

# *What's up* o Whatsapp: efecto de la conexión a internet de alta velocidad en las pruebas Saber 11

Daniel Lasso

## 1. Resumen

Este texto ahonda sobre el efecto causal de la expansión del internet de alta velocidad en indicadores de calidad de la educación como las pruebas estatales ICFES. Para esto se explota la adopción escalonada en algunos municipios en Colombia del Plan Nacional de Fibra Óptica (PNFO) con una metodología de estudio de eventos. Los resultados indican que no hay evidencia estadística de un efecto causal atribuible al PNFO sobre el puntaje en inglés de los estudiantes en municipios conectados, probablemente porque no aumentó el contenido en otros idiomas consultados por los mismos. No obstante, se estima un aumento pequeño, de en promedio 2 puntos (2.5 % respecto a la media), en el puntaje de matemáticas. Además, se estima un aumento de entre 10 y 15 puntos (5 % frente a la media) en el puntaje global del ICFES para los estudiantes en municipios con conectividad a la fibra óptica. Finalmente, se identifica como posible mecanismo que el PNFO incrementó en 65 puntos porcentuales, respecto a la media, el porcentaje de estudiantes con acceso a internet, lo que explica la mejora en estos resultados.

## 2. Introducción y motivación

Lograr una educación inclusiva y de calidad para todos es uno de los Objetivos del Desarrollo Sostenible adaptados por las Naciones Unidas en el 2015, pues la formación de capital humano es uno de los motores del crecimiento económico y tiene el poder de la transformación de los espacios de los individuos (PNUD, 2015). En particular, la educación se presenta como el medio para el progreso social y para romper las trampas de pobreza, ya que sobre todo para aquellos provenientes de los contextos más vulnerables esta permite diferentes oportunidades que rompen con la dependencia intergeneracional de los resultados socioeconómicos (García J., et al, 2015). No obstante, en Colombia, la calidad de la educación evita que en la práctica estos beneficios se cumplan. Por ejemplo, el país presentó el puesto 61 de 79 en el desempeño en las pruebas estandarizadas Pisa 2018, con puntajes en matemáticas y lectura aproximadamente 38 % inferiores al promedio de los Países OCDE (Schleicher, A., 2018). Una posible explicación de la baja calidad puede ser la brecha en el acceso a internet, pues solo el 43 % de los hogares tiene acceso a este servicio y 1 de cada 4 niños tiene un computador (DANE; Mintic, 2018).

Sin embargo, lejos de ser un debate cerrado, el rol del acceso y uso de internet en el desempeño académico presenta varias visiones contrapuestas (Torres, J., et al, 2016). Por un lado, el uso de internet mejora el desempeño académico a través de mecanismos como el E-learning, que se define como un sistema de aprendizaje con medios electrónicos que ofrece herramientas adicionales para el aprendizaje. También, la existencia de Blogs, Posts y otras formas no estructuradas de información permiten un aprendizaje continuo y autónomo (Agudelo, C., 2017). Además, facilitan la comunicación, socialización, coordinación y colaboración entre estudiantes (Heyam, 2014). No obstante, también puede presentar efectos negativos en el desempeño, principalmente por su fuente de distracción. Sobre esta se señala su uso en mensajería instantánea, las redes sociales y los juegos online, sobre los cuales el internet puede ser una fuente de adicción, desperdicio de tiempo, sobrecarga de información y aislamiento físico de la sociedad (Agudelo, C., 2017; Heyam, 2014). Ya que existen estos dos mecanismos contrarios, resulta de vital importancia establecer literatura propia del contexto colombiano referente al rol del internet en el desempeño académico.

Para el caso colombiano, un programa que resulta de interés para evaluar esta relación es el Proyecto Nacional de Fibra Óptica (PNFO). En este, el Gobierno Nacional -preocupado porque en el año 2011 solo 325 (29 %) municipios presentaban acceso a internet de fibra óptica- desarrolló un programa donde se buscaba promover la ampliación de la infraestructura de fibra óptica existente en el país, para así llegar a conectar más de 700 municipios con mejores condiciones técnicas, económicas y mejores servicios

(Mintic, s.f). Este proyecto se desarrolló bajo la licitación Pública No. 002 de 2011, donde se destinó un presupuesto de 415,837 millones de pesos provenientes del fondo FONTIC (SECOPI, 2011) para garantizar el cumplimiento del programa. Adicionalmente, este consolidó y mantuvo gratuitamente, y por 5 años, internet de alta velocidad en más de 2000 instituciones públicas en estos municipios (Mintic, s.f). Así mismo, se utilizó la fibra óptica por su facilidad de adaptarse a ambientes geográficos complejos y a su relación costo vs beneficio en el largo plazo (Compartel, 2011).

