



UNIVERSIDAD  
DE PIURA

FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

**Externalidades entre colegios públicos y privados: una  
aproximación al caso peruano**

Tesis para optar el Título de  
Economista

**Mariagracia Galvez Arrascue  
Stephanie Kimie Katayama Loayza**

**Asesor(es):  
PhD. Gabriel Natividad**

**Lima, junio de 2022**



### **Agradecimientos**

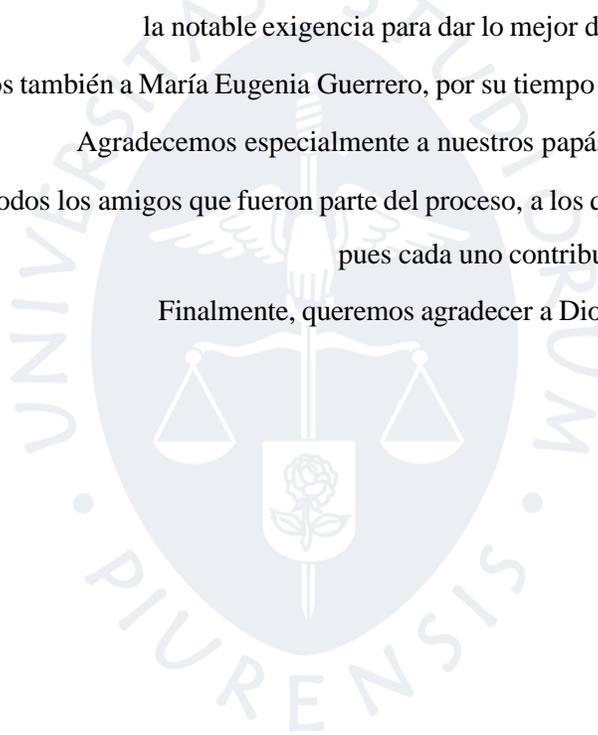
Agradecemos de manera especial a Gabriel Natividad, por el infalible apoyo, la claridad en cada consejo y la notable exigencia para dar lo mejor de nosotras en cada momento.

Agradecemos también a María Eugenia Guerrero, por su tiempo y sugerencias en el proyecto.

Agradecemos especialmente a nuestros papás, motor y motivo de la tesis.

Agradecemos también a todos los amigos que fueron parte del proceso, a los que estuvieron y a los que no: pues cada uno contribuyó a que esto salga adelante.

Finalmente, queremos agradecer a Dios, por haberlo hecho posible.





## Resumen

El presente estudio analiza una muestra de todos los colegios del Perú entre 2007 y 2016, considerando como variable independiente un índice del nivel de implementación de los laboratorios de ciencias de los colegios públicos y como dependiente, la cantidad de matriculados de un colegio cercano. Se muestra que la mejora de un laboratorio público genera un aumento del 0.28% de los matriculados en un colegio privado cercano, creciendo en más de 4% cuando dicho colegio privado es grande; y un decrecimiento del 0.37% de los matriculados de un colegio público cercano. Se atribuyen estos resultados a externalidades positivas de los colegios públicos a los privados cercanos, y a la competencia existente entre los públicos cercanos entre sí. Se analizan los posibles mecanismos subyacentes a estos resultados, presentando implicancias en políticas públicas al respecto.





## Tabla de contenido

|   |           |
|---|-----------|
| <b>Introducción.....</b>  | <b>11</b> |
| <b>1 Marco contextual</b>   | <b>17</b> |
| 1.1 Evidencia de la relevancia de la educación para las políticas en el Perú.....                 | 17        |
| 1.2 Literatura existente en torno al mercado educativo peruano .....                              | 17        |
| 1.3 Posible privatización por defecto en el Perú .....  | 18        |
| <b>2 Cuestión de investigación</b>  | <b>21</b> |
| <b>3 Base de datos</b>  | <b>27</b> |
| <b>4 Especificación</b>   | <b>31</b> |
| 4.1 Diseño empírico empleado.....   | 31        |
| 4.2 Robustez.....   | 33        |
| 4.3 Evaluación de la metodología .....  | 35        |
| <b>5 Resultados</b>   | <b>37</b> |
| 5.1 Externalidades positivas y competencia.....   | 37        |
| 5.1.1 Externalidades positivas de los colegios públicos a los privados.....                       | 37        |
| 5.1.2 Competencia entre colegios públicos .....   | 39        |
| 5.1.3 Diferencia de los resultados para varones y para mujeres .....                              | 41        |
| 5.2 Análisis de robustez.....   | 42        |
| 5.2.1 Resultados invertidos en función al tipo de gestión.....                                    | 42        |
| 5.2.2 Interacciones .....   | 43        |
| 5.3 Evaluación de la metodología .....  | 43        |
| <b>6 Implicancias para políticas públicas</b>   | <b>45</b> |
| 6.1 Justificación de la relevancia de los resultados para las políticas públicas en el Perú ..... | 45        |

6.2 Rol clave de la focalización.....45

6.3 Importancia de la mejora en la calidad de la información relacionada al sector educativo .....47

**Conclusiones.....49**

**Lista de Referencias ..... 51**

**Figuras.....55**

**Tablas ..... 57**

**Apéndice.....63**



## Lista de figuras

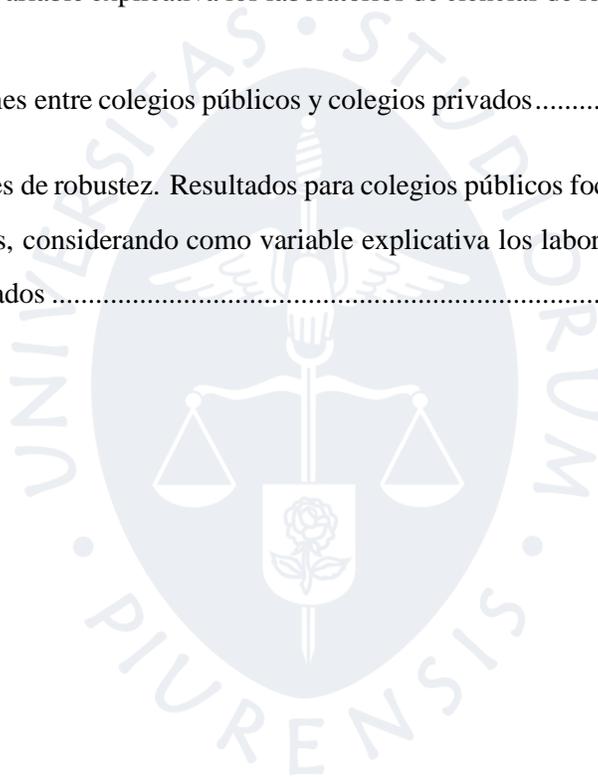
- Figura 1 Base de datos obtenida del INEI, elaboración propia. Se muestra el valor del gasto público nacional en educación como porcentaje del PBI desde 1994 hasta 2017 .....55
- Figura 2 Base de datos obtenida de ESCALE, elaboración propia. Se muestra el promedio de las notas para los respectivos colegios en los respectivos cursos de la Evaluación Censal de Estudiantes. ....55
- Figura 3 Base de datos obtenida de ESCALE, elaboración propia. Se muestra el valor promedio de la cantidad de alumnos matriculados en todos los colegios públicos y privados del Perú.....56





## Lista de tablas

|         |  |    |
|---------|--|----|
| Tabla 1 | Estadísticos sumarios de las variables.....  | 57 |
| Tabla 2 | Estadísticos sumarios de los emparejamientos.....  | 58 |
| Tabla 3 | Resultados para colegios privados focales emparejados con colegios públicos, considerando como variable explicativa los laboratorios de ciencias de los colegios públicos .....                          | 58 |
| Tabla 4 | Resultados para colegios públicos focales emparejados con colegios públicos, considerando como variable explicativa los laboratorios de ciencias de los colegios públicos .....                          | 59 |
| Tabla 5 | Interacciones entre colegios públicos y colegios privados.....   | 60 |
| Tabla 6 | Regresiones de robustez. Resultados para colegios públicos focales emparejados con colegios privados, considerando como variable explicativa los laboratorios de ciencias de los colegios privados ..... | 61 |





## Introducción

El presente trabajo de investigación estudia y muestra la existencia de externalidades generadas por los colegios públicos en el Perú, y propiciada por la cercanía geográfica con otros colegios públicos y privados. Aprovechando una variable que mide el nivel de implementación de los laboratorios de ciencias de los colegios públicos se analizan los mecanismos mediante los cuales la cantidad de matriculados de los colegios cercanos se ve afectada por la mejora de la infraestructura de los colegios públicos. Gracias a un detallado emparejamiento de los colegios cercanos y a una sólida metodología econométrica se logra mostrar la significancia estadística y la magnitud de dichas externalidades, precisando además ciertas casuísticas en las que son más relevantes.

Este análisis parte de la importancia de la educación para el país, sumada a la incertidumbre con respecto a los mecanismos subyacentes a los fenómenos educativos actuales. Desde hace más de 10 años nuestro Estado le destina un porcentaje cada vez más alto de inversión monetaria al sector educativo, pero el Perú se mantiene como el país que menos invierte en educación en Sudamérica (BID, 2017). Por otro lado, existe un fenómeno que viene captando el interés de los investigadores peruanos (cfr. Cuenca (2013), Fontdevila et al. (2018), Málaga and Velásquez (2020), entre otros): el hecho de que los colegios públicos vienen mejorando su rendimiento (hasta el punto de alcanzar o superar al sector privado), pero que cada vez tenemos más alumnos matriculados en colegios privados que en colegios públicos. Un estudio que profundice en la relación existente entre los colegios pertenecientes a estos dos tipos de gestión es de relevancia.

La literatura económica a nivel global, en lo que respecta a temas educativos, suele estar ligada a sueldos futuros y pobreza, notas, o variables difíciles de obtener (como las encontradas en Jackson et al. (2015) o Chiappori et al. (2017)). Centrándonos en el caso peruano, la mayoría de las investigaciones gira en torno al impacto del canon y regalías mineras (por ejemplo, Manrique et al. (2016)), o la inversión estatal en educación (por ejemplo, Beuermann et al. (2015)). En definitiva, “no se ha hecho evaluaciones que permitan analizar la efectividad de las inversiones” (Rodríguez-Sosa, 2013), tal vez debido a que este tipo de información es “difícil de estimar dadas las características de la información disponible” (Guadalupe et al., 2017). Este trabajo procura mostrar resultados relevantes para la literatura económica peruana, aprovechando las bases de datos disponibles y manteniendo una rigurosa metodología.

Empleando la base de datos del Censo Escolar del Perú elaborado por el Ministerio de Educación

del Perú para el periodo de años entre el 2007 y 2016 así como de la Evaluación Censal de Estudiantes (ECE) del 2016 y 2018, logramos contar con tres variables clave para el análisis.

Utilizamos como variable independiente un índice del nivel de implementación del laboratorio de ciencias de un colegio público. El índice toma el valor de 0 si el colegio no cuenta con un laboratorio de ciencias, el valor de 1 si cuenta con un laboratorio de ciencias, el valor de 2 si dicho laboratorio está equipado para realizar experimentos correspondientes al Ciclo 6 de la Educación Básica Regular (i.e., primero y segundo de secundaria) y el valor de 3 si dicho laboratorio está equipado para realizar experimentos correspondientes al Ciclo 7 de la Educación Básica Regular (i.e., tercero, cuarto y quinto de secundaria). El índice se armó de manera que las categorías 1, 2 y 3 sean mutuamente excluyentes. Para efectos de todas las interpretaciones futuras emplearemos una nomenclatura que hace referencia al índice del laboratorio de ciencias, asignándole un nivel de 1, 2 o 3 según sea el caso.

Contamos, por otro lado, con dos variables dependientes: la cantidad (en logaritmos) de alumnos matriculados en el colegio focal y una dicotómica que se activa cuando el colegio recibe más alumnos en primer grado que el año anterior. Ambas variables se distinguen según sexo, y el colegio focal puede ser privado o público según interese al análisis correspondiente.

Al disponer de coordenadas geográficas para cada colegio pudimos realizar emparejamientos mediante georreferenciación de los colegios públicos con los colegios tanto públicos como privados que se encuentren dentro de un radio de distancia de 1000 metros; a estos últimos los llamamos colegios focales. Agregamos a esta metodología dos tipos de efectos fijos (a nivel de colegio focal y a nivel de año), además de *clusters* por distrito del colegio focal. Esto nos permitió controlar por todo lo inherente al período temporal respectivo y a las características propias de los colegios focales, llegando a un análisis más fino sin necesitar una mayor cantidad de variables.

Consideramos también como parte de nuestra metodología dos interacciones que nos permitieron caracterizar mejor a los colegios que daban resultados estadísticamente más significativos. La primera variable empleada fue la de colegios grandes: una dicotómica que se activa si el número de alumnos matriculados en el colegio focal es mayor a la mediana del total de matriculados en los colegios focales para cada año. La segunda variable fue la de colegios con desempeño alto: otra dicotómica que se activa si el promedio de notas en la ECE para matemática y lengua a nivel de colegio focal es satisfactorio en el

año correspondiente. Una vez más, el colegio focal puede ser privado o público según interese al análisis correspondiente.

Por último, para reafirmar la validez de la metodología empleada, realizamos un análisis de robustez mediante una especificación que invertía el tipo de gestión de los resultados principales: teniendo como variable dependiente los matriculados del colegio público y como variable independiente el nivel de laboratorio de los colegios privados focales. Esto nos permitió evaluar la solidez de la metodología y de los resultados, analizando que mantengan un buen comportamiento a pesar de que las condiciones del mercado se vean alteradas (incluso de una forma más inverosímil). Cabe resaltar que la endogeneidad no fue considerada como un riesgo para el modelo presentado, debido a la naturaleza de las variables implicadas (por lógica temporal, la mejora en infraestructura fue implementada antes de que se vean los resultados en el Censo Escolar, y por lo tanto no puede preceder - ni en tiempo ni en causalidad - a la matrícula del colegio focal).

Debido a la desventaja de contar con bases de datos (o información) limitada nuestra metodología carece de variables de control; lo que nos dificulta el alcanzar un entendimiento exhaustivo de la relación entre los colegios públicos y privados. No obstante, rescatamos que con la información disponible y con la metodología implementada presentamos resultados interesantes y robustos que sustentan la existencia de dicha relación; y con la implementación de supuestos lógicos alcanzamos una aproximación de la descripción de la relación para profundizar en estudios futuros.

