



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

**TESIS DE GRADO**  
**MAGISTER EN ECONOMIA**

**Sasmay, Ananías, Alberto Andrés**

**Diciembre, 2017**



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

**LA BRECHA DE GÉNERO EN LAS CALIFICACIONES  
ESCOLARES: UNA DOCUMENTACIÓN ROBUSTA PARA CHILE**

**Alberto Andrés Sasmay Ananías**

Comisión

Emilio Depetris-Chauvin

Claudio Sapelli

**Santiago, Diciembre de 2017**

# La brecha de género en las calificaciones escolares: una documentación robusta para Chile

Alberto Sasmay Ananías\*

## Resumen

Utilizando datos del Ministerio de Educación de Chile, este trabajo documenta la existencia de una brecha de género, de 0.12 desviaciones estándar en favor de las mujeres, para las calificaciones escolares de matemáticas de 8<sup>vo</sup> básico. La estimación de la brecha se obtiene mediante regresiones de mínimos cuadrados ordinarios con efectos fijos por curso. Esta brecha prevalece tras incorporar las aptitudes cognitivas y características socioeconómicas como variables explicativas. Para refinar los resultados, se incorporan variables actitudinales a través de la construcción de un índice (gracias a datos auto-reportados por los estudiantes y en concordancia a lo sugerido por el marco analítico). Tras instrumentar dicho índice actitudinal con su símil del pasado para reducir potenciales sesgos, los resultados muestran que las actitudes son económicamente relevantes para explicar las calificaciones, pero sigue existiendo la misma brecha por género. Un análisis de efectos heterogéneos, en base a características de los profesores, muestra que la brecha aumenta cuando la profesora es mujer y disminuye cuando el profesor ha estado expuesto al curso por dos o más años.

---

\*Trabajo realizado para optar al grado de Magíster en Economía, Instituto de Economía, Pontificia Universidad Católica de Chile. Este trabajo utilizó como fuente de información las bases de datos del Ministerio de Educación de Chile y de la Agencia de Calidad de la Educación. Agradezco el acceso a la información. Todos los resultados del estudio son de mi responsabilidad y en nada comprometen a dichas instituciones. Agradezco especialmente a los profesores Emilio Depetris-Chauvin y Claudio Sapelli por su constante disposición y valiosa guía durante este proceso. Agradezco a mi familia por su apoyo y formación y a mis amigos por ser una fuente inagotable de felicidad. También quiero agradecer a toda persona que me ha hecho entender un poco más la vida, pues no soy nada sin el amor que ustedes han proyectado en mí: “*if there is any truth in the world, it lies when I’m with you*”. Cualquier error u omisión son de mi absoluta responsabilidad. Dudas y/o comentarios a albertosasmay@uc.cl

## 1 Introducción

El presente trabajo intenta documentar y analizar dos grandes interrogantes en el sistema educacional chileno. La primera pregunta pretende contestar si existe una brecha de género en las calificaciones escolares de matemáticas. La segunda pregunta, frente a la existencia de la brecha, intenta responder si la brecha se mantiene una vez incorporadas las variables relevantes que explican la obtención de las calificaciones. Evidencia previa en otros países ha mostrado la existencia de una brecha de género que favorece a las mujeres en distintas asignaturas y en distintos niveles educacionales (Kiss, 2013; Cornwell, Mustard, & Van Parys, 2013; Di Liberto & Casula, 2016), por lo que resulta importante documentarlo para Chile.

La relevancia pública de esto reside en que estas brechas podrían tener consecuencias en la vida futura de los estudiantes. Por un lado, las calificaciones podrían influenciar el nivel de esfuerzo académico que ejercen los estudiantes. Si las señales entregadas a los estudiantes no están alineadas con sus desempeños futuros, se estarían generando distorsiones en los esfuerzos y creencias de los propios estudiantes respecto de sus competencias académicas<sup>1</sup> (Cornwell et al., 2013). Por otro lado, las calificaciones también influenciarían en el futuro profesional de los estudiantes, generando posibles desigualdades laborales observadas, por ejemplo, en desigualdades salariales (Loury, 1997; Bonesrønning, 2008; Lavy, 2008; Fryer & Levitt, 2010; Lavy & Sand, 2015; Terrier, 2016). Como las calificaciones en enseñanza básica son la única señal constante y periódica que reciben los establecimientos, apoderados y estudiantes acerca del desempeño de estos últimos, resulta importante identificar si estas señales presentan brechas por género. Esto se hace más valioso al no existir estudios previos que lo analicen para Chile, configurando una fuente de información útil para la discusión de políticas públicas.

Para lograr responder las interrogantes sobre la brecha, el presente trabajo utiliza datos oficiales del Ministerio de Educación de Chile. Para los estudiantes en 8<sup>vo</sup> año de enseñanza

---

<sup>1</sup>En Chile, las calificaciones y mediciones en 8<sup>vo</sup> no tienen consecuencias directas para permitir a los estudiantes entrar a mejores o peores establecimientos, pero los 4 años posteriores (enseñanza media) son cruciales para definir la entrada de los estudiantes a la educación superior, pues se consideran las calificaciones escolares y los resultados de pruebas estandarizadas. Así, señales distorsionadas previas (o sea, calificaciones que no se condicen con el verdadero rendimiento académico) podrían tergiversar los niveles de esfuerzo por ejecutar y las creencias de aptitudes futuras.

básica, se muestra que las mujeres obtienen consistentemente mejores calificaciones que los hombres. Las estimaciones iniciales realizan regresiones de mínimos cuadrados ordinarios y consideran efectos fijos a nivel del curso, mostrando una brecha de 0.12 desviaciones estándar entre los géneros. Al incorporar a las especificaciones las variables que miden aptitudes cognitivas y no cognitivas, los resultados siguen mostrando una brecha de valor similar. La incorporación de las habilidades no cognitivas resulta un aporte a la literatura que evidencia la relevancia de estas actitudes respecto de las calificaciones que otorga el docente. Las estimaciones relacionadas con estas aptitudes no cognitivas son estadísticamente significativas y teóricamente relevantes.

Debido a que las actitudes son auto-reportadas y se caracterizan como un índice construido para fines de este trabajo, podrían existir sesgos en las estimaciones. Así, la documentación incorpora datos sobre las aptitudes no cognitivas de años anteriores para instrumentar las aptitudes concurrentes. Nuevamente los resultados evidencian una brecha de género en las calificaciones, favorable para las mujeres, y con valores casi iguales a los iniciales. Finalmente, el trabajo analiza si las estimaciones se mantienen al considerar especificaciones que tomen en cuenta algunas características de los profesores. Los resultados muestran que cuando la profesora de matemáticas es mujer, la brecha aumenta, y cuando el profesor ha sido profesor del curso el año anterior, la brecha disminuye. Estos resultados son estadísticamente significativos y sirven de insumo para evidenciar posibles mecanismos detrás de la brecha de género como, por ejemplo, algún tipo de sesgo discriminatorio por parte del profesor.

La literatura previa ha evidenciado y analizado brechas de género en los desempeños académicos de los alumnos<sup>2</sup>. Respecto de las calificaciones, Kiss (2013) justifica que el escenario teórico ideal contempla que los estudiantes dentro de una misma clase debiesen tener las mismas calificaciones una vez que se consideren las características relevantes para calificarlos. Dichas características, como se adelantó, serían las aptitudes cognitivas y no cognitivas que la literatura ha considerado como relevante, en especial aquellas observables. Frente a este escenario, cualquier brecha de género identificada consistentemente implicaría una anomalía en el sistema de calificación.

---

<sup>2</sup>En evaluaciones estandarizadas de matemáticas, los hombres obtienen consistentemente mejor desempeño, pero en lectura las mujeres son las que obtienen mejor desempeño (Dee, 2007; Paredes, 2014)

Al mismo tiempo, autores han explicado la dificultad de contar con datos sobre variables no cognitivas (o actitudinales). Si no se considera este aspecto, las especificaciones estarían omitiendo una variable relevante y sesgando las estimaciones pertinentes (Cornwell et al., 2013; Kiss, 2013; Golsteyn & Schils, 2014; Terrier, 2015). Este trabajo, al aprovechar los datos de las actitudes de los alumnos, estaría en concordancia con los artículos que explotan la variación actitudinal de los estudiantes para explicar las calificaciones. Finalmente, algunos autores han analizado brechas originadas en la discriminación ejercida por parte de los profesores (Price, 2010; Cornwell et al., 2013; Kiss, 2013; Sprietsma, 2013; Terrier, 2015; Terrier, 2016). Esta documentación podría ser relevante para hipotetizar sobre posibles mecanismos discriminatorios que expliquen la brecha de género, a la vez que serviría de referencia para sugerir intervenciones que busquen disminuirla. Este aspecto se analiza cuando se consideran las características de los docentes.

El presente trabajo se ordenará de la siguiente forma: en la sección 2 se contextualizará al sistema educacional chileno y se describirán los datos utilizados; en la sección 3 se abarcarán los distintos análisis empíricos y en la sección 4 se abordarán efectos heterogéneos posibles, caracterizados por cualidades de los docentes, en las estimaciones. La última sección concluye.

## 2 Contexto escolar en Chile y datos

### 2.1 Contexto escolar y marco analítico

El grupo de observación corresponde a toda la generación de estudiantes regulares pertenecientes a 8<sup>vo</sup> básico el año 2015<sup>3</sup>. Para dicha generación se obtuvieron datos oficiales otorgados por el Ministerio de Educación de Chile (MINEDUC), los cuales se pueden resumir en dos grandes grupos: el primero, correspondiente a las calificaciones oficiales otorgadas a fin de año por los profesores; y el segundo, correspondiente a los puntajes de una evaluación realizada por el MINEDUC que mide aptitudes curriculares.

---

<sup>3</sup>Cerca de 2.000 estudiantes no eran considerados regulares, pues pertenecían a unidades escolares ajenas al sistema regular.

Las calificaciones escolares son otorgadas exclusivamente por los profesores de las asignaturas. Los profesores suelen estar durante todo el año académico con el mismo curso. La escala de calificación que utiliza el sistema chileno varía entre 1 y 7, considerándose el 4 como calificación mínima de aprobación y siendo posible la utilización de hasta una cifra decimal. Los profesores pueden utilizar variadas evaluaciones durante el año escolar y distintos instrumentos de evaluación para determinar la calificación.

La evaluación realizada por el MINEDUC corresponde al SIMCE (acrónimo de Sistema de Medición de la Calidad de la Educación), el cual es un instrumento de evaluación nacional que caracteriza, con puntajes, los niveles de aprendizaje de los estudiantes. El SIMCE, como prueba estandarizada, evalúa principalmente las asignaturas de Matemáticas y Lenguaje<sup>4</sup>, pero solo evalúa las aptitudes de lectura en esta última. Cada asignatura es evaluada en pruebas independientes, las cuales son aplicadas al mismo tiempo a nivel nacional. Si bien el SIMCE no es obligatorio, las tasas de asistencia históricas son cercanas al 90 – 95% en promedio para cada asignatura (Ministerio de Educación de Chile, 2014). Para la generación de 8<sup>vo</sup> básico en 2015, también se cuentan con los resultados SIMCE de 6<sup>to</sup> básico en 2013 y de 4<sup>to</sup> básico en 2011.

