

Anexo 1

Caracterización de brechas educativas en Colombia

Este documento presenta una medición del impacto de la pandemia sobre los resultados en pruebas Saber 11 por medio de una técnica novedosa de cuantificación de los efectos. La clausura de colegios ocurrió en marzo del 2020, pero debido a los dos calendarios académicos existentes en Colombia, hubo un diferencial en tiempo de exposición al cierre de colegios. Para colegios de calendario B, hubo una pérdida de exposición directa a aprendizaje de dos meses. En contraste, los colegios de calendario A—en su mayoría, públicos—fueron fuertemente afectados por la pandemia debido a que el año escolar va de febrero a noviembre; en efecto, una pérdida de instrucción en el colegio de 8 meses. El ejercicio propuesto para medir el efecto de la pandemia en aprendizaje es la comparación en el tiempo entre los resultados de las cohortes de calendario A versus calendario B. La intuición del ejercicio de comparar los resultados entre calendario A y B es que, en ausencia de la pandemia, la diferencia de los resultados entre estas dos cohortes se hubiera mantenido en el tiempo.¹

Contrario a esta intuición, el ejercicio simple con los datos de las pruebas Saber 11 entre 2019 y el 2020 muestra un pequeño aumento en el resultado de los estudiantes de colegios de calendario A y una pérdida en los resultados de calendario B. Esta comparación simple enmascara dos efectos importantes. En primer lugar, precisamente debido a la pandemia, un porcentaje importante de estudiantes de ambos calendarios no presentaron la prueba (Lichand and Christen, 2021 documenta efectos sobre deserción en Brasil). Segundo, aún más importante, la composición de estudiantes también cambió; por consiguiente, el estimador simple no representa el verdadero efecto de la pandemia.

El documento presenta inicialmente una visión global de brechas en las pruebas por sector de la sede, zona, género y nivel socioeconómico de los estudiantes. A partir de estos datos, se documenta una serie de brechas importantes en educación.

El documento muestra tres resultados importantes. Primero, en términos de pérdida de estudiantes en la prueba, existe evidencia que la caída es mayor, en términos porcentuales, en colegios de Calendario B. Estos estudiantes presentaron la prueba en octubre, una vez graduados de colegio y varias universidades anunciaron a mediados de año la suspensión del requisito de saber 11 para admisión. Potencialmente estos cambios desincentivaron la presentación de la prueba en estudiantes de calendario B o el esfuerzo. Por otro lado, también se observa una caída (menor) porcentual de estudiantes de Calendario A. Potencialmente fueron estudiantes que se retiraron del sistema antes de graduarse (desertores).

Segundo, en términos de composición de estudiantes perdidos, la hipótesis que manejamos en este documento es que estudiantes de calendario B que hubieran tenido un puntaje alto en la prueba fueron los que dejaron de presentarla, truncando la distribución por la parte de arriba de la prueba. Es decir, en promedio, los resultados de calendario B se disminuyeron debido a que los estudiantes con potencial de mayores puntajes no presentaron la prueba. En contraste, los estudiantes que dejaron de presentar la prueba en calendario A fueron desertores, potencialmente los que hubieran tenido menores resultados en la prueba. Es decir, en promedio, los resultados de calendario A aumentaron debido a que los estudiantes que no presentaron la prueba fueron los que hubieran obtenido menores resultados. Ambos efectos inducen aumentos en la comparación simple entre calendario A y B. El documento reporta evidencia en línea con las dos hipótesis.

Tercero, debido a la pérdida de estudiantes y cambio en la composición de ellos, es necesario realizar una estimación de la pérdida que tome en cuenta estos dos factores. Haciendo uso de una técnica que “corrige” por ambos efectos (Angrist et al, 2006), encontramos que la pandemia disminuyó el desempeño de los estudiantes de colegios públicos en el rango de 4 a 7 puntos—un efecto entre 0.081 y 0.14 desviaciones estándares (DE)—con respecto a los estudiantes de calendario B.

1 En términos técnicos, este estimador es de Diferencias-en-Diferencias (DD). Los supuestos que hacen de DD un estimador insesgado es que (i) en ausencia de “tratamiento” (en este caso, la pandemia), la diferencia en el tiempo se hubiera mantenido constante; (ii) que las características no observables que afectan el resultado de la prueba son fijas en el tiempo.

Datos: Brechas en aprendizaje

Esta sección presenta los resultados en las pruebas Saber 11 desde el 2015, desagregados por sector, calendario escolar, zona de la sede educativa y género de los estudiantes. El presente análisis utilizó la base de datos de estudiantes de grado 11 que presentaron las pruebas Saber 11 entre el 2015 y el 2021. Esta base de datos cuenta con información de 3.354.572 estudiantes. Para el análisis se eliminaron estudiantes de calendario B que presentaron la prueba en noviembre y estudiantes de calendario A que presentaron la prueba en marzo. Este ejercicio resultó en una muestra final de 3.303.305 estudiantes. Adicionalmente, se tuvieron en cuenta los datos de matrícula de estudiantes en grado once a nivel de sede y clasificación de los planteles. El cruce de estas bases de datos con la de Saber 11 redujo la muestra en aproximadamente 15%.

Es importante mencionar que el análisis de la prueba tiene retos importantes. En primer lugar, la prueba no sigue a estudiantes en el tiempo; cada año tiene una nueva cohorte de estudiantes, y las cohortes pueden cambiar por accidentes estadísticos. Segundo lugar, la prueba no está anclada en preguntas comunes entre años. Si bien es cierto que los técnicos del ICFES intentan mantener la misma dificultad cada año, tal como se presentan los datos no es posible garantizar esto. Adicionalmente, debido a la pandemia, el formato de la prueba cambió en el 2020, con un menor número de preguntas. Por todo ello, es importante tener cautela al ver los datos en el tiempo.

La tabla 1 muestra características importantes de las sedes educativas en las que por lo menos 1 estudiante presentó la prueba Saber 11 en el 2019 y en el 2020, según el calendario académico del plantel. El nivel socioeconómico (NSE) de la sede se calculó de acuerdo con el puntaje del Índice de Nivel Socioeconómico (INSE) de los estudiantes en el año actual y los dos años anteriores. La clasificación fue estimada de acuerdo con el promedio ponderado del índice de cada prueba que tiene en cuenta la frecuencia relativa acumulada y la varianza de puntajes.

2 A la fecha de publicación de este documento solo han sido publicados los resultados de estudiantes de calendario B que presentaron Saber 11 en 2021.

El Panel A indica la zona de la sede y muestra que colegios de calendario B y los privados de calendario A tienden a estar ubicados en zonas urbanas, mientras que colegios públicos tienen una concentración relativamente igual en ambas zonas.

