



**“EL EFECTO VECINDARIO SOBRE EL ATRASO ESCOLAR EN
EL ÁMBITO RURAL DEL PERÚ”**

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

Presentado por

Srta. Natalia Pía Guerrero Trinidad

Asesor: Profesor Henry Espinoza, PhD.

2016

Resumen ejecutivo

Los modelos de interacción social han sido empleados con la finalidad de analizar la relación entre los pares, vecinos y entorno más cercano en temas como salud, seguridad ciudadana, elecciones, educación, entre otros. En el ámbito de educación, diferentes estudios han demostrado y reforzado la existencia de los distintos efectos de interacción social, es decir, cómo las características de las personas del entorno, sean compañeros o vecinos, tienen un efecto sobre el desarrollo de los estudiantes.

Adicionalmente, existe un componente que genera distorsiones, algunas veces no previstas, en los resultados, este es el componente espacial. Este componente está asociado a las características de cada una de las regiones, provincias, distritos y distintos clústeres de un área determinada, y la posibilidad de que estas características sean muy similares.

En el presente trabajo se plantea un modelo de interacción social y análisis espacial para estudiar el efecto vecindario o *neighborhood effect* sobre el atraso escolar entre los estudiantes de nivel primario y secundario en edad escolar de entre 6 a 18 años de edad que residen en el ámbito rural del Perú. Además, el presente estudio mide el efecto vecindario sobre el atraso escolar. Asimismo, la inclusión del análisis espacial es posible gracias al uso de bases de datos con información georreferenciada, con la cual se construyen los grupos o vecindarios a través de la distancia de hogar a hogar de los estudiantes.

Al contar con esta información es posible conocer con alta precisión las características de los hogares de los estudiantes, así como de los hogares vecinos a ellos.

Índice

Índice de tablas.....	iv
Índice de gráficos.....	v
Índice de anexos.....	vi
Capítulo I. Introducción.....	1
1. Objetivo.....	4
2. <i>Neighborhood effects</i> y componente espacial.....	5
Capítulo II. Revisión bibliográfica.....	6
Capítulo III. Modelo.....	10
Capítulo IV. Datos.....	13
Capítulo V. Descripción.....	15
Capítulo VI. Resultados.....	21
Conclusiones.....	25
Bibliografía.....	27
Anexos.....	33
Nota biográfica.....	38

Índice de tablas

Tabla 1.	Atraso escolar a nivel regional.....	2
Tabla 2.	Resultados de efecto endógeno y de contexto	21
Tabla 3.	Resultados de efectos endógeno y de contexto vía MCO.....	22
Tabla 4.	Resultados a nivel regional vía modelo SAR	23

Índice de gráficos

Gráfico 1.	Estudiantes de 6 a 18 años a nivel nacional (rural)	15
Gráfico 2.	Estudiantes de 6 a 18 años con atraso escolar	16
Gráfico 3.	Castellano y atraso escolar	16
Gráfico 4.	Quechua y atraso escolar	17
Gráfico 5.	Aimara y atraso escolar	17
Gráfico 6.	Otras lenguas originarias y atraso escolar	17
Gráfico 7.	Hogares con piso de tierra y atraso escolar	18
Gráfico 8.	Hogares sin acceso a alumbrado y atraso escolar	18
Gráfico 9.	Sexo y atraso escolar	19
Gráfico 10.	Jefe del hogar con primaria completa y atraso escolar	19
Gráfico 11.	Jefe del hogar con secundaria completa y atraso escolar	20
Gráfico 12.	Jefe del hogar con educación superior no universitaria y atraso escolar	20

Índice de anexos

Anexo 1.	Reporte de la generación de distancia entre vecinos e identificación de los vecinos más cercanos	34
Anexo 2.	Estadísticas descriptivas	35
Anexo 3.	Modelos estimados	37

Capítulo I. Introducción

El atraso escolar representa un reto importante tanto para el Perú, como para distintos países de América Latina. El atraso escolar¹ en el 2014 en los niveles de primaria y secundaria en el área rural del Perú alcanzó el 18% y 29,8%, respectivamente, según el Censo Escolar (CE) del MINEDU; mientras que en el área urbana este atraso ascendió a 5,4% y 10,7%, respectivamente.

Según cifras reportadas por el MINEDU, el atraso escolar pasó de 20,5% del total de alumnos matriculados en nivel primaria en el año 2000, a 8% en el 2014. Esta situación se repite a nivel secundaria, donde el atraso escolar pasó de 23,6% en el año 2000 a 12,9% en el 2014. Sin embargo, al considerar el nivel de atraso escolar según ámbito de residencia, se observa que si bien existe una reducción importante tanto en el ámbito urbano como en el rural, la brecha entre ambos estratos aún es amplia, dada la alta prevalencia de este problema en el ámbito rural. Durante el 2014, el número de estudiantes matriculados en el sistema educativo en los niveles de primaria y secundaria en el ámbito rural fue de 726.386 y 282.361 estudiantes, respectivamente

Ante estas cifras, es válido cuestionarse el porqué de estas diferencias, o por qué a pesar de los esfuerzos realizados por los Gobiernos central, regionales y locales esta brecha aún es significativa. Estudios como los de Alcázar y Valdivia (2005), Cavero, Montalva y Rodríguez (2011), y Cortez (2001), analizan el atraso escolar en el Perú, señalando que entre los principales determinantes de este problema se encuentran el nivel educativo de los padres de familia, el nivel socioeconómico de los hogares, la infraestructura escolar, la inadecuada oferta del servicio educativo, así como las habilidades cognitivas y no cognitivas de los estudiantes. Considerando esto, muchos de los programas, planes e iniciativas impulsadas en los últimos años tienen como eje central estos determinantes; sin embargo, los resultados parecen indicar que existen elementos no considerados en el análisis que podrían tener un rol clave en la mejora de la calidad educativa del Perú, sobre todo en el ámbito rural.

Según la Encuesta Nacional de Hogares Rurales (ENAHOR) del 2014, empleada como base principal para el presente estudio, el nivel de atraso escolar en el ámbito rural de algunas regiones supera el 40%². Entre estas se encuentran Loreto, con un 48% de atraso escolar en estudiantes de

¹ A nivel nacional, la Unidad de Estadística Educativa (UEE) del Ministerio de Educación (MINEDU) mide el atraso escolar como la proporción de la matrícula total de un cierto grado o nivel que tiene una edad mayor en dos o más años a la establecida para el grado en curso.

² Es importante señalar que las cifras expuestas superan a aquellas presentadas por el MINEDU debido a que la base de datos con la que se trabaja se centra exclusivamente en los hogares rurales del país.

primaria y secundaria, Ucayali con un 47% y Huánuco con 43%. Por otro lado, Moquegua, Ica, Arequipa y Lima son las regiones con menor nivel de atraso escolar, aun cuando este nivel supera en todos los casos el 15%.

Tabla 1. Atraso escolar a nivel regional

Región	Atraso escolar*	Región	Atraso escolar
Amazonas	35%	Lambayeque	27%
Áncash	34%	Lima	22%
Apurímac	27%	Loreto	48%
Arequipa	22%	Madre de Dios	24%
Ayacucho	34%	Moquegua	21%
Cajamarca	33%	Pasco	31%
Cusco	30%	Piura	35%
Huancavelica	32%	Puno	24%
Huánuco	43%	San Martín	33%
Ica	19%	Tacna	21%
Junín	31%	Tumbes	19%
La Libertad	35%	Ucayali	47%

* Incluye el atraso escolar a nivel primario y secundario en los estudiantes en edad escolar.
Fuente: ENAHOR 2014.

Asimismo, al considerar el atraso escolar por nivel educativo, son tanto Loreto, Ucayali y Huánuco las regiones con mayor nivel de atraso escolar tanto en primaria como secundaria. La prevalencia de este problema indicaría, como se señaló anteriormente, que existe(n) algún(os) elemento(s) no considerado(s) en el análisis de los resultados educativos.

Así, un elemento no comúnmente considerado es la interacción social entre los individuos en edad escolar. Los efectos ligados a la interacción social pueden ser llamados *peer influences*, *neighbourhood effects*, *imitation*, *herd behaviour*, o *interdependent preferences* dependiendo del contexto en el que sean analizados (Manski 1993). La interacción social analiza la forma en la que el comportamiento de un individuo se ve afectado por el comportamiento de un grupo.

En general, los modelos de interacción social analizan el cambio en la acción o decisión de una persona en relación con los cambios de las personas en su entorno, lo que tiene implicancias

importantes en distintas áreas. Según Schellig (1972), cuando los individuos pueden elegir su lugar de residencia, la presencia de la interacción entre individuos genera una clase de segregación del espacio, incluso en escenarios donde se busca la integración. Este último punto permite explicar de alguna manera los patrones de ciertos problemas en determinados lugares.

Estudios como los de Diez Roux (2001), Ioannides y Zabel (2008), y Oakes (2004) exploran el efecto de la interacción social en áreas como salud, urbanismo, y sociología. En educación, Helmers y Patnam (2014), Lin (2010), Bobonis y Finan (2009), y Jargowsky y El Komi (2009) analizan tanto el efecto vecindario como el de pares sobre el desempeño académico de los alumnos, y sobre la decisión de permanecer o no en la institución educativa.

En el caso de Perú, existe un reducido número de estudios que analizan este tipo de efecto sobre los resultados educativos, específicamente sobre el atraso escolar en los estudiantes. Estudios como el de Agüero León y Cueto (2004) confirman la existencia de un efecto de pares en el desempeño escolar, mientras que Sakellariou (2008) y Beltrán y Seinfeld (2011) refuerzan la importancia de conocer cómo la interacción entre los estudiantes afecta a sus propias habilidades cognitivas y su desempeño educativo.

Por otro lado, este efecto ha sido analizado en el Perú en áreas como salud e inclusión financiera. En el caso de salud, Bütünheim *et al.* (2014) analizan el efecto vecindario en la participación de la población urbana en campañas de control de salud de puerta a puerta en el Perú, a través de una función de infestación de los hogares. Los resultados de este estudio indican que existe un efecto vecindario significativo y que este efecto es mayor en los vecindarios nuevos. Sobre inclusión financiera, Karlan, Mobins, Rosenblat y Szeidl (2009) analizan un experimento realizado en dos barrios de Lima con la finalidad de medir la importancia de las redes sociales y el costo de pedir un préstamo.

Al igual que la interacción entre individuos, otro componente pocas veces analizado es el componente espacial. Tanto Crowder y South (2011), Dietz (2002), Durlauf (2004), Brooks-Gunn, Duncan y Aber (1997), como Kling, Liebman y Katz (2007) analizan el efecto vecindario considerando un análisis espacial dada la relevancia de la identificación de los individuos y su entorno más cercano. Este componente sirve como un complemento en países en desarrollo como Perú, donde existe un sistema de administración descentralizado, pero que por motivos como corrupción, falta de capacidad, accesibilidad, entre otros, no ha funcionado adecuadamente y, en algunos casos, contrario a su objetivo, ha ampliado la brecha de inequidad entre las regiones

(Baumol, Nelson & Wolff 1996).

