

**Relación del acceso a recursos tecnológicos con el desempeño escolar de
los colegios oficiales de educación media en Colombia**

Manuel Esteban Buitrago Botero

Trabajo de grado

Asesora: María Natalia Cantet

Universidad EAFIT

Escuela de Economía y Finanzas

Medellín, Colombia

2022

Resumen

Este trabajo bajo una metodología de propensity score matching muestra evidencia de una relación positiva de los recursos tecnológicos y el acceso a internet con el desempeño escolar de los colegios oficiales de educación media en Colombia en el puntaje promedio global y por áreas en las pruebas Saber 11 Calendario A para el año 2021, tomando como variables de tratamiento el acceso a internet mediante red LAN, WLAN o cualquiera de estas, y el acceso a recursos tecnológicos entendidos como computadores portátiles, de escritorio y tabletas. Adicional a esto, este trabajo de manera resumida contextualiza el sistema educativo del país en base al acceso a dichos recursos.

Estos resultados enfatizan sobre la importancia de garantizar el acceso a las nuevas tecnologías a todos los participantes del sistema educativo bajo unas condiciones de calidad y equidad, de tal manera que no solo se logre cerrar la brecha digital existente, sino que también se permita la transición a un modelo educativo que permita aprovechar todas las ventajas que dichas herramientas tecnológicas tienen en el proceso de enseñanza. A diferencia de otros trabajos de investigación enfocados en la misma línea temática, este trabajo aborda la problemática desde una metodología que permite hacer comparaciones entre grupos de tratamiento y tiene un enfoque diferenciador dado que busca explicar posibles asociaciones entre el desempeño escolar con el acceso a recursos tecnológicos desde una visión de la infraestructura tecnológica de la sede educativa, lo que puede convertirse en una fuente de información valiosa para los hacedores de política educativa.

Palabras clave: Economía de la educación, Propensity Score Matching, TICs, internet, desempeño escolar.

I. Introducción

La tecnología ha tenido un papel fundamental en el desarrollo de la humanidad y su creación surge de las necesidades mismas del ser humano. Herramientas en particular como los computadores e internet han permitido la generación de procesos innovadores en todos los campos del conocimiento, incluido el campo de la educación. De hecho, es aceptado que la implementación de las TIC (Tecnologías de la Información y las Comunicaciones), en el ámbito educativo permite que el sistema educativo sea lo suficientemente flexible como para realizar los cambios necesarios en la enseñanza que permitan suplir la demanda de la sociedad en términos de conocimiento (Pastrana y Saludes, s.f).

Según la OECD (2014), citado por MinTIC (2014), “los entornos más innovadores y con mejores resultados de aprendizaje son aquellos que, además de un sistema pedagógico innovador integran las nuevas tecnologías”, y es por esto que, para las entidades encargadas de formular las políticas educativas, ha tomado vital importancia garantizar que todos los participantes del ecosistema educativo cuenten con acceso a recursos tecnológicos, hagan uso de los mismos y los integren a su modelo de enseñanza.

Como evidencia de lo anterior, la evolución en el indicador de número de estudiantes por computador en las sedes educativas oficiales muestra un avance significativo con respecto a la dotación de equipos de cómputo. Mientras para el año 2010 en promedio se contaba con un computador por cada 20 estudiantes, en el año 2021 se contaba con un computador por cada 8 estudiantes (Mineducación, 2022). Sin embargo, no es claro si dichas políticas de dotación de recursos tecnológicos logran impactar por sí solas los indicadores de desempeño de las sedes educativas oficiales del país, y para ello es necesario realizar un diagnóstico de las condiciones de infraestructura tecnológica de las sedes educativas y analizar también las demás políticas educativas que se vienen implementando como complemento a las políticas de dotación y conectividad escolar.

Una de estas estrategias es la denominada Conexión total, la cual tiene como objetivo garantizar la disponibilidad, sostenibilidad y calidad de la conexión a internet en las sedes educativas oficiales aportando al propósito del Gobierno Nacional de que el país esté conectado al 100% y se cierre la brecha digital que existe actualmente (Mineducación, 2021). Dicha estrategia, “genera los lineamientos y brinda la asistencia técnica necesaria para garantizar la prestación del servicio de conectividad de las instituciones educativas oficiales con la mayor calidad posible” (Mineducación, 2017). Conexión total, en el marco del proyecto para el fortalecimiento del ecosistema de innovación educativa EAFIT-MEN, Conexión total logró en el año 2021 desarrollar un ejercicio de caracterización de necesidades de conectividad en las sedes educativas oficiales del país, permitiendo tener una visión respecto al aspecto de conectividad y energización escolar en cada una de las regiones (Universidad EAFIT, 2021).

Es importante entonces poder hacer uso de información tan valiosa y en este caso se utilizará como material esencial para la evaluación del impacto del acceso a recursos tecnológicos en el desempeño de los estudiantes en las pruebas Saber 11° para el año 2021, y que esto sirva como un insumo para los tomadores de decisiones, que permita luego orientar las políticas educativas en busca de cumplir con los objetivos propuestos en los planes de educación.

II. Marco teórico y antecedentes del problema

Las Tecnologías de la Información y las Comunicaciones, TICs, son entendidas como “la combinación de herramientas tecnológicas y recursos que son usadas para manipular y comunicar la información” (Kaware, S. S., & Sain, S. K. , 2015).

Estas tecnologías incluyen productos y servicios tales como computadoras de escritorio, portátiles, dispositivos de mano, internet alámbrico o inalámbrico, software de negocios y productividad como editores de texto y hojas de cálculo, software empresarial, almacenamiento de datos y seguridad, seguridad de red y otros. (Ashrafi & Murtaza, 2008, pp 125-138)

Según la ONU (2021), las TICs han facilitado importantes cambios económicos en las últimas décadas, permitiendo el acceso a información para la ciencia la tecnología y la innovación. En el campo de la educación, estas herramientas han permitido paulatinamente la transición de un modelo de educación tradicional a un modelo de educación constructivista, en el cual el alumno aprende a aprender. De esta manera, la dinámica del proceso educativo cambia radicalmente facilitando que el docente pueda enfocar su atención en los alumnos con más dificultades de aprendizaje, y que el estudiante no solo se muestre más interesado en sus pendientes sino, que también desarrolle aptitudes de resolución de problemas (Mao & Casas, 2004).