El Panel B muestra la distribución de las sedes por NSE, que se construye con base en el INSE, y resume las condiciones socioeconómicas del estudiante por medio de tres dimensiones: 1) nivel educativo de los padres, 2) ocupación de los padres, y 3) consumo de bienes y acceso a servicios. Este índice tiene un rango de 0 a 100: estudiantes con puntajes entre 0 y 41 están clasificados en NSE 1 y, por lo general, son estudiantes sin acceso a internet, computador, horno microondas o lavadora y la educación de sus madres es primaria incompleta. Estudiantes con puntajes entre 80 y 100 son clasificados en NSE 4, se caracterizan por tener computador, internet, consola de videojuegos, y sus madres tienen carrera profesional (Icfes, abril 2019). El NSE también es reportado para las instituciones educativas teniendo en cuenta el INSE del período actual y de los dos últimos períodos. Se observa que, en su mayoría, los colegios de calendario B tienen estudiantes con niveles socioeconómicos altos, mientras que gran parte de los colegios públicos tienen estudiantes de niveles socioeconómicos muy bajos.

El Panel C muestra la clasificación de los planteles. Esta clasificación se calcula usando los resultados en las pruebas del 80% de los estudiantes con puntajes más altos, en los últimos 3 años. Para cada sede se calcula un promedio ponderado del índice de resultados de cada componente de la prueba. Establecimientos con clasificación A+ tienen un índice mayor a 0.77, mientras que establecimientos con índice menor o igual a 0.62 tienen clasificación D. Se observa que la gran mayoría de colegios A+ y A son colegios privados, mientras que los colegios públicos típicamente tienen clasificación C o D.

Debido a las diferencias observadas entre estudiantes de colegios de calendario A y calendario B, las gráficas de brechas que se presentan en el documento diferencian a los estudiantes por calendario y por sector (oficial/no oficial).

Tabla 1*Características descriptivas de sedes en pruebas Saber 11, por calendario y sector*

	2019			2020		
	A privado	A público	B	A privado	A público	B
Panel A: Zona						
Rural	181	3,605	55	179	3709	57
Urbano	3,170	3,901	319	3135	3916	303
Panel B: NSE						
1	48	1829	2	39	1374	1
2	697	4545	12	556	5017	11
3	2014	1127	151	2187	1227	134
4	587	5	209	529	7	214
Panel C: Clasificación de los planteles						
A+	871	151	226	836	125	211
A	613	598	46	629	501	61
B	513	1686	27	514	1428	20
C	253	1736	9	283	1807	12
D	122	1928	5	133	2196	2
Total	3,351	7,506	374	3,314	7,625	360

Nota. Elaboración propia con base en datos de Saber 11 del año 2019 y 2020.

Brechas en aprendizaje por sector/calendario

El desempeño de los estudiantes de calendario A presentó una tendencia decreciente entre el 2016 y el 2019, pues su puntaje disminuyó en 9 puntos en los colegios privados y en 12 puntos en los públicos. En el 2020, sin embargo, esta tendencia se reversó, pues los estudiantes de colegios privados aumentaron en 7 puntos (aproximadamente 0.12 desviaciones estándar, DE), alcanzando niveles parecidos al 2017, y los de colegios públicos aumentaron 1 punto (equivalente a 0.02 DE). Estos resultados demostraron que la brecha entre estudiantes de colegios privados y públicos en calendario A aumentó para el 2020.

Por otro lado, con puntajes entre 314 y 325 en el periodo 2015-2019, los estudiantes de calendario B muestran un desempeño muy superior a los de calendario A. En el 2020, los puntajes caen en 7 puntos (0.14 DE), pero en el 2021 aumentan en 5, alcanzando un nivel de 312 (figura 1).

Puesto que las pruebas Saber 11 no tienen un carácter obligatorio y, en especial, el 2020 presentó desafíos para administrarlas y presentarlas, cambios en los puntajes promedio pueden estar relacionados con cambios

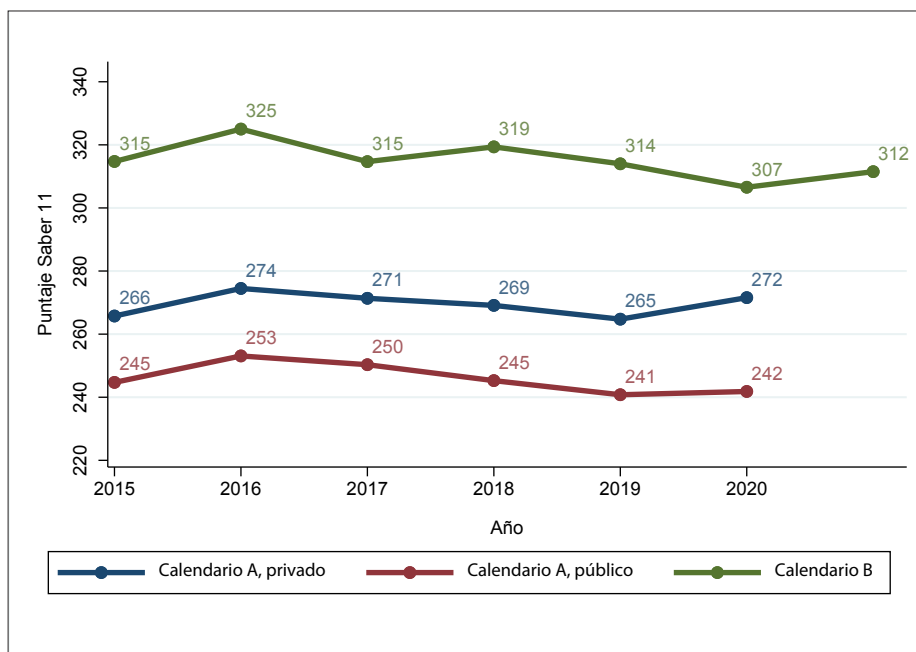
en la composición demográfica de los estudiantes a través de los años.

Zona de la sede

Estudiantes de sedes educativas privadas de calendarios A y B en zonas rurales obtienen puntajes mayores que sus pares en zonas urbanas. Este resultado puede explicarse por la ubicación de colegios privados con estudiantes de altos recursos³. En el 2020, los estudiantes de estos colegios presentaron un desempeño inferior; el puntaje promedio en calendario A bajó en 1.5 puntos (289 en 2019 y 287.6 en 2020) y en calendario B en 16 puntos (337.8 en 2019 y 321 en 2020). Estos cambios representan una disminución de 0.41 DE para el calendario B y de 0.03 DE para el A. Por otro lado, estudiantes de colegios privados de calendario A ubicados en zonas urbanas mejoraron en las pruebas, con un aumento de 263.8 a 270.8, es decir 0.13 DE. En colegios públicos, se observa que los estudiantes de sedes en zonas urbanas tienden a tener mejores puntajes que los estudiantes de sedes en zonas rurales. En el 2020, no hubo diferencias en aumento de puntajes entre sedes en zonas rurales y zonas urbanas (figura 2).