Para Perú, tanto Elías y Rey (2011) como Escobal y Torero (2000) señalan que el análisis espacial juega un rol importante en el desempeño escolar de los estudiantes ya que afecta tanto la dinámica como la tendencia de las variables educativas. La diversidad geográfica y cultural en el Perú ha generado que el acceso a los distintos servicios públicos no sea el mismo en las regiones del país; esto tiene un impacto directo e indirecto sobre el desarrollo de las regiones y su población, haciendo que el desempeño escolar represente un reto mayor.

En ese sentido, el presente estudio busca confirmar la existencia de *neighborhood effects* sobre el atraso escolar de los alumnos que residen en el ámbito rural y medir su impacto sobre el atraso escolar. Asimismo, sobre la base de la metodología empleada por Del Bello, Paacchini y Zenou (2015), Crowder y South (2011), y Lee (2007), se plantea un modelo de auto-regresión espacial (SAR, por sus siglas en inglés). Este modelo incorpora el componente espacial mencionado anteriormente y resuelve los problemas que afectan la estimación de la interacción social entre individuos. Además, permite una mejor identificación del efecto endógeno y de contexto, efectos asociados a los *neighborhood effects* (Manski 1993).

Asimismo, el presente estudio emplea la Encuesta Nacional de Hogares Rurales (ENAHOR) del 2014. Esta base fue elaborada por el Instituto Nacional de Estadística e Informática (INEI) y el Ministerio de Economía y Finanzas (MEF), y cuenta con información detallada que permite replicar la proporción de alumnos con atraso escolar siguiendo la definición planteada por el MINEDU, e identificar cada uno de los hogares sobre la base de la georreferenciación de estos, lo que permite implementar adecuadamente el modelo SAR.

De manera similar a Helmers y Patnam (2014), se emplea la proximidad de residencia entre los niños y niñas para identificar los *spatial neighborhood effects* en el atraso escolar, y se espera que los niños y niñas que viven más cerca uno de otro presenten una probabilidad más alta de interactuar respecto a aquellos que viven en zonas más alejadas, lo que podría generar un efecto vecindario.

1. Objetivo

El objetivo principal del estudio es presentar mayor evidencia acerca del efecto vecindario en el ámbito rural del Perú, y cómo este afecta el atraso escolar de los estudiantes. Un elemento importante es la inclusión del componente espacial al análisis, el cual resuelve el problema de interacción social al mismo tiempo que permite una mejor estimación de los efectos de contexto

y endógeno.

Además, al contar con información a nivel individual y georreferenciada es posible conocer con alta precisión las características de los hogares de los estudiantes, así como de los hogares vecinos a ellos.

2. *Neighborhood effects* y componente espacial

El efecto vecindario o *neighborhood effect* forma parte de los modelos de interacción social. Estas interacciones representan una forma de externalidad donde las acciones y/o decisiones de un grupo determinado (hogar, vecinos o pares) afectan las preferencias o el comportamiento de un individuo. Según Manski (1993), el efecto vecindario puede ser clasificado como efecto endógeno, efecto de contexto o exógeno y efecto correlacionado. El efecto endógeno hace referencia a la probabilidad de que el comportamiento de un individuo, en este caso, un alumno, se vea de alguna forma afectada por el comportamiento del grupo o vecinos. A diferencia del efecto endógeno, el efecto exógeno o de contexto hace referencia a la probabilidad de que el comportamiento de un alumno se vea afectado por características de sus compañeros de clase o vecinos como edad, sexo o características socioeconómicas de los hogares (Agüero León & Cueto 2004). En el caso del efecto correlacionado, el comportamiento de un alumno y de sus compañeros es similar debido a que cuentan con características individuales similares o enfrentan “*institutional environments*” similares. Sobre este último efecto, las características asociadas al lugar de residencia como acceso a servicios básicos, accesibilidad (carreteras, pistas, puentes, entre otros), ruralidad, gestión, entre otras, son compartidas por los niños y niñas que son vecinos.

La importancia de la identificación de estos tres tipos de efectos se encuentra en que cada uno tiene implicancias de política distintas.

El presente documento se divide en siete secciones. La primera sección presenta la introducción, objetivos, el concepto del efecto vecindario y el análisis espacial; la segunda sección desarrolla la revisión de literatura que abarca tanto la literatura de los efectos de interacción social, así como los determinantes más analizados sobre el atraso escolar; la tercera sección presenta el modelo SAR empleado para la incorporación del análisis espacial y plantea la especificación del modelo de interacción social; la cuarta sección presenta la descripción de las bases de datos empleadas en el estudio; la quinta sección describe las características de la población bajo análisis y presenta estos datos en mapas para observar la concentración de los estudiantes; la sexta sección presenta los resultados obtenidos bajo cada modelo estimado; y la séptima sección presenta las conclusiones respecto a los hallazgos.

Capítulo II. Revisión bibliográfica

Existen diferentes posturas acerca de cuáles son los principales determinantes del atraso escolar, donde la mayoría de estudios consideran factores ligados al nivel de pobreza del hogar, la calidad nutricional y de salud de los estudiantes, el acceso a centros educativos, participación en el mercado laboral, entre otros. Sin embargo, existe un elemento adicional que afecta la toma de decisiones y el comportamiento de los individuos en edad escolar. Este es la interacción social de los estudiantes con sus vecinos y/o compañeros. A nivel internacional, los estudios sobre *neighborhood effects* presentan resultados diversos. A nivel nacional no se encontraron estudios que consideren como una posible determinante del atraso escolar o desempeño académico a los *neighborhood effects*, aun cuando se encontró literatura relacionada a otros tipos de interacción social como los efectos de pares. Entre estos estudios se encuentran: Agüero León y Cueto (2004) y Montero, Oliart, Ames y Cabrera (2001), aun cuando el enfoque de estos últimos es etnográfico.

A nivel internacional, trabajos como los de Del Bello, Paacchini y Zenou (2015) y Vartanian y Gleason (1999) señalan que en términos de educación, la interacción entre compañeros de aula o de institución educativa es más importante respecto a la relación con amigos del mismo “vecindario”.

Asimismo, tanto en los estudios mencionados como en Crowder y South (2011), donde también se analiza los *neighborhood effects* sobre el atraso en la graduación de la escuela secundaria, se emplea un enfoque metodológico distinto. En estos estudios se incorpora la dimensión espacial de estos efectos a través de técnicas de análisis espacial, específicamente empleando un modelo SAR. El objetivo principal de incorporar esta dimensión es incluir el contexto geográfico y socioeconómico en la estimación de los efectos, ya que la interacción entre individuos se encuentra influenciada por condiciones socioeconómicas, lo que los acerca a las personas de su entorno geográfico más cercano (Sampson *et al.* 2002).

Los autores se refieren a investigaciones como las de Jencks y Mayer (1990), Crane (1991) y Aisworth (2002) donde se expone la idea de que las condiciones y características socioeconómicas de los “vecindarios” afectan e influyen el comportamiento y decisiones de los niños y niñas. Específicamente sobre educación, Wilson (1987) señaló que las percepciones sobre la importancia de asistir a la escuela o en general de la educación, y de contar con un empleo estable para mejorar el bienestar del hogar son relevantes en un contexto de “vecindario” porque se transmiten de individuo en individuo por medio de la interacción social.

Por otro lado, Gibbons (2002) analiza el impacto de los *neighborhood effects* en el desempeño académico de los adolescentes en la Gran Bretaña y encuentra que sí existe una influencia reducida de estos efectos, mientras que las características socioeconómicas de los hogares suelen tener un mayor impacto sobre el desempeño. Sin embargo, vivir en un vecindario con acceso a todos los servicios básicos y adecuada infraestructura genera beneficios sociales significativos para reducir el atraso escolar en el vecindario, en comparación a la mejora de las características socioeconómicas en el hogar.

Aun cuando a nivel nacional no se encontró estudios ligados a la interacción entre vecinos, sí existe una amplia literatura sobre los determinantes socioeconómicos del atraso escolar. Entre los determinantes más estudiados se encuentran la edad, sexo, tipo de trabajo, realización de quehaceres del hogar, área de residencia, género del jefe de hogar, nivel educativo más alto alcanzado por el jefe de hogar, activos, presencia de niños menores en el hogar, existencia de algún *shock* adverso, entre otros. Estas variables son consideradas en estudios como el de Cavero, Montalva y Rodríguez (2011), en el cual, a través de un modelo Probit se analiza la probabilidad de desertar o repetir, y a través de un modelo Logit se elabora una distinción sobre cómo afectan estas variables a la probabilidad de repetir y desertar por separado. De manera similar, Cortez (2001) estudia el atraso escolar, considerando entre los principales determinantes a las condiciones de vida de los hogares, es decir, las características socioeconómicas y demográficas de los miembros en edad escolar.

Adicionalmente, este estudio considera el gasto per cápita en programas sociales del distrito donde se ubica el hogar, la predicción de las horas dedicadas a trabajo infantil y el salario de los adultos que conforman el hogar.

De manera particular, en el primer estudio los autores concluyen que si bien las mujeres tienen menores probabilidades de repetir de grado en primaria, tienen una mayor probabilidad de desertar en la transición de primaria a secundaria, lo cual puede atribuirse a la realización de quehaceres del hogar, incrementándose en áreas rurales. Asimismo, el atraso escolar afecta más a las mujeres que a los varones, y el embarazo adolescente es una variable importante que afecta el progreso escolar en las mujeres que residen en el área rural.

Otros estudios consideran entre los principales determinantes al entorno geográfico y a las características de las instituciones educativas, como la gestión, infraestructura, número de secciones, número de docentes, entre otras. Chinen (2002) estudia las características de distintos

grupos de nivel primario sobre la base de indicadores que reflejan la calidad educativa que reciben, y analiza el progreso de estos indicadores para el periodo 1993-1998.

Asimismo, Alcázar y Valdivia (2005) concluyen que en áreas rurales los factores asociados con la pobreza son en general los principales determinantes de la deserción escolar. Asimismo, existen otros motivos como problemas familiares y de salud que son el segundo factor más importante para las mujeres, así como la calidad e importancia de la educación es el tercer factor más importante. En tanto, Chinen (2002) muestra que la inversión pública parece concentrarse en las instituciones polidocentes completas con un alto nivel de matrícula, si bien este resultado puede parecer sensato, bajo una perspectiva de equidad este resultado deja en desventaja a las instituciones más pequeñas, incrementando la brecha de oportunidades en el sistema educativo.

