la PSU. Además, en el panel A de la Tabla 3 se estima que la eliminación de la penalización por respuestas erróneas disminuye significativamente las brechas de rendimiento en 0.16 SD en la prueba Lenguaje, 0.04 SD en Matemáticas, 0.16 SD en Historia (Ciencias Sociales) y 0.12 SD en Ciencias Naturales. Por otro lado, al igual que en el caso de las preguntas omitidas, se descarta la hipótesis que estos resultados son conducidos por la inclusión de controles a nivel individual.

Al examinar la sensibilidad de los resultados ante la restricción de la muestra a dos años previos y posteriores al cambio de política se observa que la reducción de la brecha estimada para la prueba de Ciencias Naturales se cambia en magnitud y significancia cuando se considera un rango acotado de años cercanos a modificación de la PSU. Esto sugiere que el impacto fue más pequeño para este test y que la estimación del panel B de la Tabla 3 puede ser interpretado como un límite inferior del efecto de la política.

Las estimaciones presentadas anteriormente son congruentes con los resultados del modelo teórico. La regla de penalización de respuestas incorrectas actúa mediante el mecanismo de aversión al riesgo, el cual genera diferencias en la tasa de omisión de preguntas entre dos grupos de población. Una vez aplicada la reforma de la PSU desaparecen estas brechas en el número de preguntas omitidas, lo cual se traduce en una disminución en la brecha de rendimiento entre los alumnos. En otras palabras, esta regla perjudicaba a los alumnos de colegios no privados debido a su mayor grado de aversión al riesgo y luego de su remoción los estudiantes de este tipo de escuelas comenzaron a contestar un mayor número de preguntas aumentando su rendimiento promedio.

Adicional a lo anterior, se ve que el impacto sobre el Z-puntaje parece ser mayor en los dos tópicos de índole humanista. En el estudio del impacto de la penalización de respuestas incorrectas de la PSU en las brechas de género de Coffman and Klinowski (2020), la prueba de Matemáticas también presenta un comportamiento similar tras el cambio de política. Los autores argumentan que el test de Matemáticas es menos susceptible a una “adivinación educacada” en comparación al resto de las evaluaciones, debido a que en estas pruebas podría ser más fácil descartar alternativas “señuelo” mientras que en Matemáticas si una persona no conoce el concepto detrás de una determinada pregunta, descartar alternativas puede ser muy difícil. Este mismo argumento puede ser utilizado en el caso de la prueba de Ciencias Naturales, ya que ambos tests pertenecen a los tópicos de ciencias exactas y comparten características similares en la naturaleza de sus preguntas. Esto podría explicar el efecto relativamente menor para los tests de Matemáticas y Ciencias Naturales documentado en las estimaciones utilizando como medida de rendimiento el Z-puntaje.

6.2. Validez de la Estrategia Empírica

Para interpretar las estimaciones presentadas en las tablas 2 y 3 como el impacto causal del cambio de política observado en la PSU y no como una simple correlación, es importante discutir la validez interna de los resultados a través de la verificación del cumplimiento del supuesto de identificación asociado a la estrategia empírica. En este estudio, el supuesto de identificación subyacente a la estrategia econométrica radica en que las tendencias de las variables de resultado de interés (número de preguntas omitidas y rendimiento) para los grupos de personas pertenecientes a colegios privados y no privados habrían evolucionado conjuntamente de manera paralela si es que el cambio de la política no se hubiera implementado. Este es el supuesto clásico de la estrategia de diferencias en diferencias aplicado al contexto de esta investigación.

Una forma de testear el cumplimiento del supuesto de tendencias paralelas es hacer una estimación empírica al estilo de un estudio de eventos que permita comparar las brechas de preguntas omitidas y rendimiento entre los dos grupos de interés para un año de referencia con las brechas observadas para los años anteriores y posteriores al cambio de política. Si el año de referencia es el 2014, las brechas estimadas para los años del período comprendido desde el 2010 hasta el 2013 no deberían ser estadísticamente diferentes a las del año de referencia, mientras que las brechas desde el 2015 deberían ser estadísticamente distintas a la del año 2014 mostrando una reducción en el *gap* entre colegios, ya que a partir de ese momento comienza a regir la eliminación de la penalización por respuestas incorrectas. En concreto, se estima la siguiente especificación:

$$y_{icta} = \alpha_1 NoPrivado_c + \sum_{t \neq 2014} \delta_t * \mathbb{1}_{\{año=t\}} + \sum_{t \neq 2014} \mu_t \mathbb{1}_{\{año=t\}} * NoPrivado_c + \phi_c + \phi_c * t + \epsilon_{icta} \quad (8)$$

donde y_{icta} es la variable de resultado de interés obtenida por el individuo i del colegio c en el año t en la prueba de la asignatura a . La variable dummy $NoPrivado_{ic}$ toma el valor 1 si el colegio no pertenece a los colegios de dependencia particular pagada y cero en caso contrario. Por otro lado, la especificación incluye efectos fijos para todos los años excepto el año 2014 y la interacción entre el indicador $NoPrivado_{ic}$ con cada uno de estos efectos fijos, los cuales se denotan con la indicatriz $\mathbb{1}_{\{año=t\}}$. Además, se agregan efectos fijos a nivel escuela y tendencias lineales específicas de cada colegio. Finalmente, siguiendo la misma justificación de la especificación (7), los errores estándar son clusterizados a nivel colegio.

La Figura 1 muestra los resultados por tipo de prueba para la estimación de los parámetros μ_t de la especificación (8) para todos los años entre el 2010 y 2017 tomando como referencia el año 2014. La Figura 1 indica que las brechas promedios respecto a las preguntas omitidas entre colegios no privados y privados experimentaron un patrón similar para todos los instrumentos de

evaluación. Específicamente, las brechas presentan un incremento sostenido a partir del 2010 hasta el año 2014, lo cual puede ser explicado por aumentos en las preguntas omitidas por parte de los alumnos de escuelas no privados y/o disminuciones en la omisión de preguntas de los estudiantes de establecimientos privados. No obstante, en el año 2015 comienza a regir la eliminación de la penalización por respuestas erróneas y con ello se observa que la brecha de preguntas omitidas cae drásticamente. Estos resultados sugieren que los coeficientes de la Tabla 2 corresponden a límites inferiores del impacto de la política, dado que sin ella, las brechas podrían haber seguido ampliándose¹².

Por otro lado, la Figura 2 presenta las estimaciones de la especificación (8) con el fin de testear si se cumple el supuesto de tendencias paralelas cuando la variable de interés es el rendimiento de los inscritos en la PSU. Los resultados de la Figura 2 documentan evidencia a favor del supuesto de identificación de la estrategia empírica para cada una de las pruebas. En efecto, para el caso de las pruebas optativas (Historia y Ciencias) se observa que las brechas promedio de rendimiento para los años previos a la política no son estadísticamente distintas entre sí y desde el año 2015 hay una clara reducción en el *gap* de rendimiento entre los tipos de colegios. Un patrón similar se observa en las estimaciones para la prueba de Lenguaje y Comunicación, salvo que la brecha promedio del año 2012 es estadísticamente mayor que el resto de los años previos a al cambio de política. Por otro lado, el subgráfico correspondiente a la prueba de Matemáticas sugiere que las brechas promedios del período 2010-2013 son estadísticamente más grandes que la del año de referencia (2014), sin embargo, estas son estables a través del tiempo y hay una evidente reversión de la tendencia a partir del año 2015. En definitiva, para cada uno de los cuatro instrumentos de la PSU resulta verosímil asumir que la evolución de la diferencia en el rendimiento entre los dos tipos de establecimientos educacionales sería constante si no se hubiera implementado la reforma de la PSU. En consecuencia, las estimaciones de la Tabla 3 pueden ser interpretadas como el impacto causal de la eliminación de la regla de penalización por respuestas incorrectas en la PSU.

A fin de apoyar la interpretación causal de los resultados presentados en la sección anterior, a continuación se analizará la robustez de los resultados sobre el impacto de la eliminación de la penalización por respuestas incorrectas en la Prueba de Selección Universitaria a través de dos testeos empíricos. En primer lugar, se examinará la muestra de alumnos que repiten la PSU en dos años consecutivos antes y durante la implementación del cambio en la regla, con el fin de evaluar si las brechas entre colegios sobre la diferencia de rendimientos obtenidos al momento de volver a rendir las diferentes pruebas en promedio es distinta a través de los años. Posteriormente, se realizará un estudio de tratamientos placebos en donde se analizará el impacto de la penalización por respuestas incorrectas a través de distintas variables laborales y sobre el rendimiento escolar

¹²Cabe recalcar que la Figura 1 no incluye variables de control a nivel individual. Sin embargo, anteriormente se demostró que la inclusión de controles individuales no afectan las estimaciones del impacto de la política. Asimismo, la inclusión de estas variables no cambia la conclusión positiva respecto al cumplimiento del supuesto de tendencias paralelas (resultado no reportado).

que, en teoría, no deberían ser afectadas por la política.

El mejor contrafactual para evaluar la política en cuestión sería comparar a alguien que rinde la prueba antes y después de la política en las mismas condiciones de conocimientos. Si bien esta comparación es imposible debido a la inexistencia de este contrafactual perfecto, los datos proporcionados por el DEMRE permiten identificar a las personas que repiten el proceso de selección en dos años consecutivos y de esta forma construir una muestra que sólo incluye a las personas que rinden la PSU dos veces. Las razones para que un individuo decida repetir un proceso de selección son variadas: los bajos puntajes obtenidos anteriormente, el programa de estudios al que desea ingresar podría exigir un puntaje ponderado mayor al obtenido la primera vez, el incumplimiento del requisito de puntajes para la obtención de becas y/o créditos del estado chileno, entre otras. En este sentido, es muy probable que los alumnos que realizan la prueba nuevamente cambien su nivel de conocimiento y preparación entre los distintos procesos de selección. Sin embargo, es posible diseñar un análisis empírico que proporcione evidencia a favor del cumplimiento del supuesto subyacente a la estrategia de identificación.

Una vez que un alumno vuelve a rendir la PSU obtiene un diferencial en el número de preguntas omitidas y en el rendimiento respecto a la prueba anterior. Esta diferencia podría ser causada por múltiples motivos, sin embargo, si el supuesto de tendencias paralelas se cumple, la brecha según el colegio de egreso referente a este cambio en las variables de resultado entre dos procesos de selección no debería cambiar de manera significativa durante los años previos al cambio de política. Para realizar este testeo empírico, se construye una muestra para el período 2011-2015 de las personas que repiten la PSU entre dos años consecutivos y se estima la especificación (8) utilizando como variable dependiente el diferencial de preguntas omitidas y de rendimiento entre ambos procesos de selección¹³. La Tabla 4 presenta los resultados de dicha estimación para las pruebas obligatorias y optativas.

El panel A de la Tabla 4 muestra los resultados utilizando como variable dependiente el diferencial del número de preguntas omitidas obtenido por un alumno entre dos procesos de selección consecutivos. En este panel se observa que para las pruebas de Matemáticas, Historia y Ciencias Naturales, en el período previo a la eliminación de la penalización las brechas promedio del diferencial de preguntas omitidas entre colegios privados y no privados no son estadísticamente distintas a través del tiempo. Por el contrario, para el año 2015 se computa una reducción en la brecha de diferenciales estadísticamente significativa alcanzando un valor de 13.79, 9.44 y 12.87 preguntas para Matemáticas, Historia y Ciencias respectivamente. En contraste, en la prueba de Lenguaje y Comunicación se observa que para los años 2011-2013 las brechas promedios eran

¹³En esta muestra la observación del año 2011 hace referencia a una persona que rinde la PSU en el proceso de selección 2010 y 2011, es decir, se consideran los datos de la segunda vez que realizan la prueba.

significativamente más grandes que la del año de referencia. No obstante, tras la implementación de la política se observa una clara reducción de la brecha entre colegios privados y no privados, la cual es cuatro veces más grande (8.04 preguntas) que la disminución observada del año 2014 respecto al año 2013 (2.04 preguntas). Estos resultados se consideran evidencia favorable en términos del cumplimiento del supuesto de identificación y la interpretación causal de las estimaciones del impacto del cambio de política, ya que es plausible pensar que la evolución de los diferenciales de preguntas omitidas entre dos procesos de selección permanecería constante si es que el cambio de política nunca hubiese ocurrido.