Las pruebas SIMCE son generadas y calificadas por parte de la Agencia de Calidad de la Educación, dependiente del MINEDUC, y sus resultados son reportados al siguiente año lectivo de forma agregada a nivel de establecimiento (o sea, los estudiantes no saben sus puntajes individuales y el establecimiento sabe sus resultados agregados solo al siguiente año). El MINEDUC ha establecido históricamente al SIMCE como la herramienta principal para medir aptitudes cognitivas correspondientes con el currículum educacional oficial<sup>5</sup>. Luego, se esperarían altos niveles de correlación entre los puntajes del SIMCE y las calificaciones finales obtenidas para cada estudiante en cada asignatura.

---

<sup>4</sup>Algunos años y en algunos grados se evalúa SIMCE de Ciencias Naturales, Ciencias Sociales e Inglés.

<sup>5</sup>Debido a la evidencia que sugeriría instrumentos de evaluación que en sí discriminarían por género (por ejemplo, a través de lenguaje sexista o caracterización contextual más cercana a un género), el MINEDUC ha velado, desde 2008, por construir instrumentos de evaluación que no presenten esto (Ministerio de Educación de Chile, 2008).

Con los datos anteriores, este análisis se centrará en las calificaciones y puntajes SIMCE en matemáticas. La justificación para restringir el análisis a matemáticas se basa en lo planteado por Botelho, Madeira, & Rangel (2015): se espera que las calificaciones de matemáticas estén más alineadas con los niveles de logro del SIMCE, debido a que las aptitudes matemáticas son más fáciles de observar en formatos de tipo prueba. Si bien el SIMCE utiliza exclusivamente un modelo de medición basado en la Teoría de Respuestas al Ítem (IRT), dicho formato sería extensamente utilizado por los docentes en Chile, generando teóricamente una concordancia en las formas de medición de las aptitudes (Oteiza & Miranda, 1996; Parra, Palacios, & González, 2016)<sup>6</sup>. A pesar de lo anterior, se debe considerar que, explícitamente, parte de la calificación en matemáticas considera elementos básicos de comprensión lectora, por lo que será útil contar con las habilidades de comprensión lectora a la hora de explicar el resultado de una calificación en matemáticas<sup>7</sup>.

## 2.2 Datos

El universo de estudiantes con calificaciones están disponibles en la Tabla 1, Panel A. Del total de estudiantes, el 84.24% rindió las pruebas SIMCE de matemáticas y lectura, además de contar con datos válidos<sup>8</sup>. Como muestra el Panel B, dichos estudiantes son similares en términos del porcentaje de mujeres del Panel A, pero presentan mejores calificaciones<sup>9</sup>. Para facilitar la lectura de los resultados, se han estandarizado las calificaciones y los puntajes considerando a estos 202,808 estudiantes.

Con estos datos iniciales se construyó la Figura 1, que muestra la relación incondicional promedio entre la calificación de matemáticas y los puntajes SIMCE de matemáticas, ambos estandarizados. Dicha figura representa la primera irregularidad observada a partir de los

---

<sup>6</sup>En la asignatura de Lenguaje esto sería más difícil de argumentar, pues las pruebas estandarizadas como el SIMCE miden usualmente comprensión lectora, mientras que la asignatura escolar evalúa sobre estándares de comprensión lectora, escritura, expresión oral, ortografía y gramática.

<sup>7</sup>Para más información sobre comprensión lectora y su rol en matemáticas, se sugiere revisar Fuentes (1998) y Vilenius-Tuohimaa, Aunola, & Nurmi (2008).

<sup>8</sup>Errores administrativos generan que algunas evaluaciones SIMCE sean invalidadas.

<sup>9</sup>Dicha diferencia es estadísticamente significativa al 99%.



datos: para cada puntaje SIMCE, las mujeres presentan consistentemente mejores calificaciones. Esto se condice con el Panel C de la Tabla 1, en donde notamos que las mujeres presentan mejores calificaciones (equivalente a 0.14 desviaciones estándar mejor que el género opuesto), mientras que los hombres presentan mejores puntajes SIMCE (equivalente a 0.12 desviaciones estándar mejor que el género opuesto). Ambas diferencias son estadísticamente significativas a un nivel de confianza mayor al 99%.

Esta primera aproximación podría estar omitiendo características observables relevantes para el profesor a la hora de calificar, en especial respecto de las aptitudes no cognitivas como la motivación, responsabilidad y participación en clases. Por lo mismo, el presente trabajo ha considerado estos aspectos. La evaluación SIMCE no solo mide los aspectos curriculares y cognitivos, sino que incorpora una encuesta para estudiantes, padres y profesores. De la encuesta para estudiantes, se utilizaron preguntas sobre actitudes como *proxies* de las verdaderas aptitudes no cognitivas. Estas actitudes difieren por género, ya que las mujeres presentan mejores actitudes en promedio, lo cual podría explicar esta brecha inicial. En la sección 3 se analizará la incorporación de estos elementos.

De los cuestionarios realizados a los padres se extrajeron datos sociodemográficos de cada estudiante. Como esta documentación pretende hacer estimaciones robustas con el menor sesgo posible, también se utilizaron los datos de SIMCEs pasados y las respuestas a cuestionarios pasados. El Panel D de la Tabla 1 resume la muestra final a utilizar.

Debido a que ninguno de los cuestionarios es obligatorio de responder, la muestra a utilizar se ve reducida fuertemente, pues las tasas de respuesta no tienen valores muy altos. La gran cantidad de estudiantes que quedan fuera de la muestra final podría no ser aleatoria, generando sesgos de selección en los resultados de las estimaciones. Para comprender las consecuencias de contar con una muestra que representa solo el 46% de la muestra inicial, la Tabla 2 presenta un resumen de medias para las calificaciones y puntajes SIMCE<sup>10</sup>.

---

<sup>10</sup>Una situación similar se da en el trabajo de Paredes (2014), en donde un 49% de la muestra original se pierde por errores u omisiones similares.

Table 1: Estadística descriptiva de los datos

Dato	Media	Desv. Est.	Mín	Máx
<b>Panel A: con calificaciones</b>		<i>N = 240,740</i>		
Calificación	5.10	0.86	1	7
Mujeres	49.28%			
<b>Panel B: Panel A con SIMCE válido</b>		<i>N = 202,808</i>		
Calificación	5.16	0.85	1	7
Calificación estandarizada	0	1	-4.89	2.17
SIMCE matemáticas estandarizado	0	1	-2.54	2.65
Mujeres	49.30%			
<b>Panel C: Panel B analizado por género</b>				
Calificación estandarizada, mujeres	0.070	0.99	-4.18	2.17
Calificación estandarizada, hombres	-0.068	0.98	-4.89	2.17
SIMCE estandarizado, mujeres	-0.061	1.00	-2.54	2.65
SIMCE estandarizado, hombres	0.059	0.99	-2.53	2.65
<b>Panel D: muestra final</b>		<i>N = 110,871</i>		
Calificación estandarizada, mujeres	0.171	0.99	-3.95	2.17
Calificación estandarizada, hombres	0.053	0.98	-4.69	2.17
SIMCE estandarizado, mujeres	0.050	0.99	-2.51	2.65
SIMCE estandarizado, hombres	0.210	0.99	-2.51	2.65
Mujeres	51.95%			

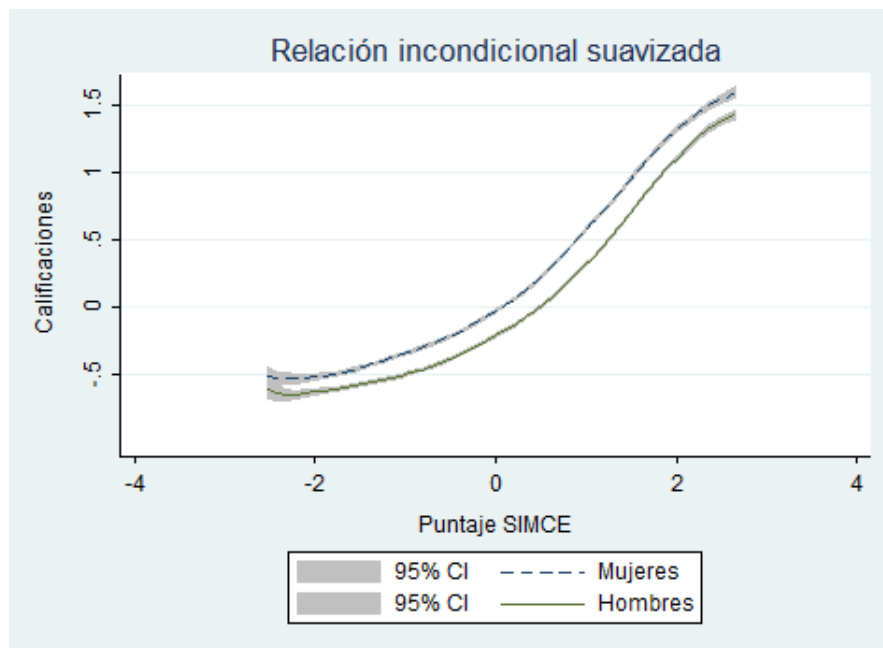


Figure 1: La relación incondicional suavizada considera los puntajes SIMCE de matemáticas estandarizados y calificaciones en matemáticas estandarizadas.

El Panel A considera como universo a todos los estudiantes con calificaciones (240,740). El subpanel A1 muestra las calificaciones promedio (y en paréntesis las desviaciones estándar) para dos grupos: el primer grupo, en la columna 1, son los estudiantes que contaban con calificaciones y SIMCE válido; el segundo grupo, en la columna 2, son los que no contaban con los datos SIMCE válidos. El subpanel A2 muestra lo mismo, pero considerando como criterio que los estudiantes pertenezcan a la muestra final. El Panel B muestra puntajes promedios SIMCE, considerando como universo al total de estudiantes con calificaciones y puntajes SIMCE válidos (202,808). El subpanel B1 utiliza un criterio de selección para los estudiantes que contestaron y entregaron cuestionarios concurrentes válidos, mientras que el subpanel B2 vuelve a reflejar la muestra final.

Como los estudiantes de la columna 2 tienen consistentemente menores promedios en calificaciones y SIMCE, es posible sugerir la existencia de alguna característica inobservable que haga que estudiantes con menores desempeños no lleguen a ser parte de la muestra final, en

especial por no responder a los cuestionarios<sup>11</sup>. Comparando los resultados de la columna 1, entre los subpaneles A1 y A2, los hombres mejoraron su media de calificación en un mayor porcentaje que las mujeres y fueron los que relativamente perdieron más observaciones. Esto podría indicar que la muestra final cuenta con una sub-representación relativa de hombres con peores calificaciones. Esta misma conclusión se obtiene de analizar la columna 1 de los subpaneles B1 y B2 (con puntajes SIMCE).