Las TICs en el campo educativo tienen los objetivos de permitir proveer oportunidades de educación mediante el aprendizaje a distancia para personas que por condiciones socioeconómicas han sido históricamente excluidas (como poblaciones en áreas rurales y estudiantes con discapacidades), generar redes de intercambio de conocimientos, complementar la labor del docente con otras habilidades y responsabilidades y mejorar la eficiencia y la eficacia de la administración y las políticas educativas. (Kaware & Sain, 2015)

Estudios realizados muestran diferentes consideraciones respecto al impacto que el acceso a recursos tecnológicos tiene sobre el rendimiento académico de los estudiantes. Por ejemplo, Barrera y Linden (2009) encuentran que el programa Computadores Para Educar, el cual realiza el proceso de remodelación de equipos de cómputo donados por el sector privado para instalarlos en escuelas públicas, además del entrenamiento para los profesores para el uso de estos en las prácticas curriculares especialmente en el área de lenguaje, tiene un pequeño efecto positivo sobre algunas pruebas. Los autores indican que

estos resultados obedecen a fallas en la implementación de los equipos en las prácticas curriculares.

Goolsbee & Guryan (2006) tomando como indicador de resultados los puntajes del Stanford Achievement Test, encuentran que el programa estatal de dotación tecnológica que se empezó a implementar en Estados Unidos en 1998, denominado E-rate program, en sí mismo no tuvo un impacto significativo en el desempeño de los estudiantes en dicho test, y lo argumentan en que los docentes no estaban preparados para poder incursionar los recursos tecnológicos que habían sido recibidos en sus prácticas pedagógicas, y también en que es probable que las políticas de dotación en tecnología tengan un impacto en los resultados tras un tiempo de rezago. Para dicho investigación, hicieron uso de un diseño de regresión discontinua y una regresión por mínimos cuadrados ordinarios.

En el mismo sentido, Beuermann et al (2015) afirma que luego de un experimento aleatorio controlado en Perú, como parte del programa One Laptop per Child, iniciativa que fomenta el aprendizaje autónomo mediante la entrega de computadores portátiles a niños en países en desarrollo, se concluye que para este caso, no se presentan diferencias significativas en cuanto a resultados académicos y habilidades de uso del computador entre los estudiantes incluidos en el grupo de control y los que no, y que por el contrario, aquellos estudiantes incluidos en el grupo de control en general, muestran una tendencia significativa a demostrar menos esfuerzo en la escuela.

Por su parte, en un estudio con el objetivo de encontrar los determinantes del rendimiento de los estudiantes del departamento de Nariño, Colombia, en las pruebas SABER para el 2021 mediante modelos probit y logit ordinales se concluye que “el acceso a herramientas tecnológicas como computadoras y conexión a la red de internet aumenta la probabilidad de obtener un mejor rendimiento académico, gracias al acercamiento del estudiante al conocimiento global y al acceso a diversas fuentes de información” Rodríguez et al (2021).

En el mismo sentido, Flórez (2021), mediante un algoritmo de emparejamiento (Propensity Score Matching) asegura que en promedio el hecho de contar en casa con un computador conexión a internet tiene un impacto positivo sobre el puntaje global de los estudiantes en las pruebas SABER para el año 2021.

A diferencia de la mayoría de las investigaciones mencionadas, este trabajo aborda la problemática mediante una metodología que permite estimar efectos en grupos de tratamientos, hace uso de una base de datos muy completa construida en el año 2021, con las respuestas de los rectores de las sedes educativas sobre la condición de las mismas en cuanto a infraestructura física y tecnológica, y la complementa con la base que recoge uno de los principales indicadores de desempeño de las sedes educativas del país como lo son las pruebas SABER 11°.

III. Metodología

Pareamiento por puntaje de propensión o Propensity Score Matching (PSM), “es una técnica estadística que permite hacer estimaciones sobre grupos de tratamientos en los que por una u otra razón no se puede hacer aleatorización” (Powell et al, 2019). Aunque es una técnica muy utilizada en la ciencia médica, este método puede ser aplicado a cualquier área del conocimiento y en este caso será aplicado en el campo de la economía de la educación.

PSM utiliza información de un conjunto de unidades que no participan en la intervención para identificar lo que les habría sucedido a las unidades participantes en la ausencia de la intervención. Al comparar cómo los resultados difieren para los participantes en relación con no participantes observacionalmente similares, es posible estimar los efectos de la intervención. (Heinrich et al, 2010)

El Propensity Score para el participante i , $PS(X_i)$ definido como la probabilidad de que dicho individuo sea asignado al grupo de tratamiento, es:

$$PS(X_i) = p(Z_i = 1 | X = X_i) \quad (1)$$

Donde $Z=1$ representa que el individuo fue asignado al tratamiento y X denota la serie de covariables observadas. (Powell et al, 2019). Para el cálculo del Propensity Score, en este caso se recurre a un modelo probit aunque bien se pudiera recurrir a un modelo logit que muestre las probabilidades de pertenecer al grupo de tratamiento dadas las covariables observadas como se explicó en el párrafo anterior.

Según Blundell & Costa Dias (2002), el cálculo del PSM, se basa en la suposición de independencia condicional, la cual se expresa como:

$$(Y(1), Y(0)) \perp Z | X \rightarrow (Y(1), Y(0)) \perp Z | P(X) \quad (2)$$

Como muestran Rosenbaum & Rubin (1983) dicha suposición sigue siendo válida si se controla por $P(X)$ en vez de X . Bajo esta suposición, “dada una serie de covariables observadas X , luego de controlar por dichas covariables, los resultados potenciales son independientes del estado de tratamiento” (IDB, 2010)

Adicional a dicho supuesto, también el cálculo del PSM está basado en la suposición de soporte común, según la cual “la proporción de individuos tratados y no tratados debe de ser más grande que cero para cada posible valor de X ” (IDB, 2010)

$$0 < P(D = 1|X) < 1 \quad (3)$$

En base a la ecuación 1 y como destaca Caliendo & Kopeinig (2008), el parámetro más usado para medir los efectos del tratamiento es el average treatment effect on the treated, el cual muestra el efecto sobre los individuos que pertenecen al grupo de tratamiento. El ATT matemáticamente se define como:

$$ATT = E(Y(1)|Z = 1) - E(Y(0)|Z = 0) \quad (4)$$

Teniendo en cuenta la ecuación 2, se tendría que finalmente el ATT es definido como:

$$ATT = E_{P(X)|Z=1}[E(Y(1)|Z = 1, P(X)) - E(Y(0)|Z = 0, P(X))] \quad (5)$$

Donde Y es la variable de resultado, que en nuestro caso corresponde al desempeño promedio de la sede educativa en las pruebas Saber 11.