Por su parte, el panel B de la Tabla 4 utiliza el diferencial de rendimiento obtenido por un alumno que rinde la PSU en dos años consecutivos como variable dependiente. Las estimaciones de este panel indican que, en el período previo a la reforma, las brechas del diferencial de Z-puntajes para las pruebas de índole humanista (Lenguaje e Historia) no son estadísticamente distintas a través del tiempo, es decir, el mayor o menor puntaje obtenido por los alumnos en la segunda vez que rinden la PSU no varía de manera significativa entre los procesos de selección y entre los tipos de colegios. Sin embargo, para el año 2015 se registra una brecha de rendimientos entre colegios significativamente menor (0.136 SD y 0.170 SD respectivamente), a favor de los establecimientos educacionales no privados. En el caso de los tests relacionados con las ciencias exactas, se observan pequeñas diferencias en las brechas del diferencial de rendimientos para algunos años del período previo a la reforma respecto al año de referencia, pero estas discrepancias son significativas solamente al diez por ciento de significancia. De esta forma, las estimaciones del panel B de la Tabla 4 también se posicionan como evidencia positiva en términos del cumplimiento del supuesto de tendencias paralelas, ya que sugieren que los diferenciales de rendimiento por alumnos de colegios privados, en promedio, han evolucionado conjuntamente de manera paralela respecto a los obtenidos por estudiantes de establecimientos no privados antes de la eliminación de la penalización por respuestas incorrectas. En definitiva, los hallazgos de la Tabla 4 suman evidencia creíble de la interpretación causal de las estimaciones relacionadas con el impacto de la política sobre el número de preguntas omitidas y el rendimiento de los alumnos, a lo largo de todas las pruebas que componen la PSU.

Si bien la política trae consigo una disminución en las brechas del número de preguntas omitidas y las brechas de rendimiento entre colegios privados y no privados, la reforma en la PSU no debería impactar otras variables que no son afectadas por la eliminación de la penalización. En este sentido, es posible realizar un análisis de tratamientos placebos, en donde se estime el impacto del cambio de política sobre diferentes variables relacionadas con la situación laboral del grupo familiar de los inscritos en los procesos de selección universitaria y en variables asociadas al rendimiento escolar durante los cuatro años de enseñanza media. En teoría, no hay un mecanismo claro mediante el cual la modificación de la regla de penalización impacte de manera significativa las decisiones laborales de los distintos integrantes del grupo familiar, por lo que el efecto de la política no debería ser

estadísticamente distinto de cero. Asimismo, el hecho que se modifique una regla específica en el diseño de los instrumentos de evaluación de la PSU, no debería impactar de manera directa el promedio de notas que obtienen los alumnos en sus respectivos colegios.

La sección (a) de la Tabla 5 presenta las estimaciones de la especificación (7) utilizando distintas variables dependientes relacionadas con el status laboral de los integrantes del hogar de cada individuo que rinde la PSU. La primera columna de esta tabla utiliza indicador que toma el valor 1 si el padre del alumno se encuentra cesante y cero en caso contrario, mientras que la segunda columna usa una variable dummy que toma el valor 1 si la madre del individuo está cesante. A su vez, la tercera columna considera un indicador que toma el valor 1 si quien rinde la PSU se encuentra trabajando y cero en cualquier otro caso. Por último, la cuarta columna de la sección (a) utiliza como variable dependiente el número de personas que trabajan en el grupo familiar del estudiante. Todas las regresiones que poseen una variable dummy como *outcome* de interés se estiman mediante el modelo de probabilidad lineal y en todas las especificaciones los errores estándar son clusterizados según el tramo de ingresos bruto familiar al cual pertenece el hogar¹⁴. Por otro lado, la sección (b) de la Tabla 5 muestra las estimaciones de la especificación (7) usando como variable dependiente el promedio de notas de enseñanza media y el puntaje NEM. En ambas estimaciones, se permite que los errores estándar estén correlacionados al nivel de establecimientos educacionales.

Los resultados de la sección (a) de la Tabla 5 indican que ninguna de las variables relacionadas con la situación laboral del grupo familiar es afectada de manera significativa por la eliminación de la penalización por respuestas erróneas. Si bien las brechas en la probabilidad que el padre de un alumno de colegios no privado se encuentre cesante (en comparación al de un alumno de escuela privada) aumenta después del cambio de política, este incremento no es estadísticamente distinto de cero a los niveles usuales de significancia. Lo mismo ocurre con el la disminución de un 0,1% en la brecha asociada a la probabilidad que la madre del inscrito esté cesante. De esta misma forma, los coeficientes relacionados con el cambio en la brecha de la probabilidad asociada a que el estudiante se encuentre trabajando de manera remunerada y el efecto sobre la brecha en el número de personas del hogar que poseen empleo tampoco son estadísticamente significativos y sus estimaciones puntuales son muy cercanas a cero.

Por su parte, los resultados de la sección (b) de la Tabla 5 sugieren que tras la incorporación de la reforma a las PSU la brecha en el rendimiento escolar entre los alumnos de colegios privados y no privados aumenta. Sin embargo, este incremento pareciera estar conducido principalmente por la inclusión de variables de control individual y sólo son estadísticamente significativas a un

¹⁴En estas especificaciones se permite que los errores estén correlacionados según el tramo de ingreso bruto familiar, en lugar de su colegio de origen, ya que es tentativo pensar que dentro de un mismo establecimiento educacional existen subgrupos socioeconómicos para los cuales las variables asociadas a la situación laboral del grupo familiar correlacionan al interior de cada subgrupo.

diez por ciento de significancia. Estos hallazgos proporcionan evidencia a favor de la robustez de las estimaciones del impacto de la modificación de la regla de penalización en la PSU sobre las preguntas omitidas y el rendimiento de los estudiantes, ya que se descarta que otras variables presentan cambios sistemáticos el año 2015, a tal punto que registran un efecto significativo tras el cambio de política que ocurre ese mismo año.

6.3. Heterogeneidad en los resultados

En el contexto chileno es posible hacer una clasificación más refinada respecto a los tipos de establecimientos educacionales de los cuales egresa un determinado alumno. Específicamente, los colegios no privados se pueden desagregar en colegios municipales y particulares subvencionados. Los primeros se caracterizan por estar administrados por la municipalidad de cada comuna y por lo general sus alumnos no deben pagar aranceles, mientras que los segundos poseen un régimen de copago, en el cual una parte de arancel es cubierto por el estado y el resto es financiado por el apoderado del estudiante.

Dado lo anterior, es posible realizar la estimación del impacto de la eliminación de la regla de penalización sobre las brechas de preguntas omitidas y rendimiento según el tipo de establecimiento educacional de egreso, de manera tal que diferencie el efecto de la política percibido por los colegios municipales y el de las escuelas particulares subvencionadas. La Tabla 6 presenta la estimaciones para la especificación (7) incluyendo una interacción adicional al modelo econométrico, que permite diferenciar el efecto para colegios municipales y particulares subvencionado. El panel A de esta tabla utiliza como variable dependiente el número de preguntas omitidas, mientras que el panel B considera el rendimiento obtenido por los alumnos. Estas estimaciones permiten examinar si dentro de los colegios no privados existe una subcategoría de establecimientos que es más afectada por el cambio de política que otra.

Los resultados del panel A de la Tabla 6 revelan que, posterior a la política, los colegios municipales experimentan una disminución de 14.3 preguntas en la brecha general de respuestas omitidas respecto a los establecimientos no privados. A su vez, se estima que la reducción de esta brecha tras la eliminación de la penalización, alcanza aproximadamente 13 preguntas en el caso de escuelas particulares subvencionadas. A pesar que esta diferencia de 1.3 preguntas en el impacto entre ambos establecimientos no privados es estadísticamente significativa, no se observan grandes diferencias en términos del impacto en las brechas de rendimiento. En efecto, el panel B de la Tabla 6 indica que, dentro de los colegios no privados, tanto las escuelas municipales como las particulares subvencionadas ven reducidas las brechas generales de rendimiento (en relación a los establecimientos privados) en 0.12 desviaciones estándar. En otras palabras, pareciera que el mayor efecto en el número de preguntas omitidas por parte de los colegios municipales no se traducen en un mayor efecto promedio en el rendimiento.

Por otro lado, la literatura referente a las reglas de penalización en pruebas de selección múltiple ha estudiado ampliamente el impacto en las brechas de género en puntajes y otras variables de resultado. En esa línea, es interesante examinar si existen resultados heterogéneos en la estimación del impacto de la penalización por respuestas incorrectas según el género de las personas que rinden la PSU. En vista de lo anterior, la Tabla 7 expone los resultados de las estimaciones de una especificación similar a la regresión (7) pero que desagrega el efecto del cambio de política según el género de los inscritos. De manera general, las brechas de preguntas omitidas se reducen en promedio en 12.94 preguntas para los hombres y en 14 preguntas para las mujeres. Asimismo, tras la implementación de la reforma a la PSU, la brecha de rendimiento experimenta una disminución de 0.09 SD para los alumnos de género masculino y de 0.12 desviaciones estándar para las estudiantes de género femenino.

Adicionalmente, en la Tabla 7 se observa un patrón claro: las pruebas que registran un mayor efecto en las preguntas omitidas para las mujeres también presentan un mayor impacto para las alumnas de género femenino en términos de rendimiento. Por ejemplo, en las pruebas de Matemáticas e Historia, las estimaciones revelan que las mujeres exhiben un impacto en la disminución de la brecha del número de preguntas omitidas mayor que los hombres en aproximadamente dos preguntas. En Ciencias esta diferencia alcanza las 0.65 preguntas y en la prueba de Lenguaje se computa un mayor efecto para los hombres. El panel B indica que justamente en las pruebas de Matemáticas e Historia las mujeres registran un efecto heterogéneo de mayor magnitud. Es más, en el test de Matemáticas solamente las estudiantes de género femenino se ven beneficiadas de manera significativa por la reforma. Asimismo, en la prueba de Lenguaje y Comunicación, se estima un mayor efecto para los hombres, aunque esta diferencia no es estadísticamente distinta de cero.

Estos resultados son congruentes con las conclusiones de las investigaciones acerca del impacto de la regla de penalización en las brechas de género presentadas en la sección 2 (por ejemplo, Baldiga 2014; Ben-Shakhar and Sinai 1991; Coffman and Klinowski 2020). Adicionalmente, si consideramos la evidencia empírica de los estudios que relacionan el grado de aversión al riesgo con características individuales y que concluyen que, en promedio, las mujeres son más aversas al riesgo que los hombres, los resultados heterogéneos entre hombres y mujeres pueden interpretarse como evidencia a favor del mecanismo de aversión al riesgo mediante el cual opera una regla de penalización. En otras palabras, dentro del grupo más averso al riesgo (colegios no privados) las mujeres son relativamente más aversas que los hombres y se ven más beneficiadas por el cambio de política.

