



EFFECTO DE LA REPITENCIA SOBRE AÑOS DE ESCOLARIDAD ALCANZADOS

TESIS PARA OPTAR AL GRADO DE
MAGÍSTER EN ECONOMÍA

Alumno: Vania Martínez Garrido
Profesor Guía: Juan Pablo Valenzuela

Santiago, noviembre 2016

Agradecimientos

En primera instancia agradezco el apoyo de CONICYT¹, ya que sin su financiamiento la obtención de este grado académico no hubiese sido posible. Agradezco también al Ministerio de Educación por toda la información facilitada, insumos que jugaron un rol importante en el desarrollo de esta investigación.

Agradezco a mi profesor guía Juan Pablo Valenzuela por todo el apoyo y paciencia mostrada. Muchas gracias por todo lo enseñado, por todo el tiempo dedicado y por creer siempre que el tema iba a perdurar e íbamos a lograr algo bueno, sin duda todo esto no hubiese sido posible sin Ud. Agradezco a su mano derecha en este proceso y alguien que aportó considerablemente a la investigación, muchas gracias Claudio Allende por todos los conocimientos transmitidos, todos los mails contestados, toda la ayuda desinteresada, fuiste una pieza importante también.

Agradezco a mis amigos de la vida y de la universidad, por todas las palabras de apoyo, porque a pesar de que fue un camino largo muchos consejos fueron dados en el momento correcto. En especial muchas gracias a Los Five, porque en esta etapa universitaria fueron importantísimos y definitivamente no hubiese culminado de esta forma si Uds. no hubiesen estado trabajando conmigo codo a codo, supimos apoyarnos en momento alegres y difíciles, les deseo todo el éxito del mundo porque capacidades tienen de sobra para lograrlo.

Infinitas gracias a mi familia por apoyarme siempre, me siento inmensamente privilegiada por tenerlos a mi lado, por todo lo que me han enseñado y entregado, este es un logro más de muchos en los que serán una pieza fundamental. En especial quiero agradecer a mi hermano y mami porque nunca han dejado de creer en mi, porque en sus palabras de aliento siempre encuentro sinceridad y amor, y es todo

¹ Apoyo otorgado a través del Programa de Formación de Capital Humano Avanzado - Concurso Magíster Nacional año 2014 y al Proyecto Basal FB0003 del Programa de Investigación asociativa.

lo que necesito para ser feliz.

Finalmente, y no por eso menos importante, agradezco a Dios por todo lo que me ha regalado, no hay argumentos para creer en ti, simplemente te siento a mi lado, apoyándome a sortear cada una de las etapas que me hacen crecer.

Índice general

1. Introducción	5
2. Revisión Bibliográfica	9
3. Sistema Legal	15
4. Metodología	18
5. Datos	22
6. Identificación	28
6.1. Identificación separando hombres y mujeres.	35
7. Efecto de la repitencia sobre los años de escolaridad alcanzados	42
7.1. Mujeres	42
7.1.1. Efectos dinámicos	48
7.2. Hombres	56
8. Conclusiones	61
Apéndices	65
Bibliografía	72

Capítulo 1

Introducción

Desde la segunda mitad del siglo pasado, la política educacional de Chile se enfocó en una ardua lucha que buscaba aumentar las tasas de cobertura nacional en todos los niveles de enseñanza. Ya en los años 90's la tasa de matrícula¹ en educación básica oscilaba en torno a un 95 %, cifra cercana a la universalidad, lo que ameritaba incluir nuevos desafíos que apuntaran a políticas cuyo objetivo principal fuese incrementar la calidad de la educación (Arellano, 2001) y desenmarañar los principales “nudos críticos” del sistema.

En el documento escrito por Bellei (2013) se señala a la repitencia escolar, deserción de los estudiantes e ingreso de los alumnos al sistema educacional a una edad inadecuada, como los principales “nudos críticos” para el cumplimiento de la universalización de la educación. En Chile, dentro de los mencionados “nudos críticos” existe uno al que no se le ha prestado una adecuada atención, el que además se ha vuelto un mecanismo de selección y estratificación muy utilizado dentro de las escuelas chilenas. Como Amadio (1996) destaca, la repitencia es “...una de las manifestaciones perceptibles de la inadecuación del sistema educacional moderno a las condiciones, posibilidades y necesidades concretas y diferenciadas de la población, en particular, de esa gran mayoría de alumnos provenientes de los sectores sociales menos favorecidos”.

Los resultados del Tercer Estudio Regional Comparativo y Explicativo (TERCE) indicaron que en Latinoamérica la repitencia *constituyó el mecanismo por excelencia de remediación de rezagos en el*

¹Tasa de Matrícula es definida como *Matrícula total de alumnos de un determinado nivel de enseñanza, educación básica y/o educación media, en relación a la población ajustada del correspondiente tramo de edad.*

aprendizaje (LLECE, 2015). Esto cobra total sentido en el contexto de los países latinoamericanos, ya que en estos sistemas escolares la repitencia no se percibe como un fracaso e irresponsabilidad de cada escuela, sino que se percibe como una práctica aceptada, masificada y necesaria para el progreso educacional, responsabilizando a cada estudiante, y su historia, de las causas por las que se vivencia y las consecuencias que ésta conlleva. LLECE (2015) también destaca a la repitencia como la *segunda variable que tiene una relación más robusta con el logro de aprendizaje*, después de nivel socioeconómico, dejando entrever un problema mayor, ¿qué niños son más susceptibles a experimentar esta práctica educacional?

En el marco de lo mencionado anteriormente podemos señalar el trabajo de Dupriez, Dumay & Vause (2008), que tomando la clasificación en Mons (2007) de cómo los sistemas escolares manejaban la heterogeneidad entre sus estudiantes, evaluaron las diferencias en que estos modelos, aplicados a países participantes de PISA 2003, hicieron frente a la diversidad manifestada entre los alumnos con peor desempeño académico. Dentro de esta investigación es interesante enfatizar el modelo de “integración uniforme”, aquel que pone en práctica un currículum común hasta los 14, 15 o 16 años, que muestra una mayor diferenciación académica y sociocultural, y en el que la repitencia escolar es uno de los principales mecanismos para controlar la heterogeneidad de los distintos estudiantes, éste sería el caso de Chile.

A modo de validar rasgos de “integración uniforme” presentes en el sistema educacional chileno, en el Cuadro 1 se encuentra la tasa de reprobación para nuestro país, considerando solo los primeros 6 años de educación básica. A partir de él se desprende que efectivamente la tasa de repitencia ha ido en aumento en la última década, pasando desde un 2,4% en el año 1999 a un 4,7% en el año 2012, esta alza casi duplicó el porcentaje de reprobación en Chile, lo que nos llevó de estar bajo el promedio mundial a superarlo en el año 2012. Además vemos que independiente del grado de desarrollo de un país, la tendencia predominante es a disminuir las tasas de repitencia, situación contraria a la experimentada en el caso chileno. Ahondando más, para el año 2012 identificamos que dentro de la distribución de porcentajes de repitencia tanto en educación básica como media, la mayor cifra se alcanzó en 7mo básico (5,86%) y 1ro medio (13,65%)², respectivamente. La tendencia fue a un aumento de las tasas de repitencia para niveles más altos de escolaridad, produciéndose una disminución solo en el último año de educación formal, es decir, 4to medio (2,09%).

²Estimaciones propias realizadas a partir de la base de datos SIGE 2012.

Cuadro 1: Repitencia en la Enseñanza Primaria.

Alumnos repetidores en todos los grados (%)	Año escolar finalizado en	
	1999	2012
Chile	2,4	4,7
Argentina	5,9	4,2 ^z
Uruguay	7,9	5,5 ^y
Brasil	24,0	8,7 ^z
México	6,6	2,9
Suiza	1,8	1,3
Alemania	1,7	0,5
Mundo	5,2 ^{**}	4,6 ^{**}
América Latina y el Caribe	4,7	5,4 ^{**}
Países desarrollados	1,0 ^{**}	0,8
Países en desarrollo	5,8 ^{**}	5,1 ^{**}

Fuente: Base de datos del Instituto de Estadística de la UNESCO (IEU). En este cuadro, la duración de la enseñanza primaria se define en función de la CINE 97 y puede diferir de la que esté vigente en el plano nacional.

(z): Los datos corresponden al año escolar finalizado en 2011.

(y): Los datos corresponden al año escolar finalizado en 2010.

(**): Estimaciones parciales del IEU.

Como vimos en el Cuadro 1, con el paso de los años la repitencia se ha hecho una práctica cada vez más común, y es en esta línea que creemos importante medir su impacto, esto con el fin de tener mayor evidencia a la hora de criticarla o potenciarla como mecanismo recurrente en el sistema educacional chileno. Por lo tanto, el objetivo de este trabajo será testear si la repitencia, como una práctica actualmente utilizada en el sistema escolar chileno, daña los subsecuentes logros académicos de los estudiantes. Si los resultados indicaran un efecto negativo, esta práctica atentaría aun más contra la calidad del sistema. Nuestra hipótesis es que la repitencia se da en perfiles de estudiantes determinados y que más allá de ser una coincidencia es una especie de “bomba de tiempo”: si estudiantes de determinadas características no repiten hoy es muy probable que lo hagan el día de mañana, esto dado que cargan con una mochila de características que el sistema educacional chileno no logra soslayar y mermar.

No obstante, aun cuando encontrásemos un efecto adverso de la repitencia nuestra investigación no sería suficiente para proponer una eliminación completa y definitiva de esta práctica, esto dado que pudie-

sen existir efectos heterogéneos en perfiles distintos de estudiantes lo que podría llevar a situaciones en donde hacer repetir a un alumno efectivamente implica una mejora en su desempeño académico, así como también a un incremento del compromiso y responsabilidad hacia sus logros educacionales. Lo anterior apoyaría la idea de una tasa de repitencia de estado estacionario, algún nivel óptimo para que el sistema educacional funcione bien y logre asegurar un nivel mínimo de contenidos para todos los estudiantes.

La estructura de la investigación será la siguiente: en el Capítulo 2 se revisarán algunos de los principales trabajos que analizan el tema de la repitencia escolar, tanto a nivel internacional como nacional; en el Capítulo 3 se describirá el marco legal en el cual se desarrolla la repitencia escolar en Chile; en el Capítulo 4 se explicará y detallará la metodología a explorar; en el Capítulo 5 se proporcionará información acerca de los datos utilizados en esta investigación; en el Capítulo 6 se mostrarán y discutirán los principales supuestos necesarios para aplicar la metodología descrita anteriormente, concluyendo que solo en la muestra de mujeres se podrá estimar el efecto de la repitencia; en el Capítulo 7 se realizarán las estimaciones, solo considerando una muestra de mujeres, y se profundizará el tema tratando de capturar la dinámica de la repitencia; y finalmente, en el Capítulo 8 se presentarán las conclusiones finales.

Capítulo 2

Revisión Bibliográfica

Entre las primeras concepciones que llevaron a validar a la repitencia como práctica común en el sistema escolar chileno se encontraban buenas intenciones de “darle” una segunda oportunidad al estudiante desaventajado que no lograba aprender un nivel mínimo de contenidos para su edad (García-Huidobro, 2000), obviando los efectos adversos que la repitencia podría generar en el sistema educacional. Por ejemplo, en el informe en que se analizan los resultados de PISA 2012 (OCDE, 2013) se destaca a la repitencia como un mecanismo de *estratificación vertical*, enfatizando la correlación negativa que existe entre repitencia y equidad, dicho de otro modo, “*los sistemas en donde más estudiantes repiten tienden a mostrar mayor impacto del nivel socioeconómico de los alumnos en su desempeño*”. Lo anterior, en el contexto del financiamiento compartido, característica que identifica al sistema educacional chileno, llevaría a exacerbar las brechas existentes entre alumnos de bajos ingresos versus aquellos educados en el seno de una familia de mayores recursos (Mizala & Torche, 2012; Epple & Romano, 1998; Valenzuela et al., 2013).

Por otra parte, en el estudio de diferencias de medias de Vera (2013) se menciona que dentro de los alumnos repitentes (versus los no repitentes) se observaría una mayor probabilidad de ser hombre, de provenir de familias con menor nivel socioeconómico, de asistir a establecimientos públicos, de presentar bajos niveles de posesión cultural, una menor cantidad de libros en el hogar y en cuyas familias sus padres tendrían menos años de escolaridad. En vista de los resultados mencionados podríamos identificar en Chile indicios de una polarización de los perfiles de alumnos y colegios, lo que estaría generando esta dicotomía en las características de niños y jóvenes repitentes versus no repitentes. Es más, Vera (2013) señala que un alumno repitente que asiste a un establecimiento de alto nivel socioeconómico puede, en

alguna medida, anular el efecto adverso de la repitencia, lo que sustentaría la hipótesis de la polarización.

Actualmente para el caso de Chile podemos mencionar el estudio de Díaz, Grau, Reyes & Rivera (2016), cuya motivación principal es cuantificar los efectos reales que tiene la repitencia escolar sobre crimen juvenil. Tomando una muestra de estudiantes que en el año 2007 se encontraba entre 4to y 8vo básico, los autores identificaron que la repitencia llevaba a un aumento en la probabilidad de cometer un crimen entre 1,6 a 3,7 puntos porcentuales. En análisis adicionales identificaron que la repitencia también aumentaba la probabilidad de desertar (1,2 a 3,2 pp.) y disminuía la probabilidad de futuras repitencias (2,3 a 10,4 pp.). Adicionalmente, podemos destacar investigaciones en las cuales la repitencia es utilizada como una variable de control en el proceso para responder a la pregunta central que las motiva. En el trabajo de Valenzuela, Allende, Sevilla & Egaña (2013) se analiza la trayectoria académica, a través del puntaje SIMCE¹, para aquellos estudiantes que experimentaron la repitencia en 4to básico. Aquí se señala que los alumnos repitentes desde antes de repetir ya tenían asociados puntajes bajos comparado con el desempeño medio nacional, y luego de la repitencia sus puntajes SIMCE aumentaron, pero en una manera insuficiente tal que no alcanzaron un nivel aceptable de logros, esto indicaría que el desempeño académico previo de los estudiantes sería un factor clave a considerar. En Mizala & Torche (2012) se examinó la distribución socioeconómica de los logros entre y al interior de los colegios, tanto del sector público como del sector particular subvencionado, los resultados apuntaron a que haber repetido un grado afectó negativa y significativamente el desempeño en la prueba SIMCE de 4to y 8vo básico para la cohorte 2002 y 2004, respectivamente. Estos son algunos de los trabajos en donde la identificación de haber repetido algún curso durante la vida escolar del estudiante es importante a la hora de abordar el tema central que atinge a la investigación².

En el ámbito internacional sí encontramos estudios cuyo objetivo principal es medir el impacto que tiene la repitencia sobre subsecuentes logros académicos y/o laborales. Un primer grupo de estudios señala que la repitencia a temprana edad no tendría resultados dañinos. Aquí destacan Schwerdt, West & Winters (2015), quienes a partir de una muestra de estudiantes de tercer grado de colegios públicos del Estado de Florida, EEUU, encontraron que los alumnos que no pasaron de curso experimentaron ganan-

¹SIMCE (Sistema de Medición de la Calidad de la Educación) es el sistema nacional de evaluación de resultados de aprendizaje aplicado a niños y jóvenes que se encuentran insertos en el sistema educacional chileno.

²Debemos tener presente que la omisión de una variable relevante siempre estará asociada a problemas econométricos que cualquier investigador quiere y debe minimizar, por ende en algunos casos se hace imprescindible la incorporación de la repitencia como variable de control.

cias de corto plazo en matemáticas y lectura: en los primeros 3 años después de repetir los estudiantes repitentes superaron a sus pares de la misma edad, pero promovidos, en 0,31 desviaciones estándar en lectura y 0,23 desviaciones estándar en matemáticas, estos efectos desaparecieron después de 5 años. Además, Schwerdt et al. (2015) encontraron que repetir en tercer grado redujo la probabilidad de volver a repetir en cada uno de los siguientes cuatro años, y mostró no tener efectos sobre la probabilidad de graduarse de secundaria³. Jacob & Lefgren (2009) analizaron datos del distrito escolar público de Chicago, EEUU, y a partir de ellos identificaron que la repitencia en estudiantes de bajo desempeño no tuvo efectos significativos sobre la probabilidad de graduarse en secundaria, tanto para estudiantes de sexto grado como para los de octavo grado, pero de mayor edad⁴. No obstante, encontraron que para estudiantes de octavo grado, pero más jóvenes, repetir conllevó a un aumento de 11,2 puntos porcentuales en la probabilidad de desertar del sistema educacional. Una explicación propuesta para el efecto no significativo de la repitencia en sexto grado descansó en que los estudiantes que no pasaron de curso tuvieron mayor probabilidad de alcanzar a su cohorte original⁵: alrededor de dos tercios de los estudiantes repitentes en sexto básico fueron capaces de alcanzar a su cohorte original. Finalmente, en la investigación de Greene & Winters (2009) se midió el impacto de recibir una exención a repetir sobre competencias futuras, para estudiantes de tercer grado del Estado de Florida, EEUU. Los resultados sugirieron que los estudiantes que recibieron una exención de portafolio (basado en recopilación de trabajos en el último año académico) tuvieron competencias en lectura, después de dos años, que fueron 0,21 desviaciones estándar menores a estudiantes con similares características demográficas, pero que no pasaron de curso. A partir de lo anterior podríamos extrapolar que haber hecho repetir al alumno hubiese sido más beneficio que otorgarle una exención.

El segundo grupo de estudios apunta a los efectos adversos asociados a la repitencia. Aquí destacamos a Manacorda (2012), quien utiliza un diseño de regresión discontinua Fuzzy⁶ para explorar la regla

³Después de 6 años los estudiantes que repitieron estuvieron, en promedio, solo 0,73 grados atrás que sus pares que pasaron de curso. Los estudiantes que repitieron y completaron la secundaria lo hicieron sólo 0,55 años después que los que pasaron de curso, a pesar de esperar que la diferencia hubiese sido de al menos 1 año entre los que pasaron versus los que no lo hicieron, indicando que existieron dinámicas que acortaron las brechas.

⁴Los autores separan en dos grupos a los estudiantes de 8vo grado, esto debido a un tratamiento diferencial entre jóvenes mayores y menores de 15 años: si un alumno con más de 15 años llegase a repetir éste no se mantendría en el colegio sino que sería trasladado a un centro de transición, en donde recibiría instrucción objetivo. Debido a que el “tratamiento” recibido por los alumnos de mayor edad es potencialmente diferente, se analizó a los grupos por separado.

⁵En el sistema de colegios públicos de Chicago todos los estudiantes, tanto repitentes como no repitentes, enfrentan una oportunidad de promoción una vez que alcanzan 8vo grado.

⁶Regresión Discontinua (RD) es un método cuasi experimental cuya asignación al tratamiento es determinada parcial o completamente por el valor de cierto predictor (variable de asignación), estando a cada lado del umbral. La diferencia entre RD Sharp y RD Fuzzy radica en la forma en que la variable de asignación y la probabilidad de recibir el tratamiento interactúan. En

de promoción escolar del sistema educacional de Uruguay y medir el impacto de la repitencia sobre la probabilidad de deserción y los años de escolaridad adicionales alcanzados. El autor concluye que la repitencia aumentaría la tasa de deserción escolar: los repitentes fueron, en promedio, 50% menos probables de estar en el colegio después de un año comparado con los no repitentes. Además, repetir disminuiría los logros académicos hasta en 4 o 5 años después de ocurrido el tratamiento: los repitentes, en promedio, lograron 0,8 años de escolaridad menos que los no repitentes. Jimerson (1999) investigó la correlación entre repitencia a temprana edad y resultados presentados al final de la adolescencia, a través de un análisis de varianzas (ANOVA) y Chi cuadrados. La muestra fue seleccionada entre niños participantes del programa “Minnesota Mother-Child Interaction Project” y sus conclusiones apuntaron a que al final del onceavo grado el grupo de repitentes mostró un menor ajuste académico que el grupo de control y el grupo de bajo desempeño, pero promovido; además, un 69% de los alumnos repitentes desertó de la secundaria versus un 46% de los estudiantes promovidos, pero de bajo desempeño; a los 20 años, sólo un 44% de los estudiantes repitentes recibió un certificado de término de secundaria versus un 72% para el grupo de promovidos, pero de bajo desempeño; los alumnos repitentes fueron menos probables a estar inscritos en la educación superior, mientras que los porcentajes para el grupo de control y para los promovidos, pero de bajo desempeño fueron mayores (23%, 41% y 56%, respectivamente).