Otros estudios analizan distintos determinantes como la incorporación de los estudiantes en el mercado laboral, las diferencias en los aprendizajes, así como el acceso a una institución educativa, en términos de tiempo. Entre estos estudios se encuentran Pariguana (2011), Glewwe Krutikova y Rolleston (2014), y Lavado y Gallegos (2006). En el primer estudio, se evalúa la relación entre las decisiones laborales y la asistencia a una institución educativa para los jóvenes de entre 12-17 años. El autor plantea que las altas tasas de actividad laboral por parte de estos jóvenes afecta la capacidad de acumular capital humano. En Glewwe, Krutikova y Rolleston (2014) se estudian las diferencias de aprendizaje entre los estudiantes y las características del entorno de estos estudiantes y el efecto que tienen sobre su desempeño educativo. Los autores elaboran una comparación entre Vietnam y Perú sobre la base de los resultados de la prueba PISA 2012, donde Perú quedó último tanto en comprensión de lectura como matemática de 65 países, a pesar de que el ingreso per cápita en Perú era 4 veces mayor que el de Vietnam para ese año. Finalmente, en el tercer estudio, Lavado y Gallegos (2006) analizan la dinámica de la deserción escolar considerando modelos de duración.

Pariguana (2011) encuentra que ante un mayor nivel de gasto en el hogar, la probabilidad de que el adolescente se dedique solo a estudiar se incrementa en 6 puntos porcentuales, es decir, la capacidad de gasto, ligada a ingresos, genera un impacto positivo sobre la acumulación de capital humano. Respecto a la variable edad señala que por cada año adicional del adolescente, la probabilidad de que solo estudie disminuye en 4 puntos porcentuales demostrando que el costo de oportunidad incrementa a medida que pasan los años; asimismo, la probabilidad de que el adolescente solo estudie incrementa ante un mayor nivel educativo alcanzado por el jefe del hogar, así como que el jefe del hogar sea mujer, la lengua materna sea castellano y el residir en área

urbana también incrementa esta probabilidad. Es importante recalcar que la variable de residencia es aquella que genera un mayor impacto en términos de magnitud.

En Lavado y Gallegos (2006), bajo un enfoque paramétrico, los resultados señalan que la distancia a un centro educativo o la falta de oferta educativa representa un problema en la decisión de enviar a las mujeres al centro educativo tanto en área rural como urbana, y son los hombres del área rural de la selva peruana quienes presentan la tasa de deserción escolar más alta por problemas de oferta. Asimismo, problemas de repetición escolar incrementan la probabilidad de deserción escolar tanto en hombres como mujeres, y el nivel educativo del padre es una variable importante para reducir la tasa de deserción en zonas urbanas, mientras que la estructura familiar representa un punto importante para hombres y mujeres en zonas rurales y urbanas. Finalmente, los autores señalan que los estudiantes que tienen hermanos en educación secundaria muestran una menor probabilidad de abandonar la escuela, y alumnos con hermanos en educación primaria muestran una mayor probabilidad de abandonar la escuela, por último encuentran que en zonas urbanas el tener hermanos menores incrementa la probabilidad de deserción en mujeres, y en zonas rurales de hombres.

La revisión expuesta presenta aún reducida evidencia en relación con los efectos vecindarios, aunque se expone la evidencia sobre los efectos de la interacción social a nivel de pares. Asimismo, se presentan características de los estudiantes y su entorno que tienen impacto sobre el atraso escolar, esto brinda información relevante para el planteamiento del modelo de interacción social a nivel de vecindario que será estimado en el presente estudio.

Capítulo III. Modelo

Esta sección comenzará definiendo algunos conceptos relevantes para el análisis espacial. Según LeSage y Pace (2009), la econometría espacial a diferencia de las herramientas clásicas considera la locación espacial de la unidad de análisis, es decir, incorpora el efecto espacial de la muestra analizada. Ante la presencia de un componente espacial, surgen dos problemas, la dependencia espacial entre las observaciones y la heterogeneidad espacial de las relaciones a modelar. La econometría tradicional ignora estos problemas, puesto que considera el cumplimiento de los supuestos de Gauss-Markov.

La dependencia o auto-correlación espacial genera que las observaciones en una determinada localidad i , dependen de las observaciones en localidades distintas a i . Existen dos motivos por los cuales podría existir esta dependencia. Según LeSage (1999), el primero está ligado a la existencia de unidades espaciales como códigos de georreferenciación de distritos, entre otros, lo que genera errores de medición. El segundo está asociado a la modelación de la dimensión socio-demográfica y económica de las observaciones.

De manera formal se representa:

$$y_i = f(y_j), i = 1, \dots, n \quad j \neq i$$

La heterogeneidad espacial está ligada a la variación de las relaciones a modelar en los distintos puntos en el espacio. De manera formal se representa:

$$y_i = f_i(X_i\beta_i + \epsilon_i)$$

Donde i representa a las observaciones en los distintos puntos en el espacio, y_i es el vector de variables dependientes, X_i es un vector de variables explicativas, asociadas al vector de parámetros β_i , y ϵ_i representa el error del modelo. El problema de heterogeneidad también se encuentra asociado a la omisión de variables relevantes, así como a otros problemas de especificación, que podrían generar errores de medición.

Como se señaló, sobre la base de la metodología planteada por Del Bello, Paacchini y Zenou (2015), Crowder y South (2011), y Lee (2007), para el análisis del efecto vecindario sobre el atraso escolar, se decidió emplear un modelo SAR. El objetivo principal de emplear un modelo SAR es incluir el contexto geográfico y socioeconómico en la estimación de los efectos, ya que la interacción entre individuos se encuentra influenciada por condiciones socioeconómicas, lo que

los acerca a las personas de su entorno geográfico más cercano, por ende se genera un problema de auto-correlación. Para el presente estudio se aplicó el estadístico de I de Morán y se encontró que no rechaza la hipótesis nula de no existencia de auto-correlación, ya que toma el valor de 0,86, lo que indica auto-correlación positiva. Esto confirma la existencia de auto-correlación espacial a nivel de vecindarios.

La auto-correlación, o “Tobler’s First Law of Geography” (Tobler, Gale & Olsson 1979), formula el concepto de auto-correlación espacial en la cual, los autores señalan que si bien todas las cosas están relacionadas entre sí, las cosas más próximas en el espacio tienen una relación mayor que las distantes. Sobre la base de esto, los resultados educativos y las características de los individuos suelen estar correlacionadas a medida que la distancia entre los hogares es menor. En *Educational performance and spatial Convergence in Peru* (Elías & Rey 2011), los autores evalúan la auto-correlación espacial de distintos indicadores sociales como complemento de los indicadores económicos del Perú, con la finalidad de comprender el desempeño socioeconómico de las distintas regiones del país. Específicamente analizan la convergencia educativa en el Perú para el periodo 1993-2005 a nivel provincial, donde los resultados indican la alta auto-correlación espacial positiva a ese nivel.

El modelo SAR emplea la metodología de máxima verosimilitud donde la(s) variable(s) dependiente(s) se encuentra(n) ponderada(s) por la matriz de pesos, la cual contiene las distintas relaciones de contigüidad o distancias entre las observaciones. La especificación del modelo SAR es la siguiente:

Ecuación 1: Especificación de modelo SAR

$$y_i = \rho W y_i + X_i \beta_i + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2 I_n)$$

Donde y_i es un vector de $n \times 1$ que contiene los resultados educativos de las variables priorizadas, X_i es un vector de $n \times k$ que contiene distintas variables explicativas como características de los individuos, características de las instituciones educativas y características de la vivienda y el hogar, ϵ es un vector de $n \times 1$ que contiene al término del error, W es una matriz inversa de $n \times n$ que contiene los pesos espaciales de la distancia de cada hogar a sus vecinos, y tanto ρ como β_i son los parámetros a estimar. El parámetro ρ es el coeficiente del rezago espacial de la variable dependiente, y β_i representa que tanto explican cada una de las variables explicativas a la variable dependiente. El término $W y_i$ es la variable dependiente de rezago espacial, el cual

genera simultaneidad en la ecuación a estimar (endogeneidad), por ende se debe emplear técnicas de máxima verosimilitud. Donde la solución de la ecuación 1 es:

$$y_i = (I - \rho W)^{-1} X_i \beta_i + (I - \rho W)^{-1} \epsilon_i$$

Donde los elementos $(I - \rho W)^{-1}$ son llamados multiplicadores espaciales (Anselin 2003). Si se realizará una estimación vía Mínimos Cuadrados Ordinarios (MCO) y existiese auto-correlación en la variable resultado, el estimador MCO sería sesgado e inconsistente (Anselin 1988).

Especificación del modelo e identificación del efecto vecindario

Sobre la base de la especificación de los modelos planteados en Del Bello, Paacchini y Zenou (2015), Crowder y South (2011), y Lee (2007), se presenta el siguiente modelo para la estimación del efecto vecindario sobre el atraso escolar en los estudiantes de edad escolar que residen en el ámbito rural del Perú:

$$y_{il} = \sigma^{EE} \sum_{j=1}^n w_{ij} y_{jl} + x'_{il} \beta + \sum_{j=1}^n w_{ij} x'_{jl} \alpha^{EE} + \lambda_l + \epsilon_{il}$$

Donde y_{il} es una variable dicotómica que toma el valor de 1 si el estudiante presenta atraso escolar en cada vecindario, x'_{jl} representa las características de los individuos, λ_l representa variables de efectos fijos del vecindario, y ϵ_{il} es el término de error. Entre las características de los individuos se encuentran sexo, lengua originaria (quechua), distancia³ del hogar a la institución educativa más cercana, distancia del hogar al centro de salud más cercano, y el nivel educativo del jefe del hogar.

Bajo este modelo, los efectos endógenos se encuentran representados por σ^{EE} , mientras que el efecto de contexto está representado por α^{EE} .

³ El análisis de distancia se basa en el modelo de accesibilidad (MAC) desarrollado por el Ministerio de Economía y Finanzas (MEF); con este modelo se estimó la distancia a los centros de salud y a las instituciones educativas (DGPP 2015). El MAC es construido con información de los valores de latitud, red vial (niveles de conexión y tipo de material de las vías), red vial urbana, rutas de navegación lacustre, ríos navegables, hidrografía, superficie de fricción, límites administrativos, precio de combustibles a nivel provincial, salarios promedio a nivel provincial, y costos de operación de transporte de carga (incluye precios de mantenimiento, salarios, permisos e impuestos).