A raíz del modelo presentado en la sección 3 y la estática comparativa respecto al nivel de conocimiento o habilidad de un individuo, se concluye que el efecto de la eliminación de la penalización puede variar dependiendo de la probabilidad de éxito que posee un alumno al momento de contestar. Esta probabilidad está muy relacionada con el nivel de habilidad y/o conocimientos de

cada alumno, por ejemplo, si un estudiante posee un nivel de conocimiento o habilidad más grande, este es relativamente más probable que conteste de manera acertada en comparación a un inscrito con menos conocimiento. En esa línea, es interesante explorar si existen efectos heterogéneos según el nivel de habilidad/conocimiento de los alumnos que rinden la PSU y determinar si las personas que poseen diferentes probabilidad de éxito se ven afectadas de distinta manera por la eliminación de la penalización.

Con el fin de analizar estos posibles efectos heterogéneos, se utilizará como *proxy* de habilidad/conocimiento el puntaje NEM (Notas de Enseñanza Media) de los postulantes. De esta forma, los inscritos son clasificados según su posición en la distribución del puntaje NEM de cada proceso de selección y posteriormente son agrupados en quintiles, fragmentando la muestra total en cinco submuestras correspondiente a cada quintil. El panel A de la Tabla 8 muestra los resultados de la estimación de la especificación (7) incluyendo las interacciones del indicador $Post_t * NoPrivado_{ic}$ con diferentes dummies correspondientes al primer quintil, los quintiles intermedios y el último quintil de la distribución del puntaje NEM. Por otra parte, el panel B utiliza como *proxy* de habilidad el promedio de notas obtenido por cada alumno durante la enseñanza media para evaluar si los resultados dependen de la definición de la *proxy*.

El panel A de la Tabla 8 muestra que los alumnos del primer quintil de habilidad ven reducida la brecha de preguntas omitidas en una mayor magnitud que el resto de los estudiantes. Esto se puede deber a que los alumnos de mayor habilidad tienen una tasa de omisión más baja y por ende, una vez que ocurre el cambio de política se percibe un menor impacto para este grupo de estudiantes. Por otro lado, en términos del rendimiento obtenido en las pruebas, se estima que los alumnos pertenecientes al último quintil de conocimiento/habilidad se ven significativamente menos beneficiados que los estudiantes del resto de la distribución. A su vez, para dos de los cuatro instrumentos que constituyen la PSU, se observa que los alumnos de los quintiles intermedios de la distribución de habilidad presentan un efecto mayor de la política que aquellos pertenecientes al primer quintil. En otras palabras, un aumento en el nivel de conocimiento no implica necesariamente un mayor impacto de la eliminación de la penalización por respuestas incorrectas. Estos resultados son coherentes (en parte) con los ejercicios de simulación de la estática comparativa respecto a la probabilidad de éxito en la contestación de una pregunta presentados en la sección A.2. Además, estos resultados contrastan con los hallazgos encontrados en la investigación de Coffman and Klinowski (2020) sobre el impacto de la reforma de la PSU sobre las brechas de género, en la cual se computa un mayor efecto para los alumnos de mayor habilidad. Finalmente, el panel B indica que los resultados no son sensibles a la elección de la *proxy* utilizada.

7. Consecuencias de la política

Si bien el Z-puntaje se posiciona como una medida adecuada para medir el rendimiento, el resultado que reciben los alumnos se llama puntaje PSU, el cual es utilizado por los estudiantes en el proceso de postulación a instituciones de educación superior y a los distintos medios de financiamientos estatales. Este puntaje es construido por el DEMRE a partir del número de preguntas correctas (ajustadas por las respuestas incorrectas en el caso de la existencia de la regla de penalización) y siguen una distribución normal con media 500 puntos, desviación estándar 110 puntos, con un mínimo 150 puntos y un máximo de 850 puntos. A fin de dar un sentido más práctico y tangible a los resultados presentados anteriormente, se estima el impacto de la reforma en la Prueba de Selección Universitaria sobre las brechas de rendimiento entre colegios privados y no privados utilizando el puntaje PSU proporcionado por el DEMRE. La Tabla 9 presenta las estimaciones utilizando esta medida de rendimiento.

El panel A de la Tabla 9 muestra que, tras la modificación en las regla de la PSU, la brecha general de puntajes por tipo de colegios disminuye significativamente en 5.24 puntos PSU. Esta reducción es mucho mayor en el caso de las pruebas humanistas, la cuales registran una disminución de la brecha de 8.08 y 9.96 puntos el test de Lenguaje e Historia respectivamente. Por otro lado, dentro de las evaluaciones relacionadas con las ciencias exactas, la prueba de Ciencias Naturales presenta un decrecimiento significativo en la brecha de 2.65 puntos PSU. Sin embargo, no es posible descartar que la reducción en la brecha de puntajes en la prueba de Matemáticas es distinto de cero a los niveles usuales de significancia. Estos resultados son compatibles con la hipótesis que existen diferencias en la dificultad promedio relativa entre los test humanistas y científicos o bien, que el nivel de conocimientos por parte de los alumnos es distinto (y favorable) para los test de Historia y Lenguaje.

Los resultados del panel B de la Tabla 9 indican que las estimaciones para las pruebas humanistas son robustas en términos de magnitud y significancia. Sin embargo, al usar una muestra restringida para el período entre los años 2013-2016, el efecto estimado de la política sobre la brecha de puntajes en la prueba de Ciencias Naturales es cercano a cero y carece de significancia estadística. Este último resultado nuevamente puede ser interpretados como que la reducción de la brecha en 2.65 puntos en el test de Ciencias documentado en el panel A probablemente está conducido por mayores diferencias entre colegios privados y no privados en los rendimientos previos al año 2013.

En Chile, los resultados de la PSU son utilizados por las universidades para seleccionar a los postulantes que ingresan a cada uno de sus programas de estudios. Por otro lado, el promedio simple de los puntajes PSU obtenidos en las pruebas de Lenguaje y Matemáticas constituye un insumo importante en términos de los requisitos de índole académicos en el proceso de pos-

tulación a becas y créditos otorgadas por el estado chileno. En esa línea, un efecto inmediato de la penalización por respuestas incorrectas es la caída en las brechas referentes al número de preguntas omitidas y al rendimiento de los alumnos según el tipo de colegio de egreso. A su vez, en el largo plazo la política impacta a la elegibilidad de los postulantes en distintas opciones de financiamiento y a variables relacionadas con la matrícula en los programas de educación superior.

El Ministerio de Educación (Mineduc) ofrece diversas becas y créditos que ayudan a cubrir los costos de arancel de los diversos programas de estudios de las instituciones de educación superior chilenas. Para acceder a estos beneficios estudiantiles, los postulantes deben cumplir una serie de requisitos académicos y socioeconómicos. Dentro del primer grupo de requisitos, el puntaje PSU promedio entre las pruebas de Lenguaje y Matemáticas se posiciona como uno de los determinantes en la elegibilidad de los alumnos a estos medios de financiamiento. Por ejemplo, para ser beneficiario del Crédito con Aval del Estado (CAE) y/o al Fondo Solidario de Crédito Universitario (FSCU) es necesario obtener un puntaje PSU promedio igual o superior a 475 puntos. Por otro lado, para acceder a algunas becas del estado es necesario contar con un puntaje PSU promedio de al menos 500 puntos. En particular, para recibir la beca Vocación de Profesor (BVP) los postulantes requieren tener un promedio PSU igual o superior a 600 puntos. A sí mismo, algunos beneficios exigen que el estudiante se encuentre matriculado en una universidad que pertenezca al Consejo de Rectores, las cuales tienen como requisito mínimo de postulación tener un puntaje PSU promedio igual o mayor que 450 puntos.

Dado que la eliminación de la regla de penalización por respuestas incorrectas reduce la brecha de rendimiento en la PSU entre colegios privados y no privados, es importante indagar si el cambio de política se materializa en la probabilidad de cumplir los distintos requisitos académicos de las becas y créditos. La Tabla 10 presenta los resultados de la estimaciones de la especificación (7) utilizando como variables dependientes cuatro indicadores que toman el valor 1 si el inscrito cumple con un determinado requisito académico relacionados al proceso de postulación de becas y créditos. La primera columna de la Tabla 10 estudia el efecto del cambio de política en las brechas sobre la probabilidad de cumplir con el requisito de obtener un puntaje promedio PSU entre las pruebas de Lenguaje y Matemáticas mayor o igual a 450 puntos. La segunda columna examina el cumplimiento del requisito para aplicar a los créditos estatales, mientras que la tercera y cuarta columna analizan los dos requerimientos generales para la obtención de algunas becas de arancel.

Las estimaciones de la Tabla 10 revelan que la política no impacta de manera significativa a las brechas en términos de la probabilidad de cumplir el requisito mínimo de puntaje promedio PSU para poder postular a las universidades del CRUCH (450 puntos). Por otro lado, la brecha en la probabilidad de cumplir el requerimiento académico de los créditos estatales se reduce en 0.08% tras la eliminación de la penalización. Si bien esta disminución es estadísticamente significativa, carece de relevancia económica considerando que la brecha promedio antes del 2015 alcanza el valor de

35,07%. No obstante lo anterior, una vez que la modificación de la PSU entra en vigencia, la brecha entre los dos tipos de colegios en términos de la probabilidad de cumplir los requisitos académicos de las becas de arancel se reduce significativamente. En efecto, la brecha en la probabilidad de obtener un promedio PSU superior a 500 puntos alcanza una disminución de 1.3 puntos porcentuales y la brecha en la probabilidad de cumplir el requisito de la beca Vocación de Profesor cae en 3.2 puntos porcentuales. En otras palabras, estos resultados sugieren que la política tiene un impacto en el cumplimiento de los requisitos académicos más exigentes y no así en los requerimientos menos estrictos. Una posible explicación para este hallazgo radica en que la eliminación de la penalización podría afectar de manera distinta a los estudiantes dependiendo de su posición en la distribución de puntajes PSU.

Como se mencionó anteriormente, los resultados obtenidos por los postulantes en cada uno de las pruebas que componen el proceso de selección universitaria son insumos importantes para determinar la admisión a los programas de estudios impartidos por las universidades. De igual forma, los ponderadores (de cada programa) para obtener el puntaje ponderado de postulación, el número total de cupos de cada plan de estudios, las vías de admisión especial, la demanda de los estudiantes y sus preferencias a la hora de postular son factores fundamentales en la admisión a las universidades chilenas. A pesar de este carácter multifactorial, es importante analizar como la eliminación de la penalización por respuestas incorrectas impacta a la brecha de matriculas por tipo de colegio de egreso y en la brecha sobre la calidad de los programas a los cuales acceden los alumnos.

En lo relativo a la calidad de los distintos programas universitarios se utilizará el puntaje de corte¹⁵ para construir una *proxy*. Específicamente, se construyen indicadores que toman el valor 1 si es que un plan de estudios califica como de “alta calidad” y cero en caso que sean clasificados como programas de “baja calidad”. Para ello, se establecen tres criterios de clasificación. El primero de ellos otorga una clasificación de alta calidad si el puntaje de corte de un determinado programa es mayor que la mediana de la distribución anual de puntajes de cortes. El segundo y tercer criterio son más estrictos y consideran el percentil 75 y 90 (respectivamente) de la distribución anual de puntajes de cortes para catalogar a un plan de estudios de alta calidad.

La Tabla 11 presenta los resultados de la estimación del impacto de la política en las brechas entre colegios privados y no privados sobre la probabilidad de estar matriculado en algún programa de estudios y en la probabilidad de que el alumno sea admitido en una carrera de “alta calidad” según los tres criterios mencionados en el párrafo anterior. En la muestra correspondiente a los procesos de admisión entre los años 2010 y 2017, sólo el 28 por ciento de los inscritos en la PSU logran matricularse en algún programa de las universidades pertenecientes al Consejo de Rectores.

¹⁵El puntaje de corte de una carrera o programa de estudios corresponde al puntaje ponderado de la última persona matriculada en dicho programa y que ingresó a este por vía regular (se excluyen las vías de ingreso BEA Y PACE).