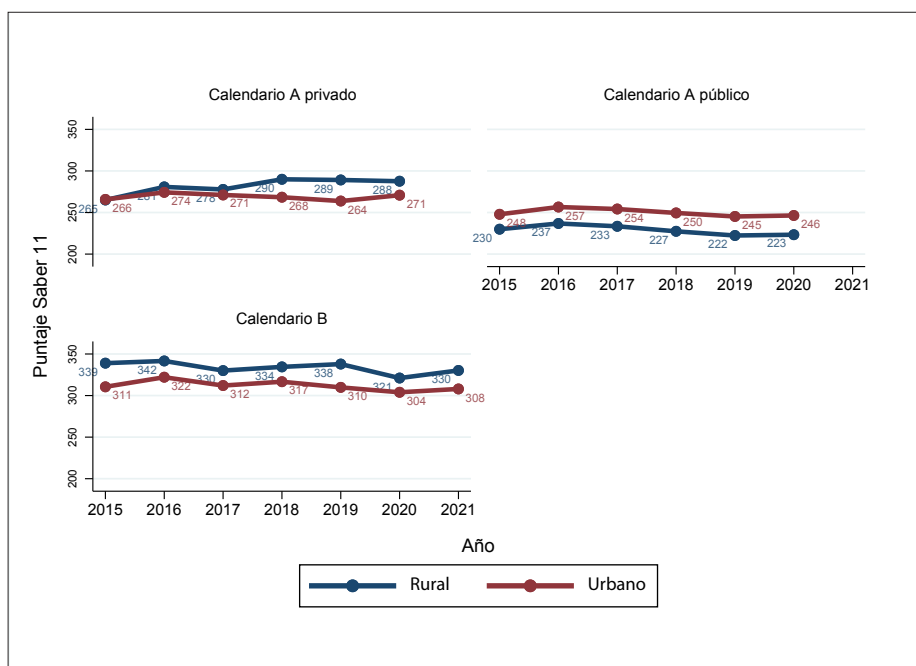
3 Este es el caso de la Sabana de Bogotá es diciente, donde se pueden encontrar colegios privados con altos puntajes en Cajicá, Chía y Cota.

Figura 1
Resultados pruebas Saber 11 por calendario y sector



Nota. Elaboración propia con base en datos de pruebas Saber 11 del año 2015 al 2021. Número de estudiantes = 3,303,305.

Figura 2
Resultados pruebas Saber 11 por zona



Nota. Elaboración propia con base en datos de pruebas Saber 11 del 2015 al 2021. Porcentaje estudiantes en zonal rural: calendario A privado = 3.9%, calendario B = 15.10%, calendario A público = 18.76%.

Género

En general, los hombres tienen mayores puntajes en las pruebas Saber 11 con respecto a las mujeres. Estas brechas, sin embargo, se cerraron levemente en el 2020, cuando las mujeres de establecimientos privados de calendario A aumentaron su desempeño en 7 puntos (263 a 270) y las de instituciones públicas en 2 puntos (237 a 239): Al mismo tiempo, los hombres mejoraron su desempeño en 6.3 puntos en los establecimientos privados y no hubo cambios en los públicos. En calendario B, el desempeño en las pruebas tanto de los hombres como de mujeres continúa con una tendencia decreciente en el 2020. Los puntajes de los hombres, sin embargo, caen más: 8.6 puntos entre hombres comparado con 6.3 puntos entre las mujeres. Figura 3.

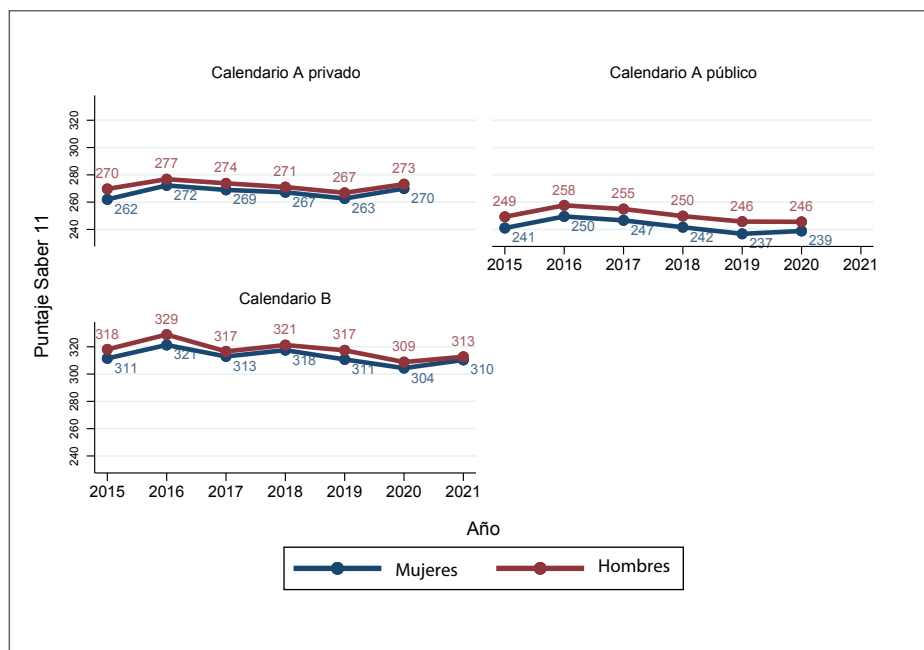
Nivel socioeconómico (NSE) de los estudiantes

Como se mencionó anteriormente, las tendencias de los puntajes según el nivel socioeconómico de

los estudiantes se hicieron con base en el puntaje de INSE. En calendario A, estudiantes de colegios clasificados por condiciones socioeconómicas en niveles 1 y 2 tendieron a mejorar su desempeño en el 2020. Se observa un aumento mucho mayor entre estudiantes de colegios privados, para los cuales los puntajes de las pruebas aumentaron en 8 puntos, cifra con la cual llegaron a niveles similares de 2017. Este aumento es sorprendente, pues se esperaría que el desempeño de estudiantes de hogares con condiciones económicas bajas disminuyera, debido a la gran afectación por la pandemia.