Capítulo IV. Datos

Como se mencionó anteriormente, la principal base de datos en el estudio es la Encuesta Nacional de Hogares Rurales (ENAHOR 2014), la cual fue elaborada por el Instituto Nacional de Estadística e Informática (INEI) en coordinación con el Ministerio de Economía y Finanzas (MEF). El objetivo de la ENAHOR es contar con información estadística, demográfica, social y económica representativa a nivel provincial de los hogares rurales del Perú. La muestra analizada consideró información estadística y documentos cartográficos del Empadronamiento Distrital de Población y Vivienda 2012 y 2013; asimismo, se consideraron las variables previamente estudiadas en la Encuesta Nacional de Hogares (ENAHO 2012) y el Censo Agropecuario (CENAGRO 2012).

Asimismo, se emplean dos bases ligadas al desempeño escolar y a las características de las instituciones educativas. Primero, la Evaluación Censal de Estudiantes (ECE 2014), la cual evalúa a los alumnos de segundo grado de primaria en las instituciones educativas de Educación Básica Regular, mientras que en las instituciones del Programa de Educación Intercultural Bilingüe (EIB), la prueba es aplicada a los alumnos de cuarto grado de primaria. Se evalúan conceptos matemáticos y de comprensión de lectura.

Es importante considerar que la prueba no es aplicada en las instituciones con menos de cinco estudiantes matriculados en segundo grado, se estima que alrededor del 4% de los estudiantes a nivel nacional son excluidos, y el 27% de las instituciones son excluidas por esta consideración. Asimismo, la cobertura de alumnos evaluados en el 2014 alcanzó el 90,5%, y en el caso de área rural fue de 88,6%.

La segunda base empleada es el Censo Escolar (CE). El CE se lleva a cabo anualmente y recoge información de cada una de las instituciones educativas públicas y privadas, y programas no escolarizados de todo el Perú. La información obtenida es declarativa y está a cargo del director de la institución educativa. El CE considera información de: matrícula, docentes y recursos; local escolar; y resultado del ejercicio. Desde 2015, se agregó la cédula ID, la cual recoge información para la identificación de cada institución y de los establecimientos en los que opera. Es importante considerar que en el CE, el 6% de las instituciones educativas tienen datos imputados, dado que algunas instituciones no reportan información, en las instituciones públicas, el nivel de imputación llega al 57%, en instituciones rurales 52%, y en instituciones con menos de diez estudiantes 54%. Para el presente análisis se usará el CE 2014.

Las bases descritas cuentan con la información necesaria para elaborar el presente análisis, puesto que describen las características de la infraestructura de las instituciones educativas, el puntaje obtenido en las pruebas de rendimiento académico y las características socioeconómicas de los estudiantes en el ámbito rural del país. Este último punto es sumamente relevante ya que la ENAHOR es aplicada solo a hogares rurales, por lo cual se consideró era la base más apropiada para el estudio. El anexo 2 presenta una tabla con los principales estadísticos.

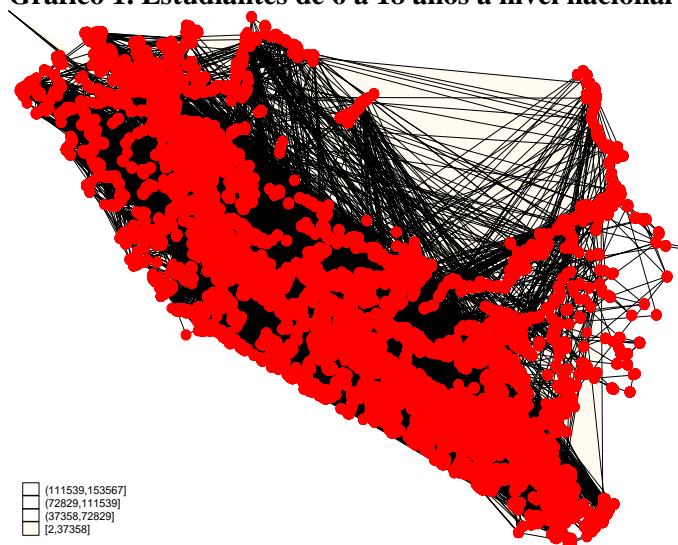
Capítulo V. Descripción

En esta sección se presenta la distribución de los estudiantes del ámbito rural a lo largo del territorio nacional. Asimismo, se presenta la concentración de estudiantes con atraso escolar según grupo o vecindario respecto a características como edad, lengua de origen, sexo, si el hogar cuenta con piso de tierra, si el hogar no cuenta con alumbrado, y el último nivel educativo alcanzado por el jefe de hogar. El principal objetivo es reportar la concentración de alumnos con atraso escolar según el vecindario para conocer si el vecindario concentra alumnos con las mismas características.

El gráfico 1 muestra la concentración de la población bajo estudio. Esta población está conformada por individuos de entre 6 y 18 años, es decir que se encuentran en edad escolar en los niveles de primaria y secundaria, y que asisten a una institución educativa estatal. Como se mencionó anteriormente, la ENAHOR solo está dirigida al ámbito rural. Los vecindarios fueron construidos sobre la base de la georreferenciación de los hogares y de la distancia entre ellos. Empleando el comando de *Stata geonear*, se identificaron a los vecinos más cercanos de cada individuo, lo que permitió la construcción de los grupos. En el anexo 1 se puede observar el reporte de este procedimiento.

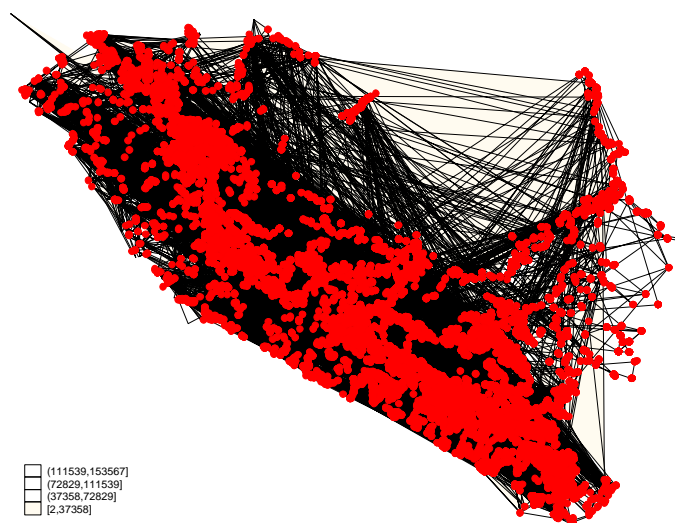
Respecto a la concentración de estudiantes con atraso escolar, es posible observar que estos se concentran con mayor intensidad en el sur del país, mientras que el norte, específicamente la zona con límite marítimo, concentra un menor número de estudiantes con atraso escolar.

Gráfico 1. Estudiantes de 6 a 18 años a nivel nacional (rural)



Fuente: ENAHOR (2014). Elaboración propia 2016.

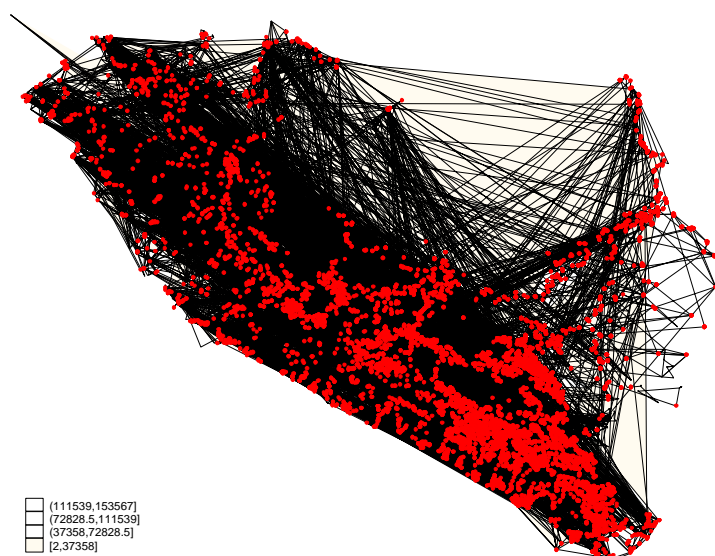
Gráfico 2. Estudiantes de 6 a 18 años con atraso escolar



Fuente: ENAHOR (2014). Elaboración propia 2016.

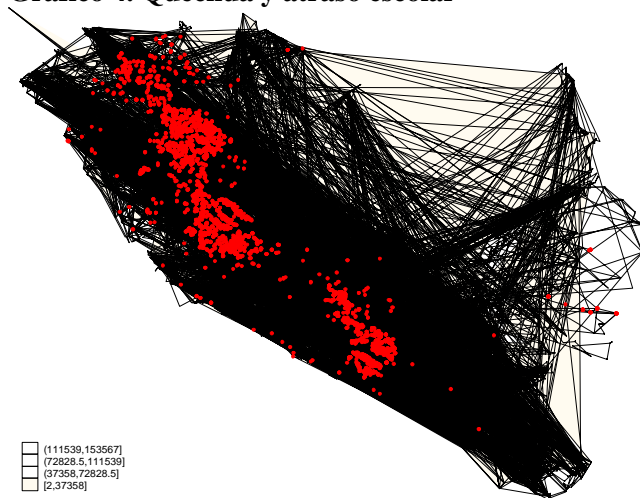
La lengua de origen, en algunos casos, suele representar un problema para los estudiantes dado que el número de escuelas bilingües aún es reducido. Esto genera que en las instituciones educativas los docentes asignados y/o contratados se comuniquen solo en castellano, causando un problema grave de entendimiento entre docente y estudiante. Asimismo, algunos estudiantes son bilingües, pero no han aprendido adecuadamente ninguna de las lenguas, ya que en la mayoría de los casos las lenguas distintas al castellano son transmitidas de manera oral y se usan frecuentemente en el entorno más cercano. Según cifras del Censo Nacional (2007), alrededor de 3,2 millones de peruanos tienen al quechua como lengua materna.

Gráfico 3. Castellano y atraso escolar



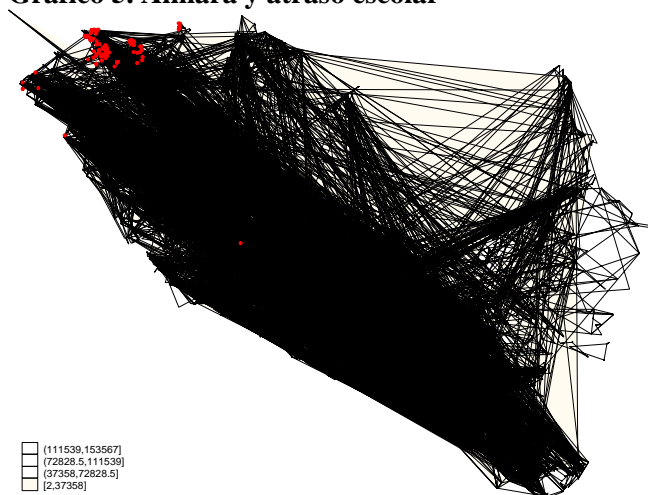
Fuente: ENAHOR (2014). Elaboración propia 2016.