Los primeros resultados hacen referencia a las externalidades de los colegios públicos hacia los privados focales. Una mejora de las instalaciones del laboratorio del colegio público se relaciona con un aumento del 0.28% de los alumnos varones matriculados en el colegio focal privado respectivo, con un intervalo de confianza del 99%. Por otro lado, dicha mejora en los laboratorios se relaciona con un ligero pero estadísticamente significativo aumento de 0.19% en la probabilidad de que ingresen alumnos varones a primer grado al colegio, con un intervalo de confianza del 99%. Además, para el caso de matriculados varones, que el colegio sea grande se relaciona con un aumento de los matriculados de un 4.3% cuando el laboratorio es de nivel 1, en 7.3% si el laboratorio es de nivel 2 y en 9.1% cuando el laboratorio es de nivel 3. Todos estos resultados cuentan con un nivel de significancia estadística del 99%.

Al tener los efectos fijos y los *clusters*, se puede afirmar que los coeficientes están libres de todo lo inherente al colegio privado en cuestión y su distrito correspondiente; por lo que existe un efecto indirecto de

la inversión pública en los colegios privados (recordando lo mencionado en párrafos anteriores con respecto a la endogeneidad). Consideramos que este efecto indirecto es una externalidad positiva por parte de los colegios públicos hacia los privados focales: si bien no era el objetivo de sus acciones, es un efecto de las mismas.

Presentamos tres canales fundamentales que podrían darse en simultáneo y ser los causantes de la existencia de estas externalidades. En primer lugar, los colegios públicos tienen una vinculación más cercana a la población de la vecindad correspondiente con potencial para ingresar al mercado educativo, por lo que una mejora en infraestructura educativa pública se puede correlacionar con una mayor motivación por estudiar. A su vez, la privatización por defecto (aquel fenómeno donde un consumidor prefiere "a ciegas" un proveedor privado por encima de un proveedor público) daría pie a que, en base a una creencia implícita en la mentalidad peruana, se escojan los colegios privados. Finalmente, que el colegio sea grande, refuerza la externalidad encauzada por la privatización por defecto. Esto es intuitivo, y acorde a la interpretación previa: se buscan colegios, se prefieren los privados, pero además se prefieren los más notorios (en este caso, los "grandes"; pues evidencian un elemento de señalización de la calidad educativa más potente).

En conclusión, las externalidades positivas halladas podrían interpretarse como el resultado de una generalización de la búsqueda de educación canalizada a través de la privatización por defecto, y reforzada por la importancia (en términos de cantidad de alumnos) que tenga el colegio privado en la vecindad respectiva.

Los siguientes resultados evidencian las externalidades de los colegios públicos hacia otros colegios públicos cercanos. Una mejora sustancial de las instalaciones en laboratorios está vinculada a un decremento del 0.37% de los matriculados varones del colegio público focal. Estos son los únicos resultados estadísticamente significativos con un intervalo de confianza del 99%. Por otro lado, que el colegio sea grande hace que su número de matriculados aumente en un 1.18% cuando el laboratorio es de nivel 1 (con una significancia estadística del 95%), que disminuya en 3% cuando dicho laboratorio es de nivel 2 (con una significancia estadística del 90%) y que disminuya en 7% cuando dicho laboratorio es de nivel 3 (con una significancia estadística del 99%).

Se puede apreciar que los colegios públicos reaccionan con una menor sensibilidad a un análisis más fino del mercado educativo: los coeficientes tienen menor magnitud y significancia estadística que en el análisis con el sector privado en lo que respecta a las interacciones. A pesar de que contamos con resultados

en cierta medida estadísticamente significativos, consideramos que no tenemos información suficiente para realizar interpretaciones al respecto, que a su vez aporten a la cuestión de la presente investigación. Para efectos de este trabajo, vemos que el análisis comparativo mencionado es de suficiente utilidad.

Otra interpretación directa que se deriva de los resultados de este trabajo es que las decisiones no son similares cuando se trata de hijos varones y de hijas mujeres, en lo que respecta a la matrícula escolar. Una mejora en el laboratorio del colegio público cercano lleva a un decrecimiento del 0.18% de alumnas matriculadas en el colegio privado focal con un nivel de significancia estadística del 90%; así como a un resultado sin significancia estadística tanto para la dicotómica de nuevas alumnas como para el caso de los emparejamientos entre colegios públicos.

Este es un tema en el que consideramos que es importante profundizar con investigaciones futuras. Los hallazgos presentados son un punto de arranque que esperamos motive al respecto.

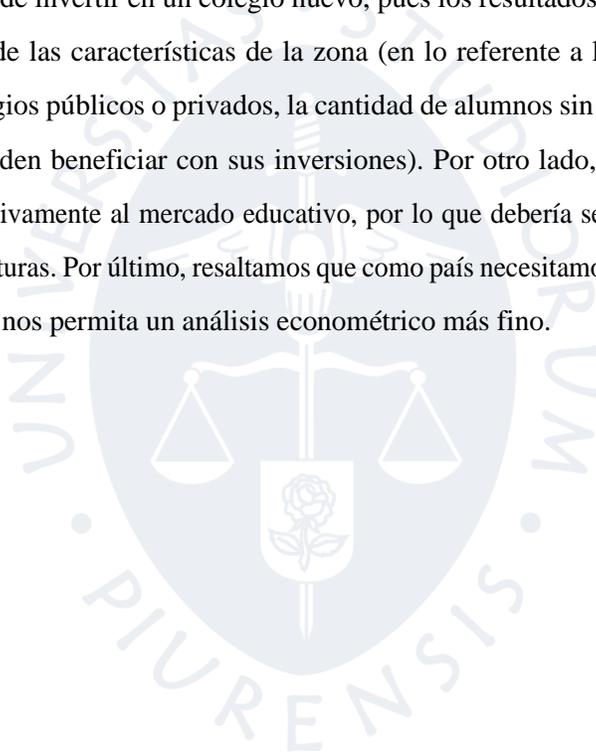
Finalmente, en lo que respecta al análisis de robustez, los resultados mostraron que una mejora sustancial en el laboratorio del colegio privado está vinculada a un decrecimiento del 0.08% de los matriculados varones del colegio público focal, con un nivel de significancia estadística del 99%. Por otra parte, los resultados para la variable dependiente de la dicotómica de nuevos alumnos muestran un decrecimiento de 0.13% para los varones y 0.12% para las mujeres con un nivel de significancia estadística del 95% en ambos casos.

Consideramos que estos resultados hablan a favor de la existencia de externalidades entre los colegios públicos y privados del Perú. Si bien la magnitud de los estimadores es menor que en el caso de colegios focales privados, el hecho de que existan resultados estadísticamente significativos para colegios focales públicos evidencia que el poder adquisitivo de las familias no es el único factor que influencia este aspecto del mercado educativo.

Un último aspecto que nos llamó la atención al realizar este trabajo fue que al incluir los efectos fijos a nivel de colegio focal y usar como variable dependiente el total de alumnos de dicho colegio focal, el  $R^2$  de las especificaciones era bastante elevado. Pero al evaluar las opciones en la metodología (retirando los efectos fijos) ponderamos como conveniente mantener dichos efectos fijos y centrarnos en lo que aporten los presentes resultados. Al comparar los estadísticos resultantes de la estimación que no considera los efectos fijos notamos grandes variaciones en signo y magnitud, que a su vez no tenían una interpretación plausible,

y por la tanto eran significativamente menos informativos. Es así que concluimos que la mejor forma de aprovechar los hallazgos, al no contar con un abanico más amplio de variables, es mantener la metodología aplicada matizando las interpretaciones respectivas: los estadísticos resultantes de la estimación que no considera los efectos fijos nos confirmaron que el análisis extraído de las regresiones que sí los incluyen supera ampliamente al de los resultados obtenidos de aquellas que los omiten.

El estudio culmina señalando ciertas implicancias para las políticas públicas, derivadas de los resultados obtenidos. Las externalidades mostradas evidencian que es necesario un análisis más focalizado al momento de decidir dónde invertir en un colegio nuevo, pues los resultados positivos de dicha inversión dependerán ampliamente de las características de la zona (en lo referente a la estructura del mercado en cuanto a presencia de colegios públicos o privados, la cantidad de alumnos sin acceso a educación y el sexo de los alumnos que pretenden beneficiar con sus inversiones). Por otro lado, la privatización por defecto puede estar afectando negativamente al mercado educativo, por lo que debería ser tomada en consideración al evaluar políticas públicas futuras. Por último, resaltamos que como país necesitamos contar con información más específica y detallada, que nos permita un análisis econométrico más fino.



## **Capítulo 1**

### **Marco contextual**

#### **1.1. Evidencia de la relevancia de la educación para las políticas en el Perú**

La evolución del gasto público en educación como un porcentaje del PBI va de la mano con el modelo económico que haya asumido la autoridad vigente (Baca et al., 2014). En el Perú esto se puede evidenciar al observar el comportamiento de esta variable en la última década. La Figura 1 muestra una tendencia al alza del gasto público en educación, entendido como el porcentaje del PBI que se destina a este fin. Si bien hemos experimentado ciertas fluctuaciones políticas y geopolíticas, la inversión estatal en educación ha mantenido un relativamente estable crecimiento en el tiempo.

A pesar de este aumento del gasto público en educación, según datos del Banco Interamericano de Desarrollo (BID, 2017) Perú es el país que menos invierte en educación en Sudamérica, tomando en cuenta el porcentaje del PBI. Mientras que la región de América Latina y el Caribe invierte en promedio un 5% entre 1990 y 2013, en el Perú se invierte un 3.7%. Chile es el país de Sudamérica que más se acerca, con un 4.6%.

Aún así, la cantidad de colegios nuevos cada año no es poca. De acuerdo a nuestra base de datos, que contempla los años 2007 a 2016, cada año nacen en promedio 277 colegios públicos y 333 colegios privados.

Esto nos lleva a preguntarnos si le rinde al país este aumento de colegios, o si existe una estrategia que pueda guiar mejor los fondos destinados a la educación. El presente trabajo se propone brindar información que sirva de base para formular dichas estrategias en el futuro.

#### **1.2. Literatura existente en torno al mercado educativo peruano**

Un primer aspecto a considerar es la literatura económica que ya existe en torno al mercado educativo a nivel global. En definitiva, el efecto que tiene el dinero que el Estado invierte sobre los resultados educativos no es claro: Bowen and Qian (2017), Calle (2017), Keller (2006), Hanushek (1996), Jackson et al. (2015), Krskova and Baumann (2017), Beuermann et al. (2015) entre otros, encuentran resultados ambiguos o incluso negativos. Por otro lado, Card and Krueger (1992), Dasso et al. (2015), Baca et al. (2014), Webber (2002), Psacharopoulos and Patrinos (2018) entre otros, encuentran los más intuitivos resultados positivos.

Acotando aún más la búsqueda, de acuerdo a nuestras investigaciones, la literatura peruana se centra en el impacto del canon y regalías mineras (Manrique et al. (2016), Calle (2017), Agüero (2016)), o en la inversión estatal en educación (Beuermann et al. (2015), Saavedra (1999), Reinikka and Smith (2004), Agüero (2016)).

Vemos así que la literatura en torno a la inversión pública en educación en el Perú, en lo que respecta al alcance de nuestra búsqueda, es todavía limitada. Encontramos aquí una ventana interesante para la investigación empírica, que coincide con lo señalado por Rodríguez-Sosa (2013): “No se ha hecho evaluaciones que permitan analizar la efectividad de las inversiones”.

### **1.3. Posible privatización por defecto en el Perú**

Un aspecto de importancia que tiene una creciente influencia en la literatura peruana respecto a la relación entre colegios públicos y privados es la privatización por defecto en el Perú.

En los años respectivos a nuestro estudio, se pueden observar dos tendencias claras en la base de datos recopilada. Por un lado, como se puede observar en la Figura 2, los colegios públicos han mejorado su rendimiento; al 2015 sus resultados en la ECE para matemáticas son muy similares a los de los colegios privados en promedio, mientras que la brecha en las notas de la ECE para lengua se ha ido reduciendo.

Por otro lado, la tendencia en la cantidad de matriculados ha dado un giro inesperado: en el 2007 había más alumnos matriculados en colegios públicos que en colegios privados; y al 2016 eso se ha invertido: tenemos ahora más alumnos en colegios privados que en colegios públicos. Esto es lo que se puede observar en la Figura 3.

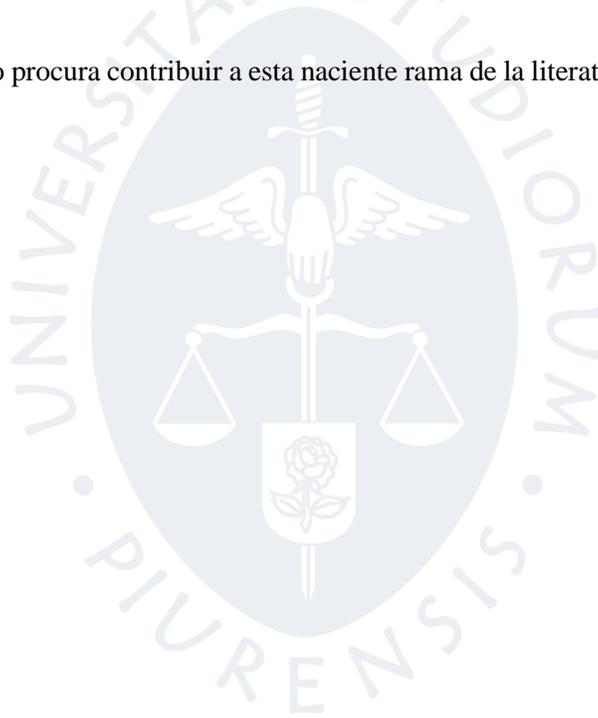
Esto nos lleva a cuestionarnos el motivo por el que los colegios públicos vienen presentando mejoras académicas pero las familias van migrando al sector privado. La literatura ha señalado la importancia de la decisión de las familias en la educación de un país, no solo a nivel peruano sino global (Black (1999), Dizon-Ross (2019) y Chiappori et al. (2017)). Por lo que queda abierta una importante puerta para investigar la relación entre dichos factores.

En lo que respecta al caso peruano, al brindar una explicación para las dos tendencias mencionadas, Guadalupe et al. (2017) señala que “esto es, claramente, el resultado conjunto de la mejora de la institución educativa estatal y de la expansión de la no estatal (...) , el deterioro medio de los resultados de aprendizaje

en la educación no estatal recientemente observado no quiere decir que las instituciones con este tipo de gestión se estén deteriorando en su conjunto, sino que se han expandido «hacia abajo» en términos de dichos resultados”.

Sin embargo, viene surgiendo otra rama de la literatura que se inclina a considerar que la privatización por defecto es un riesgo para el Perú. Cuenca (2013), Balarín (2015), Balarín (2016), Fontdevila et al. (2018), Balarín et al. (2018), Cuenca et al. (2019) y Málaga and Velásquez (2020) analizan el fenómeno del surgimiento de escuelas privadas de bajo costo que, aprovechando el desprestigio de los colegios públicos que mueve a las familias a preferir “a ciegas” un colegio privado, reducen la calidad educativa para generar mayores ingresos.

El presente estudio procura contribuir a esta naciente rama de la literatura.