Por otra parte, Hill (2014) señaló que los repitentes podían ser vistos como compañeros de clases de bajo desempeño para los estudiantes que tomaban cursos matemáticos por primera vez, en el contexto del sistema educacional norteamericano. Hill (2014) midió el impacto que tuvo la proporción de repitentes sobre los estudiantes que tomaban por primera vez el curso. A partir de los resultados se concluyó que un incremento de 10 puntos porcentuales en la participación de repitentes causó un incremento de 1,9 puntos porcentuales en la probabilidad de fracasar para los alumnos que tomaron por primera vez el curso. Los efectos de los repitentes operaron en la parte media y baja de la distribución de habilidades, y el efecto por de la habilidad fue más fuerte a nivel de clase que a nivel de grado. Los efectos distribucionales sugirieron que los repitentes fueron más probables a distraer a sus compañeros de clases quienes se localizaron en partes bajas de la distribución de desempeño y no a redireccionar recursos de profesores desde estudiantes de alta habilidad a aquellos de baja habilidad. A partir de una revisión exhaustiva de la literatura Jimerson, Anderson & Whipple (2002) pudieron concluir que la repitencia escolar fue uno de los principales

RD Sharp la probabilidad de recibir el tratamiento salta de 0 a 1 una vez que la variable de asignación traspasa cierto umbral. En RD Fuzzy la probabilidad de recibir el tratamiento experimenta un pequeño salto una vez que el predictor toma valores por sobre cierto umbral (Imbens & Wooldridge, 2009).

predictores de la posterior deserción al sistema educacional, destacando que la deserción debía ser vista como un proceso y no como un evento. Los autores señalaron que una vez que se repite, otros eventos (como por ejemplo, retiro, ausentismo, baja autoestima, entre otros) asociados con la deserción son más probables de ocurrir, reforzando las trayectorias de proceso que llevarían a desertar en secundaria.

Finalmente en esta misma línea, McCoy & Reynolds (1999) estudiaron los predictores y consecuencias de la repitencia escolar sobre logros académicos, capacidades percibidas y delincuencia al comienzo de la adolescencia. La metodología consistió en calcular los predictores de la repitencia a través de una regresión logística, luego, una regresión MCO fue utilizada para predecir los puntajes en matemáticas y lectura. Al analizar a niños provenientes de familias de bajos ingresos y en su mayoría de color negro en colegios públicos de Chicago, EEUU, los resultados apuntaron a que comparando a estudiantes de la misma edad, aquellos que no pasaron de curso lograron 9,5 y 8,9 puntos menos en la prueba de lectura y matemáticas, respectivamente.

Pero no sólo por considerar efectos positivos y negativos de la repitencia escolar ésta debiese convertirse en un tema que interese a los hacedores de políticas públicas, sino que además debiésemos tomar en cuenta los múltiples costos que ésta implica. En UNESCO (2012) señalan como costos directos la cantidad de recursos educativos que el alumno repitente consumiría en el año adicional que estaría en el establecimiento, llevando a un aumento del tamaño de clases y a una restricción de la capacidad del sistema escolar. Por otro lado, se generaría un costo de oportunidad al retrasar la entrada de los jóvenes al mercado laboral: la repitencia acortaría el tiempo que permanecerían los individuos en el mercado del trabajo. Finalmente, la retención de los alumnos en un determinado grado podría llevar a costos psicológicos asociados a un deterioro de la autoestima, y esto pudiese manifestarse ya sea en no poder seguir compartiendo con sus compañeros de clases, con los cuales ya había formado lazos de amistad, o por el bullying y presión psicológica que generaría el medioambiente social producto de este fracaso (Roderick, 1994). En el informe de OCDE (2013) estiman los costos asociados a la repitencia escolar para países que dispusieron de información relevante para los cálculos. Dentro de los países en los que la repitencia ha llevado a un mayor costo podemos destacar a Bélgica y a los Países Bajos, en cuyo caso la cifra asciende al 10% o más del total del gasto nacional anual en educación primaria y secundaria, evidenciando el enorme costo de oportunidad que conllevaría esta práctica.

Como se puede apreciar, la repitencia escolar no es una realidad que esté ajena a las preocupaciones internacionales, de hecho es un tema que ha tomado crucial importancia en los últimos años y por ende en Chile no debiese ser obviada. La aplicación de esta práctica conjuga historias privadas como también historias de políticas públicas, esto ya que considera efectos individuales en los alumnos repitentes (como por ejemplo, impacto en la autoestima, en la motivación y en la probabilidad de deserción (LLECE, 2015)) como también efectos globales (altos costos financieros e ineffectividad del sistema educacional en el logro de conocimientos básicos en todos los alumnos). Por lo tanto, creemos que por el sólo hecho de que la repitencia pudiese eventualmente representar un costo para el país es necesario situarla como un tema importante a considerar, evaluándola desde un punto de vista de política pública, independiente de los efectos positivos o negativos que podamos destacar en la literatura.

Balancear correctamente los efectos adversos y positivos, identificando casos en que los primeros pudiesen ser superados por los segundos o viceversa, es una tarea pendiente. Lo óptimo podría ser mantener una tasa de repitencia baja, algún nivel de estado estacionario, con el fin de que el sistema educacional funcione bien y logre asegurar un nivel mínimo de contenidos para todos los alumnos.

Capítulo 3

Sistema Legal

El Sistema Educacional Chileno está conformado por cuatro niveles de enseñanza escolar: preescolar, básica, secundaria y superior. La entidad máxima en materia educativa, y quien se encarga de hacer cumplir las leyes y normas, es el Ministerio de Educación. La ley que actualmente rige las prácticas educativas es la Ley General de Educación (LGE), promulgada el año 2009, cuya antecesora fue la Ley Orgánica Constitucional de Enseñanza (LOCE). Tanto en la Constitución como en el corazón de la LGE se reconoce el derecho a la educación y a la libertad de enseñanza, temas trascendentales a los que debe apuntar el sistema educacional chileno. Hoy en día, a pesar de tener una educación básica con altas tasas de cobertura¹, aún no es posible mermar del todo temas relacionados con la eficacia y calidad del sistema. Urquiola & Calderón (2006) dejan de manifiesto la discrepancia que existe entre los años que un estudiante pasa en el colegio versus los años de escolaridad que efectivamente adquiere al cumplir 18 años de edad. Chile tiene un buen desempeño en comparación con los países de América Latina y el Caribe, si analizamos a un individuo de 18 años, éste en promedio pasa 12,1 años en el colegio y consigue 10,4 años de escolaridad, ambas cifras son las más elevadas a nivel de Latinoamérica y el Caribe. Para el caso de Honduras² las cifras son más desalentadoras, en promedio un individuo de 18 años pasa 8,6 años en el colegio y logra sólo 6,2 años de escolaridad. Lo anterior nos llevaría a pensar en que si el objetivo de Chile es reducir la brecha entre años de escolaridad y años en el colegio, aún tiene que tratar materias importantes tal como la repitencia escolar.

¹En el año 2008, la Tasa Neta de matrícula en educación básica fue de 95 % para hombres, y 94,2 % para mujeres (Valladares, 2008).

²En Urquiola & Calderón (2006) comparan sistemáticamente a Chile con Honduras debido a su situación diametralmente distinta en materia educativa.

Si nos enfocamos en la repitencia escolar, ya desde la década de los 80 el marco legal que regulaba la evaluación y promoción escolar, tanto en educación básica como media, validaba la retención como práctica común frente a situaciones en que los alumnos no presentaban un “buen” desempeño académico o una alta tasa de asistencia. En el año 1988, el Decreto-146 Exento³ reconocía la promoción automática de los estudiantes de 1ro básico siempre y cuando el director del establecimiento no creyera necesario lo contrario, dejando abierta la definición de lo que el director podía considerar un caso “debidamente calificado”. No obstante, para estudiantes de mayores grados las causas de reprobación eran incluso más exigentes, en caso de obtener calificaciones insuficientes en un ramo sólo se optaba a la posibilidad de aprobar el año escolar si es que el alumno presentaba un promedio general, incluyendo el ramo con calificación insuficiente, igual o mayor a 4,5. Sumado a esta situación, el marco legal de ese entonces no permitía hacer una separación de las leyes que regulaban la promoción escolar en educación básica versus educación media, asumiendo ex-ante que tanto factores psicológicos como sociales eran comunes y se reproducían con la misma intensidad en estos dos grupos etarios.

En el año 1997 se introdujo un cambio que apuntaba en la dirección de regular con distintos decretos la promoción y evaluación de los estudiantes de enseñanza básica y media, este cambio iba de la mano de nuevos lineamientos curriculares que habían entrado en vigencia el año 1996. En 1997 se aprobó el Decreto-511 Exento⁴, válido para la educación básica⁵, que daba un tiempo de dos años a la aplicación de los contenidos curriculares de las dos primeras fases educativas, es decir, los alumnos tenían hasta 2do básico para aprender los objetivos fundamentales de la primera etapa del ciclo, luego desde 3ro a 4to básico para demostrar que manejaban los contenidos mínimos obligatorios de la siguiente etapa. De esta forma, la única causa de repetición desde 1ro a 2do básico o desde 3ro a 4to básico descansaba en motivos de asistencia (85 % o menos), pero la ley también estipulaba que podían darse situaciones en que un alumno presentaba un porcentaje de asistencia menor y tanto el director del establecimiento como el profesor jefe del estudiante creían y justificaban debidamente su aprobación. En el año 2003 hubo una modificación en el Decreto-511 Exento, la cual avalaba la práctica de no autorizar la promoción de alum-

³Ministerio de Educación (1988) “Reglamento de evaluación y promoción de alumnos de Educación General Básica, de Educación Media, de Educación de Adultos y proceso de titulación en Educación Técnico-Profesional”.

⁴Ministerio de Educación (1997) “Aprueba reglamento de evaluación y promoción escolar de niñas y niños de Enseñanza Básica”.

⁵El Decreto-112 Exento (Ministerio de Educación (1999) “Establece disposiciones para que establecimientos educacionales elaboren reglamento de evaluación y reglamenta promoción de alumnos de 1ro y 2do año de Enseñanza Media, ambas modalidades”) y Decreto-83 Exento (Ministerio de Educación (2001) “Reglamenta calificación y promoción de alumnos (as) de 3ro y 4to año de enseñanza media, ambas modalidades, y establece disposiciones para que los establecimientos educacionales elaboren su reglamento de evaluación”) son los que regulan la promoción y evaluación en Educación Media.

nos de 1ro a 2do básico o de 3ro a 4to básico si es que el director del establecimiento, previo diálogo con el profesor jefe, encontraba que el estudiante presentaba un retraso significativo en lectura, escritura y/o matemáticas, todo esto bajo los parámetros del programa de estudios que aplicaba dicho establecimiento.

Finalmente, en materia de 2do, 4to y hasta 8vo básico, y en donde están contenidos los grados que son relevantes para esta investigación, el Decreto-511 Exento define cuatro mecanismos mediante los cuales un estudiante pudiese reprobar el año escolar: primero, presentar un porcentaje de asistencia menor al 85 %⁶; segundo, reprobar una asignatura y no obtener un promedio general, incluyéndose en el promedio el ramo reprobado, igual o superior a 4,5; tercero, reprobar dos asignaturas y no obtener un promedio general, incluyéndose en el promedio los ramos reprobados, igual o superior a 5; y cuarto, reprobar tres o más asignaturas sin consideración del promedio general. El objetivo de esta investigación es explotar estas reglas de promoción para medir el efecto que tiene la repitencia escolar en la cantidad de años de escolaridad alcanzados por el estudiante, entendiéndose años de escolaridad en el sentido en que Urquiola & Calderón (2006) lo definen: máximo grado educacional alcanzado en un año (o edad) determinado(a). En la siguiente sección explicaremos la metodología desarrollada en esta investigación, poniéndose especial énfasis en los problemas y variables a utilizar.

⁶El Director del establecimiento y el Profesor Jefe podrán autorizar la promoción de los alumnos de 2do a 3ro y de 4to a 5to año, con porcentajes menores de asistencia. Desde 5to a 8vo año la autorización será tomada por el Consejo de Profesores.

Capítulo 4

Metodología

Para testear el impacto que tiene la repitencia sobre subsecuentes resultados académicos utilizaremos un diseño metodológico conocido como “Regresión Discontinua Fuzzy” (RD Fuzzy). Como Lee & Card (2008) señalan, “[...] Regresión Discontinua es apropiado para evaluación de programas en donde el estado del tratamiento (o probabilidad del tratamiento) depende de si una covariable observable excede un umbral fijo [...]”, por lo tanto la idea es comparar al grupo de personas que justo fueron afectadas por la regla de promoción escolar con aquellas que justo no fueron afectadas. El adjetivo Fuzzy responde a que la probabilidad de recibir el tratamiento no es un cambio discreto de 0 a 1 en el umbral, es decir, no depende determinísticamente de si una predictor toma valores por sobre cierto límite, sino más bien es un pequeño salto que depende *parcialmente* de si ese predictor supera cierto umbral (Lee & Lemieux, 2010; Imbens & Lemieux, 2008; Lee & Card, 2008; Imbens & Wooldridge, 2009).

La idea subyacente de RD Fuzzy es que el resultado promedio para individuos justo bajo el umbral (grupo de control) sea un contrafactual apropiado para el grupo de tratamiento (justo por sobre el umbral). De esta forma lo que se quiere lograr es que cercano al límite todas las covariables realizadas antes de la asignación sean independientes del estado del tratamiento. Por lo tanto, al igual que un experimento aleatorio, las diferencias en los resultados post-asignación no serán confundidos por variables omitidas (Lee, 2008). En nuestro caso RD Fuzzy nos permite utilizar la condición de cumplir cierta regla de promoción escolar como variable instrumental a la probabilidad de repetir del alumno, pudiendo cuantificar el efecto que tiene esta última sobre resultados académicos posteriores, siguiendo el espíritu en que Manacorda (2012) desarrolló su investigación.

En general la ventaja de esta metodología es que permite instrumentalizar la probabilidad de repetir, eliminando potenciales sesgos provenientes de la relación entre variables no observables y repitencia, y resultando en estimadores más robustos. Por otra parte, su desventaja radica en que el efecto encontrado correspondería a un Local Average Treatment Effect (LATE) alrededor del punto de discontinuidad, esto quiere decir que el estimador será el efecto del tratamiento promedio para quienes sean inducidos a participar (compliers) una vez que la regla de promoción escolar (variable instrumental) cambie. Por lo tanto, la interpretación de LATE no es independiente de la variable instrumental, y acorde a la elección de ésta los resultados cambiarán (Wooldridge, 2010; Imbens & Wooldridge, 2009)¹.

Metodología

La metodología propuesta para esta investigación descansa en el corazón de Manacorda (2012, 2010, 2006). Como se señaló en el capítulo “Revisión Bibliográfica”, el objetivo de Manacorda fue cuantificar el efecto que tenía la repitencia escolar, en el contexto del sistema educacional uruguayo, sobre posteriores logros académicos. Su estrategia de identificación fue RD Fuzzy, tomando ventaja de las reglas de promoción establecidas en Uruguay. En una primera instancia exploró la relación entre número de días faltados y condición de repitencia, pero debido a los débiles resultados encontrados la estrategia de identificación cambió en versiones posteriores, pasando a ser la cantidad de asignaturas reprobadas la nueva variable instrumental. Sus principales resultados apuntaron a que la repitencia llevaba a una mayor tasa de deserción escolar y a menores años de escolaridad.

Ahora bien, Y_i es la variable de resultado, que para términos de esta investigación será la cantidad máxima de grados alcanzados por el estudiante al cumplir 6 años de escolaridad adicional (también podríamos pensar en variables como desempeño en SIMCE, probabilidad de desertar, cantidad de años en la educación formal, entre otros). Para efectos de construcción de la variable de resultado no es necesario la permanencia completa y constante de un estudiante en el periodo analizado, esto quiere decir que un alumno puede desertar, asistir intermitentemente o “desaparecer” del sistema escolar chileno no siendo eliminado de la muestra a analizar, sino que se mantiene la observación y se imputa la cantidad de grados alcanzados al último año de registro.

¹Si se cumple la idea de que mantener una tasa de repitencia baja, de estado estacionario, logra que el sistema educacional funcione bien y asegure un nivel mínimo de contenidos para todos los alumnos, entonces un estimador LATE no capturaría esta historia. Por lo tanto, la interpretación de los resultados y la propuesta de política no debe olvidar los efectos heterogéneos que podrían presentarse alejados del punto de discontinuidad.

Supongamos que Y_i depende linealmente de R_i , que es una variable indicador igual a uno si el individuo i repitió en el año 2007 y cero en otro caso, además de otras características observables X_i que inciden en la variable de resultado², y de un término de error e_i . Tal como se muestra a continuación, la ecuación a estimar sería:

$$Y_i = \beta_0 + \beta_1 R_i + \beta_2 X_i + e_i \quad \forall i \quad (1)$$

Es lógico pensar que la Ecuación (1) presenta un problema de variable omitida relevante, esto ya que el desempeño académico anterior así como otras características no observables del estudiante podrían ayudar a explicar la variable de resultado, no obstante no están siendo controladas. Por lo tanto, si es que no se cumpliera la condición de ortogonalidad entre R_i y e_i , el estimador MCO de β_1 sería sesgado, llevando a conclusiones incorrectas por parte de la presente investigación.

Para poder resolver el sesgo del coeficiente de interés (β_1) se sugiere utilizar la discontinuidad en la cantidad de asignaturas reprobadas en el año 2007 (A_i)³, así como controlar por todas las variables observables, con el fin de instrumentalizar la probabilidad de repetir, R_i de la Ecuación (1) (Wooldridge, 2010). De esta forma asumiremos que la condición de repitencia depende de una variable indicador que toma valor 1 en caso que se haya traspasado el umbral de cursos reprobados ($U_i = I(A_i \geq 2)$) y 0 en caso contrario⁴, una función continua $z_i(A_i - 1)$ cuyo argumento es la cantidad de ramos reprobados normalizados a 0 en el punto de la discontinuidad, de variables de control X_i , interacciones entre U_i y A_i , y de un término de error η_i ⁵.

$$R_i = \gamma_0 + \gamma_1 U_i + \gamma_2^1 z_i(A_i - 1) + \dots + \gamma_2^n z_i^n(A_i - 1) + \gamma_3 X_i + \phi_1 U_i \cdot (A_i - 1) + \phi_2 U_i \cdot (A_i - 1)^2 + \dots + \phi_n U_i \cdot (A_i - 1)^n + \eta_i \quad \forall i \text{ con } z_i(0) = 0 \quad (2)$$

Con la Ecuación (2) se instrumentaliza R_i , que representa la condición de repitencia en la Ecuación

²En el Apéndice 1 se encuentran las definiciones de las variables de control y variable de resultado a utilizar.

³En este contexto A_i sería la variable de asignación, la que en inglés es referida como “forcing variable” o “assignment variable”.

⁴Como las reglas de repitencia dependen de la cantidad de asignaturas reprobadas en conjunto con el promedio general, trabajaremos sobre la base de un límite que sea independiente de las condiciones sobre promedio general. Por lo mismo definiremos a una persona que cumple la condición de reprobación como aquella que presenta dos o más asignaturas reprobadas.

⁵Tanto los errores de la Ecuación (2) como de la Ecuación (3) deben ir agrupados en función de la variable de asignación normalizada ($A_i - 1$) (Lee & Card (2008)).

(1). Ahora la ecuación de la segunda etapa quedaría:

$$Y_i = \theta_0 + \theta_1 \hat{R}_i + \theta_2^1 z_i(A_i - 1) + \dots + \theta_2^n z_i^n(A_i - 1) + \theta_3 X_i + \pi_1 U_i \cdot (A_i - 1) + \pi_2 U_i \cdot (A_i - 1)^2 + \dots + \pi_n U_i \cdot (A_i - 1)^n + \mu_i \quad \forall i \text{ con } z_i(0) = 0 \quad (3)$$

El coeficiente de interés θ_1 es interpretado como la diferencia promedio en logros educacionales entre estudiantes que repitieron versus los que aprobaron, local al umbral de la variable de asignación; θ_2 es el parámetro asociado al efecto que tiene la variable de asignación sobre Y_i ; y finalmente, θ_3 es el estimador relacionado a los efectos de las variables de control.