Una deducción inicial indica que, a la hora de evidenciar la brecha de género, se podría estar sub-estimando la diferencia entre hombres y mujeres, pues el grupo de hombres de la muestra final mejoró relativamente sus calificaciones al compararlas con las del género opuesto. Sin embargo, los datos de la Tabla 1 (Paneles C y D) muestran que la brecha de género se mantiene relativamente constante para las calificaciones y para el SIMCE, otorgando una ventaja para las especificaciones que buscan estimar la relación del género con las calificaciones, pero dejando las otras características explicativas constantes.

Finalmente, otra preocupación se originaría si los estudiantes con bajas calificaciones que no contestaron los cuestionarios pertenecían a un grupo con bajas características actitudinales. Esto dificultaría explotar la verdadera variación de niveles de actitud en la clase, generando estimaciones sesgadas. Dicho sesgo debería mostrar que las estimaciones para la relación entre actitudes y calificaciones se da con menor fuerza que lo real. A pesar de dar cuenta del mecanismo anterior para la falta de datos, y considerando que existe una variación suficiente en las variables explicativas que sí están disponibles, la utilización de los datos de la muestra final es útil para evidenciar el fenómeno de la brecha.

---

<sup>11</sup>La respuesta a los cuestionarios se da no solo en 8<sup>vo</sup> básico, sino también en 6<sup>to</sup> y 4<sup>to</sup> básico. Por otro lado, también están los cuestionarios a padres, los cuales a veces tampoco fueron respondidos válidamente por los apoderados.

Table 2: Estadística descriptiva de los datos medidos por disponibilidad de datos. Se presentan medias, desviaciones estándar en paréntesis y cantidad de observaciones.

Panel y criterio	Columna 1: Cumplen con criterio		Columna 2: No cumplen con criterio	
	Hombre	Mujer	Hombre	Mujer
<b>Panel A: Calificaciones</b>				
<b>Universo: Con calificaciones</b>				
Panel A1:	5.10 (0.84)	5.21 (0.86)	4.78 (0.83)	4.87 (0.86)
Calificación y SIMCE válido	$N = 102,824$	$N = 99,984$	$N = 19,109$	$N = 18,823$
Panel A2:	5.20 (0.84)	5.30 (0.85)	4.93 (0.83)	5.03 (0.86)
Muestra final	$N = 53,275$	$N = 57,596$	$N = 68,597$	$N = 61,272$
<b>Panel B: Puntajes SIMCE</b>				
<b>Universo: Con calificaciones y SIMCE</b>				
Panel B1:	268.6 (49.2)	262.3 (49.3)	257.7 (47.8)	251.2 (48.4)
Entregaron cuestionario 8 <sup>vo</sup>	$N = 85,638$	$N = 86,055$	$N = 17,093$	$N = 14,022$
Panel B2:	274.1 (48.8)	266.2 (48.9)	258.8 (48.3)	253.2 (48.9)
Muestra final	$N = 53,275$	$N = 57,596$	$N = 48,570$	$N = 43,367$

### 3 Análisis empírico

#### 3.1 Estimaciones iniciales

El análisis estadístico del trabajo considera dos grandes fuentes de observación por parte del profesor antes de calificar al estudiante: las habilidades cognitivas y las no cognitivas. Las primeras se relacionan con el desempeño en base a los objetivos curriculares y las segundas con aspectos actitudinales como motivación, responsabilidad y participación en clases. Este marco de referencia permite utilizar los datos SIMCE, tanto en puntajes como en respuestas de los cuestionarios sobre actitudes. Este marco de análisis solo permite documentar un hecho de forma robusta, pero no permite concluir efectos causales de alguna de las características. Esto se debe a que las especificaciones pueden estar omitiendo otros factores y dinámicas relevantes propios de datos no experimentales. Por ejemplo, no se cuenta con datos sobre los niveles particulares de esfuerzo que los estudiantes han ejercido en su vida escolar y en el hogar, ni datos específicos de la preparación de los estudiantes a la hora de rendir las evaluaciones. Estos atributos probablemente han afectado algunas de las características descritas para las especificaciones y no permiten justificar la existencia de efectos causales.

Para realizar el análisis, se considera el escenario inicial en donde las calificaciones no son explicadas por ninguna característica actitudinal o *proxy* de ello. Este escenario inicial captura la brecha por género incondicional existente. La Tabla 3 muestra regresiones de mínimos cuadrados ordinarios (MCO) en donde la variable dependiente son las calificaciones de matemáticas estandarizadas. Todas las regresiones presentadas son a nivel de estudiante.

Table 3: Variable dependiente: calificaciones de matemáticas estandarizadas.

	(1)	(2)	(3)
Mujer	0.119*** (0.006)	0.122*** (0.007)	0.123*** (0.007)
Efectos fijos	No	Escuela	Curso
$N\#$	110,871	110,871	110,871
$R^2$	0.004	0.004	0.004
$F$	395	325	337

NOTA: Columna (2) y columna (3) son regresiones con errores estándar agrupados por establecimiento y curso. Coeficiente de determinación: las especificaciones con efectos fijos consideran un coeficiente de determinación dentro del clúster (*within*). Niveles de significación: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

La primera columna considera la regresión más simple para la muestra final, donde el único regresor es una variable binaria que toma el valor 1 si el estudiante es mujer y 0 si es hombre. El valor del estimador se corresponde con la diferencia presentada en la Tabla 1, Panel D; o sea, las mujeres presentan 0.119 desviaciones estándar más que los hombres en las calificaciones en términos incondicionales. Dicha estimación se ve alterada levemente con la incorporación de efectos fijos por establecimiento, mostrado en la segunda columna: aumenta en un 3% respecto de la primera columna. Como tampoco se ve mayormente afectado en la columna 3, en donde se incorporan efectos fijos por curso (y no por establecimiento), se podría argumentar que la brecha de género es un fenómeno constante dentro de los establecimientos y cursos. Las tres especificaciones presentan coeficientes de determinación iguales. Estos coeficientes sugieren, debido a sus bajos valores, la necesidad de incorporar más variables independientes en la regresión para aumentar la explicación en la variación de las calificaciones.

La utilización de efectos fijos por curso se condice mejor con el contexto explicado previamente: dado que la asignatura suele ser enseñada por el mismo docente en un año, se puede controlar por el efecto promedio del profesor sobre el curso, así como también sobre su práctica como calificador. Por otro lado, resultará más realista considerar que algunos establecimientos realizan *tracking* académico, en donde los cursos de un mismo nivel son divididos según desempeños académicos (medidos en las calificaciones), generando heterogeneidades en, por ejemplo, los tamaños de los cursos y sus rendimientos (McEwan, 2003). Así, las prácticas docentes, incluso para un mismo docente, podrían ser distintas en distintos cursos (Woessmann & West, 2006; Himmler & Schwager, 2013; Kiss, 2013). Además de la práctica docente, también es importante notar que los cursos podrían reflejar otras condiciones y caracterizaciones propias de los estudiantes y sus familias (por ejemplo, ingresos y niveles educacionales de los padres). Por lo tanto, los efectos fijos permiten capturar todos estos aspectos inobservables que afecten al curso.

La Tabla 4 incorpora los puntajes SIMCE de matemáticas y de lectura, ambos estandarizados. También se incorpora la calificación de matemáticas estandarizada obtenida 2 años antes (en 6<sup>to</sup> básico), con la intención de capturar habilidades globales en matemáticas, para cada individuo, que sean invariantes en el tiempo y que sean percibidas como relevantes para cualquier docente. La elección de utilizar la calificación de dos años antes cumple un rol práctico para su uso en las estimaciones: ya que en el sistema educacional chileno es muy usual que un profesor de matemáticas de 8<sup>vo</sup> básico haya realizado clases al mismo curso en el año anterior, utilizar las calificaciones de dos años antes permite explotar la variación de docentes que hayan hecho clases en un mismo curso.



Table 4: Variable dependiente: calificaciones de matemáticas estandarizadas.

	(1)	(2)	(3)	(4)
Mujer	0.142*** (0.005)	0.147*** (0.005)	0.150*** (0.005)	0.150*** (0.004)
SIMCE mat	0.394*** (0.004)	0.397*** (0.005)	0.389*** (0.006)	0.407*** (0.004)
SIMCE mat <sup>2</sup>		0.145*** (0.004)	0.143*** (0.005)	0.144*** (0.004)
SIMCE mat <sup>3</sup>		0.001 (0.002)	0.002 (0.002)	0.000 (0.001)
SIMCE mat <sup>4</sup>		-0.014*** (0.001)	-0.014*** (0.001)	-0.013*** (0.001)
SIMCE lect	0.076*** (0.003)	0.076*** (0.004)	0.073*** (0.003)	0.077*** (0.002)
Calificación 6	0.460*** (0.003)	0.438*** (0.003)	0.439*** (0.004)	0.426*** (0.003)
Efectos fijos	Curso	Curso	Curso	Curso
Controles sociodemográficos	No	No	Sí	Sí
N#	110,871	110,871	110,871	175,063
R <sup>2</sup>	0.573	0.585	0.589	0.574
F	18,386	12,371	7,299	10,005

NOTA: Las especificaciones consideran errores estándar agrupados por curso. Coeficiente de determinación: las especificaciones con efectos fijos consideran un coeficiente de determinación dentro del clúster (*within*). Niveles de significación: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

La primera columna muestra una regresión lineal considerando estas variables. La estimación para el género aumenta en 15% respecto de la tercera especificación de la Tabla 3. Por otro lado, la inclusión de los puntajes SIMCE y calificaciones del pasado aumentaron sustancialmente los valores del coeficiente de determinación. En otras palabras, esta especificación se ajusta bastante mejor que las presentadas en la Tabla 3. Esto resulta intuitivo si consideramos que una parte importante de la calificación debiese ser explicada por las aptitudes cognitivas en matemáticas, medidas por los puntajes SIMCE y las calificaciones del pasado. El marco analítico sugiere que omitir el puntaje SIMCE es omitir una variable explicativa relevante.

Los datos previos analizados muestran que las mujeres tienen peor desempeño en el SIMCE, pero mejor desempeño en las calificaciones. Por lo tanto, dado que el puntaje SIMCE correlaciona positivamente con las calificaciones, omitir dicho puntaje genera un sesgo negativo en la estimación para la variable binaria de género en la Tabla 3 y explica económicamente el aumento del valor de dicha estimación en la Tabla 4.

En la columna 2 el SIMCE de matemáticas se incorpora como un polinomio de cuarto orden, siguiendo lo sugerido por Botelho et al. (2015), en donde incorporar polinomios permite capturar relaciones no lineales entre las variables. La estimación para la variable binaria de género aumenta respecto de las especificaciones de la columna 3 de la Tabla 3 en cerca de un 20%; respecto de la columna 1 de la Tabla 4, la estimación considera que la brecha de género es 0.005 desviaciones estándar mayor. El coeficiente de determinación también aumenta, aunque se mantiene en valores cercanos a los de la columna 1, indicando que la incorporación de los polinomios ayuda marginalmente en el poder explicativo de la especificación.