Algunas evaluaciones de impacto han identificado un efecto causal positivo del PNFO en la inclusión financiera de los municipios tratados (Samper, J., 2022) y en la penetración del internet fijo (Chaves, J., 2020). Por lo cual, dada la relación establecida previamente entre el acceso a internet y el desempeño educativo, resulta importante contemplar los posibles eslabonamientos de este programa en el desempeño de los estudiantes. Por todo esto, el presente texto busca estimar cuál fue el efecto causal de la expansión de infraestructura de internet ocurrida en el marco del PNFO en los niveles educativos, medidos con las pruebas Saber 11. Además, buscará resolver i) qué tipo de instituciones se vieron más afectadas por el programa y ii) qué recomendaciones de políticas públicas se pueden derivar de este aprendizaje. Para este caso en particular, se espera que los efectos positivos de la aplicación de estas tecnologías compensen los negativos, pues en el corto plazo la forma más común por la cual los estudiantes serían afectados por el PNFO es a través de la instalación de este en su colegio, lo que evitaría canales de distracción.

### 3. Datos y construcción de variables

Para dar respuesta a la pregunta de investigación se tomó información sobre tres fuentes: el Instituto Colombiano para la Evaluación de la Educación (ICFES), el Ministerio de Tecnología y Comunicaciones (MinTIC) y el SECOP I que corresponde al portal donde las Entidades Estatales registran los contratos y sus documentos anexos.

En primera instancia, se identificó un conjunto de datos en el portal de Datos Abiertos Gov con la información de la [fecha de inicio del PNFO por municipio](#) creada por el MinTIC. Con esto se pudo recopilar información de cuáles municipios habían sido tratados por el PNFO y en qué año, información que se presenta en la figura 1. Adicionalmente, era necesario identificar cuáles eran los 325 municipios que ya tenían conexión de fibra óptica. Para esto, se utilizó el anexo 10 del contrato de [licitación Pública No. 002 de 2011](#), el cual presenta información sobre las estimaciones que hizo el proveedor sobre estos municipios siempre tratados. Por último, se identificó que los municipios que no fueron conectados por el PNFO fueron adjudicados a un contrato firmado en el 2013 para un programa llamado Proyecto Nacional de Conectividad de Alta Velocidad, el cual comenzó su implementación a finales del 2016 e inicios del 2017. Este tiene como fin conectar aquellos municipios que por condiciones geográficas o dificultad de acceso no fueron conectados por los anteriores programas (MinTIC, S.F). Con base en esta información se estableció que la base de datos debía incluir periodos de tiempo hasta el 2016, pues a partir de este momento aparece un programa similar al PNFO pero con características distintas, de forma que de no tenerlo en cuenta habría un problema de identificación al no representar un buen contrafactual del grupo de tratados.

Por parte del ICFES se tomaron los resultados de las pruebas saber 11, para las áreas de inglés, matemáticas y lectura en el periodo 2009 a 2016. Adicionalmente, se tomó el puntaje global que está disponible únicamente del 2012 al 2016. Estos datos son de consulta abierta tras un registro en la [página](#). Sobre esto, es importante reconocer que las pruebas Saber 11 han tenido varias transformaciones metodológicas en los conceptos y criterios de evaluación que dificultan la comparación entre periodos (ICFES, s.f). Por ejemplo, la prueba entre el 2005 al 2014 presentaba áreas como filosofía, química, física, mientras que después del 2014 no, por lo cual no se pueden comparar los puntajes globales. No obstante, varias investigaciones en educación establecen que los puntajes de inglés, matemáticas y lectura crítica son comparables entre estos dos últimos periodos en la medida en que i) están estandarizados según la media del cohorte y ii) eran evaluadas como áreas del conocimiento únicas en los exámenes de ambos periodos. A partir de los resultados por áreas individuales de los estudiantes, su información socioeconómica y sus aspiraciones, se procedió a calcular el promedio por institución educativa de estas variables basados en el código DANE del colegio. Posteriormente, se juntó con información propia del colegio que se registra en el examen saber: como el tipo de jornada, si es privado o público, si es bilingüe o no, entre otras. Así, se pudo establecer un pseudo-panel que sigue 3115 sedes educativas (unidad de nivel) con variación del 2009 al 2016, siguiendo la metodología de secciones cruzadas repetidas de Moffitt (1993).

El cuadro 2 muestra las estadísticas descriptivas separadas por si a la institución le aplicó o no el PNFO. Por ejemplo se puede identificar menores puntajes de las pruebas Saber para aquellas instituciones en municipios donde no aplicó el PNFO. También, hubo una menor proporción de estudiantes que reportaron tener un computador o conexión a internet, lo cual hace evidente la correlación directa que hay entre acceso a internet y desempeño educativo. Así mismo, en los colegios donde aplicó el PNFO habían más colegios de jornada completa. Con respecto a la representatividad, la base de datos de puntajes de las pruebas Saber 11 representa a todo el universo de estudiantes que presentaron las pruebas, por lo cual no hay problemas en cuanto a sesgo de selección (*Selection Bias*). De la misma forma, el mapa 1 presenta la distribución espacial de los municipios afectados por el PNFO, como se puede observar corresponde a una muestra representativa de los municipios de Colombia, ya que el programa fue implementado en las cinco regiones de Colombia.

