

La Escolarización y las Ganancias de Aprendizaje según el Nivel de Integración Socioeconómica de los Estudiantes

Emmanuel Vazquez

Documento de Trabajo Nro. 299

Junio, 2022

ISSN 1853-0168

www.cedlas.econo.unlp.edu.ar

Cita sugerida: Vazquez, E. (2022). La Escolarización y las Ganancias de Aprendizaje según el Nivel de Integración Socioeconómica de los Estudiantes. Documentos de Trabajo del CEDLAS N° 299, Junio, 2022, CEDLAS-Universidad Nacional de La Plata.

La escolarización y las ganancias de aprendizaje según el nivel de integración socioeconómica de los estudiantes *

Emmanuel Vazquez **

Resumen

Este trabajo estudia el efecto de la escolaridad sobre los aprendizajes en América Latina según el nivel de integración socioeconómica de los alumnos. Utilizando datos de las distintas ediciones de PISA para 5 países de la región (Argentina, Brasil, Chile, Costa Rica y Uruguay), el trabajo explota la variación exógena en la fecha de nacimiento de los estudiantes alrededor de la fecha de corte para el ingreso a la escuela utilizando un diseño de regresión discontinua difuso. Los resultados indican que, si bien el efecto promedio de un año de escolaridad en América Latina se sitúa alrededor de los 30 puntos en la escala de PISA (0.30 desvíos estándar), difiere sustancialmente de acuerdo con el nivel de integración socioeconómica del alumnado. Así, mientras la ganancia de aprendizaje derivada de asistir a un año escolar superior para los alumnos integrados se encuentra alrededor de los 40 puntos PISA, los alumnos segregados obtienen una mejora en su aprendizaje en matemática cercana a los 20 puntos. El mayor aprendizaje que logran los estudiantes integrados con la escolaridad se da tanto para alumnos integrados con pares más ricos como para aquellos que se encuentran integrados con compañeros más pobres, proviniendo esta ganancia de los alumnos más rezagados en su aprendizaje, que son quienes más se benefician de la escolaridad al estar integrados.

Palabras clave: segregación, efectos de pares, PISA, Diseño de Regresión Discontinua.

Códigos JEL: I24, J24

* Este trabajo constituye el primer capítulo de mi tesis doctoral denominada “Ensayos sobre segregación”, realizada bajo la dirección de Mariana Marchionni en el Doctorado de Economía de la Facultad de Ciencias Económicas de la Universidad Nacional de La Plata. El trabajo aplica la metodología y extiende los resultados de un trabajo relacionado (Marchionni y Vazquez, 2019) y se benefició de la invaluable guía y apoyo de Mariana Marchionni y de los útiles comentarios y sugerencias de Facundo Albornoz, María Lucila Berniell, Walter Cont, Leonardo Gasparini, Mariana De Santis, Eugenio Giolito, David Jaume, Dolores de la Mata, Florencia Pinto, Alberto Porto, Núria Rodríguez-Planas, Jonah Rokoff, Hernán Ruffo, Christian Ruzzier, Pablo Sanguinetti y participantes del Workshop Reporte de Economía y Desarrollo 2016 (CAF), el Primer Simposio Argentino de Economía de la Educación, la Tercera Conferencia Argentina de Econometría y la 52° Reunión Anual de la Asociación Argentina de Economía Política. Los errores que puedan permanecer son responsabilidad exclusiva del autor.

** CEDLAS, Facultad de Ciencias Económicas, Universidad Nacional de La Plata. evazquez@cedlas.org

1. Introducción

Durante décadas, científicos sociales han debatido arduamente respecto a la influencia de la escolaridad sobre los logros de aprendizaje de los estudiantes. En su influyente libro *The Bell Curve*, Herrnstein y Murray (1994) revisaron la literatura existente hasta ese momento y concluyeron que la escolaridad solo tiene un efecto marginal sobre el desempeño cognitivo de los alumnos, en tanto las habilidades que determinan ese desempeño son fundamentalmente heredadas. Desde aquel entonces, la acumulación de evidencia cuasi-experimental a través del método de variables instrumentales (Neal y Johnson 1996; Hansen, Heckman, y Mullen 2004, Cascio y Lewis 2006) y de diseños de regresión discontinua (Cahan y Davis 1987, Cahan y Cohen 1989, Luyten 2006, Luyten et al. 2008, Cliffordson 2010) ha contribuido a cerrar la controversia y generar un consenso respecto a que el efecto de la escolaridad sobre el logro cognitivo de los estudiantes existe y es de tamaño considerable.¹

Las ganancias de aprendizaje obtenidas con la escolarización dependen de numerosos factores, entre los que se destaca la influencia de los compañeros de clase (Sacerdote, 2011). La importancia de los pares en el proceso educativo ha sido destacada desde hace mucho tiempo, al menos desde la publicación del bien conocido e influyente Reporte Coleman (1966), donde se argumentó que el acceso a pares más aventajados podría proveer a los estudiantes de bajos ingresos con visiones valiosas y redes de contactos al momento de entrar a la educación superior y al mundo del trabajo. Desde entonces, proliferaron trabajos que estudiaron los efectos de pares sobre el aprendizaje mirando simplemente si los estudiantes que asistían a escuelas de mayor nivel socioeconómico promedio tenían un mejor desempeño educativo, un hallazgo que fue habitual en la literatura (ver, por ejemplo, Rumberger y Palardy 2005). Aunque recientemente surgieron trabajos que explotan la asignación aleatoria para identificar un efecto causal del logro de los compañeros (Zimmerman 2003; Whitmore 2005; Carrell, Fullerton y West 2009; Lyle 2009, Duflo, Dupas y Kremer 2011; Carrell, Sacerdote y West 2013, De Giorgi y Pellizzari 2013; Booij, Leuven y Oosterbeek 2017), al momento existe solo evidencia empírica limitada sobre los canales de los efectos de pares (Feld y Zölitz 2017). Aún más importante, los estudios se han centrado en evaluar el efecto de tener compañeros más ricos o hábiles, pero no han estudiado la incidencia en los aprendizajes de tener compañeros diferentes *per se*.

La literatura que estudia las consecuencias de la segregación socioeconómica en las escuelas ha postulado frecuentemente que la integración en las aulas contribuye a la

¹ El consenso sobre el efecto de la escolaridad en los resultados cognitivos ha ido creciendo con el tiempo hasta el punto de resultar familiar el uso de una “regla del pulgar” para cuantificar las ganancias de aprendizaje que obtienen los estudiantes en un año escolar, que se coloca típicamente entre los 25 y los 40 puntos en la escala del Programa Internacional para la Evaluación de los Estudiantes (PISA) (Woesmann, 2016, OECD, 2019).

cohesión social, al promover la tolerancia y reducir los prejuicios, y a la equidad, al mejorar las oportunidades de los más pobres (Durlauf 2006, Wilson 1987). Sin embargo, la presencia de compañeros de diferente nivel socioeconómico en el aula podría tener también efectos positivos sobre los aprendizajes, los cuales no han sido explorados suficientemente. Las teorías de la psicología educativa indican que el aprendizaje en salones de clases diversos, donde los estudiantes de diferentes contextos comunican sus diferentes experiencias y perspectivas, animan a los estudiantes a pensar de maneras más complejas (Bransford y Schwartz, 1999; Wigfield et al. 2006). A su vez, como los estudiantes de diferentes contextos a menudo traen conocimientos culturales y perspectivas sociales diferentes a la escuela, se ha argumentado que las clases con grupos de estudiantes más diversos son más propensas a mejorar el pensamiento crítico al exponer a los estudiantes a nueva información y comprensión (Hawley, 1971; Orfield et al. 2008). En este sentido, la integración socioeconómica en las escuelas podría ir más allá de su contribución a la cohesión social y la equidad y resultar también en un entorno de aprendizaje más eficiente que mejore el desempeño de los alumnos de todas las clases sociales.

Este trabajo contribuye a la literatura que estudia el efecto de la escolaridad sobre los aprendizajes y las consecuencias de la segregación socioeconómica explorando la presencia de efectos compañero de una manera novedosa. Utilizando datos de las ediciones 2003, 2009, 2012, 2015 y 2018 del Programa Internacional para la Evaluación de Estudiantes (PISA, por sus siglas en inglés) para 5 países de América Latina (Argentina, Brasil, Chile, Costa Rica y Uruguay), el trabajo explota la variación exógena en la fecha de nacimiento de los estudiantes alrededor de la fecha de corte para el ingreso a la escuela utilizando un diseño de regresión discontinua difuso. Dado que en estos países existen leyes y regulaciones que establecen la obligatoriedad del ingreso al primer año de educación primaria para alumnos con determinada edad cumplida a una fecha específica, es posible encontrar diferencias en los grados escolares al que asisten estudiantes de casi la misma edad y aislar el efecto de un año de educación formal sobre el puntaje PISA comparando el rendimiento de estudiantes nacidos apenas antes y después de la fecha de corte. De esta manera, se estima el efecto promedio de asistir a un año escolar superior a los 15 años sobre los aprendizajes y la heterogeneidad de este efecto, aquí llamado “efecto de la escolaridad”, entre alumnos segregados e integrados, esto es, de acuerdo a si los estudiantes tienen compañeros parecidos o diferentes en términos de su estatus socioeconómico.

Los resultados hallados en este trabajo indican que el efecto de la escolaridad sobre el aprendizaje es de tamaño considerable y difiere sustancialmente de acuerdo al nivel de integración socioeconómica del alumnado. El efecto promedio estimado de un año de escolaridad en América Latina se sitúa alrededor de los 30 puntos en la escala de PISA (0.30 desvíos estándar), en línea con la literatura previa, pero ese promedio esconde una sustancial heterogeneidad. La mejora en las habilidades y el conocimiento que resultan de un año extra de escolaridad en los países analizados es sustancialmente mayor para los

alumnos que interactúan en la escuela con pares más diferentes en términos de su status socioeconómico que para aquellos que se encuentran segregados e interactuando con compañeros más similares. El mayor aprendizaje que logran los estudiantes integrados con la escolaridad se da tanto para alumnos integrados con pares más ricos como para aquellos que se encuentran integrados con compañeros más pobres, siendo la ganancia similar en ambos casos. A su vez, esta mejora en el aprendizaje asociada a la integración proviene fundamentalmente de los alumnos más rezagados en su aprendizaje, que son quienes más incrementan su rendimiento en la prueba PISA con la escolaridad al estar integrados. Siendo que los alumnos de bajo rendimiento segregados tienen también una mayor probabilidad de abandono escolar en el futuro, el trabajo invita a reflexionar sobre el potencial de las políticas que integren socioeconómicamente a estos estudiantes no solo desde el punto de vista de la equidad, sino también desde la eficiencia asignativa.

El resto del trabajo está organizado de la siguiente forma. La sección 2 presenta la metodología y los datos utilizados, mientras que la sección 3 discute la validez de la estrategia empírica. Los resultados son presentados en la sección 4, que muestra la heterogeneidad del efecto de la escolaridad sobre el aprendizaje en América Latina de acuerdo con el nivel de integración socioeconómica del alumnado. Finalmente, la sección 5 presenta las conclusiones y discute la relevancia y las limitaciones de los hallazgos.

2. Metodología y datos

La evaluación del impacto de la escolaridad sobre los aprendizajes requeriría idealmente la asignación aleatoria de individuos a grupos de tratamiento y control, en el que a los individuos asignados al primer grupo se les brinde de alguna manera un año adicional de escolaridad en relación al segundo grupo. Aunque no existen en el mundo experimentos sociales de este tipo, las leyes y regulaciones escolares de muchos países establecen que los niños con cierta edad específica cumplida a una fecha de corte particular (por ejemplo, 6 años cumplidos al 30 de junio) deben obligatoriamente inscribirse en la escuela para iniciar su proceso de educación formal. Esta situación genera que niños con tan solo días de diferencia de edad sean asignados exógenamente (por una ley) a iniciar la escuela un año antes o un año después según si su nacimiento tuvo lugar antes o después de esa fecha de corte. En tanto la posibilidad de manipular la fecha de nacimiento alrededor de este umbral es muy limitada, los grupos de estudiantes nacidos apenas antes y apenas después de esta fecha de corte solo diferirían en promedio en algunos días de edad y en que el grupo de estudiantes nacido antes de la fecha de corte tendría un año adicional de escolaridad, brindando una oportunidad única de evaluar de manera no experimental el efecto de la escolaridad sobre los aprendizajes comparando el aprendizaje promedio de los nacidos antes y después del umbral en un diseño de regresión discontinua *sharp* que estima:

$$\beta_S = \lim_{B \rightarrow B_0^-} E(Y_i | B_i = B) - \lim_{B \rightarrow B_0^+} E(Y_i | B_i = B) \quad (1)$$

donde Y_i es una medida del aprendizaje del estudiante i y B_i su fecha de nacimiento; B_0 es la fecha de corte y la probabilidad de tratamiento (tener un año adicional de escolaridad) es $T_i = 1 \{B_i \leq B_0\}$. Bajo el supuesto de que el aprendizaje del estudiante i hubiera sido el mismo si hubiera nacido apenas antes o apenas después de la fecha de corte, la ecuación (1) captura el efecto tratamiento promedio local en el punto de corte.