En el caso de este estudio, lo que se busca es crear un algoritmo de emparejamiento con las sedes educativas que tienen valores similares en las covariables observadas, sometiendo a dichas sedes educativas a una intervención o tratamiento, que sería el acceso por parte de las sedes educativas a recursos tecnológicos e internet, y estimando el efecto que dicho acceso tendría en el puntaje promedio de los estudiantes en las pruebas SABER ICFES calendario A para el año 2021 con respecto a las sedes educativas que no tienen acceso a dichos recursos, con el conocimiento de que la naturaleza no aleatoria de esta investigación, no permite realizar inferencia causal del efecto del tratamiento en la variable de resultados.

Para esta investigación, el algoritmo se construye bajo el modelo de emparejamiento del vecino más cercano, en inglés nearest neighbor matching, el cual “selecciona para emparejar con el individuo tratado aquel individuo no tratado con el puntaje de propensión más cercano al del tratado. Si múltiples individuos no tratados tienen puntajes de propensión cercanos al del individuo tratado, se selecciona uno de los individuos no tratados al azar”. (Austin, 2011)

a. Datos

Para la realización de dicho propósito se toman como insumos la base de datos de los resultados de pruebas Saber 11 Calendario A 2021, la base de datos resultante del formulario de caracterización de necesidades de conectividad de establecimientos educativos en el marco del programa Conexión Total en el año 2021 y la base de datos de SIMAT para el año 2021.

El formulario de caracterización de necesidades de conectividad de establecimientos educativos recogió datos en el año 2021 para 26.912 sedes educativas de todo el país en los aspectos de energía, infraestructura física, red LAN, red WLAN, conectividad e inventario de equipos. La recolección de datos se hizo mediante un formulario web diseñado por el programa Conexión total, que diligenciaban los rectores de los establecimientos educativos oficiales, con la instrucción de la Secretaría de Educación correspondiente. Además, se realizó recolección de información parcial a través de grupos focales, lo que garantiza el cruce de información desde varias fuentes de datos oficiales y confiables, disminuyendo los posibles sesgos de información. Cabe destacar que no se realizó ningún muestreo inicial para la recolección de estos datos, y más bien, el diligenciamiento de dicho formulario se hizo a voluntad de los rectores de las sedes educativas bajo el pedido de las diferentes Secretarías de Educación.

Respecto a la base de datos de las pruebas saber 11° ICFES 2021 calendario A, se cuenta con información agregada a nivel de sede educativa del puntaje promedio de los estudiantes en las áreas de lectura crítica, matemáticas, ciencias sociales y ciudadanas, ciencias naturales e inglés, además del puntaje promedio global, que es el resultado de la suma de cada una de las áreas y de una variable de clasificación socioeconómica del establecimiento educativo. Dichas pruebas son aplicadas a los estudiantes que están culminando su educación media en todos los establecimientos educativos del país por el ICFES (Instituto Colombiano para la Evaluación de la Educación), entidad encargada de realizar las mediciones de calidad en toda la trayectoria educativa de los estudiantes. Dicho examen es un requisito para el acceso a educación superior, y es a nivel país el indicador más importante de desempeño escolar.

Por último, SIMAT, es el sistema de gestión de la matrícula de los estudiantes de los colegios oficiales del país. Dicho sistema permite la organización y el control de la información referente a la matrícula de los estudiantes, convirtiéndose en una fuente de información confiable que facilita la sistematización, consolidación y análisis por parte de los tomadores de decisiones de políticas educativas en el país. (MinEducación, 2021) En este caso de estudio, dicha fuente de información es importante, porque permite realizar la caracterización de las sedes educativas desde el aspecto del número de estudiantes matriculados.

Una vez se tienen las tres bases de datos con las que se realizará la estimación se procede a realizar la consolidación de esta. En este punto, es importante resaltar que, debido al tipo de recolección de datos en el formulario de caracterización de necesidades de conectividad, a pesar de contar con un número mayor de sedes educativas que diligenciaron dicho formulario (26.912) con respecto a las sedes educativas que presentaron las pruebas Saber 11 Calendario A (11.241), por una parte no se cuenta en la totalidad de dichas sedes educativas con educación media, y por otra parte no todas las sedes educativas que presentaron las pruebas Saber 11 en el año 2021 diligenciaron el formulario de

caracterización de necesidades de conectividad, por lo que al hacer el cruce entre ambas bases de datos se cuenta con un total de 5.583 sedes educativas. Finalmente, este número se reduce a 5.440 sedes educativas, cuando se hace el cruce con la base de datos del SIMAT, debido a que la plataforma que soporta dicho sistema se actualiza constantemente, y al momento de realizar la consulta para calcular la variable de matrícula por sede educativa, algunas sedes no contaban con información.

Dado que la selección de las variables utilizadas para la estimación depende del contexto educativo del país, es importante realizar un análisis del desempeño de las sedes educativas en las pruebas Saber 11 de acuerdo con las diferentes variables que desde nuestra base de datos podrían incidir en dicho desempeño.

En primer lugar, como destacan Viana & Pinto (2018) en un estudio del desempeño de los colegios en las pruebas Saber 11 2016 para el departamento de Santander, los colegios rurales tienden a ser más ineficientes que los urbanos, y dicha ineficiencia es más notable si se tienen en cuenta los factores atribuibles al colegio. Evidencia de esto, es que como muestra la tabla 1, las sedes educativas ubicadas en área urbana presentan en promedio un puntaje global de 21.1 puntos mayor a las sedes ubicadas en área rural y a su vez presentan en promedio un acceso a internet mayor.

Tabla 1

Sedes educativas diferenciadas por área de ubicación

Área	Número de sedes educativas	Puntaje global promedio	Acceso a internet
Rural	2804	219.31	44.15%
Urbana	2636	240.41	79.62%
Total	5440	229.53	61.34%

Fuente: Conexión total (2021) – ICFES (2021) – SIMAT (2021) Cálculos propios

De la misma manera y en concordancia con lo que menciona Fernández (2017), la tabla 2 muestra evidencia de una relación de las variables que componen el índice socioeconómico de las sedes educativas con el desempeño en las pruebas Saber 11 y con el acceso a internet. En general, las sedes educativas que cuentan con un nivel socioeconómico mayor muestran un mayor puntaje global promedio y de la misma manera presentan una prevalencia en el acceso a internet.