## **Capítulo 2**

### **Cuestión de investigación**

El presente trabajo de investigación estudia las externalidades entre colegios públicos y privados en el Perú, explicadas por la cercanía geográfica entre ambos y accionada por la mejora en los laboratorios de ciencias de los colegios cercanos.

Sabemos que los colegios privados conocen las mejoras en el rendimiento de sus pares públicos y privados, y además necesitan ingresos de las familias para subsistir. Cuentan también con una mayor flexibilidad de movimientos que los colegios públicos para tomar decisiones estratégicas, y por lo tanto son más sensibles a los cambios a su alrededor. Adicionalmente la privatización por defecto es un factor que podría estar influenciando las relaciones entre colegios públicos y privados. Concluimos que queda una ventana abierta al análisis del impacto que pueden tener los colegios públicos en los colegios privados cercanos.

Ahora bien, un análisis de dicha envergadura es un reto para la investigación peruana. Los estudios educativos suelen contar con variables muy finas que permiten un análisis detallado. Por ejemplo, Chiappori et al. (2017) nos dicen que “el capital humano de un niño depende del tiempo que sus padres pasen con él y del capital humano de los padres” y Romero et al. (2020) señalan que “brindar a los padres de familia una información clara y digerible sobre el desempeño de sus hijos los lleva a actualizar sus creencias y ajustar sus inversiones”. Las bases de datos disponibles al público en el Perú no llegan a este nivel de detalle; por lo que nos enfrentamos al mismo problema que Guadalupe et al. (2017), donde menciona que este tipo de información es “difícil de estimar dadas las características de la información disponible”. La clave está en encontrar aquellas variables relevantes que nos puedan revelar información importante para la toma de decisiones, sabiendo aplicar con creatividad una metodología robusta que permita una interpretación plausible de los resultados. Esto requiere de una metodología un tanto más “rebuscada” para suplir la falta de variables, que a su vez no necesariamente brinde resultados tan potentes, sino más bien indicios que permitan saber hacia dónde ir profundizando.

El presente estudio ahondó en las bases de datos abiertas al público para el Perú, encontrando que efectivamente lo señalado por Guadalupe et al. (2017) es un reto real. Rápidamente comprobamos que aquellas variables que directamente nos hablan sobre los rendimientos educativos o las condiciones para su

respectiva mejora (ratio profesor alumno, salarios de profesores, inversiones monetarias a nivel de colegio, etc.) no se encuentran disponibles.

Sin embargo, con lo que sí contamos es con buena información en torno a un aspecto clave en la educación: la infraestructura. El aprendizaje no se desarrolla como un vaciado de información de un USB a la computadora. Existe una amplitud de factores que afectan la forma en la que un niño capta, retiene y hace propio el aprendizaje. Si bien el conocimiento en sí es fundamental, las habilidades del docente y el entorno educativo juegan un rol clave.

La literatura se ha pronunciado al respecto. Murillo (2011) realiza un estudio que “confirma la incidencia positiva y significativa en los resultados académicos de la educación primaria en América Latina para los alumnos de colegios que cuentan con gimnasios o espacios deportivos, laboratorios, librerías, salas de música y aulas de computadoras”. Del mismo modo, Branham (2004) afirma que “la calidad de la infraestructura escolar tiene un efecto significativo en la asistencia escolar y el ratio de deserción académica”, y Konstantopoulos and Borman (2011) encuentran también un impacto positivo de las instalaciones de los laboratorios de los colegios sobre el aprendizaje de los alumnos.

Zhao and Glewwe (2010) ahondan un poco más en la explicación de estos resultados. Afirman que “los laboratorios de ciencias pueden ayudar a que los estudiantes relacionen la teoría abstracta de sus libros de ciencias con la realidad, incrementando así su interés por la ciencia. También pueden ayudar a que los estudiantes aprendan por prueba y error, en lugar de solo por memoria, lo cual puede incrementar el aprendizaje y reducir la deserción escolar ocasionada por el pobre desempeño académico (...) Los laboratorios de ciencias pueden estar correlacionados con características no observables del colegio”.

Entendemos entonces que los laboratorios tienen un rol importante en la educación, pero que además pueden servir como un buen elemento de señalización del modelo de enseñanza y la calidad educativa de la institución. Es decir, si bien no se conoce el monto invertido por un colegio en su infraestructura, ni tampoco se sabe cuánto valora la institución educativa su respectiva calidad (i.e., ir más allá de cumplir con lo mínimo para seguir operando); el tener una variable que permita conocer la mejora de los laboratorios sirve como una señal de la metodología de enseñanza que existe en la institución, y por lo tanto tiene consecuencias clave en los resultados académicos.

Por otro lado, en el caso de Perú, Cuenca et al. (2017) mencionan que “la infraestructura es un

problema serio. Más de la mitad de las escuelas públicas del país no se encuentran en buen estado. Los servicios básicos son desiguales entre las escuelas urbanas y las rurales”.

De esta forma podemos apreciar que las mejoras en infraestructura no son solo un factor clave en el rendimiento académico de los alumnos, sino que para el caso peruano se trata de una necesidad fundamental que se evidencia con mayor claridad. Al tener grandes niveles de desigualdad, un colegio que mejora su infraestructura se diferencia notoriamente del resto. Luego, esta diferencia en infraestructura no se limita a la estética o acabado de sus edificios, sino que revela además su inversión en calidad educativa. Un colegio en Perú que da un salto en su nivel de infraestructura da una señal notable en relación a su calidad educativa.

Una consecuencia de lo expuesto en el párrafo anterior es que al lograr medir de cierta manera las mejoras en infraestructura educativa, uno puede ir descubriendo las repercusiones que dicha inversión tenga en el entorno. Es más: Cuesta et al. (2016) realizan una revisión de la literatura económica y educativa en torno a la inversión en educación, y nos dicen que dicha literatura “brinda evidencia sobre que los laboratorios y bebederos aumentan las matrículas escolares”. Las familias prefieren aquellos colegios que evidencian con mayor claridad su inversión en educación; pues, aunque sea involuntariamente, logran captar las señales de la importancia que la institución le brinda a la educación de sus hijos.

El presente estudio logra evaluar la relación entre una variable que mide las mejoras en los laboratorios de ciencias de los colegios públicos con los matriculados en colegios privados que se encuentran geográficamente cerca. En base a lo mencionado previamente, sabemos que los laboratorios de ciencias tienen un rol importante en la educación al ser una herramienta poderosa para aplicar los conocimientos de química y física adquiridos en la teoría, y por ello pueden servir como un buen elemento de señalización de la calidad educativa de la institución (Konstantopoulos and Borman (2011), Cuesta et al. (2016), entre otros). Es un elemento de la infraestructura educativa que es más sencillo de medir con evidencia tangible, y que a la vez puede revelar “características no observables” (Zhao and Glewwe, 2010) de los colegios.

Debido a la información disponible, no se debería esperar un resultado que abarque en su totalidad una explicación exhaustiva de la relación entre los colegios públicos y privados, sino más bien una aproximación a la existencia de dicha relación que pueda llevar a profundizar en estudios futuros.

Por ese motivo acotamos nuestro trabajo de investigación a las externalidades generadas por los colegios públicos hacia los colegios privados, empleando un análisis complementario de las externalidades

de los colegios públicos a otros colegios públicos, a modo de comparación del nivel de sensibilidad de ambos.

Consideramos de especial interés este análisis ya que las otras posibles combinaciones perdían cierto atractivo como parte sustancial de la cuestión de investigación. El análisis de externalidades del privado al público no prometía mayores relevancias, al esperar poca sensibilidad por parte del sector público; y el análisis del privado al privado o del privado consigo mismo (ya no como una externalidad, sino evaluando el rendimiento de sus inversiones) implicaba desconocer factores probablemente sustanciales en la toma de decisiones de ambos colegios (en este caso, por ejemplo, no se cumpliría lo mencionado previamente en relación a la endogeneidad: tal vez existan colegios privados que empresarialmente prevean una ampliación en la cantidad de matriculados, y estratégicamente mejoren su infraestructura. Este es el tipo de información clave para un análisis del mercado educativo privado, con el que no contamos).

Habiendo expuesto estas aclaraciones, consideramos que existen distintos resultados que se podrían esperar del presente análisis, teniendo en cuenta lo expuesto hasta el momento.

Un primer resultado que esperamos es la existencia de externalidades de los colegios públicos a los privados, lo cual se reflejaría en obtener coeficientes estadísticamente significativos en las regresiones detalladas en la sección de Especificación. Por los motivos explicados en los párrafos previos, tenemos evidencia a favor de que exista un efecto de la mejora de los laboratorios de los colegios públicos en los alumnos matriculados en colegios privados.

Un segundo resultado esperado lo fundamentamos en lo mencionado por De la Croix and Doepke (2009). Ahí se menciona que “la presencia de un amplio sector de educación privada beneficia a los colegios públicos en democracias donde los políticos buscan complacer a familias de bajos niveles económicos, pero minimiza el gasto en educación pública en sociedades políticamente dominadas por los ricos”. Esto es evidencia sobre la existencia de casos en los que se dan externalidades positivas entre el sector privado y el sector público educativo. Por lo que esperamos que la relación de la presente cuestión no sea necesariamente negativa, sino incluso positiva: existen casos en los que se benefician entre ellos, por lo que podría pasar que nuestros hallazgos vayan en esa línea, mostrando externalidades positivas.

En tercer lugar, teniendo en mente a Cuenca et al. (2017) quienes señalan que “la probabilidad de que un joven tenga un nivel de educación por encima del de sus compañeros (grupo de la misma cohorte

que se encuentra en la zona urbana/rural de la misma región) todavía depende del género, de los ingresos del hogar y de si acude a una escuela pública o a una privada”, probablemente encontremos diferencias en los resultados según tratemos con estudiantes varones o mujeres. Esto tendría sentido, considerando la diferentes metodologías de aprendizaje o las expectativas culturales para cada uno de los sexos.

Un cuarto aspecto a considerar son los mecanismos que llevarían a que se generen las externalidades. La posibilidad de la privatización por defecto, señalada en la subsección previa, podría ser un factor relevanteal respecto. Una relación positiva entre las variables hablaría a favor de su existencia: aunque los colegios públicos mejoren, las familias siguen optando por la educación privada.

Finalmente, un último tema a abordar es la reacción por parte de los colegios públicos a las actividades de otros colegios públicos. La organización del sistema educativo público nos lleva a pensar que probablemente este reaccione con una sensibilidad menor que el privado. Esto se debe a que el sistema educativo en el Perú sigue la misma estructura que la división territorial: las jurisdicciones territoriales marcan el ámbito de influencia de las distintas instancias educativas: el Ministerio de Educación (MINEDU) a nivel nacional, las Direcciones Regionales de Educación (DRE) a nivel departamental, y las Unidades de Gestión Educativa Local (UGEL) a nivel provincial. Estas últimas son las encargadas de definir el presupuesto aprobado para cada colegio público bajo su jurisdicción. Debido a la larga cadena de aprobación que requiere cada modificación del presupuesto, esperamos que la reacción por parte de los colegios públicos sea menor que la de los colegios privados.



### **Capítulo 3**

#### **Base de datos**

La base de datos empleada en el presente estudio se obtuvo del Censo Escolar del Perú elaborado por el Ministerio de Educación del Perú para el periodo de años entre el 2007 y 2016, inclusive. A partir de esta base de datos se identificó a los colegios de gestión pública y privada del nivel secundario de todo el territorio nacional. Los colegios están identificados geográficamente a nivel de la variable "distritos" y son acompañados de diversas variables que capturan características tales como la cantidad de alumnos matriculados, si la institución cuenta con un laboratorio, la ubicación de la institución a través de coordenadas geográficas, el tamaño del colegio, entre otras. La base de datos se encuentra disponible para uso público de manera *online* en la página web oficial de ESCALE.

Aprovechando la variable de "si la institución cuenta con un laboratorio" se creó la variable "Índice del Laboratorio" que consiste de un índice armado de acuerdo al nivel de experimentos que se pueden realizar en el laboratorio de ciencias del colegio. El índice toma el valor de 0 si el colegio no tiene laboratorio, 1 si el colegio tiene laboratorio, 2 si dicho laboratorio permite realizar experimentos correspondientes al Ciclo 6 de Educación Básica Regular (es decir primero y segundo de secundaria) y 3 si el laboratorio permite realizar experimentos correspondientes al Ciclo 7 de Educación Básica Regular (es decir, tercero, cuarto y quinto de secundaria). Nos explayaremos en esta variable al describir más adelante la Tabla 1.

Luego, se aprovechó la disponibilidad de coordenadas geográficas para cada colegio y la capacidad de ubicarlos en el mapa del Perú para realizar emparejamientos de los mismos mediante georreferenciación. Se utilizó como único criterio de emparejamiento la cercanía geográfica entre los colegios públicos con sus pares cercanos ya sea públicos como privados, a los cuales denominamos "colegios focales". Es así que, cada colegio público del Perú fue emparejado con los colegios públicos y privados focales que se encuentren dentro de un radio de distancia de 1000m.

Nuestro principal interés es analizar la hipótesis sobre la existencia de externalidades motivadas por la cercanía geográfica entre colegios, de tal manera que no consideramos pertinente incluir otro criterio de emparejamiento dado que distorsionaría el alcance de dichas externalidades y no se abordaría de manera precisa la hipótesis del presente trabajo.

Adicionalmente, se emplearon los resultados en Lenguaje y Matemática de la Evaluación Censal

de Estudiantes (ECE) del 2016 y 2018 aplicada por el Ministerio de Educación. Esta evaluación se realiza anualmente a estudiantes de cuarto de primaria y segundo de secundaria de colegios públicos y privados del Perú. Para el presente estudio hemos extraído los datos a nivel de colegio.

Las evaluaciones tienen un rango de puntaje que corresponde a los siguientes niveles de progreso: En inicio, En proceso y Satisfactorio. Por un lado, para las evaluaciones en matemática los intervalos de puntaje son 0 - 512, 513 - 605 y 606 - 1000, respectivamente. Y por otro lado, para las evaluaciones en lengua los intervalos de puntajes son 0 - 458, 459 - 581 y 582 - 1000, respectivamente.

Aprovechando las variables y bases de datos descritas anteriormente se armó una base de datos de panel compuesta por el Índice del Laboratorio de los colegios públicos, emparejados con la cantidad de alumnos y resultados académicos de los colegios públicos y privados focales cercanos, en un rango temporal del 2007 al 2016.

La Tabla 1 muestra los estadísticos sumarios propios de las variables empleadas en el presente estudio. Debido a la naturaleza de nuestra cuestión de investigación, consideramos enriquecedor distinguir dichas variables para el total de colegios de todo el Perú (en el rango temporal del 2007 al 2016), y luego distinguir esas mismas variables solo para los colegios públicos y los colegios privados.