Capítulo 5

Datos

Como hemos mencionado, el objetivo de este trabajo es analizar el efecto causal que tiene la repitencia en posteriores logros académicos. La imposibilidad de medir los efectos en el aprendizaje a través del SIMCE radica en que necesitamos un indicador estándar y homogéneo tanto para estudiantes repitentes como para no repitentes, pero bajo el actual sistema de medición chileno no es posible obtenerlo ya que las evaluaciones son aplicadas a estudiantes de un mismo grado y no de una misma edad, por lo tanto el efecto de la repitencia se ensuciaría con diferencias intrínsecas provenientes de pruebas estandarizadas medidas en diferentes años. En el caso del SIMCE la evaluación del grupo de tratamiento será siempre con un año de desfase debido al retraso per se de la población repitente. Por este motivo buscamos otra variable de resultado que no presente este tipo de problemas, bajo este escenario nuestra mejor alternativa fue utilizar la cantidad máxima de grados alcanzados por el estudiante al cumplir 6 años de escolaridad adicional, ya que permitiría responder a los requisitos que amerita la metodología.

Para el logro del objetivo de esta investigación se examinará a estudiantes que cursaban 4to, 5to o 6to básico en el año 2007, y que como condición necesaria no hubiesen repetido ningún curso entre el periodo que comprendía desde 1ro básico hasta el grado en el que se encontraban el año 2007. La razón para no considerar a estudiantes que estuviesen cursando grados menores fue principalmente legal, ya que los primeros ciclos básicos presentan distintas normativas de repitencia, lo que complicaría el análisis. En contraste, el motivo para no agregar a la muestra individuos que se encontraban cursando niveles superiores a 6to básico se debe meramente a tener un periodo final unificado. La muestra es tal que para el último año analizado aún permanecerán en el sistema educacional los tres cursos estudiados lo que lograría que la variable de resultado tome el mismo rango de valores en cada uno de los casos.

En esta investigación las principales fuentes de información a considerar son el Sistema de Información General de Estudiantes (SIGE)¹ y el cuestionario de padres del SIMCE, ambos puestos a disposición por el Ministerio de Educación de Chile, el que separa los registros en bases de datos por año. El SIGE contiene información acerca de la edad del estudiante, género, colegio al que asistió, tipo de dependencia del establecimiento, curso en el que se encontraba cada año, promedio de notas, porcentaje de asistencia, si pasó o reprobó el año escolar y si necesitó educación especial, hasta el año 2015. Adicionalmente, el SIGE está compuesto por datos de todo el universo de estudiantes del país, incluyendo colegios municipales, particulares subvencionados y particulares pagados, además de instituciones que imparten educación escolar para personas adultas. El cuestionario de padres del SIMCE es preguntado a los apoderados de cada estudiante evaluado, su finalidad es indagar sobre el nivel educacional de los padres, ingreso del hogar y nivel de satisfacción con el establecimiento, entre otros.

Para efectos de este análisis eliminaremos aquellas observaciones en que los alumnos ya hubiesen recibido el tratamiento (repetencia) con anterioridad al año 2007, ya que nuestro objetivo será medir el efecto que tiene la *primera* repetencia sobre la cantidad máxima de grados alcanzados al cumplir 6 años de escolaridad adicional. Además, sólo se permitirá que los estudiantes avancen grados en establecimientos educacionales que imparten clases para niños, excluyendo de todo análisis los colegios para adultos.

Teniendo en consideración que la investigación analizará el efecto que tiene repetir en 4to, 5to o 6to básico en el año escolar 2007, aquel estudiante que haya recibido el tratamiento en dicho año cargará con un grado escolar menos de por vida, lo que lo pondría en una situación de desventaja considerando la variable de resultado a utilizar. Por ende, lo que se ha querido hacer es cuantificar el efecto *adicional* que tiene reprobado, independiente de su año de atraso per se, lo que llevaría a que el periodo final de estudio sea diferente para estudiantes repitentes versus no repitentes. En el caso de los alumnos repitentes, se seguirá a los estudiantes desde 1ro básico hasta el grado alcanzado en el año 2014²; y para los promovidos el periodo final de análisis será el año 2013, de manera tal que la cantidad máxima potencial de grados alcanzados al cumplir 6 años de escolaridad adicional sea comparable entre el grupo de control y el de tratamiento.

¹Desde el año 2002 a 2008 la información de rendimiento de los estudiantes se encontraba compilada en el Registro de Estudiantes de Chile (RECH). Desde 2009 a la actualidad la fuente de información es el SIGE, que pasó a reemplazar al RECH.

²El periodo final elegido se debe meramente a tener un rango de valores unificado para la variable de resultado en cada uno de los niveles educacionales estudiados (4to, 5to y 6to básico).

A continuación presentaremos una breve estadística descriptiva para la muestra bajo análisis. En el Cuadro 2 se puede apreciar a simple vista que los repitentes presentan una posición de deficiencia en cuanto a atributos socioeconómicos se habla, esto ya que para ellos tanto escolaridad de la madre y padre como rango de número de libros en el hogar e ingreso son, en promedio, menores que para el grupo de estudiantes no repitentes, siendo las diferencias estadísticamente significativas a menos del 1%. Por otra parte, si hablamos de variables a nivel de escuela destacaría “Efecto Par”, que es aquel que hace alusión al Nivel Socioeconómico (NSE) promedio de los compañeros de curso de un estudiante en el año 2007. En el Cuadro 2 podemos ver que un estudiante repitente presenta un menor Efecto Par versus un estudiante promovido, esto indicaría que, en promedio, los estudiantes repitentes se desenvuelven en establecimientos educacionales cuyos compañeros de curso presentaban un menor NSE versus estudiantes que no repitieron, significativo al 1%.

Cuadro 2: Estadística descriptiva de estudiantes en 4to, 5to y 6to básico en 2007.

	Año 2007		
	No repitentes	Repitentes	Diferencia
Género (% Hombres)	0,49 (0,5)	0,65 (0,477)	-0,15** (0,004)
Edad	10,64 (0,96)	10,78 (0,96)	-0,13** (0,01)
Dist. Grado-Edad	-0,63 (0,52)	-0,65 (0,58)	0,02** (0,01)
Escolaridad Madre	11,50 (3,6)	9,80 (3,58)	1,7** (0,03)
Escolaridad Padre	11,54 (3,75)	9,94 (3,70)	1,6** (0,03)
No de libros:			
-Ningún libro	0,07 (0,25)	0,12 (0,32)	-0,05** (0,003)
-Menos de 10	0,38 (0,49)	0,48 (0,5)	-0,1** (0,004)
-Entre 11 y 50	0,38 (0,48)	0,31 (0,46)	0,07** (0,004)
-Más de 50	0,18 (0,38)	0,1 (0,29)	0,08** (0,002)
Ingreso	383.243 (447.313)	242.508 (282.894)	141.735** (2.282)
NSE	0,09 (0,9)	-0,34 (0,82)	0,44** (0,01)
Efecto Par	0,07 (0,69)	-0,14 (0,54)	0,21** (0,004)
Cursos Reprobados	0,08 (0,29)	2,91 (1,17)	-2,83** (0,01)
Grados Adicionales	5,31 (0,85)	4,28 (1,62)	1,03** (0,01)
Obs.	594.961	16.422	611.383

** $p < 0,01$, * $p < 0,05$

Algo interesante a mencionar del Cuadro 2 es que el porcentaje de hombres es mayor en el grupo de repitentes y también es significativo a menos del 1%, por lo que podríamos intuir que un análisis que implicase separar hombres y mujeres sería adecuado en este contexto³. Por otra parte, si hablamos de

³Si estudiantes hombres repiten en mayor medida que estudiantes mujeres, quizás hombres y mujeres no son tan parecidos en su conducta y esfuerzo escolar, existiendo factores claves que estén desencadenando esta situación dispar, y que para efectos

la variable “edad” los datos observados dejarían de manifiesto que los repitentes tienen, en promedio, levemente más edad que los promovidos, y si analizamos la distorsión Grado-Edad podemos ver que ambos grupos presentan una entrada tardía al sistema escolar, pudiendo ser un poco más notoria y estadísticamente significativa a niveles convencionales en el grupo de repitentes. Finalmente, se encuentran la variable de asignación y variable de resultado: si analizamos la cantidad de asignaturas con promedio menor a 4,0, vemos que en esta muestra los estudiantes repitentes presentan, en promedio, un mayor número de asignaturas reprobadas que los alumnos no repitentes, lo que daría indicios de una discontinuidad útil en esta variable; y por el lado de la cantidad de grados adicionales, la primera impresión sería que la repitencia trae consigo menos grados adicionales alcanzados, lo que sería una conclusión apresurada ya que se estaría ignorando que el estimador de RD Fuzzy es un LATE, por lo que al comparar los promedios simples no se estaría capturando el efecto en torno al umbral⁴.

Ahondando en que la proporción de hombres no se distribuye de igual manera entre el grupo de tratamiento y de control, y basados en la hipótesis de que las variables que caracterizan el comportamiento de hombres y mujeres quizás no son tan similares, realizaremos una estadística descriptiva separando repitentes de no repitentes, con el fin de comparar a hombres y mujeres en cada subgrupo. En esta línea podemos apreciar que en el Cuadro 3, del universo de alumnos repitentes, las diferencias entre hombres y mujeres se presentan en todas las variables de control, pero no todas son estadísticamente significativas a al menos el 5%. Por ejemplo, la variable que utilizamos como proxy de capital cultural del hogar, número de libros, no presenta diferencias significativas en ninguno de sus niveles, evidenciando que si analizamos esta dimensión entre hombres y mujeres repitentes, estos se comportarían de manera muy parecida. La misma historia se repite si consideramos variables como escolaridad del padre, ingreso y NSE, las diferencias apuntan a una leve situación de desventaja para el subgrupo de mujeres, no siendo estadísticamente significativa a niveles convencionales. Finalmente, los valores asociados a edad, distorsión Grado-Edad, escolaridad de la madre y Efecto Par sí presentan diferencias significativas a menos del 1% entre hombres y mujeres. Aquí las diferencias van principalmente a favor del grupo de mujeres: las mujeres presentan, en promedio, menos edad, menos retraso a la entrada y mayor Efecto Par. No obstante,

de esta investigación merecen ser identificados.

⁴Como se señaló en la Sección “Metodología”, la variable de asignación es la cantidad de asignaturas reprobadas estableciendo la discontinuidad en una asignatura reprobada, por lo que un análisis rápido y más acertado sería comparar al grupo de repitentes con dos asignaturas reprobadas versus el grupo de no repitentes con una asignatura reprobada. En el Apéndice 2 se encuentra la estadística descriptiva que permite realizar este cálculo para toda la muestra, a simple vista la diferencia entre estos dos grupos es pequeña, por lo que es necesario realizar la metodología paso a paso para obtener un efecto de la repitencia más preciso y fiable, bajo las condiciones antes mencionadas.

las madres de los hombres repitentes presentan, en promedio, más escolaridad que sus pares en el grupo de mujeres repitentes. Otro punto importante a analizar en el Cuadro 3 se relaciona con las variables de asignación y de resultado. Como podemos ver, ambas variables presentan un escenario mejor para mujeres repitentes versus hombres repitentes: las mujeres repitentes presentan, en promedio, menos cursos reprobados y alcanzan más grados adicionales, ambas diferencias estadísticamente significativas a menos del 1%.

Cuadro 3: Estadística descriptiva de estudiantes repitentes que cursaban 4to, 5to y/o 6to básico en 2007.

	Repitentes Año 2007		
	Mujeres	Hombres	Diferencia
Edad	10,73 (0,97)	10,80 (0,95)	-0,07** (0,02)
Dist. Grado-Edad	-0,63 (0,58)	-0,66 (0,59)	0,03** (0,01)
Escolaridad Madre	9,66 (3,58)	9,88 (3,58)	-0,224** (0,06)
Escolaridad Padre	9,94 (3,69)	9,95 (3,71)	-0,006 (0,06)
No de libros:			
-Ningún libro	0,12 (0,32)	0,12 (0,32)	0,0003 (0,01)
-Menos de 10	0,48 (0,5)	0,48 (0,5)	-0,001 (0,01)
-Entre 11 y 50	0,31 (0,46)	0,31 (0,46)	-0,01 (0,01)
-Más de 50	0,1 (0,3)	0,09 (0,29)	0,01 (0,01)
Ingreso	239.542 (282.693)	242.568 (283.010)	-3.026 (4.626)
NSE	-0,36 (0,82)	-0,33 (0,82)	-0,03 (0,01)
Efecto Par	-0,12 (0,54)	-0,15 (0,54)	0,03** (0,01)
Cursos Reprobados	2,8 (1,1)	2,97 (1,21)	-0,17** (0,02)
Grados Adicionales	4,39 (1,6)	4,22 (1,63)	0,18** (0,03)
Obs.	5.752	10.670	16.422

** $p < 0,01$, * $p < 0,05$

Pero si nuestra estadística descriptiva se basara sólo en la muestra de estudiantes promovidos, ¿nuestras conclusiones se mantendrían inalteradas? La respuesta es no. Como podemos apreciar en el Cuadro 4, gran parte de las diferencias entre hombres y mujeres son significativas a menos del 1%. Escolaridad de la madre, escolaridad del padre e ingreso son las variables que presentan diferencias con mayores magnitudes, todas en desmedro del subgrupo de mujeres. En general, si comparamos a las mujeres promovidas con los hombres promovidos podemos ver que ellas gozan de menor capital cultural, esto se ve en la mayor proporción promedio en grupos con menos de 10 libros y en la menor proporción promedio en grupos con más de 11 libros. Ahora bien, si analizamos las variables de asignación y de resultado las conclusiones son positivas para el grupo de mujeres, por un lado ellas logran, en promedio, menor número

ro de cursos reprobados, y por otro lado ellas logran avanzar más grados adicionales que los hombres promovidos, estas diferencias son significativas a al menos el 1 %.

Cuadro 4: Estadística descriptiva de estudiantes no repitentes que cursaban 4to, 5to y 6to básico en 2007.

	No repitentes Año 2007		
	Mujeres	Hombres	Diferencia
Edad	10,64 (0,97)	10,65 (0,96)	-0,01** (0,003)
Dist. Grado-Edad	-0,62 (0,52)	-0,64 (0,52)	0,02** (0,001)
Escolaridad Madre	11,43 (3,62)	11,58 (3,57)	-0,151** (0,01)
Escolaridad Padre	11,48 (3,77)	11,61 (3,73)	-0,13** (0,01)
No de libros:			
-Ningún libro	0,07 (0,25)	0,07 (0,25)	0,004** (0,001)
-Menos de 10	0,38 (0,49)	0,38 (0,48)	0,01** (0,001)
-Entre 11 y 50	0,37 (0,48)	0,38 (0,49)	-0,01** (0,001)
-Más de 50	0,18 (0,38)	0,18 (0,38)	-0,001 (0,001)
Ingreso	377.718 (445.213)	388.951 (449.402)	-11.233** (1.160)
NSE	0,07 (0,91)	0,11 (0,9)	-0,04** (0,002)
Efecto Par	0,06 (0,69)	0,07 (0,7)	-0,01** (0,002)
Cursos Reprobados	0,07 (0,27)	0,09 (0,31)	-0,02** (0,001)
Grados Adicionales	5,35 (0,82)	5,27 (0,87)	0,08** (0,002)
Obs.	302.307	292.654	594.961

** $p < 0,01$, * $p < 0,05$

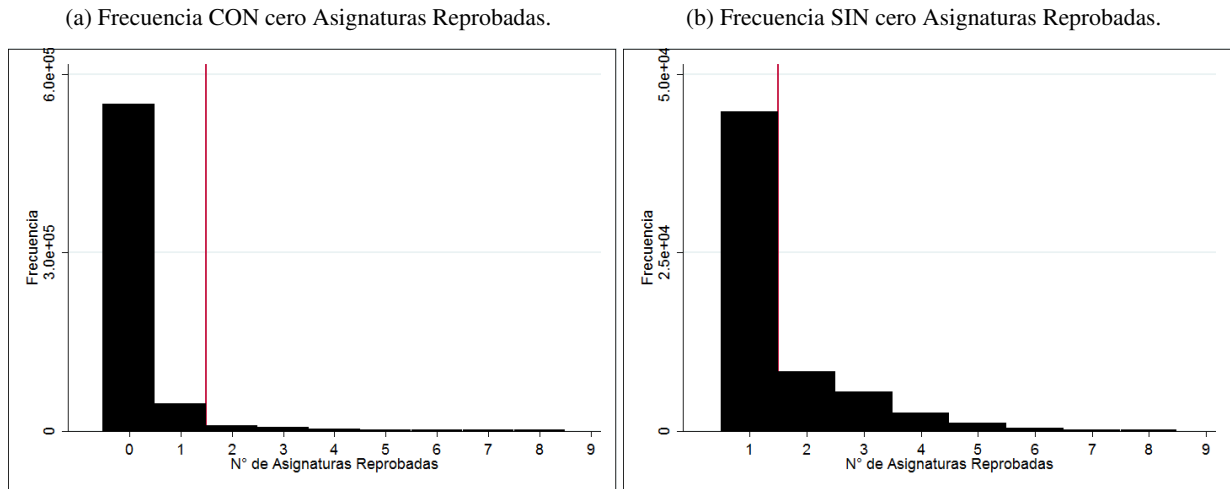
Capítulo 6

Identificación

“¿Son capaces los individuos de influir en la variable de asignación? Y si lo hacen, ¿cuál es la naturaleza de ese control?”, con estas preguntas Lee & Lemieux (2010) analizan la validez de la metodología RD. En este contexto los autores señalan que debemos tener total seguridad de que los individuos bajo análisis no gocen de control completo o “preciso” de la variable de asignación, por el contrario si es que esto llegase a ocurrir, entonces estaríamos en presencia de una auto-selección, lo que invalidaría su aplicabilidad.

Si creemos que los individuos tienen control “preciso” sobre el número de asignaturas reprobadas, y entendiendo además que estudiar conlleva a costos monetarios y no monetarios, entonces la mejor respuesta de los estudiantes sería realizar un esfuerzo tal que los situara justo a la izquierda del umbral ($A = 1$, sin normalizar). Como observamos en la Figura 1 esto no ocurre, aquí vemos que existen personas con dos y hasta ocho asignaturas reprobadas. En el panel (a) se aprecia que la mayor parte de los individuos no presentan asignaturas reprobadas para el año 2007 (89,8%), y que la frecuencia de la variable de asignación cae monótonicamente a medida que aumenta su valor. En el panel (b) se grafica la frecuencia de la variable de asignación, pero esta vez excluyendo a las personas que no presentan asignaturas reprobadas: aproximadamente el 28,5% del total de la muestra analizada, excluyendo a aquellos estudiantes que no poseen ramos deficientes, reprobó dos o más asignaturas el año 2007. Lo anterior da cuenta de un control “impreciso” por parte de los agentes, ya que de lo contrario no veríamos este comportamiento en la variable de asignación, es decir, no existirían personas que se situaran a la derecha del umbral dado que lo óptimo sería estar justo a la izquierda del límite.

Figura 1: Frecuencia de Variable de Asignación



Ahora bien, otro tema importante es ver si existe alguna discontinuidad clara y significativa entre la variable de asignación y la probabilidad de repitencia, de esta forma podemos validar gráficamente al número de asignaturas reprobadas en el año 2007 como variable instrumental¹. Un análisis preliminar informativo es realizar una inspección visual de la relación entre estas dos variables, esto es lo que recomiendan Imbens & Lemieux (2008).

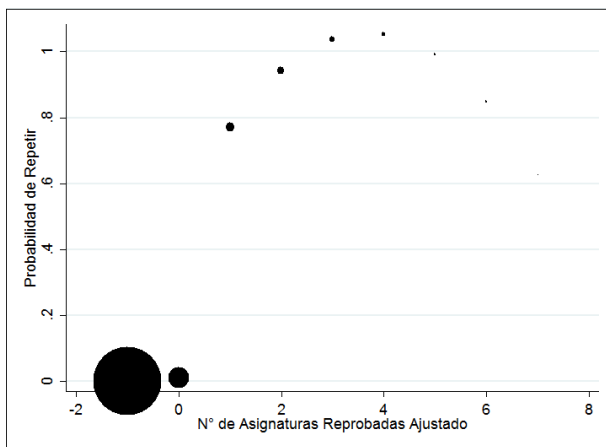


Figura 2: Probabilidad de repetir en función del Número de asignaturas reprobadas ajustado en Alumnos de 4to, 5to y 6to Básico, 2007. Se asume una relación cuadrática entre repitencia y variable de asignación. El tamaño de cada punto es proporcional a la cantidad de observaciones en cada grupo.