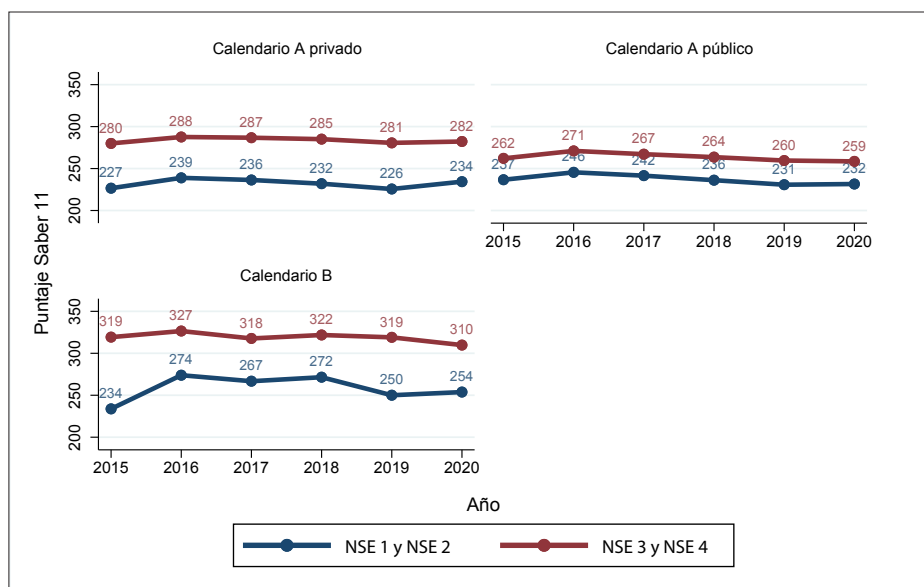
En calendario B se observa lo contrario: estudiantes de colegios con NSE 3 y 4 presentaron una disminución notable (9 puntos) en su desempeño en las pruebas, mientras que los de NSE 1 y 2 mejoraron su desempeño. Es importante tener en cuenta que el número de estudiantes en este rango es muy bajo (4.6% en todos los años), de tal forma que los resultados son sensibles a cambios entre años (figura 4).

Figura 3
Resultados pruebas Saber 11 por género



Nota. Elaboración propia con base en datos de pruebas Saber 11 del año 2015 al 2021.

Figura 4
Resultados pruebas Saber 11 por nivel socioeconómico



Nota. Elaboración propia con base en datos de pruebas Saber 11 del 2015 al 2020. El puntaje INSE para estudiantes que presentaron la prueba en el 2021 aún no está disponible.

Características de los estudiantes que presentaron las pruebas Saber 11

En esta sección se analiza el cambio en el número de estudiantes en el tiempo y su distribución de acuerdo con sus condiciones socioeconómicas. Este análisis es relevante debido a la alta correlación que existe entre NSE y resultados académicos. Si en el 2020, el número de estudiantes de familias de bajos recursos en sedes de calendario A que no presentó la prueba disminuyó desproporcionalmente, entonces aumentos relativos en la relación A/B pueden deberse a cambios en la composición de los estudiantes que tomaron la prueba.

Número de estudiantes

El número de estudiantes que presenta las pruebas Saber 11 ha declinado levemente desde el 2015, sin embargo, el declive fue más acentuado entre el 2019 y el 2020, cuando el porcentaje disminuyó en 6% para colegios públicos y en 12% para colegios privados (figura 5).

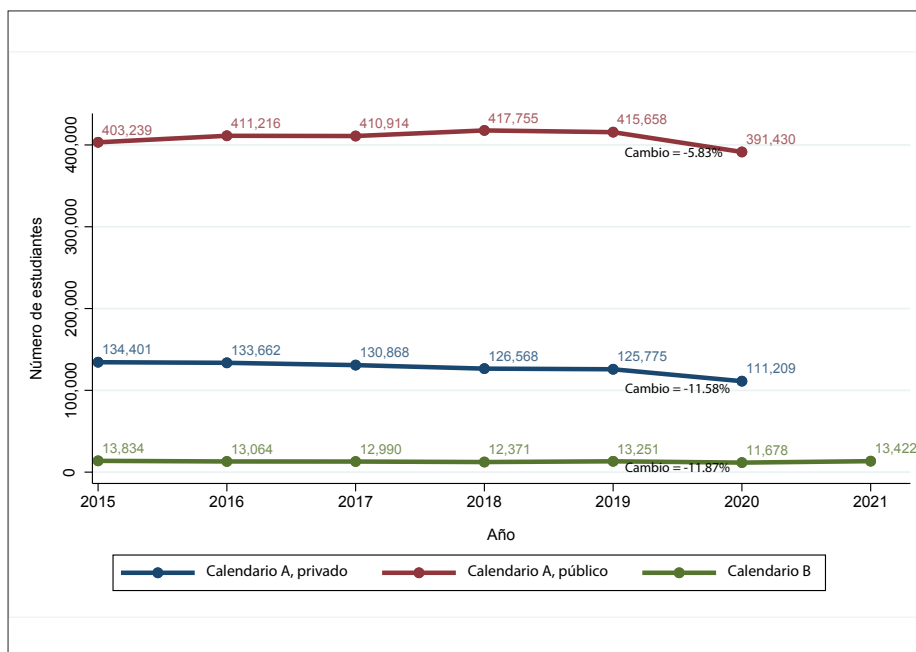
Porcentaje de estudiantes

Además de analizar el cambio en el número de estudiantes, se consideró la proporción de estudiantes de grado 11 por sede educativa que presentó las pruebas, la cual fue relativamente estable entre el 2015 y el 2019, fluctuando alrededor de 96%. En contraste, un porcentaje menor de estudiantes en promedio presentó las pruebas Saber 11 en el 2020. Esta disminución se dio principalmente entre estudiantes de calendario B, quienes ese año fueron citados para presentar el examen después de graduarse de bachilleres y los resultados no eran requisito de admisión para la universidad, lo cual posiblemente disminuyó los incentivos de presentarlo (figura 6).

A partir de estos datos, se estimó la correlación entre la pandemia y el porcentaje relativo de estudiantes por sede tomando la prueba, el cual incrementó en calendario A con respecto a calendario B en 4 puntos porcentuales entre 2019 y 2020. Es decir, la pérdida fue mayor, en términos relativos, para los colegios de calendario B, donde se esperaba una menor pérdida (menor efecto de pandemia) (figura 7).

Figura 5

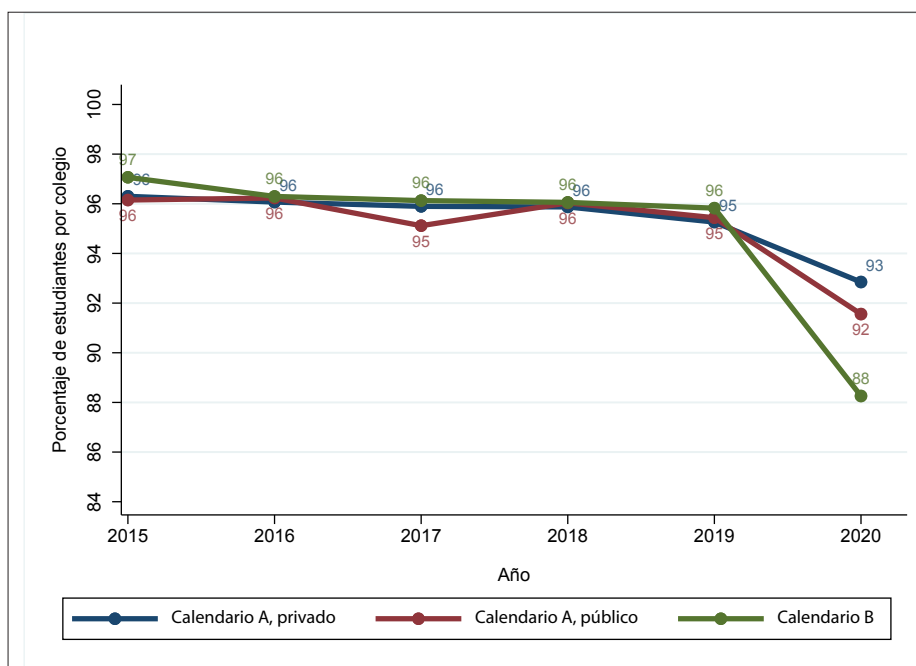
Número de estudiantes que presentaron las pruebas Saber 11



Nota. Elaboración propia con base en datos de pruebas Saber 11 del año 2015 al 2021.