La columna 3 incorpora características sociodemográficas del estudiante obtenidas del cuestionario de padres y de datos adjuntados en el SIMCE (ver Apéndice A). Las estimaciones puntuales de estos regresores no son expuestas en la tabla debido a que, a pesar de ser explicativas, las variables sociodemográficas tienen bajo impacto en los estimadores para la variable binaria de género, sugiriendo que los aspectos que correlacionan más fuertemente ya están capturados por las otras variables de la regresión, en especial por los efectos fijos. Los coeficientes de determinación para estas especificaciones muestran que, aproximadamente, la proporción de la varianza de las calificaciones es explicada en un 40% por las variables incluidas, similar a los presentados en las columnas 1 y 2. Finalmente la columna 4 replica la columna 3 para la mayor muestra disponible para estos datos<sup>12</sup>. Los resultados muestran una similitud en las estimaciones de ambas columnas, sugiriendo que, si bien la muestra final podría contener estudiantes con mejores desempeños, la brecha de género sería representativa de un fenómeno existente para una muestra mayor.

---

<sup>12</sup>La mayor muestra disponible para estos datos considera 175,063 estudiantes, los cuales tienen todas las variables explicativas presentadas en la Tabla 4. De estos, 64,192 estudiantes no cuentan con el resto de los datos para pertenecer a la muestra final utilizada.

A grandes rasgos, se puede proponer que, aun incluyendo variables explicativas relevantes, la brecha de género en las calificaciones sigue estando presente e incluso muestra un mayor peso cuando se incluyen características de aptitudes cognitivas. Esto induce a incorporar otras fuentes de información relevantes que podría utilizar un docente para calificar.

### 3.2 Índices de actitud

Para la incorporación de aspectos actitudinales, se utilizaron las encuestas SIMCE respondidas por los estudiantes como aproximación de las verdaderas aptitudes no cognitivas. Dichas encuestas incluían preguntas variadas sobre responsabilidad, participación y motivación académica. El detalle de las preguntas utilizadas se encuentra en el Apéndice B. En el mismo apéndice, Tabla 9, se hace una descripción estadística de los valores promedio para mujeres y hombres. Los resultados, agrupados por curso para corresponderse con las especificaciones utilizadas, muestran que la mayoría de las medias presentan una diferencia por género estadísticamente significativa. Para gran parte de las preguntas, los resultados arrojan mayores medias para las mujeres. En el Apéndice B también se presentan las comparaciones gráficas para estas preguntas.

Para facilitar la interpretación de la relevancia de estos aspectos sobre la calificación, se propone utilizar índices agregados. Para cada nivel se construyeron inicialmente 2 índices. Ambos fueron construidos haciendo un promedio simple de las preguntas relevantes y luego se estandarizaron para toda la muestra final de 8<sup>vo</sup>:

1. Índice de responsabilidad y motivación en clases: para 8<sup>vo</sup> básico considera las preguntas de la 1) a la 4); para 6<sup>to</sup> básico considera las preguntas a) y b); para 4<sup>to</sup> básico considera las preguntas i) y ii). Todas estas preguntas están en una escala con valores entre 1 y 4.
2. Índice de responsabilidad y motivación extra-clases: para 8<sup>vo</sup> básico considera las preguntas 5) y 6); para 6<sup>to</sup> básico considera la pregunta c); para 4<sup>to</sup> básico considera la pregunta iii). Todas estas preguntas están en una escala binaria con valores 0 o 1.

El índice de responsabilidad y motivación en clases captura las preguntas relacionadas con aspectos más fácilmente observables por el profesor, en especial porque estas preguntas expresan actitudes evidenciadas principalmente en clases. El segundo índice captura las preguntas relacionadas con aspectos relativos a actividades fuera de la sala de clases y, por lo tanto, menos observados por el profesor. La distinción de estos dos índices permite mostrar la existencia de dos contextos distintos desde donde un profesor puede extraer información acerca de los aspectos actitudinales del estudiante.

El Apéndice C presenta los resultados gráficos de la estandarización de estos índices y una tabla con las correlaciones entre los dos tipos de índices. Los resultados muestran una correlación positiva (y estadísticamente significativa al 99% de confianza), lo que podría plantear como plausible una estabilidad actitudinal de los estudiantes en el tiempo. Si bien las correlaciones varían entre 0.09 y 0.36, aquellas con mayor valor absoluto son respecto de los índices en clases, en especial aquellos que relacionan un año con su periodo anterior (o sea, la correlación entre octavo y sexto y entre sexto y cuarto). En contraste, los índices extra-clases presentan correlaciones más bajas (entre 0.09 y 0.19). Desde la perspectiva teórica, los datos mostrarían que los estudiantes presentan una mayor estabilidad en el tiempo de sus actitudes dentro de clases, en contraste con las actitudes del segundo índice.

Una primera aproximación empírica resulta de establecer relaciones entre ambos índices y el género del estudiante. La Tabla 5 muestra regresiones de mínimos cuadrados ordinarios a nivel de estudiante, considerando como variable dependiente ambos índices para 8<sup>vo</sup> básico. Las columnas 1 y 3 muestran regresiones que consideran solo el género como variable explicativa, aparte de los efectos fijos por curso. Las columnas 2 y 4 agregan las otras variables del modelo. Como notamos, para ambos índices, al agregar las otras variables aparte del género, el valor absoluto de la estimación que acompaña a la variable binaria de género se ve reducido.

Table 5: Variable dependiente: índices en clases y extra-clases construidos.

	Índice clases		Índice extra-clases	
	(1)	(2)	(3)	(4)
Mujer	0.233*** (0.007)	0.197*** (0.007)	-0.039*** (0.007)	-0.022*** (0.007)
SIMCE mat		-0.022*** (0.008)		0.015** (0.008)
SIMCE mat <sup>2</sup>		-0.008 (0.007)		0.017** (0.007)
SIMCE mat <sup>3</sup>		0.005** (0.003)		0.003 (0.002)
SIMCE mat <sup>4</sup>		-0.002 (0.002)		-0.002 (0.001)
SIMCE lect		0.066*** (0.005)		-0.074*** (0.005)
Calificación 6		0.198*** (0.004)		0.030*** (0.004)
Efectos fijos	Curso	Curso	Curso	Curso
Controles sociodemográficos	No	Sí	No	Sí
N#	110,871	110,871	110,871	110,871
R <sup>2</sup>	0.013	0.059	0.000	0.004
F	1,161	490	33	34

NOTA: Las especificaciones consideran errores estándar agrupados por curso. Coeficiente de determinación: las especificaciones con efectos fijos consideran un coeficiente de determinación dentro del clúster (*within*). Niveles de significación: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

En específico, los resultados de las primeras 2 columnas muestran una correlación positiva entre la variable binaria de género y el índice construido para capturar actitudes en clases. Esto se traduce en que las mujeres tienen, en promedio, cerca de 0,2 desviaciones estándar más de niveles actitudinales positivos en clases. Por otro lado, el índice extra-clases muestra una relación negativa.

En las cuatro especificaciones, las estimaciones para la variable de género resultan signi-

ficativas estadísticamente. Considerando los coeficientes de determinación, también notamos que las especificaciones relativas al índice de clases presentan mayores valores (también reflejado en los valores del estadístico para la prueba F de significancia global). Esto permite argumentar que la modelación econométrica explica mejor el índice en clases que el extra-clases. Se discutirá la relevancia teórica y econométrica en la siguiente sub-sección.

### 3.3 Estimaciones con índices actitudinales

Con la construcción de los índices, se pueden ampliar las especificaciones presentadas en la Tabla 4. La columna 1 de la Tabla 6 muestra la misma columna 3 de la Tabla 4, pero incorporando el índice en clases para 8<sup>vo</sup>. La estimación presenta todavía valores positivos y estadísticamente significativos para la variable binaria de género, pero esta se ve reducida en cerca de un 20%. Además, el coeficiente de determinación aumenta, indicando que la incorporación del índice en clases explica mejor la variación en las calificaciones.

El resultado anterior indicaría que la omisión del índice que captura las actitudes en clases podría sesgar hacia arriba las estimaciones de la variable de género en la Tabla 4, considerando que las mujeres tienen mejores índices en clases y que dichos índices debiesen influir positivamente en la calificación. Los autores Cornwell et al. (2013), Kiss (2013), Golsteyn & Schils (2014) y Terrier (2015) han justificado que este sesgo se originaría en la omisión de la variable actitudinal (o no cognitiva), la cual sería relevante para el docente que establece las calificaciones. Esto es concordante con el marco analítico presentado.

La estimación para el índice en clases presenta un valor positivo, estadísticamente significativo y económicamente importante. Este coeficiente se puede interpretar tal que una desviación estándar más de índice en clases correlaciona con cerca de 0,15 desviaciones estándar más de calificaciones en matemáticas. Esto es concordante con un marco analítico en que tanto las características de aptitudes cognitivas (medidas por puntajes SIMCE y calificaciones del pasado) como las no-cognitivas (en el índice en clases) explican las calificaciones.

Table 6: Variable dependiente: calificaciones de matemáticas estandarizadas.

	(1)	(2)	(3)
Mujer	0.121*** (0.005)	0.151*** (0.005)	0.121*** (0.005)
SIMCE mat	0.393*** (0.006)	0.389*** (0.006)	0.393*** (0.006)
SIMCE mat <sup>2</sup>	0.145*** (0.005)	0.143*** (0.005)	0.145*** (0.005)
SIMCE mat <sup>3</sup>	0.001 (0.002)	0.002 (0.002)	0.001 (0.002)
SIMCE mat <sup>4</sup>	-0.013*** (0.001)	-0.014*** (0.001)	-0.013*** (0.001)
SIMCE lect	0.064*** (0.003)	0.075*** (0.003)	0.064*** (0.003)
Calificación 6	0.410*** (0.003)	0.438*** (0.003)	0.410*** (0.004)
Índice en clases	0.146*** (0.002)		0.146*** (0.002)
Índice extra-clases		0.023*** (0.002)	0.003 (0.002)
Efectos fijos	Curso	Curso	Curso
Controles sociodemográficos	Sí	Sí	Sí
N#	110,871	110,871	110,871
R <sup>2</sup>	0.612	0.590	0.612
F	7,252	6,710	6,696

NOTA: Coeficiente de determinación: las especificaciones con efectos fijos consideran un coeficiente de determinación dentro del clúster (*within*).

Niveles de significación: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

En la segunda columna, la estimación para la variable binaria de género no se ve mayormente afectada cuando incluimos como control solo el índice extra-clase, en comparación a la columna 3 de la Tabla 4. El coeficiente de determinación para esta especificación es prácticamente el mismo de la columna 3 de la Tabla 4, indicando que este índice no incorpora información muy relevante para explicar la variable dependiente. Si bien la estimación de la correlación de este índice con las calificaciones es estadísticamente significativa, es de baja

magnitud en comparación a su símil de la columna 1. Una explicación teórica, en coherencia con el marco analítico, radica en que el índice extra-clase responde a situaciones académicas que el docente no necesariamente observa en particular para cada estudiante, a diferencia del índice en clases.