En la Figura 2 podemos ver la probabilidad de repetir por primera vez en el año 2007 en función del número de asignaturas reprobadas, tal como se muestra en la Ecuación (2), pero considerando una forma

¹En Wooldridge (2010) se señala que para aplicar la metodología de variable instrumental el instrumento debe cumplir dos condiciones: primero, no estar correlacionado con los errores de la primera etapa; segundo, explicar parcial, pero significativamente la variable a instrumentalizar.

funcional cuadrática y agregando efectos fijos por colegio y grado². Podemos notar que el dominio de la variable de asignación va desde -1 a 8, dado que ésta está normalizada a 0 en el punto de discontinuidad. La idea principal es que como el efecto que capturará la metodología será un LATE en torno al límite, las personas cuya variable de asignación normalizada sea igual a 0 pertenecerán al grupo de control, y aquellas cuyo valor de la variable de asignación normalizada sea igual a 1 deberán pertenecer al grupo de tratamiento. En la Figura 2 podemos apreciar que la probabilidad de repetir efectivamente es una función creciente de la variable de asignación, produciéndose un salto en el umbral en el punto en que el número de asignaturas reprobadas es normalizado a 0. Por lo tanto, podemos concluir que pasar de estar en el grupo de control a estar en el grupo de tratamiento sí tiene asociado un aumento en la probabilidad de repetir. Específicamente, la probabilidad de repetir para aquellos alumnos con menos de 2 asignaturas reprobadas es de 1,13 %, y luego salta a 77,06 % para aquellos que alcanzaron dos o más asignaturas deficientes. La discontinuidad generada en $A = 0$, normalizado, valida a la cantidad de asignaturas reprobadas como variable instrumental³.

Supuestos que se deben cumplir

En RD existen dos supuestos claves que se deben cumplir (Bellei, 2010; Lee & Lemieux, 2010; Imbens & Lemieux, 2008): primero, se debe especificar correctamente la forma funcional que relacione a la variable de resultado con la variable de asignación (número de asignaturas reprobadas normalizada, función $z(A - 1)$ en Ecuación (2) y (3)); segundo, todas las características observables (y no observables) de los individuos deben variar continuamente alrededor del umbral.

Forma Funcional

Dado que el soporte de nuestra variable de asignación es discreto, no es posible computar promedios arbitrariamente pequeños alrededor del punto límite. En contraste, debemos escoger una forma funcional paramétrica para modelar la relación entre variable de resultado y variable de asignación (Lee & Card, 2008). En este apartado analizaremos la correcta especificación de la forma funcional utilizando el modelo

²La elección del grado del polinomio será abordado con mayor profundidad en el avance del presente capítulo.

³Al realizar los test estándar para analizar si la variable escogida es una buena variable instrumental los resultados también fueron satisfactorios, se rechaza la hipótesis nula tanto en el test de sobreidentificación como en el de instrumentos débiles.

reducido de la metodología RD, y siguiendo la lógica presentada en la Ecuación (4).

$$Y_i = \alpha_0 + \alpha_1 U_i + \alpha_2^1 z_i (A_i - 1) + \dots + \alpha_2^n z_i^n (A_i - 1) + \alpha_3 X_i + \lambda_1 U_i \cdot (A_i - 1) + \lambda_2 U_i \cdot (A_i - 1)^2 + \dots + \lambda_n U_i \cdot (A_i - 1)^n + v_i \text{ con } z_i(0) = 0 \quad \forall i \quad (4)$$

La idea es escoger aquella forma funcional que conjugue de mejor manera tanto la significancia estadística de los parámetros como el criterio de información (AIC) (Lee & Card, 2008; Lee & Lemieux, 2010; Bellei, 2010), pero teniendo siempre una clara inclinación hacia polinomios de menores grados, ya que como señalan Gelman & Imbens (2014), “los resultados basados en regresiones polinomiales de alto grado son sensibles al orden del polinomio [...] Además, la inferencia basada en polinomios de alto grado frecuentemente tiene un peor desempeño”.

Considerando “Grados Adicionales” como variable de resultado, el Cuadro 5 resume la principal información a analizar. Aquí la variable que irá cambiando de grado será el número de asignaturas reprobadas normalizada, por lo tanto “Asignación²” indicará que la variable de asignación fue elevada a 2, “Interacción²” indicará que la variable de asignación fue elevada a 2 y luego multiplicada por la condición de repetencia ($U_i = I(A_i \geq 2)$), y la misma lógica pero elevadas a distintos exponentes seguirán “Asignación³”, “Asignación⁴”, “Interacción³” e “Interacción⁴”. La idea es ir ajustando diferentes relaciones entre variable de asignación y variable de resultado.

En la columna (1) del Cuadro 5 podemos ver que los coeficientes asociados a ajustar la relación por un polinomio de primer grado son significativos al 1%, implicando que la variable de asignación tenga un efecto neto negativo sobre la cantidad de grados adicionales⁴. Además vemos que el R^2 ajustado de la regresión es igual a 33,44% y el valor del criterio de información Akaike asciende a 1.339.174. Ahora bien, si comparamos los resultados de la columna (1) con los de la columna (2) podemos detectar pistas que apuntan a que agregar una relación cuadrática efectivamente mejora los resultados del modelo. Por ejemplo, vemos que tanto las nuevas como las anteriores son significativas al 1%, con excepción de $I(\text{Asig. Reprobadas} \geq 2)$. Además, podemos apreciar que el R^2 ajustado se mantiene y que AIC cae en nivel 3 puntos. Dado lo anterior, y basados principalmente en la significancia estadística de los paráme-

⁴Cuando estamos en el caso de dos o más asignaturas reprobadas los parámetros asociados a asignación e interacción se deben sumar para medir el efecto neto de A . En este caso el producto de esa suma es un número negativo, lo que corroboraría nuestra intuición de que a mayor número de asignaturas reprobadas menos grados adicionales alcanzará el individuo.

tros asociados a las nuevas variables, es que podemos concluir que efectivamente existe una relación cuadrática entre variable de asignación y variable de resultado.

Cuadro 5: En cada regresión se controla por colegio al cual asistió el alumno en 2007 y grado que cursó. Además, los errores estándar de todas las columnas son agrupados por la variable de asignación, el colegio y el grado al cual asistió el estudiante en el año 2007. La muestra total es de 611,383 observaciones.

Forma reducida (Variable dependiente: Grados Adicionales)				
	Lineal	Cuadrático	Cúbico	Cuártico
	(1)	(2)	(3)	(4)
I(Asig. Reprobadas \geq 2)	0,081** (0,023)	0,045 (0,044)	-0,034 (0,091)	-0,033 (0,218)
Asignación	-0,579** (0,006)	-0,214** (0,041)	-0,101 (0,124)	-0,104 (0,384)
Interacción	0,328** (0,012)			
Asignación ²		0,365** (0,042)	-0,05 (0,047)	-0,048 (0,214)
Interacción ²		-0,372** (0,049)		
Asignación ³			-0,527** (0,170)	0,004 (0,047)
Interacción ³			0,532** (0,175)	
Asignación ⁴				0,528 (0,643)
Interacción ⁴				-0,528 (0,646)
Constante	5,14** (0,0056)	5,14** (0,0056)	5,14** (0,0056)	5,14** (0,0056)
R^2 ajustado	0,3344	0,3344	0,3344	0,3344
AIC	1.339.174	1.339.171	1.339.168	1.339.170

** $p < 0,01$, * $p < 0,05$

Siguiendo con el análisis, para aquel modelo que ajusta la variable de asignación con un polinomio de tercer grado vemos en el Cuadro 5 que AIC aumenta levemente, pero no ocurre lo mismo con el R^2

ajustado, el cual se mantiene. Además, al incluir nuevas variables la significancia de las anteriores se pierde y la magnitud de los estimadores asociados a asignación e interacción en su grado más alto son tales que se compensan casi perfectamente, llevando a que el efecto del número de asignaturas reprobadas para personas con dos o más ramos deficientes sea cercano a cero. En el caso de un polinomio cuártico el escenario es peor, ya que sólo la constante es estadísticamente significativa, por lo que no se estaría identificando alguna relación entre el número de asignaturas reprobadas y la variable de resultado.

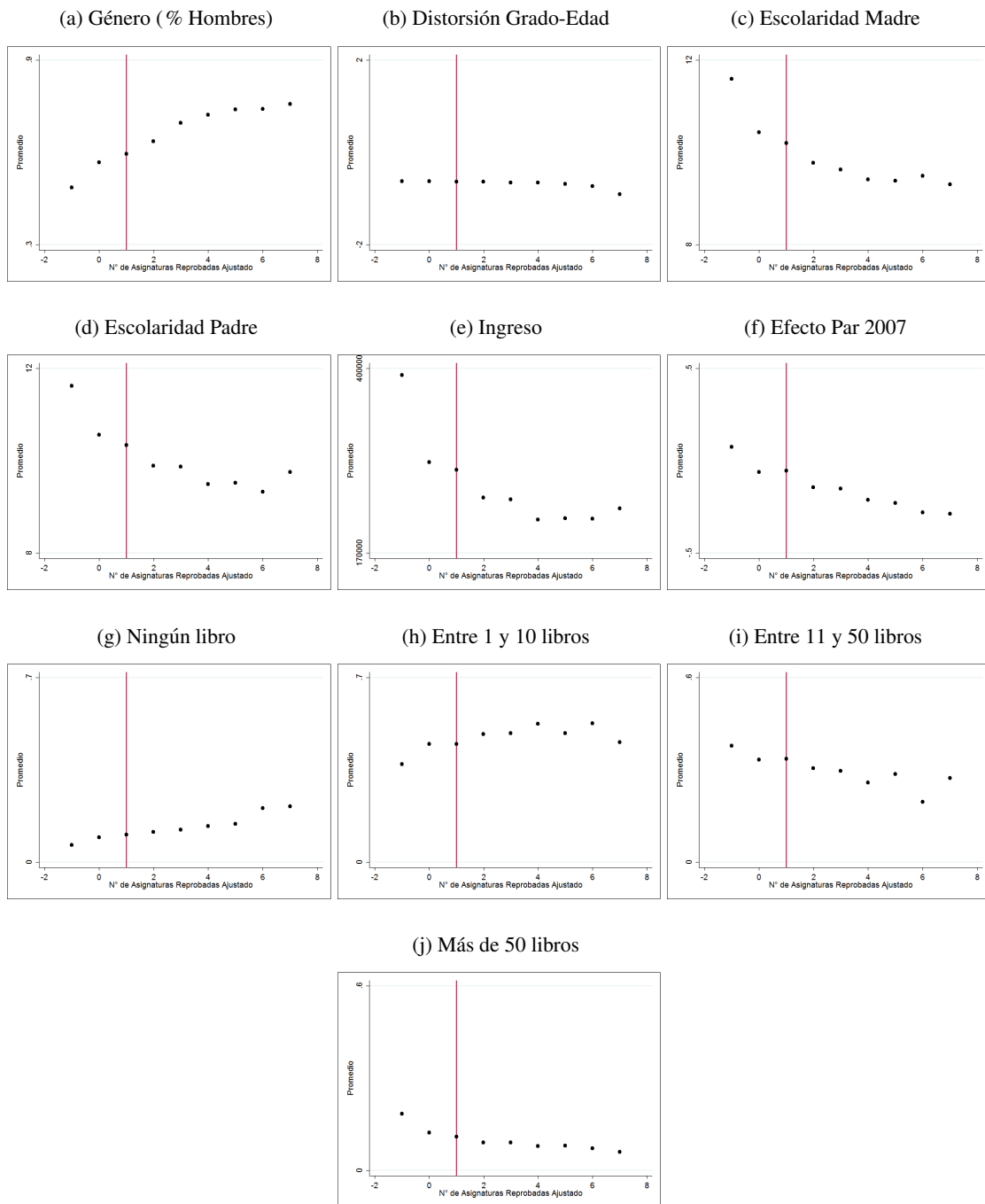
Considerando todo lo mencionado podemos concluir que el polinomio que mejor ajusta la relación entre “Grados Adicionales” y número de asignaturas reprobadas es el cuadrático. Por lo tanto, se testeará el cumplimiento de los principales supuestos sobre la muestra completa, pero teniendo en consideración esta relación cuadrática.

Continuidad alrededor del umbral

La importancia de la segunda condición radica en que al existir control impreciso por parte de los individuos, estos pueden tomar decisiones que aumenten la probabilidad de recibir o no el tratamiento, por lo tanto que las variables de control varíen continuamente alrededor del umbral implicaría que el control de los estudiantes es tal que sus decisiones o características no alcanzan a influir de manera importante en los resultados. La Figura 3 trata este tema graficando el valor promedio de las características observables de los estudiantes contra el número de asignaturas reprobadas normalizado (Jacob & Lefgren, 2009). A simple vista los resultados indican que no hay grandes discontinuidades en el umbral de la variable de asignación, lo que es necesario comprobar mediante regresiones de cada variable de control en función de $U_i = I(A_i \geq 2)$.

Siguiendo la lógica de la Ecuación (2), el Cuadro 6 resume los principales resultados. Como podemos apreciar, todas las variables con excepción de género cumplen el requisito de variar continuamente alrededor del umbral. Para género la variable explicativa $U = I(A \geq 2)$ es significativa al 1%, lo que nos permite corroborar que existen comportamientos distintos entre hombres y mujeres tal que hacen que la variable de asignación no varíe continuamente en el límite, por lo que la característica de control “impreciso” no se estaría cumpliendo, llevando a una invalidación de la metodología RD Fuzzy. Para lo anterior la solución será analizar el efecto de la repitencia de manera separada para la muestra de hombres y mujeres, tema que se desarrollará en la siguiente subsección.

Figura 3: Análisis de (Dis)Continuidad en variables de control. Observaciones a la izquierda de $A = 1$ pertenecen al grupo de control. Muestra total Obs=611,383.



Cuadro 6: Análisis de (Dis)Continuidad en variables de control considerando la muestra completa. Se asume una relación cuadrática entre la variable de asignación y la variable de control. Además en cada regresión se controla por efecto fijo de colegio y grado al que el alumno asistió en el año 2007.

	(1)	(2)	(3)	(4)	(5)
	Género	Dist. Grado-Edad	Ingreso	Esc. Madre	Esc. Padre
I(Asig. Reprobadas \geq 2)	-0,032**	-0,023	-5.079	-0,082	-0,107
S.E.	(0,012)	(0,015)	(6.2e+03)	(0,083)	(0,087)
	(6)	(7)	(8)	(9)	(10)
	Efecto Par	Ningún libro	[1,10] libros	[11, 50] libros	> 50 libros
I(Asig. Reprobadas \geq 2)	0,0008	0,011	-0,018	0,017	-0,009
S.E.	(0,002)	(0,009)	(0,014)	(0,013)	(0,008)

** $p < 0,01$, * $p < 0,05$

6.1. Identificación separando hombres y mujeres.

Tal como se señaló en la sección anterior, para poder aplicar la metodología RD Fuzzy se debe cumplir que las variables de control varíen continuamente alrededor del umbral. Como la variable dicotómica $U = I(A \geq 2)$ fue significativa a menos del 1 % en la regresión que se realizó a “Género”, en esta sección procederemos a testear las condiciones para llevar a cabo la metodología de estudio de manera separada tanto para una muestra de hombres como para una muestra de mujeres. El incumplimiento de los supuestos de identificación impedirá de manera automática la aplicación de RD Fuzzy, al menos en ese grupo en cuestión. En el Apéndice 3 se abordó el tema del mejor polinomio que relacione número de asignaturas reprobadas con “Grados Adicionales”, tanto en la muestra de hombres como en la de mujeres. En el siguiente desarrollo testaremos que, condicional a la forma funcional elegida, se cumpla el supuesto de continuidad alrededor del umbral.

Hombres

En este apartado sólo consideraremos aquella muestra de hombres que en el año 2007 estuviese cursando 4to, 5to o 6to básico. Lo primero que debemos corroborar es que la variable de asignación utilizada efectivamente presente alguna discontinuidad significativa en el umbral establecido ($A = 0$, con A normalizado)⁵. La Figura 4 muestra que la probabilidad de repetir un grado en función del número de asignaturas reprobadas normalizada exhibe una relación creciente y un salto significativo en el umbral,

⁵Como fue analizado en el Apéndice 3, la forma funcional lineal es la que relaciona correctamente variable de asignación y variable de resultado.

lo que valida la utilización de la variable instrumental. La probabilidad de repetir para aquellos alumnos con menos de 2 asignaturas deficientes es de 1,30 %, y luego salta a 84,14 % para aquellos que alcanzaron dos o más asignaturas reprobadas.

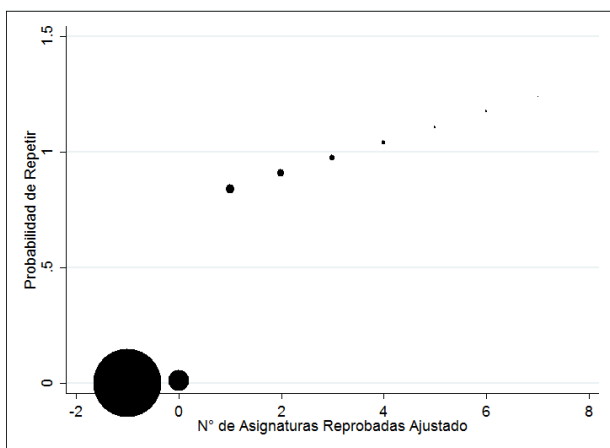
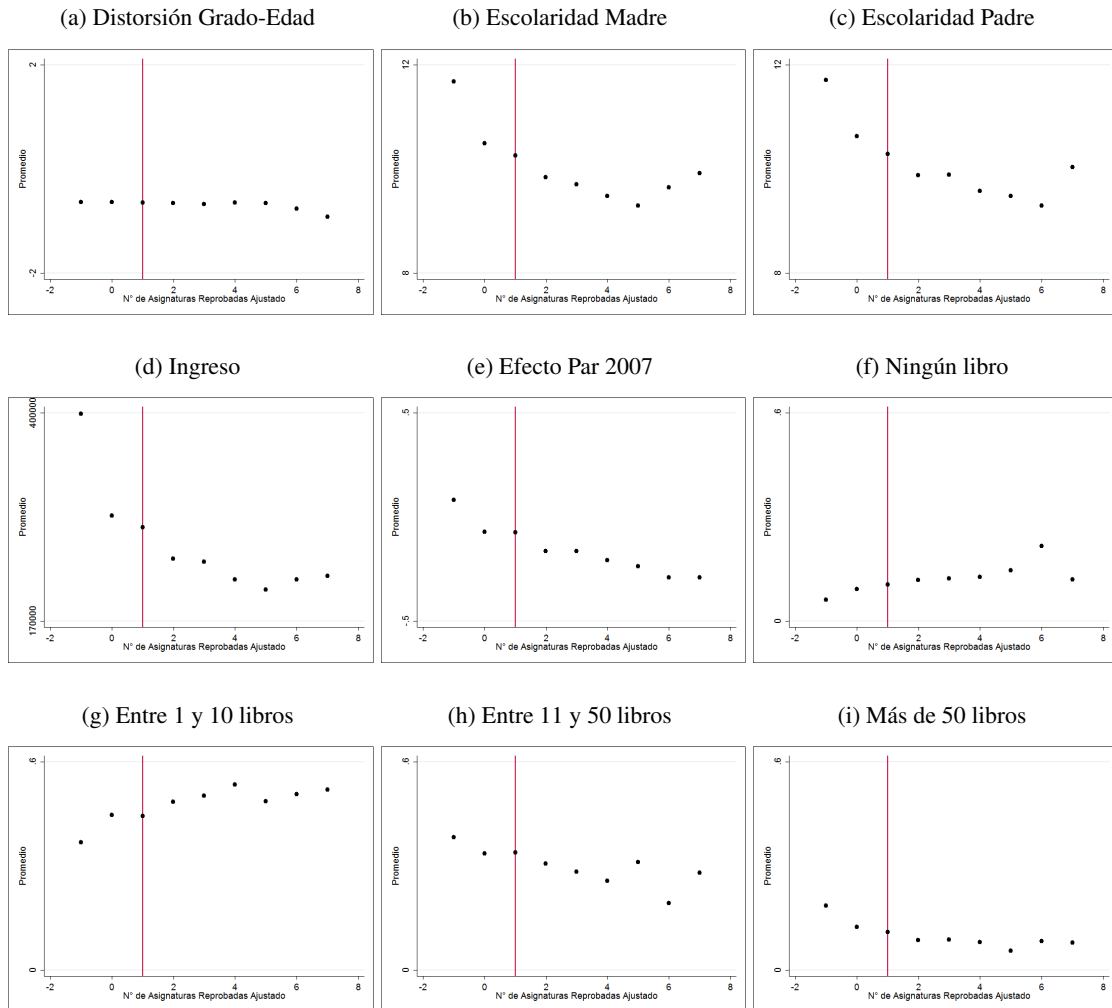


Figura 4: Probabilidad de repetir en función del Número de asignaturas reprobadas ajustado en Alumnos (Hombres) de 4to, 5to y 6to Básico, 2007. Se asume una relación lineal entre repitencia y variable de asignación. El tamaño de cada punto es proporcional a la cantidad de observaciones. La muestra solo considera hombres ($N = 303,324$)

Ahora veremos por inspección gráfica y estadística si para la muestra en cuestión el supuesto de continuidad alrededor del umbral de las características observables se cumple. Desde la Figura 5 podemos intuir que en promedio existen algunas diferencias cuando comparamos escolaridad de la madre, escolaridad del padre y nivel de ingresos del hogar del estudiante, entre un alumno que reprobó una asignatura y otro alumno que reprobó dos asignaturas.