Aunque un diseño de regresión discontinua *sharp* sería válido en un contexto en el que hubiera un cumplimiento perfecto de las leyes y regulaciones que establecen la fecha de corte para el ingreso, en la práctica estas leyes y regulaciones no se cumplen a la perfección y por ende la probabilidad de tratamiento no es una función determinística que salta de 0 a 1 en la fecha de corte como en la ecuación (1), sino que la probabilidad de tener un año adicional de escolaridad solo presenta una discontinuidad en ese umbral. Este diseño de regresión discontinua difuso (*fuzzy*) requiere para recuperar el efecto promedio local del tratamiento estimar el cociente entre el cambio en el aprendizaje promedio en el umbral y el cambio en la proporción de estudiantes tratados (con un año adicional de escolaridad) también en el umbral:

$$\beta_F = \frac{\lim_{B \rightarrow B_0^-} E(Y_i | B_i = B) - \lim_{B \rightarrow B_0^+} E(Y_i | B_i = B)}{\lim_{B \rightarrow B_0^-} \Pr(T_i = 1 | B_i = B) - \lim_{B \rightarrow B_0^+} \Pr(T_i = 1 | B_i = B)} \quad (2)$$

En un diseño difuso, la ecuación (2) es igual al efecto tratamiento promedio en el punto de corte para aquellos individuos inducidos a cambiar el estado del tratamiento en ese punto de discontinuidad, siempre que se mantengan los dos siguientes supuestos: *monotonidad* (es decir, el cruce de B_i por B_0 no causa al mismo tiempo que algunos individuos tomen el tratamiento y otros lo rechacen) y *exclusión* (es decir, el cruce de B_i por B_0 no afecta a Y_i excepto a través de su efecto sobre la recepción del tratamiento).² Estos supuestos serán discutidos en la siguiente sección.

En vista de la ausencia de datos experimentales para responder la pregunta de investigación de este trabajo, la metodología que se empleará para evaluar el efecto de la escolaridad sobre los aprendizajes es un diseño de regresión discontinua difuso, lo cual requiere de datos que provean una medida del aprendizaje de los estudiantes y a su vez capten individuos nacidos antes y después de la fecha de corte y en distintos grados escolares. La

² Para más detalles, puede consultarse Hahn et al. (2001) e Imbens y Lemieux (2008).

prueba PISA (OECD, 2019)³ proporciona una fuente de información ideal puesto que a diferencia de otros programas que también evalúan los aprendizajes, como por ejemplo TIMMS o las evaluaciones que el LLECE realiza en América Latina,⁴ PISA no define su población objetivo en términos de grados escolares, sino en función de la edad. Específicamente, para cada país PISA es representativa de los alumnos entre 15 años y 3 meses y 16 años y 2 meses (en adelante, alumnos de 15 años) que han terminado al menos 6 cursos de enseñanza obligatoria,⁵ lo que permite contar con una medida del aprendizaje diferencial obtenido por individuos nacidos antes y después de la fecha de corte en diferentes países y que asisten a distintos grados escolares.

Lamentablemente, la fecha exacta de nacimiento de cada alumno no está disponible en los datos PISA, sino solo el mes de nacimiento, por lo que, siguiendo la metodología de Marchionni y Vazquez (2019), las medias en la ecuación (2) son estimadas mediante un polinomio de grado 0 (promedios simples) para dos anchos de banda diferentes, que incluyen nacer 1 o 2 meses antes y después de la fecha de corte de ingreso a la escuela. Específicamente, β_F en la ecuación (2) es estimado por mínimos cuadrados en dos etapas (Hahn, Todd y van der Klaauw, 2001) para las observaciones cuyo nacimiento se encuentra 1 o 2 meses alrededor de la fecha de corte, regresando el puntaje en la prueba PISA en una variable *dummy* que indica la asistencia a un grado escolar superior, e instrumentando esta última con un variable *dummy* que indica si el alumno nació antes de la fecha de corte para el ingreso a la escuela en su país. Como el puntaje en PISA se mide a través de la metodología de valores plausibles, que proporciona para cada estudiante una serie de

³ PISA es un programa desarrollado por la OECD con el objetivo de medir las competencias de los estudiantes cercanos al final de su educación obligatoria en las áreas de lengua, matemática y ciencias, y más generalmente evaluar si han adquirido los conocimientos y habilidades necesarios para enfrentar los desafíos que plantea la sociedad actual. Las pruebas PISA se realizan desde el año 2000 cada 3 años en los países miembros de la OECD y en un grupo de países socios cuyo número ha ido creciendo en el tiempo. Además de los exámenes, el programa recoge información sobre los estudiantes y las escuelas empleando un cuestionario armonizado entre países para los alumnos y directores de las escuelas.

⁴ TIMMS (*Trends in International Mathematics and Science Study*), un programa pionero en la evaluación de los estudiantes a nivel internacional, evalúa los aprendizajes en matemáticas y ciencias de alumnos de cuarto y octavo grado en 64 países, mientras que las pruebas del LLECE (Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación) evalúan los aprendizajes en lectura, matemática y ciencias de alumnos de tercer y sexto grado en 18 países de América Latina.

⁵ Para obtener resultados representativos de esta población de interés, PISA utiliza un muestreo estratificado bi-etápico. En la primera etapa, se toma una muestra de todas las escuelas elegibles a nivel nacional (aquellas que tienen alumnos de 15 años), con una probabilidad de selección proporcional al tamaño de cada escuela. Antes del muestreo, las escuelas elegibles son asignadas a grupos mutuamente excluyentes llamados estratos, para mejorar la precisión de las estimaciones. En la segunda etapa, cada escuela seleccionada elabora una lista de sus alumnos de 15 años y se define para cada país un tamaño de cluster objetivo (TCO) -típicamente 35 estudiantes. Cuando la cantidad de alumnos en la lista elaborada por la escuela supera el TCO, se toma una muestra de alumnos del tamaño de cluster objetivo con igual probabilidad de selección para cada uno. En el caso de que la cantidad de alumnos en la lista elaborada por la escuela sea inferior al TCO, se toman todos los estudiantes de 15 años de esa escuela (OECD, 2010).

valores que representan el rango de habilidades que el alumno podría razonablemente tener (Wu y Adams 2002), se sigue la metodología recomendada por los diseñadores del programa, realizando estimaciones diferentes de β_F para cada valor plausible y luego promediándolas, ajustando los errores estándar en consecuencia (OECD, 2009).⁶

Aunque el programa PISA evalúa a estudiantes de muchos países que aplican reglas de ingreso a la escuela con fechas de corte específicas para el inicio de la educación formal, el énfasis del presente trabajo se encuentra en un subgrupo de estos países pertenecientes a la región de América Latina (Argentina, Brasil, Chile, Costa Rica y Uruguay) debido a la escasez de estimaciones del impacto de la escolaridad en los aprendizajes para esta región y a que los sistemas educativos de América Latina se caracterizan por ser de los más segregados socioeconómicamente del mundo (Vazquez, 2016), con fuertes disparidades en el grado de integración del alumnado en las aulas. Esta característica de los países de la región ofrece una fuente de variabilidad importante para estudiar si las ganancias de la escolaridad en los aprendizajes son diferentes en contextos donde los estudiantes se relacionan casi exclusivamente con alumnos de su mismo nivel socioeconómico que si por el contrario se encuentran integrados y relacionándose en la escuela con pares más diferentes en términos de su estatus socioeconómico.

Las reglas de entrada que determinan las leyes y regulaciones de cada país analizado e influyen en el año escolar al que asisten las diferentes cohortes de estudiantes que han sido evaluadas en PISA se muestran en la Tabla 1, la cual presenta además el número de observaciones empleadas en las estimaciones para los distintos anchos de banda utilizados en el trabajo. Así, por ejemplo, en el caso de Argentina las leyes determinan que para ingresar a la escuela primaria los individuos deben tener 6 años cumplidos al 30 de junio del año en curso, lo que hace que las cohortes evaluadas en PISA que nacieron en junio deberían haber ingresado antes que los nacidos en julio y tener un año adicional de educación que estos, siendo el año teórico al que deberían asistir los primeros el 11vo grado y los segundos el 10mo grado. En general, salvo en Costa Rica que tiene una edad de ingreso más tardía, las cohortes de alumnos evaluadas en PISA deberían estar en el grado 11 si nacieron antes de la fecha de corte y en el grado 10 si nacieron después (10 y 9, respectivamente, en Costa Rica). Lógicamente, no todos los nacidos antes de la fecha de corte asisten al año superior, debido al cumplimiento imperfecto de las reglas y a la repitencia, por lo que todas las estimaciones que se realizan corresponden a las de un diseño difuso y las especificaciones siempre incluyen un control por repitencia.^{7,8} La cantidad total

⁶ Los errores estándar se estiman como recomiendan los diseñadores del programa, utilizando el método de muestras balanceadas repetidas (BRR, por sus siglas en inglés) y aumentándolos en la estimación de la varianza de imputación asociada a la variabilidad de estos valores plausibles. Ver OECD (2009) para más detalles.

⁷ Otra forma de lidiar con el problema que plantea la repitencia involucra realizar restricciones al tamaño de la muestra. Véase por ejemplo Marchionni y Vazquez (2019).

de alumnos evaluados nacidos 1 y 2 meses alrededor del umbral son especificadas en la última columna de la tabla. Todas estas observaciones son empleadas en estimaciones del efecto de la escolaridad en los aprendizajes para el conjunto de alumnos en los 5 países con efectos fijos por país (y efectos fijos por año en las especificaciones preferidas), proveyendo luego estimaciones en muestras separadas de acuerdo al nivel de integración socioeconómica del alumnado.

A los efectos de medir el nivel de integración socioeconómica de los estudiantes y separar la muestra entre alumnos integrados y segregados, el presente trabajo utiliza como principal medida de nivel socioeconómico de los estudiantes el Índice de status económico, social y cultural (*escs*, por sus siglas en inglés) construido por PISA a partir del mayor índice de estatus ocupacional del padre o la madre,⁹ el mayor grado de educación alcanzado por los padres representado en años de escolarización, y un índice de posesiones familiares que se obtiene preguntando a los estudiantes si tienen disponible en el hogar una serie de bienes y servicios.¹⁰ En base a este índice se computan percentiles de nivel socioeconómico para cada alumno dentro de un país en cada año ($p_i=1, 2, 3, \dots, 99, 100$), se calcula el número de percentil al que pertenece el estudiante mediano (de acuerdo con el mismo indicador de bienestar) en la escuela a la que asiste cada alumno (p_{Mi}) y luego se computan las distancias percentílicas absolutas de cada estudiante respecto al estudiante mediano de su escuela como la diferencia entre ambos valores, en valor absoluto ($|p_i - p_{Mi}|$). Siguiendo a Watson (2009), mayores distancias percentílicas implican una interacción con pares más distintos, más alejados del alumno en la distribución del ingreso/nivel socioeconómico. Luego, se usa una definición de segregación *relativa* que permite dividir la muestra de una manera balanceada en términos de observaciones: se considera a un alumno relativamente integrado (segregado) si la distancia percentílica absoluta es mayor (menor) a la distancia mediana para América Latina (que es igual a 14 si se utiliza el índice *escs*).¹¹

Para ayudar a visualizar la principal definición de integración aquí empleada, el Gráfico 1 muestra un histograma suavizado de las distancias percentílicas en el conjunto de países

⁸ La ausencia de una variable que mida la repitencia previa en las rondas de PISA 2000 y 2006 hace que estas evaluaciones no sean consideradas en la muestra de estimación.

⁹ Este índice se obtiene a partir de las respuestas de los alumnos respecto a la ocupación de los padres. Específicamente, recoge los atributos de las ocupaciones que convierten la educación de los padres en ingresos y se deriva a través de una clasificación óptima de los grupos de empleo para maximizar el efecto indirecto de la educación en los ingresos laborales y para minimizar el efecto directo de la educación en los ingresos, sin tener en cuenta la ocupación (OECD, 2010). Para más información acerca de la metodología, véase Ganzeboom, De Graaf y Treiman (1992).

¹⁰ La construcción de la escala del índice de posesiones familiares y del *escs* se realizó utilizando la clasificación de Teoría de Respuesta al Ítem (IRT, por sus siglas en inglés). Véase OECD (2010, 2019) para una descripción pormenorizada de la construcción de estos índices.

