



**UNIVERSIDAD  
DEL PACÍFICO**

**Escuela de  
Postgrado**

**“¿MAYOR CONECTIVIDAD ENTRE ESCUELAS REDUCE LA  
DESERCIÓN ESCOLAR? UN ANÁLISIS GEOESPACIAL EN LA  
TRANSICIÓN EDUCATIVA PRIMARIA – SECUNDARIA EN EL  
PERÚ”**

**Trabajo de Investigación  
presentado para optar al Grado Académico de  
Magíster en Economía**

**Presentado por  
Sr. Jesús Manuel Gutiérrez Andrade**

**Asesor: Ph.D. Juan Francisco Castro Carlín**

**[0000-0003-2032-0565](tel:0000-0003-2032-0565)**

**Lima, abril de 2022**

A mi familia, por estar conmigo en todo momento y apoyarme en todas y cada una de mis metas. A mi hermana Angelina, por enseñarme lo que es la valentía y el deseo de seguir adelante. A mis amigos, quienes representaron un gran soporte emocional en esta etapa de mi vida. A Lucia, por su constante ánimo y a Morocha, por acompañarme en las largas noches de desvelo.

Agradezco también a Juan Francisco Castro por su guía y sus valiosos comentarios para la elaboración de esta investigación.

## **Resumen ejecutivo**

El presente estudio busca analizar si la conectividad terrestre entre las instituciones educativas de nivel primaria y secundaria en zonas rurales contribuye con la deserción escolar durante la transición entre estos niveles. Para ello, se propone utilizar un estimador de efectos fijos a nivel familia que explota la variación en la disponibilidad de escuelas de nivel secundaria alrededor de las primarias donde egresan los hermanos de cada familia en la muestra. A través del uso de datos administrativos y georreferenciados, se observa que egresar de escuelas primaria no conectadas a una escuela secundaria en un radio de 10 y 30 minutos incrementa la probabilidad de no transitar entre dichos niveles en 2 y 3 puntos porcentuales, respectivamente, cuando se controla por características de los estudiantes. Estos resultados son robustos a diferentes muestras, diferentes radios de conectividad y diferentes especificaciones del modelo empleado. El estudio contribuye a la literatura en tanto examina la problemática de la deserción escolar desde la perspectiva de la oferta. Además, contribuye con el debate de políticas públicas enfocadas al problema de acceso a la educación secundaria debido a la dispersión geográfica y el costo de búsqueda de escuelas.

## Contenido

Índice de figuras.....	iv
Índice de tablas .....	v
Índice de anexos .....	vi
Capítulo I. Introducción .....	1
Capítulo II. La deserción escolar en el Perú .....	4
Capítulo III. Distribución de la oferta de servicios educativos en el Perú.....	7
Capítulo IV. Revisión de literatura .....	12
Capítulo V. Metodología.....	15
5.1. Datos y descripción de la muestra.....	19
5.2. Estrategia empírica .....	24
Capítulo VI. Resultados .....	25
6.1. Conectividad y deserción escolar .....	25
6.3. Pruebas de robustez .....	28
6.4. Efectos heterogéneos .....	31
Capítulo VII. Recomendaciones de política .....	32
Capítulo VIII. Conclusiones .....	34
Bibliografía.....	37
Anexos .....	42

## Índice de figuras

<b>Figura 1. Tasa de deserción escolar interanual según nivel educativo .....</b>	<b>4</b>
<b>Figura 2. Tasa de deserción escolar interanual según área .....</b>	<b>5</b>
<b>Figura 3. Tasa de deserción escolar interanual según grado .....</b>	<b>6</b>
<b>Figura 4. Tasa de deserción escolar 6to de primaria según gradiente de ruralidad .....</b>	<b>7</b>
<b>Figura 5. Número de instituciones educativas según nivel educativo (2007 – 2019) .....</b>	<b>9</b>
<b>Figura 6. Distribución geográfica de instituciones educativas (2019).....</b>	<b>10</b>
<b>Figura 7. Deserción escolar y el número de escuelas por año y por nivel.....</b>	<b>11</b>
<b>Figura 8. Distancia entre dos puntos ante la presencia de relieve .....</b>	<b>15</b>
<b>Figura 9. Áreas de influencia en un radio de 60 minutos.....</b>	<b>17</b>
<b>Figura 10. Expansión de la oferta de secundarias y conectividad .....</b>	<b>18</b>
<b>Figura 11. Tamaño del efecto según corte de radio .....</b>	<b>29</b>

## Índice de tablas

<b>Tabla 1. Servicios educativos según ámbito y gestión (2019)</b> .....	8
<b>Tabla 2. Estadísticos descriptivos (Muestra total)</b> .....	22
<b>Tabla 3. Estadísticos descriptivos (Muestra de hermanos)</b> .....	23
<b>Tabla 4. Conectividad y deserción escolar</b> .....	27
<b>Tabla 5. Distancia a la secundaria más cercana y deserción escolar</b> .....	30
<b>Tabla 6. Efectos heterogéneos a nivel de sexo</b> .....	32

## Índice de anexos

<b>Anexo 1. Distribución del tiempo hacia la secundaria más cercana .....</b>	<b>42</b>
<b>Anexo 2. Conectividad y deserción considerando la huelga de docentes 2017 .....</b>	<b>43</b>
<b>Anexo 3. Conectividad y deserción (sin considerar escuelas integradas) .....</b>	<b>44</b>

## **Capítulo I. Introducción**

Si bien la educación es uno de los pilares más importantes para el desarrollo de cualquier sociedad, la evidencia muestra que los niveles de escolarización en los países en desarrollo son mucho más bajos que en los países desarrollados (Naz, Ejaz y Khan, 2019). En el Perú, según la Encuesta Nacional de Hogares (Enaho), en 2019, a pesar de que el 97% de niños entre 6 y 11 años tenía acceso a la educación primaria, el 13% de adolescentes entre 12 y 16 años no accedía a la educación secundaria.

Esta situación se agrava si se observa las tasas de deserción escolar. El Sistema de Información de Apoyo a la Gestión de la Institución Educativa (Siagie) del Ministerio de Educación (Minedu) mostró que, para el periodo 2018 – 2019, el nivel secundaria contaba con la mayor tasa de deserción escolar a nivel nacional (3.5%), seguido por el nivel inicial (2.2%). Si bien las tasas parecen ser relativamente bajas, el porcentaje correspondiente al nivel secundaria equivale a más de 90 mil estudiantes. Además, en línea con Alcázar (2008), el problema se acentúa en las zonas rurales, donde, en el nivel secundaria, alrededor de 11 mil estudiantes más abandonan sus estudios en comparación con las zonas urbanas.

Para entender el problema de la deserción, la literatura ha sumado esfuerzos en explicar sus principales determinantes. A través de una extensa revisión de estudios, Rumberger y Lim (2008) identificaron dos grupos de factores que la explicarían: las características del estudiante y las de su entorno (familia, escuela y comunidad). Así, se destaca que su rendimiento académico, si ha repetido un grado, sus aspiraciones, su sexo, entre otros, afectan su decisión de desertar (Rumberger, 1987; Cueto, 2002; Field, Kuczera y Pont, 2007; Markussen, 2010). Variables relacionadas a los ingresos de los padres y su nivel de educación también resultan relevantes al momento de tomar dicha decisión (Steinberg, Lin Blinde y Chan 1984; Huisman y Smits, 2015; Naz, Ejaz y Khan, 2019).

Huisman y Smits (2015) sostienen que la estructura del hogar también juega un rol crucial. Esto se alinea con Sawada y Lokshin (2001), quienes afirman que la cantidad de niños en el hogar se relaciona negativamente con mayor nivel educativo. Además, el número de hermanas o hermanos mayores podría aumentar la escolaridad gracias a los recursos que estos aportan en la familia. Así, la edad del niño termina siendo una de las determinantes de la deserción escolar al igual que su orden de nacimiento (Cueto, 2002; Alcázar, Rendón y Wachtenheim, 2002); especialmente, en la transición de la educación primaria hacia la secundaria (Huisman y Smits, 2015).

Si bien más del 72% de la variación de la deserción escolar es explicada por factores relacionados a la demanda educativa (características del estudiante y del hogar), existe una fracción de esta que se explica por la oferta (Huisman y Smits, 2015). La mayoría de estudios relacionados a la oferta se concentran en la estructura y disponibilidad de recursos de las escuelas (Neild, Stoner-Eby y

Furstenberg, 2008; Mac Iver y Mac Iver, 2009) y en su disponibilidad en el centro poblado (Alcázar, 2008). En este último, la literatura se ha enfocado en estudiar el efecto de la distancia a la escuela sobre la deserción escolar. Así, se encuentra que la distancia entre el hogar y la escuela es uno de los principales determinantes del abandono escolar (Rolleston, Akyeampong, Ampiah y Lewin, 2010; Hussain, Salfi y Khan, 2011; Henki, Nismawati y Lamadirisi, 2020).

Sharma y Levinson (2019) estudian la relación entre el tiempo de viaje hacia la escuela secundaria y la deserción escolar en dicho nivel. Los autores encuentran que mientras mayor sea la distancia, mayor es la probabilidad de que un estudiante abandone la escuela. Su principal recomendación de política es desarrollar un sistema de transporte que responda directamente al sistema educativo. Por su parte, Huisman y Smits (2015) mencionan que, debido a que la tasa de deserción se concentra en la finalización de la educación primaria, promover políticas que extiendan la duración de este nivel educativo podría mantener en las escuelas a los niños por más tiempo.

Gracias a las recomendaciones de los estudios anteriores se desprende dos cosas. Primero, se debe invertir en conectividad terrestre, ya que la distancia física es importante para no abandonar la escuela. Segundo, el costo de oportunidad “continuar” la primaria podría ser relativamente menor que el culminarla y trasladarse a una secundaria. Así, la desconexión entre los servicios educativos en ambos niveles podría ser el causante de una transición educativa con poca fluidez a lo largo de la educación obligatoria. Por ende, se requiere de una configuración de los servicios educativos que estén centrados en los estudiantes (Guadalupe y Rivera, 2020).

El objetivo del presente estudio es analizar si la conectividad terrestre entre las instituciones educativas de nivel primaria y secundaria en las zonas rurales contribuye con la deserción escolar durante la transición entre estos niveles. En particular, se busca probar si el hecho de que una escuela primaria se encuentre relativamente cerca de una escuela secundaria reduce la probabilidad de que un estudiante sea desertor en la transición primaria – secundaria.

La hipótesis de la cual se parte es que la probabilidad de que un estudiante sea desertor en la transición primaria – secundaria es mayor en caso de que termine sus estudios de la primaria en una escuela que no tiene una secundaria relativamente cerca (o no está lo suficientemente conectada). Esta probabilidad podría ser de mayor magnitud en estudiantes mujeres, en estudiantes con bajo rendimiento académico o en estudiantes que habiten en zonas altamente rurales (donde el nivel de conectividad terrestre es mucho menor).

Para responder a la pregunta de investigación y testear la hipótesis planteada, se utilizará un estimador de efectos fijos a nivel de familia que aprovecha la variación de la disponibilidad de secundarias alrededor de las primarias donde egresan los hermanos. Así, la muestra está compuesta por familias con más de un hijo que haya culminado la primaria, donde al menos uno de ellos lo haya hecho en una escuela que no tenía una secundaria lo suficientemente cerca y, al

menos uno, lo haya hecho en una escuela que sí tenía una secundaria lo suficientemente cerca. Cabe resaltar que, si bien se busca testear el efecto de la conectividad entre escuelas, esta conectividad pasa por la disponibilidad (dada la cercanía) de servicios educativos más allá de la mejora en las rutas o carreteras que conectan a estos.

De este modo, el estimador utiliza la diferencia entre hermanos para controlar por características observables y no observables de la familia que afecten la decisión de desertar del estudiante a lo largo del tiempo. A su vez, el hecho de que al menos uno haya (o no haya) tenido acceso a una escuela primaria conectada con una secundaria, permite identificar el cambio, para una misma familia, de que un niño egrese de una escuela conectada frente a una que no lo esté.

Para ello, se hará uso de datos administrativos del Siagie y de los padrones de instituciones educativas para el periodo 2016 – 2018. En cuanto a la determinación de la conectividad terrestre entre las instituciones educativas, el estudio propone utilizar la distribución geográfica de las escuelas públicas a nivel nacional. Desde el 2016, el Minedu registra la georreferenciación (latitud y longitud) de todos los locales escolares (tanto públicos como privados). Así, por medio de técnicas geoespaciales, se puede trazar áreas de influencia alrededor de cada escuela primaria para determinar el número de escuelas secundarias que tiene alrededor.

Si bien el presente estudio se alinea a la literatura relacionada con el efecto de la distancia entre el hogar y la escuela sobre la deserción escolar, dado que no se cuenta con datos georreferenciados de los hogares, se utilizará la escuela primaria como *proxy* del hogar. De este modo, el principal aporte es la contribución con el debate de políticas públicas, donde se recoge el problema de la gran dispersión geográfica de las escuelas secundaria, que, en efecto, hacen más costoso el acceso, en cuestiones de tiempo de transporte o búsqueda de escuelas. Más allá de evaluar un efecto específico, la idea de este estudio es utilizar los resultados como insumo para hacerle frente a un problema de carácter público desde las herramientas que pueden dirigir los hacedores de política.

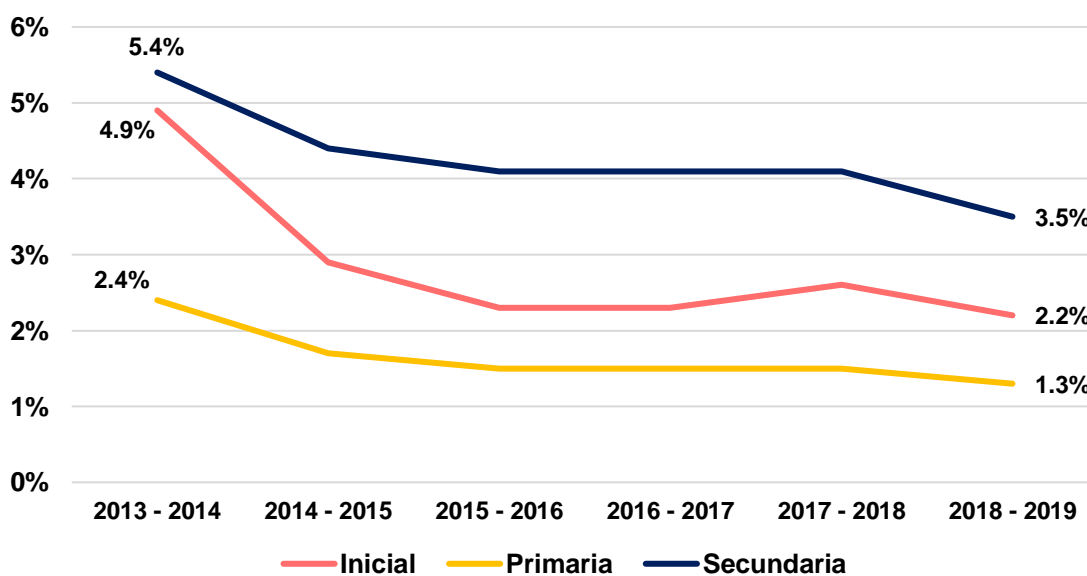
El resto del estudio se distribuye de la siguiente manera. El capítulo II describe el contexto de la deserción escolar en el Perú. El capítulo III expone la oferta educativa disponible y su distribución geográfica a nivel nacional. En el capítulo IV se muestra la revisión de literatura. El capítulo V ofrece una discusión sobre el uso de los datos y la estrategia empírica a utilizar. El capítulo VI presenta los resultados y discute sobre la robustez de los mismos, así como un análisis de heterogeneidades. El capítulo VII propone algunas recomendaciones de política y en el capítulo VIII se presenta las conclusiones del estudio.

## Capítulo II. La deserción escolar en el Perú

La deserción escolar se define como la situación en la cual los estudiantes abandonan o no completan la formación que están recibiendo (Lyche, 2010). En el presente análisis, se referirá a un estudiante desertor como aquel que estuvo inscrito en el Siagie cursando alguno de los niveles correspondientes a la Educación Básica Regular y que, en el año siguiente no figura como parte del alumnado a nivel nacional<sup>1</sup>. Así, tal y como se observa en la **Figura 1**, la deserción interanual para el último periodo registrado por el Minedu (2018 – 2019) fue de 1.3% para el nivel primaria y 3.5% para el nivel secundaria<sup>2</sup>.

La **Figura 1** muestra que, si bien la deserción escolar a nivel nacional se ha reducido durante 2013 y 2019 en todos los niveles educativos, todavía es un problema presente en el sistema educativo, especialmente, en el nivel secundaria. Antes de la pandemia, alrededor de 152 mil estudiantes habían desertado de la Educación Básica Regular (EBR), más de 61 mil correspondientes al nivel primaria y más de 90 mil correspondientes al nivel secundaria.

**Figura 1. Tasa de deserción escolar interanual según nivel educativo**



Fuente: Siagie (2013 – 2019).