Como sabemos, para realizar un análisis más acabado de la discontinuidad debemos regresionar las características observables en función de $U = I(A \geq 2)$, asumiendo una relación lineal. El Cuadro 7 resume los principales resultados, en donde podemos ver que tanto para el nivel de ingresos como la escolaridad de los padres $U = I(A \geq 2)$ es estadísticamente significativo a al menos el 5%. Lo anterior nos lleva a pensar que la situación de repitencia en la muestra de estudiantes hombres está mediada por atributos a nivel de hogar, los que inciden en el desempeño académico de estos y persisten en los pobres resultados de aprendizaje posteriores. La historia detectada aquí no es menor, estamos identificando que al grupo de hombres analizado le impacta en mayor medida las características de su entorno familiar, tanto que pueden llegar a definir el desempeño académico futuro del estudiante, influyendo hoy en la probabilidad de repitencia e impactando el día de mañana en la probabilidad de deserción (LLECE, 2015).

Figura 5: Análisis de (Dis)Continuidad en variables de control considerando sólo a una muestra de hombres. La línea vertical roja está situada en el valor $A = 1$ de la variable de asignación. Muestra total Obs=301,130.



Cuadro 7: Análisis de (Dis)Continuidad en variables de control considerando una muestra sólo de hombres. Se asume una relación lineal entre la variable de asignación y la variable de control. Además en cada regresión se controla por efecto fijo de colegio y grado al que el alumno asistió en el año 2007.

	(1)	(2)	(3)	(4)	(5)
	Efecto Par	Dist. Grado-Edad	Ingreso	Esc. Madre	Esc. Padre
I(Asig. Reprobadas \geq 2)	-0,0004	-0,006	-11.366*	-0,178**	-0,308**
S.E.	(0,001)	(0,012)	(4.4e+03)	(0,06)	(0,063)
	(6)	(7)	(8)	(9)	
	Ningún libro	[1,10] libros	[11, 50] libros	> 50 libros	
I(Asig. Reprobadas \geq 2)	0,008	-0,002	0,009	-0,016**	
S.E.	(0,006)	(0,01)	(0,009)	(0,006)	

** $p < 0,01$, * $p < 0,05$

Debido a los resultados hallados podemos concluir que el supuesto de identificación de RD Fuzzy queda inhabilitado en la muestra de hombres que se encontraba cursando 4to, 5to o 6to básico en el año 2007. El no cumplimiento de una etapa clave de la metodología impide que llevemos a cabo un análisis que arroje conclusiones robustas y no cuestionables, por ende la acción más lógica y correcta sería cambiar la estrategia de investigación. La estrategia escogida será Regresión Discontinua mezclada con Matching, de esta forma las diferencias en las características socioeconómicas para el grupo de control y de tratamiento no influirán en los resultados, ya que se controlará por ellas.

Mujeres

Al igual que en la muestra general y en la muestra considerando sólo hombres, para la muestra de mujeres que se encontraba en 4to, 5to o 6to básico en el año 2007 debemos corroborar que la probabilidad de repetir efectivamente presente alguna discontinuidad notoria en el umbral de la variable elegida para instrumentalizar. La Figura 6 grafica la probabilidad de repetir en función del número de asignaturas reprobadas, asumiendo una relación cuadrática⁶. En este caso la probabilidad de repetir para alumnas que presentaron sólo una asignatura reprobada es igual a 0,93 %, cambiando a 71,49 % si es que la estudiante tuviese dos asignaturas deficientes, en vez de una. Lo anterior es una clara muestra de discontinuidad en el umbral establecido ($A = 1$), por lo tanto podemos seguir analizando si se cumplen las condiciones necesarias para la aplicación de RD Fuzzy.

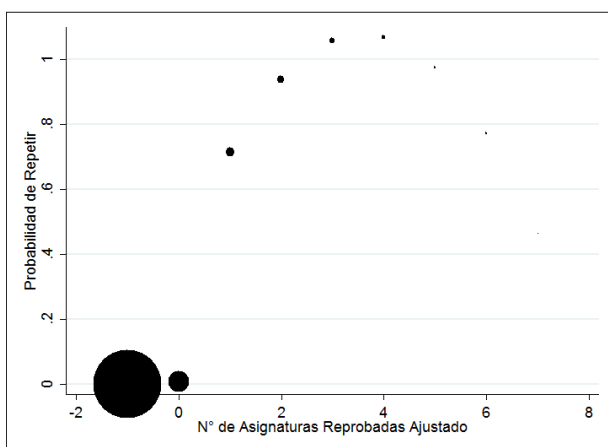


Figura 6: Probabilidad de repetir en función del Número de asignaturas reprobadas ajustado en Alumnas de 4to, 5to y 6to Básico, 2007. Asumiendo una relación cuadrática entre repitencia y variable de asignación. El tamaño de cada punto es proporcional a la cantidad de observaciones. La muestra solo considera mujeres ($N = 306,030$)

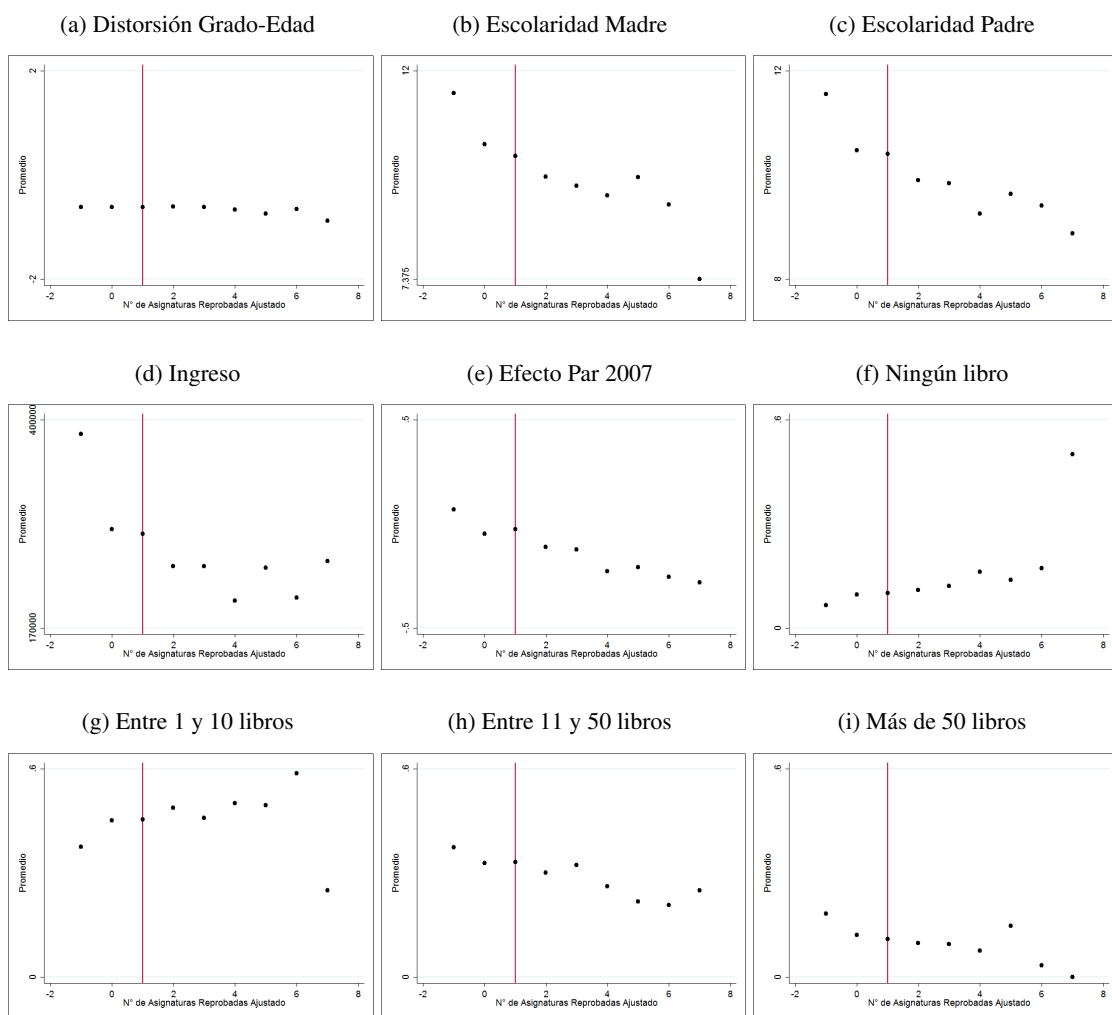
El siguiente paso es examinar si las características observables varían continuamente alrededor del umbral de la variable de asignación, número de asignaturas reprobadas, ya que como se mencionó en el capítulo anterior, las alumnas de determinadas características no debieran tener la capacidad de ordenarse estratégicamente alrededor del punto de discontinuidad, o al menos no con control “preciso”, esto con el fin de asegurarnos que no existan características no observables que estén influyendo en los resultados.

Dado lo anterior, tenemos que para la muestra de mujeres que se encontraba cursando 4to, 5to o 6to básico en el año 2007, y que al igual que los hombres impusimos que no hubiesen repetido ningún curso con anterioridad, la Figura 7 grafica el valor promedio de las características observables en función del

⁶Como fue analizado en el Apéndice 3, la forma funcional cuadrática es la que relaciona correctamente variable de asignación y variable de resultado para la muestra de mujeres.

número de asignaturas reprobadas. En general, si comparamos a las estudiantes que están justo sobre el umbral con las que están justo bajo el umbral podríamos pensar que en las variables escolaridad de la madre, efecto par e ingreso existiría alguna manipulación, por lo que el paso siguiente sería testear si existe alguna diferencia estadísticamente significativa.

Figura 7: Análisis de (Dis)Continuidad en variables de control considerando sólo a una muestra de mujeres. La línea vertical roja está situada en el valor $A = 1$ de la variable de asignación. Muestra total Obs=308,059.



El Cuadro 8 realiza un análisis de discontinuidad de las características observables, asumiendo una relación cuadrática. Aquí podemos concluir que para la muestra de mujeres la condición de elegibilidad no es estadísticamente significativa a niveles convencionales para ninguna característica observable, inclu-

yendo las variables de las que se tenían sospechas después del análisis gráfico, éstas eran escolaridad de la madre, efecto par e ingreso. Lo anterior nos permitiría descartar que existen atributos sociodemográficos y socioeconómicos que están influyendo significativamente en los resultados.

Cuadro 8: Análisis de (Dis)Continuidad en variables de control considerando una muestra sólo de mujeres. Se asume una relación cuadrática entre la variable de asignación y la variable de control. Además en cada regresión se controla por efecto fijo de colegio y grado al que el alumno asistió en el año 2007.

	(1)	(2)	(3)	(4)	(5)
	Efecto Par	Dist. Grado-Edad	Ingreso	Esc. Madre	Esc. Padre
I(Asig. Reprobadas \geq 2)	-0,00426	-0,0349	3.377	-0,163	0,021
S.E.	(0,003)	(0,0273)	(1.0e+04)	(0,1433)	(0,1501)
	(6)	(7)	(8)	(9)	
	Ningún libro	[1,10] libros	[11, 50] libros	> 50 libros	
I(Asig. Reprobadas \geq 2)	0,0168	0,00304	-0,00942	-0,0119	
S.E.	(0,0155)	(0,0234)	(0,0216)	(0,0135)	

** $p < 0,01$, * $p < 0,05$

Después de estudiar si los supuestos de identificación de RD Fuzzy se cumplen en la muestra de mujeres, podemos concluir que, a diferencia de lo hallado en la muestra de hombres, aquí sí se puede aplicar la metodología de RD Fuzzy. Lo anterior nos permite obtener un efecto local y específico de la política de repitencia, que es la motivación de la presente investigación. Lo que sigue a continuación es cuantificar el efecto que tiene la primera repitencia sólo para la muestra de mujeres, local al umbral de la variable de asignación.

Capítulo 7

Efecto de la repitencia sobre los años de escolaridad alcanzados

7.1. Mujeres

En esta sección presentaremos los estimadores tanto de la Ecuación (1), instrumentalizada y no instrumentalizada, como de la Ecuación (2), sólo para la muestra de mujeres, ya que es la muestra que cumple las condiciones necesarias para implementar la metodología RD Fuzzy.

Como hemos mencionado en el Capítulo “Identificación”, Regresión Discontinua Fuzzy descansa en una correcta especificación de la forma funcional a cada lado del punto de discontinuidad. La idea subyacente es que modelemos paramétricamente y en función de la variable de asignación normalizada y discreta, tanto la condición de repitencia como la variable de resultado. Producto de controlar por la relación entre variable de asignación y regla de promoción, el estimador θ_1 de la Ecuación (3) es interpretado como un aumento discreto en el umbral del número de asignaturas reprobadas, que es atribuido a la condición de repitencia¹. Recordar que el valor umbral de la variable de asignación está definido en *una* asignatura reprobada.

A continuación procederemos a implementar RD Fuzzy para cuantificar el efecto que tiene repetir sobre la cantidad de grados adicionales alcanzados luego de cumplir 6 años de escolaridad adicional,

¹Cabe recordar que nuestro estimador será un LATE en torno al punto de discontinuidad (una asignatura reprobada), y por lo tanto si tenemos en consideración metodología, supuestos y condiciones, debemos ser cuidadosos a la hora de interpretar los resultados y esbozar conclusiones.

considerando a una estudiante que en el año 2007 se encontraba en 4to, 5to o 6to básico.

Grados Adicionales

En el Cuadro 9 se encuentran los estimadores de la primera etapa, forma reducida y segunda etapa de RD Fuzzy, considerando “Grados Adicionales” como variable de resultado. Aquí analizamos el impacto que tiene para una estudiante repetir por primera vez, pero asumiendo una relación lineal y cuadrática entre el número de asignaturas reprobadas y la cantidad máxima de grados alcanzados al cumplir 6 años de escolaridad adicional. Recordemos que en el Apéndice 3 presentamos los estimadores de regresionar la variable de resultado en función de la condición de promoción, el número de asignaturas reprobadas y las interacciones, a partir de estas regresiones concluimos que los dos polinomios de menor grado son los correctos a la hora de realizar el análisis para la muestra que considera sólo a mujeres. En la primera y segunda columna del Cuadro 9 incluimos un polinomio de primer y segundo grado, respectivamente, sin efectos fijos ni controles. En las columnas (3) y (4) ajustamos una relación lineal en el número de asignaturas reprobadas, pero esta vez controlando por colegio y grado asistido en el año 2007, además de variables demográficas, socioeconómicas y culturales². Finalmente, en las columnas (5) y (6) incorporamos los mismos controles que los agregados en las columnas (3) y (4), pero en esta oportunidad asumiendo una relación cuadrática.

En el panel “Primera etapa” presentamos los estimadores de la primera etapa de RD Fuzzy, pero siempre teniendo en consideración cómo se diferencian los resultados de las diversas columnas. Como podemos notar, la variable instrumental es significativa a menos del 1% en todas las especificaciones. Cuando se asume una relación lineal, las estudiantes que tienen más de una asignatura reprobada presentan, en promedio, 63 puntos porcentuales más en la tasa de repitencia que las estudiantes que tienen una o ninguna asignatura deficiente, esta cifra cae 25,6 puntos porcentuales al asumir una relación cuadrática en la variable de asignación. La inclusión de efectos fijos y controles no provoca diferencias significativas en los estimadores.

En el panel “Forma reducida” del Cuadro 9 presentamos los estimadores de la ecuación en forma reducida del modelo, el método de estimación es similar al detallado en el apartado Forma Funcional, de la Sección “Identificación”. La significancia de la regla de promoción dependerá de la forma funcional

²Los controles que fueron agregados a la regresión se encuentran definidos en el Apéndice 1.

Cuadro 9: La primera etapa se estima utilizando la metodología MCO. Los errores estándar de todas las columnas son agrupados por la variable de asignación, el colegio y el grado al cual asistió el estudiante en el año 2007. La muestra solo considera mujeres ($N = 308,059$).

	(1)	(2)	(3)	(4)	(5)	(6)
Grado del Polinomio	1	2	1	1	2	2
Efecto Fijo Colegio	No	No	Sí	Sí	Sí	Sí
Efecto Fijo Grado	No	No	Sí	Sí	Sí	Sí
Controles Adicionales	No	No	No	Sí	No	Sí
Primera Etapa (Variable dependiente: Repitencia en 2007)						
I(Asignaturas Reprobadas ≥ 2)	0,633** (0,0104)	0,376** (0,0182)	0,632** (0,0114)	0,632** (0,0114)	0,376** (0,0198)	0,376** (0,0198)
Forma reducida (Variable dependiente: Grados Adicionales)						
I(Asignaturas Reprobadas ≥ 2)	0,196** (0,0406)	0,038 (0,0767)	0,073 (0,3757)	0,074* (0,0373)	-0,021 (0,0715)	-0,013 (0,0706)
Asig. Reprobadas	-0,656** (0,011)	-0,111 (0,0752)	-0,538** (0,0077)	-0,522** (0,0077)	-0,139* (0,0696)	-0,143* (0,0687)
Asig. Reprobadas ²		0,545** (0,076)			0,4** (0,07)	0,379** (0,0691)
Interacción	0,379** (0,0239)		0,301** (0,0212)	0,287** (0,021)		
Interacción ²		-0,578** (0,0904)			-0,419** (0,0833)	-0,397** (0,0821)
Segunda Etapa (Variable dependiente: Grados Adicionales)						
Repitencia en 2007	0,310** (0,0646)	0,102 (0,2043)	0,115 (0,0596)	0,117* (0,0592)	-0,057 (0,19)	-0,035 (0,1879)
Asig. Reprobadas	-0,659** (0,0112)	-0,15 (0,1513)	-0,54** (0,0079)	-0,523** (0,0078)	-0,117 (0,1402)	-0,13 (0,1385)
Asig. Reprobadas ²		0,507** (0,1492)			0,421** (0,1383)	0,392** (0,1365)
Interacción	0,347** (0,0289)		0,289** (0,0261)	0,275** (0,0259)		
Interacción ²		-0,534** (0,1737)			-0,444 (0,1609)	-0,412** (0,1589)

** $p < 0,01$, * $p < 0,05$

utilizada, ya que cuando enfrentamos un polinomio de primer grado, sin controles ni efectos fijos, el efecto es significativo y positivo, pero si asumimos un polinomio de segundo grado no ocurre lo mismo, independiente de la agregación de controles y/o efectos fijos, obteniendo efectos no significativos a lo largo de toda la muestra en cuestión. Por otra parte, los estimadores asociados a la variable de asignación en nivel y al cuadrado son significativos a al menos el 5 %, y esto se mantiene incluso cuando cambiamos la especificación.