## 4. Metodología

En primera instancia, la relación entre el acceso a internet es endógena por varias razones, la primera es la presencia de variables omitidas que explican los resultados de las pruebas saber como la motivación de los estudiantes, entre otras. También, la doble causalidad en el sentido en que una menor educación también puede evidenciar una demanda baja por recursos de información como internet. Así, los estimadores de un modelo de regresión lineal son inconsistentes y sesgados, por lo cual este modelo no reflejaría el efecto causal de la política. No obstante, este programa definió un grupo de municipios tratados y no tratados durante el periodo 2013 a 2016. Así mismo, como se mostró en la figura 1, la aplicación del programa varió a través del tiempo, por lo cual, se justifica el análisis de este problema desde la perspectiva de diferencias en diferencias escalonadas (o estudio de eventos). Por esto, se propone un modelo Two-way Fixed Effects ('TWFE') utilizando la siguiente ecuación base:

$$Y_{it} = \beta_0 + \beta_1 PNFO_{it} + \gamma X_{it} + \theta_t + \phi_i + e_{it} \quad (1)$$

Donde  $Y_{it}$  corresponde la variable objetivo para la sede del colegio  $i$  en el periodo  $t$ . Para este caso va a ser  $i$ ) el puntaje promedio de matemáticas, inglés y lectura en el colegio para el periodo 2009 a 2016. La variable  $PNFO_{it}$  corresponde a una dummy que toma el valor de 1 si en el periodo  $t$  en el municipio donde está ubicado el colegio  $i$  entró en operación el Programa y toma 0 de lo contrario. Puntualmente, se pretende hacer énfasis de que esto no significa que **todos** los colegios fueron tratados, pues en la realidad solo 730 instituciones educativas de las más de 3000 que se presentan en la base fueron tratados directamente con un nodo de conexión de fibra óptica (Suárez, F. 2018). Por lo cual, se va a explicar con cuidado la interpretación del coeficiente asociado. Igualmente, el vector  $\gamma X_{it}$  corresponde a los coeficientes asociados a las características de control a nivel del colegio, por ejemplo, si es una institución oficial o no y si es jornada completa o no. Por otra parte, las variables  $\theta_t, \phi_i$  representan los efectos fijos de tiempo y colegio respectivamente, estos son importantes porque permiten controlar por factores no observados que no varíen en el tiempo o espacio respectivamente y que estén asociados a los resultados de las pruebas Saber 11. Finalmente, el término  $e_{it}$  corresponde al término de error de la regresión, para este caso se utilizó una estimación de los errores tipo cluster por colegio debido a que el tratamiento no se da a nivel colegio, sino bajo una categoría superior correspondiente al municipio.

Para este modelo, el efecto causal de interés es capturado por el coeficiente  $\beta_1$  que identifica un ATT, pues captura la diferencia promedio en el puntaje promedio de las pruebas Saber 11 entre los colegios que fueron beneficiados por la llegada del PNFO a su municipio y el grupo de control. Nuevamente, se es enfático que no representa el efecto causal de ser tratado con un nodo de conexión de fibra óptica, hecho que era deseable pero no se encontraron los datos sobre las 735 instituciones públicas conectadas. De igual modo, al existir periodicidades distintas en el tratamiento, el grupo de control toma dos estados posibles de la naturaleza: i) colegios en municipios que nunca fueron beneficiados por el proyecto hasta el año 2016 que entra el Plan Nacional de Conectividad y ii) colegios en municipios que aún no son tratados por el PNFO.

Para que el modelo de TWFE recupere el efecto causal es importante que se cumpla el **supuesto** de identificación clave de tendencias paralelas. Este mantiene que si bien en niveles de la variable de interés puede haber diferencias entre el grupo de tratamiento y los grupos de control planteados anteriormente, esta diferencia se mantiene constante a través del tiempo. Más aún, esto es importante porque si esto ocurre en el pasado se espera que ocurriría después de la aplicación del programa, lo que valida que la

diferencia en las diferencias de los resultados potenciales sea atribuible al efecto del programa. Por otra parte, se requiere de verificar el **supuesto** de no anticipación, este implica que las unidades tratadas no identifican que van a ser tratadas en el futuro.

Otra consideración importante del modelo de TWFE se basa en el método de estimación. En particular, si se estima a través de MCO el estimador de interés  $\beta_1$  puede estar sesgado si el efecto del tratamiento presenta efectos dinámicos o heterogéneos (Godman-Baccon, 2021; Abraham y Sun, 2018). Por lo cual, si se quisiera estimar de esta forma sería necesario establecer dos supuestos adicionales: efectos constantes y efectos homogéneos. Matemáticamente, Godman-Baccon (2021) demuestra que este estimador puede ser expresado asintóticamente como:

$$\text{plim}_{n \rightarrow \infty} (\hat{\beta}_1) = VWATT + VWPT - \Delta ATT$$

Donde VWATT es un promedio ponderado de los efectos causales (ATT), VWPT es un sesgo generado en caso de que no se cumpla el supuesto de tendencias paralelas y  $\Delta ATT$  corresponde a la diferencia entre los ATT para las cohortes o tiempo. Por esto, mismo, si los efectos son dinámicos habrán diferencias en los ATT entre el tiempo, por lo cual es el estimador será sesgado y consistente. Intuitivamente, esto ocurre porque se están comparando incorrectamente unidades tratadas con unidades tratadas tarde, que podrían no ser buenos contrafactuales.