Gráfico 4. Quechua y atraso escolar



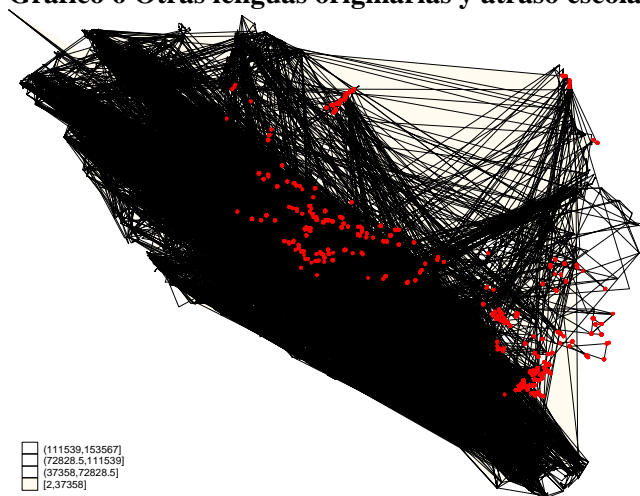
Fuente: ENAHOR (2014). Elaboración propia 2016.

Gráfico 5. Aimara y atraso escolar



Fuente: ENAHOR (2014). Elaboración propia 2016.

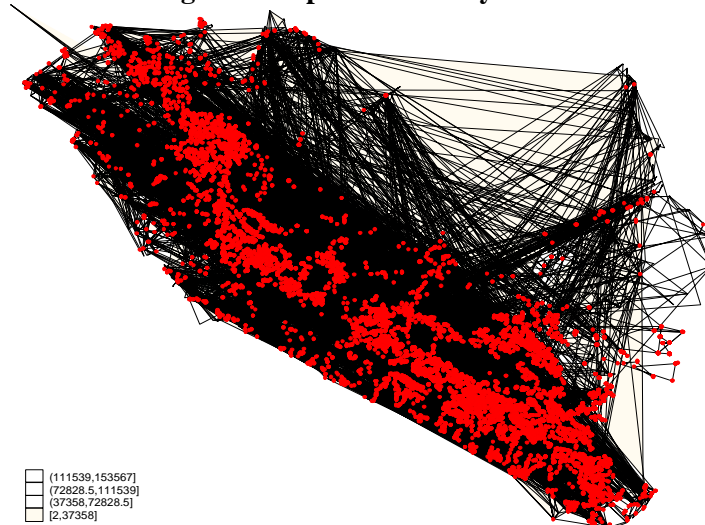
Gráfico 6 Otras lenguas originarias y atraso escolar



Fuente: ENAHOR (2014). Elaboración propia 2016.

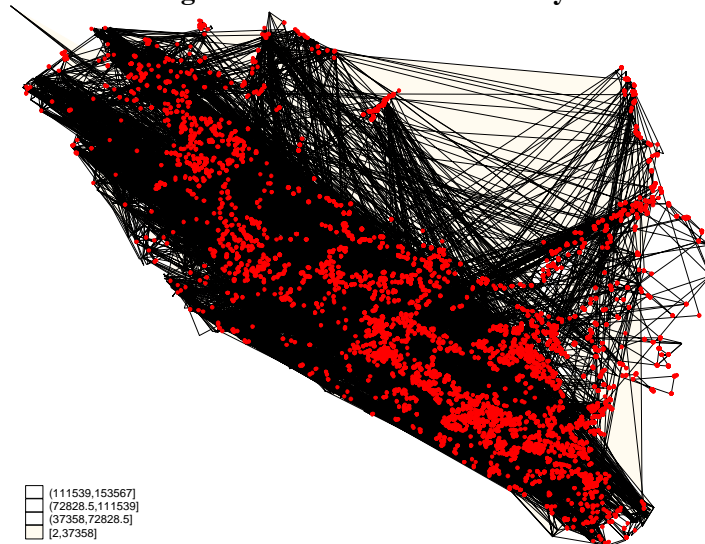
En lo que respecta a las características de los hogares de los estudiantes con atraso escolar, es posible observar que existe una alta concentración de hogares con piso de tierra en el sur del país, mientras que el no contar con acceso a electricidad es menos frecuente entre los estudiantes con atraso escolar. Es importante observar que en la mayoría de los casos, los vecindarios se pintan por completo de rojo, lo que indica que tanto el piso de tierra como el no acceso a electricidad es una característica común de estos estudiantes. Esto puede estar ligado a problemas de titulación, pobreza, salud, entre otros.

Gráfico 7. Hogares con piso de tierra y atraso escolar



Fuente: ENAHOR (2014). Elaboración propia 2016.

Gráfico 8. Hogares sin acceso a alumbrado y atraso escolar

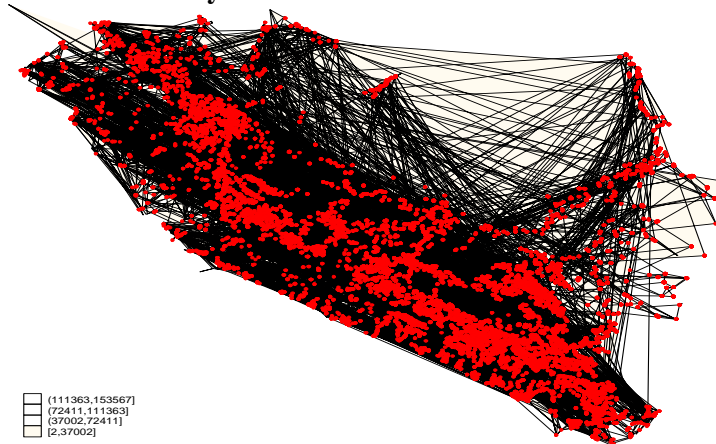


Fuente: ENAHOR (2014). Elaboración propia 2016.

Estudios como los de Cavero, Montalva y Rodríguez (2011) y Cortez (2001), indican que el atraso escolar afecta más a mujeres que a hombres, aunque si bien las mujeres tienen menores

probabilidades de repetir de grado en primaria, tienen una mayor probabilidad de desertar en la transición de primaria a secundaria, lo cual puede atribuirse a la realización de quehaceres del hogar, incrementándose en áreas rurales. Asimismo, el atraso escolar afecta más a las mujeres que a los varones, y el embarazo adolescente es una variable importante que afecta el progreso escolar en las mujeres que residen en área rural. El siguiente gráfico muestra la concentración de estudiantes con atraso escolar tanto en primaria como secundaria según sexo. Se puede observar que existen un mayor número de mujeres.

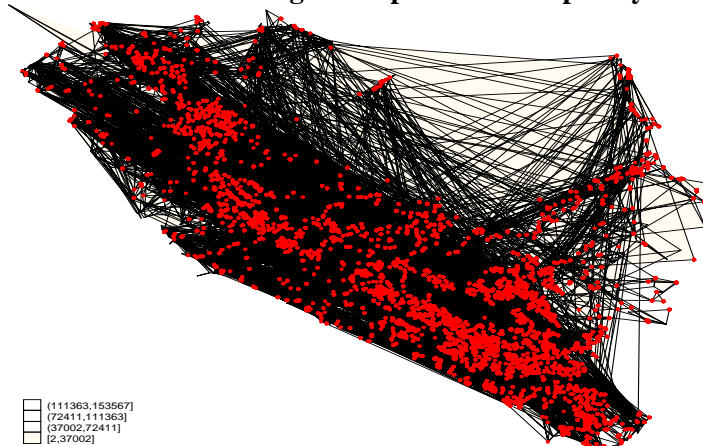
Gráfico 9. Sexo y atraso escolar



Fuente: ENAHOR (2014). Elaboración propia 2016.

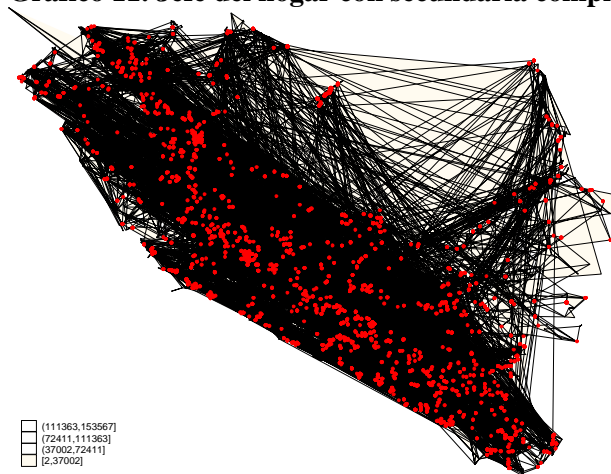
Los gráficos 10, 11 y 12 muestran como a medida que el nivel educativo del jefe del hogar es más alto, la concentración y el número de estudiantes con atraso escolar se reduce significativamente; como indica Pariguana (2011), la probabilidad de que un adolescente solo estudie incrementa ante un mayor nivel educativo alcanzado por el jefe del hogar.

Gráfico 10. Jefe del hogar con primaria completa y atraso escolar



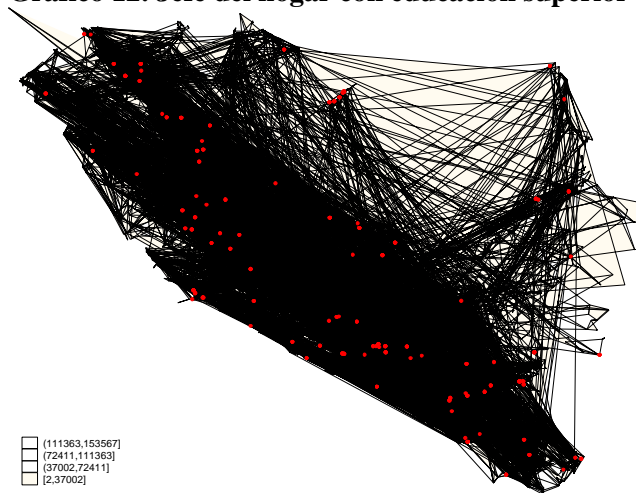
Fuente: ENAHOR (2014). Elaboración propia 2016.

Gráfico 11. Jefe del hogar con secundaria completa y atraso escolar



Fuente: ENAHOR (2014). Elaboración propia 2016.

Gráfico 12. Jefe del hogar con educación superior no universitaria y atraso escolar



Fuente: ENAHOR (2014). Elaboración propia 2016.

En el anexo 2 se presentan los principales estadísticos descriptivos de la base de datos.