El Panel A señala las variables para los colegios tanto públicos como privados. La primera variable mostrada es la variable explicativa del estudio: el índice de laboratorio de ciencias. Esta es una variable categórica que toma el valor de 0 si el colegio no cuenta con un laboratorio de ciencias, el valor de 1 si cuenta con un laboratorio de ciencias, el valor de 2 si dicho laboratorio está equipado para realizar experimentos correspondientes al Ciclo 6 de la Educación Básica Regular (i.e., primero y segundo de secundaria) y el valor de 3 si dicho laboratorio está equipado para realizar experimentos correspondientes al Ciclo 7 de la Educación Básica Regular (i.e., tercero, cuarto y quinto de secundaria). Para efectos de todas las interpretaciones futuras emplearemos una nomenclatura que hace referencia al índice del laboratorio de ciencias, asignándole un nivel 1, 2 o 3 según sea el caso.

Cabe resaltar que este índice se armó a partir de seis preguntas del Censo Escolar del Perú, todas con respuestas de "sí" o "no". La primera de ellas hace referencia a que el colegio tenga o no laboratorio, y las otras 5 hacen referencia a que dicho laboratorio permita realizar experimentos en cada uno de los grados de educación secundaria (de primero a quinto grado). La categoría 1 del índice se activa en función a la

primera pregunta descrita, las categorías 2 y 3 se activan en función a los otras cinco preguntas. Gracias a la forma como se armó el índice podemos tener certeza de que la categoría 1 es mutuamente excluyente con las categorías 2 y 3.

Así mismo, las preguntas seleccionadas del Censo nos permitieron realizar un análisis detallado de la información contenida para armar este índice. Basándonos en un estudio profundo de la Currícula Nacional para la Educación Básica Regular, logramos distinguir que el mayor cambio en la implementación de los laboratorios no se realizaba de grado a grado sino de ciclo a ciclo. Concluyendo de esta forma la mejor manera de organizar las categorías del índice.

Un aspecto a resaltar en lo referente a los estadísticos mostrados en el Panel A de la Tabla 1 para esta variable es que tiene un valor promedio para todos los colegios de 0.58: concluimos que son muchos los colegios que ni siquiera cuentan con laboratorio.

Las otras 4 variables mencionadas hacen referencia a las dependientes que reflejan la cantidad de alumnos de los colegios focales, distinguiendo según sexo. Las dos primeras, si bien son presentadas como logaritmos en las regresiones, aquí se muestran como cantidades para facilitar la interpretación. Las dos últimas son variables dicotómicas, que se activan cuando el colegio recibió más alumnos que el año anterior. Las variables se distinguen según sexo.

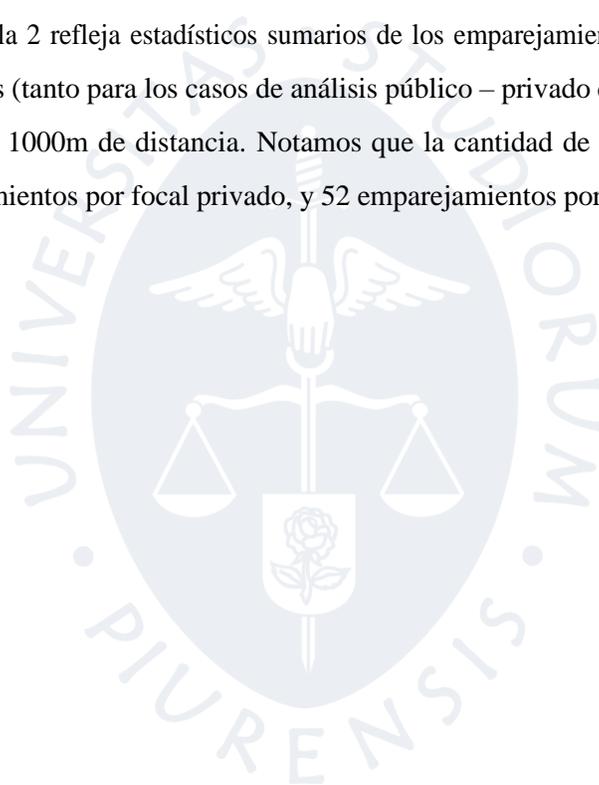
El Panel B presenta las variables para todos los colegios públicos del Perú, y el Panel C presenta las variables para todos los colegios privados del Perú. Un análisis comparativo entre los estadísticos de ambos paneles puede ser de interés: notamos una importante diferencia, por ejemplo, entre el promedio del índice de laboratorio: para los colegios públicos es de 0.5, mientras que para los colegios privados es de 0.8. Podríamos decir que en ambos casos es bajo, pero que probablemente existen más colegios privados que cuentan con laboratorios mejor implementados. Por otro lado, la cantidad de alumnos en los colegios públicos es significativamente más dispersa que en los colegios privados (el valor máximo de la variable es mucho mayor en colegios públicos, pero el promedio es bastante similar). También es de interés señalar que en promedio los colegios privados están recibiendo más alumnos nuevos.

Existe en ambos paneles una última batería de variables que consideramos relevante incluir: las variables empleadas en las interacciones. La variable “grande” es una dicotómica que toma el valor de 1 si el número de matriculados a nivel de colegio es mayor a la mediana del número de matriculados total de

colegios en cada año. Se da en función a los pares públicos o privados del colegio focal (público o privado, respectivamente), por lo que una conclusión interesante es que los colegios públicos son más similares entre sí en cuanto a cantidad de alumnos (pues la variable “grande” es una dicotómica que compara entre el mismo tipo de colegios).

Por otro lado, la variable “alto desempeño” es una dicotómica que toma el valor de 1 si el promedio de notas en la ECE en matemática y lengua a nivel de colegios es satisfactorio. Los estadísticos nos indican que, en promedio, los colegios públicos obtienen resultados más positivos en la ECE que los colegios privados.

Finalmente, la Tabla 2 refleja estadísticos sumarios de los emparejamientos realizados entre colegios públicos y colegios focales (tanto para los casos de análisis público – privado como público – público) que se encuentran a menos de 1000m de distancia. Notamos que la cantidad de emparejamientos es alta: un promedio de 59 emparejamientos por focal privado, y 52 emparejamientos por focal público.



## Capítulo 4 Especificación

### 4.1. Diseño empírico empleado

El interés del presente estudio es investigar la existencia de externalidades positivas generadas de los colegios públicos hacia los privados. Teniendo presentes las limitaciones en la cantidad de variables disponibles, y el interés por presentar resultados con información interesante, presentamos una estrategia metodológica que permita minimizar la complejidad del modelo, maximizando la relevancia de los resultados.

Para ello, el análisis distingue entre dos tipos de colegios a los que hacer referencia: el tipo de gestión (público o privada) y la función en el modelo (como colegio focal o colegio independiente). De esta manera, podemos contar con resultados para colegios públicos o privados que sean focales. Para distinguirlos entre sí, empleamos el subíndice de  $v$  para los colegios privados, y el subíndice  $b$  para los colegios públicos. Para los casos en los que tanto los colegios focales como los independientes sean públicos, se ha añadido un número al subíndice para distinguirlos ( $b1$  si es el colegio focal,  $b2$  si es el colegio independiente).

A su vez, cabe resaltar que los efectos fijos y las variables de las interacciones siempre están asociadas al colegio focal en cuestión, debido a que la metodología analiza el impacto sobre los focales. Por ende, nos interesa controlar por las características inherentes a los colegios focales (lo cual se trabaja con los efectos fijos) y profundizar en el impacto sobre los colegios focales (lo cual se trabaja con las interacciones).

Con todo esto en mente, la especificación utilizada considera un modelo de efectos fijos a nivel de año y a nivel de colegio focal, a través de la siguiente ecuación:

$$y_{v,t} = \alpha + \beta \cdot lab_{b,t} + \lambda_t + \tau_v + \mu_{v,t}, \quad (1)$$

en donde  $\beta$  describe la relación entre el índice del laboratorio del colegio público  $b$  y la variable dependiente

$y_{v,t}$  a nivel de colegio focal privado  $v$  en el año  $t$ , que se encuentra a menos de 1000m de distancia.

Con el fin comparar los resultados de los colegios focales privados con colegios focales públicos, se trabaja la siguiente ecuación:

$$y_{b_1,t} = \alpha + \beta \cdot lab_{b_2,t} + \lambda_t + \tau_{b_1} + \mu_{b_1,t}, \quad (2)$$

en donde  $\beta$  describe la relación entre el índice del laboratorio del colegio público  $b_2$  y la variable dependiente  $y_{b_1,t}$  a nivel de otro colegio focal público  $b_1$  en el año  $t$ , que se encuentra a menos de 1000m de distancia. Las variables independientes  $lab_{b,t}$  y  $lab_{b_2,t}$  son un índice armado de acuerdo al nivel de experimentos que se pueden realizar en el laboratorio de ciencias del colegio público  $b$  y  $b_2$  respectivamente, emparejado al colegio focal  $v$  y  $b_1$ , respectivamente; en el año  $t$ . El índice toma el valor de 0 si el colegio no tiene laboratorio, 1 si el colegio tiene laboratorio, 2 si dicho laboratorio permite realizar experimentos correspondientes al Ciclo 6 de Educación Básica Regular (es decir primero y segundo de secundaria) y 3 si el laboratorio permite realizar experimentos correspondientes al Ciclo 7 de Educación Básica Regular (es decir, tercero, cuarto y quinto de secundaria).

Las variables dependientes  $y_{v,t}$  y  $y_{b_1,t}$  reflejan los alumnos matriculados en el colegio focal privado o público, respectivamente, en el año  $t$ ; tomando dos formas distintas: el logaritmo del total de matriculados en secundaria y una dicotómica que se activa si el colegio recibió más alumnos en primer grado que el año anterior, haciendo la distinción según sexo del alumno para ambos casos.

La ecuación (1) incluye  $\lambda_t$  efectos fijos a nivel de año y  $\tau_v$  efectos fijos a nivel de colegio focal privado, utilizando observaciones para los años 2007 al 2016. La ecuación (2) incluye  $\lambda_t$  efectos fijos a nivel de año y  $\tau_{b_1}$  efectos fijos a nivel de colegio focal público, utilizando observaciones para los años 2007 al 2016. Estos efectos fijos nos permitieron controlar por distintos factores inherentes al mercado en cuestión, a pesar de no contar con variables específicas para este fin. De tal manera que preocupaciones relacionadas al crecimiento poblacional o a la definición de tipos de mercado no fue tan acuciante. Por último, como no se puede asumir que las observaciones dentro de un mismo distrito son independientes, los errores estándares  $\mu_{v,t}$  y  $\mu_{b_1,t}$  son calculados utilizando robustez por *clusters* a nivel de distrito del colegio focal privado o público, respectivamente.

Con la finalidad de comprender mejor las características de los colegios privados focales que afectan a las externalidades derivadas de la ecuación (1), el análisis es enriquecido con dos especificaciones que consideran interacciones a través de la ecuación:

$$y_{v,t} = \alpha + \beta_1 \cdot lab_{b,t} + \beta_2 \cdot grande_{v,t} + \beta_3 \cdot lab_{b,t} \cdot grande_{v,t} + \lambda_t + \tau_v + \mu_{v,t} \quad (3)$$

en donde  $\beta_3$  describe la relación entre cada nivel del índice del laboratorio del colegio público  $b$  y la variable

dependiente  $y_{v,t}$  para un colegio focal grande  $v$  en el año  $t$ , y de la ecuación:

$$y_{v,t} = \alpha + \beta_1 \cdot lab_{b,t} + \beta_2 \cdot alto_{v,t} + \beta_3 \cdot lab_{b,t} \cdot alto_{v,t} + \lambda_t + \tau_v + \mu_{v,t} \quad (4)$$

en donde  $\beta_3$  describe la relación entre cada nivel del índice del laboratorio del colegio público  $b$  y la variable dependiente  $y_{v,t}$  para un colegio focal  $v$  con alto desempeño en el año  $t$ .

La variable dependiente  $y_{v,t}$  refleja los alumnos matriculados en el colegio focal, tomando las dos formas descritas anteriormente: el logaritmo del total de matriculados en secundaria y una dicotómica que se activa si el colegio recibió más alumnos en primer grado que el año anterior, haciendo la distinción según sexo del alumno para ambas formas.

Por un lado,  $grande_{v,t}$  es una dicotómica que se activa si el número de alumnos matriculados en el colegio focal es mayor a la mediana del total de matriculados en los colegios focales para cada año (siguiendo lo propuesto por Harrison and Wicks (2013)). Por otro lado,  $alto_{v,t}$  es una dicotómica que se activa si el promedio de notas en la ECE para matemática y lengua a nivel de colegio focal es satisfactorio en el año correspondiente.

La ecuación (3) considera interacciones con la dicotómica  $grande_{v,t}$ , además, incluye  $\lambda_t$  efectos fijos a nivel de año y  $\tau_v$  efectos fijos a nivel de colegio focal privado, utilizando observaciones para los años 2016 y 2018. La ecuación (4) considera interacciones con la dicotómica  $alto_{v,t}$ , además, incluye  $\lambda_t$  efectos fijos a nivel de año y  $\tau_v$  efectos fijos a nivel de colegio focal privado, utilizando observaciones para los años 2016 y 2018. Como no se puede asumir que las observaciones dentro de un mismo distrito son independientes, los errores estándar  $\mu_{v,t}$  son calculados utilizando robustez por *clusters* a nivel de distrito del colegio privado.

## 4.2. Robustez

La robustez del modelo podría estar en juego por dos factores: la endogeneidad o las alteraciones en las variables empleadas.

La endogeneidad no fue considerada como un riesgo para el modelo presentado debido a la naturaleza de las variables implicadas. Por lógica temporal, la mejora en infraestructura debió ser implementada antes de que se vean los resultados en el Censo Escolar (es decir, para cuando se realizó el Censo la inversión

habría culminado), y por lo tanto no puede preceder - ni en tiempo ni en causalidad - a la matrícula del colegio focal.

Por este motivo centramos nuestra evaluación de robustez en una especificación que nos permita observar cómo se comporta el modelo frente a alteraciones en las variables empleadas.

Fundamentamos la elección de las variables en el posible riesgo de que las externalidades entre colegios públicos y privados se limiten a la facilidad económica de los padres. Con esto nos referimos a la posibilidad de que los resultados hallados sean simple reflejo de la situación socioeconómica de la zona: en determinado distrito existen familias que pueden pagar un colegio privado y mandan ahí a sus hijos, y existen familias que no pueden pagar un colegio privado y mandan a sus hijos a colegios públicos. En tal caso, sencillamente los distritos con mayor poder adquisitivo contarían con más recursos para mejorar su infraestructura (incluyendo la de los colegios públicos); y también contarían con una población con mayores posibilidades de enviar a sus estudiantes a un colegio privado; ergo, la relación entre colegios públicos y colegios privados sería inexistente (incluyendo aquí las externalidades de la presente cuestión de investigación).