Para el caso de estudio, por construcción teórica no se cumplen los supuestos de efectos constantes ni homogéneos. Primero, la introducción de tecnologías como el internet genera eslabonamientos hacia adelante que son utilizados por empresas e individuos en periodos posteriores a su implementación (Bonnet, J., 2005). Por lo cual, los efectos de esta adopción tecnológica pueden variar con el tiempo. Segundo, los efectos de estos cambios estructurales varían entre cohortes en la medida en que algunas cohortes fueron tratadas después porque el proveedor no pudo generar conexiones por dificultades principalmente geográficas (Compartel, 2011). Estas pueden estar asociadas con heterogeneidades que expliquen el puntaje saber, por lo cual tampoco se cumple el supuesto de efectos constantes.

Para solucionar este problema se va a adoptar el estimador propuesto por Callaway y Sant'Anna (2020), que soluciona eliminando las estimaciones provenientes de las comparaciones incorrectas. Esto no implica un cambio en la especificación dada en la ecuación (1), solamente la forma de estimación. Adicionalmente, para verificar el supuesto de tendencias paralelas del modelo TWFE se plantea el siguiente modelo de efectos dinámicos:

$$Y_{it} = \alpha_0 + \sum_{j=-7}^{-2} \tau_j PNFO_{it}^j + \sum_{j=-0}^3 \tau_j PNFO_{it}^j + \gamma X_{it} + \theta_t + \phi_i + e_{it} \quad (2)$$

Donde la única diferencia con la ecuación (1) es que en esta se crea una medida del tiempo relativo al tratamiento para cada cohorte tratado y se evalúa el efecto del programa en ese tiempo relativo. Puntualmente,  $\sum_{j=-3}^{-2} \tau_j PNFO_{it}^j$  corresponde a los efectos anticipatorios y permite verificar estadísticamente la existencia de tendencias paralelas, pues si se cumple estos coeficientes ( $\tau_i$ ) no deberían ser estadísticamente significativos, lo que determina que no hay diferencias en la tendencia entre los grupos tratados y control para periodos previos al tratamiento. Análogamente, el resto de variables se interpretan de la forma que se comentó antes en la ecuación (1), con errores cluster por colegio.

## 5. Resultados y Pruebas

El cuadro 3 presenta la estimación realizada por Callaway y Sant'Anna (2020) del efecto causal promedio del PNFO en los puntajes de las pruebas Saber 11 en las áreas de matemáticas, inglés y lectura crítica. En particular, se puede encontrar que en promedio el PNFO aumentó en 1.2 puntos el puntaje promedio de matemáticas en los colegios tratados y 1.5 puntos en lectura. Estos resultados son estadísticamente significativos con una confianza del 99 %. Esto provee evidencia a la literatura de que el acceso a internet en los estudiantes puede, en el neto, ser positivo. Respecto a la significancia económica, el puntaje promedio para matemáticas y lectura era aproximadamente 46 puntos, por lo cual este efecto corresponde a un incremento del 2.1 %. De esta forma, no parece tener un efecto grande en el puntaje de los estudiantes. Además, no se encuentra evidencia estadística para sustentar que el programa tuvo un incremento en el puntaje de Inglés, esto indicaría que el acceso a internet no derivó en que los estudiantes consultaran

fuentes de información en otros idiomas. Finalmente, en el cuadro 1 se puede identificar la existencia de efectos heterogéneos entre los cohortes. Por ejemplo, el cohorte del 2013, los primeros tratados, se ven mucho más afectados que los últimos en ser tratados. Esto se puede deber a que el tratamiento se dio con base en priorizaciones como la inexistencia del acceso a internet o telecomunicaciones, por lo cual, estos primeros pudieron aprovechar la incursión total de estas tecnologías ([Compartel, 2011](#)).

Por otra parte, la figura 2 presenta los efectos dinámicos del PNFO en el desempeño del Saber 11 en inglés y matemáticas siguiendo la estimación de la ecuación (2). A nivel general esta gráfica ofrece evidencia visual del cumplimiento de tendencias paralelas entre los grupos de tratamiento y control, pues no se encuentran diferencias en las tendencias de los puntajes de inglés y matemáticas. Además, el cuadro 4 presenta evidencia estadística que muestra que las tendencias paralelas se mantienen, pues como se puede identificar solo un rezago es estadísticamente significativo al 5 %, lo que indica que 4 periodos antes del tratamiento hay una perturbación en la tendencia de las pruebas saber entre los grupos de tratamiento y control. Así mismo, como se puede ver no hay evidencia estadística de la existencia de efectos anticipatorios. Esto se explica porque en el programa no se definió completamente cuales municipios serían tratados, sino que se le dio libertad al contratista de seleccionarlos según sus estudios, basado en una lista de priorización que entregó el Gobierno.