Figura 6

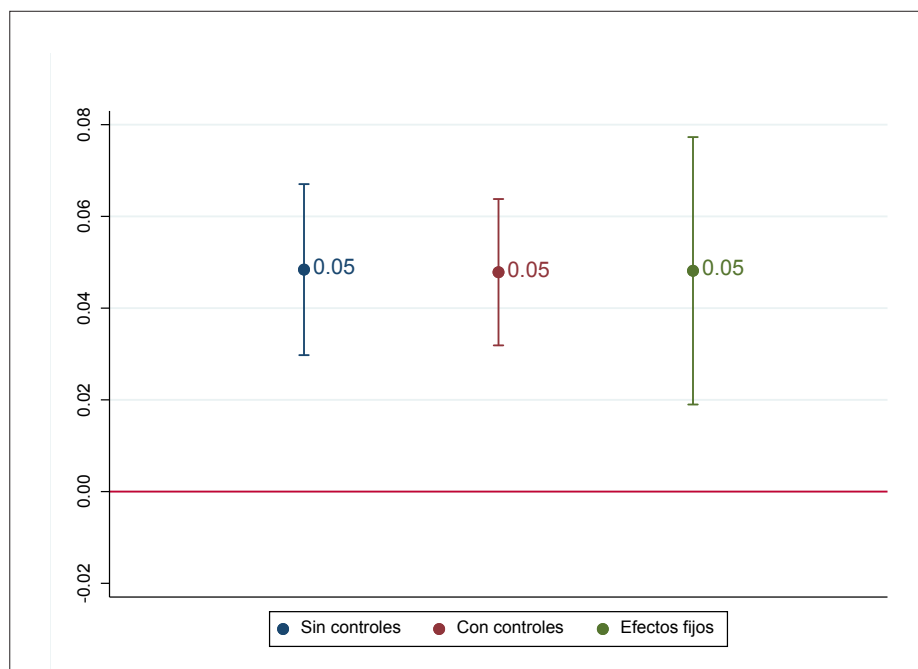
Porcentaje promedio de estudiantes que presentaron las pruebas Saber 11



Nota. Elaboración propia con base en datos de pruebas Saber 11 del 2015 al 2020 y datos de estudiantes matriculados en grado 11 del Ministerio de Educación Nacional, Datos Abiertos.

Figura 7

Cambio en el porcentaje de estudiantes que presentó Saber 11 entre el 2019 y el 2020 y entre colegios de calendario A y calendario B



Nota. Elaboración propia con base en estimación de DID del porcentaje de estudiantes por sede que presentó la prueba Saber 11⁴. Distribución de índice socioeconómico entre estudiantes que presentaron las pruebas.

La figura 8 muestra la densidad de estudiantes por índice socioeconómico, medido por el INSE. Esta gráfica sugiere que hubo un cambio en la composición de estudiantes que presentaron las pruebas por calendario entre el 2019 y el 2020. El panel de la izquierda revela un desplazamiento a la derecha de la densidad, es decir, en términos relativos, estudiantes de colegios públicos de calendario A con mayor índice socioeconómico presentaron la prueba en el 2020. La diferencia entre años es menos evidente en el caso de los colegios de calendario B (panel de la derecha); sin embargo, se puede ver un leve desplazamiento a la izquierda, lo que indica una pérdida de estudiantes de mayor índice socioeconómico en la presentación la prueba en el 2020. En resumen, estos resultados sugieren que hubo mayor representación de estudiantes de nivel socioeconómico alto en colegios públicos en calendario A, y menor representación de estos estudiantes en colegios de calendario B.

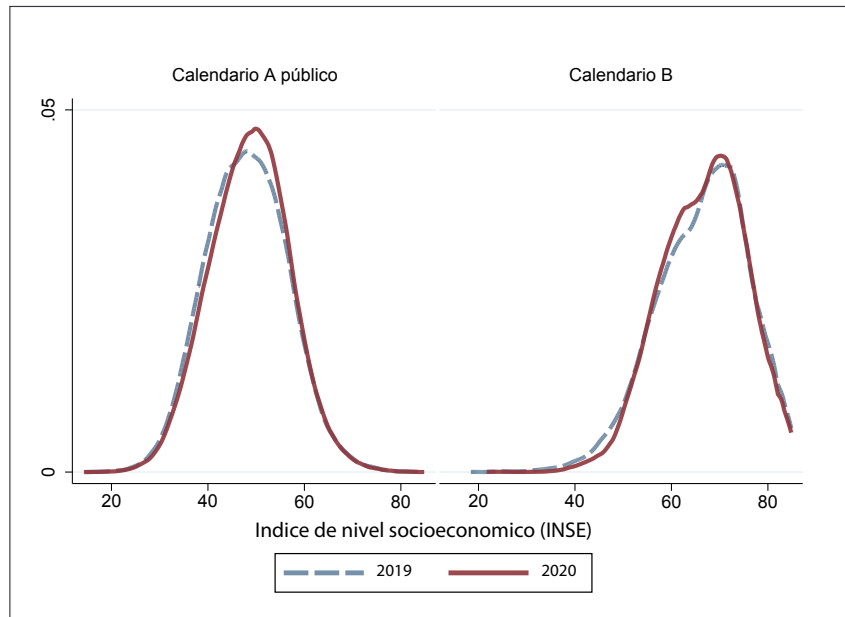
Clasificación de planteles

Además de diferencias en la distribución de estudiantes por nivel socioeconómico, se observa una diferencia según la clasificación de la sede educativa. El porcentaje promedio de estudiantes en sedes de calendario B clasificados como A+ y A disminuyó notablemente en el 2020. Por el contrario, en calendario A, la mayor disminución se dio entre sedes de clasificación C y D. Estos resultados indican que estudiantes de sedes de calendario A, quienes posiblemente hubieran podido tener un muy bajo desempeño en las pruebas Saber 11 (dada la calidad de la sede), no la presentaron en el 2020. Asimismo, estudiantes que se esperaba que tuvieran muy buenos resultados en sedes de calendario B, tampoco presentaron la prueba (figura 9).