Este análisis lo hacemos mediante la siguiente especificación, que invierte el tipo de gestión del colegio focal y del colegio explicativo de la ecuación (1):

$$y_{b,t} = \alpha + \beta \cdot \ln ab_{v,t} + \lambda_t + \tau_b + \mu_{b,t}, \quad (5)$$

La variable independiente  $\ln ab_{v,t}$  se mantiene como el índice armado de acuerdo al nivel de implementación del laboratorio del colegio privado  $v$  en el año  $t$ , pero ahora hace referencia a los colegios privados. Mientras que las variables dependientes  $y_{b,t}$  reflejan los alumnos matriculados en el colegio focal público, tomando las dos formas diversas mencionadas anteriormente: el logaritmo del total de matriculados en secundaria según sexo y una dicotómica que se activa si el colegio recibió más alumnos en primer grado que el año anterior, haciendo la distinción según sexo del alumno.

La ecuación (5) incluye, además,  $\lambda_t$  efectos fijos a nivel de año y  $\tau_b$  efectos fijos a nivel de colegio focal público, utilizando observaciones para los años 2016 y 2018. Los errores estándar  $\mu_{b,t}$  son calculados utilizando robustez por *clusters* a nivel de distrito.

Consideramos también que las interacciones presentadas en la subsección anterior podrían contribuir

a un análisis de robustez del modelo, al presentar implicancias que modifican las condiciones del mercado educativo.

Mediante estas dos herramientas esperamos reafirmar la solidez del modelo, así como la existencia de las externalidades entre colegios públicos y privados.

### 4.3. Evaluación de la metodología

Teniendo en mente el objetivo del presente trabajo de investigación por minimizar la complejidad del modelo, maximizando la relevancia de los resultados dada la base de datos accesible, para verificar la eficiencia de los efectos fijos trabajamos las siguientes especificaciones:

$$y_{v,t} = \alpha + \beta \cdot lab_{b,t} + \mu_{v,t}, \quad (6)$$

y

$$y_{b_1,t} = \alpha + \beta \cdot lab_{b_2,t} + \mu_{b_1,t}, \quad (7)$$

Las variables dependientes  $y_{v,t}$  y  $y_{b_1,t}$  reflejan los alumnos matriculados en el colegio focal privado o público, respectivamente, en el año  $t$ ; tomando las dos formas diversas descritas anteriormente: el logaritmo del total de matriculados en secundaria y una dicotómica que se activa si el colegio recibió más alumnos en primer grado que el año anterior, haciendo la distinción según sexo del alumno para ambas formas.

Las ecuaciones (6) y (7) son similares a las ecuaciones (1) y (2), pero sin incluir efectos fijos. Este análisis se realizó con la finalidad de evaluar la importancia de la metodología en los hallazgos presentados.

Como no se puede asumir que las observaciones dentro de un mismo distrito son independientes, los errores estándar  $\mu_{v,t}$  y  $\mu_{b_1,t}$  son calculados utilizando robustez por *clusters* a nivel de distrito del colegio focal privado o público, respectivamente.



## Capítulo 5

### Resultados

#### 5.1. Externalidades positivas y competencia

##### 5.1.1. *Externalidades positivas de los colegios públicos a los privados.*

La Tabla 3 muestra los resultados correspondientes a la especificación (1). La variable dependiente refleja los alumnos matriculados en el colegio privado focal, tomando dos formas diversas: el logaritmo del total de matriculados en secundaria (según sexo) y una dicotómica que se activa si el colegio recibió más alumnos en primer grado que el año anterior, haciendo la distinción según sexo del alumno.

La variable independiente es un índice armado de acuerdo al nivel de experimentos que se pueden realizar en el laboratorio del colegio público emparejado al colegio focal privado. El índice toma el valor de 0 si el colegio no tiene laboratorio, 1 si el colegio tiene laboratorio, 2 si dicho laboratorio permite realizar experimentos correspondientes al Ciclo 6 de Educación Básica Regular (es decir primero y segundo de secundaria) y 3 si el laboratorio permite realizar experimentos correspondientes al Ciclo 7 de Educación Básica Regular (es decir, tercero, cuarto y quinto de secundaria). El índice se armó de manera que las categorías 1, 2 y 3 sean mutuamente excluyentes. Para efectos de todas las interpretaciones futuras emplearemos una nomenclatura que hace referencia al índice del laboratorio de ciencias, asignándole un nivel de 1, 2 o 3 según sea el caso.

Los resultados de esta investigación señalan una relación positiva y estadísticamente significativa entre la mejora de los laboratorios de colegios públicos y el número de matriculados varones de los colegios privados que se encuentran a menos de 1000m de distancia. En concreto, como se muestra en la Tabla 3, una mejora sustancial de las instalaciones del laboratorio del colegio público se relaciona con un aumento del 0.28% de los alumnos varones matriculados en el colegio focal respectivo, con un intervalo de confianza del 99%. Por otro lado, dicha mejora en los laboratorios se relaciona con un ligero pero estadísticamente significativo aumento de 0.19% en la probabilidad de que ingresen alumnos varones a primer grado al colegio, con un intervalo de confianza del 99%. Los resultados para las mujeres fueron no significativos, y serán analizados en la sección subsiguiente.

Consideramos que estos resultados muestran la existencia de externalidades positivas de los colegios

públicos a los privados cercanos. Al contar con los efectos fijos, se puede afirmar que los coeficientes están libres de todo lo inherente al colegio privado en cuestión; por lo que el aumento de matriculados en el colegio privado no se debe a factores “privados” ni geográficos, sino que está vinculado a la mejora del laboratorio del colegio público. Concluimos entonces que existe un efecto indirecto en los colegios privados.

Analizamos diversos mecanismos, basados en supuestos lógicos, que podrían ser los causantes de la existencia de estas externalidades. Consideramos dos canales fundamentales: el primero es que los colegios públicos motivan de una forma más cercana a la población de la vecindad correspondiente, que tiene el potencial de ingresar al mercado educativo. Por lo que una mejora en la infraestructura de los colegios públicos que es bastante evidente a las familias cercanas desencadena o por lo menos se correlaciona con una mayor motivación por estudiar. En consecuencia, más personas buscan colegios.

El segundo canal fundamental sería la privatización por defecto. Esta, per se, haría que las familias escojan colegios privados independientemente de la mejor o peor calidad de los colegios públicos. Sin embargo, dado que se buscan colegios en la vecindad, esta creencia implícita en la mentalidad daría pie a que se escojan los colegios privados.

En conclusión, las externalidades positivas halladas podrían interpretarse como el resultado de una generalización de la búsqueda de educación canalizada a través de la privatización por defecto. La ausencia de variables que permitan una medición de esta valorización de la educación en la zona o de la privatización por defecto no permite brindar una fundamentación econométrica a nuestra interpretación. Sin embargo, la evidencia desarrollada en base a los siguientes hallazgos refuerza la potencia lógica de nuestra argumentación (son interpretaciones plausibles).

En primera instancia, las interacciones empleadas nos permiten reforzar esta interpretación de los mecanismos subyacentes a las externalidades halladas, pues encontramos que estas son más potentes (el coeficiente crece en magnitud) cuando el privado focal es grande (es decir, tiene un número de matriculados mayor a la mediana de colegios privados del año correspondiente). Estos resultados se pueden apreciar en la Tabla 5.

Vemos que para el caso de matriculados varones, que el colegio sea grande hace que su número de matriculados aumente en un 4.3% cuando el colegio público al que está emparejado tiene un laboratorio de nivel 1, en 7.3% cuando dicho laboratorio es de nivel 2 y en 9.1% cuando dicho laboratorio pasa a ser de

nivel 3. Todos estos resultados cuentan con un nivel de significancia estadística del 99%. Para el caso de matriculadas mujeres los resultados son estadísticamente significativos para los colegios de alto desempeño, y serán descritos en la sección subsiguiente.

Estos resultados dan pie a distinguir los dos canales mediante los cuales las variables se relacionan, presentando incluso un tercer canal que los refuerza: que el colegio sea grande, refuerza la externalidad encauzada por la privatización por defecto. Esto es intuitivo, y acorde a la interpretación previa: se buscan colegios, se prefieren los privados, pero además se prefieren los más notorios (en este caso, los "grandes").

A su vez, la gradación en el aumento del impacto según cada una de las categorías del índice refuerza la existencia de los tres canales: conforme el laboratorio es más completo, más alumnos se ven atraídos. Se podría decir, además, que mientras mayores sean las mejoras en la infraestructura, mayor es la motivación a estudiar y mayor es el beneficio de los colegios privados.

Los resultados presentados no pretenden dar una explicación exhaustiva acerca de la relación entre los colegios públicos y privados, sino contribuir al análisis planteado en secciones previas. Si bien es claro que existen mejoras en la educación pública peruana, estos resultados nos muestran que todavía tenemos camino por recorrer: las tendencias del mercado y el desprestigio de estas instituciones públicas juegan un papel importante en el rendimiento que tienen las inversiones realizadas. Siendo asimismo innegable que existe una relación latente entre los colegios públicos y privados del Perú, que estas externalidades sencillamente han empezado a delinear.

### **5.1.2. Competencia entre colegios públicos**

A continuación exponemos los resultados presentados en la Tabla 4, que evalúa el efecto de la mejora de los laboratorios públicos sobre otros colegios públicos cercanos.

La variable dependiente refleja los alumnos matriculados en el colegio público focal, tomando dos formas diversas: el logaritmo del total de matriculados en secundaria y una dicotómica que se activa si el colegio recibió más alumnos en primer grado que el año anterior, haciendo la distinción según sexo del alumno para ambas formas.

La variable independiente es, de nuevo, un índice armado de acuerdo al nivel de experimentos que se pueden realizar en el laboratorio del colegio público emparejado al colegio focal público. El índice

toma el valor de 0 si el colegio no tiene laboratorio, 1 si el colegio tiene laboratorio, 2 si dicho laboratorio permite realizar experimentos correspondientes al ciclo 6 de Educación Básica Regular (es decir primero y segundo de secundaria) y 3 si el laboratorio permite realizar experimentos correspondientes al ciclo 7 de Educación Básica Regular (es decir, tercero, cuarto y quinto de secundaria). El índice se armó de manera que las categorías 1, 2 y 3 sean mutuamente excluyentes. Para efectos de todas las interpretaciones futuras emplearemos una nomenclatura que hace referencia al índice del laboratorio de ciencias, asignándole un nivel de 1, 2 o 3 según sea el caso.

Los resultados de la investigación señalan una relación negativa entre el nivel de laboratorio de los colegios públicos, y la cantidad de alumnos matriculados en los colegios públicos que se encuentran a menos de 1000m de distancia. En concreto, como se muestra en la Tabla 4, una mejora sustancial de las instalaciones en laboratorios está vinculada a un decrecimiento del 0.37% de los matriculados varones del colegio público focal. Estos son los únicos resultados estadísticamente significativos, con un intervalo de confianza del 99%. Para el caso de la variable dicotómica de nuevos alumnos se obtuvo que una mejora sustancial de las instalaciones del laboratorio está vinculada a un decrecimiento del 0.16% en la probabilidad de recibir nuevos alumnos, con una significancia estadística del 95%. Los resultados para matriculadas mujeres serán descritos en la sección siguiente.

En primer lugar, cabe resaltar que los colegios públicos reaccionan con una sensibilidad bastante similar a la del sector privado. En segundo lugar, los coeficientes negativos y estadísticamente significativos que mostramos hablan a favor de la existencia de externalidades negativas entre los colegios públicos, que se pueden interpretar como competencia entre los mismos: la gente que opta por este tipo de educación claramente escogerá aquellos colegios que ve que mejoran, especialmente si se encuentran tan cerca.

Se realizaron las mismas interacciones que para los emparejamientos de colegios públicos con colegios privados, considerando las variables de colegio grande y colegio de alto desempeño, pero los resultados no aportan tanto a la investigación. La Tabla I del Apéndice muestra las estimaciones con mayor significancia estadística que se obtuvieron para la variable dependiente del logaritmo de matriculados varones, con las interacciones de colegio grande para varones en secundaria y alto desempeño en matemática.

Por un lado, que el colegio sea grande hace que su número de matriculados aumente en un 1.18% cuando el colegio público al que está emparejado tiene un laboratorio de nivel 1 (con una significancia

estadística del 95%), que disminuya en 3% cuando dicho laboratorio es de nivel 2 (con una significancia estadística del 90%) y que disminuya en 7% cuando dicho laboratorio es de nivel 3 (con una significancia estadística del 99%).

Por otro lado, que el colegio sea de alto desempeño en matemática hace que su número de matriculados aumente en un 2% cuando el colegio público al que está emparejado pasa a ser de nivel 2 (con una significancia estadística del 95%) y en 1% cuando dicho laboratorio pasa a ser de nivel 3 (con una significancia estadística del 95%).

En este caso, se puede apreciar que los colegios públicos reaccionan con una menor sensibilidad a un análisis más fino del mercado educativo: los coeficientes tienen menor magnitud y significancia estadística que en el análisis con el sector privado. A pesar de que encontramos resultados en cierta medida significativos, consideramos que no contamos con la información suficiente para realizar interpretaciones al respecto, que a su vez aporten a la cuestión del presente trabajo de investigación. Para efectos del presente trabajo, vemos que el análisis comparativo mencionado en relación al nivel de sensibilidad es de suficiente utilidad.

### **5.1.3. Diferencia de los resultados para varones y para mujeres**

Las tablas 3, 4 y 5 muestran resultados interesantes en cuanto a la diferencia de los estimadores entre alumnos varones y alumnas mujeres.

En el caso de los emparejamientos entre colegios públicos y colegios privados, una mejora en el laboratorio del colegio público cercano lleva a un decrecimiento del 0.18% de alumnas matriculadas en el colegio privado focal con un nivel de significancia estadística del 90%; así como a un estadístico de 0.06% sin significancia estadística para la dicotómica de nuevas alumnas.

En el caso de los emparejamientos entre colegios públicos, una mejora en el laboratorio del colegio público cercano lleva a resultados sin significancia estadística tanto para el logaritmo de alumnas matriculadas (estadístico de 0.07% sin significancia estadística) como para la dicotómica de nuevas alumnas (estadístico de 0.01% sin significancia estadística).

En el caso de las interacciones para los emparejamientos entre colegios públicos y privados, los resultados crecen en significancia estadística cuando se considera la interacción con la variable de alto

desempeño. Que el colegio focal sea alto hace que su número de matriculadas disminuya en 7.9% cuando el laboratorio de su colegio público emparejado pasa a ser de nivel 2, y en 6.4% cuando pasa a ser de nivel 3. Ambos resultados cuentan con un nivel de significancia estadística del 99%.