No obstante, en últimas la corroboración de las tendencias paralelas corresponde a una definición teórica. En particular, Cunningham, S. (2021) evidencia que para que se cumplan las tendencias paralelas es necesario que la variable de interés sea independiente de la asignación del tratamiento ( $Y_{it} \perp D_{it}$ ). Esto quiere decir que la asignación al PNFO no depende de los resultados potenciales del ICFES. Como lo menciona el anexo técnico del proyecto, el tratamiento de los municipios se debió a los resultados de posibilidades de acceso y conexiones de los informes técnicos elaborados por el contratista. En ese sentido, el principal criterio de asignación fue particularmente la geografía ([Compartel, 2011](#)). En principio, se podría pensar que la geografía condiciona el desarrollo educativo a través del acceso y transporte de profesores y estudiantes a las escuelas (Ramírez Salazar, 2008). No obstante, al menos en la muestra actual la mayoría de colegios corresponden a establecimientos urbanos. Además, es importante entender que el proyecto se realizó porque buscaba llevar internet de fibra óptica a regiones de difícil acceso, por lo cual varios de estos municipios tenían las mismas condiciones geográficas adversas.

Del mismo modo, un posible sesgo puede estar dado por la configuración de los puntajes de las pruebas. Como se mencionó anteriormente, por más de que los puntajes estén estandarizados puede que no sean comparables, por esto mismo, resulta importante identificar que el efecto persiste aún cuando se utilizan pruebas comparables. Así, la figura 3 presenta que tras limitar la muestra al periodo 2012-2016, que comprenden un mismo periodo de características del examen, se estima que el PNFO tuvo un efecto causal de entre 10 y 15 puntos en el puntaje global promedio de la sede educativa. Así mismo, se observa que hay evidencia del cumplimiento de tendencias paralelas, no obstante, hay pocos periodos pre-tratamiento para asegurar la comparabilidad. Igualmente, en cuanto a su significancia económica, el puntaje global promedio durante ese periodo fue de 253 puntos, por lo cual este incremento corresponde a un aumento 5 %, un aumento mayor que el estimado anteriormente.

En general, como lo muestra la gráfica 4 se encuentra que un posible mecanismo por el cual incrementan estos puntajes es a través del acceso a internet por parte de los estudiantes. Puntualmente, se encuentra un incremento promedio de 5 puntos porcentuales en la cantidad de estudiantes que reportaban tener acceso a internet en la encuesta socio-demográfica del ICFES en los municipios donde se implementó el PNFO. Este incremento es en magnitud grande, pues respecto a la media es un incremento de 65 % en el acceso a internet en estudiantes. También, fue posible identificar que hubo un mayor efecto para aquellas instituciones educativas oficiales y no bilingües (ver cuadro 1), esto puede ser porque en algunas de estas instituciones educativas se establecieron los nodos de conexión de la fibra óptica y se permitieron los planes de 5 años. Además, las instituciones bilingües ya tenían acceso a internet, entonces el efecto marginal del programa era bajo.

Finalmente, la figura 4 presenta un chequeo de robustez de la metodología frente a distintos métodos de estimación. Por ejemplo, se exploran otros métodos como el postulado de Abraham y Sun (2020). En este se puede identificar que los resultados son parecidos para los métodos de estimación que eliminan la comparación entre tratados con tratados tarde como Chaisemartin and D'Haultfoeuille (2020) y Callaway y Santana (2020). Así mismo, en estos se encuentra un efecto positivo del programa y el cumplimiento

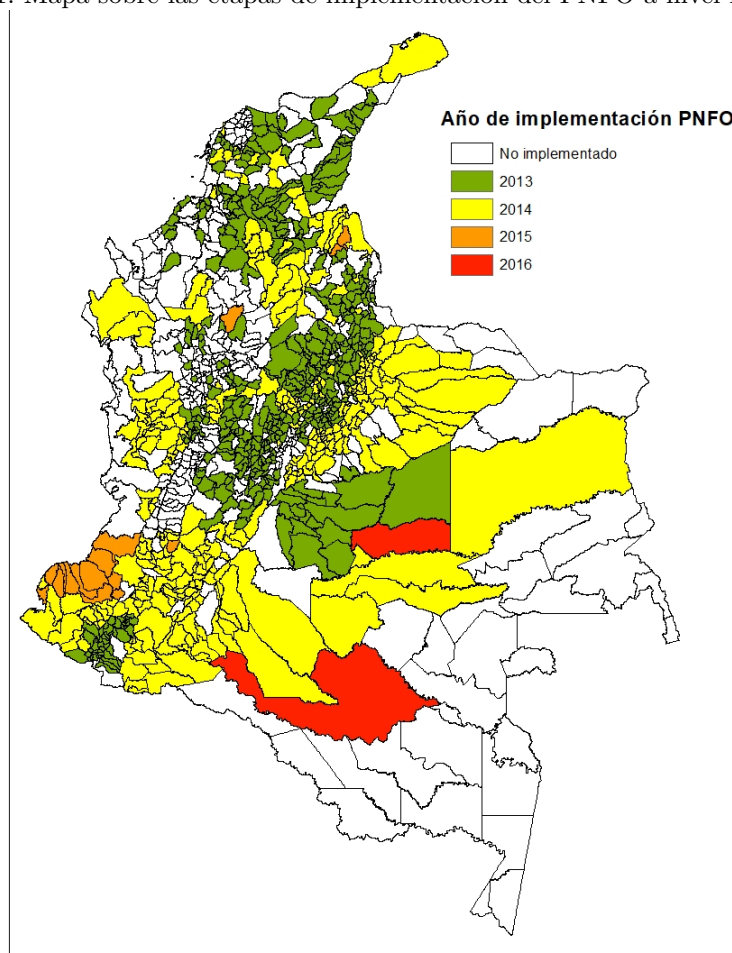
de tendencias paralelas para estos métodos.