Es intuitivo pensar que la reducción en la brechas de puntajes entre los tipos de establecimientos educacionales se traduzca en una menor brecha en las tasas de matriculas. La primera columna de la Tabla 11 revela que tras la modificación en la regla de penalización, la brecha en términos de la probabilidad de estar matriculado en algún plan de estudios se reduce significativamente en 15 puntos porcentuales a favor de los alumnos egresados de escuelas no privadas. Sin embargo, dentro del grupo de individuos que se matriculan en alguna universidad, bajo ninguno de los tres criterios de clasificación se observa un efecto positivo del cambio de política sobre la calidad de los programas a los cuales ingresan los alumnos de colegios no privados. Por el contrario, la brecha en la probabilidad de ser admitido en programa que posee un puntaje de corte en el 10 por ciento más alto de la distribución aumenta en dos puntos porcentuales¹⁶. Esto es consistente con los resultados heterogéneos por nivel de habilidad/conocimiento, donde los alumnos más beneficiados son aquellos que pertenecen a la cola inferior de la distribución del puntaje NEM.

8. Conclusiones

Este estudio se posiciona como una de los primeros en examinar el cambio de la norma de descuento por respuestas erradas en la PSU en términos del análisis de las brechas de los rendimientos desde una perspectiva socioeconómica. A lo largo del análisis, se encuentra que la eliminación de la penalización por respuestas incorrectas redujo significativamente la cantidad de respuestas omitidas de los colegios no privados en comparación a la de establecimientos privados. Efectivamente, desde el año 2015 las discrepancias en la tasa de omisión entre los dos tipos de colegios desaparece por completo y con ello se observa una disminución de la brecha promedio de rendimiento correspondiente a un 9% de la brecha promedio histórica. Esta reducción beneficia en mayor parte a mujeres que a hombres y no necesariamente es creciente en el nivel de habilidad o conocimiento de los alumnos.

Las brechas en el rendimiento se manifiestan en diferencias en la tasa de matrícula entre alumnos de colegios privados y no privados. En esa línea, la política reduce la brecha en la probabilidad que un alumno se matricule en algún programa de estudios de las universidades del Consejo de Rectores en quince puntos porcentuales. Además, tras la reforma de la PSU los estudiantes de colegios no privados aumentan relativamente la probabilidad de cumplir los requisitos para la obtención de becas estatales de arancel, el cual representa uno de los principales medios de financiamiento para este grupo de la población.

La evidencia presentada en este estudio es un aporte relevante para las instituciones encargadas del diseño e implementación de futuros instrumentos utilizados en el proceso de selección universi-

¹⁶Si bien la Tabla 11 muestra otros efectos estadísticamente significativos en las especificaciones que no incluyen controles individuales, su significancia estadística desaparece ante la inclusión de estas variables de control. Por esta razón, no se discute la presencia de estos efectos.

tario chileno y si bien la validez externa de los resultados de esta investigación es cuestionable, la experiencia de Chile puede ser un punto de partida importante para elaborar posibles reformas en los procesos de selección de otros países similares.

Por último, es importante recalcar que a raíz del estallido social en Chile en Octubre del 2019, la PSU fue cuestionada ampliamente en la opinión pública y el proceso de selección universitaria para el año 2020 tuvo una serie de complicaciones relacionadas con manifestaciones en contra de la prueba de selección universitaria. A pesar que el reemplazo de la PSU por un nuevo instrumento de selección es inminente, los resultados de esta investigación pueden ser un insumo importante para decisiones de políticas públicas educacionales independiente si la PSU es o no reemplazada, y a su vez, apunta a las principales críticas de la actual prueba: la desigualdad y segregación.

Referencias

- Akyol, Ş. P., Key, J., and Krishna, K. (2016). Hit or miss? test taking behavior in multiple choice exams. *National Bureau of Economic Research Working Paper No. 22401*.
- Baldiga, K. (2014). Gender differences in willingness to guess. *Management Science*, 60(2):434–448.
- Ben-Shakhar, G. and Sinai, Y. (1991). Gender differences in multiple-choice tests: the role of differential guessing tendencies. *Journal of Educational Measurement*, 28(1):23–35.
- Beyer, H. (2009). Igualdad de oportunidades y selección a las universidades. *Puntos de Referencia*, 330:1–10.
- Coffman, K. B. and Klinowski, D. (2020). The impact of penalties for wrong answers on the gender gap in test scores. *Proceedings of the National Academy of Sciences*.
- Dohmen, T., Falk, A., Huffman, D., and Sunde, U. (2010). Are risk aversion and impatience related to cognitive ability? *American Economic Review*, 100(3):1238–60.
- Eckel, C. C. and Grossman, P. J. (2008). Men, women and risk aversion: Experimental evidence. *Handbook of experimental economics results*, 1:1061–1073.
- Guiso, L. and Paiella, M. (2008). Risk aversion, wealth, and background risk. *Journal of the European Economic association*, 6(6):1109–1150.
- Halek, M. and Eisenhauer, J. G. (2001). Demography of risk aversion. *Journal of Risk and Insurance*, pages 1–24.
- Hartog, J., Ferrer-i Carbonell, A., and Jonker, N. (2002). Linking measured risk aversion to individual characteristics. *Kyklos*, 55(1):3–26.
- Jerrim, J., Parker, P., and Shure, D. (2019). Bullshitters. who are they and what do we know about their lives?
- Jianakoplos, N. A. and Bernasek, A. (1998). Are women more risk averse? *Economic inquiry*, 36(4):620–630.
- Koljatic, M. and Silva, M. (2010). Algunas reflexiones a siete años de la implementación de la PSU. *Estudios públicos*, 120:125–146.
- Koljatic, M., Silva, M., and Cofré, R. (2013). Achievement versus aptitude in college admissions: A cautionary note based on evidence from chile. *International Journal of Educational Development*, 33(1):106–115.

Mas-Colell, A., Whinston, M. D., Green, J. R., et al. (1995). In *Microeconomic theory*, volume 1, chapter 6, pages 167–205. Oxford university press New York.

Tapia Griñen, P. (2012). Estudios basados en la aversión al riesgo.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*, *Second edition*. Princeton university press.

9. Tablas

Tabla 1: Promedio de preguntas omitidas y Z-puntajes por colegio de egreso

Panel A: Número de preguntas omitidas								
Año	Lenguaje		Matemáticas		Historia y C. Sociales		Ciencias	
	Privado	No Privado	Privado	No Privado	Privado	No Privado	Privado	No Privado
2010	11,01	19,66	17,18	32,81	14,43	24,85	20,41	32,22
2011	11,81	21,33	17,49	34,05	14,80	26,70	20,31	33,92
2012	11,84	22,07	18,38	36,25	14,48	27,68	20,46	35,34
2013	11,54	22,60	20,86	38,96	15,94	30,30	21,71	37,45
2014	10,93	20,21	19,97	36,34	15,32	28,09	21,09	35,82
2015	0,68	1,54	0,93	1,94	0,39	0,90	0,47	1,52
2016	0,36	0,78	0,77	1,31	0,31	0,57	0,38	0,98
2017	0,43	0,82	0,64	1,09	0,23	0,48	0,32	0,73

Panel B: Rendimiento (Z-puntajes)								
Año	Lenguaje		Matemáticas		Historia y C. Sociales		Ciencias	
	Privado	No Privado	Privado	No Privado	Privado	No Privado	Privado	No Privado
2010	1,06	-0,11	1,30	-0,13	1,03	-0,10	1,10	-0,12
2011	1,06	-0,11	1,33	-0,14	1,04	-0,10	1,11	-0,13
2012	1,09	-0,12	1,30	-0,15	1,05	-0,11	1,14	-0,14
2013	1,03	-0,12	1,30	-0,15	1,09	-0,12	1,16	-0,15
2014	1,00	-0,12	1,27	-0,15	1,08	-0,13	1,19	-0,14
2015	0,86	-0,10	1,22	-0,14	0,94	-0,11	1,13	-0,13
2016	0,86	-0,10	1,25	-0,14	0,99	-0,11	1,05	-0,12
2017	0,89	-0,10	1,26	-0,14	1,02	-0,11	1,11	-0,12

Notas: Preguntas omitidas y Z-puntajes en promedio por cada tipo de colegio separados por prueba. Esta tabla considera promedios simples al interior de cada grupo de colegios.

Tabla 2: Resultados de las estimaciones de la especificación (7) utilizando las preguntas omitidas como variable dependiente
Panel A: Por tipo de prueba y muestra completa

	General	LyC	Mate	C.S	C.N	General	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	-13,51*** (0,56)	-9,51*** (0,42)	-16,37*** (0,70)	-13,96*** (0,60)	-15,16*** (0,67)	-13,63*** (0,57)	-9,57*** (0,42)	-16,55*** (0,71)	-14,03*** (0,61)	-15,30*** (0,67)
Controles	Sí	Sí	Sí	Sí	Sí	No	No	No	No	No
R^2	0,59	0,48	0,65	0,56	0,66	0,58	0,47	0,64	0,56	0,66
Obs	4.640.091	1.459.469	1.449.195	865.844	865.167	6.176.646	1.942.245	1.927.346	1.160.674	1.145.931

Panel B: Por tipo de prueba y muestra restringida para los años 2013-2016

	General	LyC	Mate	C.S	C.N	General	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	-11,32*** (0,51)	-7,37*** (0,37)	-14,31*** (0,64)	-11,28*** (0,56)	-13,02*** (0,61)	-11,29*** (0,51)	-7,31*** (0,37)	-14,29*** (0,64)	-11,28*** (0,56)	-12,99*** (0,60)
Controles	Sí	Sí	Sí	Sí	Sí	No	No	No	No	No
R^2	0,65	0,54	0,72	0,65	0,72	0,64	0,53	0,71	0,64	0,71
Obs	2.297.204	723.079	717.728	419.595	436.312	3.050.156	960.090	952.413	561.817	575.325

Notas: Estimaciones para el coeficiente de interés de la especificación (7) utilizando como variable dependiente el número de preguntas omitidas para cada prueba. El panel A y B muestran las estimaciones con la data "General" (la cual consiste en la muestra que incluye a todas las pruebas) y con las submuestras resultantes de dividir la data "General" por tipo de prueba. En el panel B, se restringe la muestra conservando los datos para los dos años previos al cambio de política y los dos años posteriores a la misma. Los controles incluyen la edad, el género, el promedio de notas de enseñanza media, estado civil, situación laboral, el plan de residencia en caso de continuar estudios superiores, número de integrantes del grupo familiar, número de personas que trabajan en el grupo familiar, miembro de la familia que es el jefe de hogar, el ingreso familiar bruto, el tipo de cobertura de salud, educación de los padres, situación laboral de los padres, provincia del establecimiento educacional del alumno y el número de veces que ha rendido la PSU anteriormente. Además, para todas las regresiones con la data "General" se incluyen efecto fijos por tipo de prueba. Los errores estándar clusterizados por establecimiento educacional se muestran entre paréntesis.
 * $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$ "

Tabla 3: Resultados de las estimaciones de la especificación (7) utilizando como variable dependiente los Z-puntajes