La columna 3 incorpora los dos índices como variables independientes. Los resultados reflejan estimaciones prácticamente iguales a las de la columna 1. El índice en clases resulta relevante en términos de la magnitud y significancia estadística, pero el índice extra-clase se vuelve completamente irrelevante: se ve reducido en su magnitud (cerca de un 87% respecto de la columna 2) y se vuelve no significativo estadísticamente. Con estas estimaciones, y basándonos en el marco analítico en que el profesor observa más fuertemente señales dentro del aula escolar, las siguientes especificaciones no utilizarán índices extra-clase.

Considerando los resultados presentados hasta el momento, una preocupación surge respecto del insesgamiento de las estimaciones. Una fuente de sesgo podría originarse en que las calificaciones y los índices en clases presenten problemas de causalidad reversa. En otras palabras, es posible que los estudiantes hayan respondido que tienen mejores actitudes en clases como forma de justificar sus buenas calificaciones.

Por otro lado, el índice en clases solo captura aspectos relacionados con las preguntas seleccionadas para su construcción, haciendo que este índice sea una aproximación (*proxy*) de la verdadera variable actitudinal. Como se discutió anteriormente, omitir el aspecto actitudinal generaría un sesgo por variable omitida. Sin embargo, incorporar un *proxy* del verdadero aspecto actitudinal no necesariamente deja las estimaciones libres de sesgo debido a que no es posible afirmar que se cumplen los supuestos para ello<sup>13</sup>. A pesar de lo anterior, usar el índice en clases presenta dos ventajas: primero, las estimaciones resultantes de las especificaciones en la Tabla 6 son teórica y económicamente relevantes debido a que capturan la relevancia de la dimensión actitudinal; y segundo, se podría esperar que los sesgos originados por la incorporación del índice son menores en comparación a omitir las actitudes, como en las especificaciones de la Tabla 4.

---

<sup>13</sup>Los supuestos son los correspondientes a la “*plug-in solution to the omitted variables problem*” presentado por Wooldridge (2015).



Finalmente, debido a la naturaleza de los cuestionarios SIMCE, los índices podrían contar con errores de medición. En otras palabras, el auto-reporte de actitudes podría no representar la realidad de las actitudes en clases. En otros contextos se ha sugerido que los errores de medición no son clásicos, pues los estudiantes tienden a reportar consistentemente mejores actitudes (Kiss, 2013). Sin embargo, en los datos de este documento no se puede descartar que los errores de medición sean clásicos. La siguiente subsección realizará un refinamiento del índice de clases a través del uso de respuestas de los cuestionarios del pasado y discutirá sobre estos errores de medición.

### **3.4 Refinamiento del índice actitudinal en clases: uso de variables instrumentales**

Aprovechando los cuestionarios y respuestas del pasado, tras observar correlaciones entre los índices en clases para distintos momentos del tiempo (ver Apéndice C), una propuesta útil es utilizar dichos índices del pasado como variables instrumentales de los índices en clases concurrentes. De esta forma, se pueden mejorar las estimaciones al corregir los errores de medición asociados al modelo y la endogeneidad asociada a la causalidad reversa explicada anteriormente.

La Tabla 7 muestra, en la columna 1, el mismo modelo de regresión de mínimos cuadrados ordinarios (MCO) de la columna 1 en la Tabla 6, pero considerando los índices en clases de 6<sup>to</sup> y 4<sup>to</sup> básico en vez del concurrente. Los resultados muestran una relación estadísticamente significativa para cada índice y las calificaciones concurrentes. El valor del estimador para la variable binaria de género se asimila a las magnitudes presentadas en las estimaciones de la Tabla 4. Esta columna sirve como ejercicio práctico para ver la relevancia de los índices en la explicación de las calificaciones.

Table 7: Variable dependiente: calificaciones de matemáticas estandarizadas.

	(1)	(2)	(3)
Mujer	0.145*** (0.005)	0.118*** (0.005)	0.119*** (0.005)
SIMCE mat	0.397*** (0.006)	0.393*** (0.006)	0.393*** (0.006)
SIMCE mat <sup>2</sup>	0.145*** (0.005)	0.145*** (0.005)	0.145*** (0.005)
SIMCE mat <sup>3</sup>	0.002 (0.002)	0.001 (0.002)	0.001 (0.002)
SIMCE mat <sup>4</sup>	-0.014*** (0.001)	-0.013*** (0.001)	-0.013*** (0.001)
SIMCE lect	0.073*** (0.003)	0.063*** (0.003)	0.063*** (0.003)
Calificación 6	0.421*** (0.003)	0.407*** (0.004)	0.407*** (0.004)
Índice en clases VI		0.161*** (0.006)	0.159*** (0.007)
Índice en clases 6	0.047*** (0.002)		
Índice en clases 4	0.015*** (0.002)		
Efectos fijos	Curso	Curso	Curso
Controles sociodemográficos	Sí	Sí	Sí
N#	110,871	110,871	110,871
R <sup>2</sup>	0.593	0.612	0.612
F	6,284	7,023	7,014
F de primera etapa	-	4,830	9,020

NOTA: Regresiones con errores estándar agrupados por curso. La columna 1 estima utilizando una regresión MCO. La columna 2 estima utilizando los índices en clases de 6<sup>to</sup> y 4<sup>to</sup> básico como variables instrumentales. La columna 3 estima utilizando los índices en clases de 6<sup>to</sup> básico como variables instrumentales. Las primeras etapas se encuentran en el Apéndice D. Coeficiente de determinación: las especificaciones con efectos fijos consideran un coeficiente de determinación dentro del clúster (*within*). F de primera etapa: corresponde a la prueba de instrumentos excluidos. Niveles de significación: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

La columna 2 muestra una estimación similar a la columna 1 de la Tabla 6, pero instrumentando el índice de clases de 8<sup>vo</sup> básico con los índices en clases de 6<sup>to</sup> y 4<sup>to</sup> básico. La primera etapa de la estimación, disponible en el Apéndice D, muestra que los instrumentos son relevantes para la especificación. Al comparar esta columna con la columna 3 de la Tabla 4, la agregación del índice en clases instrumentado produce una reducción cercana al 21% de la estimación para la variable de género.

La magnitud de la estimación es similar a la obtenida en las especificaciones de la Tabla 6, columnas 1 y 3. Sin embargo, esta especificación con variables instrumentales muestra una mayor estimación, de aproximadamente 0.015 desviaciones estándar (equivalentes a un aumento del 10%), para el índice en clases, otorgándole una mayor relación entre las actitudes en clases y la calificación del estudiante. Este resultado también se puede adaptar a lo obtenido en la columna 3, la cual realiza la misma estimación que la columna 2, pero considerando como instrumento solo el índice en clases de 6<sup>to</sup> básico. Estos valores sugieren que el sesgo por error de medición y/o por causalidad reversa de la Tabla 6 no distorsiona mayormente las estimaciones de las especificaciones presentadas.

Debido a que ambas especificaciones presentan estimaciones similares, se puede argumentar que existe evidencia robusta sobre la presencia de una brecha de género en las calificaciones escolares de matemáticas. Además, el rol de las variables de actitud (medidas de forma aproximada con el índice) se muestra como relevante cuando se incorpora como variable explicativa en las especificaciones, obteniéndose una especificación que explica de mejor forma la variación en las calificaciones.

## 4 Efectos heterogéneos

A pesar de lo encontrado anteriormente, las brechas podrían diferir en contextos escolares distintos. Utilizando los mismos datos provistos por los cuestionarios SIMCE realizados a profesores, se consideran cuatro escenarios particulares, todos relativos a características de los docentes. La importancia de analizar las características de los profesores radica en ex-

aminar nuevas hipótesis que consideren potenciales mecanismos explicativos de la brecha, en especial aquellos relacionados con discriminación por sesgos previos (Hanna & Linden, 2012; Schneeweis & Zweimüller, 2012; Lavy & Sand, 2015).

En primera instancia, se tomará en cuenta si el profesor de matemáticas es de género femenino. En segunda instancia, se tomará en cuenta si el profesor de matemáticas es profesor jefe<sup>14</sup>. En tercera instancia, se tomará en cuenta si el profesor tiene estudios de post-grado de cualquier tipo<sup>15</sup>. Finalmente, se tomará en cuenta si el profesor fue también el profesor de matemáticas el año anterior para dicho curso. Estas características son recurrentes en la literatura que analiza brechas por género, origen étnico u origen socioeconómico. En el Apéndice E se muestra una tabla con los porcentajes de estudiantes de la muestra final que tienen profesores con dichas características.

Para analizar estos efectos, la Tabla 8 muestra las mismas estimaciones econométricas de la Tabla 7 (columna 2), pero incluyendo interacciones entre el género del estudiante y la nueva variable. Se han omitido de este trabajo las primeras etapas de las estimaciones. Cada columna corresponde a estimaciones con las interacciones mencionadas, respectivamente.

La primera columna muestra estimaciones para la interacción del género del estudiante con el género del profesor (donde el valor de 1 se corresponde con que ambas sean mujeres). Si bien la estimación no es estadísticamente significativa a nivel del 95% de confianza, sí lo es para el 90%. Esto indicaría que, en promedio, la brecha de género se vería extendida en 0.015 desviaciones estándar para cursos con profesoras mujeres de matemáticas. La estimación para la variable de género del estudiante se ve reducida en un 6%, pero sigue siendo positiva y estadísticamente significativa. Los datos no permiten establecer el mecanismo exacto que explique esta relación de la interacción con las calificaciones, pero evidencia sugerida

---

<sup>14</sup>El profesor jefe es un profesor que acompaña al curso durante el año, siendo un guía más directo en el proceso educativo de los estudiantes. El profesor jefe usualmente está encargado de ser el nexo entre el establecimiento y los apoderados, además de ser el que reporta sobre el desempeño y comportamiento de cada estudiante de su curso a los directivos escolares.

<sup>15</sup>Dentro de las posibilidades de estudios de post-grado están los diplomados, los estudios de magíster y los doctorados.

por Paredes (2014), en un contexto similar, indicaría que las estudiantes se esforzarían más cuando la profesora del curso es mujer, por efecto de tener un “modelo a seguir”.

Table 8: Variable dependiente: calificaciones de matemáticas estandarizadas.