Una interpretación directa que se deriva de estos resultados es que las decisiones no son similares cuando se trata de hijos varones y de hijas mujeres, por lo menos en lo que respecta a la matrícula escolar. Si bien es de gran interés ahondar en lo subyacente a estos resultados, consideramos que no contamos con la información suficiente para sacar conclusiones acertadas al respecto. Es por ello que nos limitaremos a mostrar nuestros hallazgos, esperando que esto motive a investigaciones futuras que puedan profundizar al respecto.

## **5.2. Análisis de robustez**

### **5.2.1. Resultados invertidos en función al tipo de gestión**

Recordando lo mencionado en la sección 5.2 en referencia a la ausencia de endogeneidad del modelo, procederemos a analizar las herramientas escogidas como fundamentación de la robustez del modelo.

La Tabla 6 muestra los resultados de la ecuación (5). Encontramos resultados estadísticamente significativos a la relación negativa entre los laboratorios privados y los matriculados varones en el colegio público cercano. En concreto, una mejora sustancial en el laboratorio del colegio privado está vinculada a un decrecimiento del 0.08% de los matriculados varones del público focal, con un nivel de significancia estadística del 99%. Por otra parte, los resultados para la variable dependiente de la dicotómica de nuevos alumnos muestra un decrecimiento de 0.13% para los varones y 0.12% para las mujeres con un nivel de significancia estadística del 95% en ambos casos. El resultado para las alumnas matriculadas mujeres no es estadísticamente significativo. Se puede concluir que la mejora del privado perjudica al público: hay evidencias a favor de externalidades negativas y competencia.

Consideramos que esta información muestra que la facilidad económica no es el único factor en la relación de colegios públicos con privados, fortaleciendo la confianza en el modelo empleado al evidenciar la existencia de externalidades que serían aún más inverosímiles entre ambos. Si bien la magnitud de los estimadores es menor que en el caso de colegios focales privados, el hecho de que existan resultados estadísticamente significativos para colegios focales públicos evidencia la existencia de una vinculación

entre los colegios privados y los colegios públicos cercanos; vinculación en la que, en base a los hallazgos mostrados, el privado se beneficia en general del progreso del público, pero el público se ve perjudicado en general por el progreso del privado.

Una vez más, reiteramos que nuestro análisis no pretende agotar la explicación de esta relación sino mostrar que se requiere de mayores análisis en el futuro para lograr comprenderla mejor. Consideramos que nuestros hallazgos contribuyen en esta línea a la literatura vigente.

### **5.2.2. Interacciones**

Las interacciones presentadas en la Tabla 5 y explicadas en la subsección 6.1.1 también brindan información que favorece la robustez del modelo empleado.

Al presentar especificaciones más concretas del colegio focal considerado en cada regresión (ya sea en función a su número de alumnos o a su desempeño en la ECE) y brindar resultados no solo significativos estadísticamente al 99% sino también más grandes en magnitud, consideramos que presentan evidencia a favor de la validez del modelo y de la fiabilidad de los resultados.

Lo expuesto en el presente trabajo no es fruto de simples coincidencias fortuitas, sino de relaciones económicas subyacentes al mercado educativo peruano.

### **5.3. Evaluación de la metodología**

Teniendo en mente el objetivo del presente trabajo de investigación por minimizar la complejidad del modelo, maximizando la relevancia de los resultados dada la base de datos accesible, notamos que una característica saltante en este estudio es el alto valor del  $R^2$  en los resultados de las Tablas 3, 4 y 5 para la variable dependiente del logaritmo de matriculados.

Al ahondar en esta cuestión, pudimos ver que al quitar los efectos fijos evidentemente el  $R^2$  cambiaba, y los coeficientes también. La Tabla II del Apéndice muestra los resultados de las ecuaciones (6) y (7), que son similares a las (1) y (2) pero sin los efectos fijos.

Para el caso del panel i, que considera los emparejamientos de colegios públicos con privados, el signo positivo y la significancia estadística para los casos de alumnos varones se mantiene, pero aumenta al 99% y pasa de negativo a positivo para el caso de alumnas mujeres. A su vez, el  $R^2$  disminuye drásticamente,

no superando el 10% en ninguna de las especificaciones.

Para el caso del panel ii, que considera los emparejamientos de colegios públicos, el signo cambia de negativo a positivo en todas las especificaciones, aumentando además la significancia estadística hasta el 99% en el caso de alumnas mujeres y manteniendo dicha significancia elevada para el caso de alumnos varones. A su vez, el  $R^2$  disminuye drásticamente, llegando al 10% solo en el caso de la variable dependiente que toma el logaritmo de matriculados varones.

La validez de un  $R^2$  alto o bajo en la estimación depende principalmente de la naturaleza de las variables implicadas. Si bien estamos considerando decisiones humanas (matrícula escolar) que tienen inherentemente un nivel de predictibilidad bajo, el incluir efectos fijos a nivel de colegio focal a la par que empleamos como variable dependiente el logaritmo de alumnos matriculados en dicho colegio, probablemente es lo que repercute en el valor tan alto del  $R^2$ . Esto se puede apreciar con más claridad en las Tablas III y IV del Apéndice, donde mostramos los resultados de las especificaciones (1) y (2) pero considerando solo efectos fijos de año o solo efectos fijos de colegio focal, respectivamente. El valor del  $R^2$  en aquellas especificaciones que consideran solo efecto fijo de año es significativamente menor en todos los casos; mientras que el valor del  $R^2$  en aquellas especificaciones que consideran solo el efecto fijo de colegio focal es muy similar al de los resultados de las tablas 3 y 4.

Sin embargo, dada la naturaleza de nuestra investigación y la mínima posibilidad de encontrar variables que nos permitan hacer un análisis más fino, ponderamos como conveniente mantener dichos efectos fijos y centrarnos en lo que aporten los presentes resultados. Al comparar los estadísticos resultantes de la estimación que no considera los efectos fijos notamos grandes variaciones en signo y magnitud, que a su vez no tenían una interpretación plausible, y por lo tanto eran significativamente menos informativos. Es así que concluimos que la mejor forma de aprovechar los hallazgos, al no contar con un abanico más amplio de variables, es mantener la metodología aplicada matizando las interpretaciones respectivas: los estadísticos resultantes de la estimación que no considera los efectos fijos nos confirmaron que el análisis extraído de las regresiones que sí los incluyen supera ampliamente al de los resultados obtenidos de aquellas que los omiten.

## **Capítulo 6**

### **Implicancias para políticas públicas**

#### **6.1. Justificación de la relevancia de los resultados para las políticas públicas en el Perú**

Si bien el presente estudio no contempla realizar una explicación exhaustiva de las externalidades encontradas entre los colegios públicos y privados del Perú, consideramos que los resultados presentados pueden brindar información enriquecedora para la toma de decisiones en términos de políticas públicas del país.

Como se mostró en la sección 2, la educación es un ámbito en el que el Perú invierte cada vez más, y donde la presión mediática y las necesidades nacionales probablemente seguirán presionando para que se siga en esa línea. Pero no basta invertir en mayor cantidad para obtener resultados educativos, es necesario saber cómo invertir esos fondos.

El crear nuevos colegios es un buen punto de inicio, y es lo que se viene realizando hasta la actualidad. Pero estamos llegando a un punto en el que, o se realiza una inversión más estratégica, o el mercado educativo avanzará según el criterio de la producción de riqueza, y la labor del Estado como ente regulador terminará siendo obsoleta. Los datos del presente estudio, al evidenciar de una forma cuantificable la manera como se relaciona el mercado educativo privado con la provisión pública de bienes educativos, contribuyen con información para dicha inversión estratégica.

#### **6.2. Rol clave de la focalización**

Tanto las externalidades positivas que se dan de los colegios públicos a los colegios privados (Tabla 3 del presente estudio) como la competencia entre colegios públicos (Tabla 4 del presente estudio) brindan información relevante a la hora de decidir dónde invertir en un colegio nuevo.

Las externalidades positivas evidencian que cuando un colegio público mejora su nivel de infraestructura, el colegio privado recibe más alumnos. Esto es señal de que se requiere de un análisis más focalizado para conocer la realidad de cada espacio territorial antes de realizar inversiones en la infraestructura de los colegios.

En una zona donde ya existe un colegio privado, y hay muchos niños que no acceden a educación, lo mejor será invertir en un colegio público que se encuentre a menos de 1000m de distancia de dicho colegio

privado. Esto se debe a que la educación pública no tiene como naturaleza la competición con el sector privado, sino el apoyo a las personas que no pueden acceder por sus propios medios al sistema educativo. De esta forma, se aprovecharían las externalidades mencionadas para mejorar el nivel educativo de la zona (esto sujeto a una causalidad subyacente al primer canal mencionado en la sección de resultados, que habría que comprobar).

En una zona donde ya existe un colegio privado, y hay pocos niños que no acceden a educación, lo mejor será no invertir en un colegio público que se encuentre a menos de 1000m de distancia de dicho colegio privado (especialmente si es grande en función a sus alumnos varones, basándonos en los resultados de la Tabla 5 del presente estudio). Probablemente sea más eficiente acudir a otro tipo de políticas públicas para apoyar a esas familias en concreto, en lugar de realizar inversiones masivas que finalmente no brindarán mayores resultados.

Por otro lado, la competencia entre colegios públicos cercanos nos evidencia que no es conveniente poner muchos colegios públicos cerca. Esto puede parecer lo más sencillo: si ya tienes un colegio público en una determinada zona, la UGEL correspondiente cuenta con la estructura logística necesaria, por lo que tener un colegio más en su jurisdicción es más barato que invertir en un colegio alejado que implique coordinaciones y gestiones más complejas (en el primer caso, el costo fijo se reduce).

Sin embargo, recordando la naturaleza de la educación pública, su función no está en ser barata, sino en brindar educación buena a aquellos que no pueden acceder a ella por sus propios medios. Invertir en un colegio público que tiene a menos de 1000m a otro colegio público (especialmente si es grande en función a sus alumnos varones en secundaria, basándonos en los resultados de la Tabla I del Apéndice del presente estudio) terminará ocasionando competencia entre ambos; es mejor tener pocos colegios bien implementados y bien ubicados, que muchos colegios de poca calidad y concentrados.

Por último, la inversión en educación para mejorar el panorama en alumnas mujeres se mantiene como una incógnita importante. El presente estudio evidencia que las decisiones que toman las familias no son iguales en relación a los hijos varones y a las hijas mujeres (Tablas 3, 4 y 5 del presente estudio), pero en ninguno de los casos presentados acertamos con encontrar una inversión que arroje resultados positivos y estadísticamente significativos. Tal vez el aspecto más preocupante a resaltar es que invertir en colegios públicos cuando se tiene cerca un privado con alto desempeño académico es perjudicial para las matriculadas

mujeres de dicho colegio privado. No queda claro que sea porque se las matricula en un colegio público, o porque abandonan los estudios.

Lo que sí queda claro es que las políticas públicas en educación deberán considerar una focalización más fina en cuanto al sexo de los estudiantes que pretenden beneficiar con sus inversiones. El presente trabajo evidencia que este no es un aspecto que se pueda pasar por alto.

### **6.3. Importancia de la mejora en la calidad de la información relacionada al sector educativo**

Las externalidades positivas existentes entre colegios públicos y privados (Tabla 3 del presente estudio) nos llevan a evaluar la posibilidad de que la privatización por defecto esté afectando negativamente al mercado educativo. Tal vez sea conveniente que se cuente con mejores mecanismos de señalización en cuanto a la calidad educativa de las diversas instituciones, de manera que las familias puedan tomar decisiones guiadas por una información más completa. Se tiene la información de las pruebas de la ECE, pero esta información no llega con facilidad al ciudadano promedio que es quien la necesita.

Finalmente, un último aporte que consideramos que este estudio brinda a la toma de decisiones en políticas públicas es la clara necesidad de contar con información más específica y detallada. El análisis brindado ha procurado ser riguroso en su metodología, pero los resultados podrían ser mucho más certeros y enriquecedores si se contara con bases de datos que permitan realizar un análisis comparativo con lo estudiado en otros países en temas de educación.



## Conclusiones

La relación entre los colegios públicos y privados en el Perú se mantiene como una puerta abierta a la investigación económica, que en los últimos años va cobrando creciente importancia debido a la posibilidad de experimentar una privatización por defecto del mercado educativo peruano. El presente trabajo, al utilizar la metodología de efectos fijos, logra una aproximación a esta casuística, encontrando externalidades entre colegios cercanos entre sí.

Analizando una muestra de todos los colegios públicos y privados del Perú para el período del 2007 al 2016, tomamos como variable independiente un índice del nivel de implementación de los laboratorios de los colegios públicos y como variable dependiente la cantidad de matriculados de un colegio cercano. Los resultados muestran que una mejora sustancial del índice se relaciona con un aumento del 0.28% de los alumnos varones matriculados en un colegio privado cercano, aumentando en por lo menos 4% cuando dicho colegio focal es grande en comparación a sus pares.

Por otro lado, dicha mejora está a su vez vinculada a un decrecimiento del 0.37% de los matriculados varones de un colegio público cercano.

Atribuimos estos resultados a la existencia de externalidades positivas de los colegios públicos a los colegios privados cercanos, y a la competencia existente entre los colegios públicos cercanos entre sí. De esta forma concluimos que nos encontramos frente a un nuevo panorama de posibilidades para la investigación económica en beneficio del Perú.

Dichas investigaciones futuras podrían ahondar en la causa de los diferentes resultados para varones y para mujeres, el radio de alcance geográfico de las externalidades positivas o de la competencia, un análisis para el efecto fijo por colegio independiente, las cualidades de las familias que optan por colegios privados, o la información sobre bibliotecas, laboratorios de inglés o aulas de cómputo de los colegios.

Por otro lado, un aspecto final a considerar es la posibilidad de que las familias busquen algo más que desempeño académico en la educación. Es posible que la elección de colegios privados no responda solo a un interés por mejores notas, sino a una búsqueda de educación afectiva y social. Mayores investigaciones al respecto brindarían luces al sector público sobre otras brechas a trabajar.

En definitiva, un factor clave para profundizar en esta rama de la literatura es seguir mejorando la calidad y cantidad de bases de datos levantadas en torno a ella. Es esta una condición clave para que las

investigaciones brinden información que efectivamente guíe una mejor toma de decisiones en el país.