## 6. Conclusiones

En conclusión se puede identificar un efecto causal del PNFO en el desempeño de las pruebas Saber 11 de matemáticas y los puntajes globales de los estudiantes. Aunque, en inglés no se encuentra ningún beneficio del programa. Estos resultados son robustos y principalmente, están explicados por los eslabonamientos del acceso a internet, puesto que en principio el programa no ofrecía internet a las familias. Así mismo, se encuentra que el efecto es económicamente pequeño, pues aumentó en 2.1 % los puntajes en pruebas específicas, lo cual es poco para una inversión total de más de 415 mil millones de pesos. En este sentido, una recomendación de política pública para mejorar el resultado de las pruebas estaría orientada a ofrecer las conexiones directamente a los hogares de los estudiantes y no en sus instituciones públicas. También, a establecer programas que se especialicen en atender directamente las necesidades educativas y que esta no sea una externalidad o eslabonamiento.

## 7. Anexo y tablas

Figura 1: Mapa sobre las etapas de implementación del PNFO a nivel municipal



Nota: Elaboración propia con los datos del PNFO de MinTIC

Cuadro 1: Efecto en el puntaje de matemáticas separado por características de los colegios

	No oficiales	Oficiales	Bilingue	No bilingue
ATT	0.899 (0.526)	1.256*** (0.260)	-0.178 (1.302)	1.364*** (0.258)
<i>N</i>	341	21816	2030	20125

**Nota:** esta tabla corresponde a la estimación de la ecuación (2) por el método de CS (2020). Se presentan los errores por cluster de colegio en paréntesis con la significancia estadística de la forma: \*  $p < 0,05$ , \*\*  $p < 0,01$ , \*\*\*  $p < 0,001$

Cuadro 2: Efecto en el puntaje de matemáticas separado por características de los colegios

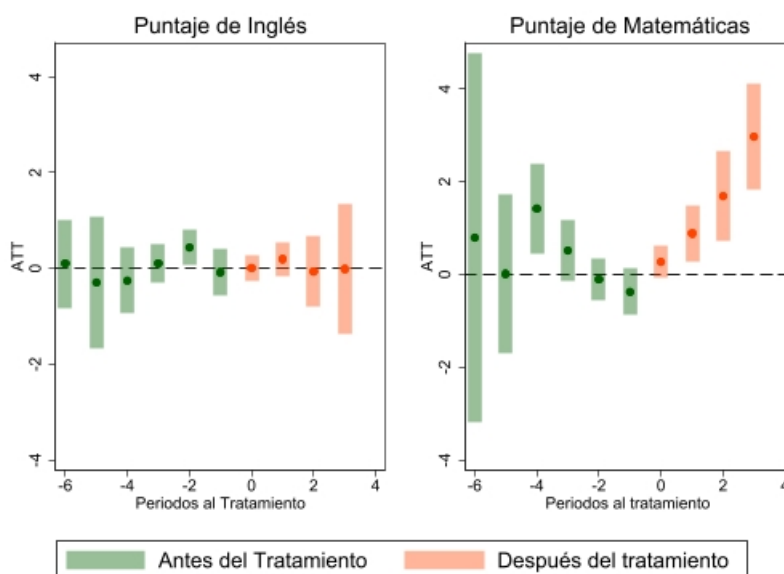
No aplicó el PNFO					
	N	Mean	D.E	Min	Max
puntaje matematicas	673	42.978	5.413	21.167	60.821
puntaje ingles	673	44.885	4.355	35	68.727
puntaje lectura critica	673	44.005	4.975	26.8	61.071
puntaje global	448	222.619	23.038	173.455	305.536
tipo jornada completa	673	.087	0.278	0	1
Colegio Oficial	673	.944	0.231	0	1
Cuenta con Internet	673	.076	0.139	0	1
Cuenta con computador	673	.211	0.229	0	1
Aplicó el PNFO					
puntaje matematicas	22356	46.035	5.249	24.5	82
puntaje ingles	22356	45.453	4.585	28.286	93.097
puntaje lectura critica	22356	46.318	4.519	27.7	77
puntaje global	14932	236.124	22.028	159.467	384.087
tipo jornada completa	22356	.292	0.445	0	1
Colegio Oficial	22356	.901	0.298	0	1
Cuenta con Internet	22356	.113	0.165	0	1
Cuenta con computador	22356	.248	0.230	0	1