	(1)	(2)	(3)	(4)
<i>Interacción:</i>	Profesora mujer	Profesor jefe	Con post-grado	Año anterior
Mujer	0.111*** (0.007)	0.121*** (0.005)	0.120*** (0.006)	0.123*** (0.007)
Interacción de interés	0.015* (0.009)	-0.011 (0.009)	-0.004 (0.009)	-0.014* (0.009)
SIMCE mat	0.393*** (0.006)	0.393*** (0.006)	0.393*** (0.006)	0.393*** (0.006)
SIMCE mat <sup>2</sup>	0.145*** (0.005)	0.145*** (0.005)	0.145*** (0.005)	0.145*** (0.005)
SIMCE mat <sup>3</sup>	0.001 (0.002)	0.001 (0.002)	0.001 (0.002)	0.001 (0.002)
SIMCE mat <sup>4</sup>	-0.013*** (0.001)	-0.013*** (0.001)	-0.013*** (0.001)	-0.013*** (0.001)
SIMCE lect	0.063*** (0.003)	0.063*** (0.003)	0.063*** (0.003)	0.062*** (0.003)
Calificación 6	0.407*** (0.004)	0.407*** (0.004)	0.407*** (0.004)	0.406*** (0.004)
Índice en clases VI	0.161*** (0.006)	0.161*** (0.006)	0.161*** (0.006)	0.161*** (0.006)
Efectos fijos	Curso	Curso	Curso	Curso
Controles sociodemográficos	Sí	Sí	Sí	Sí
N#	110,871	110,871	110,871	110,871
R <sup>2</sup>	0.612	0.612	0.612	0.612
F	6,484	6,485	6,488	6,490
F de primera etapa	4,830	4,830	4,831	4,830

NOTA: Regresiones con errores estándar agrupados por curso. Todas las columnas estimna utilizando los índices en clases de 6<sup>to</sup> y 4<sup>to</sup> básico como variables instrumentales. Las primeras etapas no se presentan en el trabajo. Coeficiente de determinación: las especificaciones con efectos fijos consideran un coeficiente de determinación dentro del clúster (*within*). F de primera etapa: corresponde a la prueba de instrumentos excluidos. Niveles de significación: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Las columnas 2 y 3 presentan estimaciones que no son estadísticamente significativas. Considerando que las estimaciones para todas las otras variables independientes se mantienen relativamente estables al compararlas con la columna 2 de la Tabla 7, se podría argumentar que estas interacciones no tendrían un rol estadísticamente explicativo para la brecha de género documentada.

La columna 4 muestra estimaciones para la interacción del género del estudiante con que el docente haya sido profesor de matemáticas el año anterior (donde el valor de 1 se corresponde con que el estudiante sea mujer y el profesor haya estado el año anterior como profesor de matemáticas). Al igual que la columna 1, la estimación es estadísticamente significativa al 90% de confianza. En este caso, la interacción puede interpretarse como que la brecha de género disminuye, en promedio, 0.014 desviaciones estándar cuando el docente ha hecho clases durante dos (o más) años al mismo curso. Si bien este resultado no permite establecer un mecanismo explícito detrás de las estimaciones, evidencia estudiada por Lavy (2008) y Botelho et al. (2015), en contextos similares, muestra que los docentes podrían traer sesgos previos a la hora de calificar, pero que estos van disminuyendo a medida que la exposición al curso se extiende. Al comparar la estimación de la variable de género con su símil de la Tabla 7 (columna 2), notamos que esta aumenta un 4%, lo cual es concordante con que la estimación para la interacción reporte un signo negativo. Por último, si bien la estimación de la columna 2 no resulta estadísticamente significativa, sí presenta el signo esperado bajo esta misma argumentación: los profesores jefe debiesen estar más expuestos a los alumnos y eso disminuiría la brecha de género.

De forma general, los resultados sugieren que solo las estimaciones de las interacciones en las columnas 1 y 4 son estadísticamente explicativas para niveles de confianza al 90%. En las cuatro especificaciones se mantiene una estimación de la variable binaria de género positiva y estadísticamente significativa. Como los valores de las estimaciones para las otras variables explicativas son muy similares a los obtenidos en la Tabla 7, existiría evidencia de que las columnas 1 y 4 representan interacciones de importante análisis económico.

## 5 Conclusiones

El presente trabajo realiza un análisis exhaustivo acerca de la brecha de género en las calificaciones escolares de matemáticas. Utilizando bases de datos oficiales del MINEDUC, y caracterizando especificaciones econométricas adecuadas al marco analítico, se ha podido documentar de forma robusta que existe una brecha de género, la cual muestra que las mujeres tienen mejores calificaciones en matemáticas que sus pares hombres. La diferencia por género es, aproximadamente, equivalente a 0,12 desviaciones estándar en la especificación preferida. Dicha brecha no tiene explicación en variables observadas como las características de habilidades cognitivas y actitudinales, y representa cerca del 30% de la estimación obtenida tanto para el SIMCE como para la calificación del pasado en la asignatura. La brecha también es identificada al hacer un análisis de efectos heterogéneos en base a características de los docentes, reforzando la idea de que la brecha existe en prácticamente todos los establecimientos y en todos los contextos explicados.

Si bien la documentación de la brecha es uno de los aportes de este trabajo, otro aporte radica en la incorporación de las habilidades no-cognitivas o actitudinales extraídas desde los cuestionarios SIMCE. Dicha información contribuye a la literatura que ha evidenciado y explicado la importancia de las actitudes en clases en pos del desempeño de los estudiantes. Como mostraron los resultados de este trabajo, las actitudes, medidas en una variable del tipo *proxy*, resultan económica y estadísticamente significativas para explicar parte de las calificaciones obtenidas por los estudiantes. La utilización de datos actitudinales del pasado como variables instrumentales para disminuir sesgos también aporta a mejorar las estimaciones de estas aptitudes y muestran que la brecha es un fenómeno consistente entre las distintas especificaciones.

Desde una perspectiva de políticas públicas, esta documentación resulta relevante para entender cómo es el proceso de calificación por parte de los profesores. Los resultados presentados en la sección de efectos heterogéneos consideran características de los docentes interactuadas con la variable de género. En estas especificaciones, se concluye que cuando la profesora es mujer la brecha de género aumenta y cuando el profesor lleva por lo menos dos años realizando clases al curso, la brecha disminuye. Estos resultados son estadísticamente significativos y son consistentes con evidencia analizada por otros autores en situaciones similares. Frente a ello,

podrían sugerirse intervenciones o recolección de datos para ahondar en los mecanismos y variables que podrían explicar esta brecha de género. Las aproximaciones de algunos autores, como Lavy (2008), considera que los mecanismos a estudiar son relevantes para afirmar si la discriminación es a favor de las mujeres (en donde, por ejemplo, los docentes tienen creencias de que las mujeres podrían tener mejores desempeños) o en contra de los hombres (en donde, por ejemplo, los docentes tratan de compensar una discriminación percibida en contra de las mujeres, disminuyendo las calificaciones de los hombres). Con lo anterior, resultaría interesante que los actores públicos pudieran hacer experimentos en donde una misma evaluación sea calificada tanto por el docente como por un corrector externo, para focalizar el estudio de posibles discriminaciones por género; o explotar la variación de género entre hermanos que tienen al mismo docente, para analizar determinantes de una posible brecha de género, tal como es presentado por Autor, Figlio, Karbownik, Roth, & Wasserman (2016).

Sumado a lo anterior, y considerando la estabilidad de la brecha por género, resulta importante cuestionar al sistema de calificación, sobre todo desde la perspectiva del estudiante: algunos podrían estar recibiendo señales erróneas acerca de sus competencias, así como otros podrían estar ejerciendo esfuerzos sub-óptimos tras recibir información incompleta acerca de sus aptitudes (Kiss, 2013). Esto podría aportar para entender el puzzle empírico que presentan Hinnerich, Höglin, & Johannesson (2011): a pesar de que las mujeres tienden a tener mejores calificaciones al finalizar la educación secundaria, ellas reciben menores salarios al entrar al mercado laboral. Las consecuencias de la brecha podrían impactar en variados ámbitos del futuro de los estudiantes, por lo que ahondar en su estudio resulta útil para diseñar políticas que la mitiguen. Finalmente, los aportes que puedan extraerse desde esta documentación no solo podrían tener impacto en la discusión sobre políticas públicas educativas, sino también podrían aportar a la discusión teórica sobre los modelos de función de producción escolar, en especial respecto de la consideración de habilidades cognitivas y no cognitivas.



## References

- Autor, D., Figlio, D., Karbownik, K., Roth, J., & Wasserman, M. (2016). School quality and the gender gap in educational achievement. *American Economic Review*, 106(5), 289–95. doi:10.1257/aer.p20161074
- Bonesrønning, H. (2008). The effect of grading practices on gender differences in academic performance. *Bulletin of Economic Research*, 60(3), 245–264.
- Botelho, F., Madeira, R. A., & Rangel, M. A. (2015). Racial discrimination in grading: Evidence from Brazil. *American Economic Journal: Applied Economics*, 7(4), 37–52.
- Cornwell, C., Mustard, D. B., & Van Parys, J. (2013). Noncognitive skills and the gender disparities in test scores and teacher assessments: Evidence from primary school. *Journal of Human Resources*, 48(1), 236–264.
- Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, 42(3), 528–554.
- Di Liberto, A. & Casula, L. (2016). Teacher assessments versus standardized tests: Is acting “girly” an advantage?
- Fryer, R. G. & Levitt, S. D. (2010). An empirical analysis of the gender gap in mathematics. *American Economic Journal: Applied Economics*, 2(2), 210–240.
- Fuentes, P. (1998). Reading comprehension in mathematics. *The Clearing House*, 72(2), 81–88.
- Golsteyn, B. H. & Schils, T. (2014). Gender gaps in primary school achievement: A decomposition into endowments and returns to iq and non-cognitive factors. *Economics of Education Review*, 41, 176–187.
- Hanna, R. N. & Linden, L. L. (2012). Discrimination in grading. *American Economic Journal: Economic Policy*, 4(4), 146–168.
- Himmler, O. & Schwager, R. (2013). Double standards in educational standards—do schools with a disadvantaged student body grade more leniently? *German Economic Review*, 14(2), 166–189.
- Hinnerich, B. T., Höglin, E., & Johannesson, M. (2011). Are boys discriminated in swedish high schools? *Economics of Education review*, 30(4), 682–690.
- Kiss, D. (2013). Are immigrants and girls graded worse? Results of a matching approach. *Education Economics*, 21(5), 447–463.

- Lavy, V. (2008). Do gender stereotypes reduce girls' or boys' human capital outcomes? Evidence from a natural experiment. *Journal of public Economics*, 92(10), 2083–2105.
- Lavy, V. & Sand, E. (2015). *On the origins of gender human capital gaps: Short and long term consequences of teachers' stereotypical biases*. National Bureau of Economic Research.
- Loury, L. D. (1997). The gender earnings gap among college-educated workers. *ILR Review*, 50(4), 580–593.
- McEwan, P. J. (2003). Peer effects on student achievement: Evidence from Chile. *Economics of education review*, 22(2), 131–141.
- Ministerio de Educación de Chile. (2008). Marco para la buena enseñanza cpeip. *Santiago de Chile: C&C impresores*.
- Ministerio de Educación de Chile. (2014). Informe técnico SIMCE 2014. Retrieved August 28, 2017, from [http://archivos.agenciaeducacion.cl/InformeTecnicoSimce\\_2014.pdf](http://archivos.agenciaeducacion.cl/InformeTecnicoSimce_2014.pdf)
- Oteiza, F. & Miranda, H. (1996). La matemática en el aula: Contexto y evaluación. *Santiago de Chile: MINEDUC*.
- Paredes, V. (2014). A teacher like me or a student like me? Role model versus teacher bias effect. *Economics of Education Review*, 39, 38–49.
- Parra, M. P., Palacios, L. G., & González, X. M. (2016). La lógica evaluativa de los aprendizajes matemáticos en la enseñanza básica: ¿certificación de logros de aprendizaje u oportunidad para el desarrollo de las habilidades cognitivas superiores de los estudiantes? *PAIDEIA, REVISTA DE EDUCACIÓN*, (46).
- Price, J. (2010). The effect of instructor race and gender on student persistence in STEM fields. *Economics of Education Review*, 29(6), 901–910.
- Schneeweis, N. & Zweimüller, M. (2012). Girls, girls, girls: Gender composition and female school choice. *Economics of education review*, 31(4), 482–500.
- Sprietsma, M. (2013). Discrimination in grading: Experimental evidence from primary school teachers. *Empirical Economics*, 1–16.
- Terrier, C. (2015). Giving a little help to girls? Evidence on grade discrimination and its effect on students' achievement.
- Terrier, C. (2016). *Boys lag behind: How teachers' gender biases affect student achievement*. Institute for the Study of Labor (IZA).