Si se desagrega la evolución de las tasas de deserción, se puede observar que existe una brecha entre las zonas rurales y las zonas urbanas. La **Figura 2** muestra la evolución de las tasas de deserción según ámbito geográfico. Como se observa, en el nivel inicial, en el periodo 2013 –

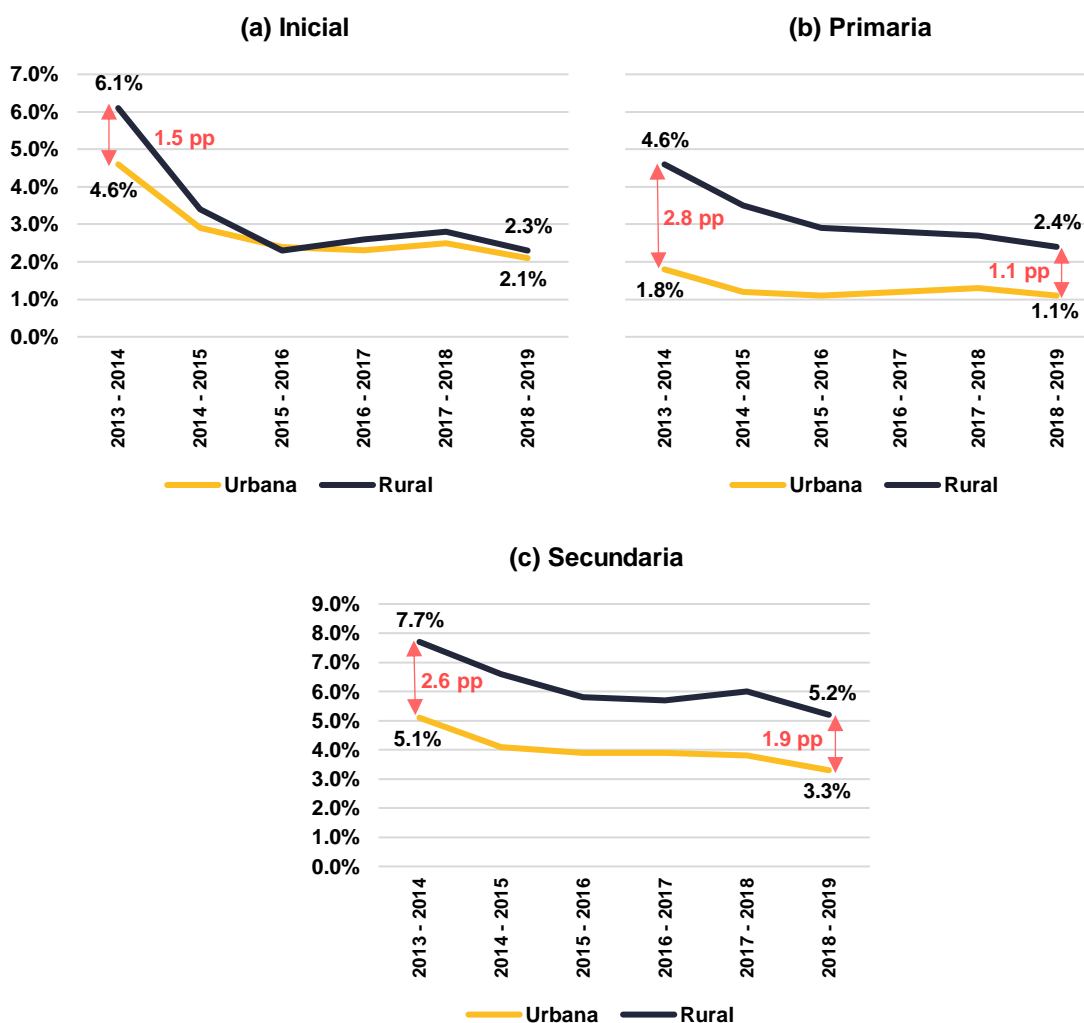
<sup>1</sup> En el caso de los alumnos que cursaban el quinto año de nivel secundaria, se definió como desertor a aquel estudiante que no aprobó dicho grado y que, al año siguiente, no figuró en los registros del Siagie.

<sup>2</sup> El presente análisis se desarrolla en un contexto de educación presencial (pre-pandemia), por lo que no se considera las tasas de deserción en el contexto de pandemia debido a que podría generar ruido en la muestra.

2014, la brecha en la tasa de deserción interanual entre las zonas rurales y urbanas equivalía a 1.5 puntos porcentuales (pp.). No obstante, para el último periodo en análisis, se observa un cierre de estas.

En el caso del nivel primaria, si bien se observa un decrecimiento en las tasas de deserción en ambos ámbitos, la caída de las zonas rurales es mayor (2.2 pp. frente a 0.7 pp. en zonas urbanas). Sin embargo, se muestra una brecha persistente de 1.1 pp. entre ambas zonas. Algo similar sucede en el nivel secundaria, donde la brecha pasó de 2.6 pp. en el periodo 2013 – 2014 a 1.9 pp. en el periodo 2018 – 2019. De este modo, se observa que las tasas de deserción son mayores en las áreas rurales, especialmente, en los niveles primaria y secundaria. En términos de estudiantes, se puede decir que en el último periodo, en las zonas rurales desertaron alrededor de 13 mil estudiantes más que en las zonas urbanas en el nivel primaria y cerca de 11 mil estudiantes más lo hicieron en el nivel secundaria.

**Figura 2. Tasa de deserción escolar interanual según área**

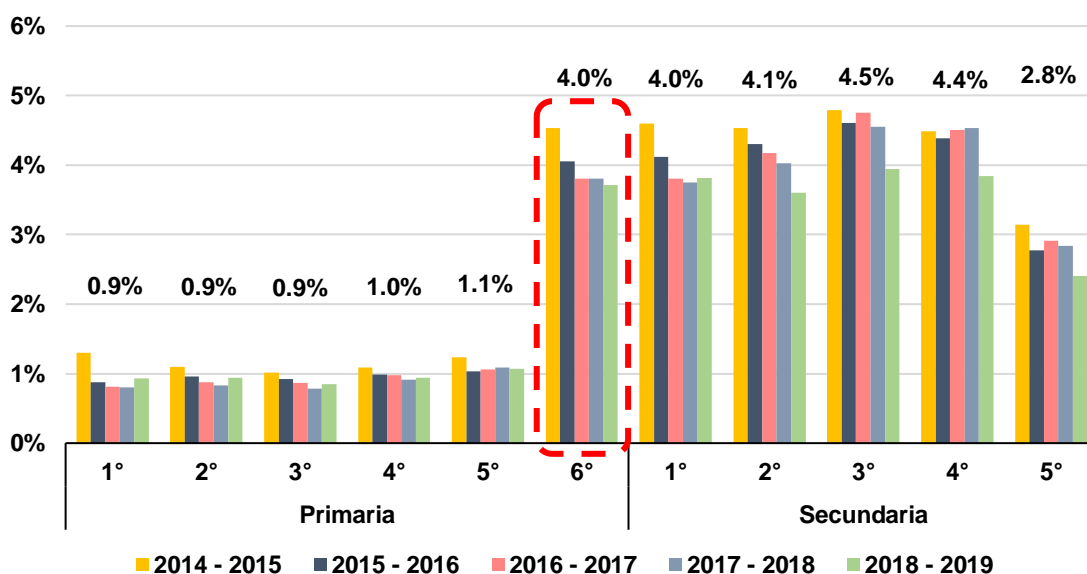


Fuente: Siagie.

Ahora bien, siguiendo a Huisman y Smits (2015), la **Figura 3** muestra la tasa de deserción escolar por grado en los dos niveles de interés: primaria y secundaria. Tal y como se menciona en la literatura, la mayor concentración de abandono escolar se encuentra en el nivel secundaria (Steinberg, Lin Blinde y Chan, 1984; Alcázar, 2008; Henki, Nismawati y Lamadiridi, 2020); específicamente, en la transición entre la educación primaria y secundaria. Entre 2014 y 2019, la tasa de deserción escolar promedio de niños que concluyen la primaria es de 4%, lo que equivale a decir que, anualmente, alrededor de 22 mil estudiantes, en promedio, terminan la primaria y no inician la secundaria en todo el Perú.

El salto en la deserción escolar entre los niveles primaria y secundaria es bastante marcado. Esta diferencia se alinea con Huisman y Smits (2015), quienes encuentran que hay un momento crítico al finalizar la educación primaria en el cual el riesgo de abandonar el sistema educativo es bastante elevado. Ello hace que, en países donde la duración de la educación primaria es mayor, se observe menor deserción escolar en los siguientes grados. Por otro lado, según los datos del Siagie, entre 2016 y 2018, alrededor del 62% de los estudiantes que continúan sus estudios de nivel secundaria en zonas rurales, transitan a la escuela secundaria más cercana a su escuela primaria.

**Figura 3. Tasa de deserción escolar interanual según grado**

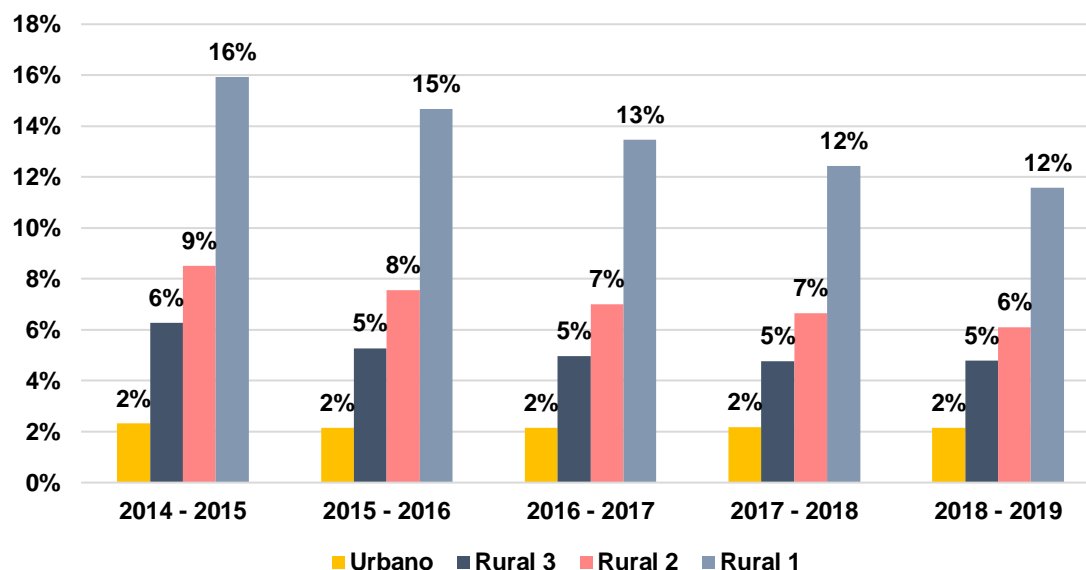


**Nota:** Las etiquetas del gráfico corresponden a la tasa de deserción promedio de todo el periodo en análisis para el grado correspondiente.  
Fuente: Siagie.

Finalmente, el problema de la deserción en la transición primaria – secundaria se agudiza en las zonas rurales. La **Figura 4** muestra las tasas de abandono escolar para el sexto de primaria según

gradiente de ruralidad<sup>3</sup>. Como se observa, existe cierto grado de correlación positiva entre el grado de ruralidad de una zona y las tasas de deserción. Así, las zonas que no son urbanas, pero son menos rurales, presentan una tasa de deserción de alrededor del 5%; mientras que las zonas más rurales presentan tasas de 14%, en promedio. Por ende, existe una urgencia de desarrollar políticas públicas enfocadas en atender el problema de abandono escolar, especialmente, en la transición primaria – secundaria en las zonas rurales.

**Figura 4. Tasa de deserción escolar 6to de primaria según gradiente de ruralidad**



**Nota:** Rural 1 refiere a instituciones educativas ubicadas en centros poblados (CCPP) rurales con máximo de 500 habitantes y pueden acceder a la capital provincial más cercana en un tiempo mayor de 2 horas. Se considera Rural 2 a aquellas ubicadas en zonas rurales con más de 500 habitantes y pueden acceder a la capital provincial más cercana en un tiempo mayor a 2 horas, o instituciones ubicadas en CCPP de máximo 500 habitantes y pueden acceder a la capital provincial más cercana en un tiempo mayor a 30 minutos y menor igual a 2 horas. Finalmente, la categoría Rural 3 responde a instituciones ubicadas en zonas rurales con más de 500 habitantes y pueden acceder a la capital provincial más cercana en no más de 2 horas, o aquellas ubicadas en CCPP con menos de 500 habitantes y pueden acceder a la capital provincial más cercana en no más de 30 minutos.

Fuente: Siagie (2014 – 2019).

### Capítulo III. Distribución de la oferta de servicios educativos en el Perú

En 2019, Perú contaba con alrededor de 107,142 servicios educativos, considerando los programas de educación inicial no escolarizado. La **Tabla 1** muestra el número de servicios por nivel educativo. Se observa que, el número de instituciones educativas de nivel primaria son más del doble que las instituciones de nivel secundaria. Además, la distribución de los servicios por ámbito muestra que, si bien cerca del 60% de las primarias se concentran en las zonas rurales, sólo el 30% de las secundarias están en estas zonas. Esto podría deberse a la gran dispersión en

<sup>3</sup> La gradiente de ruralidad es un instrumento de política elaborada por la Unidad de Estadística del MINEDU en base a la densidad poblacional (tamaño de la población de los centros poblados) y la dispersión de las escuelas (grado de accesibilidad de los centros poblados a la capital provincial más cercana). La principal utilidad de esta gradiente es la focalización de diversas intervenciones y políticas educativas dentro del ministerio.

las zonas rurales y no necesariamente a la existencia de mayor demanda, lo que genera la atomización de los servicios educativos (Guadalupe y Rivera, 2020).

**Tabla 1. Servicios educativos según ámbito y gestión (2019)**

<b>Nivel Educativo</b>	<b>Total</b>	<b>Urbano</b>	<b>Rural</b>	<b>Público</b>	<b>Privado</b>
<b>Inicial*</b>	53,706	54%	46%	80%	20%
<b>Primaria</b>	38,755	42%	58%	77%	23%
<b>Secundaria</b>	14,969	70%	30%	65%	35%
<b>Total</b>	<b>107,430</b>	<b>52%</b>	<b>48%</b>	<b>77%</b>	<b>23%</b>

(\*) Incluye los programas no escolarizados

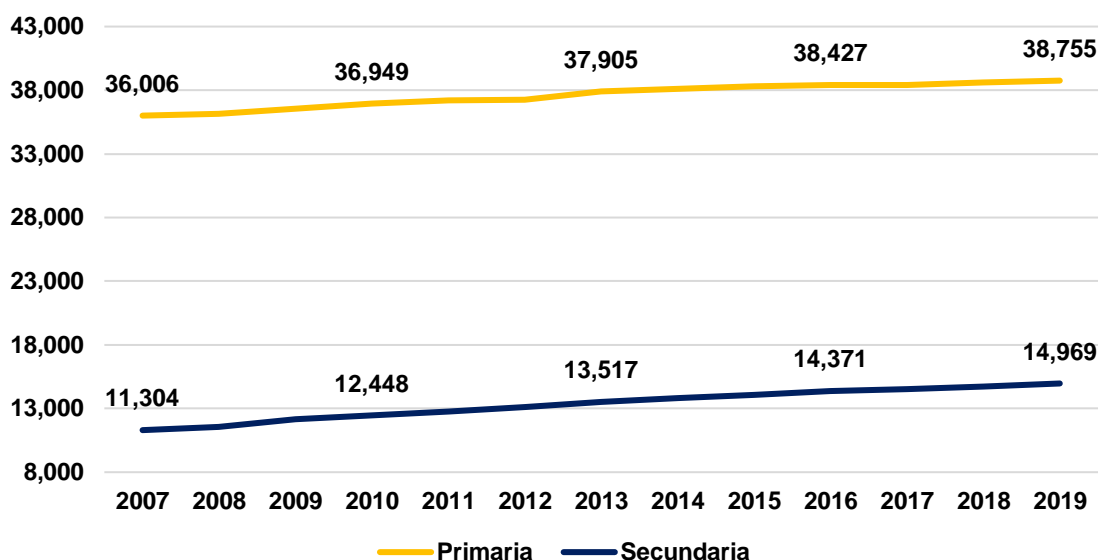
Fuente: ESCALE.

Otro dato que se puede resaltar de la **Tabla 1** es que el 77% de las instituciones educativas de nivel primaria son gestionadas por el sector público, ya sea de gestión directa o privada<sup>4</sup>. En las zonas rurales, menos del 1% de las instituciones educativas de nivel secundaria son de gestión privada. Así, el principal responsable de la oferta de los tres niveles en dichas zonas es el Estado. Nuevamente, esto podría deberse a la gran dispersión de la demanda en las zonas rurales, lo que haría costoso para el sector privado encargarse de la oferta.