Finalmente, en el panel “Segunda etapa” del Cuadro 9 se encuentra el estimador que responde a la pregunta que motiva esta investigación, ¿cuál es el efecto de repetir por primera vez en el año 2007, sobre la cantidad de grados alcanzados al cumplir 6 años de escolaridad adicional? Podemos ver que en la columna (1) el efecto de la repitencia es positivo y estadísticamente significativo: las estudiantes que repiten en el año 2007 alcanzan, en promedio, 0,31 grados adicionales más que el grupo de control, local al umbral. La cantidad de asignaturas reprobadas en el año 2007 impacta negativamente en la variable de resultado, disminuyendo en promedio 0,659 grados por cada asignatura extra reprobada. El efecto neto de la variable de asignación es negativo y significativo en torno al umbral, lo que quiere decir que no solamente importa repetir, sino que también importa el desempeño que la alumna presente en cada asignatura rendida. En la columna (2) se asume una relación cuadrática entre la variable de resultado y la variable de asignación, en este caso el efecto de repetir en el año 2007 también aumenta la cantidad de grados adicionales, pero no de manera estadísticamente significativa. La hipótesis del efecto de la cantidad de asignaturas reprobadas sobre la variable de resultado se mantiene, pero ahora sólo es significativo su impacto cuadrático. Vemos que los coeficientes asociados a “Asig. Reprobadas²” e “Interacción²” son significativos a menos del 1 %, lo que indicaría que en estudiantes que reprobaban dos asignaturas o más el efecto neto de la variable de asignación (la suma de estos dos coeficientes) es, en promedio, -0,027 grados adicionales.

Ahora bien, considerar correctos y adecuados los resultados arrojados por los modelos de las columnas (1) y (2) sería incurrir en un error, ya que en estos se estarían omitiendo variables explicativas importantes, lo que sesgaría el coeficiente de interés, $\hat{\theta}_1$ de la Ecuación (3). La solución sería controlar por variables que creemos explicarían la cantidad de grados adicionales, de manera tal que $\hat{\theta}_1$ no esté capturando un efecto indirecto no atribuible a la condición de repitencia, en esta línea los modelos más completos y correctos serían aquellos estimados en las columnas (4) y (6). Como podemos apreciar en

la columna (4), repetir por primera vez en el año 2007 tiene un efecto que, en promedio, es positivo, pequeño en magnitud y estadísticamente significativo a menos del 5 %, el cual asciende a 0,117 grados adicionales. Por otra parte, el efecto neto que tiene el número de asignaturas reprobadas sobre la variable de resultado es negativo y significativo a menos del 1 %, esto se puede ver en los coeficientes asociados a “Asig. Reprobadas” e “Interacción”, cuya diferencia promedio es igual a -0,248. Lo anterior quiere decir que mientras más asignaturas reprueba una estudiante, menores logros académicos consigue al finalizar 6 años de escolaridad adicional, estos medidos en grados adicionales. La historia es un poco distinta si consideramos los estimadores de la columna (6), ya que como podemos observar en el Cuadro 9 repetir por primera vez en el año 2007 tiene un efecto que es negativo, pequeño en magnitud y no significativo a niveles convencionales. En promedio, una estudiante que en el año 2007 repite logra -0,035 grados adicionales, este efecto negativo es potenciado por un efecto neto negativo del número de asignaturas reprobadas: por una parte, si consideramos el efecto en nivel del número de asignaturas reprobadas, variable “Asig. Reprobadas” en el Cuadro 9, éste es negativo pero no significativo a menos del 5 % (-0,13); por otra parte, si consideramos el efecto neto cuadrático que tiene el número de asignaturas reprobadas, variables “Asig. Reprobadas²” e “Interacción²” del Cuadro 9, encontramos que para las personas que cumplen con tener dos o más asignaturas reprobadas ($U_i = I(A_i \geq 2)$), el efecto promedio de reprobar una asignatura adicional es negativo e igual a -0,2, el cual es significativo a al menos el 1 %.

Además de identificar los efectos de la primera repetencia en la muestra de mujeres, también es importante señalar cómo está distribuida la cantidad de alumnas en función del número de asignaturas reprobadas, esto nos ayudará a dimensionar para cuántas estudiantes es válido el resultado encontrado. En el Cuadro 10 se aprecia el porcentaje de alumnas según el número de asignaturas reprobadas, aproximadamente el 7,36 % de la muestra se encuentra alrededor del umbral; específicamente, el 6,27 % de las alumnas reprobó una asignatura en el año 2007, proporción que cae a 1,09 % cuando consideramos dos asignaturas reprobadas.

La mayor parte de las estudiantes no reprueba ninguna asignatura. Para las estudiantes que reprueban una asignatura, el 0,03 % de ellas repite de curso, mientras que para las estudiantes que reprueban dos asignaturas, la tendencia cambia y el 70,03 % de ellas reprueba el respectivo curso. En general, los efectos encontrados afectan a una porción muy pequeña de la población estudiada, solo 22.685 alumnas se encuentran alrededor del punto de discontinuidad. Por lo tanto, los resultados encontrados debiesen

ser interpretados con cautela y siempre teniendo en mente la magnitud de la población abarcada por la metodología de estudio.

Cuadro 10: Número y porcentaje de alumnas en función del número de asignaturas reprobadas en el año 2007.

No de Asig. Reprobadas	No de Alumnas	% de Alumnas
0	282.210	91,61 %
1	19.322	6,27 %
2	3.363	1,09 %
3	1.963	0,64 %
4	761	0,25 %
5	302	0,10 %
>6	101	0,04 %
Total	308.059	100,0 %

7.1.1. Efectos dinámicos

Como pudimos ver en el apartado anterior, los efectos que tiene repetir en el año 2007 son débiles y siempre pequeños en magnitud, independiente de la forma funcional utilizada. Lo anterior deja entrever que si comparamos al grupo de tratamiento y al grupo de control en torno al umbral de la variable de asignación, encontraremos que estos son muy similares en la cantidad de grados alcanzados al cumplir 6 años de escolaridad adicional. Una posible explicación sería que el grupo de control eventualmente repetirá en años posteriores, igualando el valor de la variable de resultado obtenido por el grupo de tratamiento. Una manera simple de avalar nuestra hipótesis es comparando al grupo que justo no repitió en 2007, es decir, aquellas estudiantes que no repitieron y presentaron una asignatura reprobada (grupo de control), con aquellas que justo repitieron en 2007, es decir, aquel grupo que repitió y presentó dos asignaturas reprobadas (grupo de tratamiento). Los Cuadros 11 y 12 nos ayudan con este análisis descriptivo, aquí debemos tener presente que el valor máximo que puede tomar la variable de resultado es igual a 6, por lo tanto, en el caso de las alumnas repitentes, cualquier valor menor a 6 años implicaría una reincidencia en la repitencia, y en el caso de las alumnas promovidas cualquier valor menor a 6 años implicaría haber repetido en posteriores grados.

Cuadro 11: Estadística descriptiva de estudiantes en 4to, 5to y 6to básico, considerando sólo una muestra de mujeres no repitentes en 2007.

	Ninguna asignatura reprobada	Una asignatura reprobada	Dos asignaturas reprobadas	Tres o más asignaturas reprobadas
Edad	10,62 (0,965)	10,84 (0,950)	10,77 (0,953)	10,95 (1,09)
Dist. Grado-Edad	-0,619 (0,512)	-0,611 (0,548)	-0,604 (0,549)	-0,591 (0,59)
Escolaridad Madre	11,50 (3,61)	10,38 (3,63)	10,83 (3,78)	9,86 (4,1)
Escolaridad Padre	11,55 (3,76)	10,48 (3,72)	11,7 (3,75)	11,32 (3,73)
No de libros:				
-Ningún libro	0,067 (0,249)	0,097 (0,296)	0,074 (0,263)	0,091 (0,294)
-Menos de 10	0,375 (0,484)	0,5 (0,497)	0,434 (0,496)	0,591 (0,503)
-Entre 11 y 50	0,374 (0,484)	0,33 (0,47)	0,361 (0,481)	0,227 (0,429)
-Más de 50	0,183 (0,386)	0,122 (0,327)	0,128 (0,334)	0,091 (0,294)
Ingreso	384.479 (451.202)	279.862 (333.043)	342.316 (393.331)	434.404 (589.201)
NSE	0,093 (0,906)	-0,198 (0,845)	-0,05 (0,893)	-0,083 (1,026)
Efecto Par	0,069 (0,696)	-0,046 (0,581)	0,156 (0,649)	0,038 (0,668)
Grados Adicionales	5,39 (0,775)	4,74 (1,08)	4,75 (1,043)	4,55 (1,438)
Obs.	282.135	19.142	1.008	64

Cuadro 12: Estadística descriptiva de estudiantes en 4to, 5to y 6to básico, considerando sólo una muestra de mujeres repitentes en 2007.

	Ninguna asignatura reprobada	Una asignatura reprobada	Dos asignaturas reprobadas	Tres o más asignaturas reprobadas
Edad	10,47 (1,004)	10,72 (0,987)	10,72 (0,969)	10,75 (0,96)
Dist. Grado-Edad	-0,613 (0,543)	-0,694 (0,626)	-0,626 (0,568)	-0,62 (0,578)
Escolaridad Madre	11,11 (3,875)	9,12 (3,76)	9,803 (3,48)	9,55 (3,64)
Escolaridad Padre	11,59 (3,956)	9,38 (3,97)	10,12 (3,58)	9,8 (3,73)
No de libros:				
-Ningún libro	0,107 (0,311)	0,133 (0,341)	0,112 (0,316)	0,12 (0,325)
-Menos de 10	0,4 (0,493)	0,544 (0,499)	0,463 (0,499)	0,482 (0,5)
-Entre 11 y 50	0,307 (0,464)	0,228 (0,421)	0,32 (0,466)	0,299 (0,458)
-Más de 50	0,187 (0,392)	0,089 (0,285)	0,102 (0,303)	0,096 (0,294)
Ingreso	383.330 (402.947)	219.178 (278.161)	245.112 (277.169)	233.101 (282.719)
NSE	0,057 (0,912)	-0,496 (0,84)	-0,322 (0,793)	-0,389 (0,822)
Efecto Par	-0,006 (0,742)	-0,305 (0,481)	-0,103 (0,534)	-0,131 (0,541)
Grados Adicionales	4,91 (1,38)	4,27 (1,71)	4,57 (1,503)	4,26 (1,65)
Obs.	75	180	2.355	3.142

En el Cuadro 11 podemos ver que la cantidad de grados adicionales alcanzados por la muestra de mujeres no repitentes y que reprobaron sólo una asignatura en el año 2007 es igual a 4,74, lo que se puede interpretar como que este grupo eventualmente fue expuesto al tratamiento en grados posteriores, ya que en promedio lograron 1,26 grados menos que el límite superior de la cantidad de grados adicionales. Por otra parte, en el Cuadro 12 apreciamos que la cantidad de grados adicionales alcanzados por la muestra de mujeres repitentes y que reprobaron dos asignaturas asciende a 4,57, lo que nos quiere decir que en este grupo la repitencia volvió a ser experimentada. La diferencia simple entre estas dos cifras es igual a 0,17, en desmedro del grupo de tratamiento, y una vez que realizamos un test de diferencia de medias rechazamos la hipótesis nula de que ambos grupos presentan el mismo promedio de grados adicionales alcanzados en el año 2007. Nuestra creencia es que en el umbral, a pesar de existir diferencias significativas, la situación de riesgo de repetir y/o volver a repetir en años posteriores es común a ambos grupos, esto sustentado en el bajo valor de la diferencia de los promedios, lo que explicaría los resultados encontrados (efectos del tratamiento cercanos a cero).

En el siguiente análisis regresionaremos distintas variables de resultado en función de la condición de repitencia, esta última instrumentalizada por la condición de promoción y número de asignaturas reprobadas. La idea subyacente es indagar en la dinámica con la que la repitencia influye en el desempeño

académico de las alumnas, en el Cuadro 13 presentamos los resultados. Desde la columna (1) a la (6) la variable dependiente es la probabilidad de haber repetido uno, dos, tres, cuatro, cinco y hasta seis años, respectivamente, después de la aplicación del tratamiento³. En las columnas (7) a la (12) la variable de resultado es la cantidad de grados adicionales alcanzados al transcurrir uno, dos, tres, cuatro, cinco y hasta seis años, respectivamente, desde el momento en que recibió el tratamiento⁴. Finalmente, en la columna (13) la variable de resultado es una variable indicador que toma valor 1 si la estudiante presentó una asistencia intermitente, es decir, si la estudiante estuvo un determinado año y al buscarla en el año siguiente no la encontramos, y 0 en caso contrario, considerando un periodo total de 6 años después de ocurrido el tratamiento.

Una de las conclusiones que se pueden desprender del Cuadro 13 apuntan a que la probabilidad de repetir al año inmediatamente siguiente a la aplicación del tratamiento puede ser explicada por la condición de repitencia en el año 2007 de forma estadísticamente significativa a menos del 5%: las alumnas repitentes presentan una menor probabilidad de repetir al año siguiente, en promedio, 9,8 puntos porcentuales menos que el grupo de control en torno al valor de umbral de la variable de asignación. Para posteriores años identificamos que el grupo de repitentes en el año 2007 tiene asociados probabilidades menores de repetir, pero debemos tener en consideración que los estimadores no son significativos a niveles convencionales, por lo que no podríamos rechazar la hipótesis de que el riesgo de repitencia es común a ambos grupos. Ahora bien, si consideramos las variables grados adicionales alcanzados después de uno, dos y hasta seis años posteriores, vemos que en promedio las alumnas que repitieron el año 2007 persistentemente muestran un retraso en su progreso escolar comparado con las alumnas promovidas en 2007, esta brecha se va acortando a medida que pasan los años, evidenciando que al sexto año el grupo de tratamiento consigue, en promedio, 0,73 grados adicionales menos⁵. Finalmente, si analizamos la va-

³En una primera instancia impusimos que la muestra no debía incluir a estudiantes que ya hubiesen recibido el tratamiento, esto en el marco de querer capturar un efecto que se identificara como un hecho aislado y no como una situación de reincidencias. En esta ocasión no realizaremos este filtro precisamente porque nos interesa identificar la dinámica y definir cuáles son las causas del efecto que responde a la pregunta central de esta investigación.

⁴El efecto que tiene la repitencia en el año 2007 sobre la cantidad de grados adicionales al año inmediatamente después debe ser igual a uno o muy cercano a uno, ya que es un periodo lo suficientemente corto como para que opere alguna dinámica o “reordenamiento” entre las alumnas, por lo tanto el grupo de control se encontrará en una posición de desventaja frente al grupo de tratamiento. Estado del Arte.

⁵Durante el desarrollo de esta investigación el periodo final de análisis ha sido diferenciado tanto para la muestra de repitentes como para la muestra de promovidos, esto ya que nuestro objetivo es que ambos grupos se encuentren en igualdad de condiciones y puedan lograr la misma cantidad de grados adicionales, de modo tal que la condición de repitencia sea vista como un pequeño tropiezo en el progreso escolar. Para efectos del análisis de la dinámica quisimos dejar de lado lo anterior, y estando conscientes del grado menos con el que carga el grupo de repitentes, hacer un seguimiento de la cantidad de grados adicionales alcanzados año a año.

Cuadro 13: Se asume una relación cuadrática entre la variable de asignación y la variable de control. Además en cada regresión se controla por efecto fijo de colegio y grado al que el alumno asistió en el año 2007, más características observables. Se considera sólo la muestra de mujeres.

	(1)	(2)	(3)	(4)	(5)	(6)
	Probabilidad de repetir en t=					
	1	2	3	4	5	6
Repitencia en t=0	-0,098*	-0,034	-0,043	-0,027	-0,013	-0,012
S.E.	(0,042)	(0,04)	(0,0405)	(0,043)	(0,042)	(0,038)
	(7)	(8)	(9)	(10)	(11)	(12)
	Grados adicionales en t=					
	1	2	3	4	5	6
Repitencia en t=0	-0,987**	-0,9**	-0,860**	-0,871**	-0,845**	-0,730**
S.E.	(0,007)	(0,047)	(0,071)	(0,097)	(0,127)	(0,16)
	(13)					
	Asistencia intermitente					
Repitencia en t=0	-0,110					
S.E.	(0,071)					

** $p < 0,01$, * $p < 0,05$

riable “Asistencia Intermitente” podemos ver en el Cuadro 13 que haber repetido en el año 2007 no está asociado a un efecto estadísticamente significativo, lo que nos dice el valor del estimador es que las alumnas que experimentaron la repitencia en el año 2007 presentaron, en promedio, una menor probabilidad de asistir de manera intermitente al colegio versus el grupo de alumnas que no repitió durante el 2007, de todas maneras no podemos rechazar la hipótesis de que tanto las alumnas promovidas como las alumnas repitentes sean parecidas alrededor del umbral, al menos en lo que a esta variable respecta.

A continuación veremos la dinámica de distribución de las personas al interior del grupo de repitentes para uno, dos y hasta seis años después del momento en que se efectuó el tratamiento. Separaremos el análisis entre alumnas promovidas y repitentes en el año 2007, y que presentaron una, dos o tres asignaturas reprobadas, esto último se realizará con el fin de identificar alguna dinámica alrededor del umbral. El Cuadro 14 resume la principal información a analizar. Lo que nos dice la primera fila del apartado “No Repitentes en t=0” es cuál es la distribución de alumnas no repitentes en el año 2007 con uno, dos o tres asignaturas reprobadas y que repitieron al año siguiente, subsiguiente o 6 años después, es decir, del universo de alumnas que en 2007 no repitieron, qué porcentaje repitió al año siguiente, subsiguiente o hasta 6 años después del 2007? Por ejemplo, aquí vemos que la distribución más elevada se observa al cuarto año después del tratamiento, en donde el porcentaje ascendió al 16,92%, esto quiere decir que el 16,92% de las alumnas que no repitieron el año 2007 sí lo hizo cuatro años después. La segun-

da fila de “No Repitentes en t=0” apunta a la concentración del grupo de repitentes uno, dos y hasta seis años después de ocurrido el tratamiento, es decir, del total de alumnas que repitieron 1, 2 o hasta 6 años después del 2007, qué porcentaje corresponde a estudiantes que en el año 2007 no repitieron, pero presentaron una, dos o tres asignaturas reprobadas? Por ejemplo, apreciamos que la concentración más baja de repitentes en t=0 se concentra al sexto año después de efectuado el tratamiento, cuyo porcentaje asciende a 77,63 %, esto quiere decir que el 77,63 % de las alumnas que repitieron el año 2013 no habían repetido el año 2007. La misma lógica cuenta para el apartado “Repitentes en t=0”, pero esta vez considerando al grupo de alumnas que no pasaron de curso en el año 2007 como la principal unidad de análisis.

Cuadro 14: Porcentaje de repitentes en año 2008, 2009, 2010, 2011, 2012 y 2013, considerando sólo a una muestra de mujeres que el año 2007 presentaban 1, 2 o 3 asignaturas reprobadas.

	t=1	t=2	t=3	t=4	t=5	t=6
No Repitentes en t=0	14,69	9,74	13,47	16,92	13,77	8,43
	91	82,65	86,41	84,77	82,72	77,63
Obs.	2.961	1.963	2.715	3.411	2.776	1.700
Repitentes en t=0	6,53	9,18	9,52	13,66	12,93	10,92
	9	17,35	13,59	15,23	17,28	22,37
Obs.	293	412	427	613	580	490

En general, lo que desprendemos del Cuadro 14 es que consistentemente la concentración del grupo de repitentes en cada año se caracteriza por una mayor proporción de alumnas que en el año 2007 fueron promovidas, esto también puede obedecer a que es un grupo más grande, por ende, es más probable que el día de mañana alguien que pertenezca a éste puede experimentar la repitencia. Lo que sí es más interesante ver es cómo la proporción de alumnas no repitentes en el 2007 y que repite en años posteriores es mayor que la proporción de alumnas repitentes en el año 2007 y que vuelve a repetir en años siguientes, con excepción del último periodo. Lo anterior quiere decir que, considerando al grupo de estudiantes que en 2007 presentaron uno, dos o tres asignaturas reprobadas, es más probable que la que repita después de efectuado el tratamiento sea alguien que haya sido promovida en el año 2007, pero que dicho año estuvo muy cerca de ser parte del grupo de tratamiento.