Panel A: Por tipo de prueba y muestra completa										
	General	LyC	Mate	C.S	C.N	General	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	0,12*** (0,01)	0,16*** (0,01)	0,04*** (0,01)	0,16*** (0,02)	0,12*** (0,02)	0,12*** (0,01)	0,16*** (0,01)	0,05*** (0,01)	0,15*** (0,02)	0,13*** (0,02)
Controles	Sí	Sí	Sí	Sí	Sí	No	No	No	No	No
R^2	0,58	0,56	0,67	0,51	0,63	0,49	0,41	0,49	0,36	0,43
Obs	4.640.091	1.459.469	1.449.195	865.844	865.167	6.176.646	1.942.245	1.927.346	1.160.674	1.145.931
Brecha Promedio antes del 2015	1,28	1,16	1,44	1,16	1,27	1,28	1,16	1,44	1,16	1,27
Panel B: Por tipo de prueba y muestra restringida para los años 2013-2016										
	General	LyC	Mate	C.S	C.N	General	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	0,11*** (0,01)	0,16*** (0,01)	0,07*** (0,01)	0,18*** (0,02)	0,02 (0,02)	0,11*** (0,01)	0,15*** (0,01)	0,08*** (0,02)	0,19*** (0,02)	0,04* (0,02)
Controles	Sí	Sí	Sí	Sí	Sí	No	No	No	No	No
R^2	0,58	0,55	0,67	0,51	0,63	0,42	0,40	0,49	0,37	0,44
Obs	2.297.204	723.079	717.728	419.595	436.312	3.050.156	960.090	952.413	561.817	575.325

Notas: Estimaciones para el coeficiente de interés de la especificación (7) utilizando como variable dependiente el rendimiento del alumno medido como el Z-puntaje. El panel A y B muestran las estimaciones con la data "General" (la cual consiste en la muestra que incluye a todas las pruebas) y con las submuestras resultantes de dividir la data "General" por tipo de prueba. En el panel B, se restringe la muestra conservando los datos para los dos años previos al cambio de política y los dos años posteriores a la misma. Los controles incluyen la edad, el género, el promedio de notas de enseñanza media, estado civil, situación laboral, el plan de residencia en caso de continuar estudios superiores, número de integrantes del grupo familiar, número de personas que trabajan en el grupo familiar, miembro de la familia que es el jefe de hogar, el ingreso familiar bruto, el tipo de cobertura de salud, educación de los padres, situación laboral de los padres, provincia del establecimiento educacional del alumno y el número de veces que ha rendido la PSU anteriormente. Además, para todas las regresiones con la data "General" se incluyen efecto fijos por tipo de prueba. Los errores estándar clusterizados por establecimiento educacional se muestran entre paréntesis. * $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$.

Tabla 4: Diferencias en el cambio en el número de preguntas omitidas y en el rendimiento (Z-puntajes) de alumnos que repiten la PSU en dos años consecutivos durante el período de 2010-2015.

	Panel A: Preguntas Omitidas				Panel B: Z-puntajes			
	Repite Lyc	Repite Mate	Repite C.S	Repite C.N	Repite Lyc	Repite Mate	Repite C.S	Repite C.N
<i>NoPrivado_{ic}</i> * 2011	3,53** (1,75)	0,39 (1,52)	1,73 (3,28)	2,80 (3,54)	-0,013 (0,126)	0,034 (0,077)	-0,078 (0,135)	-0,202 (0,183)
<i>NoPrivado_{ic}</i> * 2012	3,32*** (1,20)	0,87 (1,07)	2,70 (2,22)	2,45 (2,38)	-0,080 (0,086)	0,098* (0,054)	-0,118 (0,094)	-0,150 (0,122)
<i>NoPrivado_{ic}</i> * 2013	2,42*** (0,63)	0,23 (0,58)	1,40 (1,17)	2,04 (1,24)	0,053 (0,045)	0,027 (0,030)	-0,046 (0,051)	-0,108* (0,066)
<i>NoPrivado_{ic}</i> * 2015	-8,04*** (0,69)	-13,79*** (0,80)	-9,44*** (1,33)	-12,87*** (1,39)	0,136*** (0,044)	-0,019 (0,027)	0,170*** (0,054)	0,120* (0,065)
R^2	0,32	0,57	0,46	0,54	0,07	0,09	0,12	0,13
Obs	168.167	166.648	70.789	100.543	168.167	166.648	70.789	100.543

Notas: El panel A muestra las estimaciones para la especificación (8) utilizando como variable dependiente la diferencia de preguntas omitidas entre dos procesos de selección. El panel B usa como variable dependiente la diferencia de rendimiento entre procesos de selección, definida como la resta entre los Z-puntajes obtenido en la segunda oportunidad que el alumno rinde la prueba menos los Z-puntajes del proceso anterior. La muestra está restringida únicamente para las personas que rinden en dos oportunidades la PSU durante dos años consecutivos en el período 2010-2015. La primera observación de esta muestra comienza el año 2011 con el diferencial de puntajes entre el 2010 y 2011. Los controles incluyen las mismas variables a nivel individual utilizadas en las tablas 2 y 3. Las estimaciones consideran como año de referencia el 2014 y los errores estándar son clusterizados a nivel colegio. Finalmente, todas las regresiones consideran tendencias lineales específicas para cada colegio. * $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$ ”

Tabla 5: Tratamiento placebo. Efecto de la política en variables relacionadas a la situación laboral y al rendimiento escolar

	(a)				(b)	
	Situación laboral				Rendimiento escolar	
	Padre Cesante	Madre Cesante	Inscrito Trabaja	Número de personas del hogar que trabajan	Promedio Notas	Puntaje NEM
$Post_t^*$	0,005	-0,001	0,000	-0,002	-0,011*	-2,32*
$NoPrivado_{ic}$	(0,007)	(0,002)	(0,003)	(0,011)	(0,006)	(1,22)
Controles	Sí	Sí	Sí	Sí	Sí	Sí
R^2	0,05	0,03	0,22	0,36	0,23	0,23
Obs	1.514.267	1.538.188	1.522.960	1.507.638	1.507.638	1.507.638
$Post_t^*$	0,003	-0,001	-0,001	-0,008	-0,004	-0,781
$NoPrivado_{ic}$	(0,005)	(0,001)	(0,003)	(0,008)	(0,006)	(1,170)
Controles	No	No	No	No	No	No
R^2	0,01	0,008	0,05	0,07	0,21	0,20
Obs	2.004.350	2.004.350	2.004.350	2.004.350	2.004.396	2.004.396

Notas: Estimaciones para el coeficiente de interés de la especificación (7) considerando diferentes variables dependientes sobre el situación laboral del inscrito y su grupo familiar. Dentro de la sección (a), la columna “Padre Cesante” utiliza como variable de interés una dummy que toma el valor 1 si el padre del inscrito se encuentra cesante y cero en caso contrario. La columna “Madre Cesante” considera como variable dependiente un indicador que toma el valor 1 si la madre del alumno se encuentra cesante y cero en caso contrario. La columna “Inscrito Trabaja” utiliza una dummy que toma el valor 1 si quien rinde la PSU se encuentra trabajando y cero en caso contrario. Finalmente, la última columna de la sección (a) utiliza como variable independiente el número de personas del hogar que trabajan. Por otro parte, las dos columnas de la sección (b) utilizan el promedio de notas y el puntaje NEM como variables dependientes. Los controles incluyen las mismas variables sociodemográficas a nivel individual utilizadas en las tablas anteriores, sin embargo, para cada caso se excluyen las que son utilizadas como variable dependiente. Todas las regresiones de la sección (a) utilizan errores estándar clusterizados según el tramo de ingreso bruto familiar y las las especificaciones de la sección (b) permiten correlación de los errores a nivel colegio. * $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$ ”

Tabla 6: Resultados heterogéneos de las estimaciones de la especificación (7) dentro de los colegios no privados

Panel A: Preguntas Omitidas					
	Gral	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	-14,30*** (0,69)	-9,79*** (0,51)	-17,62*** (0,82)	-15,00*** (0,70)	-16,05*** (0,87)
$Post_t * NoPrivado_{ic}$ $*Subvencionado_{ic}$	1,29*** (0,48)	0,45 (0,39)	2,07*** (0,54)	1,75*** (0,47)	1,44** (0,66)
Controles	Sí	Sí	Sí	Sí	Sí
R^2	0,59	0,48	0,65	0,56	0,66
Obs	4.640.091	1.459.469	1.449.195	865.844	865.167
valor-p	0,01	0,2452	0,00	0,00	0,03
Panel B: Z-puntajes					
	Gral	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	0,12*** (0,01)	0,15*** (0,01)	0,05*** (0,01)	0,16*** (0,02)	0,12*** (0,02)
$Post_t * NoPrivado_{ic}$ $*Subvencionado_{ic}$	0,00 (0,007)	0,02* (0,009)	-0,02** (0,008)	0,00 (0,010)	0,00 (0,010)
Controles	Sí	Sí	Sí	Sí	Sí
R^2	0,58	0,56	0,66	0,51	0,63
Obs	4.640.091	1.459.469	1.449.195	865.844	865.167
valor-p	0,89	0,05	0,05	0,82	0,95

Notas: El panel A y B presentan los resultados de las especificación (7) utilizando un interacción adicional entre la variable $Post_t * NoPrivado_{ic}$ con una dummy que toma el valor 1 si el individuo pertenece a un colegio de dependencia particular subvencionada y cero en caso contrario. Para las regresiones con la data “General” se incluyen efecto fijos por tipo de prueba. Los controles incluyen las mismas variables a nivel individual utilizadas en las tablas 2 y 3. Además, para todas las regresiones con la data “General” se incluyen efecto fijos por tipo de prueba. Los errores estándar clusterizados por establecimiento educacional se muestran entre paréntesis. * $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$ ”

Tabla 7: Efectos heterogéneos por género de la penalización por respuestas incorrectas.

Panel A: Preguntas Omitidas					
	Gral	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	-12,94*** (0,57)	-9,67*** (0,43)	-15,28*** (0,71)	-12,96*** (0,61)	-14,80*** (0,69)
$Post_t * NoPrivado_{ic} * Mujer$	-1,06*** (0,11)	0,29*** (0,09)	-2,01*** (0,13)	-1,90*** (0,11)	-0,65*** (0,15)
Controles	Sí	Sí	Sí	Sí	Sí
R^2	0,59	0,48	0,65	0,56	0,66
Obs	4.640.091	1.459.469	1.449.195	865.844	865.167
valor-p	0,00	0,00	0,00	0,00	0,00
Panel B: Z-puntajes					
	Gral	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	0,09*** (0,01)	0,16*** (0,01)	0,02 (0,01)	0,13*** (0,02)	0,11*** (0,02)
$Post_t * NoPrivado_{ic} * Mujer$	0,03*** (0,002)	-0,001 (0,003)	0,05*** (0,003)	0,07*** (0,005)	0,02*** (0,004)
Controles	Sí	Sí	Sí	Sí	Sí
R^2	0,58	0,56	0,67	0,51	0,63
Obs	4.640.091	1.459.469	1.449.195	865.844	865.167
valor-p	0,00	0,70	0,00	0,00	0,00

Notas: El panel A y B presentan los resultados de las especificación (7) utilizando un interacción adicional entre la variable $Post_t * NoPrivado_{ic}$ con una dummy que toma el valor 1 si el individuo es de género femenino y cero en caso contrario. Para las regresiones con la data “General” se incluyen efecto fijos por tipo de prueba. Los controles incluyen las mismas variables a nivel individual utilizadas en las tablas 2 y 3. Además, para todas las regresiones con la data “General” se incluyen efecto fijos por tipo de prueba. Los errores estándar clusterizados por establecimiento educacional se muestran entre paréntesis. * $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$ ”

Tabla 8: Resultados heterogéneos de las estimaciones de la especificación (7) según quintil de habilidad/conocimiento
Panel A: Utilizando quintiles de la distribución del puntaje NEM