En cuanto a la evolución de la oferta, la **Figura 5** muestra que, si bien la tendencia ha sido positiva en los últimos años, no necesariamente hay un incremento sustancial; especialmente en el nivel primaria (8%). No obstante, esto no necesariamente es un problema, pues, según la Enaho (2019), el 97% de niños entre 6 y 11 años tiene acceso a la educación primaria. En el caso de las escuelas secundaria, se observa que entre 2007 y 2019 hubo un crecimiento de más de 30%. Sin embargo, la tasa de jóvenes entre 12 y 16 años que accede a esta educación es de 87%. Así, en términos de políticas públicas, para el cierre de brechas en el acceso se requiere de estrategias que i) expandan la oferta y ii) reduzca las tasas de deserción en este nivel educativo.

<sup>4</sup> Las instituciones educativas públicas de gestión directa son aquellas que están a cargo del Ministerio de Educación; mientras que las públicas de gestión privada son aquellas que, vía convenio con el ministerio, son gestionadas por el sector privado (Asociación Público – Privada). Un ejemplo de este último son las escuelas Fe y Alegría.

**Figura 5. Número de instituciones educativas según nivel educativo (2007 – 2019)**

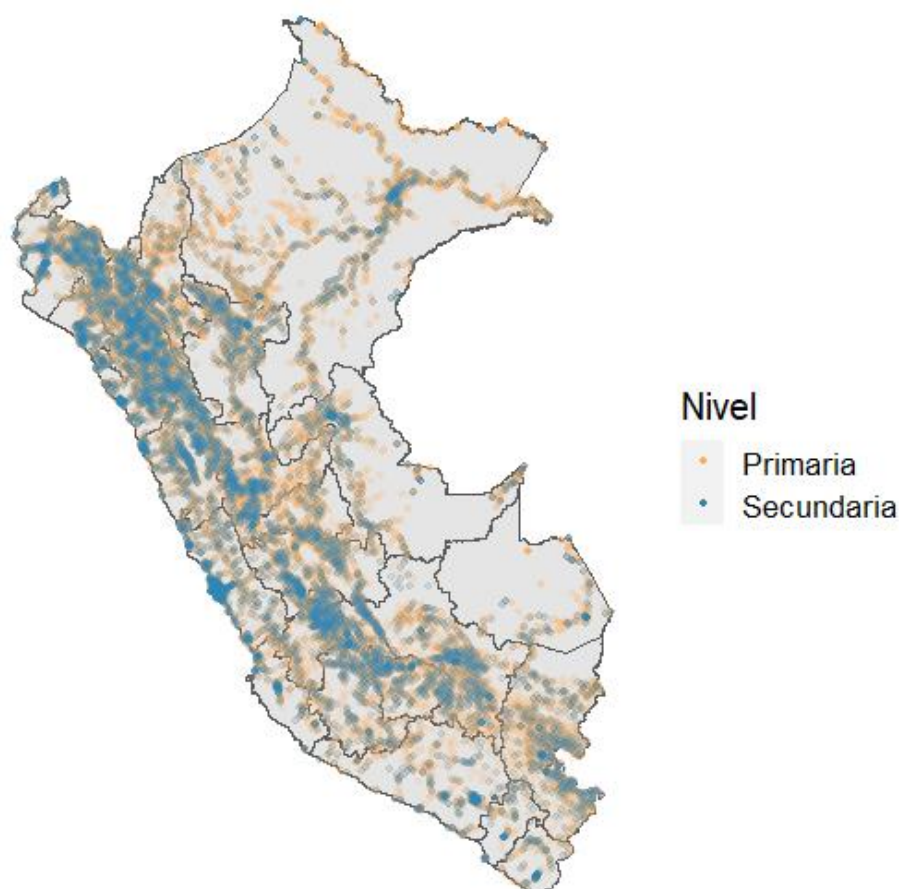


Fuente: ESCALE.

La **Figura 6** muestra la distribución geográfica de la oferta pública de instituciones educativas primarias y secundarias a nivel nacional al 2019. Como se observa, la oferta de instituciones educativas de nivel primaria cubre gran parte del territorio nacional, mientras que la oferta de instituciones de nivel secundaria se concentra en la costa norte del país. Así, regiones como Cajamarca, Lambayeque y La Libertad son las que más concentran servicios educativos de nivel secundaria. Algo similar se observa en El Callao, Huancavelica y parte de Ayacucho.

Al lado opuesto, se encuentran las regiones de Loreto, Ucayali y Madre de Dios con menor concentración de servicios educativos de nivel secundaria. Cabe resaltar que Loreto es la región que concentra la mayor cantidad de pueblos indígenas (18.3%) y, por ende, de servicios de Educación Intercultural Bilingüe (Ministerio de Cultura, 2021). Al sur del país, Arequipa, Moquegua y Tacna también son regiones que presentan muy poca concentración de oferta de nivel secundaria. Así, queda clara la dispersión entre la oferta de servicios educativos en las zonas rurales y la poca conexión entre estas en algunas regiones del país.

**Figura 6. Distribución geográfica de instituciones educativas (2019)**

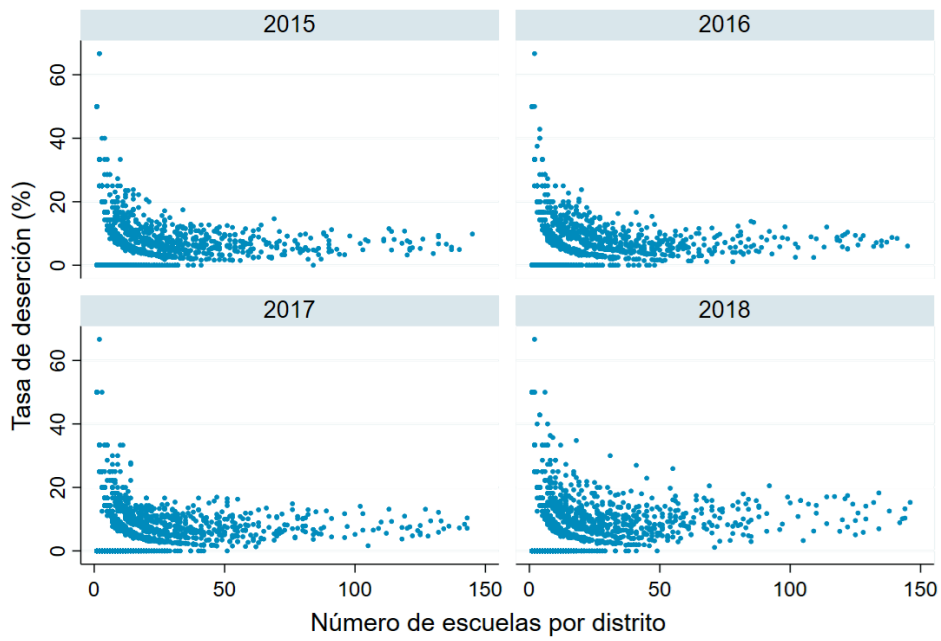


Fuente: ESCALE (2019).

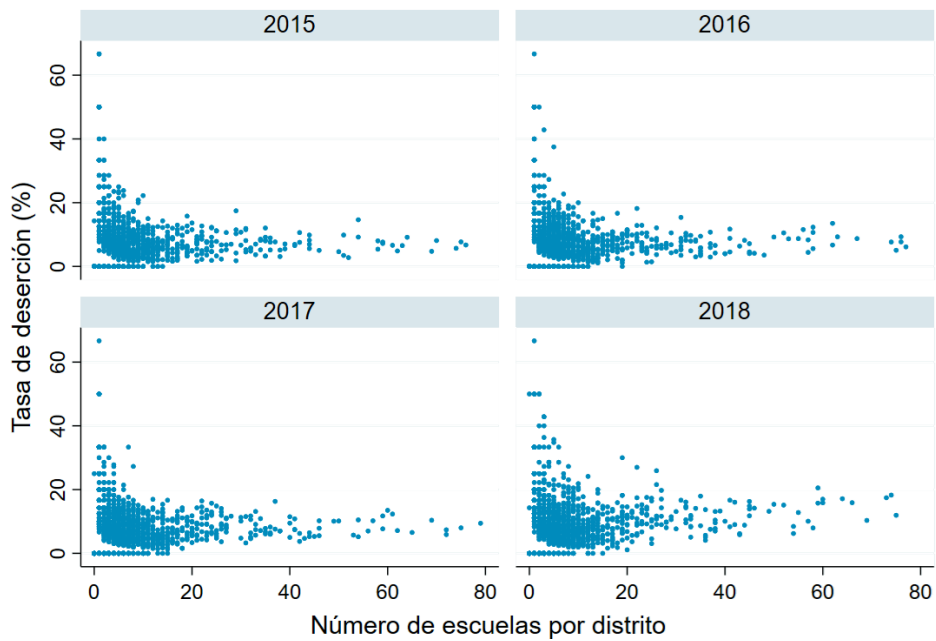
La oferta educativa es de vital importancia, pues según los datos de Siagie, existe una correlación negativa entre las tasas de deserción escolar y el número de escuelas en un determinado distrito. La **Figura 7** muestra dicha correlación para el periodo 2015 – 2018. Se observa que, independientemente del nivel (primaria o secundaria) o del año, la relación se mantiene. En distritos donde la oferta educativa es escasa, se puede observar tasas de deserción escolar por encima del 60%; mientras que en distritos con mayor oferta educativa se observa tasas por debajo del 10%. Esto se alinea con la literatura relacionada al acceso de servicios educativos, donde se relaciona la disponibilidad de las escuelas con la continuidad de la educación (Alcázar, 2008; Henki, Nismawati y Lamadirisi, 2020).

**Figura 7. Deserción escolar y el número de escuelas por año y por nivel**

**a. Primaria**



**b. Secundaria**



Fuente: Siagie (2015 – 2018).

#### **Capítulo IV. Revisión de literatura**

La deserción escolar es un fenómeno común en los sistemas educativos que toma mayor fuerza en el nivel secundaria, aunque con una tendencia decreciente (Roman, 2013). En un estudio longitudinal para cuatro países, incluido Perú, Cueto, Felipe y León (2020) encuentran que este es un problema relacionado con la pobreza. La mayoría de los niños que abandonan la escuela mencionan que lo hicieron por la necesidad de trabajar, cuidar o mantener a su familia. Por el lado de las mujeres, esto también se relaciona con el embarazo adolescente (Dagenais, Montmarquette y Viennot-Briot, 2002; Alcázar, 2008).

No obstante, el factor socioeconómico no lo es todo (Dowrick y Crespo, 2005). En una revisión de literatura de más de 389 estudios, Rumberger y Lim (2008) concluyen que la deserción escolar se determina por factores individuales del propio alumno y por factores de su entorno, como características de su familia, su escuela y su comunidad. Lyche (2010) segmenta los factores en tres categorías: individuales (o sociales), de la escuela y sistémicos.

La primera categoría coincide en que el rendimiento académico de los estudiantes es clave para la continuidad de los estudios. Así, estudiantes con mejor rendimiento son quienes tienen más probabilidad de concluir sus estudios (Steinberg, Lin Blinde y Chan, 1984; Rumberger, 1987; Rumberger y Lim, 2008; Markussen, 2010). Al respecto, Cueto, Felipe y León (2019) sostienen que un mayor desempeño académico está relacionado con ser varón y el hecho de haber migrado de zona rural a urbana antes de los 8 años. Entre otros determinantes asociados al factor individual, Cueto (2002) menciona que la edad y ser mujer guardan relación positiva con la probabilidad de desertar. Por su parte, Lyche (2010) menciona que el compromiso del estudiante y su contexto (experiencias pasadas, salud, y características de su familia) resultan relevantes sobre la decisión de abandonar la escuela.

Respecto al último, además del aspecto socioeconómico, también es importante la estructura del hogar. Así, hogares con más niños estarían relacionados con menor probabilidad de asistencia escolar en alguno de ellos debido a que la asignación de recursos para cada niño es menor. No obstante, la presencia de hermanos mayores también podría favorecer a la generación de más recursos, lo cual permitiría mayor escolaridad entre los menores (Sawada y Lokshin, 2001; Orazem y Gunnarson, 2004).

En la segunda categoría, se tiene factores relacionados a los recursos y estructura de la escuela y las prácticas pedagógicas. En el primer caso, Mac Iver y Mac Iver (2009) muestran que, por lo general, la deserción escolar se concentra en escuelas públicas. Esto se relaciona bastante con el contexto socioeconómico de quienes asisten a cada tipo de escuela. Además, es común observar que quienes se retiran de una escuela privada, no salen del sistema educativo; sino que migran a escuelas de gestión pública (Rumberger y Lim, 2008). Otro factor adicional que se atribuye a la

escuela es el efecto de los pares. Markussen, Frøseth, Lødding y Sandberg (2008) y Rumberger y Lim (2008) coinciden que las relaciones entre estudiantes, así como el contexto y características de los pares están correlacionadas con las tasas de deserción.

Por el lado de las prácticas pedagógicas, como se mencionó anteriormente, el compromiso del estudiante con sus estudios es un determinante del abandono escolar. Si bien este no es muy fácil de medir, se puede generar indicadores que midan la motivación del estudiante como el número de veces que participa en clase (Newmann, Wehlage y Lamborn, 1992), lo cual estaría asociado tanto con la estructura de la escuela como con las relaciones con los docentes (Neild, Stoner-Eby y Furstenberg, 2008; Rumberger y Lim, 2008). Así, un clima positivo tanto en el salón de clases como en la escuela, en general, reduciría las tasas de deserción en estas.

En la tercera categoría se tiene a los factores sistémicos o estructurales del sistema educativo, los cuales influyen sobre las tasas de deserción, especialmente, en el nivel secundaria (Lyche, 2010). Bridgeland, Dilulio y Morison, (2006) y Field, Kuczera y Pont (2007) sostienen que el hecho de repetir un año escolar se correlaciona fuertemente con la deserción. Rumberger y Lim (2008) encontraron que esto también estaría correlacionado con la edad relativa de los estudiantes, pues en un salón de clases, los mayores tienen una probabilidad de desertar más alta que los menores.

Sin embargo, una de las causas estructurales más frecuentes por la que los estudiantes dejan la escuela es la falta de oferta educativa (Alcázar, 2008; Markussen, 2010). Este problema se agrava en países en desarrollo, donde la distancia entre el hogar y la escuela es uno de los principales determinantes del abandono escolar; especialmente para las mujeres (Rolleston, Akyeampong, Ampiah y Lewin, 2010; Husain, Salfi y Khan, 2011; Sharma y Levinson, 2019). A pesar de ello, Huisman y Smits (2009) muestran que la distancia a la escuela no es un factor relevante en hogares donde el padre no se dedica a la agricultura y la madre trabaja.

Si bien la deserción es un problema en todos los grados educativos, por lo general, suele concentrarse en la transición primaria – secundaria (Huisman y Smits, 2015). Neild, Stoner-Eby y Furstenberg (2008) sostienen que ello se explicaría, en parte, por el rendimiento académico del estudiante en el último grado de primaria. Esto se debe a que representa un *proxy* de lo que sería su rendimiento académico en el siguiente nivel educativo y, por ende, los potenciales retornos de su educación. Por ello, ante mejor rendimiento, mayor será la probabilidad de que el estudiante pueda transitar a la educación secundaria. Huisman y Smits (2015) mencionan que este hecho se explicaría debido a que tanto los niños como los padres tienden a esperar que terminen la primaria por completo para detener su educación .

En cuanto al uso de datos georreferenciados, hay una literatura reciente sobre *location as an asset*, en el cual se considera a la ubicación del individuo como un potencial activo, cuyos retornos futuros se basan en las mejores oportunidades que pueda obtener por el simple hecho de estar en

un determinado lugar (Bilal y Rossi-Hansberg, 2021). Las oportunidades que recibe el individuo por vivir en una ciudad frente a otra podrían estar en función de mayor escolaridad, mejores trabajos o menores costos del hogar.

Chetty y Hendren (2018) presentan evidencia cuasi-experimental sobre que el cambio de ubicación de las familias genera diferentes tasas de acumulación de capital humano intergeneracional. De este modo, mudarse a una mejor ubicación podría ofrecerte mejores opciones de escolarización, lo cual impacta directamente sobre variables de resultados como la tasa de asistencia a la universidad. Los autores utilizan la variación en la edad de los niños cuando las familias se mudan para capturar los efectos de i) mudarse a un mejor vecindario y ii) cómo mejoran los niños que ya viven ahí en proporción a como lo estaban haciendo antes de la mudanza. Así, Chetty y Hendren (2018) sostienen que la ubicación de las familias podría cambiar los resultados de los hijos a lo largo de su vida.

Si bien el uso de datos espaciales en la literatura sobre movilidad ha sido bastante utilizada para evaluar sus efectos sobre variables económicas (i.e. Wilson, 1987; Massey y Denton, 1993; Katz, Kling y Liebman, 2001; Ludwig y otros, 2013; Sharkey y Faber, 2014), su uso para determinar la deserción escolar en la transición de niveles educativos (de manera causal), a conocimientos del autor, parece ser inexistente. No obstante, sí existe evidencia del uso de estos datos en el desarrollo de políticas educativas como la construcción de escuelas según la demografía o de centros de capacitación para docentes (Mendelsohn, 1996).