Después del análisis expuesto, nuestra principal conclusión apunta a que alrededor del umbral, el

grupo de alumnas promovidas en el año 2007 tiende a converger a tasa lenta al grupo de tratamiento, retrasando su progreso escolar y obteniendo logros similares a los obtenidos por el grupo de alumnas repitentes en 2007⁶. Debido a esto es que los estimadores identificados a través de la metodología RD Fuzzy fueron pequeños, y en ocasiones no significativos, ya que estarían reflejando que en el límite las alumnas son muy parecidas, tanto que la situación de vulnerabilidad y de repitencia se replicaría tarde o temprano en ambos grupos. Es importante destacar que el 23,3% del total de alumnas analizadas repitió algún grado entre el periodo escolar 2007-2013, una cifra que refleja la alta recurrencia de esta práctica educacional en las escuelas chilenas.

Finalmente, otro tema interesante es identificar si la repitencia efectivamente tiene efectos heterogéneos sobre la población estudiada, ya que podría darse el caso en que para las estudiantes que están más alejadas del punto de discontinuidad los efectos adversos sean más nocivos, o al contrario, se benefician más de la política dado que los conocimientos que habían adquirido no alcanzaban a cumplir un estándar mínimo necesario para un determinado grado. Para abordar este tema utilizaremos la metodología de Regresión Discontinua “Donut-hole” (Díaz et al., 2016) explorando dos modelos: uno en donde se compara a estudiantes con una o menos asignaturas reprobadas versus alumnas con tres o más asignaturas reprobadas, y otro donde se compara a estudiantes con una o menos asignaturas reprobadas versus alumnas con cuatro o más asignaturas reprobadas. Si efectivamente existen efectos heterogéneos, entonces debiéramos encontrar distintos efectos de la primera repitencia a medida que nos alejamos más del umbral.

En el Cuadro 15 se encuentran los resultados del ejercicio en que se quita del análisis a las alumnas con 2 asignaturas reprobadas. Podemos ver que los efectos de la primera repitencia son negativos, de mayor magnitud que los encontrados en el Cuadro 9, pero estadísticamente no significativos, local al umbral. Además, el efecto del número de asignaturas reprobadas en nivel es más nocivo que en el análisis con la muestra completa, aunque en este caso no es estadísticamente significativo. Por otro lado, el efecto neto cuadrático sigue siendo negativo y significativo a niveles convencionales. Lo anterior nos permite concluir que los efectos de la repitencia son más nocivos bajo esta nuevo escenario de estudio, lo que indicaría alguna pista de efectos heterogéneos.

⁶El ritmo de convergencia es lento, ya que en él se encuentra un número significativamente mayor de personas, por lo que cualquier cambio sobresaliente debiese realizarse en un proporción mayor.

Cuadro 15: Se compara a estudiantes con una o menos asignaturas reprobadas versus alumnas con tres o más asignaturas reprobadas. Se asume una relación cuadrática entre repitencia y variable de asignación, controlando por grado y colegio asistido en 2007. Los errores estándar de todas las columnas son agrupados por la variable de asignación, el colegio y el grado al cual asistió el estudiante en el año 2007.

Controles Adicionales	No	Sí
Primera Etapa (Variable dependiente: Repitencia en 2007)		
I(Asignaturas Reprobadas \geq 2)	0,975** (0,0105)	0,975** (0,0105)
Segunda Etapa (Variable dependiente: Grados Adicionales)		
Repitencia en 2007	-0,172 (0,1233)	-0,167 (0,1216)
Asig. Reprobadas	-0,188 (0,1297)	-0,189 (0,1276)
Asig. Reprobadas ²	0,349** (0,1289)	0,331** (0,1267)
Interacción ²	-0,367* (0,1557)	-0,347* (0,153)

** $p < 0,01$, * $p < 0,05$

Para concluir el análisis centraremos nuestra atención en la muestra de alumnas que están un poco más alejadas del punto de discontinuidad, específicamente en aquellas con cuatro o más asignaturas reprobadas. En el Cuadro 16 se puede apreciar que tanto el efecto de la primera repitencia como de la cantidad de asignaturas reprobadas aumentan en magnitud y siguen siendo negativos, pero estadísticamente no significativos a niveles convencionales. Estos resultados implicarían efectos heterogéneos y más nocivos a medida que nos alejamos del umbral establecido (una asignatura reprobada). La no significancia estadística podría venir por el hecho de que mientras más a la derecha del punto de discontinuidad nos encontremos, existe una menor cantidad de alumnas, lo que hace que el grupo de tratamiento sea muy pequeño en comparación al grupo de control, llevando a una pérdida de significancia en los efectos identificados.

Cuadro 16: Se compara a estudiantes con una o menos asignaturas reprobadas versus alumnas con cuatro o más asignaturas reprobadas. Se asume una relación cuadrática entre repitencia y variable de asignación, controlando por grado y colegio asistido en 2007. Los errores estándar de todas las columnas son agrupados por la variable de asignación, el colegio y el grado al cual asistió el estudiante en el año 2007.

Controles Adicionales	No	Sí
Primera Etapa (Variable dependiente: Repitencia en 2007)		
I(Asignaturas Reprobadas \geq 2)	0,978** (0,0225)	0,978** (0,0225)
Segunda Etapa (Variable dependiente: Grados Adicionales)		
Repitencia en 2007	-0,339 (0,2353)	-0,330 (0,2318)
Asig. Reprobadas	-0,266 (0,2595)	-0,269 (0,2549)
Asig. Reprobadas ²	0,269 (0,2575)	0,249 (0,253)
Interacción ²	-0,278 (0,3156)	-0,256 (0,3098)

** $p < 0,01$, * $p < 0,05$

7.2. Hombres

Dado que en el capítulo “Identificación” concluimos que el supuesto de identificación de RD Fuzzy no se cumple en la muestra de hombres, no podemos aplicar dicha metodología a la hora de calcular el efecto de la primera repitencia sobre la cantidad de grados adicionales alcanzados luego de 6 años. Esto ocurre porque existen características sociodemográficas que diferencian a los alumnos repitentes de los no repitentes desde antes de que ocurra el tratamiento, en este contexto los hombres que casi repiten de los que casi no repiten no son parecidos en el límite. Por lo tanto, cuando existen indicios de una fuerte autoselección en los límites de interés, una alternativa es combinar diseños y asumir que, después de condicionar por covariables, la asignación al tratamiento logra ser lo suficientemente aleatoria para aquellos que se encuentran cerca del límite.

Basado en lo anterior, una solución es diseñar un grupo de control y uno de tratamiento lo más parecido posible utilizando la metodología de Propensity score Matching (PSM). PSM opera bajo el supuesto básico de que dentro de la subpoblación con igual valor de propensity score, las covariables serán independientes de la variable indicador del tratamiento por lo que no existiría sesgo (Imbens & Wooldridge, 2009). Los pasos de la metodología son:

1. En una primera instancia estimamos el propensity score de la condición de estar sobre o bajo el umbral de una asignatura reprobada, en función de la escolaridad de la madre, la escolaridad del padre, el ingreso del hogar, la edad en 2007 y el colegio y curso al que asistió en 2007. La idea es hacer que las distribuciones empíricas de las covariables observables de los dos grupos sean lo más similares posibles.
2. A continuación realizamos un match 1 a 1 de los individuos que repitieron en 2007 (grupo de tratamiento) con los que no repitieron (grupo de control), en función del propensity score calculado antes.
3. Para finalizar, aplicamos la metodología RD Fuzzy solo en el soporte común de las unidades que fueron matcheadas.

A continuación procederemos a implementar RD Fuzzy-Matching para cuantificar el efecto que tiene repetir sobre la cantidad de grados, considerando la muestra de estudiantes hombres que en el año 2007 se encontraban en 4to, 5to o 6to básico.

Grados Adicionales

En el Cuadro 17 se presentan los estimadores de la primera etapa, forma reducida y segunda etapa de RD Fuzzy, solo considerando a hombres que se encuentran en el soporte común. Aquí analizamos el impacto que tiene para un estudiante repetir por primera vez, pero asumiendo una relación lineal entre el número de asignaturas reprobadas y la cantidad máxima de grados alcanzados al cumplir 6 años de escolaridad adicional⁷.

En la primera y segunda columna del Cuadro 17 se asume un polinomio de primer grado, se incluyen efectos fijos por curso y grado y el soporte común es identificado a través de un Matching 1 a 1, las columnas se diferencian por la omisión o incorporación de variables de control⁸.

En el panel “Primera etapa” del Cuadro 17 presentamos los estimadores de la primera etapa de RD Fuzzy-Matching. Como podemos notar, la variable instrumental es significativa a menos del 1% en las dos especificaciones. Cuando no se incluyen variables de control los estudiantes que tienen más de una asignatura reprobada presentan, en promedio, 71 puntos porcentuales más en la tasa de repitencia que los estudiantes que tienen una o ninguna asignatura deficiente, esta cifra cae en 1,7 puntos porcentuales al incluir variables de control en la regresión.

En el panel “Forma reducida” del Cuadro 17 presentamos los estimadores de la ecuación en forma reducida del modelo, el método de estimación es similar al detallado en el apartado *Forma Funcional*, de la Sección “Identificación”. La regla de promoción es estadísticamente significativa a al menos un 1% en ambas regresiones, aumentando su impacto de 2 a 4,76 puntos porcentuales una vez que incorporamos variables de control. Por otra parte, los estimadores asociados a la variable de asignación en nivel son negativos y significativos a al menos el 5%, indicando que a mayor número de asignaturas reprobadas en el año 2007 menores grados adicionales se alcanzan al final del periodo.

Finalmente, en el panel “Segunda etapa” del Cuadro 17 se encuentra el estimador que responde a la pregunta central de esta investigación, ¿cuál es el efecto de repetir por primera vez en el año 2007, sobre

⁷Recordemos que en el Apéndice 3 presentamos los estimadores de regresionar la variable de resultado en función de la condición de promoción, el número de asignaturas reprobadas y las interacciones, a partir de estas regresiones concluimos que el polinomio de primer grado es el correcto a la hora de realizar el análisis para la muestra que considera sólo a hombres.

⁸Los controles que fueron agregados a la regresión se encuentran definidos en el Apéndice 1.

Cuadro 17: La primera etapa se estima utilizando la metodología MCO. Los errores estándar de todas las columnas son agrupados por la variable de asignación, el colegio y el grado al cual asistió el estudiante en el año 2007. La muestra solo considera hombres que se encuentran en el soporte común ($N = 21,507$).

	(1)	(2)
Grado del Polinomio	1	1
Efecto Fijo Colegio	Sí	Sí
Efecto Fijo Grado	Sí	Sí
Controles Adicionales	No	Sí
Primera Etapa (Variable dependiente: Repitencia en 2007)		
I(Asignaturas Reprobadas ≥ 2)	0,709** (0,013)	0,692** (0,0167)
Forma reducida (Variable dependiente: Grados Adicionales)		
I(Asignaturas Reprobadas ≥ 2)	0,2** (0,064)	0,476** (0,073)
Asig. Reprobadas	-0,419** (0,0499)	-0,433** (0,0503)
Interacción	0,1677** (0,052)	0,1871** (0,0524)
Segunda Etapa (Variable dependiente: Grados Adicionales)		
Repitencia en 2007	0,282** (0,0902)	0,688** (0,106)
Asig. Reprobadas	-0,418** (0,05)	-0,431** (0,05)
Interacción	0,149 (0,0485)	0,121** (0,0487)

** $p < 0,01$, * $p < 0,05$

la cantidad de grados alcanzados al cumplir 6 años de escolaridad adicional? Podemos ver que en la columna (1) el efecto de la repitencia es positivo y estadísticamente significativo: los estudiantes que repiten en el año 2007 alcanzan, en promedio, 0,282 grados adicionales más que el grupo de control, local al umbral. La cantidad de asignaturas reprobadas en el año 2007 impacta negativamente en la variable de resultado, disminuyendo en promedio 0,418 grados por cada asignatura extra reprobada en la especifica-

ción sin controles. El efecto neto de la variable de asignación es negativo en torno al umbral, lo que quiere decir que no solamente importa repetir, sino que también importa el desempeño que el alumno presente en cada asignatura rendida. La lógica de los resultados de la columna (2) no difieren en gran medida de los presentados en la columna (1), existen dos cambios importantes: el efecto de la primera repitencia aumenta en 0,406 grados adicionales y el efecto neto de la variable de asignación es estadísticamente significativo a al menor un 1 % (efecto neto=0,31).

Además de identificar los efectos de la primera repitencia en la muestra de hombres también es importante señalar cómo está distribuida la cantidad de alumnos en función del número de asignaturas reprobadas, esto nos ayudará a dimensionar para cuántos estudiantes es válido el resultado encontrado. En el Cuadro 18 se aprecia el porcentaje de alumnos según el número de asignaturas reprobadas. Aproximadamente el 23,00 % de la muestra se encuentra alrededor del umbral; específicamente, el 1,13 % de los alumnos reprobó una asignatura en el año 2007, proporción que aumenta a 21,87 % cuando consideramos dos asignaturas reprobadas⁹.

Cuadro 18: Número y porcentaje de alumnos en función del número de asignaturas reprobadas en el año 2007.

No de Asig. Reprobadas	No de Alumnos	% de Alumnos
0	11.035	48,87 %
1	256	1,13 %
2	4.938	21,87 %
3	3.438	15,22 %
4	1.739	7,70 %
5	1.739	3,46 %
>6	782	1,74 %
Total	22.582	100,0 %

La mayor parte de los estudiantes no reprueba ninguna asignatura (48,87 %). Para los estudiantes que reprueban una asignatura, ninguno de ellos repite de curso, mientras que para los alumnos que reprueban

⁹Las diferencias significativas respecto a este mismo análisis pero aplicado en la muestra de mujeres se producen principalmente porque la metodología de estudio es distinta, en la muestra de hombres aplicamos PSM en primera instancia para definir un soporte común.

dos asignaturas, la tendencia cambia y el 79,34% de ellos reprueba su respectivo curso. En general, los efectos encontrados afectan a casi un cuarto de la muestra final de estudio, y el grupo de control, es decir, aquellos que no repiten y presentan una asignatura reprobada, corresponde solo al 4,9% de la muestra en torno al umbral. Por lo tanto, los resultados deben ser interpretados con cautela y siempre teniendo en mente la magnitud de la población abarcada por la metodología de estudio.

Capítulo 8

Conclusiones

El principal objetivo de esta investigación fue contribuir con la literatura que cuantifica el(los) efecto(s) que tiene la primera repitencia sobre subsecuentes logros académicos, dejando abierta la definición de logro académico según las herramientas que cada investigador tuviese a mano. En nuestro caso la variable de resultado fue definida como la cantidad máxima de grados alcanzados al cumplir 6 años de escolaridad adicional.

Como vimos en la parte inicial de este estudio, la repitencia es una práctica que en Chile se ha hecho cada vez más común, y debido a esto se hace imprescindible aumentar los estudios con el fin de obtener mayor evidencia a la hora de criticarla o potenciarla como un mecanismo recurrente en el sistema educacional chileno.

El resultado en este trabajo indicaría que para la muestra de mujeres¹, la repitencia pudiese tener un efecto ambiguo, local al umbral y muy pequeño sobre la cantidad máxima de grados alcanzados al cumplir 6 años de escolaridad adicional, el estimador dependerá del grado del polinomio asumido para controlar por la relación entre variable de resultado y variable de asignación. Además, identificamos que este pequeño efecto responde a una dinámica en que al pasar los años las trayectorias del grupo de control y grupo de tratamiento comienzan a asemejarse, logrando obtener valores similares en cuanto a variable de resultado se hable, y llevando a que la primera repitencia alcance efectos pequeños en la muestra analizada. Por lo tanto, los resultados avalan la idea de que “si hoy no te caes, pero estuviste a

¹Recordar que debido a los análisis de robustez realizados para comprobar la aplicabilidad de la metodología RD Fuzzy, se concluyó que los resultados de este estudio sólo serían válidos bajo la muestra de mujeres que en el año 2007 se encontraba cursando 4to, 5to o 6to básico.

punto de hacerlo, lo más probable es que el día de mañana sí repitas, sí sufras un tropiezo”. Además, los resultados de análisis adicionales avalan la existencia de efectos heterogéneos en la muestra de mujeres, lo que observamos aquí es que existen efectos negativos y no significativos que aumentan en magnitud a medida que nos alejamos del umbral establecido (una asignatura reprobada). La no significancia estadística podría venir por el hecho de que la cantidad de alumnas disminuye mientras más a la derecha del punto de discontinuidad nos encontramos, lo que hace que el grupo de tratamiento sea muy pequeño en comparación al grupo de control.

Para la muestra de hombres el escenario fue incluso peor, concluimos que existen factores socio-económicos que diferencian desde antes a los alumnos repitentes y no repitentes, la historia aquí es un poco más compleja y junto con ella el análisis de una posible “recuperación”, existe un entorno familiar que los impacta en mayor medida, desencadenando en situaciones que aumentan su vulnerabilidad, tal como la repitencia, y éstas terminan incidiendo en sus logros educacionales. Una forma de sortear esta situación y continuar con la investigación fue aplicar la metodología PSM sobre la muestra de hombres, de esta forma condicionamos por covariables observadas y construimos una muestra donde se cumple el supuesto de aleatoriedad en torno al umbral. Los resultados encontrados apuntan a un efecto positivo y estadísticamente significativo de la primera repitencia, sin embargo, como hemos identificado a lo largo de esta investigación, el resultado más crítico y relevante es el efecto neto negativo de la cantidad de asignaturas reprobadas sobre los grados adicionales alcanzados, ya que este resultado avalaría la hipótesis de que si hoy no repetiste, pero estuviste muy cerca, es muy probable que el día de mañana sí repitas, esto dado que cargas con una mochila de características que el sistema educacional chileno no logra soslayar y mermar.

En general, se han identificado en otros países efectos negativos de la repitencia, especialmente asociados a una mayor tasa de abandono y menores desempeños en los aprendizajes. Para el caso de Chile hemos identificado evidencia robusta que apunta a que más que la repitencia de un grado, el indicador más relevante de política pública vinculado a este fenómeno académico debiese ser la reprobación de cualquier asignatura, puesto que de suceder, y aunque en ese momento no implicase la repitencia de grado, anticiparía un alto riesgo de repitencia futura, profundizando los potenciales efectos nocivos para las oportunidades educacionales y laborales del estudiante. En este sentido el trabajo de cada establecimiento se debiese enfocar en una política de apoyo escolar constante para que “ningún estudiante repruebe

ninguna asignatura”.

Más allá del enfoque objetivo y técnico que hemos querido otorgarle a esta investigación, para poder plantear temas de políticas públicas debemos complementar esta visión y reconocer realidades y concepciones que nos llevan hoy en día a estar bajo este escenario. En esta línea García-Huidobro (2000) identificó tres supuestos básicos errados en torno al tema de repitencia y que creemos debemos mencionar: primero, en la percepción colectiva se suele considerar que la culpa de la repitencia es del alumno que no aprende; segundo, las buenas intenciones de la creación de esta práctica hacen suponer que es una segunda oportunidad que se le está dando a los estudiantes repitentes; y finalmente, se cree que la repitencia ayuda a potenciar un sistema educacional de excelencia. Debemos tener cuidado a la hora de sustentar programas bajo premisas que aún no han sido comprobadas, lo que el famoso sociólogo James S. Coleman señala es que “Un programa deficiente afecta más adversamente a los alumnos provenientes de familias de bajos ingresos, que a los procedentes de hogares de medianos ingresos, ya que estos últimos dependen menos del colegio para adquirir conocimientos. Pero un buen programa escolar es inherentemente igualitario debido al efecto compensador que tiene, precisamente, en los alumnos de hogares de bajos ingresos” (Hirsch Jr, 2007). Por lo tanto, debemos dar a la repitencia la importancia y análisis que amerita: la repitencia es un fenómeno multidimensional, por lo que se deben considerar todos sus costos, tanto directos como indirectos.