¹¹ En el texto, se utilizará por simplicidad la expresión América Latina para referirse a los 5 países de la región considerados en el análisis: Argentina, Brasil, Chile, Costa Rica y Uruguay.

analizados: los alumnos en la cola inferior de las distribuciones (con $p_i - p_{Mi}$ menor a -14) y en la cola superior (con $p_i - p_{Mi}$ mayor a 14) se consideran (relativamente) integrados, mientras que el resto es considerado (relativamente) segregado. La robustez de esta medida es evaluada utilizando alternativamente el indicador de riqueza (*wealth*) y de máximos años de escolaridad de los padres (*pared*) construido por PISA,¹² y considerando a un alumno relativamente integrado (segregado) si el índice Gini de los máximos años de escolaridad de los padres en la escuela es mayor (menor) a la mediana correspondiente para América Latina.¹³

3. Validez del diseño

3.1. Logros de aprendizaje y probabilidad de asistir al año escolar superior por mes de nacimiento

A efectos de motivar el análisis, el Gráfico 2 muestra el puntaje promedio en matemática por mes de nacimiento de los *compliers*, esto es, aquellos alumnos que por su mes de nacimiento asisten al año teórico al que deberían asistir de acuerdo a las leyes y regulaciones en los 5 países analizados. Cada punto en el gráfico representa el puntaje promedio para los alumnos nacidos en un mes relativo a la fecha de corte, que se indica con una línea vertical. Puntos más cerca de esta línea representan un mes de nacimiento más cercano al corte, con los puntos a la derecha indicando un nacimiento anterior a la fecha de corte y los puntos a la izquierda señalando nacimientos en los meses posteriores a dicha fecha. Como puede observarse en el gráfico, aquellos alumnos cuyo mes de nacimiento se encuentra a la derecha de la barra vertical presentan como es esperable un rendimiento superior a los alumnos cuyo mes de nacimiento se encuentra a la izquierda del umbral, a quienes se les exigía demorar al siguiente año su ingreso a la escuela. El salto al cruzar el umbral en el nivel de aprendizaje demostrado en la prueba PISA en el área de matemáticas no es homogéneo, sino que es mayor para los alumnos integrados que para aquellos que se encuentran segregados, esto es, para quienes se encuentran interactuando con pares más diferentes en términos de su nivel socioeconómico.

La existencia de un salto en los aprendizajes al cruzar el umbral podría deberse a numerosas razones, pero aquí se argumenta que la única característica de los alumnos relevante para los aprendizajes que difiere entre los individuos nacidos apenas antes y

¹² El índice *wealth* se construye en base a las respuestas de los estudiantes sobre la posesión de una lista de bienes en el hogar de menor tamaño que el empleado para capturar la posesión de bienes y servicios en el hogar en el *escs*, mientras que el índice *pared* toma el valor máximo entre los años de escolaridad del padre y la madre del estudiante. Véase OECD (2010, 2019) para los detalles sobre la construcción de estos índices.

¹³ Por las características de los índices *escs* y *wealth*, que toman valores negativos y han sido estandarizados, no se utiliza como prueba de robustez adicional el Gini de estas medidas, sino solo el de los máximos años de escolaridad.

apenas después de la fecha de corte es su grado de exposición al tratamiento, esto es, su probabilidad de asistir al año escolar superior (el 11vo en general, salvo en Costa Rica donde es el 10mo). Antes de proveer evidencia que apoye este supuesto de *exclusión*, es necesario constatar en primera instancia que existe una discontinuidad en la función que relaciona la probabilidad de asistir al año escolar superior y la fecha de nacimiento, esto es, constatar que en efecto la exposición al tratamiento difiere entre las observaciones a la izquierda y a la derecha del umbral. El Gráfico 3 muestra la proporción de estudiantes que asisten al año escolar superior por mes de nacimiento. Así, mientras que casi ningún estudiante de los evaluados en las cohortes participantes en PISA que nacieron después de la fecha de corte se encuentran en el 11vo grado (10mo en Costa Rica), casi un 40% de los nacidos en el mes anterior al corte y cerca de un 50% de los nacidos un mes más atrás asiste a este grado. Esta discontinuidad en la exposición al tratamiento se encuentra también en las poblaciones de alumnos integrados y segregados, siguiendo en ambos casos un patrón muy similar: al cruzar el umbral, la proporción de individuos que toman el tratamiento es siempre mayor, pero el salto es siempre inferior a 1, lo cual fundamenta la necesidad de utilizar un diseño de regresión discontinua difuso.

3.2. Balance de covariables predeterminadas

Una de las pruebas más importantes para evaluar la validez del diseño de regresión discontinua implica examinar si, cerca del punto de corte, las unidades a la derecha del umbral son similares a las de la izquierda en términos de sus características observables. Formalmente, si se encuentra que X_i cruzando x_0 tiene un efecto sobre una covariable que impacta a Y_i , entonces el supuesto de *exclusión* no se mantendría y el diseño de regresión discontinua no sería válido. Cuando los RDD son interpretados como un experimento local, es natural requerir que las unidades justo por encima del umbral sean, en promedio, idénticas a aquellas unidades apenas por debajo de él en todas sus características observables previas al tratamiento, excepto por el hecho de que las primeras están desproporcionadamente expuestas al tratamiento. La presencia de diferencias significativas en las covariables predeterminadas entre estos dos grupos pondría en duda la validez del diseño.

La Tabla 2 muestra para el conjunto de los 5 países de América Latina analizados los valores promedio de un conjunto de covariables potencialmente importantes para el aprendizaje para los alumnos nacidos antes y después de la fecha de corte utilizando anchos de banda de 1 y 2 meses alrededor del umbral, junto con los p-valores del test de diferencia de medias para cada variable. A nivel general, tanto para el agregado de estudiantes en los 5 países (panel a) como en las submuestras de alumnos segregados e integrados (paneles b y c), no se encuentran diferencias significativas en estas covariables, a excepción de la existencia de una proporción significativamente más baja de alumnos no repitentes para los estudiantes nacidos antes de la fecha de corte (aquellos que de acuerdo a las leyes y

regulaciones debían ingresar a la escuela más temprano). La existencia de una mayor repitencia entre estos alumnos era esperable, en tanto quienes ingresaron antes a la escuela tuvieron una chance adicional de repetir un año que quienes ingresaron a la escuela al año siguiente. En tanto el resto de las características observables de los alumnos son en promedio las mismas, no existen razones para pensar que los alumnos nacidos con posterioridad a la fecha de corte no sean un contrafactual válido para quienes nacieron un poco antes. De hecho, de existir un sesgo, sería a la baja, siendo las estimaciones obtenidas una cota inferior de las ganancias de aprendizaje. Aún más, como la brecha de repitencia antes y después del umbral es más alta para los alumnos integrados que para los segregados, de encontrarse una mayor ganancia de aprendizaje para los integrados, la brecha hallada también sería una cota inferior, y esto solo reforzaría los resultados encontrados.¹⁴ De cualquier manera, tal como se especificó en la sección metodológica, todas las especificaciones controlaran por repitencia. En su conjunto, la evidencia presentada en la Tabla 2 indica la ausencia de desbalance de covariables y apoya la validez del diseño de regresión discontinua.

3.3. No manipulación

La validez del diseño de regresión discontinua descansa también en parte en que la manipulación exacta de la variable que se utiliza para determinar la asignación de tratamiento alrededor del punto de corte sea improbable o imposible. Aunque la imprecisión de este control es algo que no se puede probar y, a menudo, no es más que una conjetura, tiene claras predicciones observables (Lee y Lemieux, 2009). Si los individuos pueden manipular con precisión la fecha de nacimiento, la regla de asignación al tratamiento es de conocimiento público y el tratamiento es deseable (o indeseable), presumiblemente habrá algún tipo de aglomeración de individuos alrededor del umbral y, por lo tanto, un salto en la densidad de la variable fecha de nacimiento en el valor de corte. Por el contrario, si no es posible manipular con exactitud la fecha de nacimiento alrededor del punto del corte, la densidad de esta variable debería ser continua en un vecindario local cercano al punto de corte. El Gráfico 4 muestra la distribución de las fechas de nacimiento en la última edición de PISA con observaciones antes y después del corte para cada país analizado (la imagen es muy similar en los restantes años). Como era esperable, las distribuciones son relativamente uniformes en la mayoría de los países, sin una clara

¹⁴ Note además, al comparar los paneles (b) y (c) de la Tabla 2, que los alumnos integrados en la muestra de estimación son bastante parecidos en promedio a los alumnos integrados, a excepción de que los integrados tienen una tasa de repitencia mayor, un nivel socioeconómico levemente inferior, y que están mucho más concentrados en escuelas públicas (solo un 33 por ciento asiste a escuelas privadas a diferencia del 45 por ciento que concurre a estas escuelas entre los segregados). En tanto todos ellos son factores que en general se asocian a un menor rendimiento en América Latina, lo esperable a partir de las diferencias entre estos grupos no vinculadas a la segregación, sería que estos alumnos se beneficiaran menos de la educación formal, algo que veremos que no ocurre al estimar las ganancias de aprendizaje de la escolaridad para ambos grupos.

discontinuidad en la fecha de corte. La excepción principal es Argentina, donde es posible apreciar una caída un poco mayor en la densidad al cruzar el umbral.

Discontinuidades en la distribución de las fechas de nacimiento alrededor de la fecha de corte pueden deberse a algún control de los padres de la fecha de nacimiento de sus hijos cerca del umbral o también a diferencias en las tasas de abandono escolar entre los grupos de alumnos a ambos lados del umbral. La primera hipótesis parece menos probable, ya que los padres no tienen un control preciso de esta fecha, e incluso si lo tuvieran, no está claro si es deseable para los padres que sus hijos ingresen a la educación primaria con anterioridad. Discontinuidades en la distribución de las fechas de nacimiento también pueden surgir si los estudiantes abandonan la escuela después de terminar el grado anterior al año escolar superior (esto es, luego del 10mo grado, a excepción de Costa Rica donde sería el 9no). Si los estudiantes nacidos antes de la fecha de corte tienen más probabilidades de abandonar la escuela que los nacidos después de esa fecha por esta razón, las estimaciones que aquí se muestran podrían sufrir un sesgo al alza. La muestra de estudiantes nacidos antes del punto de corte sería más seleccionada en ese caso, con una mayor intención de permanecer en la educación y presumiblemente un mayor rendimiento. Aunque la inspección de los histogramas sugiere que este no es el caso, al replicar el análisis de RD utilizando la asistencia escolar en lugar de las puntuaciones de PISA para todos los países donde los datos del censo o de las encuestas de hogares lo permiten, no se encuentra evidencia a favor de una discontinuidad en las tasas de abandono escolar alrededor de la fecha de corte.¹⁵

Adicionalmente, se realizó el test formal de manipulación propuesto por Frandsen (2017), que es más adecuado para una variable discreta como es aquí el mes de nacimiento que la prueba tradicional de discontinuidad de la densidad en el umbral (McCrary, 2008). El test se basa en tres puntos del soporte de la variable de asignación: el umbral (el punto más pequeño en el soporte de los tratados) y sus dos puntos inmediatamente adyacentes. Limitaciones en los datos disponibles impiden realizar una prueba de este tipo en Argentina y para algunos años en Uruguay (los correspondientes a PISA 2012, 2015 y 2018), pero los resultados para Brasil, Chile, Costa Rica y Uruguay (en este último caso en base a PISA 2003 y 2009) no muestran evidencia de manipulación, aunque la evidencia para Brasil 2015 y Costa Rica 2018 es menos concluyente.¹⁶ Sin embargo, en vista de la evidencia general

¹⁵ Específicamente, se utiliza el Censo de Población 2010 en Argentina, la *Pesquisa Nacional por Amostra de Domicilios* (PNAD) de Brasil 2012, y el Censo de Población 2011 en Uruguay. Estas fuentes poseen una medida tanto de asistencia escolar como de mes de nacimiento para los estudiantes de 15 años. En todos los casos, las diferencias en las tasas de asistencia escolar antes y después de la fecha de corte con los anchos de banda de 1 y 2 meses son pequeñas (alrededor de un punto porcentual) y no estadísticamente significativas. Los resultados están disponibles a pedido.

¹⁶ Específicamente, se realizó el test para todas las posibles cotas k y no fue posible rechazar la hipótesis nula que implica que no hay manipulación para ningún k en ningún país y año, excepto en Brasil 2015 y Costa Rica 2018 donde es necesario asumir una mayor curvatura de la función de densidad en el umbral (k debe ser

de todos los años, el análisis en base a los censos y encuestas de hogares y la inspección gráfica de los histogramas, no parece haber justificación para sugerir que los resultados deban interpretarse con precaución por estas excepciones menores. En conjunto, la evidencia presentada en esta sección respalda entonces la validez de la estrategia de identificación, por lo que a continuación se presentan los resultados hallados.