En efecto, las herramientas de análisis espacial pueden ser muy útiles para el planeamiento de la provisión de servicios educativos de países en desarrollo (Ndaendelao, 2007; Odhiambo e Imwati, 2014). Adeyemi y Edeki (2018) resaltan la importancia de la distribución de las universidades sobre las aspiraciones futuras de los jóvenes. Galabawa, Obeleagu y Miyazawa (2002) sostienen que el desarrollo cartográfico de las escuelas primaria en Tanzania afectó positivamente sobre algunos indicadores de escolaridad como la matrícula y asistencia y redujo las tasas de deserción escolar. Sin embargo, la mayoría de análisis en esta rama de la literatura carecen de una metodología rigurosa que permita evaluar la causalidad entre la distribución espacial de las escuelas y la correcta transición educativa desde la primaria hacia la secundaria.

Con respecto a la estrategia de identificación, Holland y Johnson (1984) sostienen que la similitud entre hermanos depende mucho del sexo. Sin embargo, al compartir el mismo ambiente, suelen ser muy parecidos entre sí. Gracias a ello, diversos estudios han explorado la importancia del orden de nacimiento y la edad de los hermanos (mayores o menores) sobre la asistencia escolar, el rendimiento académico o, incluso, la participación política (Blake, 1989; Xiong, Nicoletti y Rabe, 2019; Zang, Zhou y Liu, 2020; Symeonides y otros, 2021).

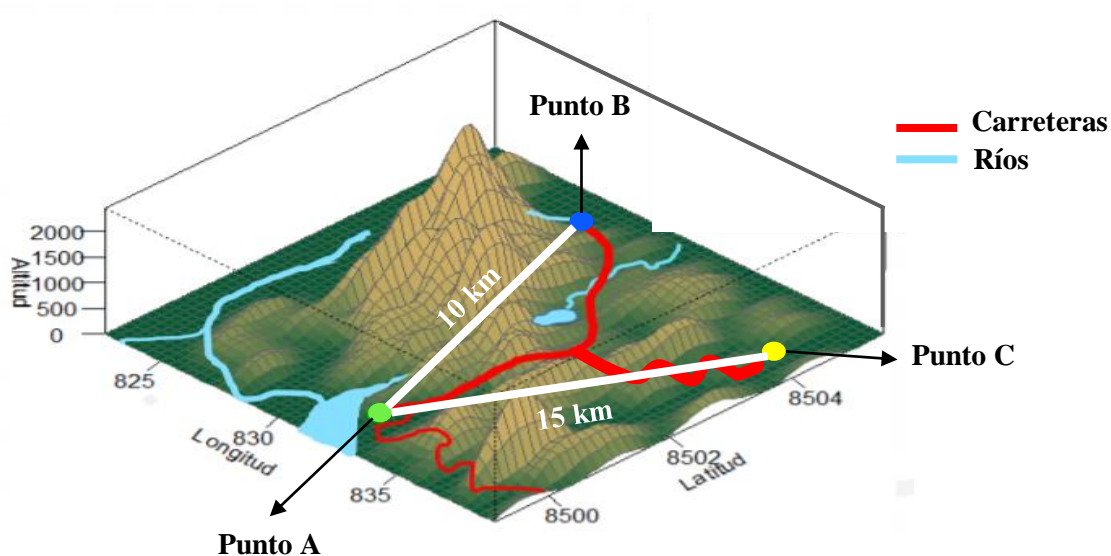
Duflo (2001) utiliza una estrategia similar para evaluar el impacto de la expansión de oferta educativa en Indonesia. Si bien el incremento de la oferta en el tiempo podría no afectar directamente a un mismo individuo, sí puede hacerlo con alguien muy similar en cuanto a características socioeconómicas, pero de una cohorte menor. Así, la comparación entre hermanos como herramienta para probar hipótesis de causalidad es válido si se quiere evaluar el ambiente al cual se exponen los niños sobre algún indicador de comportamiento como la asistencia escolar (Lahey y D'Onofrio, 2010).

## Capítulo V. Metodología

El presente estudio busca evaluar si existe alguna relación causal entre la conectividad terrestre de las escuelas de nivel primaria y secundaria y la probabilidad de que un estudiante abandone sus estudios al finalizar el primer nivel educativo. En este contexto, se define a un desertor como aquel estudiante que en el año  $t$  aprobó el sexto grado de primaria y en el año  $t + 1$  no presenta registro alguno en el Siagie. Además, como se mostró en la sección anterior, la decisión de desertar depende de factores individuales, socioeconómicos, de la escuela y sistemáticos.

Para evaluar la conectividad entre escuelas, primero se observa la distribución geográfica de las instituciones educativas de nivel primaria y se generan áreas de influencia alrededor de cada una, con un radio de “ $x$ ” minutos. En este caso, no se utiliza la distancia euclidiana entre dos puntos debido a la existencia de accidentes geográficos que no permiten trazar rutas de transporte para la correcta movilidad entre los puntos.

**Figura 8. Distancia entre dos puntos ante la presencia de relieve**



Fuente: Minedu (2019).

Un ejemplo bastante sencillo se encuentra en las zonas rurales, donde la escuela A puede estar a una distancia euclidiana (lineal) de 10 kilómetros de la escuela B y de 15 kilómetros de la escuela C. Sin embargo, entre la escuela A y la escuela B hay una montaña; por lo que no existe un medio de transporte que permita recorrer linealmente desde la escuela A hacia la escuela B sin tener que cruzar dicha montaña. Así, considerando el hecho de rodear la montaña, la escuela B dista de la escuela A en aproximadamente 20 kilómetros. Finalmente, se concluye que la escuela C está relativamente más cerca de la escuela A que la escuela B (ver **Figura 8**).

Para evitar el problema generado por la distancia euclidiana, se considera la distancia entre dos puntos medida en minutos si se viaja con el medio de transporte más rápido disponible. Esto quiere decir que, si entre los puntos existe una carretera, se considera la distancia en minutos movilizándose a través de un auto; si los puntos tienen un río de por medio, se utilizará un bote como medio de transporte; si no hay vías ni ríos, entonces, el medio de transporte es a pie.

La medición de distancia se ajusta a la metodología utilizada por el Minedu para la elaboración de los padrones de asignaciones temporales a docentes por ruralidad. En este sentido, se emplea el mismo modelo de accesibilidad que considera la ruta óptima entre dos puntos según elementos de acceso como carreteras y elementos de fricción como ríos, quebradas, relieves, entre otros. Cabe resaltar que dicho modelo no considera información como el acceso a transporte público, el estado de la vía o su accesibilidad según el clima debido a que no está disponible (Minedu, 2019).

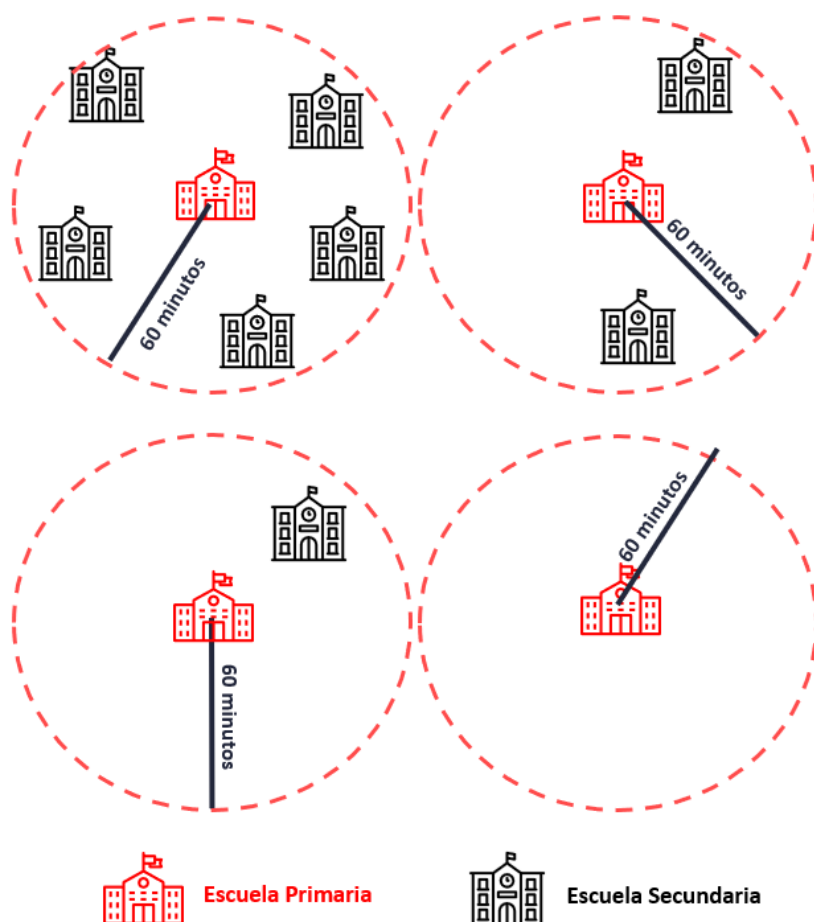
Adicionalmente, con el fin de definir la conectividad entre escuelas debido a la disponibilidad de la oferta, se utiliza como insumo el mismo ráster<sup>5</sup> de distancia para todos los años. Esto equivale a decir que el modelo no considera mejoras en las carreteras o la construcción de nuevas carreteras que permitan la reducción de tiempos de movilidad entre dos puntos geográficos. Si bien el presente estudio se enfoca en la conectividad entre escuelas, esta conectividad se enmarca única y exclusivamente en la disponibilidad de infraestructura educativa. Es preciso mencionar que no se incluye dichas mejoras en las carreteras para el análisis debido a que se desea capturar el efecto de la variación en la disponibilidad de la oferta dado todo lo demás constante. Por lo que el uso del mismo ráster en todos los periodos de análisis es un buen *proxy* al hecho de mantener cualquier otro hecho sin cambio alguno.

La **Figura 9** muestra un ejemplo simplificado de lo explicado, considerando áreas de influencia circulares con un radio de 60 minutos.

---

<sup>5</sup> Un ráster es una matriz compuesta por celdas (o píxeles) que se organizan en filas y columnas (también llamadas cuadrículas). Cada una de las celdas en la matriz contiene valores que representan cierta información, como la temperatura. A su vez, el tamaño de cada celda puede representar la dimensión de un determinado espacio. Así, el moverse de un punto hacia otro en un mapa, indicaría recorrer cierta cantidad de celdas, lo cual, a través de algunas conversiones, resulta en el tiempo de movilidad entre puntos geográficos (ArcMap, 2021).

Figura 9. Áreas de influencia en un radio de 60 minutos

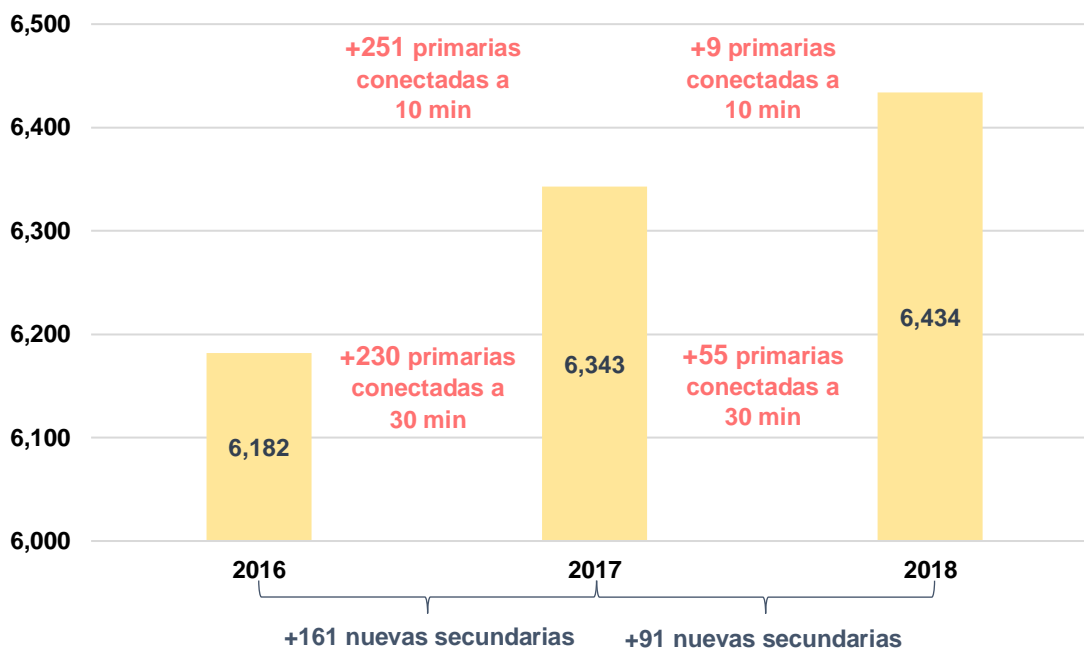


De esta manera, se considera que una institución educativa de nivel primaria está conectada a una de nivel secundaria si en su área de influencia con un radio de “x” minutos hay, por lo menos, una escuela secundaria. En el ejemplo de la **Figura 9**, sólo la escuela ubicada en la parte inferior derecha se considera como no conectada, mientras que las otras tres sí lo estarían. Así, se espera que el hecho de transitar desde la educación primaria hacia la educación secundaria en una escuela no conectada implique un mayor costo (en términos de esfuerzo) para las familias de los estudiantes.

Ahora bien, en el presente estudio, la conectividad de escuelas es un concepto que está ligado a la disponibilidad de la oferta educativa. Como se mostró anteriormente en la **Figura 5**, la oferta educativa de nivel secundaria se ha expandido en los últimos años. Para las zonas rurales, entre 2016 y 2018 se registró un aumento de 252 nuevas escuelas de nivel secundaria (71% de la expansión total). Esta expansión significó un aumento de 260 en el número de escuelas de nivel primaria conectadas a una secundaria en un rango de 10 minutos y de 285 en un rango de 30 minutos.

El desglose del aumento de escuelas conectadas se observa en la **Figura 10**. Mientras que el mayor incremento está entre 2016 y 2017; en el siguiente año, las cifras en conectividad y expansión de la oferta educativa de nivel secundaria son relativamente bajas. Esto se puede deber al fenómeno del niño ocurrido en 2017, el cual afectó a más de 3 mil escuelas en todo el Perú, de las cuales 318 quedaron destruidas e inhabitables (Instituto Nacional de Defensa Civil, 2017).

**Figura 10. Expansión de la oferta de secundarias y conectividad**



Fuente: ESCALE (2016 – 2018).

Si bien la mayoría de estudios que analizan la disponibilidad de la oferta asumen que esta es exógena, Mendelsohn (1996) menciona que la distribución geográfica de las escuelas es un elemento endógeno, pues responde a variables de la demanda que el Estado considera al momento de su construcción (Ndaendelao, 2007; Odhiambo e Imwati, 2014). Así, se espera que exista algún tipo de sesgo de selección a nivel de centro poblado, ya que se estaría comparando áreas con mayor acceso a escuelas (o con mayor número de escuelas) con áreas con menor acceso. Por ejemplo, no es lo mismo la conectividad entre escuelas de Lima y la conectividad entre escuelas de Loreto.

En cuanto a la identificación del efecto causal, se debe distinguir entre el efecto de la selección y el de acceder a una escuela primaria conectada sobre la deserción escolar. El primero respondería al acceso que enfrentan las familias producto de su ubicación y sus preferencias entre enviar o no al hijo a la escuela. Incluso si se controla por características observables, cabe la posibilidad de que existan características no observables (como la prioridad que le dan los padres a la educación)

que motiven a ciertas familias a esforzarse más por buscar una escuela conectada con el fin de brindarle continuidad a la educación de sus hijos. Por ende, una especificación con efectos fijos resulta útil para reducir este tipo de sesgo.

Así, la línea de decisión entre las familias es la siguiente. Los hogares envían a sus hijos a una escuela primaria, la cual puede o no estar conectada a una secundaria. Una vez que el niño *i* termina la primaria, la familia decide si enviarlo o no a la secundaria dada la disponibilidad actual de la oferta alrededor de la primaria donde concluyó sus estudios. En otro periodo de tiempo, el hermano del niño *i* también concluye la primaria y la familia vuelve a observar la disponibilidad de la oferta alrededor de esa primaria para tomar la decisión de enviar a su hijo o no a la secundaria. Cabe resaltar que, la disponibilidad de la oferta puede estar dada por dos motivos: (i) la construcción de nuevas secundarias o (ii) la migración de las familias.

### **5.1. Datos y descripción de la muestra**

El presente trabajo utiliza los datos de tres fuentes de información. El primero es el Siagie, el cual registra los datos de todos los niños que forman parte del sistema educativo peruano. En este se puede encontrar el código de la escuela en la que asisten los niños, el nivel, el grado, la sección y algunos datos de sus apoderados como edad, nivel educativo y si trabaja.