Tabla 2

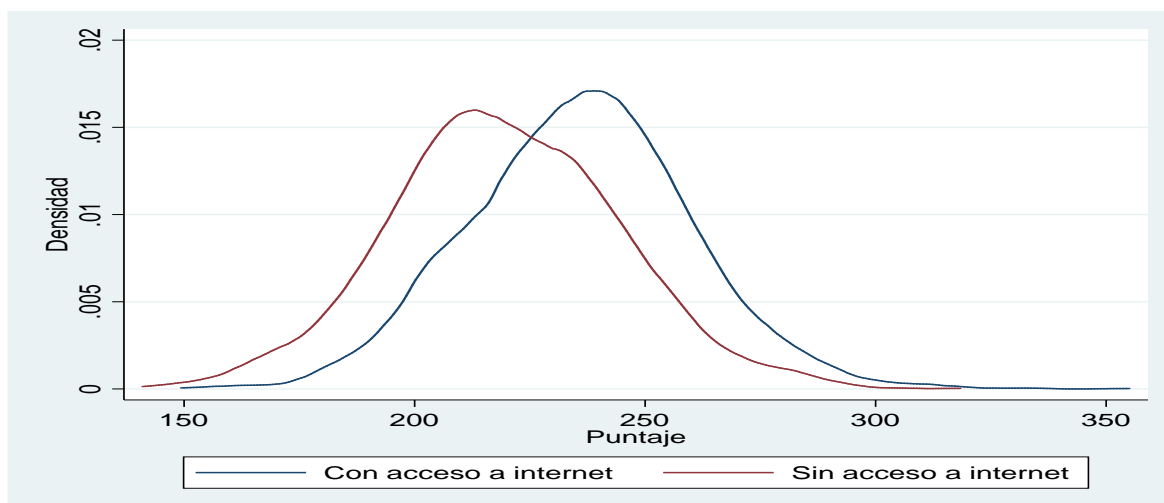
Sedes educativas diferenciadas por NSE del establecimiento

NSE	Número de sedes educativas	Puntaje global promedio	Acceso a internet
1	1079	214.273	39.01%
2	3595	227.730	62.33%
3	766	259.5301	88.12%
4	0	N/A	N/A
Total	5440	229.53	61.34%

Fuente: Conexión total (2021) – ICFES (2021) – SIMAT (2021) Cálculos propios

Además, como se puede observar en la figura 1, tomando como referencia el promedio del puntaje global, hay una diferencia en la curva de densidad de dicha variable cuando se controla por las sedes educativas que cuentan con acceso a internet ya sea por medio de red inalámbrica o cableada frente a aquellas sedes educativas que no cuentan con dicho acceso. En general, las sedes educativas con acceso a internet presentan un mejor desempeño en las pruebas Saber 11. Sin embargo, no necesariamente la diferencia que se presenta entre ambos grupos de estudio se deba únicamente al acceso a recursos tecnológicos, dado que puede haber otras variables no concebidas aún, que tienen relevancia en la estimación.

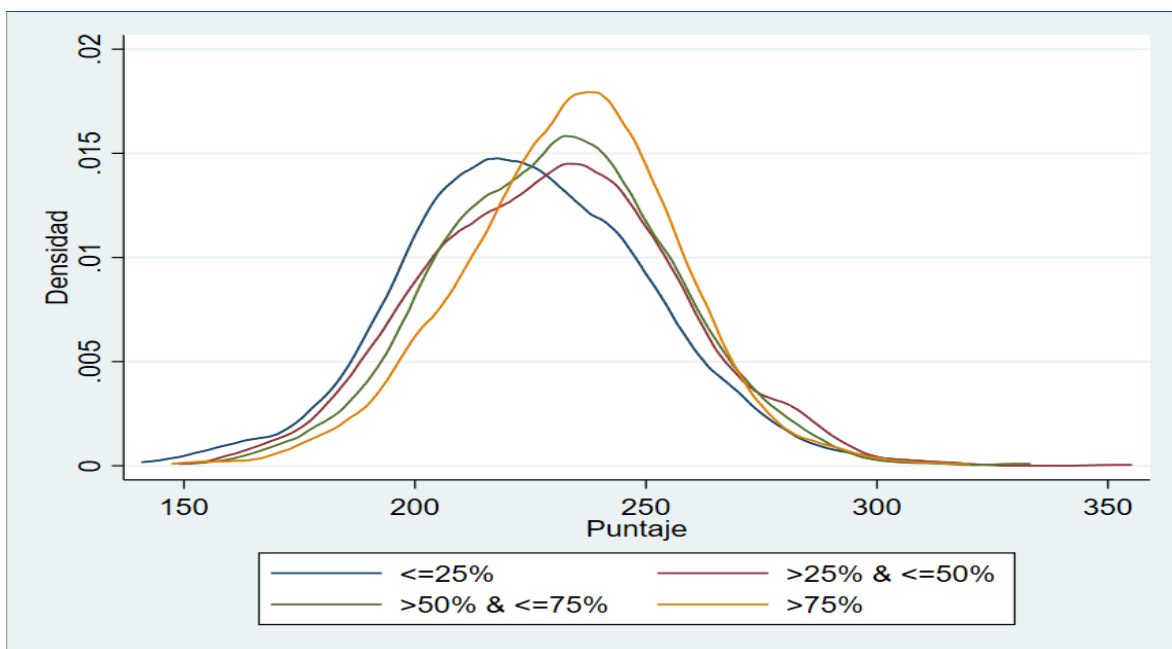
Figura 1. Densidad del puntaje promedio global en pruebas Saber 11 para sedes educativas con y sin acceso a internet



Fuente: Elaboración propia (2022)

Se decidió crear una variable que mostrara el número de recursos tecnológicos entendidos como la suma de computadores de escritorio, portátiles y tabletas en relación con la matrícula de cada sede educativa. La figura 2 muestra la distribución en el puntaje global de las sedes educativas, categorizando de acuerdo con la distribución porcentual de la variable creada. Las sedes educativas con menos recursos tecnológicos por estudiante muestran una tendencia a presentar menor rendimiento en las pruebas Saber 11.

Figura 2. Densidad del puntaje promedio global en pruebas Saber 11 para sedes educativas de acuerdo con el porcentaje en la distribución de la variable de equipos de cómputo por estudiante



Fuente: Elaboración propia (2022)

Por último, con el fin de caracterizar a las sedes educativas de una manera más completa de acuerdo con la base de datos con la que se cuenta, se deciden incluir las variables de matrícula, aulas para uso educativo, aulas con recursos tecnológicos y riesgos del establecimiento, la cual es una variable equivalente a la suma de los riesgos y limitaciones en términos de infraestructura, seguridad, sociales, naturales y de otra naturaleza bajo los cuales la sede educativa está expuesta.

En la tabla 3 se muestran las variables utilizadas para realizar las estimaciones con su tipo de variable y su fuente.