Capítulo VI. Resultados

Sobre la base del modelo planteado, se estimaron 3 versiones. La diferencia entre estas versiones se encuentra en las características de contexto o exógenas. Así, entre las características socioeconómicas de los estudiantes considerados en los modelos se encuentran el sexo, el nivel educativo del jefe de familia, lengua de origen, distancia del hogar de cada estudiante a la institución educativa más cercana, distancia del hogar de cada estudiante al centro de salud más cercano. Asimismo, los efectos fijos considerados a cada grupo o vecindario se encuentran ligados al acceso a servicios básicos: alumbrado, agua potable y desagüe, siendo este el modelo base (modelo 1), así como algunas características de las instituciones educativas de cada vecindario (aulas en buenas condiciones (modelo 2) y bibliotecas (modelo 3)).

Como se mencionó anteriormente, el efecto endógeno y de contexto se encuentran representados por σ^{EE} y α^{EE} , respectivamente. Ambos efectos representan los *neighborhood effects* y son el centro de presente estudio. A continuación, se presentan los resultados obtenidos en ambos efectos en las 3 versiones estimadas.

Tabla 2. Resultados de efecto endógeno y de contexto

Efecto	Variable	(1)	(2)	(3)
Endógeno		0,0016 ***	0,0002***	0,0014***
	Sexo	0,000***	0,000***	0,001**
	Quechua	0,005**	0,003**	0,003**
Contexto	Distancia a la inst. educativa	0,044***	0,016***	0,025***
	Distancia al centro de salud	0,048***	0,019***	0,026***
	Nivel educativo del jefe del hogar	-	-	-0,037***
	Alumbrado	-0,016***	-0,0196***	-0,0008***
	Agua potable	-0,021***	-0,016**	-0,0025***
Efectos fijos	Desague	-0,001**	-0,000***	-0,001**
	Bibliotecas	-	-	0,000**
	Aulas en buenas condiciones	-	-0,039***	-
	Rho	0,002	0,009	0,0005
	F-Test	10,820***	10,439***	11,273***
	Log Likelihood Fuction	4,701,189	5,887,694	4,815,485

El número de observaciones en cada una de los modelos es 4.570.994.

*p < 0,10, ** p< 0,05, ***p< 0,01

Fuente: Elaboración propia 2016.

El efecto endógeno en los tres modelos resultó significativo. En otras palabras, aun controlando

por el sexo de los estudiantes, la lengua de origen, la distancia a la institución educativa y el centro de salud más cercana, y por el acceso a servicios básicos, se confirma la existencia de un efecto vecindario en el atraso escolar. El resultado muestra la existencia de un impacto de los vecinos sobre el atraso escolar, es decir, que los vecinos más cercanos de un estudiante tengan una edad mayor en dos o más años a la establecida para el grado en el que se encuentran, tiene un efecto sobre la probabilidad de que el estudiante presente atraso escolar.

Este resultado se mantiene en los tres modelos analizados, los cuales se controla adicionalmente por el porcentaje de aulas en buenas condiciones, si la institución educativa cuenta con al menos una biblioteca y el nivel educativo del jefe del hogar.

Asimismo, se plantearon los tres modelos empleando la técnica de Mínimos Cuadrados Ordinarios (MCO). Bajo MCO, el efecto endógeno sigue siendo significativo pero los coeficientes superan ampliamente los obtenidos vía el modelo SAR. Esto se debe a que al no considerar el componente espacial, se sobreestima el efecto endógeno ante la presencia de la auto-correlación, lo que ocasiona que no se cumpla con uno de los supuestos de Gauss-Markov.

Tabla 3. Resultados de efectos endógeno y de contexto vía MCO

Efecto	Variable	Coefficientes
Endógeno		0,895***
	Sexo	-0,027
Contexto	Quechua	-0,083***
	Distancia a la inst. educativa	0,000
	Distancia al centro de salud	0,000***
	Alumbrado	-0,030
Efectos fijos	Agua potable	-0,099***
	Desague	-0,048**
	Constante	0,163***
	R cuadrado	0,080

El número de observaciones en cada una de los modelos es 4.570.994.

*p < 0,10, ** p< 0,05, ***p< 0,01

Fuente: Elaboración propia 2016.

De otro lado, bajo el modelo SAR, los efectos de contexto demuestran que en los estudiantes con una mayor distancia de sus hogares al centro educativo y al centro de salud se incrementa la probabilidad de presentar atraso escolar tanto en el nivel primario como secundario en cada uno de los grupos o vecindarios. Asimismo, el nivel educativo del jefe del hogar parece generar una reducción importante en la probabilidad del atraso escolar, específicamente, a medida que el nivel

educativo mejora, el efecto también lo hace alcanzando los 0,037 puntos porcentuales cuando el último nivel alcanzado es educación superior completa. Es importante señalar que la inclusión de los efectos fijos por vecindario reduce los coeficientes del efecto endógeno, asimismo, la alta significancia de las variables de contexto podría estar ligada a que estas características se encuentran correlacionadas.

Al considerar el análisis a nivel regional a través del modelo SAR, se puede apreciar que el efecto endógeno, es decir el efecto del atraso escolar de los vecinos, es positivo, aun cuando los resultados y la significancia de estos atrasos varían a lo largo del país. En regiones como Pasco, Puno, Cajamarca y Huancavelica, donde el ámbito rural es predominante, se confirma la existencia de un efecto vecindario sobre el atraso escolar en los estudiantes de primaria y secundaria con un 1% de significancia. Sin embargo, es importante indicar que en Arequipa se encontró un efecto nulo, mientras que tanto en Ica como en Madre de Dios los resultados no son estadísticamente significativos, y tanto Loreto como Huancavelica presentan un efecto bastante por encima al del resto de regiones.

Tabla 4 Resultados a nivel regional vía modelo SAR

Región	Efecto endógeno		Región	Efecto endógeno	
Amazonas	0,0033	***	Lambayeque	0,0006	*
Ancash	0,0007	***	Lima	0,0015	***
Apurímac	0,0002	**	Loreto	0,0098	***
Arequipa	0,0000	***	Madre de Dios	0,0004	
Ayacucho	0,0024	**	Moquegua	0,0013	***
Cajamarca	0,0018	***	Pasco	0,0010	***
Cusco	0,0005	***	Piura	0,0024	***
Huancavelica	0,0075	***	Puno	0,0014	***
Huánuco	0,0003	***	San Martín	0,0017	***
Ica	0,0001		Tacna	0,0005	*
Junín	0,0020	**	Tumbes	0,0032	***
La Libertad	0,0008	**	Ucayali	0,0029	*

* p < 0,10, ** p < 0,05, ***p < 0,01
Fuente: Elaboración propia 2016.

Por último, se debe señalar que para realizar estas estimaciones se empleó el comando *spautoreg* de *Stata*, el cual, al trabajar bajo un enfoque espacial y bayesiano, genera una serie de iteraciones hasta lograr la convergencia del modelo. Luego de la convergencia, emplea la matriz de pesos espaciales para estimar los modelos.

Finalmente, se empleó un test de especificación con el objetivo de analizar la omisión o inclusión de variables. El test empleado fue un test de especificación de los errores (RESET); los resultados de este test rechazan la hipótesis nula de omisión de variables con un *p-value* de 0,215.

En el anexo 3 se presentan los resultados encontrados en los tres modelos estimados. En estos se presentan tanto el efecto endógeno, el efecto exógeno como los efectos fijos del vecindario.

Conclusiones

Sobre la base de los resultados y al análisis de la literatura existente, se puede concluir que existen efectos vecindarios significativos. Tanto el efecto endógeno como de contexto resultaron significativos con un nivel de confianza del 99%. Sin embargo, en ambos casos el coeficiente es reducido, y a veces cercano a cero. Esto indicaría que si bien el efecto vecindario existe, el impacto sobre el atraso escolar es reducido.

Este resultado es similar al expuesto por autores como Ginther, Haveman y Wolfe (2000) y Aaronson (1998), en este sentido, aun cuando el coeficiente encontrado es reducido se puede confirmar que existe un “multiplicador social” en el sentido de que el atraso escolar en el vecindario genera que los estudiantes de este vecindario se encuentren más propensos a tener atraso escolar entre primaria y secundaria. El efecto endógeno encontrado va de 0,0014 a 0,0022.

En el caso del efecto de contexto, los resultados son en la mayoría de los casos significativos a un nivel de confianza del 99%, pero el coeficiente encontrado es bastante reducido, por lo que, en base a lo dicho por Durlauf (2004), es posible que esto se encuentre ligado a la alta correlación entre los estudiantes en cada uno de los vecindarios analizados.

Los resultados indican que este efecto debe ser considerado entre la serie de factores que generan impacto sobre el atraso escolar de los estudiantes. Asimismo, al ser un factor aparentemente importante, sería relevante desarrollar políticas que permitan difundir o sensibilizar la importancia de permanecer y asistir a las instituciones educativas, especialmente, difundir las implicancias del atraso escolar tanto a los padres de familia como a los estudiantes.

Esta difusión y sensibilización debe realizar no solo en las aulas, sino a nivel nacional. El principal problema del atraso escolar, es que, una vez un estudiante se atrasa un año, este año no podrá ser recuperado, es decir, el estudiante mantendrá este atraso de un año a lo largo de su vida escolar y/o estudiantil. Esto, muchas veces, coloca a los alumnos con atraso escolar en una seria desventaja respecto aquellos que finalizaron sus estudios a la edad establecida por el MINEDU. Esta desventaja genera un impacto negativo sobre el desarrollo de estas personas lo cual repercute el nivel de bienestar al cual pueden acceder.

En este sentido, resulta importante prestar importancia a lo que sucede fuera de las instituciones educativas, si bien esto puede encontrarse al margen de las funciones del MINEDU y/o de los

gobiernos regionales y locales, la difusión de la importancia de la educación, especialmente en el ámbito rural donde se concentra la mayor proporción de alumnos con atraso escolar, es resulta relevante ya que la interacción entre vecinos parece tener un efecto importante sobre el atraso escolar.