	Preguntas Omitidas			Z-Puntajes						
	General	LyC	Mate	C.S	C.N	General	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	-15,82*** (0,56)	-11,53*** (0,42)	-19,92*** (0,69)	-16,13*** (0,59)	-17,07*** (0,65)	0,12*** (0,01)	0,15*** (0,01)	0,07*** (0,01)	0,17*** (0,02)	0,11*** (0,02)
$Post_t * NoPrivado_{ic} * Q_{2-4}$	1,52*** (0,11)	1,50*** (0,10)	2,48*** (0,13)	1,90*** (0,11)	0,45*** (0,14)	0,005 (0,004)	0,02*** (0,004)	-0,03*** (0,004)	0,00 (0,005)	0,04*** (0,005)
$Post_t * NoPrivado_{ic} * Q_5$	7,67*** (0,20)	6,15*** (0,16)	11,58*** (0,25)	6,70*** (0,17)	7,23*** (0,27)	-0,03** (0,007)	-0,01** (0,007)	-0,05*** (0,009)	-0,05*** (0,009)	-0,03*** (0,01)
Controles	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
R^2	0,59	0,48	0,66	0,56	0,67	0,58	0,56	0,67	0,51	0,64
Obs	4.640.091	1.459.469	1.449.195	865.844	865.167	4.640.091	1.459.469	1.449.195	865.844	865.167

Panel B: Utilizando quintiles de la distribución del promedio de notas de enseñanza media

	Preguntas Omitidas			Z-Puntajes						
	General	LyC	Mate	C.S	C.N	General	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	-15,81*** (0,55)	-11,50*** (0,42)	-19,93*** (0,68)	-16,16*** (0,59)	-17,03*** (0,65)	0,12*** (0,01)	0,15*** (0,01)	0,07*** (0,01)	0,17*** (0,02)	0,11*** (0,02)
$Post_t * NoPrivado_{ic} * Q_{2-4}$	1,49*** (0,11)	1,50*** (0,10)	2,48*** (0,13)	1,93*** (0,11)	0,38*** (0,14)	0,006 (0,004)	0,02*** (0,004)	-0,03*** (0,004)	-0,00 (0,005)	0,04*** (0,005)
$Post_t * NoPrivado_{ic} * Q_5$	7,69*** (0,20)	6,12*** (0,16)	11,62*** (0,25)	6,74*** (0,17)	7,22*** (0,27)	-0,04*** (0,007)	-0,02*** (0,007)	-0,05*** (0,009)	-0,05*** (0,009)	-0,03*** (0,01)
Controles	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí	Sí
R^2	0,59	0,48	0,66	0,56	0,67	0,58	0,56	0,67	0,51	0,64
Obs	4.640.091	1.459.469	1.449.195	865.844	865.167	4.640.091	1.459.469	1.449.195	865.844	865.167

Notas: Estimaciones para los coeficientes $Post_t * NoPrivado_{ic}$ de la especificación (7) y sus interacciones con los indicadores correspondiente a los quintiles de la distribución de habilidad. En el panel A se utiliza como proxy al nivel de habilidad/conocimiento el puntaje NEM de cada alumno, mientras que el panel B usa el promedio de notas de enseñanza media como proxy. La categoría de referencia es el primer quintil de la distribución. Todas las estimaciones incluyen los mismos controles individuales de las tablas expuestas anteriormente y efectos fijos para cada quintil de habilidad. Para las regresiones con la data “General” se incluyen efecto fijos por tipo de prueba. Los errores estándar clusterizados por establecimiento educacional se muestran entre paréntesis. * $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$.

Tabla 9: Resultados de las estimaciones de la especificación (7) utilizando como variable dependiente el puntaje PSU
Panel A: Por tipo de prueba y muestra completa

	General	LyC	Mate	C.S	C.N	General	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	5,24*** (0,89)	8,98*** (1,08)	0,51 (0,89)	9,96*** (1,42)	2,65*** (1,18)	5,22*** (1,03)	8,68*** (1,18)	0,97 (1,04)	9,15*** (1,42)	2,87** (1,35)
Controles	Sí	Sí	Sí	Sí	Sí	No	No	No	No	No
R^2	0,56	0,55	0,61	0,49	0,61	0,40	0,40	0,44	0,34	0,42
Obs	4.642.789	1.459.580	1.450.677	866.126	865.990	6.180.525	1.942.416	1.929.531	1.161.080	1.147.048
Brecha Promedio antes del 2015	118,07	115,06	126,83	111,25	113,41	118,07	115,06	126,83	111,25	113,41

Panel B: Por tipo de prueba y muestra restringida para los años 2013-2016

	General	LyC	Mate	C.S	C.N	General	LyC	Mate	C.S	C.N
$Post_t * NoPrivado_{ic}$	5,06*** (1,04)	9,08*** (1,36)	1,48 (1,17)	9,83*** (1,98)	-0,14 (1,48)	5,46*** (1,26)	8,99*** (1,46)	2,19* (1,31)	10,02*** (2,04)	0,54 (1,76)
Controles	Sí	Sí	Sí	Sí	Sí	No	No	No	No	No
R^2	0,56	0,54	0,61	0,49	0,61	0,40	0,39	0,45	0,35	0,43
Obs	2.298.392	723.113	718.285	419.738	436.766	3.051.887	960.153	953.250	562.026	575.949

Notas: Estimaciones para el coeficiente de interés de la especificación (7) utilizando como variable dependiente el rendimiento del alumno medido como el puntaje PSU. El panel A y B muestran las estimaciones con la data "General" (la cual consiste en la muestra que incluye a todas las pruebas) y con las submuestras resultantes de dividir la data "General" por tipo de prueba. En el panel B, se restringe la muestra conservando los datos para los dos años previos al cambio de política y los dos años posteriores a la misma. Los controles incluyen la edad, el género, el promedio de notas de enseñanza media, estado civil, situación laboral, el plan de residencia en caso de continuar estudios superiores, número de integrantes del grupo familiar, número de personas que trabajan en el grupo familiar, miembro de la familia que es el jefe de hogar, el ingreso familiar bruto, el tipo de cobertura de salud, educación de los padres, situación laboral de los padres, provincia del establecimiento educacional del alumno y el número de veces que ha rendido la PSU anteriormente. Además, para todas las regresiones con la data "General" se incluyen efecto fijos por tipo de prueba. Los errores estándar clusterizados por establecimiento educacional se muestran entre paréntesis. * $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$.

Tabla 10: Efectos del cambio de política en el cumplimiento de requisitos académicos para becas y créditos

Requisitos académico para becas y créditos				
	Promedio de puntaje Lenguaje y Matemáticas			
	Mayor a 450 puntos	Mayor a 475 puntos	Mayor a 500 puntos	Mayor a 600 puntos
$Post_t * NoPrivado_{ic}$	0,002 (0,003)	0,008** (0,004)	0,013*** (0,004)	0,032*** (0,005)
Controles	Sí	Sí	Sí	Sí
R^2	0,36	0,41	0,45	0,41
Obs	1.507.638	1.507.638	1.507.638	1.507.638
	Promedio de puntaje Lenguaje y Matemáticas			
	Mayor a 450 puntos	Mayor a 475 puntos	Mayor a 500 puntos	Mayor a 600 puntos
$Post_t * NoPrivado_{ic}$	0,000 (0,003)	0,006* (0,003)	0,010*** (0,004)	0,032*** (0,005)
Controles	No	No	No	No
R^2	0,27	0,31	0,33	0,30
Obs	2.004.396	2.004.396	2.004.396	2.004.396

Notas: Estimaciones para el coeficiente de interés de la especificación (7) utilizando como variable dependiente diferentes indicadores relacionadas con diferentes requisitos de becas y créditos estatales. Estos indicadores toman el valor 1 si el puntaje promedio PSU entre las pruebas de Lenguaje y Matemáticas superan un determinado umbral (450, 475, 500 y 600 puntos). Todas las estimaciones son basadas en modelos de probabilidad lineal. Los controles incluyen las mismas variables a nivel individual de las tablas 2 y 3. Los errores estándar clusterizados por establecimiento educacional se muestran entre paréntesis. * $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$.

Tabla 11: Efectos del cambio de política en variables relacionadas a la matricula en programas de estudios

VARIABLES DE RESULTADO RELACIONADO A MATRICULAS				
	Probabilidad de estar matriculado en algún programa	Probabilidad de estar matriculado en programa de alta calidad (1)	Probabilidad de estar matriculado en programa de alta calidad (2)	Probabilidad de estar matriculado en programa de alta calidad (3)
$Post_t * NoPrivado_{ic}$	0,15*** (0,007)	0,003 (0,006)	-0,008 (0,006)	-0,02*** (0,005)
Controles	Sí	Sí	Sí	Sí
R^2	0,32	0,37	0,46	0,36
Obs	1.507.638	430.143	430.143	430.143
$Post_t * NoPrivado_{ic}$	0,15*** (0,007)	-0,031*** (0,006)	-0,046*** (0,007)	-0,043*** (0,007)
Controles	No	No	No	No
R^2	0,21	0,17	0,20	0,16
Obs	2.004.396	561.658	561.658	561.396

Notas: Estimaciones para el coeficiente de interés de la especificación (7) utilizando diferentes variables dependientes relacionadas con la matricula de los inscritos en la PSU. La primera columna utiliza un indicador que toma el valor 1 si el inscrito está matriculado en algún programa de estudio de las universidades pertenecientes al consejo de rectores. Las siguientes tres columnas utilizan una dummy que toma el valor 1 si el programa al que se matricula un individuo califica como de “alta calidad” según los tres diferentes criterios detallados en la sección 7. Todos las estimaciones se realizan a través de modelos de probabilidad lineal. Los controles incluyen las mismas variables a nivel individual de las tablas 2 y 3. Los errores estándar clusterizados por establecimiento educacional se muestran entre paréntesis. * $p < 0,1$, ** $p < 0,05$, *** $p < 0,01$.

10. Figuras

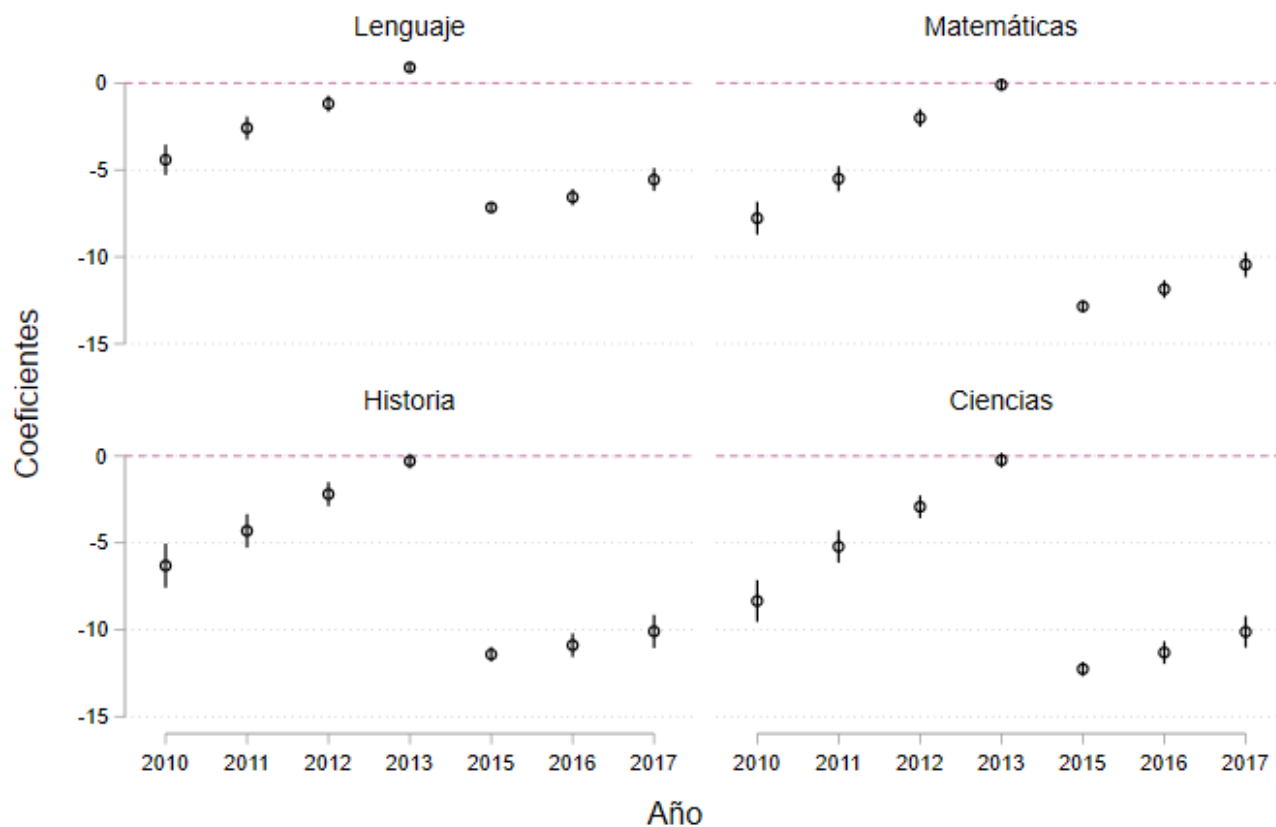


Figura 1: Brechas de preguntas omitidas entre colegios no privados y privados respecto al año la brecha del año 2014.