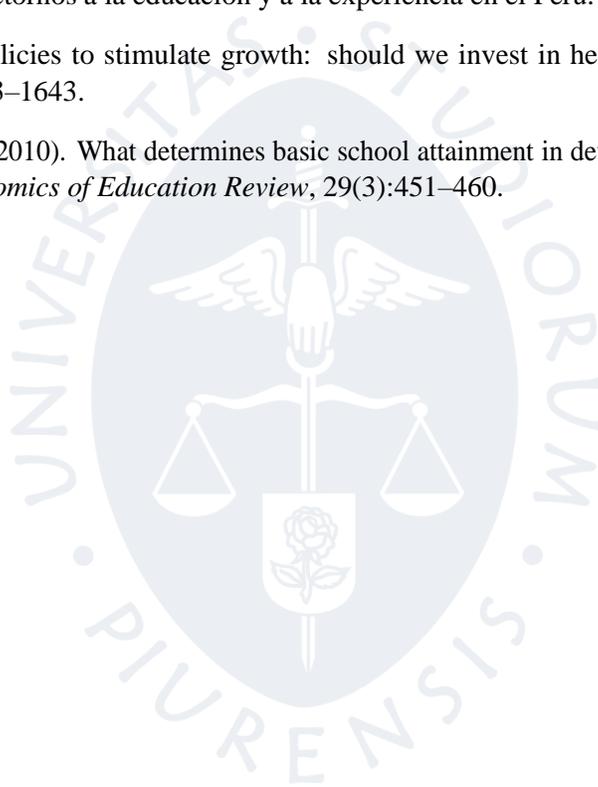


## Lista de Referencias

- Agüero, J. M. (2016). Evaluación de impacto de la jornada escolar completa.
- Baca, J. F., Peschiera, J., and Mesones, J. (2014). The impact of public expenditures in education, health, and infrastructure on economic growth and income distribution in Perú. *BID*.
- Balarín, M. (2015). Las múltiples formas y efectos de la participación del sector privado en la educación. *FORGE*.
- Balarín, M. (2016). La privatización por defecto y el surgimiento de las escuelas privadas de bajo costo en el Perú. ¿cuáles son sus consecuencias? *Asociación de Sociología de la Educación*.
- Balarín, M., Kitmang, J., Ñopo, H., and Rodríguez, M. F. (2018). *Mercado privado, consecuencias públicas. Los servicios educativos de provisión privada en el Perú*.
- Beuermann, D. W., Cristia, J., Cueto, S., Malamud, O., and 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.
- BID (2017). Aprender mejor: Políticas públicas para el desarrollo de habilidades. *Washington DC: Publicaciones del BID*.
- Black, S. E. (1999). Do better schools matter? Parental valuation of elementary education. *The Quarterly Journal of Economics*, 114(2):577–599.
- Bowen, W. M. and Qian, H. (2017). State spending for higher education: Does it improve economic performance? *Regional Science Policy and Practice*, 9(1):7–23.
- Branham, D. (2004). The wise man builds his house upon the rock: The effects of inadequate school building infrastructure on student attendance. *Social Science Quarterly*, 85(5):1112–1128.
- Calle, F. A. (2017). ¿Maldición o bendición de los recursos naturales?: los efectos del canon minero sobre los resultados educativos en los distritos del Perú. *Pontificia Universidad Católica del Perú*.
- Card, D. and Krueger, A. B. (1992). Does school quality matter? Returns to education and the characteristics of public schools in the United States. *Journal of Political Economy*, 100(1):1–40.
- Chiappori, P., Salanié, B., and Weiss, Y. (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review*, 107(8):2109–67.
- Cuenca, R. (2013). La escuela pública en Lima Metropolitana. ¿una institución en extinción? *Sociedad de Investigación Educativa Peruana*.
- Cuenca, R., Carrillo, S., de los Ríos, C., Reátegui, L., and Ortiz, G. (2017). La calidad y equidad de la educación secundaria en el Perú. *Instituto de Estudios Peruanos*.

- Cuenca, R., León, L., Reátegui, A., and Oré, S. (2019). Itinerarios de la educación privada en Perú. *Educação and Sociedade*, 40.
- Cuesta, A., Glewwe, P., and Krause, B. (2016). School infrastructure and educational outcomes: a literature review, with special reference to Latin America. *Economía*, 17(1):95–130.
- Dasso, R., Fernandez, F., and Ñopo, H. (2015). Electrification and educational outcomes in rural Peru. *IZA Discussion Paper*.
- De la Croix, D. and Doepke, M. (2009). To segregate or to integrate: Education politics and democracy. *The Review of Economic Studies*, 76(2):597–628.
- Dizon-Ross, R. (2019). Parents' beliefs about their children's academic ability: Implications for educational investments. *American Economic Review*, 109(8):2728–65.
- Fontdevila, C., Marius, P., Balarín, M., and Rodríguez, M. F. (2018). Educación privada de 'bajo coste' en el Perú: un enfoque desde la calidad. *Internacional de la Educación*.
- Guadalupe, C., León, J., Rodríguez, J. S., and Vargas, S. (2017). Estado de la educación en el Perú: análisis y perspectivas de la educación básica. *FORGE*.
- Hanushek, E. A. (1996). Measuring investment in education. *Journal of Economic Perspectives*, 10(4):9–30.
- Harrison, J. S. and Wicks, A. C. (2013). Stakeholder theory, value, and firm performance. *Business Ethics Quarterly*, 23(1):97–124.
- Jackson, C. K., Johnson, R. C., and Persico, C. (2015). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *National Bureau of Economic Research*, 1:2016.
- Keller, K. (2006). Investment in primary, secondary, and higher education and the effects on economic growth. *Contemporary Economic Policy*, 24(1):18–34.
- Konstantopoulos, S. and Borman, G. D. (2011). Family background and school effects on student achievement: A multilevel analysis of the Coleman data. *Teachers College Record*, 113(1):97–132.
- Krskova, H. and Baumann, C. (2017). School discipline, investment, competitiveness and mediating educational performance. *International Journal of Educational Management*.
- Málaga, O. and Velásquez, D. (2020). Reflexiones en torno a los colegios de bajo costo en el Perú. *Educación*, 26(2):139–146.
- Manrique, G., Contreras, A., and Guerrero, N. (2016). Evaluación del efecto de transferencias de canon minero en los resultados educativos: análisis a nivel de instituciones educativas de las regiones de Arequipa, Moquegua y Tacna. *CIES*.
- Murillo, F. J. y Román, M. (2011). School infrastructure and resources do matter: analysis of the incidence of school resources on the performance of Latin American students. *School Effectiveness and School Improvement*, 22(1):29–50.

- Psacharopoulos, G. and Patrinos, H. A. (2018). Returns to investment in education: a decennial review of the global literature. *Education Economics*, 26(5):445–458.
- Reinikka, R. and Smith, N. (2004). *Public expenditure tracking surveys in education*.
- Rodríguez-Sosa, J. (2013). La investigación nacional en educación 2007-2011. Balance y agenda. *Propósitos y Representaciones*, 1(1):131–172.
- Romero, M., Sandefur, J., and Sandholtz, W. A. (2020). Outsourcing education: Experimental evidence from Liberia. *American Economic Review*, 110(2):364–400.
- Saavedra, J. (1999). Los retornos a la educación y a la experiencia en el Perú: 1985-1997. *GRADE*.
- Webber, D. J. (2002). Policies to stimulate growth: should we invest in health or education? *Applied Economics*, 34(13):1633–1643.
- Zhao, M. and Glewwe, P. (2010). What determines basic school attainment in developing countries? evidence from rural China. *Economics of Education Review*, 29(3):451–460.





## Figuras

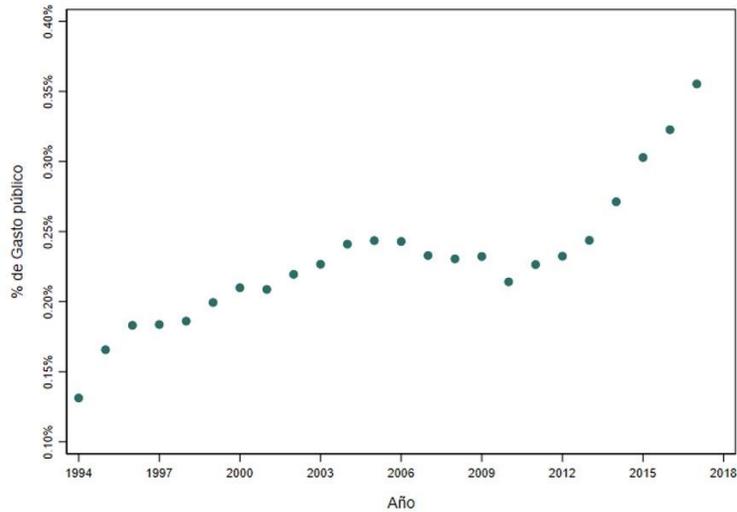


Figura 1: Base de datos obtenida del INEI, elaboración propia. Se muestra el valor del gasto público nacional en educación como porcentaje del PBI desde 1994 hasta 2017.

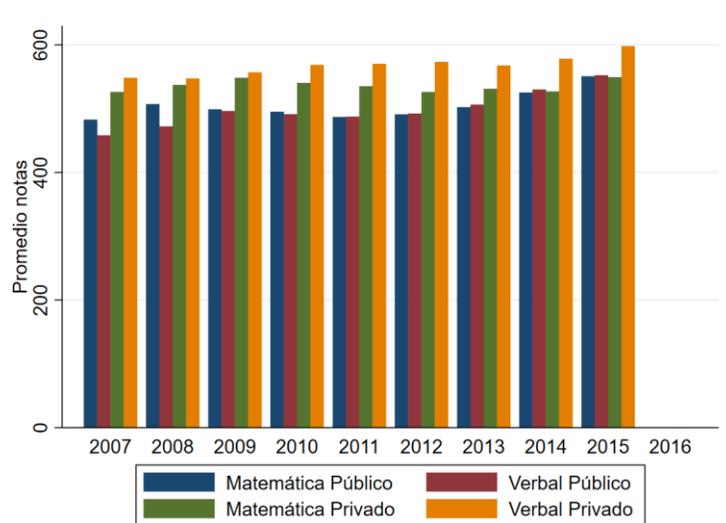


Figura 2: Base de datos obtenida de ESCALE, elaboración propia. Se muestra el promedio de las notas para los respectivos colegios en los respectivos cursos de la Evaluación Censal de Estudiantes.

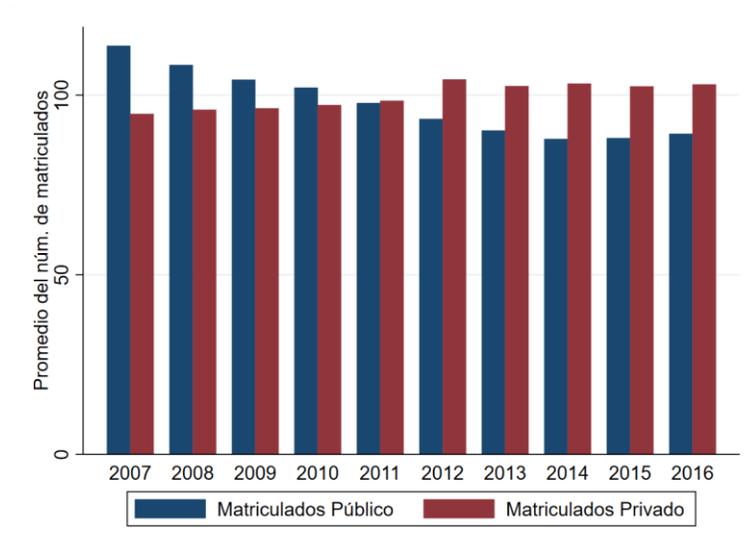


Figura 3: Base de datos obtenida de ESCALE, elaboración propia. Se muestra el valor promedio de la cantidad de alumnos matriculados en todos los colegios públicos y privados del Perú.

## Tablas

Tabla 1: Estadísticos sumarios de las variables

|  | Obs.    | Media | Std. Dv. | Mín. | Máx. |
|--|---------|-------|----------|------|------|
| Panel A: Variables de colegios públicos y privados |         |       |          |      |      |
| Índice del laboratorio                             | 491,910 | 0.58  | 0.94     | 0    | 3    |
| Varones matriculados en secundaria                 | 491,790 | 25.18 | 89.85    | 0    | 4570 |
| Mujeres matriculadas en secundaria                 | 491,790 | 29.17 | 87.15    | 0    | 3194 |
| Dicotómica de nuevos alumnos varones               | 491,790 | 0.55  | 0.50     | 0    | 1    |
| Dicotómica de nuevas alumnas mujeres               | 491,790 | 0.55  | 0.50     | 0    | 1    |
| Panel B: Variables solo de colegios públicos       |         |       |          |      |      |
| Índice del laboratorio                             | 369,659 | 0.50  | 0.86     | 0    | 3    |
| Varones matriculados en secundaria                 | 369,617 | 25.67 | 97.58    | 0    | 4570 |
| Mujeres matriculadas en secundaria                 | 369,617 | 30.19 | 95.65    | 0    | 3194 |
| Dicotómica de nuevos alumnos varones               | 369,617 | 0.52  | 0.50     | 0    | 1    |
| Dicotómica de nuevas alumnas mujeres               | 369,617 | 0.52  | 0.50     | 0    | 1    |
| Colegio grande                                     | 369,617 | 0.46  | 0.50     | 0    | 1    |
| Colegio grande para varones                        | 369,617 | 0.46  | 0.50     | 0    | 1    |
| Alto desempeño                                     | 369,617 | 0.41  | 0.49     | 0    | 1    |
| Alto desempeño en matemática                       | 369,617 | 0.43  | 0.50     | 0    | 1    |
| Alto desempeño en lengua                           | 369,617 | 0.44  | 0.50     | 0    | 1    |
| Panel C: Variables solo de colegios privados       |         |       |          |      |      |
| Índice del laboratorio                             | 122,251 | 0.81  | 1.10     | 0    | 3    |
| Varones matriculados en secundaria                 | 122,173 | 23.71 | 60.68    | 0    | 1911 |
| Mujeres matriculadas en secundaria                 | 122,173 | 26.07 | 53.68    | 0    | 1668 |
| Dicotómica de nuevos alumnos varones               | 122,173 | 0.66  | 0.47     | 0    | 1    |
| Dicotómica de nuevas alumnas mujeres               | 122,173 | 0.66  | 0.48     | 0    | 1    |
| Colegio grande                                     | 122,173 | 0.59  | 0.49     | 0    | 1    |
| Colegio grande para varones                        | 122,173 | 0.59  | 0.49     | 0    | 1    |
| Alto desempeño                                     | 122,173 | 0.29  | 0.45     | 0    | 1    |
| Alto desempeño en matemática                       | 122,173 | 0.29  | 0.45     | 0    | 1    |
| Alto desempeño en lengua                           | 122,173 | 0.44  | 0.50     | 0    | 1    |

Para el período 2016 y 2018 en el caso de las variables de alto desempeño, y 2007 a 2016 para el resto de variables. La muestra abarca todo el Perú.