4 La muestra incluye las sedes que están en todos los años, entre 2015 y 2020. La estimación con controles incluye dummies de año, sector (oficial/no oficial), zona (rural/urbana), promedio del puntaje INSE por sede, porcentaje de mujeres y de estudiantes pertenecientes a etnia que presentaron la prueba, y puntaje promedio en la prueba Saber 11. La estimación de efectos fijos incluye municipio. Errores estándares robustos.

Figura 8

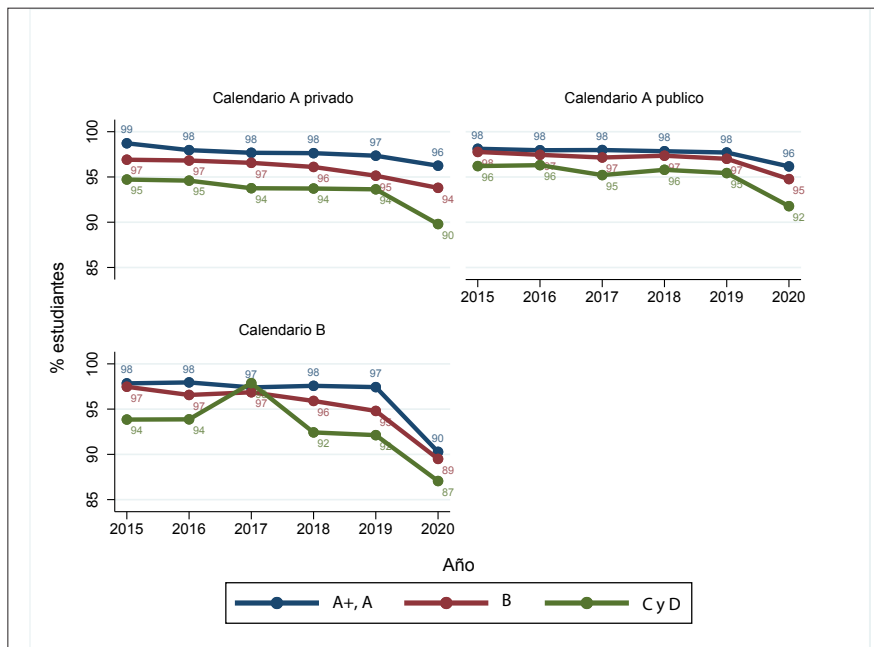
Cambio en la distribución de índice de nivel socioeconómico de estudiantes que presentaron Saber 11 en 2019 y 2020



Nota. Elaboración propia con base en datos de pruebas Saber 11 de 2019 y 2020.

Figura 9

Porcentaje promedio de estudiantes que presentaron Saber 11 según clasificación de la sede



Nota. Elaboración propia con base en datos de pruebas Saber 11 del 2015 al 2020, y datos de clasificación de planteles del Icfes.

Relación entre puntajes y nivel socioeconómico

Para explorar la relación entre el cambio en la composición de estudiantes por nivel socioeconómico y los puntajes entre el 2019 y el 2020 se compararon los resultados de ambos años teniendo en cuenta el número de estudiantes en distintos niveles INSE. Específicamente, se dividieron los estudiantes de ambos años en 20 grupos iguales de acuerdo al puntaje INSE, y se calculó el puntaje promedio en la prueba Saber 11 de cada grupo. Si hubo cambios en la composición de estudiantes, se habría esperado que el número de estos en cada grupo fuera igual para los dos años. Sin embargo, esto no ocurrió. El Panel A de los siguientes tres gráficos muestra un cambio importante en el número de estudiantes que presentó la prueba para diferentes niveles INSE, hallazgo que fue diferente para sedes de calendario A y B.

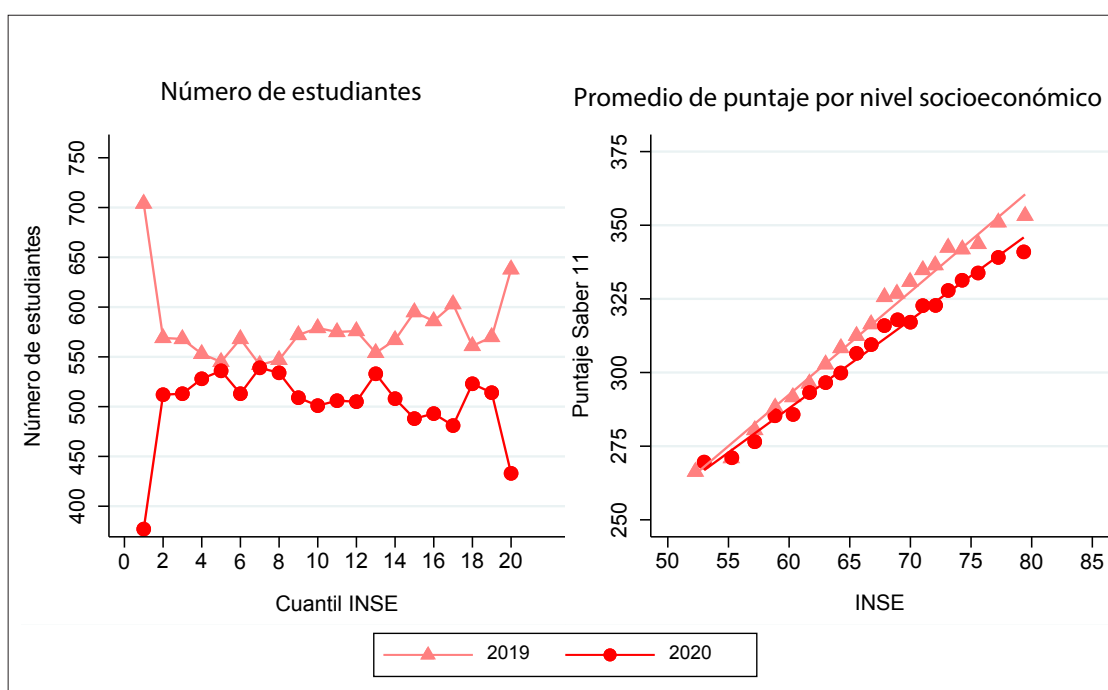
En comparación al 2019, en el 2020 hubo menos estudiantes con niveles altos de INSE en las sedes de

calendario B, estos estudiantes presentan menores puntajes en la prueba Saber 11 (figura 10a). El menor desempeño de estudiantes de familias con niveles INSE alto puede deberse a menor esfuerzo en las pruebas debido a que el puntaje no era requisito para ser admitidos a una universidad.

Entre estudiantes de sedes privadas de calendario A se observan puntajes muy similares, excepto para los de INSE bajo, quienes tienen puntajes un poco más altos, en promedio, pero a la vez un número menor de estudiantes presentó las pruebas en este grupo (figura 10b).