Cuadro 3: Efectos por cohortes de tratamiento del PNFO sobre el desempeño Saber 11

	Puntaje Matemáticas	Puntaje Inglés	Puntaje Lectura
Efecto causal Promedio	1.200*** (0.320)	0.0671 (0.232)	1.508*** (0.256)
Efecto causal en el cohorte 2013	1.455*** (0.300)	-0.127 (0.261)	1.697*** (0.339)
Efecto causal en el cohorte 2014	0.933* (0.500)	0.300 (0.295)	1.337*** (0.383)
Efecto causal en el cohorte 2015	0.220 (0.574)	0.102 (0.396)	0.0567 (0.461)
Efecto causal en el cohorte 2016	-0.363 (2.645)	0.334 (1.118)	0.528 (1.088)
Puntaje Promedio	45.94	45.43	46.25
<i>N</i>	22157	22157	22157
Controles	Sí	Sí	Sí

**Nota:** esta tabla corresponde a la estimación de la ecuación (1) por el método de CS (2020) por cohortes de tratamiento usando el comando CSDID de Stata. La estimación corresponde a una ponderación por cohortes del ATT, medida con una dummy que toma el valor de 1 si el colegio estuvo en un municipio con acceso al PNFO. Se tomaron como controles si la jornada era de tiempo completo o diurna y si el colegio era privado o público. Se presentan los errores por cluster de colegio en paréntesis con la significancia estadística de la forma: \*  $p < 0,05$ , \*\*  $p < 0,01$ , \*\*\*  $p < 0,001$

Figura 2: Efecto del PNFO sobre el desempeño Saber 11, estudio de eventos



Nota: estimación con la metodología de CS (2020) de la ecuación (2) sobre efectos dinámicos en el modelo TWFE. SE utilizó el comando CSDID en Stata. Se aplicaron errores estandar tipo cluster a nivel de sede educativa. Se adicionaron como controles si la jornada era de tiempo completo o diurna/nocturna y si el colegio era privado o público.

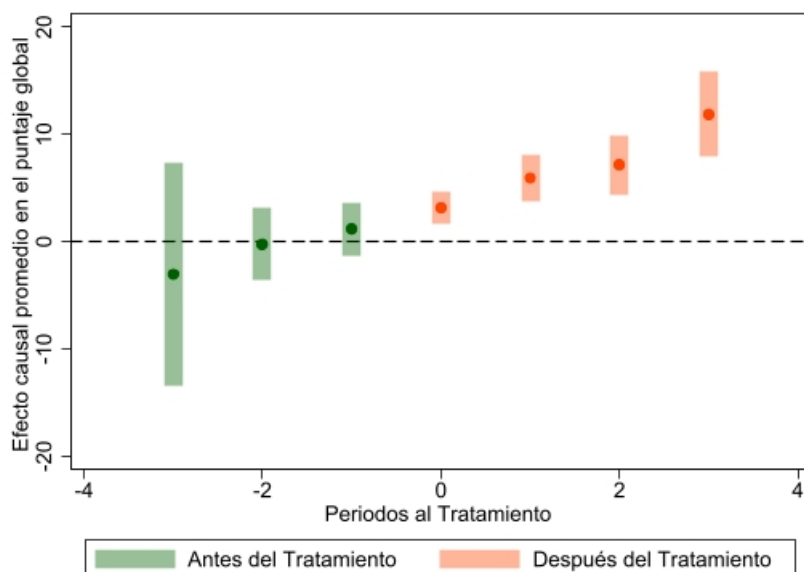
Cuadro 4: Estimación de las tendencias paralelas entre colegios tratados y control para los puntajes Saber

	Puntaje Matemáticas	Puntaje Inglés
Estimación de los leads		
6 Periodos antes	0.792 (2.033)	0.102 (0.469)
5 Periodos antes	0.0107 (0.872)	-0.293 (0.698)
4 Periodos antes	1.423** (0.492)	-0.254 (0.353)
3 Periodos antes	0.515 (0.330)	0.104 (0.211)
2 Periodos antes	-0.106 (0.223)	0.435* (0.193)
1 Periodo antes	-0.374 (0.252)	-0.0821 (0.249)
Promedio	0.377	0.00189
Antes del tratamiento	(0.390)	(0.193)
N	22157	22157
Controles	Sí	Sí

**Nota:** esta tabla corresponde a la estimación de la ecuación (2) por el método de CS (2020). Solo se presentan las estimaciones de los leads para evidenciar estadísticamente las tendencias paralelas. Se tomaron como controles si la jornada era de tiempo completo o diurna y si el colegio era privado o público. Se presentan los errores por cluster de colegio en paréntesis con la significancia estadística de la forma: \*  $p < 0,05$ , \*\*  $p < 0,01$ , \*\*\*  $p < 0,001$