4. Resultados

Esta sección muestra, para el agregado de los estudiantes en los 5 países de América Latina analizados, las estimaciones de regresión discontinua difusa de asistir a un año escolar superior sobre el puntaje PISA en matemáticas. Se presentan los resultados correspondientes a seis especificaciones. Las estimaciones (1) a (3) utilizan las observaciones en un ancho de banda de 2 meses alrededor de la fecha de corte de ingreso a la escuela, mientras que las estimaciones (4) a (6) emplean las observaciones en la ventana de 1 mes alrededor de dicho corte. Todas las especificaciones utilizan efectos fijos por país y controlan por las otras variables que pueden capturar diferencias en la trayectoria académica de los estudiantes (asistencia a uno o más de un año de educación preescolar y repitencia), así como también el género y el nivel socioeconómico del estudiante medido por el índice PISA de riqueza. Las especificaciones (2) y (5) agregan además controles a nivel de escuela (un indicador de si el clima educativo de la escuela a la que asiste el alumno es alto, definido en base a si el promedio de la máxima educación de los padres de la escuela es mayor a 12 años), mientras que las estimaciones que se muestran en las columnas (3) y (6) de las tablas a continuación adicionan efectos fijos por año para capturar heterogeneidades entre las distintas ediciones de PISA. Se evita el uso de más controles para evitar la pérdida de observaciones por datos faltantes, especialmente en las variables a nivel escuela, siendo la especificación preferida la que emplea controles a nivel país, estudiante, escuela y efectos fijos por año, con el ancho de banda de 1 mes alrededor de la fecha de corte (especificación (6)).

Los resultados más generales muestran que, en línea con la literatura reciente, el efecto de la escolaridad sobre los aprendizajes de los estudiantes existe y es de tamaño considerable. La Tabla 3 muestra que la asistencia a un año escolar superior otorga un incremento promedio levemente superior a los 30 puntos PISA o 0.30 desvíos estándar,¹⁷ en línea el resto de la literatura. En efecto, en tanto algunos autores, basados en que las ganancias obtenidas durante un año escolar en la mayoría de los exámenes nacionales e

mayor a 0.058 y 0.037 respectivamente para evitar el rechazo a un 5% de significatividad). Como reconoce Frandsen (2017) “un k más pequeño implica un test con mayor potencia, pero también puede detectar manipulación cuando no la hay”.

¹⁷ En la primera edición de PISA (año 2000) se escalaron los puntajes de la prueba para que 100 puntos PISA sean equivalentes a 1 desvío estándar entre los países de la OECD.

internacionales se encuentran entre un cuarto y un tercio de desvío estándar, estiman una ganancia de aprendizaje de entre 25 y 33 puntos PISA (Woesmann, 2016), otros autores utilizan las estimaciones de modelos multinivel de la OECD, que colocan la ganancia de aprendizaje en un rango de 35 a 40 puntos PISA entre grados adyacentes (OECD, 2019). La evidencia provista en este trabajo para América Latina es consistente con estos hallazgos, aunque, como se verá a continuación, este promedio esconde una notable heterogeneidad.

Una de las heterogeneidades más importantes que tiene el potencial de moldear el proceso de aprendizaje y ha sido escasamente evaluada, es aquella que resulta de la integración socioeconómica en las aulas. La escasa evidencia existente en general indica que los niños de familias menos aventajadas se benefician académicamente de la integración socioeconómica mientras que aquellos que provienen de las clases medias no se ven perjudicados y podrían beneficiarse de ella (Mickelson y Bottia, 2010). Los resultados para América Latina que se encuentran en este trabajo están en línea con estos hallazgos, en tanto la ganancia de aprendizaje que obtienen con la escolaridad los alumnos integrados es superior a la que obtienen los alumnos segregados, esto es, aquellos que se encuentran interactuando en la escuela con pares más parecidos a ellos en términos de su nivel socioeconómico. Para ilustrar este punto, la Tabla 4 muestra las estimaciones de regresión discontinua difusa que resultan de separar la muestra total de los 5 países de la región entre alumnos integrados y segregados de acuerdo a si la distancia percentílica absoluta al mediano de la escuela es mayor o menor a la distancia mediana de América Latina. Los resultados que se muestran en esta Tabla ponen en evidencia que la ganancia de aprendizaje derivada de la escolaridad para los alumnos integrados se encuentra alrededor de los 40 puntos PISA, mientras que los alumnos segregados obtienen una mejora en su aprendizaje en matemática cercana a los 20 puntos.¹⁸

La existencia de un efecto heterogéneo en las ganancias de aprendizaje de acuerdo al nivel de integración socioeconómica prevaleciente en el aula es un resultado robusto, que no depende de la medida de bienestar utilizada para capturar el nivel socioeconómico de los estudiantes ni de la medida que se emplee para capturar el grado de integración en la escuela. La Tabla 5 muestra que la brecha encontrada se mantiene si en lugar de usar el Índice de status económico, social y cultural (*escs*) como medida de bienestar, se utiliza el Índice de riqueza de PISA (*wealth*) o el Índice de máximos años de educación de los padres (*pared*). Asimismo, la Tabla 6 permite apreciar que los resultados tampoco cambian si se utiliza otra medida de integración relativa, como el coeficiente de Gini de los años de educación de los padres de la escuela, ya que los estudiantes que asisten a establecimientos

¹⁸ Al abrir la muestra por país y por nivel de integración, puede observarse que si bien el impacto estimado de la escolaridad es superior para los alumnos integrados en los 5 países analizados, la brecha es especialmente grande en Argentina y Uruguay (resultados disponibles a pedido).

donde existe una desigualdad educativa de los padres mayor a la mediana de América Latina obtienen una ganancia de aprendizaje mayor que el resto.

La pregunta natural por realizarse luego de observar estos resultados es la siguiente: ¿De dónde provienen las ganancias de la integración? Entre los estudiantes considerados integrados en este estudio hay quienes se encuentran integrados con pares más ricos y quienes se encuentran integrados con pares más pobres (los primeros con distancia percentílica al mediano de la escuela negativa y los segundos con distancia positiva, siempre mayor en valor absoluto a la mediana de América Latina). Sin embargo, al estimar los efectos de la escolaridad separadamente para estos dos grupos, se encuentra una ganancia similar en ambos casos (Tabla 7), independientemente del indicador de bienestar utilizado para medir el nivel socioeconómico (Tabla 8), lo que sugiere que las mayores ganancias de la escolaridad entre alumnos integrados no provienen de la integración con pares más aventajados socioeconómicamente, sino que se da tanto para estudiantes integrados con alumnos mucho más ricos como para aquellos que están integrados con compañeros mucho más pobres. Este es un resultado importante, dado que la literatura previa ha enfatizado el efecto positivo para los más pobres de interactuar con pares de mayor nivel socioeconómico, pero ignorado el potencial efecto positivo sobre los más ricos de la diversidad que les otorga interactuar con pares más pobres.

Aunque no se observa un efecto diferente de la escolaridad entre alumnos integrados con pares más ricos o más pobres, sí puede apreciarse una mayor ganancia de aprendizaje para los alumnos de bajo rendimiento que se encuentran integrados socioeconómicamente. La Tabla 9 muestra que mientras que las ganancias de la escolaridad son pequeñas (e incluso no significativamente distintas de cero) para los alumnos de rendimiento aceptable en PISA (nivel 2 o más), las ganancias entre los alumnos de bajo rendimiento (considerados analfabetos funcionales por estar debajo del nivel 2 de rendimiento en PISA), son mucho más altas, particularmente entre aquellos estudiantes que se encuentran integrados. Así, mientras que entre alumnos segregados o de buen rendimiento integrados el efecto de la escolaridad tiende a ser pequeño y no significativo (paneles a, b y c de la Tabla 9), es entre aquellos estudiantes de bajo rendimiento que se encuentran integrados (panel d) donde se observa el mayor efecto de la escolaridad. En efecto, la Tabla 10 confirma que aún si se utilizan otros indicadores de bienestar, la mayor ganancia de aprendizaje que se encuentra entre los estudiantes integrados proviene de aquellos alumnos que se hallan en la parte inferior de la distribución de puntajes. Este resultado tiene notables implicancias desde el punto de vista de la equidad, al ser estos alumnos quienes más se benefician de la educación formal en contextos de integración.

5. Discusión final y conclusiones

Este trabajo estimó de manera causal el efecto de la escolaridad sobre las ganancias de aprendizaje para los alumnos de un conjunto de países de América Latina (Argentina, Brasil, Chile, Costa Rica y Uruguay). Para ello, hizo uso de los datos provenientes de las distintas ediciones de la prueba PISA y explotó la variación exógena en la fecha de nacimiento de los estudiantes alrededor de la fecha de corte de ingreso a la escuela utilizando un diseño de regresión discontinua difuso. Los resultados hallados confirman que si bien el efecto promedio de asistir a un año escolar superior se sitúa alrededor de los 30 puntos en la escala de PISA (0.30 desvíos estándar), ese promedio esconde una sustancial heterogeneidad de acuerdo al nivel de integración socioeconómica del alumnado.

El principal aporte de este trabajo proviene de la exploración de las diferencias en las ganancias de aprendizaje derivadas de la escolaridad entre alumnos segregados e integrados, esto es, de acuerdo a si los estudiantes tienen compañeros parecidos o diferentes en términos de su estatus socioeconómico. Los resultados hallados sugieren la existencia de un mayor efecto de la escolaridad en los aprendizajes de los alumnos integrados que entre los estudiantes segregados, siendo el efecto promedio de alrededor de los 40 puntos PISA (0.40 desvíos estándar) entre los primeros y cercano a los 20 puntos PISA (0.20 desvíos estándar) entre los segundos. Las mayores ganancias de la escolaridad entre alumnos integrados no provienen de la integración con pares más aventajados socioeconómicamente, sino que se da tanto para estudiantes integrados con alumnos mucho más ricos como para aquellos que están integrados con compañeros mucho más pobres. Aún más importante desde el punto de vista de la equidad, esta mejora en el aprendizaje asociada a la integración proviene fundamentalmente de los alumnos más rezagados en su aprendizaje, que son quienes más incrementan su rendimiento en la prueba PISA con la escolaridad al estar integrados.

A diferencia de la literatura previa que ha estudiado los efectos de pares, el presente trabajo exploró la presencia de efectos compañeros de una manera novedosa. En efecto, el énfasis ha estado en evaluar si existe un efecto mayor de la escolaridad sobre el aprendizaje no para aquellos que tienen compañeros más aventajados socioeconómicamente, sino simplemente diferentes. Aunque de los resultados hallados no puede argumentarse que una mayor integración cause mayor aprendizaje, el presente trabajo provee una perspectiva diferente que invita a investigar con más profundidad si las consecuencias de la integración van más allá de su contribución a la cohesión social y a la equidad, y puede resultar también en un entorno de aprendizaje más eficiente que mejore el desempeño de los alumnos de todas las clases sociales, especialmente de los más rezagados en el proceso educativo.

Referencias

- Booij, A. S., Leuven, E. y Oosterbeek, H. (2017). Ability Peer Effects in University: Evidence from a Randomized Experiment. *Review of Economic Studies* 84(2), 547–78.
- Bransford, J. D. y Schwartz, D. L. (1999). Rethinking Transfer: A Simple Proposal with Multiple Implications. En *Review of Research in Education*, ed. Asghar Iran-Nejad y P. David Pearson. Washington, D.C.: American Educational Research Association, 61-101.
- Cahan, S. y Cohen, N. (1989). Age versus Schooling Effects on Intelligence Development. *Child-development* 60(5), 1239-49.
- Cahan, S. y Davis, D. (1987). A Between-Grade-Levels Approach to the Investigation of the Absolute Effects of Schooling on Achievement. *American Educational Research Journal* 24(1), 1- 12.
- Carrell, S. E., Fullerton, R. L. y West, J. E. (2009). Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics* 27(3), 439–64.
- Carrell, S. E., Sacerdote, B. I. y West, J.E. (2013). From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica* 81(3), 855–82.
- Cascio, E. U. y Lewis, E. G. (2006). Schooling and the AFQT: Evidence from School Entry Laws. NBER Working Paper 11113. Cambridge, Mass.: National Bureau of Economic Research.
- Cliffordson, C. (2010). Methodological issues in investigations of the relative effects of schooling and age on school performance: the between-grade regression discontinuity design applied to Swedish TIMSS 1995 data. *Educational Research and Evaluation* 16, 39-52.
- Coleman, J. S. (1966). *Equality of educational opportunity*. Washington: U.S. Govt. Print. Off., 1966. [summary report].
- De Giorgi, G. y Pellizzari, M. (2013). Understanding Social Interactions: Evidence from the Class Room.” *Economic Journal* 124 (579), 917–53.
- Duflo, E., Dupas, P. y Kremer, M. (2011). Peer effects and the impacts of tracking: evidence from a randomized evaluation in Kenya. *American Economic Review* 101(5), 1739–1774.