Medidas de corto plazo apuntan a un sistema de promoción más inclusivo y de mayor calidad. Poner en práctica un sistema escolar que se lleve a cabo mediante ciclos, tal como se pensó en 1ro y 3ro básico, podría generar resultados positivos ya que los alumnos tendrían más tiempo para asimilar las materias aprendidas sin necesidad de recibir una penalización. En el trabajo llevado a cabo por UNESCO (2012) se destaca a Ndaruhutse (2008), la cual señala que cuando la implementación de metas por subciclos se aplica en un sistema educacional que identifica y brinda apoyo a los alumnos con menor desempeño, esta práctica logra mejorar la calidad de toda la educación.

Otro punto muy criticado es lo que García-Huidobro (2000) señala como “dar el remedio, pero pedirle al alumno que se lo tome en otra parte”, una práctica común es que las escuelas hagan un continuo uso de la repitencia escolar como “medida” sanadora, pero no permitan que el estudiante continúe su educación en el mismo establecimiento, llevando a que se rompan lazos afectivos y generando mayores

efectos adversos sobre los niños y jóvenes a nivel psicológico.

Debemos recordar que en este trabajo se analizó el efecto de repetir por primera vez durante el ciclo básico, los efectos negativos y/o positivos se podrían estar encontrando en otros grupos. Abolir la aceptación o rechazo de esta práctica en determinadas circunstancias permitiría cambiar la interpretación de lo que hoy en día acompaña a la repitencia, y avanzaría en el desafío de transformar el paradigma desde una escuela selectiva o expulsora, a una escuela inclusiva o atractiva. Una alternativa factible sería mantener una tasa de repitencia baja, algún nivel de estado estacionario, con el fin de que el sistema educacional logre asegurar un nivel mínimo de contenidos para todos los alumnos, incluso para aquellos con altos niveles de ausentismo o inmadurez. Como señala Vera (2013) “El desafío para la educación en Latinoamérica es el de proveer educación de calidad a todos, incluyendo a aquellos que se educan en contextos adversos sin generar más segregación de la que ya, de por sí, sus condiciones de vida implican”. Por lo tanto, el refinamiento de la repitencia apuntaría a que en aquellos grupos en que se hace necesaria su aplicación efectivamente repitan, y por el contrario, salvaría a aquellos estudiantes que se encuentran en esa situación solo porque el sistema educacional no supo cómo actuar frente a determinadas características.

Apéndices

Apéndice 1

Cuadro 19: Definición de variables utilizadas en la investigación. Las últimas dos variables de control definidas sólo responden a la estadística descriptiva básica y no a la regresión.

Variable	Definición
Variable de Resultado	
Grados adicionales	Cantidad de cursos adicionales alcanzados (no necesariamente aprobados) en el año 2014 para los repitentes, y 2013 para los alumnos promovidos. La repitencia que se analiza es la del año 2007.
Variable de Asignación	
Número de Asignaturas reprobadas	Cantidad de cursos que para el año 2007 su promedio estuvo por bajo la nota 4.
Variables de Control	
Género	Variable binaria igual a 1 en caso que el estudiante sea hombre y 0 en caso contrario.
Distorsión Grado-Edad	Variable continua que indica la diferencia entre la edad que debiese tener un alumno que asiste a un determinado grado y la edad real en año 2007. Distorsiones negativas dan cuenta de estudiantes mayores en un grado.
Escolaridad de la Madre	Años de educación de la madre del estudiante.
Escolaridad del Padre	Años de educación del padre del estudiante.
No de libros en el hogar	Existen 4 variables binarias que indican en qué rango de número de libros en el hogar se categoriza al estudiante. (1) NUM-LIBROS0 es igual a 1 si en el hogar del alumno no hay libros y 0 en otro caso, (2) NUM-LIBROS1 toma valor 1 si existen entre 1 a 10 libros y 0 en otro caso, (3) NUM-LIBROS2 es igual a 1 si en el hogar del estudiante hay entre 11 a 50 libros y 0 en otro caso, y finalmente (4) NUM-LIBROS3 toma valor 1 si hay más de 50 libros en el hogar del estudiante y 0 en otro caso.
Ingreso	Es una variable continua que indica el ingreso mensual percibido en el hogar del estudiante (considerando los ingresos de todos los miembros que aportan al hogar).
Efecto Par	NSE promedio de los compañeros de un estudiante, para el año 2007. Se consideran cursos con más de 15 alumnos.
Variables para Estadística Descriptiva	
Edad	Edad que tenía el estudiante en el año 2007.
NSE	Índice de Nivel Socioeconómico creado considerando la escolaridad de los padres y el ingreso familiar para el año 2007. Se basa en la metodología de Análisis de Factores.

Apéndice 2

Cuadro 20: Estadística descriptiva de estudiantes en 4to, 5to y 6to básico en 2007, considerando sólo una muestra de no repitentes.

	Ninguna asignatura reprobada	Una asignatura reprobada	Dos asignaturas reprobadas	Tres o más asignaturas reprobadas
Género (% Hombres)	0,486 (0,5)	0,567 (0,496)	0,503 (0,5)	0,656 (0,479)
Edad	10,63 (0,964)	10,87 (0,937)	10,82 (0,952)	10,875 (0,984)
Dist. Grado-Edad	-0,628 (0,513)	-0,624 (0,549)	-0,620 (0,555)	-0,719 (0,629)
Escolaridad Madre	11,59 (3,58)	10,45 (3,59)	10,83 (3,72)	10,05 (4,14)
Escolaridad Padre	11,62 (3,74)	10,57 (3,7)	11,0 (3,78)	10,83 (3,59)
No de libros:				
-Ningún libro	0,065 (0,246)	0,094 (0,292)	0,086 (0,28)	0,109 (0,315)
-Menos de 10	0,371 (0,483)	0,448 (0,497)	0,43 (0,495)	0,547 (0,502)
-Entre 11 y 50	0,379 (0,485)	0,333 (0,471)	0,35 (0,477)	0,266 (0,445)
-Más de 50	0,184 (0,387)	0,123 (0,329)	0,131 (0,338)	0,078 (0,270)
Ingreso	391.393 (454.358)	283.927 (334.668)	342.866 (407.746)	353.465 (446.554)
NSE	0,114 (0,902)	-0,177 (0,84)	-0,057 (0,893)	-0,158 (0,905)
Efecto Par	0,075 (0,7)	-0,059 (0,586)	0,122 (0,674)	-0,046 (0,630)
Grados Adicionales	5,37 (0,797)	4,67 (1,1)	4,66 (1,08)	4,61 (1,31)
Obs.	548.687	44.182	2.028	64

Cuadro 21: Estadística descriptiva de estudiantes en 4to, 5to y 6to básico en 2007, considerando sólo una muestra de repitentes.

	Ninguna asignatura reprobada	Una asignatura reprobada	Dos asignaturas reprobadas	Tres o más asignaturas reprobadas
Género (% Hombres)	0,599 (0,491)	0,646 (0,479)	0,625 (0,484)	0,668 (0,471)
Edad	10,56 (0,956)	10,79 (1,02)	10,76 (0,966)	10,79 (0,944)
Dist. Grado-Edad	-0,652 (0,588)	-0,715 (0,601)	-0,643 (0,575)	-0,647 (0,586)
Escolaridad Madre	11,19 (4,031)	9,39 (3,70)	10,0 (3,54)	9,67 (3,59)
Escolaridad Padre	11,57 (4,03)	9,65 (3,84)	10,12 (3,64)	9,81 (3,71)
No de libros:				
-Ningún libro	0,08 (0,272)	0,122 (0,327)	0,11 (0,313)	0,122 (0,328)
-Menos de 10	0,396 (0,490)	0,491 (0,5)	0,454 (0,498)	0,491 (0,5)
-Entre 11 y 50	0,31 (0,464)	0,297 (0,457)	0,331 (0,471)	0,295 (0,456)
-Más de 50	0,214 (0,411)	0,088 (0,284)	0,103 (0,304)	0,089 (0,284)
Ingreso	394.219 (464.461)	222.836 (268.759)	251.397 (282.997)	232.930 (277.752)
NSE	0,067 (0,995)	-0,43 (0,829)	-0,298 (0,809)	-0,376 (0,816)
Efecto Par	0,003 (0,743)	-0,266 (0,473)	-0,11 (0,547)	-0,161 (0,535)
Grados Adicionales	4,96 (1,38)	4,21 (1,71)	4,50 (1,53)	4,12 (1,66)
Obs.	187	509	6.273	9.453

Apéndice 3

Hombres

Para el caso de la muestra de hombres, la principal información a analizar se encuentra contenida en el Cuadro 22. En la primera columna del Cuadro 22 podemos ver que los coeficientes asociados a ajustar la relación por un polinomio lineal son significativos a menos del 1%. Además vemos que el R^2 ajustado de la regresión es igual a 32,06% y el valor AIC asciende a 694.574. Ahora bien, si comparamos los anteriores resultados con los arrojados al controlar por una relación cuadrática podemos apreciar que tanto las nuevas como las anteriores variables son significativas al 1%, con excepción de I(Asig. Reprobadas \geq 2). No obstante, el R^2 ajustado no aumenta, contrario a lo que sucede con AIC, el cual aumenta marginalmente. Por otra parte, los coeficientes asociados a la variable de asignación al cuadrado y su respectiva interacción con la condición de elegibilidad son tales que se compensan casi perfectamente, llevando a que para el grupo con dos o más asignaturas reprobadas el impacto de la variable de asignación sea simplemente en nivel. Dado los antecedentes encontrados, nuestra conclusión es diferente a la esbozada en el caso de la muestra completa, creemos que no existe una relación no lineal entre el número de asignaturas reprobadas y “Grados Adicionales”, lo cual se sustenta principalmente en la magnitud y significancia estadística de los parámetros asociados a una relación cuadrática.

En el caso de un polinomio de tercer grado ocurre el mismo fenómeno identificado al analizar la muestra completa, la significancia de los parámetros asociados a menores grados se pierde, y la magnitud de los estimadores asociados a asignación e interacción en su grado más alto es tal que se compensan casi perfectamente, llevando a que el efecto del número de asignaturas reprobadas para hombres con dos o más ramos deficientes sea cercano a cero. Finalmente, al controlar por un polinomio de cuarto grado podemos ver que ninguna variable es significativa, a excepción de la constante, por lo tanto descartamos completamente este modelo, ya que en él no se encuentra relación entre la variable de asignación y la variable de resultado.

En el caso de la muestra de hombres nuestra conclusión sería que el polinomio que mejor ajusta la relación entre “Grados Adicionales” y número de asignaturas reprobadas es el lineal, ya que captura la misma información que una relación cuadrática a un menor esfuerzo. Entonces, basados principalmente en la significancia estadística de los parámetros, y obviando el leve aumento en AIC, nos inclinaremos

por un polinomio de primer grado para la realización de posteriores análisis de consistencia.

Cuadro 22: En cada regresión se controla por colegio al cual asistió el alumno en 2007 y grado que cursó. Además, los errores estándar de todas las columnas son agrupados por la variable de asignación, el colegio y el grado al cual asistió el estudiante en el año 2007. La muestra total de hombres es de 303,324 observaciones.

Forma reducida (Variable dependiente: Grados Adicionales)				
	Lineal	Cuadrático	Cúbico	Cuártico
	(1)	(2)	(3)	(4)
I(Asig. Reprobadas \geq 2)	0,083** (0,03)	0,074 (0,056)	-0,035 (0,113)	-0,037 (0,266)
Asignación	-0,60** (0,007)	-0,24** (0,05)	-0,086 (0,152)	-0,081 (0,464)
Interacción	0,351** (0,015)			
Asignación ²		0,361** (0,051)	-0,059 (0,057)	-0,062 (0,257)
Interacción ²		-0,362** (0,06)		
Asignación ³			-0,573** (0,208)	0,007 (0,056)
Interacción ³			0,58** (0,214)	
Asignación ⁴				0,588 (0,774)
Interacción ⁴				-0,58 (0,778)
Constante	5,08** (0,007)	5,08** (0,007)	5,08** (0,007)	5,08** (0,007)
R^2 ajustado	0,3206	0,3206	0,3206	0,3206
AIC	694.574	694.576	694.572	694.574

** $p < 0,1$, * $p < 0,05$

Mujeres

Para el caso de la muestra de mujeres es el Cuadro 23 el que resume la principal información a analizar. En la columna (1) podemos ver que los coeficientes asociados a ajustar la relación por un polinomio lineal son significativos al 1%. Además vemos que el R^2 ajustado de la regresión es igual a 35,44% y el valor del AIC asciende a 631.182. Ahora bien, si comparamos los resultados de la columna (1) con los arrojados al controlar por una relación cuadrática, podemos apreciar que $I(\text{Asig. Reprobadas} \geq 2)$ ya no es significativa, sin embargo los coeficientes asociados a una relación lineal y cuadrática de la variable de asignación son significativos a al menos un 5%. Además AIC cae 10 puntos en nivel cuando se trata de una relación cuadrática.

Por otra parte, los escenarios en que ajustamos por un polinomio de tercer y cuarto grado llevan a conclusiones distintas, aquí podemos ver que ninguna variable es significativa, a excepción de la constante, lo que nos estaría diciendo que al agregar nuevas variables se pierde toda relación entre la variable de asignación y la variable de resultado, invalidando el modelo.

Por lo tanto, basados en la significancia estadística de los parámetros y el criterio de información de Akaike, concluimos que el polinomio que mejor ajusta la relación entre “Grados Adicionales” y número de asignaturas reprobadas es el lineal o cuadrático. En esta ocasión nos inclinaremos por un polinomio de segundo grado para la realización de posteriores análisis de consistencia, ya que es el que presenta menor criterio de información y si analizamos el efecto neto de la variable de asignación para el caso de una persona con dos o más asignaturas reprobadas, vemos que éste es distinto comparado con la relación lineal (se está capturando otra relación) y es estadísticamente significativo.

Cuadro 23: En cada regresión se controla por colegio al cual asistió el alumno en 2007 y grado que cursó. Además, los errores estándar de todas las columnas son agrupados por la variable de asignación, el colegio y el grado al cual asistió el estudiante en el año 2007. La muestra total de mujeres es de 308,059 observaciones.

Forma reducida (Variable dependiente: Grados Adicionales)				
	Lineal	Cuadrático	Cúbico	Cuártico
	(1)	(2)	(3)	(4)
I(Asig. Reprobadas \geq 2)	0,073 (0,038)	-0,021 (0,072)	-0,032 (0,149)	-0,051 (0,378)
Asignación	-0,538** (0,008)	-0,139* (0,070)	-0,123 (0,21)	-0,089 (0,674)
Interacción	0,301** (0,021)			
Asignación ²		0,4** (0,07)	-0,026 (0,081)	-0,045 (0,381)
Interacción ²		-0,419** (0,083)		
Asignación ³			-0,442 (0,29)	0,005 (0,084)
Interacción ³			0,442 (0,298)	
Asignación ⁴				0,499 (1,14)
Interacción ⁴				-0,499 (1,14)
Constante	5,22** (0,008)	5,22** (0,008)	5,22** (0,008)	5,22** (0,008)
R^2 ajustado	0,354	0,354	0,354	0,354
AIC	631.182	631.172	631.174	631.176

** $p < 0,01$, * $p < 0,05$

Bibliografía

- Amadio, M. (1996). *La repetición escolar en la enseñanza primaria: una perspectiva global*. Oficina Internacional de Educación.
- Arellano, J. P. (2001). La reforma educacional chilena. *Revista de la CEPAL*.
- Bellei, C. (2010). Using a regression-discontinuity design to evaluate the causal impact of a compensatory educational program: Evidence on the effect of in-school professional development in low-performing schools in Chile. *Harvard Graduate School of Education*.
- Bellei, C. c. (2013). Situación educativa de América Latina y el Caribe: hacia la educación de calidad para todos al 2015. Informe UNESCO. *Santiago de Chile: Imbunche Ltda*.
- Díaz, J., Grau, N., Reyes, T., & Rivera, J. (2016). The impact of grade retention on juvenile crime.
- Dupriez, V., Dumay, X., & Vause, A. (2008). How do school systems manage pupils' heterogeneity? *Comparative Education Review*, 52(2):245–273.
- Epple, D. & Romano, R. E. (1998). Competition between private and public schools, vouchers, and peer-group effects. *American Economic Review*, pages 33–62.
- García-Huidobro, J. (2000). Notas para exposición en Seminario “Educación, pobreza y deserción escolar”, Santiago. pages 61–69.
- Gelman, A. & Imbens, G. (2014). Why high-order polynomials should not be used in regression discontinuity designs. Technical report, National Bureau of Economic Research.
- Greene, J. P. & Winters, M. A. (2009). The effects of exemptions to Florida's test-based promotion policy: Who is retained? Who benefits academically? *Economics of Education Review*, 28(1):135–142.

- Hill, A. J. (2014). The costs of failure: Negative externalities in high school course repetition. *Economics of Education Review*, 43:91–105.
- Hirsch Jr, E. (2007). La comprensión lectora requiere conocimiento de vocabulario y del mundo: Hallazgos científicos sobre el bajón de cuarto grado y el estancamiento en los puntajes nacionales de comprensión. *Estudios Públicos*, (108):229–252.
- Imbens, G. W. & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635.
- Imbens, G. W. & Wooldridge, J. M. (2009). Recent developments in the Econometrics of program evaluation. *Journal of Economic Literature*, 47(2):5–86.
- Jacob, B. A. & Lefgren, L. (2009). The effect of grade retention on high school completion. *American Economic Journal: Applied Economics*, pages 33–58.
- Jimerson, S. R. (1999). On the failure of failure: Examining the association between early grade retention and education and employment outcomes during late adolescence. *Journal of School Psychology*, 37(3):243–272.
- Jimerson, S. R., Anderson, G. E., & Whipple, A. D. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools*, 39(4):441–457.
- Lee, D. S. (2008). Randomized experiments from non-random selection in US House elections. *Journal of Econometrics*, 142(2):675–697.
- Lee, D. S. & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2):655–674.
- Lee, D. S. & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature*, 48(2):281–355.
- LLECE (2015). TERCE en la mira: ¿Es la repitencia efectiva? 1.
- Manacorda, M. (2006). Grade failure, drop out and subsequent school outcomes: Quasi-experimental evidence from uruguayan administrative data.

- Manacorda, M. (2010). The cost of grade retention. CEPR Discussion Papers 7889.
- Manacorda, M. (2012). The cost of grade retention. *Review of Economics and Statistics*, 94(2):596–606.
- McCoy, A. R. & Reynolds, A. J. (1999). Grade retention and school performance: An extended investigation. *Journal of School Psychology*, 37(3):273–298.
- Mizala, A. & Torche, F. (2012). Bringing the schools back in: the stratification of educational achievement in the Chilean voucher system. *International Journal of Educational Development*, 32(1):132–144.
- Mons, N. (2007). Les nouvelles politiques éducatives: La France fait-elle les bons choix.
- Ndaruhutse, S. (2008). Grade repetition in primary schools in Sub-Saharan Africa: An evidence base for change. *London Review*.
- OCDE (2013). *PISA 2012 Results: What makes schools successful?: Resources, policies and practices*, volume IV. Organización para la Cooperación y el Desarrollo Económicos.
- Roderick, M. (1994). Grade retention and school dropout: Investigating the association. *American Educational Research Journal*, 31(4):729–759.
- Schwerdt, G., West, M. R., & Winters, M. A. (2015). The effects of test-based retention on student outcomes over time: Regression discontinuity evidence from Florida. (21509).
- UNESCO (2012). *Oportunidades perdidas: El impacto de la repetición y de la salida prematura de escuela*. Instituto de Estadística de la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura.
- Urquiola, M. & Calderón, V. (2006). Apples and oranges: Educational enrollment and attainment across countries in Latin America and the Caribbean. *International Journal of Educational Development*, 26(6):572–590.
- Valenzuela, J. P., Allende, C., Sevilla, A., & Egaña, P. (2013). La (in) movilidad del desempeño educativo de los estudiantes chilenos: Realidad, oportunidades y desafíos. *Documento de trabajo*, (8).
- Valladares, M. I. (2008). Indicadores de la Educación en Chile 2007-2008. *Santiago de Chile. Departamento de Estudios y desarrollo, Ministerio de Educación*.

Vera, G. G. (2013). Los efectos de la repitencia en tanto que política pública en cuatro países del cono sur: Argentina, Brasil, Chile y Uruguay. Un análisis en base a PISA 2009. *Revista Latinoamericana de Educación Comparada*, 4(4):59–70.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.