La segunda fuente de información es el padrón de instituciones educativas del Minedu, el cual muestra la relación de todas las escuelas (activas o inactivas) del sistema educativo peruano. La importancia de este padrón radica en que recoge la información georreferenciada de todas las escuelas, ya sean de gestión pública o privadas, además del gradiente de ruralidad y el tipo de escuela que es (unidocente, polidocente o multigrado). La tercera fuente es el Censo Escolar, el cual recoge las características de las escuelas a nivel nacional (infraestructura, matrícula, número de docentes, etc.).

El periodo de análisis en el cual se desarrolla el estudio corresponde a los años 2016 – 2019. Esto debido a que, antes de 2016, el registro nominal de padres de familia y la georreferenciación de las escuelas no garantizan la confiabilidad necesaria para el estudio. Según el mismo Minedu, los datos no tienen la calidad suficiente como para llevar a cabo un estudio de este tipo. Así, la muestra está compuesta por grupos de hermanos en zonas rurales que, entre 2016 y 2018 terminaron la educación primaria en escuelas públicas y que, entre 2017 y 2019 debieron haber asistido a la educación secundaria. Además, dentro de cada familia en la muestra debe haber al menos un hermano cuya escuela primaria haya estado (o no) conectada a una secundaria, pues esta es la principal condición para tener variabilidad en la variable explicativa.

La muestra consta de 1,101 familias con dos o más hijos<sup>6</sup>. Dentro de cada familia, al menos un niño debió terminar en una primaria conectada (o no conectada) y, al menos uno, no lo haya hecho. De este modo, se encuentra que alrededor de 1,315 niños terminaron la educación primaria en una escuela no conectada en el periodo 2016 – 2018, mientras que 945 lo hicieron en una escuela conectada a una secundaria en un radio de 10 minutos. Para definir el corte de tiempo en el cual se clasifica a una escuela conectada o no conectada, se observó la distribución de la distancia de tiempo y se eligió puntos de corte en un intervalo de 10 y 20 minutos (ver **Anexo 1**). Así, según el punto de corte, la clasificación de una escuela podría variar. Por ejemplo, puede que, al considerar un radio de 10 minutos, una escuela no esté conectada; sin embargo, al considerar un radio de 30 minutos, sí lo esté.

Sin embargo, si bien la estrategia pasa por definir una muestra de hermanos, con el fin de comparar los resultados con la muestra total de estudiantes, se utilizará como insumo dicha muestra. De este modo, la **Tabla 2** muestra los estadísticos descriptivos de la muestra total de estudiantes en zonas rurales que culminaron sexto grado de primaria en una escuela pública entre los años 2016 – 2018. Estos estadísticos se determinan para dos puntos de corte de conectividad: un radio de 10 minutos y uno de 30 minutos.

Lo primero que se observa es que, a medida que se amplía el radio de distancia entre escuelas, la muestra se concentra más hacia el grupo de instituciones educativas conectadas, lo cual tiene sentido, pues se estaría ampliando el margen del área de influencia de cada primaria. A su vez, a medida que el radio se amplía, la tasa de deserción en las escuelas no conectadas aumenta de 7% a 10%; mientras que en las escuelas conectadas se mantiene alrededor de 2%. La muestra presenta cierto equilibrio a nivel de género, pero se observa que las escuelas no conectadas albergan más del doble de estudiantes que repitieron al menos un grado entre 2015 y 2018.

Asimismo, se observa que las escuelas conectadas están mucho más cerca a sus Unidades de Gestión Educativa Local (Ugel). Además, si bien toda la muestra parte de escuelas rurales, más del 30% de las escuelas conectadas suelen concentrarse en las áreas menos rurales (rural 3); mientras que más del 50% de las escuelas no conectadas están muy concentradas en las áreas más rurales (rural 1). Por otra parte, el 60% de las escuelas no conectadas en la muestra son multigrado y están ligeramente más focalizadas por alguna intervención del Minedu<sup>7</sup>. Finalmente, se observa que las escuelas conectadas tienen mayor acceso a servicios básicos y son más grandes (mayor número de alumnos matriculados).

---

<sup>6</sup> Si bien la idea es comparar hermanos, la muestra también considera a medio hermanos debido a que se cumple con el requisito de que ambos niños están expuestos al mismo ambiente familiar.

<sup>7</sup> Las intervenciones que son consideradas en este caso son: Soporte Pedagógico y Acompañamiento Pedagógico Multigrado o de Educación Intercultural Bilingüe (EIB).

Por su parte, la **Tabla 3** corresponde a los estudiantes con las mismas características de la tabla anterior, pero que cuyo hermano menor también cursó el sexto grado de primaria en una escuela pública rural en el mismo periodo. En este caso, los resultados son muy similares a los hallados en la tabla anterior y también se agrega que el orden de nacimiento de los estudiantes en escuelas conectadas suele ser ligeramente mayor que el de los estudiantes en escuelas no conectadas (hermanos menores).

**Tabla 2. Estadísticos descriptivos (Muestra total)**

Variables	Radio de 10 minutos		Radio de 30 minutos		N
	Conectada	No conectada	Conectada	No conectada	
<b>Tasa de deserción de la IE</b>	0.018 (0.000)	0.078 (0.001)	0.025 (0.000)	0.102 (0.001)	365,572
<b>Sexo del estudiante (1 si es mujer)</b>	0.488 (0.001)	0.486 (0.001)	0.487 (0.001)	0.486 (0.002)	365,572
<b>El estudiante repitió algún Grado entre 2015 y 2018</b>	0.047 (0.000)	0.103 (0.001)	0.053 (0.000)	0.130 (0.001)	365,572
<b>Rendimiento en Lenguaje en 5to de primaria (max = 4)</b>	3.092 (0.001)	3.060 (0.001)	3.087 (0.001)	3.049 (0.001)	362,491
<b>Rendimiento en Matemáticas en 5to de primaria (max = 4)</b>	3.098 (0.001)	3.066 (0.001)	3.093 (0.001)	3.055 (0.001)	362,492
<b>Edad del apoderado</b>	41.824 (0.040)	41.901 (0.072)	41.849 (0.032)	41.856 (0.146)	365,451
<b>Distancia de la IE a la Ugel (en minutos)</b>	196.102 (0.716)	295.566 (1.220)	189.846 (0.603)	431.165 (2.142)	365,299
<b>IE es Rural 1</b>	0.267 (0.001)	0.599 (0.001)	0.291 (0.001)	0.824 (0.002)	365,572
<b>IE es Rural 2</b>	0.380 (0.001)	0.323 (0.001)	0.400 (0.001)	0.166 (0.002)	365,572
<b>IE es Rural 3</b>	0.353 (0.001)	0.078 (0.001)	0.309 (0.001)	0.010 (0.000)	365,572
<b>IE es Multigrado</b>	0.340 (0.001)	0.670 (0.001)	0.415 (0.001)	0.638 (0.002)	365,572
<b>IE cuenta con alguna intervención del Ministerio</b>	0.301 (0.001)	0.352 (0.001)	0.304 (0.001)	0.392 (0.002)	365,572
<b>IE tiene paredes de material noble</b>	0.910 (0.001)	0.835 (0.001)	0.907 (0.001)	0.768 (0.002)	359,871
<b>IE tiene techos de material noble</b>	0.948 (0.000)	0.954 (0.001)	0.950 (0.000)	0.950 (0.001)	359,877
<b>IE tiene pisos de material noble</b>	0.792 (0.001)	0.744 (0.001)	0.785 (0.001)	0.729 (0.002)	359,857
<b>IE tiene acceso a electricidad</b>	0.892 (0.001)	0.733 (0.001)	0.884 (0.001)	0.604 (0.002)	360,316
<b>IE tiene acceso a red pública de agua</b>	0.737 (0.001)	0.542 (0.001)	0.711 (0.001)	0.469 (0.002)	359,290
<b>IE tiene acceso a desagüe</b>	0.567 (0.001)	0.296 (0.001)	0.519 (0.001)	0.255 (0.002)	358,997
<b>Total de alumnos en la IE</b>	124.509 (0.198)	41.216 (0.096)	108.288 (0.171)	35.962 (0.104)	365,572
<b>N</b>	241,805	123,767	305,025	60,547	
<b>Proporción</b>	66%	34%	83%	17%	

**Notas:** 1) La muestra total consta de 365,572 estudiantes. En caso de que alguna variable muestre un menor número de observaciones, se infiere de que no existe información en dicha variable para un número de estudiantes. 2) IE = Institución Educativa. 3) Para el nivel primaria, las notas registradas en Siagie tienen una escala desde C hasta AD, donde C es la nota más baja y AD es la más alta. Para el presente trabajo se transformó dicha escala a una numérica donde: AD = 4, A = 3, B = 2 y C = 1.

**Tabla 3. Estadísticos descriptivos (Muestra de hermanos)**

Variables	Radio de 10 minutos		Radio de 30 minutos		N
	Conectada	No conectada	Conectada	No conectada	
<b>Tasa de deserción</b>	0.029 (0.005)	0.081 (0.008)	0.047 (0.005)	0.092 (0.012)	2,260
<b>Sexo del estudiante (1 si es mujer)</b>	0.503 (0.016)	0.486 (0.014)	0.498 (0.012)	0.478 (0.020)	2,260
<b>Orden de nacimiento del estudiante</b>	1.654 (0.017)	1.457 (0.015)	1.575 (0.013)	1.444 (0.023)	2,260
<b>El estudiante repitió algún grado</b>	0.102 (0.010)	0.115 (0.009)	0.102 (0.007)	0.129 (0.013)	2,260
<b>Rendimiento en Lenguaje en 5to de primaria (max = 4)</b>	3.054 (0.009)	3.047 (0.007)	3.057 (0.007)	3.033 (0.009)	2,238
<b>Rendimiento en Matemáticas en 5to de primaria (max = 4)</b>	3.054 (0.009)	3.055 (0.008)	3.060 (0.007)	3.041 (0.011)	2,238
<b>Edad del apoderado</b>	42.822 (0.350)	42.596 (0.298)	42.483 (0.262)	43.239 (0.452)	2,260
<b>Distancia de la IE a la Ugel (en minutos)</b>	257.232 (11.368)	273.585 (9.449)	229.214 (7.799)	365.807 (15.955)	2,260
<b>IE es Rural 1</b>	0.431 (0.016)	0.611 (0.013)	0.446 (0.012)	0.773 (0.017)	2,260
<b>IE es Rural 2</b>	0.362 (0.016)	0.297 (0.013)	0.369 (0.012)	0.206 (0.016)	2,260
<b>IE es Rural 3</b>	0.207 (0.013)	0.091 (0.008)	0.185 (0.010)	0.021 (0.006)	2,260
<b>IE es Multigrado</b>	0.486 (0.016)	0.643 (0.013)	0.547 (0.012)	0.657 (0.019)	2,260
<b>IE cuenta con alguna intervención del Ministerio</b>	0.274 (0.015)	0.353 (0.013)	0.295 (0.011)	0.385 (0.020)	2,260
<b>IE tiene paredes de material noble</b>	0.870 (0.011)	0.833 (0.010)	0.875 (0.008)	0.779 (0.017)	2,225
<b>IE tiene techos de material noble</b>	0.952 (0.007)	0.956 (0.006)	0.955 (0.005)	0.952 (0.009)	2,225
<b>IE tiene pisos de material noble</b>	0.806 (0.013)	0.775 (0.012)	0.801 (0.010)	0.753 (0.018)	2,225
<b>IE tiene acceso a electricidad</b>	0.867 (0.011)	0.709 (0.013)	0.852 (0.009)	0.570 (0.020)	2,224
<b>IE tiene acceso a red pública de agua</b>	0.742 (0.014)	0.563 (0.014)	0.698 (0.011)	0.479 (0.020)	2,212
<b>IE tiene acceso a desagüe</b>	0.577 (0.016)	0.309 (0.013)	0.495 (0.012)	0.224 (0.017)	2,209
<b>Total de alumnos en la IE</b>	102.792 (2.877)	49.901 (1.274)	82.813 (1.973)	43.522 (1.231)	2,260
<b>N</b>	945	1,315	1,639	621	
<b>Proporción</b>	42%	58%	73%	28%	

**Notas:** 1) La muestra total consta de 1,101 familias con un total de 2,260 estudiantes. Cada familia puede tener entre dos a cuatro hermanos. En caso de que alguna variable muestre un menor número de observaciones, se infiere de que no existe información en dicha variable para un número de estudiantes. 2) IE = Institución Educativa. 3) Para el nivel primaria, las notas registradas en Siagie tienen una escala desde C hasta AD, donde C es la nota más baja y AD es la más alta. Para el presente trabajo se transformó dicha escala a una numérica donde: AD = 4, A = 3, B = 2 y C = 1.

## 5.2. Estrategia empírica

Formalmente, se plantea la siguiente especificación:

$$Abandono_{ijdt} = \alpha_j + \lambda_d + \beta SC_{ijt} + X'\gamma + \varepsilon_{ijt} \quad (1)$$

Donde la variable de interés es la probabilidad de que el niño  $i$  de la familia  $j$  ubicada en el distrito  $d$  culmine sus estudios de nivel primaria en el año  $t$  y no inicie sus estudios de nivel secundaria en el año  $t + 1$ .  $\alpha_j$  representa un efecto fijo por familia.  $\lambda_d$  es un efecto fijo a nivel de distrito.  $SC_{ijt}$  es una variable dicotómica que toma el valor de 1 en caso la escuela donde el niño  $i$  de la familia  $j$  culminó sus estudios de nivel primaria no esté conectada a una escuela de nivel secundaria (o no tenga una escuela de nivel secundaria lo suficientemente cerca) y cero de otro modo.  $X$  es un vector de variables de control que varían a nivel de familia y entorno, como la edad del niño, su sexo, edad del apoderado, entre otros. Finalmente,  $\varepsilon_{ijt}$  es el error de regresión asociado a la especificación.

De este modo, el efecto fijo a nivel familia controla todas aquellas características observables o no observables que influyen sobre la decisión de que el niño continúe con sus estudios y transite hacia el nivel secundaria y son constantes en la familia. Estas características incluyen el nivel socioeconómico, la educación de los padres, el sexo del jefe del hogar, entre otros. Asimismo, el efecto fijo a nivel de distrito controla por todas las características observables o no observables que comparten determinados distritos en cuanto a políticas educativas y que podrían influenciar sobre la decisión de continuar los estudios secundarios para un niño. La idea detrás de este efecto fijo es controlar por la heterogeneidad que tienen distintos puntos geográficos a pesar de compartir la característica de ruralidad.

No se considera un efecto fijo por año debido a que se busca aprovechar la variabilidad de la oferta educativo entre años. Como se observó en la **Figura 10**, la fuente de variación exógena está estrictamente relacionada con el tiempo; especialmente, entre los años 2016 y 2017. Esto debido a que la comparación entre hermanos se da, precisamente, entre diferentes años. Para entender mejor esto, se puede pensar en cuál sería un buen experimento para medir el efecto de la conectividad entre escuelas. En ese caso, lo ideal sería tener al mismo niño y que egrese de la misma escuela en dos contextos diferentes: el primero, cuando la escuela no está conectada a una secundaria y, el segundo, cuando sí lo está. Debido a que este experimento no es posible, lo más cercano a tener al mismo niño en la muestra es tener a su hermano, ya que ambos se enfrentan el mismo contexto familiar (o muy similar).

Por otro lado, dado que no se puede tener a la misma escuela con distinta condición de conectividad a la vez, se puede comparar a la misma escuela (o escuelas similares) en dos momentos diferentes en el tiempo: uno cuando no estaba conectada a una secundaria y el otro

cuando sí lo estaba. Finalmente, no es necesario controlar por un efecto de tiempo, pues el objetivo del estudio es aprovechar cualquier evento que genere cambios en la conectividad de escuelas en el tiempo y, de hacerlo, se podría no identificar correctamente al parámetro de interés.

En cuanto a la estimación del modelo, dado que la variable de resultado califica como una binaria (1 si es desertor, 0 de otro modo), se esperaría que esta se desarrolle a través de un modelo logístico, pues los modelos de probabilidad lineal suelen presentar algunos problemas (Cameron y Triveli, 2005; Baltagi, 2011). No obstante, Angrist y Pischke (2008) sostienen que la preferencia a los modelos no lineales como *logit*, Probit o Tobit se mantienen siempre que se tenga fines predictivos. En cambio, si se busca evaluar el efecto causal promedio, no se requiere *hacer* algo especial.