Tabla 3
Descripción de las variables

Variable	Descripción	Tipo de variable	Fuente
Puntaje global	Promedio del puntaje global por sede educativa	Continua	ICFES
Puntaje Matemáticas	Promedio del puntaje en el área de matemáticas por sede educativa	Continua	ICFES
Puntaje Lectura crítica	Promedio del puntaje en el área de lectura crítica por sede educativa	Continua	ICFES
Puntaje Naturales	Promedio del puntaje en el área de ciencias naturales por sede educativa	Continua	ICFES
Puntaje Sociales	Promedio del puntaje en el área de ciencias sociales por sede educativa	Continua	ICFES
Puntaje inglés	Promedio del puntaje en el área de inglés por sede educativa	Continua	ICFES
NSE Establecimiento	Nivel Socioeconómico del establecimiento	Categórica (1-4)	ICFES
Área	Área de ubicación de la sede educativa	Dicotómica (1 urbana, 0 rural)	Conexión total
Matrícula	Número de estudiantes matriculados reportados	Continua	SIMAT
Computadores de escritorio	Número de computadores de escritorio con los que cuenta la sede educativa	Continua	Conexión total
Computadores portátiles	Número de computadores portátiles con los que cuenta la sede educativa	Continua	Conexión total
Tabletas	Número de tabletas con los que cuenta la sede educativa	Continua	Conexión total
Aulas uso educativo	Número de aulas para el uso educativo con las que cuenta la sede educativa	Continua	Conexión total
Aulas recursos tecnológicos	Número de aulas con recursos tecnológicos con las que cuenta la sede educativa	Continua	Conexión total
Riesgos establecimiento	Número de tipo de riesgos y limitaciones naturales, sociales, de infraestructura, de seguridad, y de otra naturaleza a los que la sede está expuesta	Continua	Conexión total
Internet WLAN	¿Cuenta la sede educativa con internet inalámbrico?	Dicotómica (1 sí, 0 no)	Conexión total)
Internet LAN	¿Cuenta la sede educativa con internet alámbrico?	Dicotómica (1 sí, 0 no)	Conexión total)
Internet	¿Cuenta la sede educativa con acceso a internet por medio de cualquiera de los dos tipos de red?	Dicotómica (1 sí, 0 no)	Conexión total – Cálculo a partir de las variables Internet LAN e internet WLAN

Fuente: Conexión total (2021) – ICFES (2021) – SIMAT (2021) Elaboración propia

Como se observa en la tabla 4, el promedio a nivel general del puntaje global para las sedes educativas participantes de este caso de estudio es de 229,54, los promedios de los puntajes para cada área del conocimiento son cercanos, donde lectura crítica con 48,93 es el área en la que mejor muestran desempeño y ciencias sociales y ciudadanas la peor con 43,41. Las sedes educativas por su carácter oficial presentan un nivel socioeconómico bajo si se tiene en cuenta que dicha variable tiene un rango entre 1 y 4, siendo uno muy bajo y 4 muy alto. En promedio una sede educativa cuenta con 669 alumnos matriculados y a pesar de contar con una desviación estándar tan alta, la mediana no está tan alejada de la media. Con respecto al inventario de equipos tecnológicos con los que cuentan las sedes educativas, se encuentra que en promedio una sede educativa cuenta con 14 computadores de escritorio, 50 computadores portátiles y 73 tabletas. En promedio una sede educativa cuenta con 15 aulas para uso educativo, 2 con recursos tecnológicos y está expuesta a 3 diferentes tipos de riesgos. Con excepción a la variable de número de tipo de riesgos, todas las variables presentan un valor mayor en la media cuando se cuenta con acceso a internet a cuando no se tiene en cuenta la diferenciación por el acceso a internet y a cuando no se cuenta con acceso a internet.

Tabla 4

Estadística descriptiva

Variable	General		Acceso a internet		Sin acceso a internet	
	Media	D.E	Media	D.E	Media	D.E
Puntaje global	229.53	25.72	235.85	24.16	219.51	24.94
Puntaje Matemáticas	46.43	5.87	47.76	5.40	44.33	5.96
Puntaje Lectura crítica	48.92	5.22	50.25	4.82	46.81	5.14
Puntaje Naturales	45.42	4.75	46.48	4.45	43.75	4.74
Puntaje Sociales	43.41	5.50	44.70	5.17	41.37	5.41
Puntaje inglés	44.24	5.59	45.64	5.51	41.98	4.95
Matrícula	669.13	623.66	824.00	679.32	423.39	418.98
Computadores de escritorio	14.28	37.60	18.41	28.64	7.72	47.81
Computadores portátiles	50.47	98.99	61.40	116.86	33.14	56.50
Tabletas	73.18	124.36	88.86	139.55	48.32	90.00
Aulas uso educativo	15.59	12.07	18.08	13.18	11.65	8.74
Aulas recursos tecnológicos	1.99	2.54	2.41	2.91	1.33	1.60
Riesgos establecimiento	3.43	2.21	3.18	2.17	3.83	2.22
NSE	1.94	.57	2.07	.56	1.73	.53

Fuente: Conexión total (2021) – ICFES (2021) – SIMAT (2021) Cálculos propios

b. Aplicación de la metodología

Se realizan dos estimaciones en las que se toman como variables de resultado, los promedios tanto del puntaje global en las pruebas Saber 11 y los puntajes para cada una de las asignaturas. De la misma manera, con el fin de tener una caracterización completa de las sedes educativas, se incluyen como variables de control, la variable de nivel socioeconómico del establecimiento, matrícula, área, número de aulas tanto de recursos tecnológicos como de uso educativo, número de recursos tecnológicos, entendiéndolos como computadores de escritorio, portátiles y tabletas. Finalmente se eligen como variables de tratamiento para la primera estimación, las variables de acceso a internet, ya sea por medio de red LAN, WLAN o por cualquiera de estas y en el caso de la segunda estimación, una variable de binarización que indica si una sede educativa está ubicada o no en el cuartil inferior aluciente a la distribución de la variable creada equivalente al número de recursos tecnológicos en relación con la matrícula de la sede educativa. Dado que en este último caso se está tomando como variable de tratamiento una transformación de una de las covariables incluida en la primera estimación y adicional a esto, la covariable de matrícula tiene correlación directa con la variable de tratamiento, se decide eliminar dichas variables.

Antes del cálculo del PSM y del ATT, de acuerdo con Powell et al (2019), es importante validar el balance del modelo. Para validar el balance del modelo, es comúnmente usado el parámetro de diferencia medias estandarizada o Standardized Mean Difference (SMD) en inglés, introducido por Cohen (1988). Dicho parámetro se calcula para cada una de las covariables elegidas en el modelo teniendo en cuenta la media y la varianza (o en su defecto la desviación estándar elevada al cuadrado) para el grupo de tratamiento y el grupo de control en cada una de dichas covariables.