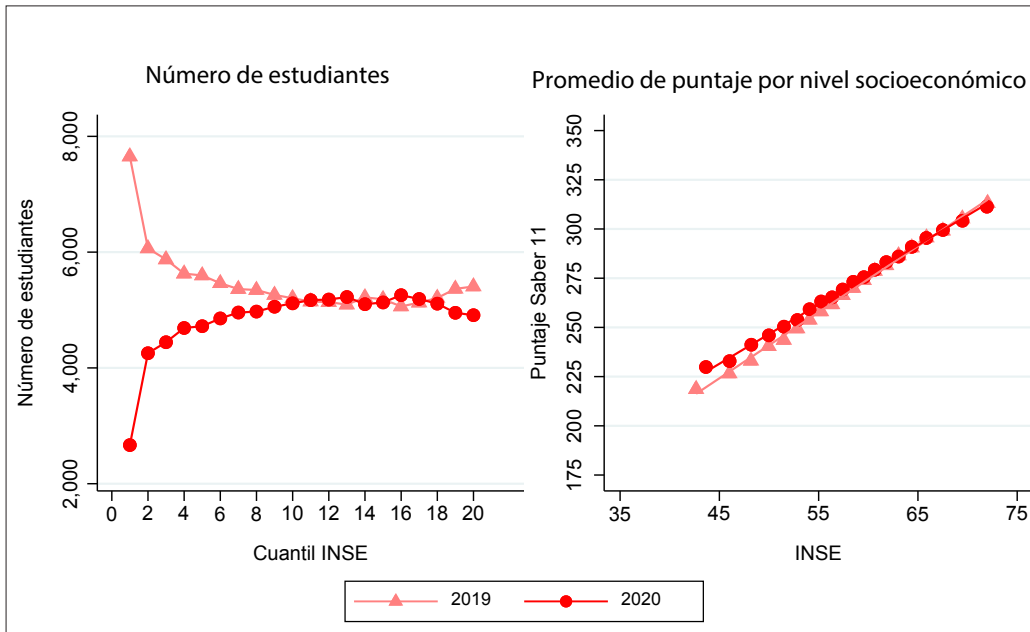
En sedes públicas también se observa un menor número de estudiantes de bajo INSE que presentó las pruebas, pero no se evidencian cambios en los puntajes. Es posible que aquellos estudiantes que no presentaron las pruebas fueran los más afectados por la pandemia y si hubiesen presentado el examen sus puntajes bajarían el promedio (figura 10c).

Figura 10a
Calendario B



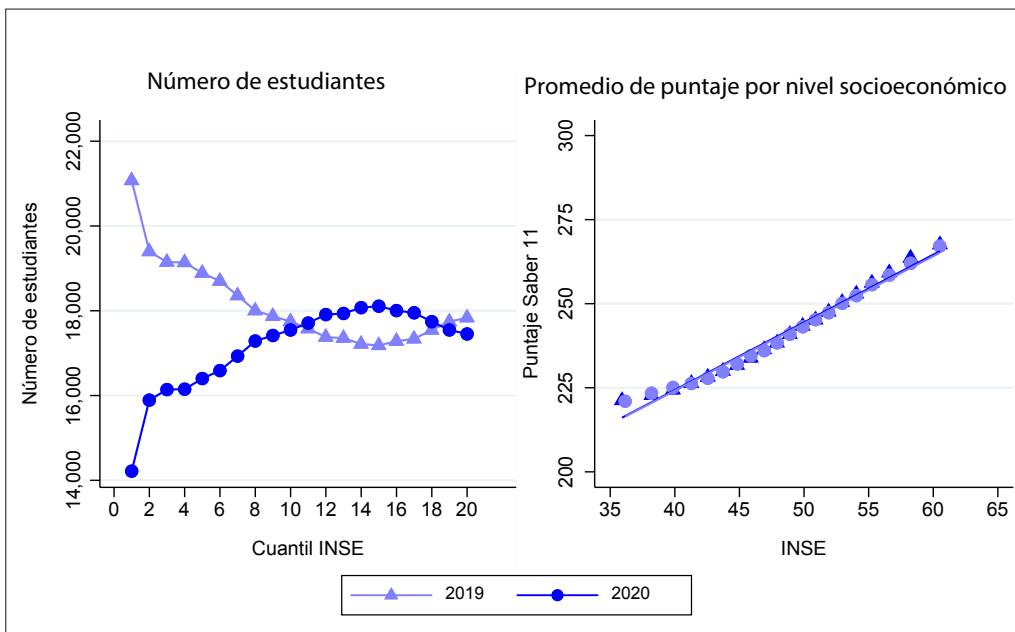
Nota. Elaboración propia con base en datos de pruebas Saber 11 de 2019 y 2020. La distribución de cada año excluye el 5% de estudiantes con INSE más bajo y el 5% de estudiantes con INSE más alto.

Figura 10a
Calendario A, privados



Nota. Elaboración propia con base en datos de pruebas Saber 11 de 2019 y 2020. La distribución de cada año excluye el 5% de estudiantes con INSE más bajo y el 5% de estudiantes con INSE más alto.

Figura 10c
Calendario A, públicos



Nota. Elaboración propia con base en datos de pruebas Saber 11 de 2019 y 2020. La distribución de cada año excluye el 5% de estudiantes con INSE más bajo y el 5% de estudiantes con INSE más alto.

Resultados prueba saber 11 de estudiantes de calendario A, 2021:

Los resultados de las pruebas Saber 11 de estudiantes de colegios de calendario A no están incluidos en los análisis anteriores porque no han sido publicados definitivamente por el ICFES. Sin embargo, análisis realizados sugieren que las diferencias entre el puntaje global de la prueba Saber 11 entre estudiantes de colegios oficiales y no oficiales se ampliaron levemente entre 2020 y 2021. Por ejemplo, en 2020 el primer grupo tuvo un puntaje global promedio de 242 puntos; mientras que el segundo grupo tuvo un puntaje promedio de 272 puntos (30 puntos de diferencia a favor de los estudiantes de colegios no oficiales). Y para el 2021, esta diferencia creció a 32 puntos. No se observaron grandes diferencias en los puntajes de estudiantes de colegios rurales o urbanos o en otras características sociodemográficas.