Bibliografía

- Aaronson, D. (1998). "Using sibling data to estimate the impact of neighborhoods on children's educational outcomes". *Journal of Human Resources*, Vol. 33, 915-946.
- Agüero, J. M. (2014). *Education and teenage pregnancy in rural Peru: Evidence from nearly three decades*. Department of Economics and EI Institute University of Connecticut.
- Agüero León, J. & Cueto, S. (2004). *Dime con quien estudias y te diré cómo rides: Peer-effect como determinantes del rendimiento escolar*. Consorcio de Investigación Económica y Social (CIES).
- Ainsworth, J. W. (2002). *Why Does It Take a Village? The Mediation of Neighborhood Effects on Educational Achievement*. *Social Forces*, Vol. 81(1), 117-152.
- Alcázar, L., & Valdivia, N. (2005). *Análisis de la deserción escolar en el Perú: Evidencia a partir de encuestas y de técnicas cualitativas*. Lima: Reporte de Investigación: GRADE.
- Anselin, L. (1988). *Spatial econometrics methods and models*. *Economic Inquiry*.
- Anselin, L. (2003). "Spatial externalities and spatial econometrics. *International regional science review*", Vol. 26(2), 153 - 166.
- Barde, J. A., & Walkiewicz, J. (2013). *The impact of access to piped drinking water on human capital formation: Evidence from brasilian primary schools*. Beiträge zur Jahrestagung des Vereins für Socialpolitik.
- Barro, R. J. (1991). "Economic growth in cross-section of countries". *Quarterly Journal of Economics*, 407-443.
- Bassi, M., Busso, M., & y Muñoz, J. S. (2014). *Is the glass half empty or half full?: Enrollment, graduation and dropout rates in Latin America*. Inter-American Development Bank.
- Baumol, W. J., Nelson, R. R., & Wolff, E. N. (1996). "Converge of productivity: Cross-national studies and historical evidence". *Southern Economic Association*, Vol. 63 N° 3, 809 - 811.
- Beck, N., Gleditsch, K. S., & Beardsley, K. (2006). "Space Is More than Geography: Using Spatial Econometrics in the Study of Political Economy". *International Studies Quarterly*, Vol. 50, 27-44.
- Becker, G., Murphy, K. M., & Tamura, R. (1990). "Human capital, fertility and economic growth". *Journal of Political Economy*, 12 - 37.
- Beltrán, A., & Seinfeld, J. (2011). *Desnutrición crónica infantil en el Perú: Un problema persistente*. Lima: Centro de Investigación de la Universidad del Pacífico.
- Beltrán, A., & Seinfeld, J. (2011). *Hacia una educación de calidad: La importancia de los recursos pedagógicos en el rendimiento escolar*. Lima: Centro de Investigación de la

Universidad del Pacífico, Consorcio de Investigación Económica y Social.

- Black, S. E., Devereux, P. J., & Salvanes, K. G. (2004). *The more the merrier?: The effect of family composition on children's education*. National Bureau of Economic Research (NBER), Working Paper Series 10720.
- Bobonis, G. J., & Finan, F. (2009). "Neighborhood peer effects in secondary school enrollment decisions". *Review of Economics and Statistics*, Vol. 91 N° 4, 695-716.
- Brooks-Gunn, J., Duncan, G., & Aber, J. (1997). *Neighborhood Poverty, Volume 2: Policy Implications in Studying Neighborhoods*. Russell Sage Foundation.
- Buitenhuis, A. M., Paz-Soldanb, V., Barbu, C., Skoivira, C., Quintanilla Calderon, J., Mollesaca Rieros, L. M., y Levy, M. Z. (2014). "Is participation contagious? Evidence from a household vector control campaign in urban Peru". *Journal of Epidemiology Community Health*, Vol. 68 N° 2, 103-109.
- Cavero, D., Montalva, V., & y Rodríguez, J. (2011). *Determinantes socioeconómicos de las transiciones entre niveles educativos: Un enfoque sobre género y ruralidad en el Perú*. Lima: Departamento de Economía - Pontificia Universidad Católica del Perú.
- IV Censo Nacional Agropecuario 2012 (CENAGRO 2012), Instituto Nacional de Estadística e Informática.
- Chinen, M. (2002). *Indicadores de equidad educativa en el Perú: Un análisis de los Censos Escolares de 1993 y 1008*. Lima: Consorcio de Investigación Económica y Social (CIES) - Grupo de Análisis para el Desarrollo (GRADE).
- Cortez, R. (2001). *El atraso escolar en el Perú: Lecciones para una agenda de política pública*. Lima: Centro de Investigación de la Universidad del Pacífico (CIUP) - Consorcio de Investigación Económica y Social (CIES).
- Crane, J. (1991). *The Epidemic Theory of Ghettos and Neighborhood Effects on Dropping Out and Teenage Childbearing*. *American Journal of Sociology*. Vol. 96 N°1, 226–259.
- Crowder, K., & South, S. J. (2011). "Spatial and temporal dimensions of neighborhood effects on high school graduation". *Social Science Research*, Vol. 40 N° 1, 87 - 106.
- Del Bello, C. L., Paacchini, E., & Zenou, Y. (2015). *Neighborhood Effects in Education*. Institute for the Study of Labor (IZA). Discussion Paper N° 8956.
- DeSilva, S., Pham, A., & Smith, M. (2010). *Racial Preferences in a Small Urban Housing Market: A Spatial Econometric Analysis of Microneighborhoods in Kingston, New York*. Economics Working Paper Archive wp_599, Levy Economics Institute.
- DGPP, D. G. (2015). *Construcción de modelos de accesibilidad: Una aplicación para herramientas de priorización de gasto público*. Lima: Ministerio de Economía y Finanzas (MEF).

- Dietz, R. D. (2002). "The estimation of neighborhood effects in the social sciences: An interdisciplinary approach". *Social Science Research*, Vol. 31 N° 4, 539-575.
- Diez Roux, A. V. (2001). "Investigating neighborhood and area effects on health". *American Journal of Public Health*, Vol. 91 N° 11, 1783-1789.
- Durlauf, S. N. (2004). "Neighborhood effects". *Handbook of Regional and Urban Economics*, Vol. 4 N° 4, 2173-2242.
- Evaluación Censal de Estudiantes 2014 (ECE 2014), Ministerio de Educación del Perú.
- Elías, M., & Rey, S. J. (2011). "Educational performance and spatial Convergence in Peru". *Region et Development*, Issue: 33, 107-135.
- Encuesta Nacional de Hogares sobre Condiciones de Vida y Hogares Rurales 2014 (ENAHOR 2014), Instituto Nacional de Estadística e Informática.
- Encuesta Nacional de Hogares sobre Condiciones de Vida y Hogares 2012 (ENAHO 2012), Instituto Nacional de Estadística e Informática.
- Escobal, J., & Torero, M. (2000). *¿Cómo enfrentar una geografía adversa?: El rol de los activos públicos y privados*. Grupo de Análisis para el Desarrollo (GRADE), Documento de trabajo N° 29.
- Gibbons, S. (2002). *Neighbourhood Effects on Educational Achievement: Evidence from the Census and National Child Development Study*. London: Centre for the Economics of Education, London School of Economics and Political Science.
- Ginther, D., Haveman, R., & Wolfe, B. (2000). "Neighborhood attributes as determinants of children's outcomes: How robust are the relationships". *Journal of Human Resources*, Vol. 35, 603-642.
- Glewwe, P., Krutikova, S., & Rolleston, C. (2014). *Do Schools Reinforce or Reduce Learning Gaps between Advantaged and Disadvantaged Students? Evidence from Vietnam and Peru*. Department of International Development at the University of Oxford, 46. Young Lives Working Paper 133.
- Helmers, C., & Patnam, M. (2014). "Does the rotten child spoil his companion?: Spatial peer effects among children in rural India". *Quantitative Economics*, Vol. 5, 67 - 121.
- Ilahi, N. (2001). "Children's work and schooling: Does gender matter? Evidence from Peru LSMS". *Economist*, 44.
- Ioannides, Y. M., & Zabel, J. E. (2008). "Interactions, neighborhood selection and housing demand". *Journal of Urban Economics*, 229 - 252.
- Jargowsky, P. A., & El Komi, M. (2009). *Before or after the Bell? School Context and Neighborhood Effects on Student Achievement*. National Center for Analysis of Longitudinal Data in Education Research, Working Paper 28.

- Jencks, C., & Mayer, S. E. (1990). *The Social Consequences of Growing Up in a Poor Neighborhood*. Inner-City Poverty in the United States, 111-186.
- Kalenkosi, C. M., & Lacombe, D. J. (2014). *Using spatial econometric techniques to analyze the joint employment decisions of spouses*. IZA Discussion Paper N° 8050.
- Kangjuan, L., Tao, W., & Siyi, G. (2012). “Study on the spatial effect of provincial education investment based on spatial statistics”. *International Journal of Information and Education Technology*, Vol. 2(4).
- Karlan, D., Mobins, M. M., Rosenblat, T. S., & Szeidl, A. (2009). *Measuring Trust in Peruvian Shantytowns*. Innovation for Poverty Action (IPA).
- KAS. (2011). *Inversión en infraestructura pública y reducción de la pobreza en América Latina*.
- Kling, J. R., Liebman, J. B., & Katz, L. F. (2007). “Experimental Analysis of Neighborhood Effects”. *Econometrica*, Vol. 75 N° 1, 83-119.
- Lavado, P., & Gallegos, J. (2006). *La dinámica de la deserción escolar en el Perú: Un enfoque usando modelos de duración*. Lima: Consorcio de Investigación Económica y Social (CIES) - Centro de Investigación de la Universidad del Pacífico (CIUP).
- Lee, L.-f. (2007). “Identification and estimation of econometric models with group interactions, contextual factors and fixed effects”. *Journal of Econometrics*, Vol. 140 N° 2, 333 - 374.
- LeSage, J. L. (1999). *The theory and practice of spatial econometrics*. Departamento de Economía, Universidad de Toledo.
- LeSage, J., & Pace, R. K. (2009). “Introduction to Spatial Econometrics”. *International Statistical Review*, Vol. 77(3), 1 - 472.
- Lin, X. (2010). “Identifying Peer Effects in Student Academic Achievement by Spatial Autoregressive Models with Group Unobservables”. *Journal of Labor Economics*, Vol. 28 N° 4, 825 - 860.
- Lucas, R. E. (1988). “On the mechanics of economic development”. *Journal of Monetary Economics*, Vol. 22, 3 - 42.
- Macroconsult. (2014). *Elaboración de la evaluación de impacto y la ampliación de la línea de base del Programa de Transporte Rural Descentralizado (PTRD)*. Lima: Macroconsult.
- Manski, C. F. (1993). “Identification of endogenous social effects: The reflection problem”. *Review Economic Studies*, Vol. 60 N° 3, 531 - 542.
- Montero, C., Oliart, P., Ames, P., & Cabrera, Z. U. (2001). *La Escuela Rural: Modalidades y prioridades de intervención*. Programa Especial Mejoramiento de la Calidad de la Educación Peruana (MECEP), Ministerio de Educación.