Para ello, se parte de la siguiente premisa. Considere que la variable de resultados es  $Y_i$  y se busca evaluar el efecto de la variable  $D_i$  sobre esta, donde  $D_i$  es una variable dicotómica que toma el valor de 1 si se cumple una determinada condición y cero de otro modo. Bajo el supuesto de tener datos experimentales, para identificar el efecto de  $D_i$  sobre  $Y_i$  se plantea:

$$\begin{aligned}
 E[Y_i|D_i = 1] - E[Y_i|D_i = 0] \\
 = \underbrace{(E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 1])}_{\text{Efecto promedio del tratamiento sobre el tratado}} + \underbrace{(E[Y_{0i}|D_i = 1] - E[Y_{0i}|D_i = 0])}_{\text{Sesgo de selección}}
 \end{aligned}$$

Dado que se cumple el supuesto de independencia entre  $D_i$  sobre la variable de resultado potencial ( $Y_i$ ),

$$E[Y_i|D_i = 1] - E[Y_i|D_i = 0] = E[Y_{1i}|D_i = 1] - E[Y_{0i}|D_i = 1] = E[Y_{1i} - Y_{0i}] = \hat{\beta}$$

Donde  $\hat{\beta}$  es el coeficiente de la regresión de  $Y_i$  sobre  $D_i$ . Asimismo, se observa que, la estimación de efectos causales no presenta cambios especiales si  $Y_i$  es binario, no negativo o distribuida de manera continua, siempre que se cumpla el supuesto de independencia. Por lo tanto, el presente modelo se estimará de manera lineal controlando por efectos fijos a nivel de familia a fin de garantizar la exogeneidad de la variable explicativa.

## Capítulo VI. Resultados

### 6.1. Conectividad y deserción escolar

La **Tabla 4** muestra los resultados de cómo afecta el hecho de terminar la primaria en una escuela no conectada sobre la deserción escolar para la muestra total y dos submuestras. La primera submuestra considera el conjunto de hermanos en las 1,101 familias (columnas 2 y 3) y la segunda

submuestra sólo considera el conjunto de hermanos que egresaron de las mismas escuelas primarias (columna 4). Cabe resaltar que, para la primera submuestra, se generó dos regresiones: una sin considerar efectos fijos a nivel familia (columna 2) y otra considerando dichos efectos fijos (columna 3).

En cada panel se muestran las regresiones para las muestras detalladas anteriormente, pero con diferentes especificaciones. El panel A muestra los resultados de una regresión sin controles para áreas de influencia de un radio de 10 y 30 minutos en las primarias. El panel B muestra los resultados de una regresión que incluye controles a nivel de estudiante como su sexo, si repitió algún grado y su rendimiento académico en los cursos de comunicación integral y matemáticas en quinto grado de primaria. El panel C muestra los resultados de una regresión que incluye controles a nivel de estudiante y de escuela como si es una multigrado, si cuenta con al menos una intervención del ministerio, el total de estudiantes en la escuela, entre otros.

Así, los resultados para la muestra total de estudiantes que egresaron de la primaria en zonas rurales entre 2016 y 2018 muestran que, egresar de una IE no conectada en un radio de 10 minutos, incrementa la probabilidad de no continuar con los estudios de nivel secundaria en 6 pp. No obstante, al controlar por características de estudiantes o de escuela, el parámetro estimado se reduce a 3 pp. Si se considera la información detallada en la **Figura 3**, la probabilidad base de que un estudiante que termina sexto de primaria no continúe sus estudios de nivel secundaria es de 4%. Bajo esa premisa, el hecho de egresar de una escuela no conectada a una secundaria en un radio de 10 minutos hace que su probabilidad incremente a 7%, un número bastante cercano a la tasa de deserción promedio mostrado en la **Tabla 2**. Los resultados son similares si se considera un radio de 30 minutos.

Al limitar la muestra al grupo de hermanos, se observa que el tamaño del efecto cae en el panel A, cuando no se consideran controles o se mantiene como en el panel C, cuando se consideran controles a nivel de estudiantes y escuela. Asimismo, se observa que la caída de la magnitud es mayor cuando se controla por efectos fijos de familia. Esto indicaría que la muestra total estaría sobreestimando el parámetro y la dirección del sesgo sería positivo. No obstante, la significancia se mantiene en todos los casos por lo que se puede afirmar que el hecho de que un hermano egrese de una primaria no conectada en un radio de 10 minutos aumenta la probabilidad de que no continúe sus estudios entre 2 y 4 pp. En el caso de un radio de 30 minutos, este aumento es de entre 3 y 5 pp.

**Tabla 4. Conectividad y deserción escolar**

Variables	Muestra total	Muestra hermanos (sin EF de familia)	Muestra hermanos (con EF de familia)	Muestra hermanos de misma IE (con EF de familia)
<i>Panel A: Sin controles</i>				
Egresó de una IE no conectada (10 min)	0.06*** (0.001)	0.05*** (0.011)	0.04*** (0.011)	0.04*** (0.014)
Egresó de una IE no conectada (30 min)	0.07*** (0.001)	0.06*** (0.013)	0.05*** (0.014)	0.06*** (0.017)
N	<b>365,572</b>	<b>2,260</b>	<b>2,260</b>	<b>1,188</b>
<i>Panel B: Controles de estudiantes</i>				
Egresó de una IE no conectada (10 min)	0.03*** (0.001)	0.04*** (0.008)	0.03*** (0.009)	0.03** (0.011)
Egresó de una IE no conectada (30 min)	0.04*** (0.001)	0.04*** (0.010)	0.04*** (0.011)	0.04*** (0.013)
N	<b>362,490</b>	<b>2,238</b>	<b>2,238</b>	<b>1,178</b>
<i>Panel C: Controles de estudiantes y escuela</i>				
Egresó de una IE no conectada (10 min)	0.03*** (0.001)	0.03*** (0.009)	0.02** (0.010)	0.02 (0.012)
Egresó de una IE no conectada (30 min)	0.03*** (0.001)	0.03** (0.010)	0.03*** (0.012)	0.04*** (0.014)
N	<b>355,306</b>	<b>2,184</b>	<b>2,184</b>	<b>1,149</b>

**Nota:** Los errores estándar robustos en paréntesis.

\*p < 0:1; \*\*p < 0:05; \*\*\*p < 0:01.

Todas las regresiones se controlan por efectos fijos de distrito.

**Controles de estudiantes:** Orden de nacimiento del estudiante (sólo muestra hermanos), sexo, si repitió un grado, desempeño en comunicación integral y matemáticas.

**Controles de escuela:** es multigrado, tiene al menos una intervención por parte del ministerio, tiene piso de material noble, tiene acceso a agua de una red pública, tiene acceso a desagüe y total de alumnos.

Dado que la probabilidad base de la muestra de hermanos es 6%, un incremento de 2 pp. o 3 pp. implicaría que su probabilidad de deserción en la transición de la primaria hacia la secundaria pase a 8% en caso egrese de una primaria no conectada en un radio de 10 minutos o a 9% en caso egrese de una primaria no conectada en un radio de 30 minutos. Ambos porcentajes son muy similares a los observados en la **Tabla 3**. En todos los casos, se observa que la introducción de variables de control reducen el tamaño del efecto, por lo que estarían corrigiendo la sobreestimación inicial.

Finalmente, al limitar la muestra considerando sólo grupos de hermanos que egresaron de una misma escuela primaria, se observa que los efectos son muy similares a los hallados en la columna anterior. Así, para aquellos hermanos que egresaron de una escuela primaria cuando esta no estaba conectada con una secundaria en un radio de 10 minutos, la probabilidad de no continuar sus estudios de nivel secundaria incrementa en 4 pp. Este efecto incrementa a 6 pp. cuando el radio del área de influencia pasa de 10 a 30 minutos. Al incluir variables de control, el tamaño de los efectos se reducen a 2 y 4 pp., respectivamente. Si bien en el segundo caso no cambia la significancia del parámetro estimado, en el primer caso sí lo hace. No obstante, al mantener el

mismo tamaño y signo que en la columna anterior, se concluye que este cambio podría deberse a la pérdida de poder estadístico debido a la reducción de la muestra.

La importancia de esta última submuestra se alinea con Bilal y Rossi-Hansberg (2021), quienes mencionan que la migración es un factor importante al momento de evaluar la oferta de oportunidades que tienen las personas. De este modo, si el grupo de hermanos termina la primaria en la misma escuela, se reduce la probabilidad de que la conectividad a la escuela secundaria más cercana se deba, específicamente, a un cambio de escuela entre hermanos (migración). Por ende, cualquier cambio en la conectividad se debería a un choque de oferta en las escuelas secundarias, lo cual es el principal foco del presente estudio.

### 6.3. Pruebas de robustez

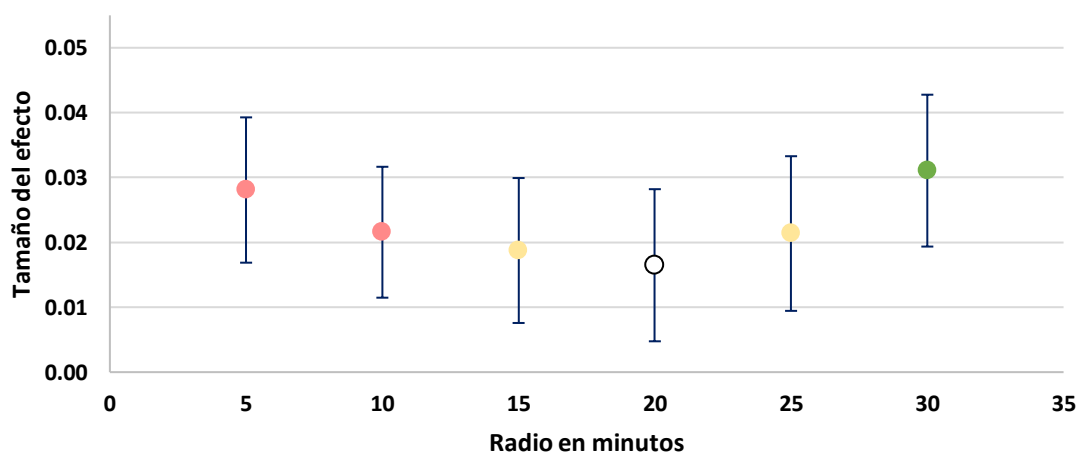
La sección anterior muestra que, independientemente de la muestra que se utilice y del tamaño del radio que se utilice para definir la no conectividad, esta afecta positiva y significativamente sobre la probabilidad de deserción en la transición primaria – secundaria. Adicionalmente, se observa que incrementar el radio del área de influencia de una escuela primaria aumenta ligeramente el tamaño del efecto. No obstante, a fin de probar la robustez en los resultados, se procederá a hacer dos ejercicios adicionales: (i) cambiar el corte (en minutos) en el que se evalúa la conectividad para revisar si se mantiene la dirección y la significancia del efecto; y (ii) definir conectividad en función de la distancia en tiempos entre las escuelas primarias y las secundarias más cercanas a estas. Adicionalmente, en los **Anexos 2 y 3** se hace dos ejercicios más en los que se cambia la muestra y la especificación de los modelos a fin de probar robustez.

En primer lugar, la **Figura 11** muestra el tamaño del efecto encontrado para diversos radios del área de influencia de las primarias para evaluar su conectividad. Así, en el eje horizontal, se traza seis (6) radios de áreas de influencia de las escuelas primarias y en el eje vertical se tiene el tamaño del parámetro estimado ( $\beta$ ) que captura el efecto sobre la deserción escolar de terminar la primaria en una escuela no conectada a una secundaria en un radio de  $M$  minutos. La muestra utilizada en este ejercicio corresponde a la primera submuestra de hermanos de zonas rurales que egresaron de la primaria entre los años 2016 – 2018. La especificación del modelo utilizado para estimar el parámetro en cuestión considera efectos fijos a nivel familia y distrito. Además, controla por las características propias de cada estudiante detalladas anteriormente y de las escuelas de las que egresan.

Se observa que, independientemente del radio que se utilice para definir las áreas de influencia de las escuelas primarias, el hecho de egresar de una escuela no conectada a una secundaria incrementa la probabilidad de no continuar los estudios de nivel secundaria entre 0.5 y 4 pp. El

parámetro estimado mantiene tanto su dirección como su significancia en casi todos los casos. Asimismo, a medida que el radio se va ampliando, se observa una especie de caída en el tamaño del efecto hasta cuando el corte llega a 20 minutos y también de la significancia. De hecho, en el corte de 20 minutos, el efecto deja de ser significativo. No obstante, a partir de este, el efecto se hace más grande y también incrementa su significancia.

**Figura 11. Tamaño del efecto según corte de radio**



Significativo al:  
● 99% de confianza      ● 95% de confianza      ● 90% de confianza

**Nota:** Los puntos representan el tamaño del efecto estimado mientras que las bandas representan los intervalos de confianza de la estimación.

Lo último se podría explicar debido a la definición misma de la variable. Dado que la variable de interés hace referencia a que la escuela no tiene una secundaria cerca de un radio de 5 minutos, no se especifica si esta tiene una secundaria cerca de un radio de 6 minutos o de 20 minutos. Así, la distribución en la muestra podría ser muy similar en los primeros cuatro casos (escuelas que no tienen una secundaria cerca de un radio de 5 minutos, tampoco la tienen en un radio de 10, 15 o 20 minutos), lo cual hace que la importancia del corte se vaya reduciendo hasta el corte de 20 minutos. Luego, el cambio en la conectividad en cada radio sí representaría un cambio en la distribución de la muestra, lo cual generaría mayor importancia a dichos radios.

En línea con lo anterior y, considerando la información de la **Tabla 3**, el máximo corte que se ha considerado en este análisis es el de 30 minutos (donde el 73% de la muestra egresa de una escuela conectada y el 27% de una no conectada). Pues de utilizar radios más amplios, se podría desbalancear fuertemente la distribución entre escuelas conectadas y no conectadas y tener consecuencias sobre la captura del efecto en cuestión.

En segundo lugar, para establecer la definición de conectividad, se utilizó como insumo el tiempo en minutos entre una primaria y la escuela secundaria más cercana en un radio máximo de 60

minutos. En esta línea, un primer ejercicio es ver qué sucede con la variable a explicar (deserción escolar) a medida que una primaria se distancia más (o menos) de una secundaria. La **Tabla 5** muestra los resultados de la regresión entre deserción escolar en la transición primaria – secundaria y la distancia en minutos entre la escuela primaria donde el estudiante egresa y la escuela secundaria más cercana a esta.

La dinámica de la tabla es similar a la de la **Tabla 4**, cada panel incluye una especificación distinta ya sea sin considerar variables de control o considerándolas. Asimismo, para este caso sólo se consideran dos muestras: la muestra total y la muestra de hermanos controlando por efectos fijos a nivel de familia. Esto debido a que los resultados son más limpios en este caso y muy similares a la muestra de hermanos que egresan de la misma institución educativa. Cabe resaltar que no se considera la última muestra debido a que es muy reducida y podría generar problemas con el poder estadístico. Además, dado que los resultados son muy similares con la muestra inicial de hermanos, no debería haber mayor problema entre usar una muestra u otra.

De esta manera, se observa que, independientemente de la muestra, al utilizar una especificación que incluya variables de control, un minuto adicional de distancia generaría un incremento en la probabilidad de no continuar los estudios del nivel secundaria de 0.1 pp. Este valor se duplica en la muestra total cuando no se incluye variables de control. Por lo cual parecería estar sobrestimada en dicho escenario.

**Tabla 5. Distancia a la secundaria más cercana y deserción escolar**

<b>Variab</b> les	<b>Muestra total</b>	<b>Muestra hermanos (con EF de familia)</b>
<i>Panel A: Sin controles</i>		
<b>Tiempo en minutos a la secundaria más cercana</b>	0.002*** (0.000)	0.001*** (0.000)
<b>N</b>	<b>336,831</b>	<b>2,014</b>
<i>Panel B: Controles de estudiantes</i>		
<b>Tiempo en minutos a la secundaria más cercana</b>	0.001*** (0.000)	0.001*** (0.000)
<b>N</b>	<b>334,288</b>	<b>1,997</b>
<i>Panel C: Controles de estudiantes y escuela</i>		
<b>Tiempo en minutos a la secundaria más cercana</b>	0.001*** (0.000)	0.001*** (0.000)
<b>N</b>	<b>327,990</b>	<b>1,945</b>

**Nota:** Los errores estándar robustos en paréntesis.

\*p < 0:1; \*\*p < 0:05; \*\*\*p < 0:01.

**Controles de estudiantes:** Orden de nacimiento del estudiante (sólo muestra hermanos), sexo, si repitió un grado, desempeño en comunicación integral y matemáticas.

**Controles de escuela:** es multigrado, tiene al menos una intervención por parte del ministerio, tiene piso de material noble, tiene acceso a agua de una red pública, tiene acceso a desagüe y total de alumnos.

Los resultados se alinean con lo hallado por Sharma y Levinson (2019) y Rolleston, Akyeampong, Ampiah y Lewin (2010), quienes encuentran que la distancia física entre el hogar y la escuela es un factor clave para explicar el abandono escolar. En este caso, si se toma como supuesto que la escuela primaria a la que un niño asiste está relativamente cerca a su hogar, la distancia entre esta y la escuela secundaria más cercana tendría similar importancia al momento de tomar la decisión de ir a la escuela y continuar con sus estudios<sup>8</sup>.