En el caso de las covariables continuas, la SMD se calcula de la siguiente manera:

$$SMD = \frac{\mu_{G1} - \mu_{G2}}{\sqrt{\frac{S^2_{G1} + S^2_{G2}}{2}}} \quad (6)$$

Donde μ es la media para cada covariable diferenciada por G1 el cual es el grupo de tratamiento y G2 entendido como el grupo de control. Además, S es la varianza para la respectiva covariable diferenciada por grupo de tratamiento y control.

Para las variables dicotómicas, el cálculo de la SMD es el siguiente:

$$SMD = \frac{\rho_1 - \rho_2}{\sqrt{\frac{\rho_1(1 - \rho_1) + \rho_2(1 - \rho_2)}{2}}} \quad (7)$$

Donde ρ es la prevalencia de la covariable diferenciada por grupo de control y tratamiento.

Tabla 5

SMD antes y después del emparejamiento - Internet

Variable	SMD antes	SMD después
NSE Establecimiento	.6279	.0048
Área	.8118	.0110
Matrícula	.7098	.0062
Computadores de escritorio	.2712	.0028
Computadores portátiles	.3077	.0123
Tabletas	.3452	-.0072
Aulas uso educativo	.5746	-.0093
Aulas recursos tecnológicos	.4577	-.0078
Energía	.2361	.0185
Riesgos establecimiento	-.2920	.0281

Fuente: Conexión total (2021) – ICFES (2021) – SIMAT (2021) Cálculos propios

Tabla 6

SMD antes y después del emparejamiento – Recursos tecnológicos

Variable	SMD antes	SMD después
NSE Establecimiento	.0468	-.0066
Área	.1711	-.0266
Aulas uso educativo	.0694	.0579
Aulas recursos tecnológicos	-.2086	-.0065
Energía	-.0565	.0142
Riesgos establecimiento	.0742	-.0127

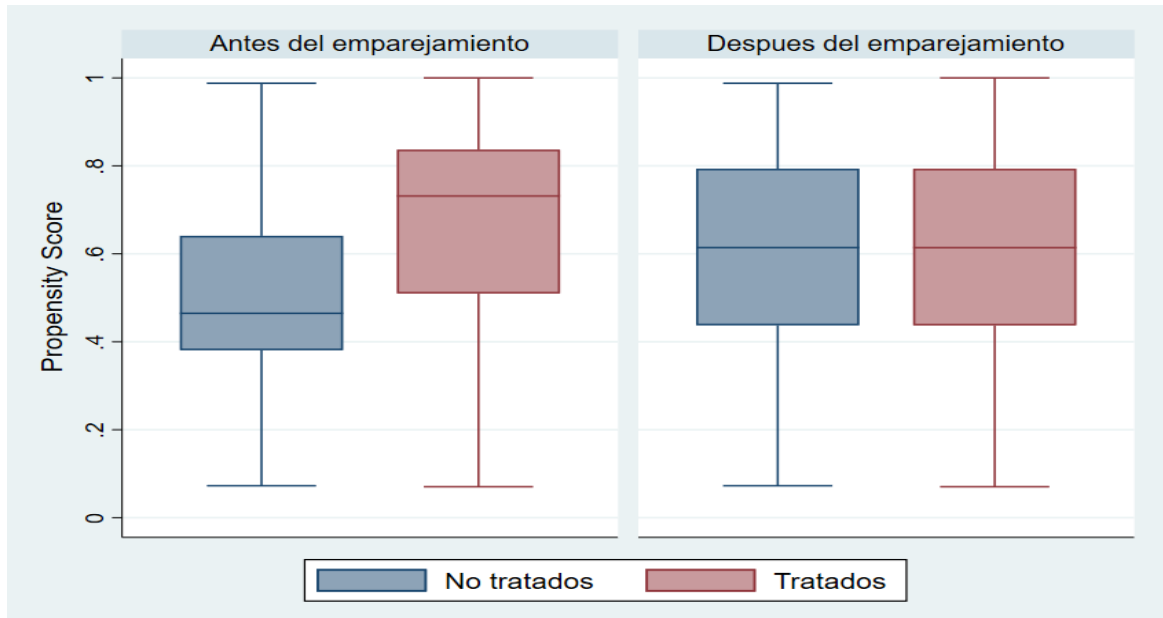
Fuente: Conexión total (2021) – ICFES (2021) – SIMAT (2021) Cálculos propios

Según Zhang et al (2019), un valor de SMD superior a 0.1 puede ser considerado un signo de imbalance en el modelo. Como se ve en la tabla 5 y la tabla 6, los modelos corrigen el imbalance en todas las covariables respectivas.

Adicional a esto, Garrido et al (2014) menciona que la condición de soporte común es subjetivamente evaluada mediante una gráfica de distribución de los propensity score para el grupo de tratamiento y para el grupo de control. Para garantizar dicha condición debería existir una distribución similar de dichos propensity score entre los dos grupos de comparación.

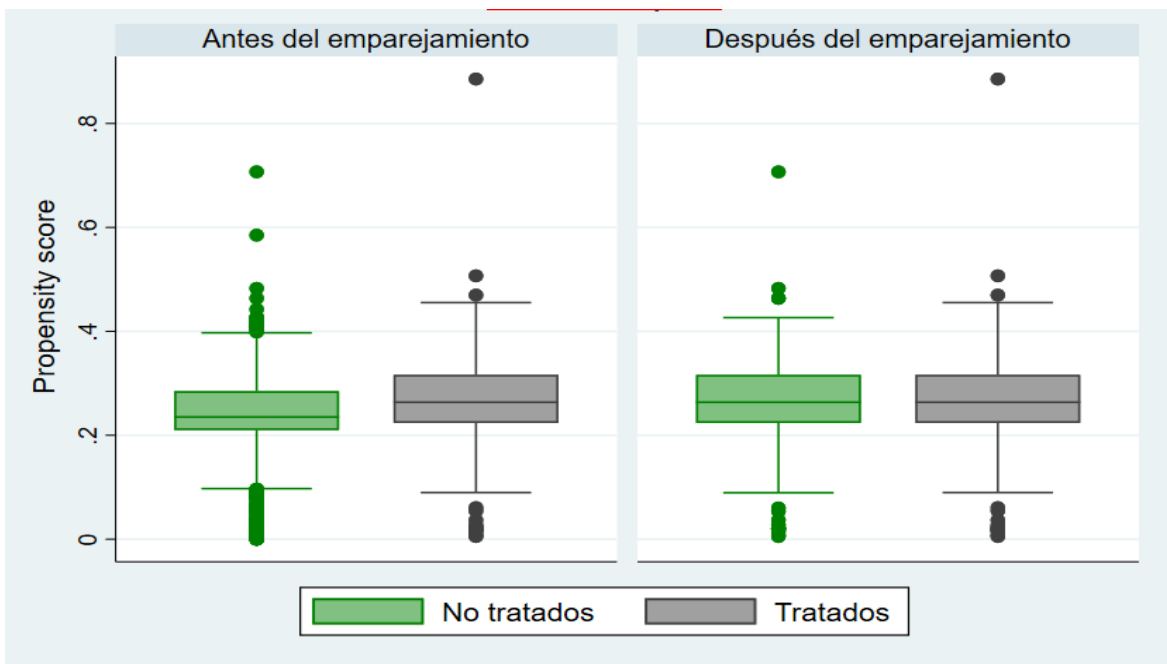
De acuerdo con la figura 3 y a la figura 4, a pesar de existir una diferencia en el gráfico de caja de los dos grupos, una vez se realizan los emparejamientos, dichas diferencias se corrigen y las distribuciones de los dos grupos es similar, por lo que se muestra evidencia a favor del cumplimiento de la condición de soporte común.