Durlauf, S. (2006). Groups, Social Influences, and Inequality: A Memberships Theory Perspective on Poverty Traps. En *Poverty Traps*, eds. S. Bowles, S. Durlauf, y K. Hoff. Princeton: Princeton University Press.

Feld, J., y Zölitz, U. (2017). Understanding peer effects: On the nature, estimation, and channels of peer effects. *Journal of Labor Economics* 35(2), 387-428.

Frandsen, B. (2017). Party bias in union representation elections: Testing for manipulation in the regression discontinuity design when the running variable is discrete. En M. D. Cattaneo y J. C. Escanciano (eds.), *Regression discontinuity designs: Theory and applications. Advances in econometrics* 38, 281–315. Bingley: Emerald Publishing Limited.

Frenette, M. (2008). The returns to schooling on academic performance: Evidence from large samples around school entry cut-off dates. Analytical Studies Research Paper Series No. 317.

Ganzeboom, H.B.G., De Graaf, P. M. y Treiman, D. J. (1992). A Standard International Socio-economic Index of Occupational Status. *Social Science Research* 21(1), 1-56.

Hahn, J., Todd, P., y Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica* 69(1), 201–209.

Hansen, K., Heckman, J. J. y Mullen, K. J. (2004). The Effect of Schooling and Ability on Achievement Test Scores. *Journal of Econometrics* 121(1-2), 39-98.

Hawley, (1971). *Designing Schools. Jean Piaget, Biology and Knowledge*. Chicago: Univ. of Chicago Press.

Heck, R. H. y Moriyama, K. (2010). Examining relationships among elementary schools' contexts, leadership, instructional practices, and added-year outcomes: a regression discontinuity approach. *School Effectiveness and School Improvement* 21(4), 377-408.

Herrnstein, R. J., y Murray, C. A. (1994). *The bell curve: Intelligence and class structure in American life*. Free Press.

Imbens, G., y Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142, 615–635.

Lee, D., y Lemieux, T. (2009). Regression discontinuity designs in economics. *Journal of Economic Literature* 48(2), 281–355.

Luyten, H., Peschar, J. y Coe, R. (2008). Effects of Schooling on Reading Performance, Reading Engagement, and Reading Activities of 15-Year- Olds in England. *American Education Research Journal* 45(2), 319-342.

Luyten, J. W. (2006). An empirical assessment of the absolute effect of schooling: regression-discontinuity applied to TIMSS-95. *Oxford Review of Education* 32, 397-429.

Lyle, D. S. (2009). The Effects of Peer Group Heterogeneity on the Production of Human Capital at West Point. *American Economic Journal: Applied Economics* 1(4), 69–84.

Marchionni, M. y Vazquez, E. J. (2019). The Causal Effect of an Extra Year of Schooling on Skills and Knowledge in Latin America. Evidence from PISA. *Assessment in Education: Principles, Policy & Practice* 26(4), 489-515. DOI: 10.1080/0969594X.2018.1454401.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142, 698–714.

Mickelson, R. A., y Bottia, M. (2009). Integrated education and mathematics outcomes: A synthesis of social science research. *North Carolina Law Review* 88, 993-1090.

Neal, D. y Johnson, W. R. (1996). The Role of Premarket Factors in Black- White Wage Differences. *Journal of Political Economy* 104(5), 869-95.

OECD (2009). PISA Data Analysis Manual: SPSS, Second Edition, PISA. OECD Publishing, Paris, <https://doi.org/10.1787/9789264056275-en>.

OECD (2010). PISA 2009 Results: What Students Know and Can Do – Student Performance in Reading, Mathematics and Science. OECD Publishing, Paris.

OECD (2019). PISA 2018 Results (Volume I): What Students Know and Can Do, PISA, OECD Publishing, Paris, <https://doi.org/10.1787/5f07c754-en>.

Orfield, G., Frankenberg, E., y Garces, L. M. (2008). Statement of American social scientists of research on school desegregation to the US Supreme Court in *Parents v. Seattle School District* and *Meredith v. Jefferson County*. *The Urban Review* 40(1), 96-136.

Rumberger, R. y Palardy, G. (2005). Does segregation still matter? The impact of student composition on academic achievement in high school. *Teachers college record* 107(9), 1999-2045.

Sacerdote, B. (2011). Peer effects in education: How might they work, how big are they and how much do we know thus far? En *Handbook of the Economics of Education* (Vol. 3, pp. 249-277). Elsevier.

- Tyumeneva, Y., y Kuzmina, Y. (2012). The Effect of One Extra Year of Schooling on Pisa Results: a Case of Countries with Different Tracking Systems. Higher School of Economics Research Paper WP BRP, 8.
- Vazquez, E. J. (2016). Segregación escolar por nivel socioeconómico. *Económica* 62, 121-184.
- Watson, T. (2009). Inequality and the measurement of residential segregation by income in American neighborhoods. *Review of Income and Wealth* 55(3), 820-844.
- Whitmore, D. (2005). Resource and Peer Impacts on Girls' Academic Achievement: Evidence from a Randomized Experiment. *American Economic Review* 95(2), 199–203.
- Wigfield, A. Eccles, J. S., Schiefele, U., Roeser, R.W. y Davis-Kean, P. (2006). Development of Achievement Motivation. En *Social, Emotional, and Personality*, vol. 3, 6th ed. of Handbook of Child Psychology, ed. William Damon, Richard M. Lerner y Nancy Eisenberg. New York: Wiley.
- Wilson, W. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.
- Woessmann, L. (2016). The Importance of School Systems: Evidence from International Differences in Student Achievement. *Journal of Economic Perspectives*, 30(3), 3-32.
- Wu, M. L., y Adams, R. J. (2002). Plausible Values: Why they are important. Paper presented at the International Objective Measurement workshop, New Orleans, LA.
- Zimmerman, D. J. (2003). Peer Effects in Academic Outcomes: Evidence from a Natural Experiment. *Review of Economics and Statistics* 85(1), 9–23.

Tablas

Tabla 1. Características de los sistemas educativos de los países de América Latina que inciden sobre el año escolar y características de las muestras PISA utilizadas

País	Edad de ingreso a la escuela primaria	Fecha de corte para el ingreso a la escuela	Cohortes en las muestras PISA utilizadas *	Año escolar teórico para los nacidos:		Cantidad de observaciones totales	
				Antes de la fecha de corte	Después de la fecha de corte	Ancho de banda +/- 2 meses	Ancho de banda +/- 1 mes
Argentina	6 años	30 de Junio	06/1993 - 06/1994 06/1996 - 05/1997 06/1999 - 05/2000 07/2002 - 06/2003	11	10	4,799	2,821
Brasil **	6 años	Varía por estado (referencia: 30 de Junio)	01/1993 - 12/1993 01/1996 - 12/1996 01/1999 - 12/1999	11	10	1,683	839
Chile	6 años	30 de Junio	05/1993 - 04/1994 05/1996 - 04/1997 06/1999 - 04/2000 05/2002 - 04/2003	11	10	6,995	3,708
Costa Rica	7 años ***	31 de Octubre	03/9194 - 02/1995 03/1996 - 02/1997 03/1999 - 02/2000 03/2002 - 02/2003	10	9	5,938	2,961
Uruguay	6 años	30 de Abril	04/1987 - 03/1988 04/1993 - 03/1994 05/1996 - 04/1997 05/1999 - 04/2000 05/2002 - 04/2003	11	10	6,564	3,798

Fuente: Elaboración propia en base a leyes y regulaciones detalladas en el Apéndice 1 de Marchionni y Vazquez (2019) y microdatos de PISA.

Notas: * Las cohortes que participaron de las evaluaciones PISA 2000 y PISA 2006 no han sido consideradas al no disponer de información sobre la repitencia previa del alumno, una variable de control importante que es utilizada en todas las estimaciones.

** En Brasil, el año escolar por un estudiante en el grado 11 (10) en la muestra es de hecho su 10mo (9no) año de educación formal, dado que la edad de ingreso a la primaria para las cohortes analizadas es de 7 años. El análisis en ese país se restringe a los 3 estados en los que los datos revelan que una fecha de corte uniforme (30 de junio) tuvo un cumplimiento fuerte, esto es, Amazonas, Distrito Federal y Roraima.

*** La edad de ingreso a la primaria en Costa Rica es de 6 años y 3 meses. Para las cohortes de estudiantes en la muestra, el requisito de que tengan al menos 6 años y 3 meses cumplidos al 31 de enero es equivalente a tener al menos 7 años de edad al 31 de octubre.

Tabla 2. Balance de covariables.

(a) Total América Latina. Media antes y después del corte y test de diferencia de medias.

Variable	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 1 mes		
	Media antes del corte	Media después del corte	P-valor	Media antes del corte	Media después del corte	P-valor
1 = Mujer	0.52	0.52	0.63	0.53	0.50	0.05
1 = No asistió a educación preescolar	0.10	0.07	0.73	0.09	0.07	0.26
1 = Asistió un año a educación preescolar	0.37	0.30	0.14	0.34	0.31	1.00
1 = Asistió más de un año a educación preescolar	0.53	0.63	0.09	0.57	0.63	0.54
1 = No repitente	0.62	0.70	0.00*	0.60	0.70	0.00*
1 = Vive en familia nuclear (con ambos padres)	0.72	0.72	0.38	0.72	0.72	0.76
1 = Vive con algún hermano/a	0.84	0.86	0.44	0.85	0.87	0.27
1 = Inmigrante (1ra o 2da generación)	0.03	0.03	0.41	0.04	0.03	0.43
1 = Madre económicamente activa	0.65	0.66	0.75	0.65	0.65	0.82
1 = Padre económicamente activo	0.93	0.94	0.49	0.93	0.94	0.57
Status ocupacional más alto de los padres (<i>hisei</i>)	44.21	44.24	0.95	44.07	43.77	0.51
Nivel educativo más alto de los padres (<i>pared</i>)	12.53	12.58	0.52	12.57	12.52	0.15
Riqueza familiar (<i>wealth</i>)	-0.84	-0.87	0.53	-0.86	-0.87	0.42
1 = Tiene computadora en el hogar	0.78	0.78	0.68	0.79	0.78	0.81
1 = Tiene más de 100 libros en el hogar	0.16	0.16	0.66	0.16	0.15	0.93
1 = Asiste a una escuela privada	0.39	0.38	0.55	0.38	0.39	0.59
Clima educativo de la escuela (promedio <i>pared</i>)	12.47	12.59	0.66	12.47	12.58	0.97

(b) Segregados en América Latina. Media antes y después del corte y test de diferencia de medias.

Variable	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 1 mes		
	Media antes del corte	Media después del corte	P-valor	Media antes del corte	Media después del corte	P-valor
1 = Mujer	0.53	0.53	0.38	0.54	0.51	0.10
1 = No asistió a educación preescolar	0.09	0.07	0.25	0.08	0.07	0.94
1 = Asistió un año a educación preescolar	0.37	0.29	0.24	0.34	0.31	0.69
1 = Asistió más de un año a educación preescolar	0.54	0.63	0.59	0.58	0.63	0.64
1 = No repitente	0.65	0.72	0.00*	0.64	0.72	0.00*
1 = Vive en familia nuclear (con ambos padres)	0.73	0.72	0.44	0.74	0.72	0.46
1 = Vive con algún hermano/a	0.85	0.86	0.79	0.86	0.86	0.77
1 = Inmigrante (1ra o 2da generación)	0.03	0.03	0.13	0.04	0.03	0.29
1 = Madre económicamente activa	0.65	0.66	0.91	0.65	0.68	0.35
1 = Padre económicamente activo	0.94	0.94	0.71	0.94	0.94	0.83
Status ocupacional más alto de los padres (<i>hisei</i>)	44.79	44.07	0.36	44.51	44.21	0.53
Nivel educativo más alto de los padres (<i>pared</i>)	12.94	12.85	0.04*	12.99	12.87	0.08
Riqueza familiar (<i>wealth</i>)	-0.74	-0.79	0.93	-0.77	-0.79	0.71
1 = Tiene computadora en el hogar	0.80	0.79	0.59	0.80	0.80	0.88
1 = Tiene más de 100 libros en el hogar	0.17	0.17	0.65	0.17	0.17	0.77
1 = Asiste a una escuela privada	0.46	0.44	0.65	0.45	0.46	0.99
Clima educativo de la escuela (promedio <i>pared</i>)	12.75	12.77	0.26	12.75	12.77	0.26

(c) Integrados en América Latina. Media antes y después del corte y test de diferencia de medias.