- Oakes, M. (2004). "The (mis)estimation of neighborhood effects: Causal inference for a practicable social epidemiology". *Social Science and Medicine*, Vol. 58 N° 10, 1929 - 1952.
- Ord, K. (1975). "Estimation methods for models of spatial interaction". *Journal of the American Statistical Association*, Vol. 70(349), 120 - 126.
- Pariguana, M. (2011). *Trabajo adolescente y deserción escolar en el Perú*. Lima: Consorcio de Investigación Económica y Social (CIES) - Grupo de Análisis para el Desarrollo (GRADE).
- Pavanello, S., & Othieno, T. (2008). *Improving the provision of basic services for the poor in fragile environments*. Londres: Overseas Development Institute.
- Plenzler, Nicole Lynn (2004). "Student performance and educational resources: a spatial econometric examination". *Theses and Dissertations. Paper 1535*. <<http://utdr.utoledo.edu/theses-dissertations/1535>>
- Rebelo, S. (1991). "Long-run policy analysis and long-run growth". *Journal of Political Economy*, 500 - 521.
- Romer, Paul M. (1990). "Human capital and growth: Theory and evidence". *Elsevier*, vol. 32(1), January, Carnegie-Rochester Conference Series on Public Policy, pages 251-286.
- Sakellariou, C. (2008). "Peer effects and the indigenous/non-indigenous early test-score gap in Peru". *Education Economics*, 371-390.
- Salazar, V., Quispe, V., & Choque, R. (2015). "Educación y movilidad social en el Perú". *Evidencia para políticas públicas en educación superior*. Lima: Programa Nacional de Becas y Crédito Educativo, MINEDU.
- Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). *Assessing "Neighborhood Effects": Social Processes and New Directions in Research*. *Annual Review of Sociology*, Vol. 28(1), 443-478.
- Schelling, T. (1972). *A Process of Residential Segregation: Neighborhood Tipping*. In *Racial Discrimination in Economic Life*, ed. Anthony Pascal (Lexington, Mass.: Lexington Books) 157-184.
- Skoufias, E. (2005). "PROGRESA and its impacts on the welfare of rural households in Mexico". *Research Report of the International Food Policy Research Institute*, 1-98.
- Tobler, W., Gale, S., & Olsson, G. (1979). "Cellular geography". *Philosophy in Geography*, ed. S. Gale and G. Olsson, Dordrecht: Riedel, 379-386.
- Vartanian, T. P., & Gleason, P. M. (1999). "Do neighborhood conditions affect high school dropout and college graduation rates?" *Journal of Socio-Economics*, Vol. 28 N° 1, 21-41.
- Viton, P. A. (2010). "Notes on Spatial Econometric Models". *City and regional planning 870.03*. The Ohio State University.

- Wang, H. C. (2010). “Institutions of higher education and the regional economy: A long-term spatial analysis”. *Economics Research International*, 19.
- Wilson, W.J. (1987). *The Truly Disadvantaged*. Chicago: University of Chicago Press.
- Yadavalli, A. P., Waldorf, B. S., & Florax, R. J. (2013). *The effect of U.S. school quality on migration: A spatial econometric approach*. Department of Agricultural Economics, Purdue University, Working Paper.
- Yamada, G., & Castro, J. F. (2007). *Retornos a la educación superior en el mercado laboral: ¿vale la pena el esfuerzo?* Lima: Consorcio de Investigación Económica y Social - Centro de Investigación de la Universidad del Pacífico.

Anexos

Anexo 1. Reporte de la generación de distancia entre vecinos e identificación de los vecinos más cercanos

Reporte de generación de distancia entre vecinos (comando geonear - Stata)						
Current re- gion	Base locs	Nbor locs	Ops in re- gion	Cumul ops (m)	Cumul % % done	Remaining (hh:mm:ss)
40	32	466	14.912	0,3	3,5	00:04:39
87	61	169	10.309	0,5	7,5	00:04:06
140	23	113	2.599	0,8	11,8	00:03:46
200	47	288	13.536	1,2	16,4	00:03:25
256	47	171	8.037	1,5	20,8	00:03:11
308	28	81	2.268	1,8	24,5	00:03:07
347	57	215	12.255	2,1	27,7	00:03:05
393	38	225	8.550	2,4	31,1	00:03:00
427	38	313	11.894	2,6	33,9	00:02:58
486	16	83	1.328	3	38,1	00:02:45
537	12	505	6.060	3,2	41,9	00:02:35
594	34	75	2.550	3,6	46,5	00:02:20
659	25	149	3.725	4	51,3	00:02:05
725	48	131	6.288	4,4	56,3	00:01:50
791	3	3	9	4,9	61,9	00:01:33
852	35	251	8.785	5,2	66,8	00:01:20
917	44	137	6.028	5,6	71,7	00:01:08
977	57	165	9.405	6	76,7	00:00:55
1038	16	496	7.936	6,4	81,8	00:00:43
1098	29	449	13.021	6,8	86,7	00:00:31
1158	58	203	11.774	7,1	91	00:00:21
1224	67	169	11.323	7,6	96,6	00:00:08

Unique base locations = 47.120 Unique neighbor locations = 47.120

Bases *Neighbors (M) = 2220,3 Number of regions = 1.268

Computed distances (M) = 7,79 Total run time (seconds) = 231.454

Anexo 2. Estadísticas descriptivas

Variable	Significado	Mean	Std. Dev	Min	Max
atraso	Variable <i>dummy</i> . El 33% de la población analizada presenta atraso escolar encontrándose aún en edad escolar.	0,33	0,47	-	-
c1p25	Variable <i>dummy</i> , toma el valor de 1 si es hombre y 0 si es mujer. El 50% de la población analizada es mujer y 50% hombre.	0,50	0,50	-	-
c1p27_anio	Edad. La edad promedio de los estudiantes analizados es 11 años.	11,41	3,38	6,00	17,00
dist_educa	Distancia del hogar al centro educativo. La distancia es medida en metros.	1.720,07	3.364,08	8,82	55.096,43
dist_salud	Distancia del hogar al centro de salud. La distancia es medida en metros.	5.666,33	6.691,54	11,53	64.413,03
M500_C_2014	Puntaje obtenido por los estudiantes en las instituciones educativas del centro poblado donde residen los estudiantes en la prueba de comprensión de lectura.	526,14	54,20	358,60	759,20
M500_M_2014	Puntaje obtenido por los estudiantes en las instituciones educativas del centro poblado donde residen los estudiantes en la prueba de matemáticas.	528,31	71,80	314,07	810,33
elect14	Indica el acceso a electricidad en la institución educativa. El 75% de las instituciones ubicadas en los centros poblados donde se encuentran los vecindarios cuenta con electricidad.	0,75	0,26	-	-
aguapot14	Indica el acceso a agua potable en la institución educativa. El 93% de las instituciones ubicadas en los centros poblados donde se encuentran los vecindarios cuenta con agua potable.	0,93	0,10	-	-
desague14	Indica el acceso a desagüe en la institución educativa. El 93% de las instituciones ubicadas en los centros poblados donde se encuentran los vecindarios cuenta con desagüe.	0,93	0,10	-	-
conint14	Indica el acceso a internet en la institución educativa. El 80% de las instituciones ubicadas en los centros poblados donde se encuentran los vecindarios cuenta con internet.	0,80	0,18	-	-
comp14	Indica el número de computadoras en la institución educativa de uso estrictamente pedagógico. El número promedio de computadoras en las instituciones ubicadas en los centros poblados donde se encuentran los vecindarios es 18.	18,47	10,85	-	89,00

Variable	Significado	Mean	Std. Dev	Min	Max
biblioteca14	Indica el porcentaje de instituciones educativas que cuentan con bibliotecas. El 5% de las instituciones ubicadas en los centros poblados donde se encuentran los vecindarios cuentan con bibliotecas.	0,05	0,08	-	-
labciencia14	Indica el porcentaje de instituciones educativas que cuentan con laboratorios de ciencia. El 2% de las instituciones ubicadas en los centros poblados donde se encuentran los vecindarios cuentan con laboratorios de ciencias.	0,02	0,04	-	-
taller14	Indica el porcentaje de instituciones educativas que cuentan con talleres. El 3% de las instituciones ubicadas en los centros poblados donde se encuentran los vecindarios cuentan con talleres.	0,03	0,05	-	-
salaprof14	Indica el porcentaje de instituciones educativas que cuentan con salas de profesores. El 3% de las instituciones ubicadas en los centros poblados donde se encuentran los vecindarios cuentan con salas de profesores.	0,03	0,06	-	-
salacomp14	Indica el porcentaje de instituciones educativas que cuentan con salas de cómputo. El 10% de las instituciones ubicadas en los centros poblados donde se encuentran los vecindarios cuentan con salas de cómputo.	0,10	0,12	-	-
aulas14	Indica el número de computadoras en la institución educativa de aulas. El número promedio de aulas en las instituciones ubicadas en los centros poblados donde es 5.	5,02	2,56	1,22	21,89
aulas14_rmy	Indica el número de computadoras en la institución educativa de aulas que requieren reparaciones mayores. El número promedio de aulas que requieren reparaciones mayores en las instituciones ubicadas en los centros poblados donde es 1.	0,96	0,79	-	6,83
aulas14_rmn	Indica el número de computadoras en la institución educativa de aulas que requieren reparaciones menores. El número promedio de aulas que requieren reparaciones menores en las instituciones ubicadas en los centros poblados donde es 3.	2,79	1,51	-	12,11
aulas14_bcon	Indica el número de computadoras en la institución educativa de aulas en buenas condiciones. El número promedio de aulas en buenas condiciones en las instituciones ubicadas en los centros poblados donde es 1.	1,24	1,31	-	10,17
km_to_nid	Distancia entre los hogares, específicamente identifica a los hogares más cercanos. La distancia es medida en km.	0,22	0,60	-	12,97

Anexo 3. Modelos estimados

Resultados de los tres modelos

Efectos		Modelo 1	Modelo 2	Modelo 3
Endógeno		0,0016 ***	0,0002***	0,0014***
Contexto	Sexo	0,000***	0,000***	0,001**
	Quechua	0,005**	0,003**	0,003**
	Distancia a la inst. educativa	0,044***	0,016***	0,025***
	Distancia al centro de salud	0,048***	0,019***	0,026***
	Nivel educativo del jefe del hogar	-	-	-0,037***
Efectos fijos	Alumbrado	-0,016***	-0,0196***	-0,0008***
	Agua potable	-0,021***	-0,016**	-0,0025***
	Desagüe	-0,001**	-0,000***	-0,001**
	Bibliotecas	-	-	0,000**
	Aulas en buenas condiciones	-	-0,039***	-
Rho		0,002	0,009	0,0005
F-Test		10,820***	10,439***	11,273***
Log Likelihood Function		4701,189	5887,694	4815,485

El número de observaciones en cada una de los modelos es 4.570.994.

*p < 0,10, ** p< 0,05, ***p< 0,01

Fuente: Elaboración propia 2016.

Nota biográfica

Natalia Pía Guerrero Trinidad

Nació en Lima, el 29 de diciembre de 1992. Bachiller en Economía, egresada de la Universidad del Pacífico. Cuenta con experiencia en consultoría e investigación. Actualmente desempeña el cargo de Asistente de Investigación en la Universidad del Pacífico.