Finalmente, los **Anexos 2 y 3** muestran dos ejercicios adicionales para revisar la robustez del parámetro estimado. En el primer caso, se replica las columnas 1 y 3 de la **Tabla 4**, pero se agrega una variable adicional que controla por un evento sucedido en el año 2017: la huelga magisterial de profesores. De acuerdo con Jaume y Willén (2019), la exposición de los niños a eventos como huelgas de docentes incide sobre la escolaridad reduciéndola. Por lo que podría considerarse que afecta, de algún modo, a la deserción escolar. En el segundo caso, se replica las columnas 1 y 3 de la **Tabla 4** pero se excluye de la muestra a las escuelas integradas<sup>9</sup>. Esto con el fin de “limpiar” la muestra de aquellas que están conectadas automáticamente por el hecho de impartir más de un nivel educativo. En ambos escenarios se observa resultados muy similares a los de la **Tabla 4**. Por lo tanto, se podría afirmar que los hallazgos son robustos a diferentes muestras, diferentes definiciones de conectividad y diferentes especificaciones del modelo.

#### 6.4. Efectos heterogéneos

Si bien evaluar el efecto promedio en la muestra es interesante; evaluar efectos diferenciados por grupos poblacionales podría ser aún más interesante. Por ello, un ejercicio adicional que se propone es el análisis de efectos heterogéneos según el sexo de los estudiantes y el nivel de ruralidad de las zonas en las que habitan. La muestra utilizada corresponde a los grupos de hermanos que egresaron de la primaria en zonas rurales. Esta es la misma muestra y especificación que la de la columna 3 en la **Tabla 4**.

La **Tabla 6** muestra los efectos heterogéneos de egresar de una escuela primaria no conectada a una secundaria en un radio de 10 y 30 minutos. Se observa que el tamaño del efecto cae tanto para niños como para niñas y se pierde la significancia en ambos casos. Por ende, no parece haber heterogeneidad a nivel de sexo. No obstante, en cuanto a la dirección del efecto estimado, se

---

<sup>8</sup> El supuesto toma mayor fuerza debido a que la población objetivo del presente estudio se concentra en zonas rurales, donde hay mayor dispersión de la población y los servicios educativos y menor movilidad escolar (Guadalupe y Rivera, 2020).

<sup>9</sup> Las escuelas integradas son aquellas que, en un mismo local, imparten enseñanza de más de un nivel educativo. En este contexto, se hace referencia a las escuelas que enseñan primaria y secundaria.

observa que esta se mantiene únicamente en el caso de las mujeres; mientras que en el caso de los varones, este cae hasta incluso volverse negativo.

**Tabla 6. Efectos heterogéneos a nivel de sexo y ruralidad**

<i>Muestra de hermanos (FE de familia + controles)</i>				
Variables	Sexo		Ruralidad	
	Hombre	Mujer	No rural 1	Rural 1
<b>Egresó de una IE no conectada (10 min)</b>	-0.01 (0.014)	0.01 (0.022)	0.06*** (0.016)	-0.00 (0.013)
<b>Egresó de una IE no conectada (30 min)</b>	-0.00 (0.016)	0.00 (0.026)	0.07*** (0.024)	0.02 (0.014)
<b>N</b>	<b>1,111</b>	<b>1,073</b>	<b>1,015</b>	<b>1,169</b>

**Nota:** Los errores estándar robustos en paréntesis.

\*p < 0:1; \*\*p < 0:05; \*\*\*p < 0:01.

**Controles de estudiantes:** Orden de nacimiento del estudiante (sólo muestra hermanos), sexo, si repitió un grado, desempeño en comunicación integral y matemáticas.

**Controles de escuela:** es multigrado, tiene al menos una intervención por parte del ministerio, tiene piso de material noble, tiene acceso a agua de una red pública, tiene acceso a desagüe y total de alumnos.

Por otra parte, en el caso de la ruralidad, se divide la muestra entre estudiantes que están en zonas rural 1 (las más rurales) y los que están en zonas rural 2 y 3 (semi urbanas). Se observa que el efecto de la conectividad a un radio de 10 y 30 minutos es significativo sólo en el caso de los que no están en las zonas más rurales. Esto probablemente se deba a que los radios de conectividad son relativamente bajos para el contexto de los estudiantes que habitan en zonas rural 1, donde la escuela más cercana podría estar a más de 30 minutos del hogar del niño o de la primaria a la que acude. No obstante, en el caso de los niños que habitan en zonas semiurbanas, un cambio en la conectividad entre escuelas de 5, 10 o 30 minutos sí es relevante en la toma de decisiones de continuar sus estudios de nivel secundaria.

Lo último no necesariamente quiere decir que la conectividad no es un factor relevante al tomar la decisión de continuar los estudios para las familias de zonas rural 1. De hecho, estos resultados sugieren que existen factores más allá de la conectividad que son importantes en dichas zonas. Por ello, es necesario plantear políticas alternativas que permitan reducir las brechas de desigualdad generadas por el reducido acceso a la educación en estas zonas.

## **Capítulo VII. Recomendaciones de política**

Los resultados mostrados en el capítulo anterior muestran que tener escuelas conectadas en los niveles educativos de primaria y secundaria son vitales para cerrar las brechas de acceso a la educación en la transición de ambos niveles en las zonas rurales. Adicionalmente, los resultados sugerirían que, probablemente, el efecto promedio se concentra en los niños y niñas de zonas semiurbanas, quienes presentan un tamaño del efecto por encima al del promedio.

En este sentido, esta sección busca plantear recomendaciones de política que: (i) mejoren la conectividad entre escuelas y (ii) reduzcan las brechas de acceso a la educación básica. Si bien la primera propuesta de política que puede venir a la mente es la construcción de más escuelas secundarias para expandir la oferta, la viabilidad de esta política podría ser un problema debido a los altos costos que demanda. Sin embargo, una política de reordenamiento territorial de los servicios educativos sí podría ser efectiva, en tanto se centre en ofrecer una transición limpia a lo largo de toda la educación básica regular (Guadalupe y Rivera, 2020).

Una alternativa para conectar las escuelas ya existentes es a través de la mejora de las rutas de transporte o promoviendo sistemas de transporte que respondan directamente al sistema educativo, tal y como sugieren Sharma y Levinson (2019). Sin embargo, una limitante en este estudio es la agregación de información concerniente a la mejora de rutas de transporte o la construcción de nuevas carreteras. El principal foco detrás de la idea de la conectividad es el acceso a infraestructura educativa y no necesariamente el acceso a más y mejores rutas de transporte. De esta manera, queda como agenda futura incluir este análisis y evaluar la contribución que tienen dichas mejoras sobre las familias y el acceso a la educación.

Si bien el problema en cuestión afecta a las zonas rurales, gracias al análisis de efectos heterogéneos se sabe que este problema se enfoca, principalmente, en zonas semiurbanas. Así, este tipo de política se enfocaría en las zonas semiurbanas, donde los sistemas de transporte están más desarrollados y es plausible conectar escuelas a través de los medios de transporte.

Una segunda alternativa podría ser el uso de modelos de educación híbrida para reducir las brechas de acceso a educación secundaria en estas zonas. Debido a la pandemia generada por el COVID-19, el sistema educativo se vio forzado a responder a una educación a distancia a través de recursos tecnológicos. De esta manera, se puede hacer uso de los recursos otorgados por las estrategias de Cierre de Brechas Digitales (donde se entrega *tablets* a los estudiantes y docentes) y Aprendo en Casa (una multiplataforma de educación a distancia) para “llevar” la educación secundaria a las casas de los estudiantes de escuelas cuyas primarias no estén conectadas. La evidencia internacional sugiere que estrategias similares como las Telesecundarias en México tuvieron éxito no sólo en el cierre de brechas de acceso a educación, sino también en los retornos de largo plazo de los mismos (Navarro-Sola, 2019).

Cabe resaltar que, si bien la educación híbrida ha sido bien aprovechada en algunos contextos, se debe considerar que es una opción alternativa sólo en caso no se pueda acceder a una escuela secundaria de manera presencial. Ello debido a la existencia de evidencia que muestra que la educación a distancia puede ser perjudicial para los aprendizajes (Figlio, Rush y Yin, 2013; Alpert, Couch y Harmon, 2016; Bettinger, Fox, Loeb y Taylor, 2017; Cacault, Hildebrand, Laurent-Lucchetti y Pellizzari, 2021). Además, según el Instituto Nacional de Estadística e

Informática (2020), en zonas rurales de Perú, sólo el 6% de hogares cuentan con acceso a internet. A ello se suma los resultados de un reciente estudio del Banco Mundial (2021), en el que se muestra que el cierre de las escuelas ha generado un retroceso en los aprendizajes de miles de niños y jóvenes en el país. Así, el 76% de niños podrían no entender un texto a la edad de 10 años.

Una alternativa complementaria a la anterior es la expansión de estrategias como las Formas de Atención Diversificadas (FAD) que brinda el Ministerio de Educación. En este caso, se podría implementar una estrategia similar a las Secundarias Tutoriales en las que se imparte educación semipresencial a estudiantes que se encuentran a más de 75 minutos a pie de una IE secundaria pública o privada o a las Secundaria con Residencia Estudiantil, donde los estudiantes pueden alojarse en las mismas IIEE (Guerrero, 2018). Si bien estas estrategias se emplean para niños cuya localización supera las distancias evaluadas en este estudio, se podría considerar la idea de flexibilizar dichas distancias a fin de aumentar el acceso a la educación secundaria. Este tipo de política se adaptaría adecuadamente al contexto de las zonas rurales 1.

Finalmente, para los casos en los que no sea posible la movilidad de los estudiantes, una alternativa es la movilización de docentes hacia las casas y/o comunidades de los estudiantes que terminan la primaria en escuelas no conectadas a una secundaria. Si bien el acceso a docentes es un problema latente en las zonas rurales, la contratación de “para-docentes” podría ser un recurso interesante en estas zonas a fin de cerrar las brechas de acceso a la educación (Banerjee, Duflo, Cole y Linden, 2007; Eble et al., 2021). A su vez, esta última estrategia debería estar acompañada de una adecuada capacitación y monitoreo por parte del Ministerio, a fin de que los “para-docentes” puedan llevar educación de calidad a los estudiantes (Campos y Gutiérrez, 2021).

## **Capítulo VIII. Conclusiones**

El presente trabajo tiene como finalidad evaluar si la conectividad terrestre entre las instituciones educativas de nivel primaria y secundaria a nivel nacional contribuye con la deserción escolar durante la transición entre estos niveles. Para identificar el efector marginal de la conectividad terrestre, se propuso un modelo lineal que controle por efectos fijos de familias, de distrito y otras variables explicativas a nivel de estudiantes y de escuelas. Así, el efecto fijo a nivel familia controla todas aquellas características observables o no observables que influyen sobre la decisión de que el niño continúe con sus estudios y transite hacia el nivel secundaria y son constantes en la familia. En otras palabras, gracias a este estimador, se podría reducir el potencial sesgo de selección que en la muestra.

Para la estimación se utilizó datos administrativos y georreferenciados del Minedu como el Siagie y el padrón de instituciones educativas. La muestra utilizada está compuesta por grupos de

hermanos que culminaron el sexto grado de primaria en escuelas públicas de zonas rurales entre los años 2016 y 2018. Además, es importante que en la muestra de evaluación al menos un hermano haya egresado de una escuela primaria que no tenga una escuela secundaria en un radio de 10 o 30 minutos y al menos un hermano haya egresado de una escuela primaria con dichas características. De este modo, la muestra consta de 1,101 familias con dos o más hijos, donde 1,315 niños terminaron la educación primaria en una escuela no conectada y 945 lo hicieron en una escuela conectada a una secundaria en un radio de 10 minutos.

Los resultados muestran que los estudiantes que concluyen la primaria en escuelas rurales no conectadas a escuelas secundarias en un radio de 10 minutos son 6 pp. más probables a no continuar sus estudios de nivel secundaria. Este aumento en la probabilidad es de 7 pp. en caso la primaria no esté conectada a una escuela secundaria en un radio de 30 minutos. Además, se observa que estos efectos se reducen a 3 pp. al controlar por características de los estudiantes y las escuelas.

Al utilizar la muestra de grupos de hermanos, se encuentra que los hermanos que terminan la educación primaria en escuelas no conectadas en un radio de 10 minutos son 5 pp. más probables a no continuar en el sistema educativo que aquellos que terminan su primaria en escuelas conectadas a dicho radio. Esta probabilidad de desertar incrementa en 6 pp. cuando el radio de análisis es de 30 minutos. Al controlar por características de los estudiantes y de las escuelas de las que egresan, el efecto se reduce a 2 y 3 pp., respectivamente. Estos resultados son robustos a diferentes especificaciones y diferentes cortes en los radios de las zonas de influencia.

De manera gráfica, los resultados indicarían lo siguiente. En una zona rural del Perú hay dos hermanos que han asistido a la misma escuela primaria en dos momentos diferentes. Cada uno tiene una probabilidad base de no continuar los estudios de nivel secundaria de ambos hermanos es de 6% y características propias (como su sexo, su orden de nacimiento y su rendimiento académico). Si se considera que el primero terminó la primaria cuando esta no tenía una escuela de nivel secundaria a un radio de 30 minutos y el segundo lo hizo cuando esta sí la tenía; la probabilidad de que el primero no continúe en el sistema educativo pasaría de 6% a 10%, mientras que la del segundo quedaría intacta. Esto por el simple hecho de que el primero egresó de una escuela que no estaba conectada a una secundaria mientras que el segundo lo hizo cuando esta estaba conectada.

Una explicación al mecanismo detrás de la conectividad sería los costos de búsqueda de las escuelas generado por la disponibilidad de la oferta de secundarias alrededor de las primarias. Además, se observa que si se desagrega el efecto promedio, el tamaño del mismo es mucho mayor en las zonas semiurbanas (7 pp.), donde la conectividad en un radio de entre 10 y 30 minutos

parece ser bastante relevante. Sin embargo, no se encuentra que haya heterogeneidad según el sexo de los hermanos.

A partir de los resultados, se concluye que la conectividad terrestre de las escuelas sí es un factor clave para la decisión de continuidad de los estudios. Si bien el estudio queda limitado a definir conectividad como la disponibilidad de infraestructura educativa alrededor de un punto geográfico, la agenda futura debe trabajar con metodologías complementarias que permitan agregar la información de la mejora de las rutas de transporte y la construcción de nuevas carreteras. Esto con el fin de mejorar la definición de conectividad y evaluar su efecto sobre la decisión de las familias sobre mayor escolaridad, así como de otras variables.

Dado que una de las principales contribuciones de este estudio es ser un insumo para la mejora en el debate de políticas públicas, es importante plantear algunas recomendaciones de política a fin de mejorar la conectividad entre escuelas y ampliar el acceso a la educación básica en las zonas rurales. Si bien la expansión de la oferta educativa a través de la construcción de nuevas escuelas es una política clave a raíz de este problema, también se debe considerar algunas otras alternativas como el reordenamiento territorial para una educación obligatoria fluida.

Adicionalmente, se plantea desarrollar sistemas de transporte que respondan directamente al sistema educativo, especialmente en las zonas semiurbanas. Una alternativa adicional es utilizar los recursos de las estrategias de Cierre de Brechas Digitales y Aprendo en Casa para llevar la educación secundaria a las casas de los estudiantes o flexibilizar los requerimientos para aplicar a las Formas de Atención Diversificadas que brinda el ministerio. Además, se puede promover una política de educación que incluya la movilidad de los docentes hacia las casas o comunidades de los estudiantes en las zonas más alejadas.

Finalmente, si bien el estudio contribuye a la literatura examinando la problemática de la deserción escolar desde la perspectiva de la oferta, es importante resaltar que este es solo uno de los factores que contribuyen a la persistencia de diversas brechas en cuanto al acceso de la educación y las desigualdades que eso conlleva. Por lo tanto, invertir en conectar escuelas no sólo brindaría mejores oportunidades a los estudiantes, sino que representaría uno de los pilares de desarrollo y un paso importante para el cierre de brechas que es tan necesario en el país.