Efectos de la no presencialidad

Aprendizaje de los estudiantes

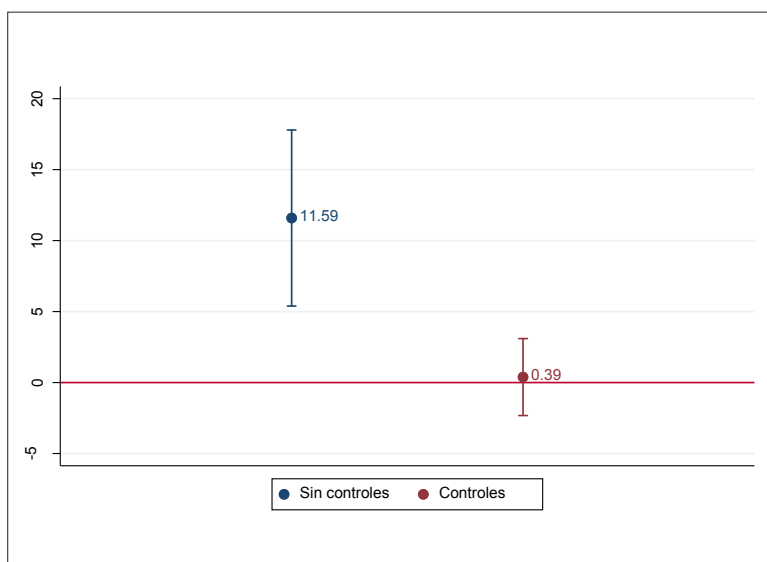
Los resultados de la anterior sección sugieren primero, que un porcentaje importante de estudiantes dejó de

presentar la prueba y, segundo, que la composición de aquellos que la presentaron cambió. Los datos indican que los estudiantes de mayores ingresos de calendario B fueron menos propensos a presentar la prueba en ese período, mientras que los estudiantes de menores ingresos de calendario A no presentaron la prueba en ese ciclo. Si no se toman ambas consideraciones al estimar el efecto de la pandemia en la prueba, los resultados indicarían que existió un aumento relativo en los resultados de los estudiantes más afectados por la pandemia (e.g., los estudiantes de calendario A).

A continuación, se presenta el ejercicio llevado cabo para medir el impacto de la pandemia en las pruebas Saber 11. En primer lugar, se realiza una comparación simple en el tiempo, sin tomar en cuenta estos dos elementos, entre los resultados de colegios de calendario A vis a vis los de calendario B. La forma de estimar el efecto parte de un modelo estadístico de DID. En términos simples, se comparan en el tiempo las diferencias entre los dos tipos de colegios (A públicos versus B). El supuesto clave es que la diferencia se hubiera mantenido igual en ausencia de la pandemia. Para el análisis se toman los colegios observados en todos los periodos (observación de panel por colegio), y se comparan los colegios públicos de calendario A con los colegios privados de calendario B.

Figura 11

Estimación del cambio en el puntaje promedio en Saber 11 entre colegios de calendario B y colegios públicos de calendario A, entre 2019 y 2020



Nota. Elaboración propia con base en estimación de DID del puntaje promedio en pruebas Saber 11, entre colegios públicos y colegios de calendario B. La muestra incluye las sedes con estudiantes que presentaron la prueba en todos los años, entre 2015 y 2020. La estimación con controles incluye dummies de año, zona (rural/urbana), promedio del puntaje INSE por sede, jornada única, porcentaje de mujeres y de estudiantes pertenecientes a etnia que presentaron la prueba, número de estudiantes matriculados en grado 11 y municipio de la sede. Errores estándares robustos.

En un segundo lugar, se presenta una forma de estimar el efecto, tomando en cuenta la pérdida de estudiantes y el cambio de composición (Angrist, Bettinger, & Kremer, 2006). La intuición del ejercicio es la siguiente: para calendario B, se asume que los estudiantes de puntaje más bajo tienen un puntaje igual al del 10% de la distribución de la prueba. Esto mueve la distribución de calendario B “hacia arriba”, acercándola a la observada en los años de no pandemia. Para los estudiantes de calendario A se hace el ejercicio análogo: se asume que los estudiantes de puntaje más alto tienen un puntaje igual al del 90% de la distribución de la prueba, es decir, un puntaje menor al observado. Esto mueve la distribución hacia abajo, imitando la de años anteriores al de la pandemia.⁵

Estimación simple: diferencias en el tiempo entre calendario B y colegios públicos de calendario A

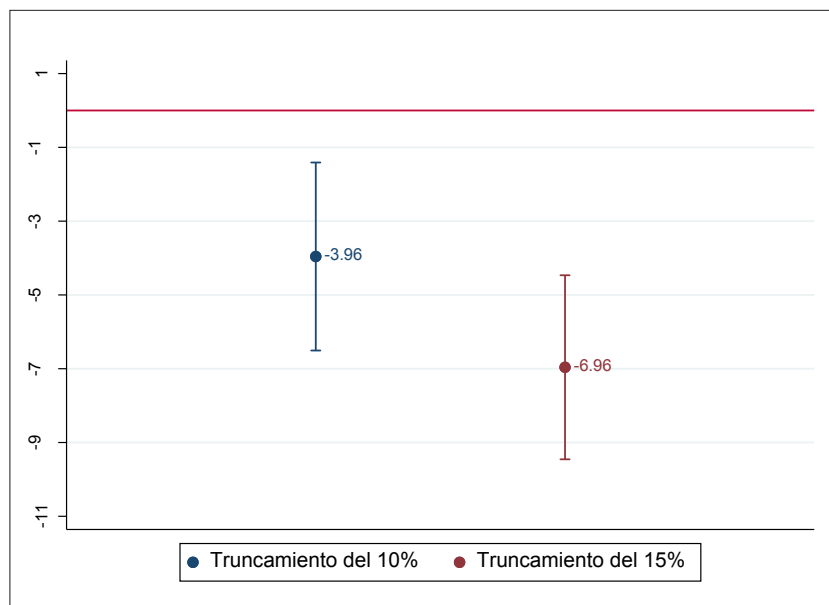
Corrección por número y composición de estudiantes

Los resultados de estos dos ejercicios muestran que la pandemia tuvo un efecto positivo o nulo sobre estudiantes de calendario A, en comparación con el desempeño de los de calendario B. No obstante, una vez se tiene en cuenta el cambio en la composición socioeconómica de los estudiantes entre 2019 y 2020, los efectos cambian. Específicamente, se observa que al truncar las distribuciones de puntajes al 10%, la pandemia disminuyó el desempeño de los estudiantes de calendario A en aproximadamente 4 puntos con respecto a los de calendario B, aumentando las brechas académicas entre estudiantes de bajos y altos recursos económicos.

⁵ En términos técnicos, se estima un modelo Tobit, con truncamiento en el 10% y el 90% de estudiantes y colegios. La estimación adapta el procedimiento en (Angrist, Bettinger, & Kremer, 2006).

Figura 12

Estimación del efecto del cambio en el puntaje promedio en Saber 11 entre colegios de calendario B y colegios públicos de calendario A, entre 2019 y 2020



Nota. Elaboración propia con base en estimación de DID del puntaje promedio en pruebas Saber 11 entre colegios públicos y colegios de calendario B usando un modelo Tobit. Los puntajes de estudiantes de sedes públicas de calendario A se truncaron al 85 y 90 percentil. Los puntajes de estudiantes de calendario B se truncaron al 10 y 15 percentil. La muestra incluye las sedes que tienen estudiantes que presentaron la prueba en todos los años, entre 2015 y 2020. La estimación incluye controles por dummies de año, zona(rural/urbana), promedio del puntaje INSE por sede, jornada única, porcentaje de mujeres y de estudiantes pertenecientes a etnia que presentaron la prueba, número de estudiantes matriculados en grado 11 y municipio de la sede. Errores estándares robustos.