Variable	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 1 mes		
	Media antes del corte	Media después del corte	P-valor	Media antes del corte	Media después del corte	P-valor
1 = Mujer	0.51	0.52	0.89	0.51	0.50	0.31
1 = No asistió a educación preescolar	0.11	0.07	0.15	0.10	0.07	0.14
1 = Asistió un año a educación preescolar	0.37	0.30	0.37	0.35	0.31	0.72
1 = Asistió más de un año a educación preescolar	0.52	0.63	0.12	0.55	0.62	0.26
1 = No repitente	0.59	0.68	0.00*	0.57	0.68	0.00*
1 = Vive en familia nuclear (con ambos padres)	0.71	0.72	0.60	0.71	0.73	0.78
1 = Vive con algún hermano/a	0.83	0.86	0.32	0.84	0.88	0.10
1 = Inmigrante (1ra o 2da generación)	0.03	0.04	1.00	0.04	0.04	0.76
1 = Madre económicamente activa	0.64	0.66	0.73	0.66	0.63	0.17
1 = Padre económicamente activo	0.93	0.93	0.59	0.92	0.94	0.33
Status ocupacional más alto de los padres (<i>hisei</i>)	43.63	44.40	0.45	43.65	43.33	0.71
Nivel educativo más alto de los padres (<i>pared</i>)	12.13	12.32	0.50	12.18	12.17	0.58
Riqueza familiar (<i>wealth</i>)	-0.93	-0.93	0.47	-0.94	-0.94	0.50
1 = Tiene computadora en el hogar	0.76	0.76	0.90	0.77	0.77	0.81
1 = Tiene más de 100 libros en el hogar	0.14	0.15	0.30	0.14	0.14	0.79
1 = Asiste a una escuela privada	0.33	0.33	0.20	0.32	0.33	0.49
Clima educativo de la escuela (promedio <i>pared</i>)	12.21	12.42	0.06	12.21	12.39	0.28

Fuente: Elaboración propia en base a microdatos de PISA.

Notas: (1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) * p-valor < 0.05. (3) Los p-valores de cada fila corresponden al coeficiente de una regresión de la variable en cuestión en la variable *dummy* que identifica el nacimiento anterior a la fecha de corte para las observaciones en el ancho de banda definido en el encabezado de la columna, empleando efectos fijos por país y año. (4) Estudiantes segregados (integrados): distancia percentilica absoluta al mediano de la escuela inferior (superior) a la mediana para América Latina (computada en base al índice *escs* de PISA). (5) Errores estándar estimados por BRR. (6) Brasil = Amazonas, Distrito Federal y Roraima.

Tabla 3. Efecto promedio del tratamiento (asistir al año superior) sobre el puntaje PISA en matemática. Resultados de regresión discontinua difusa para América Latina.

Variable dependiente: Puntaje PISA en matemática	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 1 mes		
	(1)	(2)	(3)	(4)	(5)	(6)
Tratamiento (1= Asiste al año escolar superior)	34.2*** (5.2)	34.3*** (5.1)	30.5*** (5.1)	36.4*** (6.9)	36.8*** (6.8)	31.8*** (6.8)
Controles:						
Efecto fijo por país	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI
Número de observaciones:	25,979	25,978	25,978	14,127	14,126	14,126

Fuente: Elaboración propia en base a microdatos de PISA.

Notas: (1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) * p-valor <0.10 ** p-valor <0.05 *** p-valor <0.01. (3) Los controles a nivel estudiante incluyen género, asistencia a un año de educación preescolar, asistencia a más de un año de educación preescolar, un indicador de repitencia y el nivel socioeconómico del estudiante (índice *wealth*). A nivel de escuela se incluye como control un indicador de si la escuela a la que el estudiante asiste tiene un clima educativo alto (promedio de máxima educación de los padres de la escuela mayor a 12 años). Se evita el uso de más controles para evitar la pérdida de observaciones por datos faltantes, especialmente en las variables a nivel escuela. (4) Errores estándar computados por BRR. (5) Brasil = Amazonas, Distrito Federal y Roraima.

Tabla 4. Efecto promedio del tratamiento (asistir al año superior) sobre el puntaje PISA en matemática. Resultados de regresión discontinua difusa para América Latina por nivel de integración socioeconómica del estudiante.

(a) Alumnos integrados

Variable dependiente: Puntaje PISA en matemática	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 1 mes		
	(1)	(2)	(3)	(4)	(5)	(6)
Tratamiento (1= Asiste al año escolar superior)	39*** (6.4)	40.3*** (6.3)	35.6*** (6.2)	44.3*** (9.1)	46.7*** (8.8)	40.4*** (8.8)
Controles:						
Efecto fijo por país	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI
Número de observaciones:	12,796	12,796	12,796	6,985	6,985	6,985

(b) Alumnos segregados

Variable dependiente: Puntaje PISA en matemática	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 1 mes		
	(1)	(2)	(3)	(4)	(5)	(6)
Tratamiento (1= Asiste al año escolar superior)	28.1*** (6.9)	26.6*** (6.7)	23.9*** (6.9)	26.8*** (9.1)	24.8*** (8.9)	20.9** (9.0)
Controles:						
Efecto fijo por país	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI
Número de observaciones:	13,113	13,112	13,112	7,100	7,099	7,099

Fuente: Elaboración propia en base a microdatos de PISA.

Notas: (1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) * p-valor <0.10 ** p-valor <0.05 *** p-valor <0.01. (3) Un estudiante se considera relativamente integrado si la distancia percentilica absoluta al estudiante mediano de su escuela es mayor a la distancia absoluta mediana para América Latina (14) y relativamente segregado si la distancia es menor a este valor. (4) Los controles a nivel estudiante incluyen género, asistencia a un año de educación preescolar, asistencia a más de un año de educación preescolar, un indicador de repitencia y el nivel socioeconómico del estudiante (índice *wealth*). A nivel de escuela se incluye como control un indicador de si la escuela a la que el estudiante asiste tiene un clima educativo alto (promedio de máxima educación de los padres de la escuela mayor a 12 años). Se evita el uso de más controles para evitar la pérdida de observaciones por datos faltantes, especialmente en las variables a nivel escuela. (5) Errores estándar computados por BRR. (6) Brasil = Amazonas, Distrito Federal y Roraima.

Tabla 5. Robustez del efecto promedio del tratamiento (asistir al año superior) sobre el puntaje PISA en matemática para América Latina por nivel de integración socioeconómica del estudiante, con distintos indicadores de bienestar (ancho de banda: +/- 1 mes).

(a) Alumnos integrados

Variable dependiente: Puntaje PISA en matemática	Indicador de bienestar								
	Nivel socioeconómico (escs)			Riqueza (wealth)			Educación padres (pared)		
Tratamiento (1= Asiste al año escolar superior)	44.3*** (9.1)	46.7*** (8.8)	40.4*** (8.8)	40.8*** (10.7)	41*** (10.3)	35.9*** (10.1)	45.8*** (8.3)	47.1*** (8.1)	39.4*** (8.5)
Controles:									
Efecto fijo por país	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI	NO	NO	SI
Número de observaciones:	6,985	6,985	6,985	6,692	6,691	6,691	6,869	6,869	6,869

(b) Alumnos segregados

Variable dependiente: Puntaje PISA en matemática	Indicador de bienestar								
	Nivel socioeconómico (escs)			Riqueza (wealth)			Educación padres (pared)		
Tratamiento (1= Asiste al año escolar superior)	26.8*** (9.1)	24.8*** (8.9)	20.9** (9.0)	30.4*** (8.1)	31.1*** (8.1)	27.9*** (8.5)	28.7*** (10.2)	29.2*** (9.5)	26.7*** (9.6)
Controles:									
Efecto fijo por país	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI	NO	NO	SI
Número de observaciones:	7,100	7,099	7,099	7,433	7,433	7,433	7,099	7,099	7,099

Fuente: Elaboración propia en base a microdatos de PISA.

Notas: (1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) * p-valor <0.10 ** p-valor <0.05 *** p-valor <0.01. (3) Un estudiante se considera relativamente integrado si, utilizando la variable de bienestar indicada en el encabezado de la columna, la distancia percentilica absoluta al estudiante mediano de su escuela es mayor a la distancia absoluta mediana para América Latina y relativamente segregado si la distancia es menor a este valor. (4) Los controles a nivel estudiante incluyen género, asistencia a un año de educación preescolar, asistencia a más de un año de educación preescolar, un indicador de repitencia y el nivel socioeconómico del estudiante (índice *wealth*). A nivel de escuela se incluye como control un indicador de si la escuela a la que el estudiante asiste tiene un clima educativo alto (promedio de máxima educación de los padres de la escuela mayor a 12 años). Se evita el uso de más controles para evitar la pérdida de observaciones por datos faltantes, especialmente en las variables a nivel escuela. (5) Errores estándar computados por BRR. (6) Brasil = Amazonas, Distrito Federal y Roraima.

Tabla 6. Robustez del efecto promedio del tratamiento (asistir al año superior) sobre el puntaje PISA en matemática para América Latina por nivel de integración socioeconómica del estudiante, con definición alternativa de integración (ancho de banda: +/- 1 mes).

(a) Alumnos integrados

Variable dependiente: Puntaje PISA en matemática	Medida de integración					
	Distancia percentílica			Desigualdad educación padres		
Tratamiento (1= Asiste al año escolar superior)	44.3*** (9.1)	46.7*** (8.8)	40.4*** (8.8)	50.5*** (12.3)	51.6*** (12.1)	42.5*** (11.5)
Controles:						
Efecto fijo por país	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI
Número de observaciones:	6,985	6,985	6,985	7,031	7,031	7,031

(b) Alumnos segregados

Variable dependiente: Puntaje PISA en matemática	Medida de integración					
	Distancia percentílica			Desigualdad educación padres		
Tratamiento (1= Asiste al año escolar superior)	26.8*** (9.1)	24.8*** (8.9)	20.9** (9.0)	26.4*** (7.6)	26.2*** (7.6)	22.4*** (7.8)
Controles:						
Efecto fijo por país	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI
Número de observaciones:	7,100	7,099	7,099	7,095	7,095	7,095

Fuente: Elaboración propia en base a microdatos de PISA.

Notas: (1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) * p-valor <0.10 ** p-valor <0.05 *** p-valor <0.01. (3) Se define a un alumno como integrado (segregado) si la distancia percentílica absoluta al mediano o el gini de la máxima educación de los padres de la escuela a la que asiste es mayor (menor) a la mediana correspondiente para América Latina. (4) Los controles a nivel estudiante incluyen género, asistencia a un año de educación preescolar, asistencia a más de un año de educación preescolar, un indicador de repitencia y el nivel socioeconómico del estudiante (índice *wealth*). A nivel de escuela se incluye como control un indicador de si la escuela a la que el estudiante asiste tiene un clima educativo alto (promedio de máxima educación de los padres de la escuela mayor a 12 años). Se evita el uso de más controles para evitar la pérdida de observaciones por datos faltantes, especialmente en las variables a nivel escuela. (5) Errores estándar computados por BRR. (6) Brasil = Amazonas, Distrito Federal y Roraima.

Tabla 7. Efecto promedio del tratamiento (asistir al año superior) sobre el puntaje PISA en matemática. Resultados de regresión discontinua difusa para América Latina por tipo de integración socioeconómica del estudiante.

(a) Alumnos integrados con pares más ricos

Variable dependiente: Puntaje PISA en matemática	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 1 mes		
	(1)	(2)	(3)	(4)	(5)	(6)
Tratamiento (1= Asiste al año escolar superior)	44.6*** (10.4)	44.4*** (10.2)	40.5*** (10.0)	51.5*** (15.4)	50.7*** (15.0)	42.9*** (14.9)
Controles:						
Efecto fijo por país	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI
Número de observaciones:	6,404	6,404	6,404	3,534	3,534	3,534

(b) Alumnos integrados con pares más pobres

Variable dependiente: Puntaje PISA en matemática	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 1 mes		
	(1)	(2)	(3)	(4)	(5)	(6)
Tratamiento (1= Asiste al año escolar superior)	34*** (8.5)	36.3*** (8.5)	31*** (8.5)	39.4*** (11.4)	43.5*** (11.3)	38.8*** (11.4)
Controles:						
Efecto fijo por país	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI
Número de observaciones:	6,392	6,392	6,392	3,451	3,451	3,451

Fuente: Elaboración propia en base a microdatos de PISA.