- Vilenius-Tuohimaa, P. M., Aunola, K., & Nurmi, J.-E. (2008). The association between mathematical word problems and reading comprehension. *Educational Psychology, 28*(4), 409–426.
- Woessmann, L. & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in timss. *European Economic Review, 50*(3), 695–736.
- Wooldridge, J. M. (2015). *Introductory econometrics: A modern approach*. Nelson Education.

## Apéndice

### Apéndice A

Elementos sociodemográficos incorporados:

1. Nivel educacional de la madre, caracterizado en 21 rangos.
2. Rangos en que se encuentra la suma de los ingresos de todas las personas que aportan al hogar donde vive el estudiante, caracterizado en 15 rangos.
3. Rangos de cantidad de libros disponibles en el hogar, caracterizado en 5 rangos.
4. estudiante es beneficiario de subvención escolar preferencial (SEP), caracterizado como variable binaria.

## Apéndice B

Preguntas de las encuestas SIMCE respondidas por estudiantes. Las preguntas eran distintas cada año, pero compartían sentido. Las preguntas eran de selección de alternativas, con escalas categorizadas de 1 a 4, y fueron respondidas luego de realizar las pruebas SIMCE. Las escalas de 1 a 4 siempre representan mejores niveles de habilidades no cognitivas o actitudinales a medida que el valor crece; por ejemplo, para la pregunta 1), un valor de 4 es que siempre el estudiante hace tareas y trabajos.

Año 2015 (8<sup>vo</sup> básico):

- 1) “Siempre hago las tareas y trabajos.” (Escala de 1 a 4)
- 2) “Cuando falto a clases, me importa mucho ponerme al día con la materia.” (Escala de 1 a 4)
- 3) “Siempre trato de entender bien lo que me enseñan en clases.” (Escala de 1 a 4)
- 4) “Si me saco una mala nota, estudio más para la próxima prueba.” (Escala de 1 a 4)
- 5) “He participado en actividades académicas y culturales (por ejemplo, obras de teatro, exposiciones de arte, feria de ciencia y tecnología, concursos de debate, etc.)” (Escala binaria)
- 6) “He ayudado a organizar actividades académicas y culturales.” (Escala binaria)

Año 2013 (6<sup>to</sup> básico):

- a) “Siempre hago mis tareas.” (Escala de 1 a 4)
- b) “Hago las tareas aunque sean difíciles.” (Escala de 1 a 4)
- c) “¿Has participado en las siguientes actividades este año? Actividades culturales que organiza la escuela.” (Escala binaria)

Año 2011 (4<sup>to</sup> básico):

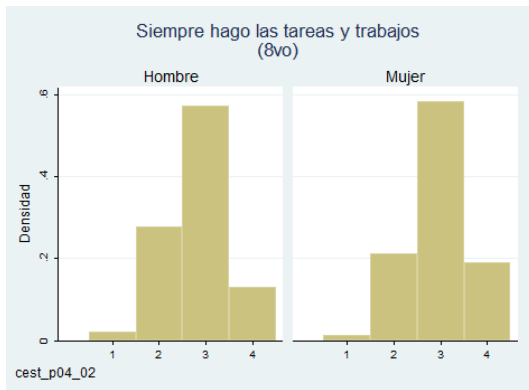
- i) “¿Cuántos días a la semana, de lunes a viernes, haces las siguientes actividades fuera del horario escolar? Estudiar, hacer tareas o trabajos escolares.” (Escala de 1 a 4)
- ii) “Hago las tareas aunque me resulten difíciles.” (Escala de 1 a 4)
- iii) “¿Participas en las siguientes actividades en tu establecimiento? Actividades académicas y culturales (por ejemplo: concursos de debate, feria de ciencias y tecnología, encuentros literarios, exposiciones de arte, etc.)” (Escala binaria)

Table 9: Diferencias por género en las preguntas con medias agrupadas por curso. Se presentan las medias y los errores estándar en paréntesis. Nivel de confianza al 99%:

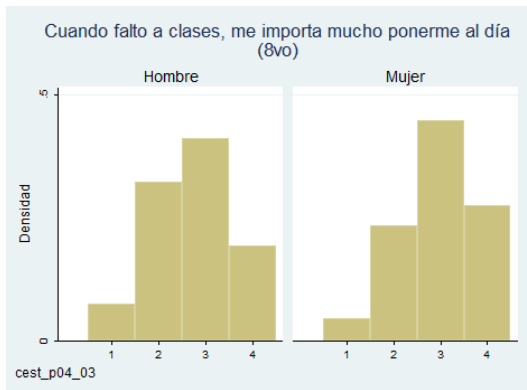
Pregunta	Mujeres	Hombres	¿Diferencia significativa?
1)	2,95 (0,003)	2,81 (0,003)	Sí
2)	2,95 (0,004)	2,72 (0,004)	Sí
3)	3,33 (0,003)	3,30 (0,003)	Sí
4)	3,15 (0,004)	3,03 (0,004)	Sí
5)	0,62 (0,003)	0,63 (0,003)	No
6)	0,50 (0,002)	0,54 (0,002)	Sí
a)	2,95 (0,003)	2,84 (0,003)	Sí
b)	3,09 (0,003)	3,07 (0,003)	Sí
c)	0,75 (0,002)	0,76 (0,002)	Sí
i)	3,07 (0,004)	2,91 (0,004)	Sí
ii)	3,51 (0,003)	3,39 (0,003)	Sí
iii)	0,79 (0,002)	0,78 (0,002)	Sí

Comparaciones gráficas de las preguntas del cuestionario SIMCE.

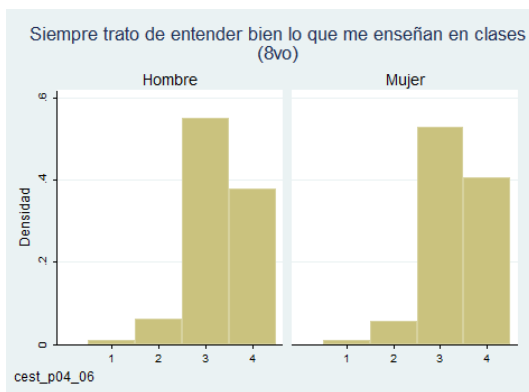
Preguntas de 8<sup>vo</sup>:



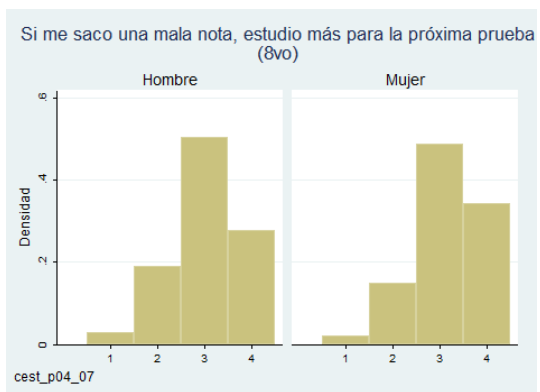
1)



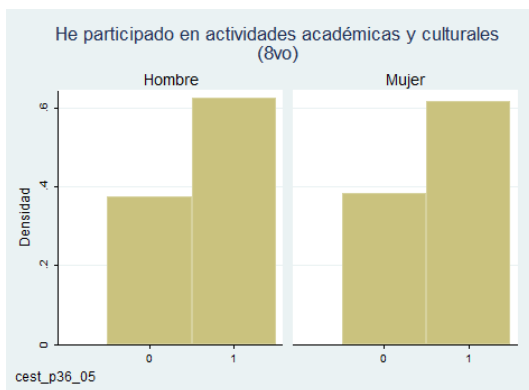
2)



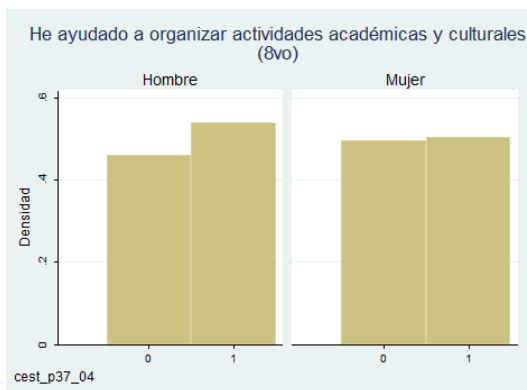
3)



4)



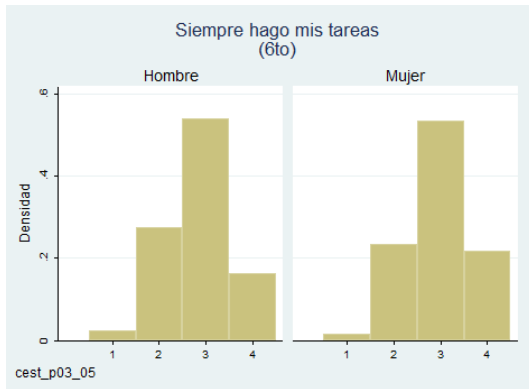
5)



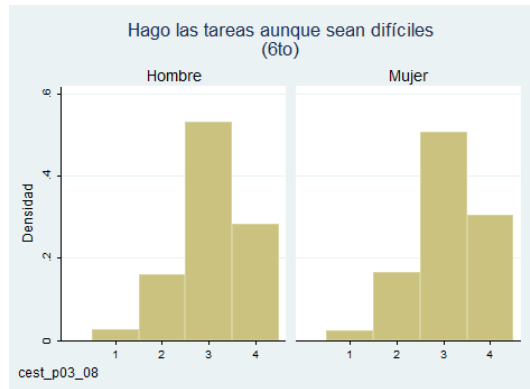
6)



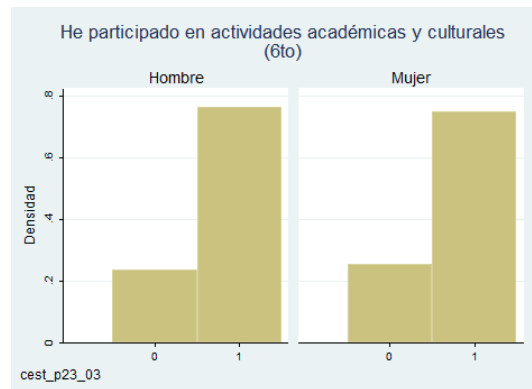
Preguntas de 6<sup>to</sup>:



a)