Notas: Esta figura muestra de manera gráfica los resultados para la estimación de los coeficientes μ_t de la especificación (8) para los años 2010-2017 excluyendo el año de referencia (2014). Se considera como variable dependiente el número de preguntas omitidas. Además, se presentan los intervalos de confianza al 95%. Las especificaciones estimadas no utilizan controles individuales y consideran errores estándar clusterizados a nivel colegio.

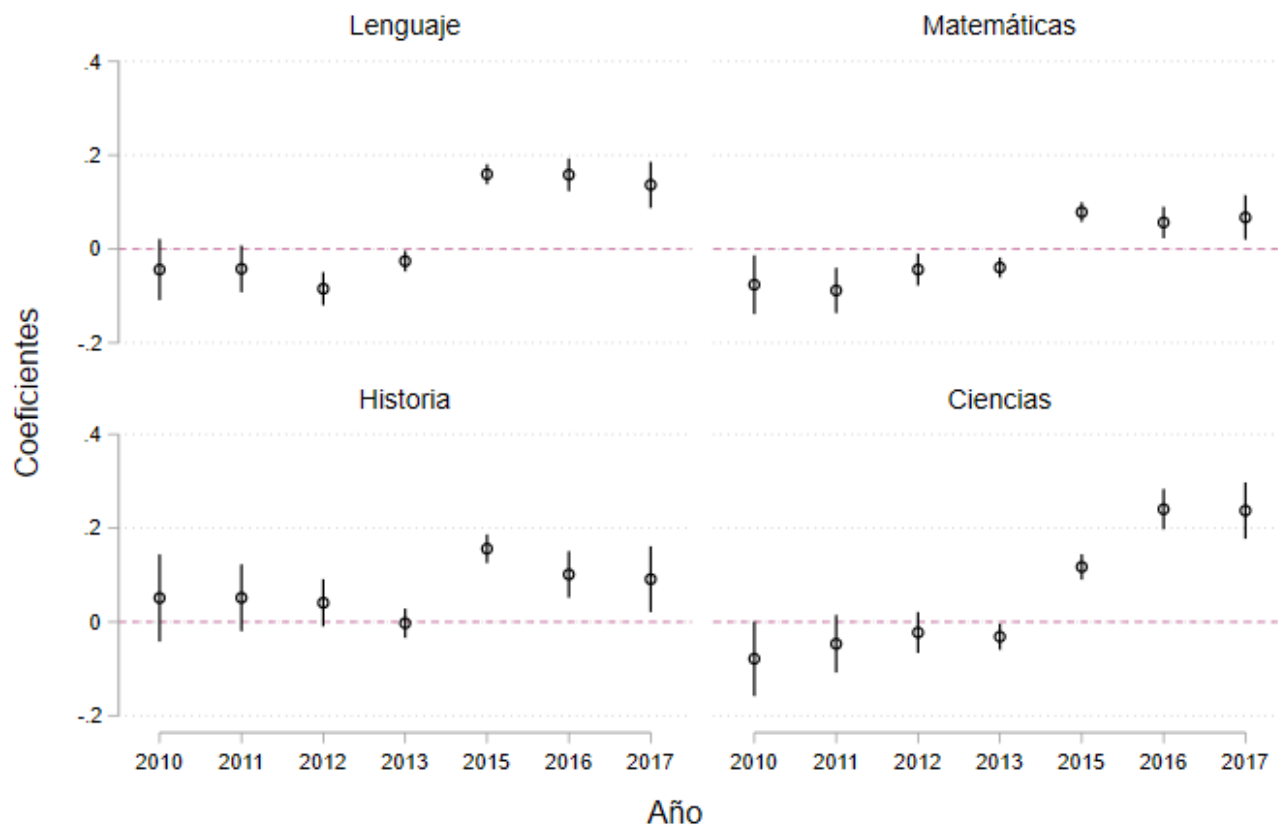


Figura 2: Brechas de Z-puntajes entre colegios no privados y privados respecto al año la brecha del año 2014.

Notas: Esta figura muestra de manera gráfica los resultados para la estimación de los coeficientes μ_t de la especificación (8) para los años 2010-2017 excluyendo el año de referencia (2014). Se considera como variable dependiente el rendimiento de un alumno medido como el Z-puntaje. Además, se presentan los intervalos de confianza al 95%. Las especificaciones estimadas no utilizan controles individuales y consideran errores estándar clusterizados a nivel colegio.

A. Apéndice

A.1. Ejemplo de función u

Un ejemplo para la función $u_i = u(x, \sigma_i)$ donde x representa un pago y σ_i es un parámetro de aversión al riesgo, es la siguiente función:

$$u(x, \sigma_i) = \begin{cases} \frac{x^{1-\sigma_i}}{1-\sigma_i} & \sigma_i \geq 0, \sigma_i \neq 1 \\ \ln(x) & \sigma_i = 1 \end{cases} \quad (\text{A.1})$$

En este ejemplo, si σ_i toma el valor cero entonces la función bernoulli es lineal, es decir, $u(x, \sigma_i) = x$ y por lo tanto el individuo sería un agente neutral al riesgo. A medida que el parámetro σ_i aumenta, la función se hace más cóncava por lo que el grado de aversión del agente es creciente en σ_i .

A.2. Simulaciones de estática comparativa

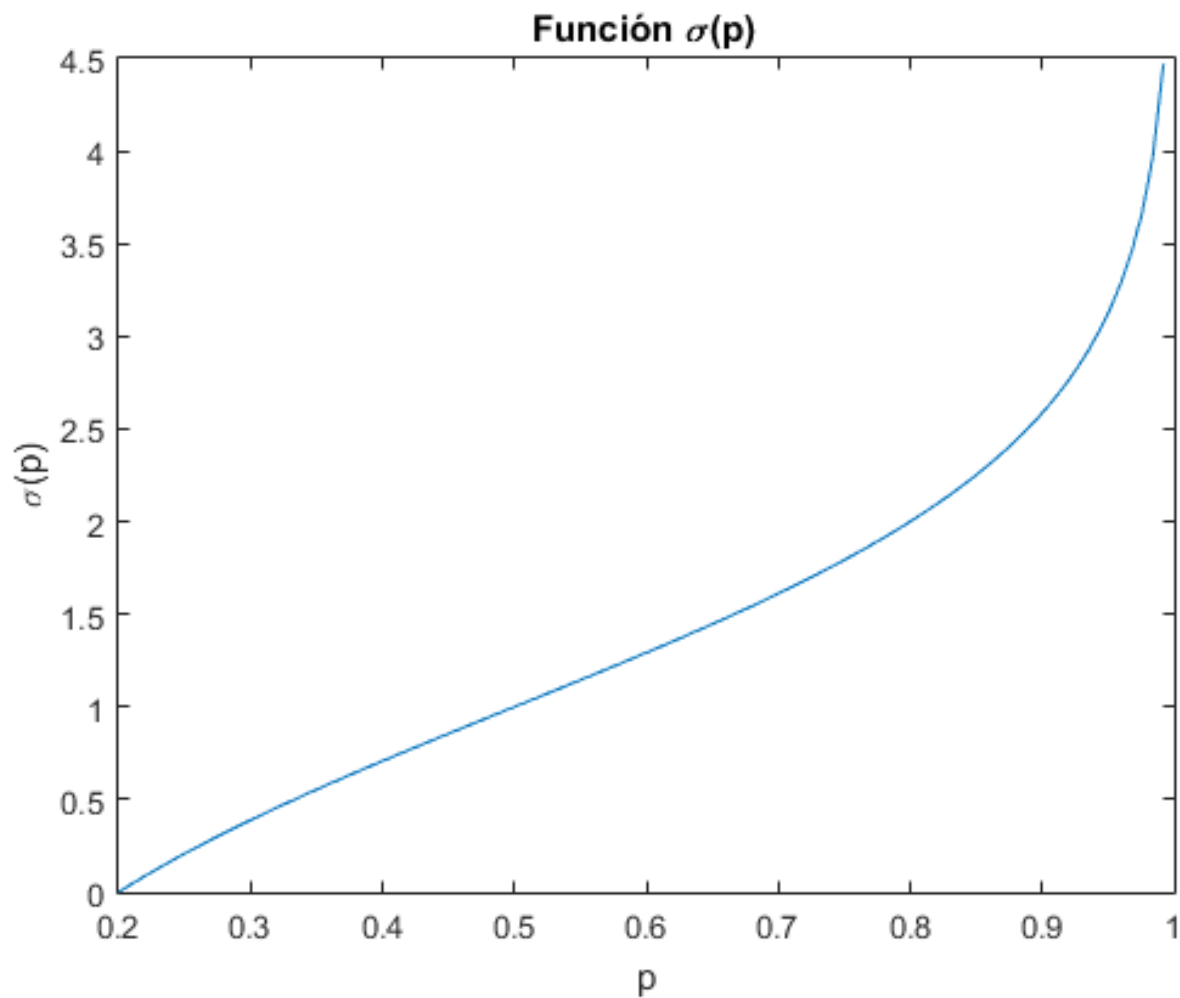


Figura A.1: Gráfico de la función $\bar{\sigma}(p)$ para la utilidad $u(x, \sigma_i)$ del ejemplo.

Notas: Esta figura representa la función $\bar{\sigma}(p)$ utilizando la función de utilidad descrita en (A.1). Este gráfico considera una grilla de 100 valores equidistantes para p en el rango $[0.2, 1]$.