Notas: (1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) * p-valor <0.10 ** p-valor <0.05 *** p-valor <0.01. (3) Un estudiante se considera integrado con pares más ricos si la distancia percentilica al estudiante mediano de su escuela es negativa y mayor (en valor absoluto) a la distancia mediana para América Latina (14). Un alumno se considera integrado con pares más pobres si la distancia percentilica al estudiante mediano de su escuela es positiva y mayor (en valor absoluto) a la distancia mediana para América Latina (14). (4) Los controles a nivel estudiante incluyen género, asistencia a un año de educación preescolar, asistencia a más de un año de educación preescolar, un indicador de repitencia y el nivel socioeconómico del estudiante (índice *wealth*). A nivel de escuela se incluye como control un indicador de si la escuela a la que el estudiante asiste tiene un clima educativo alto (promedio de máxima educación de los padres de la escuela mayor a 12 años). Se evita el uso de más controles para evitar la pérdida de observaciones por datos faltantes, especialmente en las variables a nivel escuela. (5) Errores estándar computados por BRR. (6) Brasil = Amazonas, Distrito Federal y Roraima.

Tabla 8. Efecto promedio del tratamiento (asistir al año superior) sobre el puntaje PISA en matemática. Resultados de regresión discontinua difusa para América Latina por tipo de integración socioeconómica del estudiante para diferentes indicadores de bienestar (ancho de banda: +/- 1 mes).

(a) Alumnos integrados con pares más ricos

Variable dependiente: Puntaje PISA en matemática	Indicador de bienestar								
	Nivel socioeconómico (escs)			Riqueza (wealth)			Educación padres (pared)		
Tratamiento (1= Asiste al año escolar superior)	51.5*** (15.4)	50.7*** (15.0)	42.9*** (14.9)	38.3*** (13.0)	36.5*** (13.3)	32.9** (13.4)	55.8*** (12.2)	52.3*** (11.9)	44.4*** (11.9)
Controles:									
Efecto fijo por país	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI	NO	NO	SI
Número de observaciones:	3,534	3,534	3,534	3,440	3,439	3,439	3,554	3,554	3,554

(b) Alumnos integrados con pares más pobres

Variable dependiente: Puntaje PISA en matemática	Indicador de bienestar								
	Nivel socioeconómico (escs)			Riqueza (wealth)			Educación padres (pared)		
Tratamiento (1= Asiste al año escolar superior)	39.4*** (11.4)	43.5*** (11.3)	38.8*** (11.4)	42.1*** (14.1)	43.9*** (14.0)	40.2*** (13.9)	39*** (12.3)	43*** (12.2)	35.8*** (12.6)
Controles:									
Efecto fijo por país	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI	NO	NO	SI
Número de observaciones:	3,451	3,451	3,451	3,252	3,252	3,252	3,315	3,315	3,315

Fuente: Elaboración propia en base a microdatos de PISA.

Notas: (1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) * p-valor <0.10 ** p-valor <0.05 *** p-valor <0.01. (3) Un estudiante se considera integrado con pares más ricos si, utilizando la variable de bienestar indicada en el encabezado de la columna, la distancia percentilica al estudiante mediano de su escuela es negativa y mayor (en valor absoluto) a la distancia mediana para América Latina. Un alumno se considera integrado con pares más pobres si, utilizando la variable de bienestar indicada en el encabezado de la columna, la distancia percentilica al estudiante mediano de su escuela es positiva y mayor (en valor absoluto) a la distancia mediana para América Latina. (4) Los controles a nivel estudiante incluyen género, asistencia a un año de educación preescolar, asistencia a más de un año de educación preescolar, un indicador de repitencia y el nivel socioeconómico del estudiante (índice *wealth*). A nivel de escuela se incluye como control un indicador de si la escuela a la que el estudiante asiste tiene un clima educativo alto (promedio de máxima educación de los padres de la escuela mayor a 12 años). Se evita el uso de más controles para evitar la pérdida de observaciones por datos faltantes, especialmente en las variables a nivel escuela. (5) Errores estándar computados por BRR. (6) Brasil = Amazonas, Distrito Federal y Roraima.

Tabla 9. Efecto promedio del tratamiento (asistir al año superior) sobre el puntaje PISA en matemática. Resultados de regresión discontinua difusa para América Latina por nivel de rendimiento e integración socioeconómica del estudiante.

(a) Alumnos de buen rendimiento integrados

Variable dependiente: Puntaje PISA en matemática	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 1 mes		
	(1)	(2)	(3)	(4)	(5)	(6)
Tratamiento (1= Asiste al año escolar superior)	6.6 (5.3)	8.1* (5.3)	8.1* (5.3)	6.8 (7.4)	9.4 (7.5)	9.2 (7.4)
Controles:						
Efecto fijo por país	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI
Número de observaciones:	3,534	3,534	3,534	1,972	1,972	1,972

(b) Alumnos de buen rendimiento segregados

Variable dependiente: Puntaje PISA en matemática	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 1 mes		
	(1)	(2)	(3)	(4)	(5)	(6)
Tratamiento (1= Asiste al año escolar superior)	7.3 (6.1)	8.3 (6.2)	8.6 (6.3)	4.6 (8.7)	5.0 (8.8)	5.7 (8.9)
Controles:						
Efecto fijo por país	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI
Número de observaciones:	5,105	5,105	5,105	2,840	2,840	2,840

(c) Alumnos de bajo rendimiento integrados

Variable dependiente: Puntaje PISA en matemática	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 2 meses		
	(1)	(2)	(3)	(1)	(2)	(3)
Tratamiento (1= Asiste al año escolar superior)	30.8*** (8.7)	31.6*** (8.6)	27.8*** (8.7)	37.1*** (12.1)	38.4*** (12.0)	33.4*** (11.9)
Controles:						
Efecto fijo por país	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI
Número de observaciones:	9,262	9,262	9,262	5,013	5,013	5,013

(d) Alumnos de bajo rendimiento segregados

Variable dependiente: Puntaje PISA en matemática	Ancho de banda: +/- 2 meses			Ancho de banda: +/- 2 meses		
	(1)	(2)	(3)	(1)	(2)	(3)
Tratamiento (1= Asiste al año escolar superior)	18.4** (8.6)	16.9** (8.5)	15.5* (8.5)	19.1* (11.6)	16.9* (11.4)	14.1 (11.2)
Controles:						
Efecto fijo por país	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI
Número de observaciones:	8,008	8,007	8,007	4,260	4,259	4,259

Fuente: Elaboración propia en base a microdatos de PISA.

Notas: (1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) * p-valor <0.10 ** p-valor <0.05 *** p-valor <0.01. (3) Se define a un alumno como integrado (segregado) si la distancia percentilica absoluta al mediano de la escuela a la que asiste es mayor (menor) a la mediana correspondiente para América Latina. (4) Un estudiante se considera de bajo rendimiento si alguno de sus valores plausibles en matemática se encuentra por debajo del puntaje de corte para pertenecer al nivel 2 de rendimiento en PISA, y de buen rendimiento si no cumple esa condición. Siendo que encontrarse por debajo del nivel 2 se asocia al analfabetismo funcional, el concepto de "buen" y "bajo" rendimiento es relativo al nivel que muestra América Latina. (5) Los controles a nivel estudiante incluyen género, asistencia a un año de educación preescolar, asistencia a más de un año de educación preescolar, un indicador de repitencia y el nivel socioeconómico del estudiante (índice *wealth*). A nivel de escuela se incluye como control un indicador de si la escuela a la que el estudiante asiste tiene un clima educativo alto (promedio de máxima educación de los padres de la escuela mayor a 12 años). Se evita el uso de más controles para evitar la pérdida de observaciones por datos faltantes, especialmente en las variables a nivel escuela. (6) Errores estándar computados por BRR. (7) Brasil = Amazonas, Distrito Federal y Roraima.

Tabla 10. Efecto promedio del tratamiento (asistir al año superior) sobre el puntaje PISA en matemática. Resultados de regresión discontinua difusa para América Latina por nivel de rendimiento e integración socioeconómica del estudiante, para diferentes indicadores de bienestar (ancho de banda: +/- 1 mes).

(a) Alumnos de buen rendimiento integrados

Variable dependiente: Puntaje PISA en matemática	Indicador de bienestar								
	Nivel socioeconómico (escs)			Riqueza (wealth)			Educación padres (pared)		
Tratamiento (1= Asiste al año escolar superior)	6.8 (7.4)	9.4 (7.5)	9.2 (7.4)	13.5 (9.9)	15.1* (10.0)	15.9* (10.1)	10.0 (7.7)	12.2* (7.8)	12.2* (7.8)
Controles:									
Efecto fijo por país	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI	NO	NO	SI
Número de observaciones:	1,972	1,972	1,972	2,016	2,016	2,016	1,899	1,899	1,899

(b) Alumnos de buen rendimiento segregados

Variable dependiente: Puntaje PISA en matemática	Indicador de bienestar								
	Nivel socioeconómico (escs)			Riqueza (wealth)			Educación padres (pared)		
Tratamiento (1= Asiste al año escolar superior)	4.6 (8.7)	5.0 (8.8)	5.7 (8.9)	-0.2 (7.4)	0.6 (7.4)	0.8 (7.5)	1.9 (8.7)	3.0 (8.8)	3.2 (8.8)
Controles:									
Efecto fijo por país	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI	NO	NO	SI
Número de observaciones:	2,840	2,840	2,840	2,799	2,799	2,799	2,891	2,891	2,891

(c) Alumnos de bajo rendimiento integrados

Variable dependiente: Puntaje PISA en matemática	Indicador de bienestar								
	Nivel socioeconómico (escs)			Riqueza (wealth)			Educación padres (pared)		
Tratamiento (1= Asiste al año escolar superior)	37.1*** (12.1)	38.4*** (12.0)	33.4*** (11.9)	32.9*** (12.2)	32*** (12.1)	28.7** (11.6)	40.4*** (9.8)	40.9*** (9.7)	37.2*** (10.1)
Controles:									
Efecto fijo por país	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI	NO	NO	SI
Número de observaciones:	5,013	5,013	5,013	4,676	4,675	4,675	4,970	4,970	4,970

(d) Alumnos de bajo rendimiento segregados

Variable dependiente: Puntaje PISA en matemática	Indicador de bienestar								
	Nivel socioeconómico (escs)			Riqueza (wealth)			Educación padres (pared)		
Tratamiento (1= Asiste al año escolar superior)	19.1* (11.6)	16.9* (11.4)	14.1 (11.2)	23.9** (10.6)	24.2** (10.5)	20.7* (11.1)	18.5* (12.6)	18.9* (12.2)	15.6 (11.8)
Controles:									
Efecto fijo por país	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de los estudiantes	SI	SI	SI	SI	SI	SI	SI	SI	SI
Características de las escuelas	NO	SI	SI	NO	SI	SI	NO	SI	SI
Efecto fijo por año	NO	NO	SI	NO	NO	SI	NO	NO	SI
Número de observaciones:	4,260	4,259	4,259	4,634	4,634	4,634	4,208	4,208	4,208

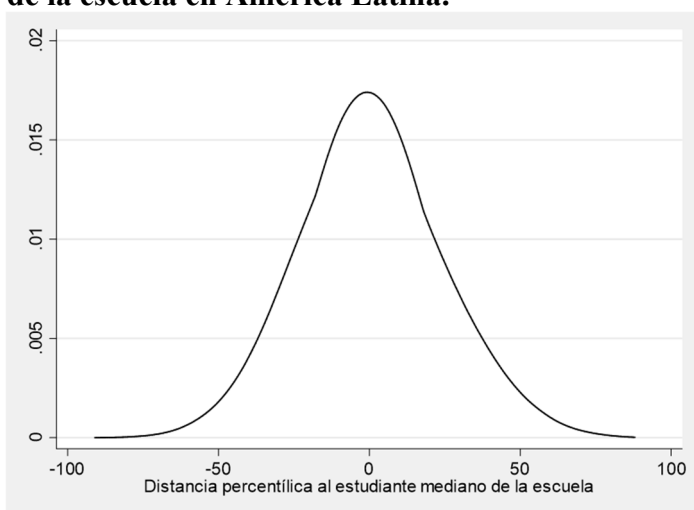
Fuente: Elaboración propia en base a microdatos de PISA.

Notas: (1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) * p-valor <0.10 ** p-valor <0.05 *** p-valor <0.01. (3) Se define a un alumno como integrado (segregado) si, utilizando la variable de bienestar indicada en el encabezado de la columna, la distancia percentilica absoluta al mediano de la escuela a la que asiste es mayor (menor) a la mediana correspondiente para América Latina. (4) Un estudiante se considera de bajo

rendimiento si alguno de sus valores plausibles en matemática se encuentra por debajo del puntaje de corte para pertenecer al nivel 2 de rendimiento en PISA, y de buen rendimiento si no cumple esa condición. Siendo que encontrarse por debajo del nivel 2 se asocia al analfabetismo funcional, el concepto de "buen" y "bajo" rendimiento es relativo al nivel que muestra América Latina. (5) Los controles a nivel estudiante incluyen género, asistencia a un año de educación preescolar, asistencia a más de un año de educación preescolar, un indicador de repitencia y el nivel socioeconómico del estudiante (índice *wealth*). A nivel de escuela se incluye como control un indicador de si la escuela a la que el estudiante asiste tiene un clima educativo alto (promedio de máxima educación de los padres de la escuela mayor a 12 años). Se evita el uso de más controles para evitar la pérdida de observaciones por datos faltantes, especialmente en las variables a nivel escuela. (6) Errores estándar computados por BRR. (7) Brasil = Amazonas, Distrito Federal y Roraima.