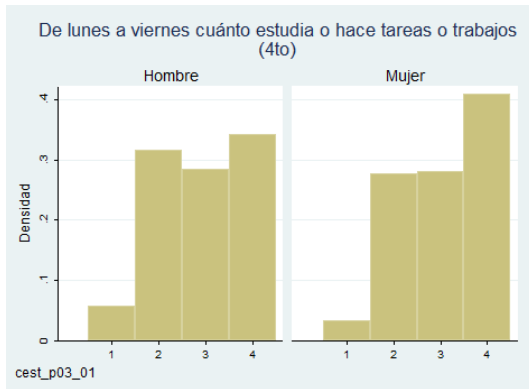


b)

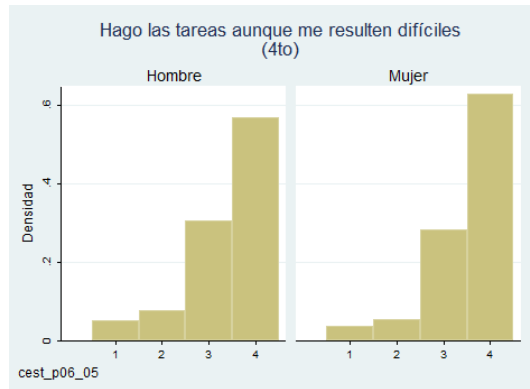


c)

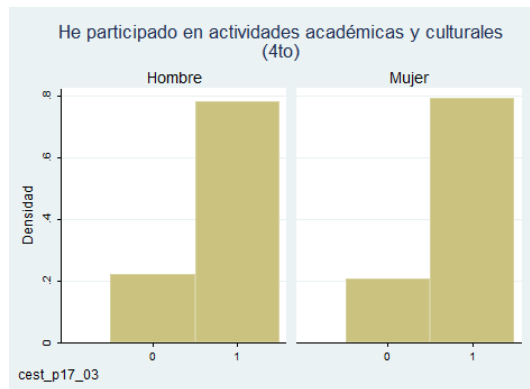
Preguntas de 4<sup>to</sup>:



i)



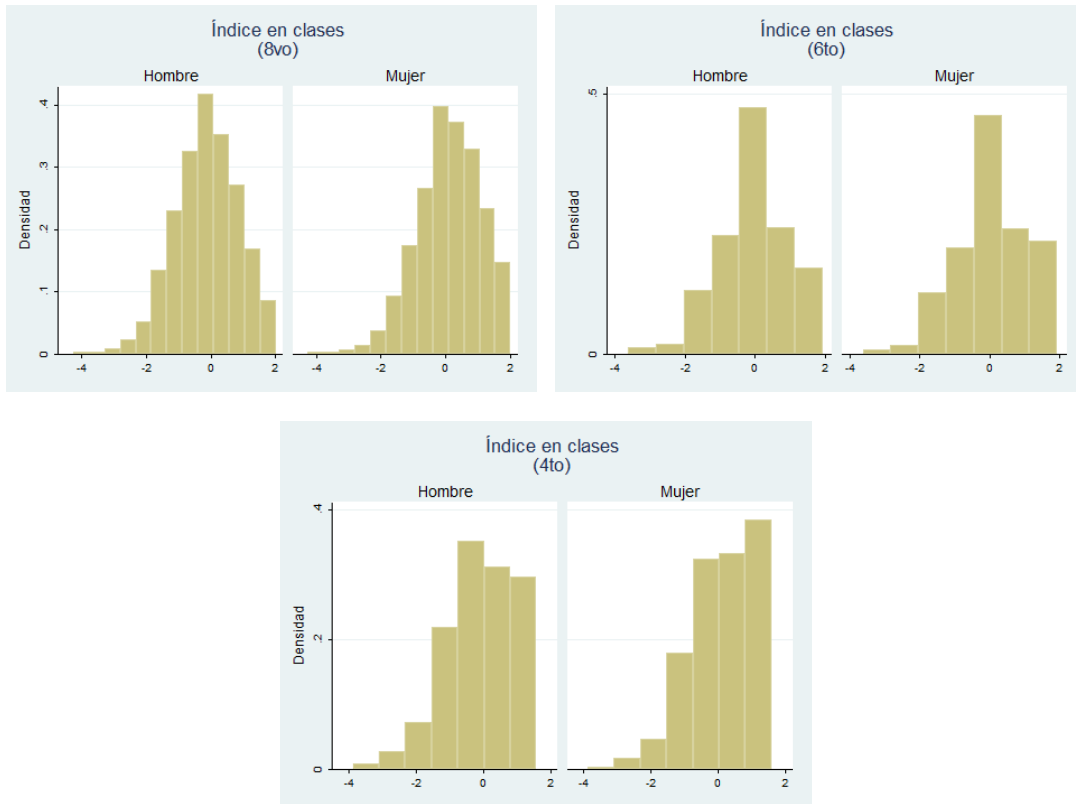
ii)



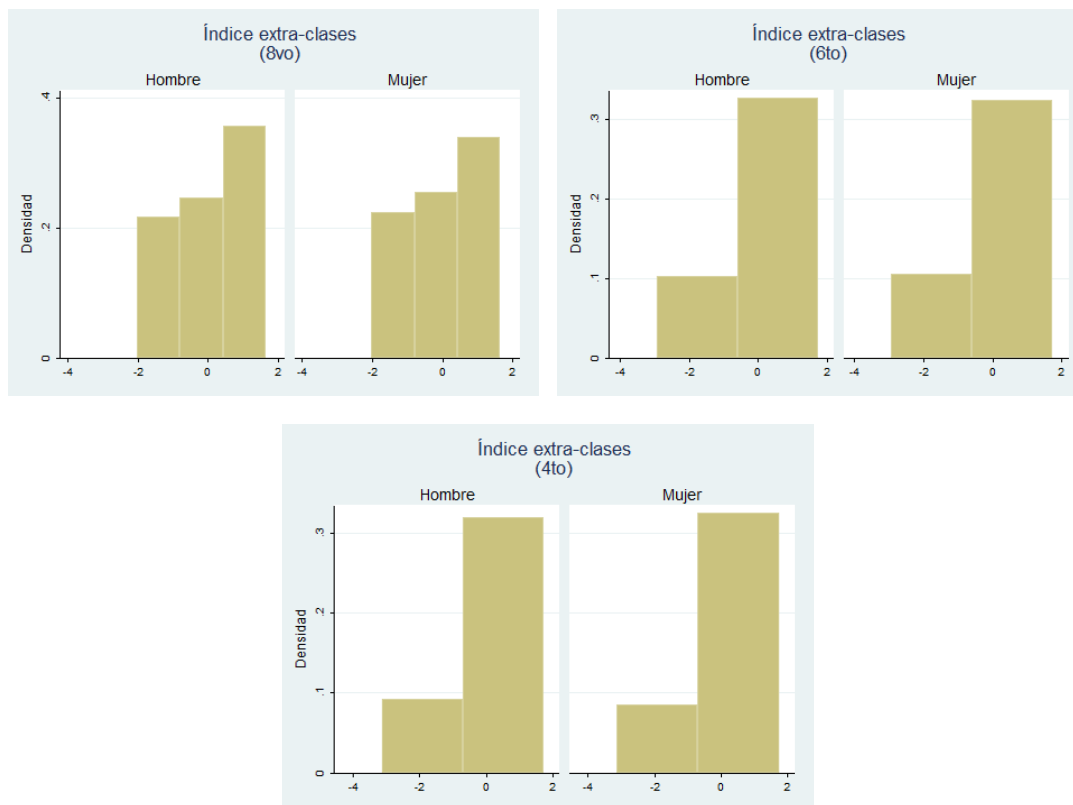
iii)

## Apéndice C

Índices construidos y estandarizados respecto de la muestra del nivel. Índices en clases:



Índices construidos y estandarizados respecto de la muestra del nivel. Índices extra-clases:



Correlación para índices (todos los valores son significativos al 99% de confianza):

	Índice en clases			Índice extra-clases		
	8 <sup>vo</sup>	6 <sup>to</sup>	4 <sup>to</sup>	8 <sup>vo</sup>	6 <sup>to</sup>	4 <sup>to</sup>
8 <sup>vo</sup>	1			1		
6 <sup>to</sup>	0.359	1		0.186	1	
4 <sup>to</sup>	0.171	0.210	1	0.089	0.093	1

## Apéndice D

Primeras etapas de regresiones con Variables Instrumentales y efectos fijos de la Tabla 7.

Table 10: Variable dependiente: índices en clases para 8<sup>vo</sup> básico.

	(1)	(2)
	Columna 2 Tabla 7	Columna 3 Tabla 7
Mujer	0.168*** (0.007)	0.180*** (0.007)
SIMCE mat	0.025*** (0.008)	0.025*** (0.008)
SIMCE mat <sup>2</sup>	-0.002 (0.007)	-0.004 (0.007)
SIMCE mat <sup>3</sup>	0.003* (0.002)	0.003* (0.002)
SIMCE mat <sup>4</sup>	-0.002 (0.002)	-0.002 (0.002)
SIMCE lect	0.066*** (0.004)	0.067*** (0.004)
Calificación 6	0.086*** (0.004)	0.092*** (0.004)
Índice en clases 6	0.302*** (0.003)	0.315*** (0.003)
Índice en clases 4	0.073*** (0.003)	
Efectos fijos	Curso	Curso
Controles sociodemográficos	Sí	Sí
N#	110,871	110,871
<i>F</i> de instrumentos excluidos	4,830	9,020
<i>F</i> Cragg-Donald Wald de instrumentos débiles	5,445	10,264

NOTA: Regresiones con errores estándar agrupados por curso. Niveles de significación: \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

## Apéndice E

Tabla descriptiva para los datos de la Tabla 7: porcentaje de estudiantes pertenecientes a cursos con profesores con la característica relevante.

Table 11: Estadística descriptiva de los datos para la Tabla 8

Característica del docente	N =	Porcentaje del total (N = 110,871)
<b>Columna 1: profesora mujer</b>		
Cantidad de estudiantes	<i>58,283</i>	52.57%
Estudiantes mujeres dentro del sub-grupo	<i>30,781</i>	27.76%
<b>Columna 2: profesor jefe</b>		
Cantidad de estudiantes	<i>20,766</i>	18.73%
Estudiantes mujeres dentro del sub-grupo	<i>10,643</i>	9.60%
<b>Columna 3: profesor con post-grado</b>		
Cantidad de estudiantes	<i>52,203</i>	47.08%
Estudiantes mujeres dentro del sub-grupo	<i>26,975</i>	24.33%
<b>Columna 4: profesor hizo clases año anterior</b>		
Cantidad de estudiantes	<i>79,488</i>	71.69%
Estudiantes mujeres dentro del sub-grupo	<i>40,656</i>	36.67%