**Tabla 2:** Estadísticos sumarios de los emparejamientos

| Variable                         | Total de colegios | Total de colegios focales | Total de independientes | Media de emparejamientos | Mín. | Máx. |
|----------------------------------|-------------------|---------------------------|-------------------------|--------------------------|------|------|
| Universo                         | 55,412            | -                         | -                       | -                        | -    | -    |
| Emparejamiento Público - Privado | 20,567            | 14,870                    | 5,697                   | 59                       | 1    | 192  |
| Emparejamiento Público - Público | 9,507             | 9,507                     | 9,507                   | 52                       | 1    | 180  |

Las observaciones están a nivel de colegio público emparejado al colegio focal (privado o público) a un radio de distancia de 1000m para todos los colegios de la muestra. Universo es el total de colegios aptos para ser emparejados. Emparejamiento Público - Privado está a nivel de colegio público emparejado al colegio focal privado dentro de un radio de 1000m. Emparejamiento Público - Público está a nivel de colegio público emparejado al colegio focal público dentro de un radio de 1000m.

**Tabla 3:** Resultados para colegios privados focales emparejados con colegios públicos, considerando como variable explicativa los laboratorios de ciencias de los colegios públicos

|                           | Logaritmo de matriculados en secundaria |                     | Dicotómica de nuevos alumnos en primer grado de primaria |                    |
|---------------------------|---|---------------------|--|--------------------|
|                           | Varones                                 | Mujeres             | Varones  | Mujeres            |
| Índice del laboratorio    | 0.00285 ***<br>(6.30)                   | -0.0018*<br>(-1.86) | 0.0019 ***<br>(4.24)                                     | 0.000675<br>(1.45) |
| Efecto Fijo año           | Sí                                      | Sí                  | Sí   | Sí                 |
| Efecto Fijo colegio focal | Sí                                      | Sí                  | Sí   | Sí                 |
| R <sup>2</sup>            | 0.98                                    | 0.88                | 0.31   | 0.31               |
| N. clusters (distrito)    | 544                                     | 544                 | 544  | 544                |
| N. muestra                | 535,136                                 | 535,136             | 535,136  | 535,136            |

Las observaciones están a nivel de colegio público emparejado al colegio focal privado dentro de un radio de distancia de 1000m. El modelo estima la Eq.(1) incluyendo efectos fijos de año y colegio focal privado. El estadístico t basado en errores estándar clusterizados por distrito está en paréntesis. \*\*\*, \*\*, \* hacen referencia a niveles de significancia estadística de 1%, 5% y 10%, respectivamente.

**Tabla 4:** Resultados para colegios públicos focales emparejados con colegios públicos, considerando como variable explicativa los laboratorios de ciencias de los colegios públicos

|                           | Logaritmo de matriculados en secundaria |                      | Dicotómica de nuevos alumnos en primer grado de primaria |                      |
|---------------------------|---|----------------------|--|----------------------|
|                           | Varones                                 | Mujeres              | Varones  | Mujeres              |
| Índice del laboratorio    | -0.00371 ***<br>(-5.72)                 | -0.000762<br>(-0.66) | -0.00167 **<br>(-2.18)                                   | -0.000177<br>(-0.26) |
| Efecto Fijo año           | Sí                                      | Sí                   | Sí   | Sí                   |
| Efecto Fijo colegio focal | Sí                                      | Sí                   | Sí   | Sí                   |
| R <sup>2</sup>            | 0.99                                    | 0.95                 | 0.29   | 0.28                 |
| N. clusters (distrito)    | 1,155                                   | 1,155                | 1,155  | 1,155                |
| N. muestra                | 240,590                                 | 240,590              | 240,590  | 240,590              |

Las observaciones están a nivel de colegio público emparejado al colegio focal público dentro de un radio de distancia de 1000m. El modelo estima la Eq.(2) incluyendo efectos fijos de año y colegio focal público. El estadístico t basado en errores estándar clusterizados por distrito está en paréntesis. \*\*\*, \*\*, \* hacen referencia a niveles de significancia estadística de 1%, 5% y 10%, respectivamente.

**Tabla 5:** Interacciones entre colegios públicos y colegios privados

|   | Logaritmo de matriculados<br>en secundaria |                        |
|---|--|------------------------|
|   | Varones                                    | Mujeres                |
| Índice del Laboratorio si el nivel es 1 × Colegio grande para varones | 0.0432 ***<br>(11.08)                      |                        |
| Índice del Laboratorio si el nivel es 2 × Colegio grande para varones | 0.0731 ***<br>(5.91)                       |                        |
| Índice del Laboratorio si el nivel es 3 × Colegio grande para varones | 0.091 ***<br>(12.14)                       |                        |
| Índice del Laboratorio si el nivel es 1 × Alto desempeño              |  | 0.0037<br>(0.43)       |
| Índice del Laboratorio si el nivel es 2 × Alto desempeño              |  | -0.0798 ***<br>(-3.64) |
| Índice del Laboratorio si el nivel es 3 × Alto desempeño              |  | -0.0646 ***<br>(-7.29) |
| Efecto Fijo año   | Sí   | Sí                     |
| Efecto Fijo colegio focal   | Sí   | Sí                     |
| R <sup>2</sup>  | 0.98                                       | 0.88                   |
| N. clusters (distrito)  | 544  | 544                    |
| N. muestra  | 534,645                                    | 534,645                |

Las observaciones están a nivel de colegio público emparejado al colegio focal privado dentro de un radio de distancia de 1000m. El modelo estima la Eq.(3) y la Eq.(4) incluyendo efectos fijos de año y colegio focal privado. Las interacciones se realizan considerando variables dicotómicas que se activan para el caso en el que el colegio focal tiene más alumnos hombres matriculados que sus pares, o el caso en el que el colegio focal tiene alto desempeño. El estadístico t basado en errores estándar clusterizados por distrito está en paréntesis. \*\*\*, \*\*, \* hacen referencia a niveles de significancia estadística de 1%, 5% y 10%, respectivamente.

**Tabla 6:** Regresiones de robustez. Resultados para colegios públicos focales emparejados con colegios privados, considerando como variable explicativa los laboratorios de ciencias de los colegios privados

|                           | Logaritmo de matriculados en secundaria |                     | Dicotómica de nuevos alumnos en primer grado de primaria |                        |
|---------------------------|---|---------------------|--|------------------------|
|                           | Varones                                 | Mujeres             | Varones  | Mujeres                |
| Índice del laboratorio    | -0.000887 ***<br>(-3.01)                | -0.00032<br>(-0.36) | -0.0013 **<br>(-2.39)                                    | -0.00124 **<br>(-2.19) |
| Efecto Fijo año           | Sí                                      | Sí                  | Sí   | Sí                     |
| Efecto Fijo colegio focal | Sí                                      | Sí                  | Sí   | Sí                     |
| R <sup>2</sup>            | 0.99                                    | 0.94                | 0.33   | 0.31                   |
| N. clusters (ubigeo)      | 522                                     | 522                 | 522  | 522                    |
| N. muestra                | 534,995                                 | 534,995             | 534,995  | 534,995                |

Las observaciones están a nivel de colegio privado emparejado al colegio focal público dentro de un radio de distancia de 1,000 metros. El modelo estima la Eq.(5) incluyendo efectos fijos de año y colegio focal público. El estadístico t basado en errores estándar clusterizados por distrito está en paréntesis. \*\*\*, \*\*, \* hacen referencia a niveles de significancia estadística de 1%, 5% y 10%, respectivamente.



## Apéndice

TABLA I Interacciones entre colegios públicos

|   | Logaritmo de matriculados<br>varones en secundaria |                     |
|---|--|---------------------|
| Índice del Laboratorio si el nivel es 1 × Colegio grande para varones en secundaria | 0.0118 **  | (2.10)              |
| Índice del Laboratorio si el nivel es 2 × Colegio grande para varones en secundaria | -0.0301 *  | (-1.80)             |
| Índice del Laboratorio si el nivel es 3 × Colegio grande para varones en secundaria | -0.0701 ***  | (-9.39)             |
| Índice del Laboratorio si el nivel es 1 × Alto desempeño en matemática              |  | -0.00418<br>(-2.03) |
| Índice del Laboratorio si el nivel es 2 × Alto desempeño en matemática              |  | 0.0204 **<br>(2.11) |
| Índice del Laboratorio si el nivel es 3 × Alto desempeño en matemática              |  | 0.0104 **<br>(2.5)  |
| Efecto Fijo año   | Sí   | Sí                  |
| Efecto Fijo colegio focal   | Sí   | Sí                  |
| R <sup>2</sup>  | 0.99   | 0.99                |
| N. clusters (distrito)  | 1,155  | 1,155               |
| N. muestra  | 240,509  | 240,509             |

Las observaciones están a nivel de colegio público emparejado al colegio focal público dentro de un radio de distancia de 1000m. El modelo estima la Eq.(3) y la Eq. (4), pero con todas las variables de colegios públicos incluyendo efectos fijos de año y colegio focal público, y considerando las variables dependientes solo para varones. El estadístico t basado en errores estándar clusterizados por distrito está en paréntesis. \*\*\*, \*\*, \* hacen referencia a niveles de significancia estadística de 1%, 5% y 10%, respectivamente.

**TABLA II** Evaluación de la metodología sin considerar efectos fijos

|   | Logaritmo de matriculados<br>en secundaria |                      | Dicotómica de nuevos<br>alumnos en primer grado<br>de primaria |                       |
|---|--|----------------------|--|-----------------------|
|   | Varones                                    | Mujeres              | Varones  | Mujeres               |
| i. Colegio público con colegio focal privado  |  |                      |  |                       |
| Índice del laboratorio                        | 0.396 ***<br>(18.35)                       | 0.249 ***<br>(17.46) | 0.0541 ***<br>(16.09)  | 0.054 ***<br>(15.93)  |
| R <sup>2</sup>                                | 0.09                                       | 0.05                 | 0.02   | 0.02                  |
| N. clusters (distrito)                        | 557  | 557                  | 557  | 557                   |
| N. muestra                                    | 535,305                                    | 535,305              | 535,305  | 535,305               |
| ii. Colegio público con colegio focal público |  |                      |  |                       |
| Índice del laboratorio                        | 0.559 ***<br>(36.81)                       | 0.382 ***<br>(20.96) | 0.067 ***<br>(37.97)   | 0.0655 ***<br>(38.13) |
| R <sup>2</sup>                                | 0.10                                       | 0.08                 | 0.03   | 0.03                  |
| N. clusters (distrito)                        | 1168                                       | 1168                 | 1168   | 1168                  |
| N. muestra                                    | 240,809                                    | 240,809              | 240,809  | 240,809               |
| Efecto Fijo año                               | No   | No                   | No   | No                    |
| Efecto Fijo colegio focal                     | No   | No                   | No   | No                    |

Las observaciones están a nivel de colegio público emparejado al colegio focal (público o privado) dentro de un radio de distancia de 1000m. El modelo estima la Eq.(6) y la Eq.(7), que no incluyen efectos fijos de año ni colegio focal. El estadístico t basado en errores estándar clusterizados por distrito está en paréntesis. \*\*\*, \*\*, \* hacen referencia a niveles de significancia estadística de 1%, 5% y 10%, respectivamente.

**TABLA III** Evaluación de la metodología considerando solo el efecto fijo para año

|   | Logaritmo de matriculados<br>en secundaria |                      | Dicotómica de nuevos<br>alumnos en primer grado<br>de primaria |                       |
|---|--|----------------------|--|-----------------------|
|   | Varones                                    | Mujeres              | Varones  | Mujeres               |
| i. Colegio público con colegio focal privado  |  |                      |  |                       |
| Índice del laboratorio                        | 0.453 ***<br>(19.54)                       | 0.279 ***<br>(18.09) | 0.0656 ***<br>(18.5)   | 0.0654 ***<br>(18.46) |
| R <sup>2</sup>                                | 0.10                                       | 0.05                 | 0.03   | 0.03                  |
| N. clusters (distrito)                        | 557  | 557                  | 557  | 557                   |
| N. muestra                                    | 535,305                                    | 535,305              | 535,305  | 535,305               |
| ii. Colegio público con colegio focal público |  |                      |  |                       |
| Índice del laboratorio                        | 0.638 ***<br>(36.98)                       | 0.456 ***<br>(22.24) | 0.0766 ***<br>(37.87)  | 0.0764 ***<br>(36.69) |
| R <sup>2</sup>                                | 0.11                                       | 0.10                 | 0.04   | 0.04                  |
| N. clusters (distrito)                        | 1168                                       | 1168                 | 1168   | 1168                  |
| N. muestra                                    | 240,809                                    | 240,809              | 240,809  | 240,809               |
| Efecto Fijo año                               | Sí   | Sí                   | Sí   | Sí                    |
| Efecto Fijo colegio focal                     | No   | No                   | No   | No                    |

Las observaciones están a nivel de colegio público emparejado al colegio focal (público o privado) dentro de un radio de distancia de 1000m. El modelo estima la Eq.(6) y la Eq.(7), pero incluyendo sólo efectos fijos de año. El estadístico t basado en errores estándar clusterizados por distrito está en paréntesis. \*\*\*, \*\*, \* hacen referencia a niveles de significancia estadística de 1%, 5% y 10%, respectivamente.

**TABLA IV** Evaluación de la metodología considerando solo el efecto fijo para colegios focales

|   | Logaritmo de matriculados en secundaria |                        | Dicotómica de nuevos alumnos en primer grado de primaria |                         |
|---|---|------------------------|--|-------------------------|
|   | Varones                                 | Mujeres                | Varones  | Mujeres                 |
| i. Colegio público con colegio focal privado  |   |                        |  |                         |
| Índice del laboratorio                        | 0.00638 ***<br>(8.16)                   | 0.0107 ***<br>(6.37)   | -0.00437 ***<br>(-8.11)                                  | -0.00533 ***<br>(-9.71) |
| R <sup>2</sup>                                | 0.98                                    | 0.87                   | 0.31   | 0.31                    |
| N. clusters (distrito)                        | 544                                     | 544                    | 544  | 544                     |
| N. muestra                                    | 535,136                                 | 535,136                | 535,136  | 535,136                 |
| ii. Colegio público con colegio focal público |   |                        |  |                         |
| Índice del laboratorio                        | -0.0057 ***<br>(-6.27)                  | -0.0236 ***<br>(-9.62) | -0.0019 *<br>(-1.88)                                     | -0.00245 **<br>(-2.25)  |
| R <sup>2</sup>                                | 0.99                                    | 0.95                   | 0.29   | 0.27                    |
| N. clusters (distrito)                        | 1155                                    | 1155                   | 1155   | 1155                    |
| N. muestra                                    | 240,590                                 | 240,590                | 240,590  | 240,590                 |
| Efecto Fijo año                               | No                                      | No                     | No   | No                      |
| Efecto Fijo colegio focal                     | Sí                                      | Sí                     | Sí   | Sí                      |

Las observaciones están a nivel de colegio público emparejado al colegio focal (público o privado) dentro de un radio de distancia de 1000m. El modelo estima la Eq.(6) y la Eq.(7), pero incluyendo solo efectos fijos de colegio focal. El estadístico t basado en errores estándar clusterizados por distrito está en paréntesis. \*\*\*, \*\*, \* hacen referencia a niveles de significancia estadística de 1%, 5% y 10%, respectivamente.