Figura 3. Distribución de los propensity score antes y después del emparejamiento -Internet



Fuente: Elaboración propia (2022)

Figura 4. Distribución de los propensity score antes y después del emparejamiento -Recursos tecnológicos



Fuente: Elaboración propia (2022)

IV. Resultados

Tabla 7

Average Treatment Effect on the Treated (ATT) Internet LAN o WLAN

Variable de resultado	ATT	Errores estándar	Z
Puntaje global	6.11***	.97	6.27
Lectura crítica	1.17***	.17	6.87
Matemáticas	1.14***	.22	5.12
Ciencias naturales	1.11***	.19	5.70
Ciencias sociales y ciudadanas	1.34***	.21	6.32
Inglés	1.44***	.16	8.72

*p<0.1, **p<0.05, ***p<0.01

Fuente: Elaboración propia

Tabla 8

Average Treatment Effect on the Treated (ATT) Internet LAN

Variable de resultado	ATT	Errores estándar	Z
Puntaje global	3.16***	.90	3.50
Lectura crítica	.63***	.17	3.67
Matemáticas	.56***	.20	2.76
Ciencias naturales	.55***	.17	3.18
Ciencias sociales y ciudadanas	.73***	.19	3.80
Inglés	1.00***	.19	5.26

*p<0.1, **p<0.05, ***p<0.01

Fuente: Elaboración propia

Tabla 9

Average Treatment Effect on the Treated (ATT) Internet WLAN

Variable de resultado	ATT	Errores estándar	Z
Puntaje global	5.05***	.90	5.58
Lectura crítica	1.01***	.17	5.96
Matemáticas	.97***	.21	4.61
Ciencias naturales	.92***	.17	5.28
Ciencias sociales y ciudadanas	1.08***	.19	5.63
Inglés	1.12***	.18	5.95

*p<0.1, **p<0.05, ***p<0.01

Fuente: Elaboración propia

Como se evidencia en la tabla 7, en la tabla 8 y en la tabla 9, hay un coeficiente positivo en el efecto promedio sobre los tratados para las estimaciones donde se toma como tratamiento tanto el acceso a red LAN, como el acceso a red WLAN y como el acceso a red LAN o red WLAN. Dichos coeficientes varían entre 3.16 y 6.11, siendo mayor en el caso que el tratamiento es el acceso a uno de los dos tipos de redes. En el mismo sentido, tomando como variables de resultado el puntaje promedio para cada una de las áreas evaluadas, se encuentra una relación positiva entre el acceso por parte de las sedes educativas a internet por cualquiera de los tipos de red. En general, esta relación es más pronunciada para las áreas de ciencias sociales y ciudadanas e inglés, y deja de serlo para el área de ciencias naturales. Todos los coeficientes de la tabla 7, de la tabla 8 y de la tabla 9 son estadísticamente significativos al 99%.

Tabla 10

Average Treatment Effect on the Treated (ATT) Recursos tecnológicos por matrícula $\leq 25\%$

Variable de resultado	ATT	Errores estándar	Z
Puntaje global	-6.98***	.89	-7.79
Lectura crítica	-1.18***	.18	-6.54
Matemáticas	-1.70***	.20	-8.23
Ciencias naturales	-1.33***	.17	-7.80
Ciencias sociales y ciudadanas	-1.48***	.19	-7.72
Inglés	-1.04***	.20	-5.04

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Fuente: Elaboración propia

En la tabla 10 se encuentra una asociación negativa de -6.98 en el puntaje global y entre -1.04 y -1.70 por áreas de los resultados promedio en las pruebas Saber 11 2021, por el hecho de que una sede educativa se encuentre en el percentil 25 sobre la distribución de la variable de recursos tecnológicos en relación con la matrícula en comparación a las sedes educativas ubicadas en un percentil mayor al 25. Los resultados mencionados anteriormente dan indicios de una relación negativa entre el hecho de contar con menos recursos tecnológicos por número de estudiantes y el puntaje promedio en las pruebas Saber 11 2021 tanto para el puntaje global como por áreas. A diferencia de la estimación en la que se toma como tratamiento el acceso a internet, en este caso las áreas de ciencias matemáticas, ciencias sociales y ciencias naturales presentan un efecto negativo mayor en comparación a las áreas de inglés y lectura crítica. Los coeficientes son estadísticamente significativos al 99%.

V. Conclusiones

Según la evaluación del contexto educativo del país teniendo en cuenta las sedes educativas que hacen parte de la muestra utilizada en esta investigación, es evidente la brecha existente entre sedes educativas ubicadas en áreas rurales y áreas urbanas, no solo en el desempeño en las pruebas Saber 11, sino también en el acceso de estas sedes a recursos tecnológicos e internet. Esto sumado a la relación sobre ambas variables asociadas al nivel socioeconómico del establecimiento educativo, hace inminente el llamado a las autoridades encargadas de la política educativa del país a garantizar el acceso a la educación bajo condiciones de calidad y equidad, de tal manera que se permita que todos los participantes del sistema educativo tengan la posibilidad de adaptarse a una sociedad cambiante, demandante de personas capacitadas para afrontar los retos en términos de transformación digital.