## Bibliografía

- Adeyemi, G., & Edeki, S. (2018). Dataset on spatial distribution and location of universities in Nigeria. *Data in Brief*, 18: 1602–1608.
- Alcázar, L. (2008). Asistencia y deserción en escuelas secundarias rurales del Perú. En M. Benavides (ed.), *Análisis de programas, procesos y resultados educativos en el Perú: contribuciones empíricas para el debate* (págs. 41-82). Lima: GRADE.
- Alcázar, L., Rendón, S., & Wachtenheim, E. (2002). *Working and Studying in Rural Latin America: Critical Decisions of Adolescence*. Washington D.C.: Banco Interamericano de Desarrollo, Research Network Working Paper R-469.
- Alpert, W., Couch, K., & Harmon, O. (2016). A Randomized Assessment of Online Learning. *American Economic Review*, 106(5), 378–382.
- Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.
- ArcMap. (10 de 05 de 2021). *ArcGIS for Desktop*. Obtenido de <https://desktop.arcgis.com/es/arcmap/10.3/manage-data/raster-and-images/what-is-raster-data.htm>
- Baltagi, B. H. (2011). *Econometrics*. Springer Texts in Business and Economics.
- Banco Mundial. (2021). *Repensar el futuro del Perú. Notas de política para transformar al Estado en un gestor del bienestar y el desarrollo*. Lima: Grupo Banco Mundial.
- Banerjee, A., Duflo, E., Cole, S., & Linden, L. (2007). Remedying Education: Evidence from two randomized experiments in India. *The Quarterly Journal of Economics*, Vol. 122 (3).
- Bettinger, E. P., Fox, L., Loeb, S., & Taylor, E. S. (2017). Virtual Classrooms: How Online College Courses Affect Student Success. *American Economic Review*, 107(9), 2855–2875.
- Bilal, A., & Rossi-Hansberg, E. (2021). *Location as an Asset*. Princeton: Princeton University.
- Blake, J. (1989). Number of Siblings and Educational Attainment. *Science, New Series*, Vol. 245, No. 4913: 32-36.
- Bridgeland, J., Dilulio, J., & Morison, K. (2006). *The Silent Epidemic: Perspectives of High School Dropouts*. Washington.: Civic Enterprises, LLC.
- Cacault, M. P., Hildebrand, C., Laurent-Lucchetti, J., & Pellizzari, M. (2021). Distance Learning in Higher Education: Evidence from a Randomized Experiment. *Journal of the European Economic Association*.
- Cameron, A., & Trivelpic, P. (2005). *Microeconometrics: Methods and Applications*. Cambridge: Cambridge University Press.
- Campos, A., & Gutiérrez, J. (2021). *¿Focalizar para mejorar? Un análisis de las políticas y esquemas de incentivos de atracción docente en el Perú*. Lima: CIES.
- Chetty, R., & Hendren, N. (2018). The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. *The Quarterly Journal of Economics*, 133(3):1107–1162.

- Cueto, S. (2002). *Factores predictivos del rendimiento escolar, deserción e ingreso a educación secundaria en una muestra de estudiantes de zonas rurales del Perú*. Lima: Consorcio de Investigación Económica y Social (CIES).
- Cueto, S., Felipe, C., & León, J. (2020). *Predictors of school dropout across Ethiopia, India, Peru and Vietnam*. Lima: GRADE.
- Dagenais, M., Montmarquette, C., & Viennot-Briot, N. (2002). *Dropout, School Performance and Working While in School: An Econometric Model with Heterogeneous Groups*. Montreal: University of Montreal / CIRANO.
- Dowrick, P. W., & Crespo, N. (2005). School failure. En T. P. Gullotta, & G. R. Adams (eds.), *Handbook of adolescent behavioral problems: Evidence-based approaches to prevention and treatment* (págs. 589-610). New York: Springer Science and Business Media.
- Duflo, E. (2001). Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment. *The American Economic Review*, 91(4), 795-813.
- Eble, A., Frost, C., Camara, A., Bouy, B., Bah, M., Sivaraman, M., . . . Elbourne, D. (2021). How much can we remedy very low learning levels in rural parts of low-income countries? Impact and generalizability of a multi-pronged para-teacher intervention from a cluster-randomized trial in the Gambia. *Journal of Development Economics*, Vol 148.
- Field, S., Kuczera, M., & Pont, B. (2007). *No More Failures: Ten Steps to Equity in Education*. París: OECD.
- Figlio, D., Rush, M., & Yin, L. (2013). Is It Live or Is It Internet? Experimental Estimates of the Effects of Online Instruction on Student Learning. *Journal of Labor Economics*, 31(4), 763-784.
- Galabawa, J., Obeleagu, A., & Miyazawa, I. (2002). The impact of school mapping in the development of education in Tanzania: an assessment of the experiences of six districts. *Evaluation and Program Planning*, 25: 23-33.
- Guadalupe, C., & Rivera, A. (2020). *Cómo crear centros educativos integrados de Educación Básica Regular*. Lima: Universidad del Pacífico.
- Guerrero, G. (2018). *Evaluación del diseño e implementación de las formas de atención diversificada en el nivel de educación secundaria de la educación básica regular en el ámbito rural (FAD)*. Lima: GRADE.
- Henki, Z., Nismawati, & Lamadirisi, M. (2020). School Dropout (Case Study of Farmers Children in Tounsaru District of South Tondano Minahasa Regency). *Advances in Social Science, Education and Humanities Research*, 473: 73-78.
- Holland, M., & Johnson, D. R. (1984). Sibling Similarities in Educational Attainment: A Comparison of Like-Sex and Cross-Sex Sibling Pairs. *Sociology of Education*, 57(1), 11-21.
- Huisman, J., & Smits, J. (2009). Effects of household and district-level factors on primary school enrollment in 30 developing countries. *World Development*, 37, 179-193.
- Huisman, J., & Smits, J. (2015). Keeping Children in School: Effects of Household and Context Characteristics on School Dropout in 363 Districts of 30 Developing Countries. *SAGE Open*, 1-16.

- Hussain, A., Salfi, N. A., & Khan, T. M. (2011). Causes of Students' Dropout at Primary Level in Pakistan: An Empirical Study. *International Journal of Humanities and Social Science*, 1(12): 143-151.
- Instituto Nacional de Defensa Civil. (2017). *Compendio Estadístico del INDECI 2017 - Gestión Reactiva*. Lima: Instituto Nacional de Defensa Civil - INDECI.
- Instituto Nacional de Estadística e Informática. (2020). *Estadísticas de las tecnologías de la información en los hogares. Informe técnico 2*. Obtenido de [https://www.inei.gob.pe/media/MenuRecursivo/boletines/boletin\\_tics.pdf](https://www.inei.gob.pe/media/MenuRecursivo/boletines/boletin_tics.pdf)
- Jaume, D., & Willén, A. (2019). The Long-Run Effects of Teacher Strikes: Evidence from Argentina. *Journal of Labor Economics*, Vol. 37.
- Katz, L., Kling, J., & Liebman, J. B. (2001). Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment. *Quarterly Journal of Economics*, 116: 607–654.
- Lahey, B. B., & D'Onofrio, B. M. (2010). All in the Family: Comparing Siblings to Test Causal Hypotheses Regarding Environmental Influences on Behavior. *Current directions in psychological science*, 19(5), 319-323.
- Ludwig, J., Duncan, G. J., Genetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R., & Sanbonmatsu, L. (2013). Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity. *American Economic Review Papers and Proceedings*, 103: 226–231.
- Lyche, C. S. (2010). *Taking on the Completion Challenge: A Literature Review on Policies to Prevent Dropout and Early School Leaving*. OECD Education Working Papers No. 53.
- Mac Iver, D., & Mac Iver, M. A. (2009). *Beyond the Indicators: An Integrated School-level Approach to Dropout Prevention*. Arlington: The George Washington University Center for Equity and Excellence in Education.
- Markussen, E. (2010). Dropout from Education among 16-20-year-olds in the Nordic Countries. En E. Markussen (ed.), *Dropout from Education among 16-20-year-olds in the Nordic Countries* (págs. 193-226). Copenhagen: Nordic Council of Ministers.
- Markussen, E., Frøseth, M. W., Lødding, B., & Sandberg, N. (2008). Completion, drop-out and attainment of qualification in upper secondary vocational education in Norway. En H. Høst (ed.), *Continuity and Change in Norwegian Vocational Education and Training (VET)* (págs. 31-53). Oslo: Norwegian Institute for Studies in Innovation, Research and Education (NIFU STEP).
- Massey, D., & Denton, N. (1993). *American Apartheid: Segregation and the Making of the Underclass*. Cambridge, MA: Harvard University Press.
- Mendelsohn, J. M. (1996). *Education planning and management, and the use of geographical information systems*. Paris: UNESCO.
- Minedu. (2019). *Propuesta Metodológica para elaborar padrones de asignaciones temporales a docentes por ruralidad*. Lima: Repositorio Minedu.
- Ministerio de Cultura. (20 de 05 de 2021). *Lista de pueblos indígenas u originarios*. Obtenido de Base de Datos de Pueblos Indígenas u Originarios (BDPI): <https://bdpi.cultura.gob.pe/pueblos-indigenas>

- Navarro-Sola, L. (2019). *Secondary School Expansion through Televised Lessons: The Labor Market Returns of the Mexican Telesecundaria*. Working Paper.
- Naz, U., Ejaz, Z., & Khan, N. (2019). Determinants of dropout and child school enrollment: a case study from rural Islamabad. *Journal of Quantitative Methods*, 3(2), 77-89.
- Ndaendelao, E. (2007). *The Implementation of Geographic Information Systems in Namibia*. Joensuu: University of Joensuu, Publications in Social Sciences.
- Neild, R. C., Stoner-Eby, S., & Furstenberg, F. (2008). Connecting Entrance and Departure: the Transition to Ninth Grade and High School Dropout. *Education and Urban Society*, 40(5), pp. 543-569.
- Newmann, F., Wehlage, C., & Lamborn, S. (1992). The significance and sources of student engagement. En F. Newman (ed.), *Student engagement and achievement in American secondary schools* (págs. 11-39). New York: Teachers College Press.
- Nicoletti, C., & Rabe, B. (2019). Sibling spillover effects in school achievement. *Journal of Applied Econometrics*, 34(4), 482-501.
- Odhiambo, O., & Imwati, A. T. (2014). Use of Geo-Information Systems for Educational Services Provision and Planning in Asal Areas: A Case Study of Garissa County - Kenya. *International Journal of Science and Research (IJSR)*, 3(9): 2432-2446.
- Orazem, P., & Gunnarson, V. (2004). *Child Labor, School Attendance and Performance: A Review*. Ames: Iowa State University, Working Paper 4001.
- Rolleston, C., Akyeampong, K., Ampiah, J. G., & Lewin, K. M. (2010). *Educational Access in Ghana. Country Research Summary*. Consortium for Educational Access, Transitions and Equity.
- Roman, M. (2013). Factores asociados al abandono y la deserción escolar en América Latina: una mirada en conjunto. *Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, 11(2), 33-59.
- Rumberger, R. (1987). High School Dropouts: A Review of Issues and Evidence. *Review of educational research*, vol. 57, No. 2, pp. 101-121.
- Rumberger, R. W., & Lim, S. A. (2008). Why Students Drop Out of School: A Review of 25 Years of Research. *California Dropout Research Project Report*, 15.
- Sawada, Y., & Lokshin, M. (2001). *Household Schooling Decisions in Rural Pakistan*. Washington D.C.: Banco Mundial, Working Paper No. 2541.
- Sharkey, P., & Faber, J. W. (2014). Where, When, Why, and for Whom Do Residential Contexts Matter? Moving Away from the Dichotomous Understanding of Neighborhood Effects. *Annual Review of Sociology*, 40: 559-579.
- Sharma, S., & Levinson, D. (2019). Travel cost and dropout from secondary schools in Nepal. *Transportation Research Part A*, 385-397.
- Steinberg, L., Lin Blinde, P., & Chan, K. S. (1984). Dropping out Among Language Minority Youth. *Review of Educational Research*, vol. 54, No. 1, pp. 113-132.
- Symeonides, C., Vuillermin, P. J., Sciberras, E., Senn, 2. E., Thomson, S. M., Wardrop, N., . . . Ponsonby, A.-L. (2021). Importance of accounting for sibling age when examining the

association between family size and early childhood cognition, language and emotional behaviour: a birth cohort study. *BMJ Open*, 11(3), e041984.

Wilson, W. (1987). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Chicago: University of Chicago Press.

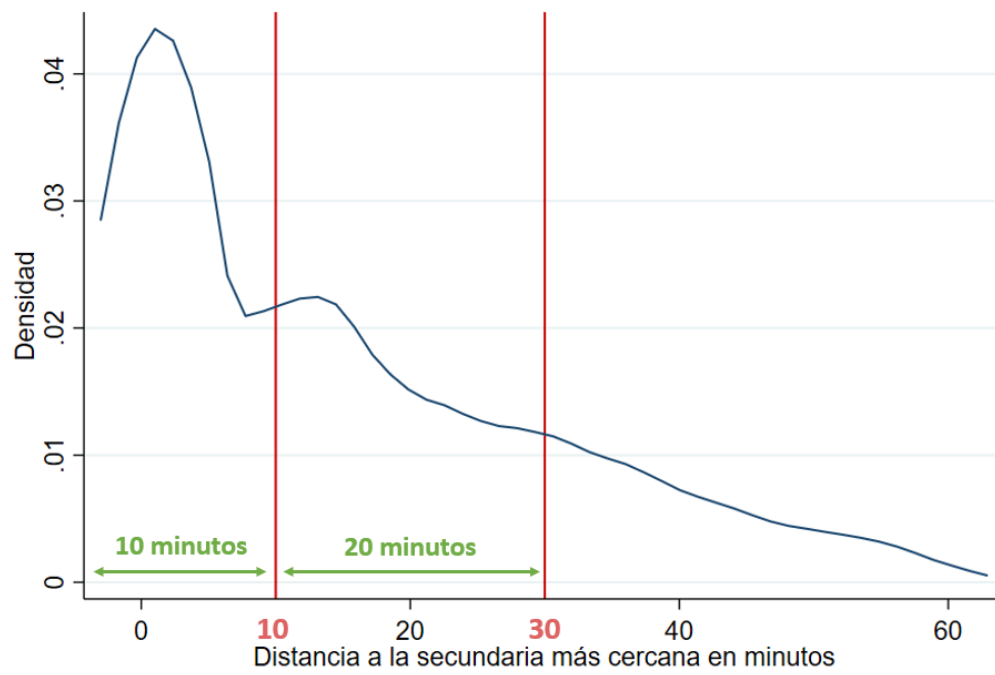
Xiong, F., Zang, L., Zhou, L., & Liu, F. (2020). The effect of number of siblings and birth order on educational attainment: Empirical Evidence from Chinese General Social Survey. *International Journal of Educational Development*, 78: 102270.

**Bases de datos consultadas:**

Instituto Nacional de Estadística e Informática (2019). *Encuesta Nacional de Hogares*. Lima.  
Consultado en: <http://inei.inei.gob.pe/microdatos/>

## Anexos

### Anexo 1. Distribución del tiempo hacia la secundaria más cercana



## Anexo 2. Conectividad y deserción considerando la huelga de docentes 2017

VARIABLES	Muestra total	Muestra hermanos (con EF de familia)
<i>Panel A: Sin controles</i>		
Egresó de una IE no conectada (10 min)	0.06*** (0.001)	0.03*** (0.012)
Egresó de una IE no conectada (30 min)	0.07*** (0.001)	0.05*** (0.014)
<b>N</b>	<b>365,572</b>	<b>2,260</b>
<i>Panel B: Controles de estudiantes</i>		
Egresó de una IE no conectada (10 min)	0.03*** (0.001)	0.02*** (0.009)
Egresó de una IE no conectada (30 min)	0.04*** (0.001)	0.04*** (0.011)
<b>N</b>	<b>362,490</b>	<b>2,238</b>
<i>Panel C: Controles de estudiantes y escuela</i>		
Egresó de una IE no conectada (10 min)	0.03*** (0.001)	0.01 (0.010)
Egresó de una IE no conectada (30 min)	0.03*** (0.001)	0.02** (0.012)
<b>N</b>	<b>355,306</b>	<b>2,184</b>

**Nota:** Los errores estándar robustos en paréntesis.

\*p < 0:1; \*\*p < 0:05; \*\*\*p < 0:01.

**Controles de estudiantes:** Orden de nacimiento del estudiante (sólo muestra hermanos), sexo, si repitió un grado, desempeño en comunicación integral y matemáticas.

**Controles de escuela:** es multigrado, tiene al menos una intervención por parte del ministerio, tiene piso de material noble, tiene acceso a agua de una red pública, tiene acceso a desagüe y total de alumnos.

### Anexo 3. Conectividad y deserción (sin considerar escuelas integradas)

Variab <span>l</span> es	Muestra total	Muestra hermanos (con EF de familia)
<i>Panel A: Sin controles</i>		
Egresó de una IE no conectada (10 min)	0.05*** (0.001)	0.04** (0.016)
Egresó de una IE no conectada (30 min)	0.06*** (0.001)	0.05*** (0.019)
<b>N</b>	<b>266,265</b>	<b>1,823</b>
<i>Panel B: Controles de estudiantes</i>		
Egresó de una IE no conectada (10 min)	0.03*** (0.001)	0.03** (0.012)
Egresó de una IE no conectada (30 min)	0.03*** (0.001)	0.04*** (0.014)
<b>N</b>	<b>264,007</b>	<b>1,806</b>
<i>Panel C: Controles de estudiantes y escuela</i>		
Egresó de una IE no conectada (10 min)	0.02*** (0.001)	0.02 (0.013)
Egresó de una IE no conectada (30 min)	0.02*** (0.001)	0.03** (0.015)
<b>N</b>	<b>258,594</b>	<b>1,764</b>

**Nota:** Los errores estándar robustos en paréntesis.

\*p < 0:1; \*\*p < 0:05; \*\*\*p < 0:01.

**Controles de estudiantes:** Orden de nacimiento del estudiante (sólo muestra hermanos), sexo, si repitió un grado, desempeño en comunicación integral y matemáticas.

**Controles de escuela:** es multigrado, tiene al menos una intervención por parte del ministerio, tiene piso de material noble, tiene acceso a agua de una red pública, tiene acceso a desagüe y total de alumnos.