Gráficos

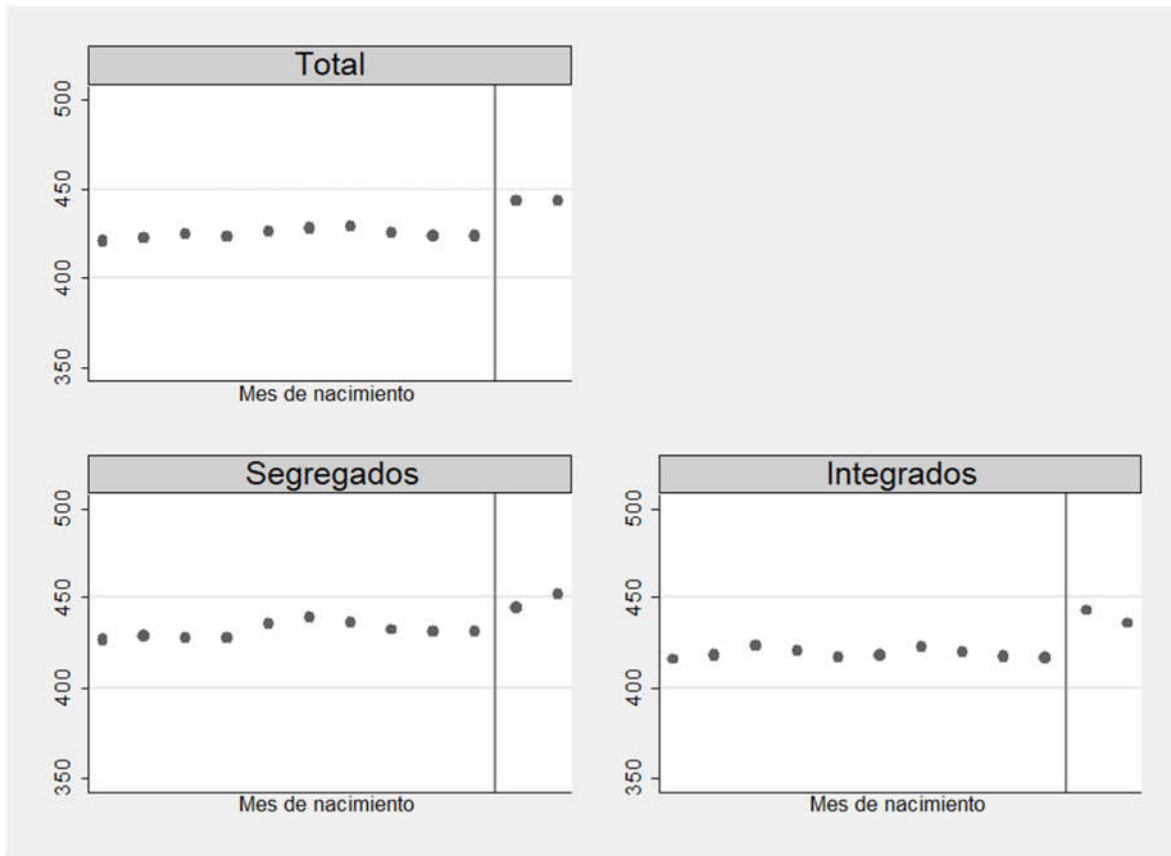
Gráfico 1. Histograma suavizado de las distancias percentílicas al estudiante mediano de la escuela en América Latina.



Fuente: Elaboración propia en base a microdatos de PISA.

Notas: (1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) La distancia percentílica es computada como la diferencia entre el número de percentil al que pertenece el estudiante en cada país-año de acuerdo al indicador de nivel socioeconómico (índice *escs*) y número de percentil al que pertenece el estudiante de nivel socioeconómico mediano en la escuela a la que asiste el estudiante. (3) Brasil = Amazonas, Distrito Federal y Roraima.

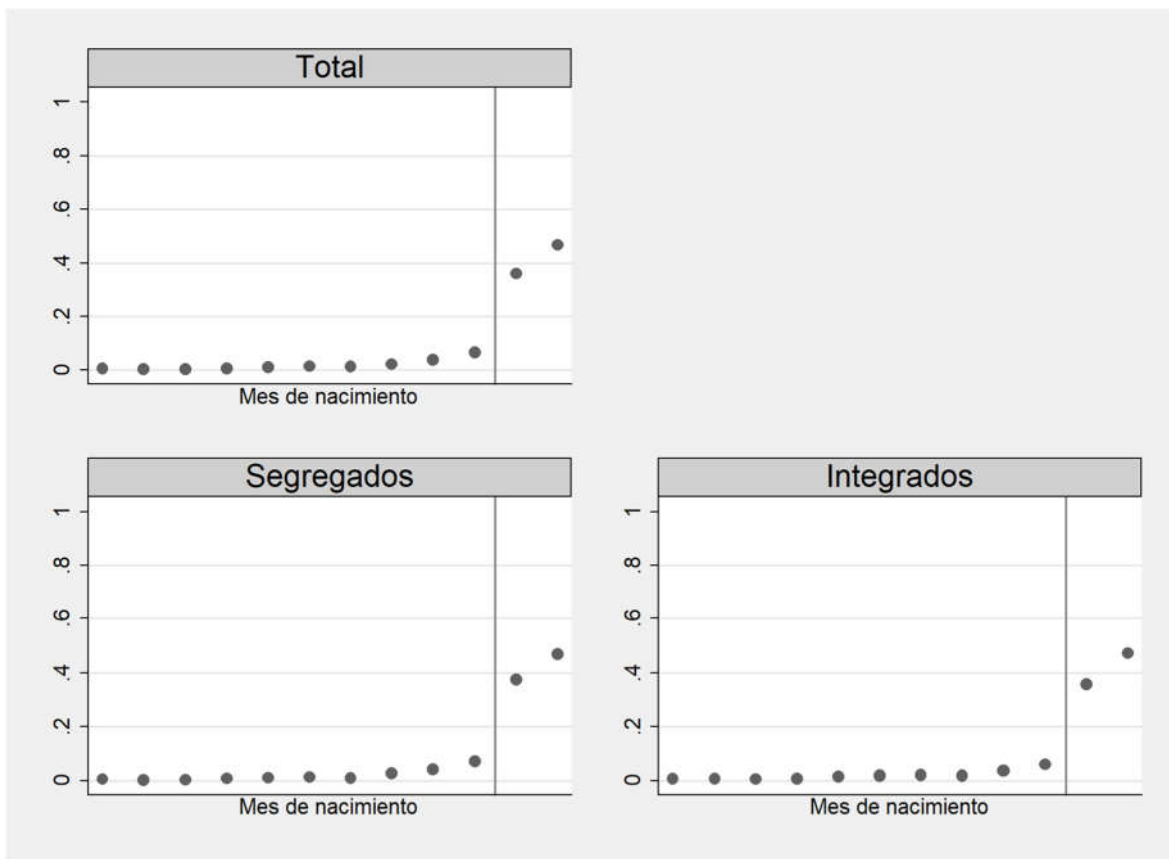
Gráfico 2. Puntaje PISA medio en matemática de los *compliers* por mes de nacimiento para el total de América Latina, estudiantes segregados e integrados.



Fuente: Elaboración propia en base a microdatos de PISA.

Notas: (1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) Los *compliers* son aquellos alumnos que asisten al año escolar al que deberían asistir dada su fecha de nacimiento. (3) Cada punto en el gráfico representa el puntaje PISA promedio en matemática para los alumnos nacidos en un mes relativo a la fecha de corte, que se indica con una línea vertical. Puntos más cerca de esta línea representan un mes de nacimiento más cercano al corte, con los puntos a la derecha indicando un nacimiento anterior a la fecha de corte y los puntos a la izquierda señalando nacimientos en los meses posteriores a dicha fecha. (4) Estudiantes segregados (integrados): distancia percentilica absoluta al mediano de la escuela inferior (superior) a la mediana para América Latina (computada en base al índice *escs* de PISA). (5) Brasil = Amazonas, Distrito Federal y Roraima.

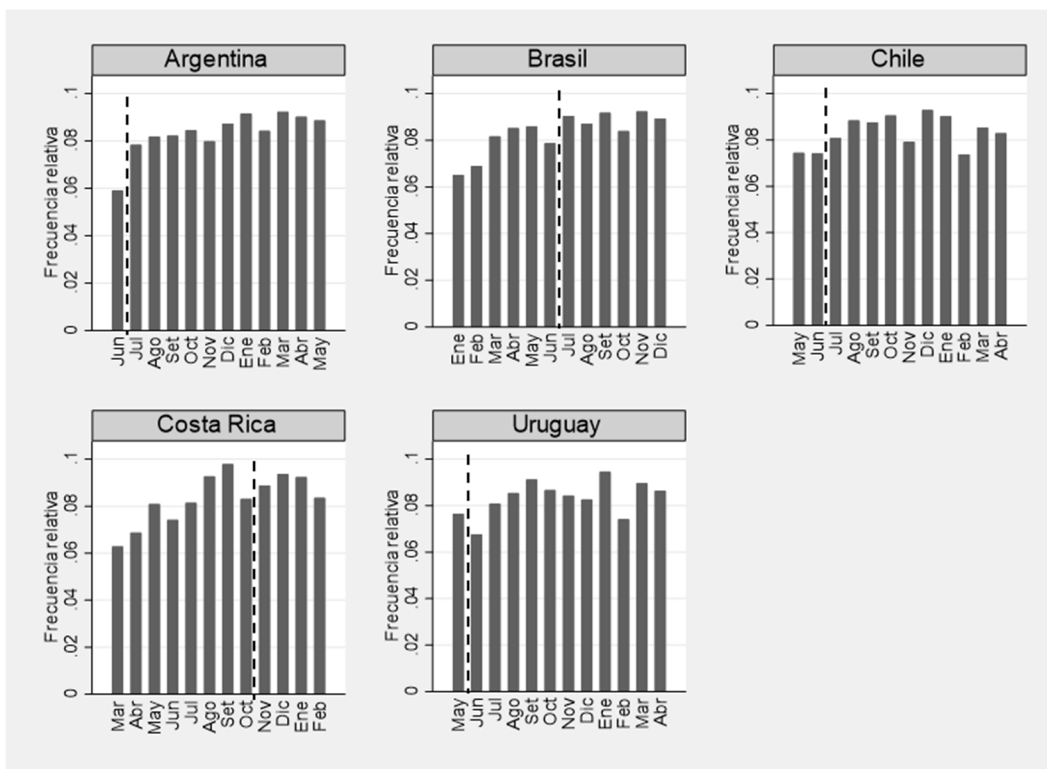
Gráfico 3. Probabilidad de tratamiento (asistir al año escolar superior) por mes de nacimiento para el total de América Latina, estudiantes segregados e integrados.



Fuente: Elaboración propia en base a microdatos de PISA.

Nota: 1) Se considera como América Latina al agregado de los 5 países analizados en este trabajo: Argentina, Brasil, Chile, Costa Rica y Uruguay. (2) El año escolar superior es el 11vo grado en todos los países, salvo en Costa Rica donde es el 10mo grado. (3) Cada punto en el gráfico representa la proporción de estudiantes en el grado superior para los nacidos en un mes relativo a la fecha de corte, que se indica con una línea vertical. Puntos más cerca de esta línea representan un mes de nacimiento más cercano al corte, con los puntos a la derecha indicando un nacimiento anterior a la fecha de corte y los puntos a la izquierda señalando nacimientos en los meses posteriores a dicha fecha. (4) Estudiantes segregados (integrados): distancia percentílica absoluta al mediano de la escuela inferior (superior) a la mediana para América Latina (computada en base al índice *escs* de PISA). (5) Brasil = Amazonas, Distrito Federal y Roraima.

Gráfico 4. Histograma por país y mes de nacimiento.



Fuente: Elaboración propia en base a microdatos de PISA 2018, excepto Argentina y Brasil donde se utilizan datos de 2015.

Notas: (1) La línea punteada señala la fecha de corte para el ingreso a la escuela de acuerdo a las leyes y regulaciones vigentes en el país. (2) Brasil = Amazonas, Distrito Federal y Roraima.