Si bien no hay un acuerdo con respecto a la postura frente al impacto de los recursos tecnológicos e internet en los resultados de desempeño escolar, esta investigación encuentra una asociación positiva de poca magnitud que puede sugerir una etapa temprana en la implementación de las TICS en el sistema educativo del país, lo que también invita a las autoridades encargadas de la política educativa a generar estrategias que no exclusivamente se limiten a garantizar el acceso a dichas herramientas. Es innegable que las TICs por sí mismas tienen enormes beneficios en el aula de clases, pero no basta con el componente tecnológico si lo que se busca es una verdadera transformación de la realidad educativa y, por el contrario, es imperativa la necesidad de integrar dicho componente con los aspectos pedagógicos y didácticos de las instituciones teniendo en cuenta el contexto de la dinámica de aprendizaje (Restrepo, 2005).

Limitaciones importantes para tener en cuenta, están dadas en que no todas las sedes educativas oficiales del país diligenciaron el formulario de caracterización y además no todas las sedes que diligenciaron dicho formulario presentan pruebas saber ICFES, debido a que solo cuentan con educación básica, lo que puede suponer cierto sesgo de información. Es importante también establecer que debido a la limitación temporal de los datos (dado que solo se cuenta con información para el 2021) se puede realizar una errónea interpretación de los resultados. Por último, debido a que la metodología empleada puede presentar coeficientes sesgados dado que no hay evidencia de corrección del sesgo sobre variables no observadas, es erróneo hacer inferencia causal con los resultados de esta investigación.

VI. Referencias

- Ashrafi, R. and Murtaza, M. "Use and Impact of ICT on SMEs in Oman." *The Electronic Journal Information Systems Evaluation* 11(3) 2008, 125 – 138.
- Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate behavioral research*, 46(3), 399-424.
- Barrera-Osorio, F., & Linden, L. L. (2009). The use and misuse of computers in education: evidence from a randomized experiment in Colombia. *World Bank Policy Research Working Paper*, (4836).
- Beuermann, D. W., Cristia, J., Cueto, S., Malamud, O., & Cruz-Aguayo, Y. (2015). One laptop per child at home: Short-term impacts from a randomized experiment in Peru. *American Economic Journal: Applied Economics*, 7(2), 53-80.
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, 22(1), 31-72.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (No. 300.72 C6).
- Fernández García, S. (2017). Resultados en las Pruebas de Estado ICFES Saber 11-¿ qué características socioeconómicas de los estudiantes y particulares de los colegios pueden explicar las diferencias entre los resultados de colegios públicos y privados?.
- Flórez Wandurraga, J. J. (2021). ¿ Podría el acceso a un computador e internet en casa hacer la diferencia en el desempeño de los estudiantes en las pruebas de Estado-SABER 11°? Una aproximación desde el propensity score matching.
- Garrido, M. M., Kelley, A. S., Paris, J., Roza, K., Meier, D. E., Morrison, R. S., & Aldridge, M. D. (2014). Methods for constructing and assessing propensity scores. *Health services research*, 49(5), 1701-1720.
- Goolsbee, A., & Guryan, J. (2006). The impact of Internet subsidies in public schools. *The Review of Economics and Statistics*, 88(2), 336-347.
- Heinrich, C., Maffioli, A., & Vazquez, G. (2010). A primer for applying propensity-score matching. *Inter-American Development Bank*.
- Kaware, S. S., & Sain, S. K. (2015). ICT Application in Education: An Overview. *International Journal of Multidisciplinary Approach and Studies*, 2, 25-32.
- Mao, R. E., & Casas, L. S. (2004). La integración de las TIC en la educación: Apuntes para un modelo pedagógico pertinente The integration of ICTs in education: Notes for. *Rev. Ped*, 25, 74.

- Mineducación (2017) *Plan nacional decenal de educación 2016-2026 El camino hacia la calidad y equidad*. Recuperado de https://siteal.iiep.unesco.org/sites/default/files/sit_accion_files/siteal_colombia_0404.pdf
- Mineducación (2017, 13 febrero). *Programa Conexión Total*. Recuperado 12 de mayo de 2022 de <https://www.mineducacion.gov.co/portal/micrositios-institucionales/Conexion-Total/Programa-Conexion-Total/321608:Programa-Conexion-Total>
- Mineducación (2021) *Lineamiento técnico 2021 Programa conexión total*. Recuperado de: https://www.mineducacion.gov.co/1759/articles-321649_recurso_2.pdf
- Mineducacion (2021, 12 enero) *Sistema Integrado de Matrícula*. Recuperado 26 de abril de 2022, de <https://www.mineducacion.gov.co/portal/micrositios-institucionales/Sistemas-de-Informacion/168883:SIMAT>
- Mineducación. (2022, 24 enero). *Número de estudiantes promedio por Computador*. Recuperado 2 de marzo de 2022, de <https://www.mineducacion.gov.co/portal/micrositios-institucionales/Conexion-Total/Conexion-Total-en-Cifras/354999:Numero-de-estudiantes-promedio-por-Computador>
- MinTIC. (2014). *TIC y educación - MINTIC - Vive Digital*. Recuperado 2 de marzo de 2022, de <https://mintic.gov.co/portal/vivedigital/612/w3-article-19513.html>
- Pastrana Gil, J. H., & Saludes Avila, S. G. (s. f.). *El Papel Y La Relevancia Del Avance De La Tecnología Para El Desarrollo Humano* calameo.com. <https://es.calameo.com/read/00535832855fec4871300>
- Powell, M. G., Hull, D. M., & Beaujean, A. A. (2020). Propensity score matching for education data: Worked examples. *The Journal of Experimental Education*, 88(1), 145-164.
- Rodríguez Rosero, D. D., Ordoñez Ortega, R. E., & Hidalgo Villota, M. E. (2021). Academic Performance Determinants of High School Students in the Department of Nariño, Colombia. *Lecturas de Economía*, (94), 87-126.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Universidad EAFIT (2021) *Caracterización de las necesidades de conectividad de las sedes educativas incluido el contexto regional*. Recuperado de https://www.mineducacion.gov.co/1780/articles-406755_recurso_3.pdf
- Viana Barceló, R. A., & Pinto Prieto, H. M. (2018). Eficiencia de los estudiantes urbanos y rurales de Santander: “Saber 11” 2016. *Suma de Negocios*, 9(20), 111-119.
- Zea Restrepo, C. M., Trujillo Vargas, J. A., Atuesta Venegas, M. del R., & Foronda Velásquez, N. (2012). Características de los procesos de gestión en los contextos E-

Learning. *Revista Universidad EAFIT*, 41(140), 43–57. Recuperado de <https://publicaciones.eafit.edu.co/index.php/revista-universidad-eafit/article/view/82>

- [Zhang, Z., Kim, H. J., Lonjon, G., & Zhu, Y. \(2019\). Balance diagnostics after propensity score matching. *Annals of translational medicine*, 7\(1\).](#)