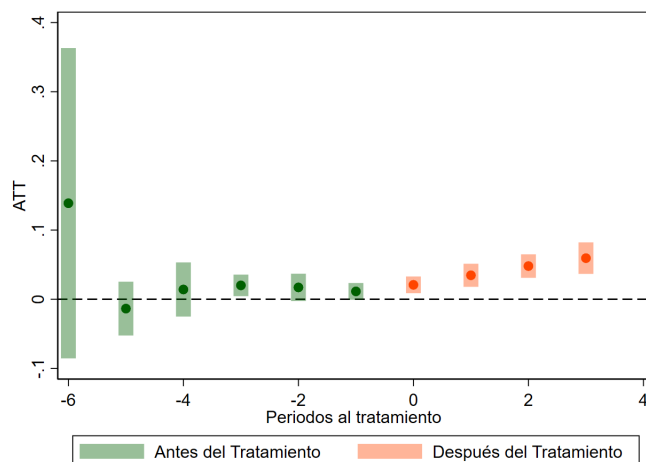


Figura 3: Efecto del PNFO sobre el puntaje global Saber 2012-2016, estudio de eventos



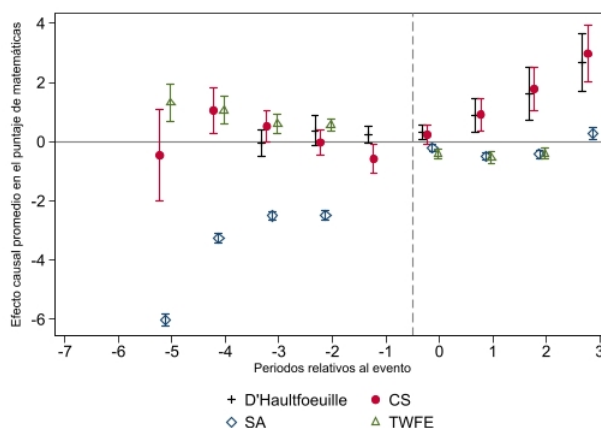
Nota: estimación con la metodología de CS (2020). Se limitó la muestra al puntaje global reportado para el ICFES en periodos comparables del 2012-2016. Se aplicaron errores estandar tipo cluster a nivel de sede educativa. Se adicionaron como controles si la jornada era de tiempo completo o diurna/nocturna y si el colegio era privado o público

Figura 4: Efecto del PNFO sobre la proporción de estudiantes con acceso a internet



Nota: estimación con la metodología de CS (2020). Se aplicaron errores estandar tipo cluster a nivel de sede educativa. Se adicionaron como controles si la jornada era de tiempo completo o diurna/nocturna y si el colegio era privado o público

Figura 5: Cálculo del efecto causal en el puntaje de matemáticas a través de distintos métodos de estimación



Nota: la gráfica resume los siguientes métodos de estimación. En Azul Abraham y Sun (2018), en rojo Callaway y Santana (2020), en negro Chaisemartin and D'Haultfoeuille (2020) y en verde la estimación por MCO del modelo de TWFE. Se aplicaron errores estándar tipo cluster a nivel de sede educativa. Para esta comparación no se adicionaron controles.

## 8. Referencias

Abraham, S., Sun, L. (2018). Estimating Dynamic Treatment Effects in Event Studies With Heterogeneous Treatment Effects. SSRN Electronic Journal. doi:10.2139/ssrn.3158747

Agudelo, C., (2017) influencia del acceso a internet en los resultados de las pruebas Saber 11 de los años 2011, 2012, 2013 y 2014 en el área de lenguaje en dosquebradas, Risaralda.

Bonet, J., 2005 Cambio estructural regional en Colombia: una aproximación con matrices insumo-producto <https://www.banrep.gov.co/es/cambio-estructural-regional-colombia-aproximacion-con-matrices-insumo-producto>

Callaway, B., Sant'Anna, P. H. (2020). Difference-in-Differences with multiple time periods. Journal of Econometrics. doi:10.1016/j.jeconom.2020.12.001

Cunningham, S. (2021). Causal Inference: The Mixtape. Yale University Press.

García J., S, Rodríguez, C., Sánchez, F. y Bedoya, J. (2015). La lotería de la cuna : la movilidad social a través de la educación en los municipios de Colombia. Universidad de los Andes, Facultad de Economía, CEDE.

Godman-Baccon, Difference-in-Differences with Variation in Treatment Timing, Journal of Econometrics, March 2021, 225(2).

Heyam, A. (2014). The Influence of Social Networks on Students' Performance. Journal of Emerging Trends in Computing and Information Sciences, 5(3), 200-205. <http://doi.org/10.3916/C41-2013-19>

ICFES, s.f, Documentación del examen Saber 11

Moffitt, R. (1993). Identification and Estimation of Dynamic Models with Time Series of Repeated Cross-Sections, Journal of Econometrics, 59,99-123.

Schleicher, A. (2018) Insights and Interpretations, PISA. Recuperado de: <https://www.oecd.org/pisa/PISA>

Torres-Díaz, J. C., Duarte, J. M., Gómez-Alvarado, H. F., Marín-Gutiérrez, I., Segarra-Faggioni, V. (2016). Internet use and academic success in university students. Comunicar, 24(48), 61-70. <https://doi.org/10.3916/C48-2016-06>