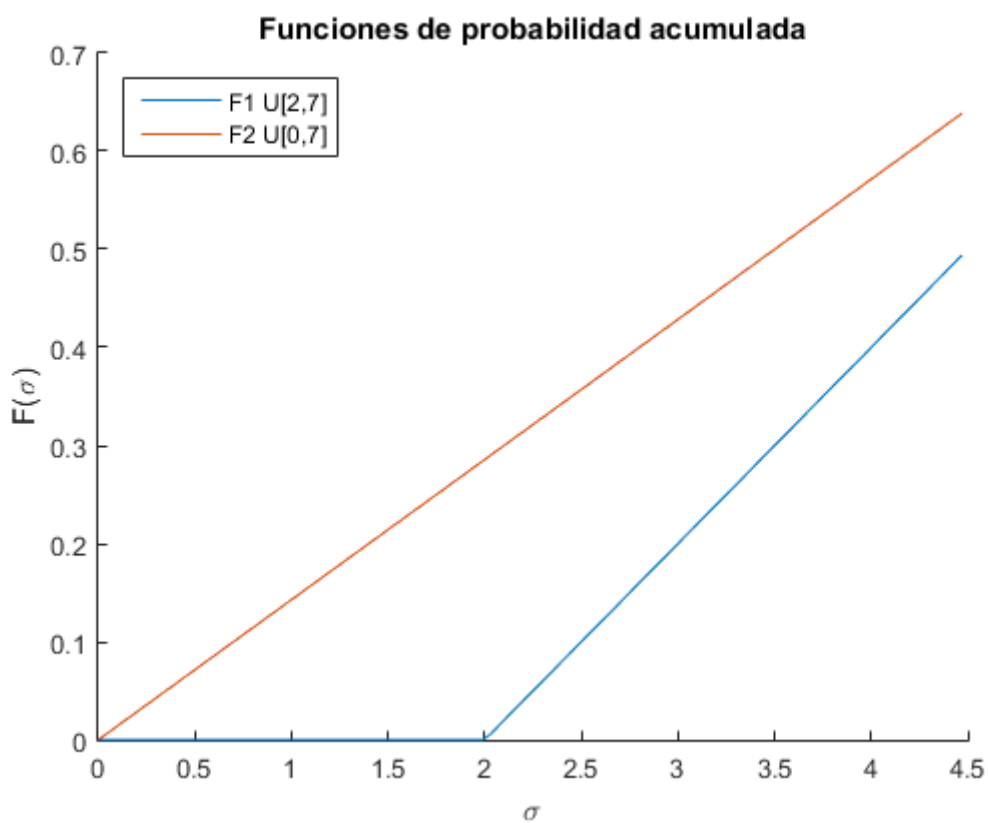


Figura A.2: Funciones de probabilidad acumulada utilizando distribuciones uniformes para σ

Notas: Esta figura muestra los resultados de simulaciones para las funciones de probabilidad acumulada para dos distribuciones uniformes que cumplen con la condición de dominancia estocástica de primer orden. La función de color naranja proviene de una distribución uniforme para el parámetro σ con soporte $[0,7]$. La función de color azul proviene de una distribución uniforme para el parámetro σ con soporte $[2,7]$. Este gráfico considera una grilla de 100 valores para σ .

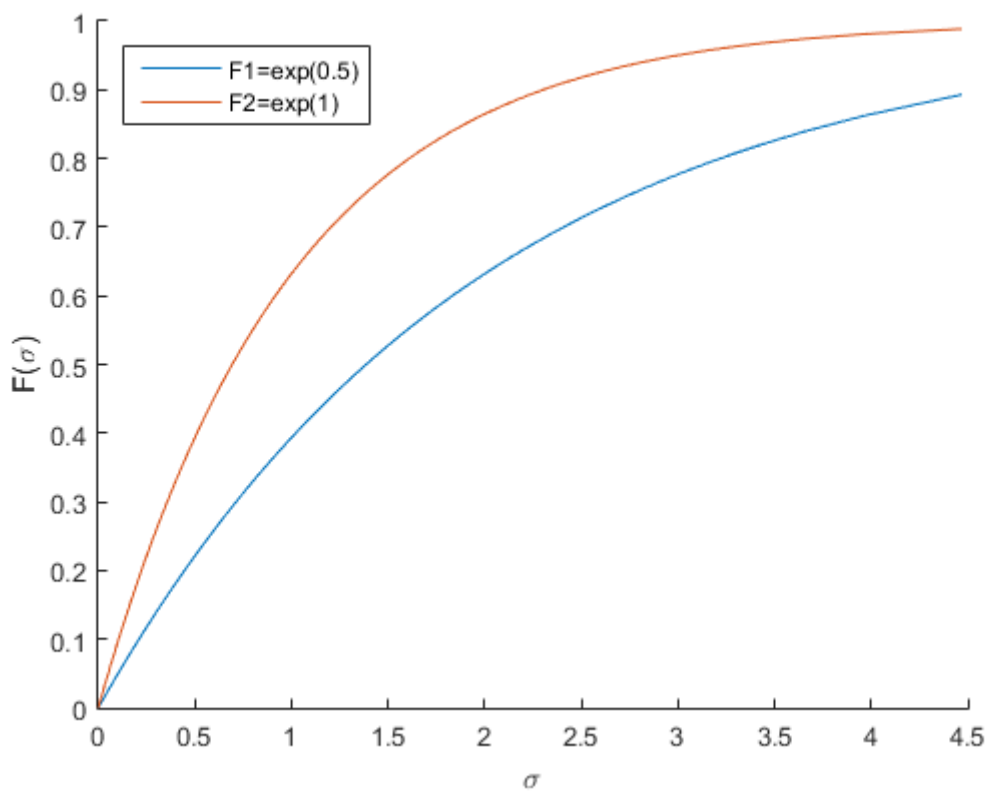


Figura A.3: Funciones de probabilidad acumulada utilizando distribuciones exponenciales para σ

Notas: Esta figura muestra los resultados de simulaciones para las funciones de probabilidad acumulada para dos distribuciones exponenciales que cumplen con la condición de dominancia estocástica de primer orden. La función de color naranja proviene de una distribución exponencial para σ con parámetro 1. La función de color azul proviene de una distribución exponencial para σ con parámetro 0,5 . Este gráfico considera una grilla de 100 valores para σ .

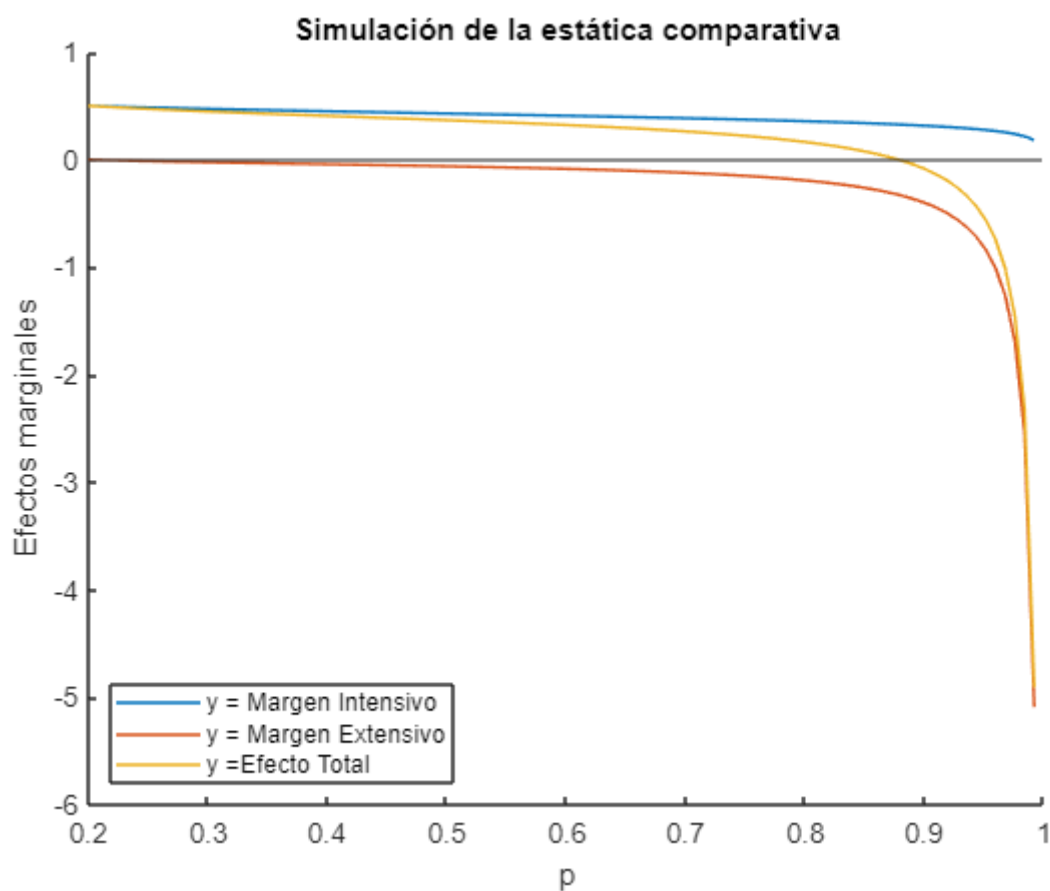


Figura A.4: Efectos marginales de un aumento en p sobre la magnitud del impacto de la política

Notas: Esta Figura muestra los resultados de la simulación del efecto marginal utilizando las distribuciones exponenciales descritas anteriormente. El margen intensivo hace referencia al primer término de la ecuación (6) mientras que el margen extensivo se refiere al segundo término de dicha ecuación. El efecto total es la suma del margen intensivo y el margen extensivo.

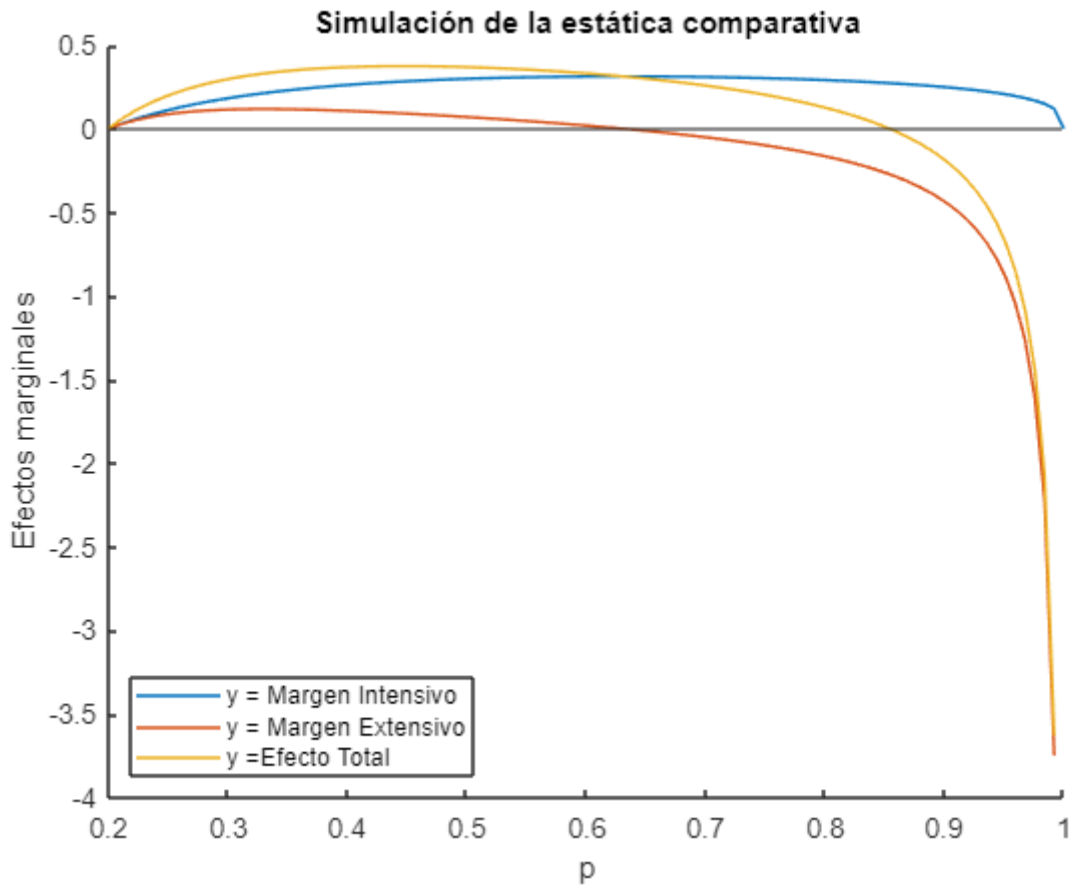


Figura A.5: Efectos marginales de un aumento en p sobre la magnitud del impacto de la política

Notas: Esta Figura muestra los resultados de la simulación del efecto marginal utilizando las distribuciones exponenciales descritas anteriormente. El margen intensivo hace referencia al primer término de la ecuación (6) mientras que el margen extensivo se refiere al segundo término de dicha ecuación. El efecto total es la suma del margen intensivo y el margen extensivo.